Constrained Variational Analysis of Sounding Data Based on Column-Integrated Budgets of Mass, Heat, Moisture, and Momentum: Approach and Application to ARM Measurements

M. H. ZHANG AND J. L. LIN

Institute for Terrestrial and Planetary Atmospheres, Marine Sciences Research Center, State University of New York at Stony Brook, Stony Brook, New York

(Manuscript received 14 June 1996, in final form 4 November 1996)

ABSTRACT

For the purpose of deriving grid-scale vertical velocity and advective tendencies from sounding measurements, an objective scheme is developed to process atmospheric soundings of winds, temperature, and water vapor mixing ratio over a network of a small number of stations. Given the inevitable uncertainties in the original data, state variables of the atmosphere are adjusted by the smallest possible amount in this scheme to conserve column-integrated mass, moisture, static energy, and momentum. The scheme has the capability of incorporating a variety of supplemental measurements to constrain large-scale vertical velocity and advective tendencies derived from state variables.

The method has been implemented to process the Atmospheric Radiation Measurement Program's (ARM) soundings of winds, temperature, and water vapor mixing ratio at the boundary facilities around the Cloud and Radiation Testbed site in northern Oklahoma in April 1994. It is found that state variables are adjusted by an amount comparable to their measurement uncertainties to satisfy the conservation requirements of mass, water vapor, heat, and momentum. Without these adjustments, spurious residual sources and sinks in the column budget of each quantity have the same magnitudes as other leading components. Sensitivities of the diagnosed vertical velocity and apparent heat, moisture, and momentum sources to the number of conservation constraints are presented. It is shown that constraints of column budget of moisture and dry static energy can make large differences to these diagnostics, especially when some original sounding data are missing and have to be interpolated.

Analysis of the moisture budget shows that large-scale convergence often corresponds to precipitation, but there are occasions when precipitation corresponds to a large reduction of column precipitable water and column-moisture divergence. Analysis of momentum budget shows large magnitudes of subgrid-scale momentum sources and sinks (about 4 m s^{-1} h^{-1}) in the convective events.

1. Introduction

Data from a sounding array can be used to calculate vertical velocity and advective tendencies for an atmospheric column. These diagnostics are often used to 1) describe the structure of the large-scale systems and budgets (e.g., Thompson et al. 1979); 2) drive semi-prognostic GCM physical parameterizations (e.g., Lord 1982; Wang and Randall 1996) or single-column models (Randall et al. 1996); and 3) diagnose effects of subgrid circulations on the resolvable-scale fields (Yanai and Johnson 1993). A review of relevant studies can be found in Cotton and Anthes (1989) and Emanuel (1994).

When measurements of surface pressure, sensible heat flux, latent heat flux, precipitation, wind stress, and radiative fluxes are available, the column-integrated budgets of mass, heat, moisture, and momentum provide checks on accuracies of the vertical velocity and advective tendencies. Because of inevitable uncertainties in atmospheric measurements, wind divergence calculated from sounding measurements typically does not meet mass conservation; therefore, in previous studies (e.g., O'Brien 1970; Lin and Johnson 1996) winds were customarily adjusted to conserve air mass. Similarly, observed surface fluxes and boundary fluxes of moisture and heat diagnosed from sounding measurements typically do not satisfy their column-integrated budgets (e.g., Yanai et al. 1973; Thompson et al. 1979; Frank 1979; Kuo and Anthes 1984). Yet corrections to these unbalanced budgets and their impact on diagnosed atmospheric vertical velocity and advective tendencies have not received enough attention in the past. This is probably due to lack of or inaccuracies in measurements of relevant surface fluxes of moisture and heat associated with sounding measurements. These unbalanced budgets are particularly worrisome if the diagnosed ver-

Corresponding author address: Dr. Minghua Zhang, Institute for Terrestrial and Planetary Atmospheres, Marine Sciences Research Center, State University of New York, Stony Brook, NY 11794-5000. E-mail: mzhang@atmsci.msrc.sunysb.edu

tical velocity and advective tendencies are intended to drive physical parameterizations and models.

The lack of surface and top-of-the-atmosphere (TOA) measurements associated with sounding measurements has changed recently. Several experiments have included the direct measurements of surface fluxes by using the eddy correlation or the Bowen ratio methods. These projects include the Convection and Precipitation Experiment (CaPE) in 1991 (Wakimoto and Atkins 1994), the Tropical Ocean Global Atmosphere (TOGA) Coupled Ocean–Atmosphere Response Experiment (COARE) from 1992 to 1993 (Webster and Lukas 1992), and the ongoing Atmospheric Radiation Measurement (ARM) project (Stokes and Schwartz 1994). With surface and TOA measurements of the flux components, it is possible to check the column-integrated budgets using data from these experiments. Figure 1 gives an example in which the observational data are taken from the ARM April 1994 Intensive Observation Period (IOP) around the Cloud and Radiation Testbed (CART) facility in northern Oklahoma.[1] The wind soundings have been adjusted to meet column-integrated mass conservation. Plotted in the figure is the spurious moisture source diagnosed from the residual of the column-integrated moisture budget. Also plotted is the moisture source due to evaporation minus precipitation. The spurious residual term is as large as the precipitation, thereby limiting the utility of the data. It should be noted that dynamically and thermodynamically consistent analyses can be derived from the output of a numerical model such as the operational weather forecast model (e.g., Daley 1991). In that case, moisture and energy are all balanced. Such a balance, however, is achieved by using model-calculated precipitation and model variables, which are not the same as observations.

In this study, by making minimum adjustments to the original sounding data, we force atmospheric state variables to satisfy conservations of mass, moisture, energy, and momentum through a variational technique. The purpose of the study is twofold. The first is to examine the amount of adjustments required to make the data meet the conservation constraints, in particular the relative magnitude of the adjustment with respect to measurement and instrument uncertainties. The second is to study the sensitivity of diagnostic variables such as vertical velocity, apparent heat, and moisture sources to the conservation constraints. We show that the magnitude of adjustments required to meet the conservations of mass, moisture, energy, and momentum at each individual station is comparable to that of measurement uncertainties. Several important diagnostic variables are affected by the implementation of conservation constraints. A by-product of this study is a dataset that is dynamically and thermodynamically consistent in terms of vertically integrated quantities, with adjustments not far exceeding uncertainties of the original measurements. The dataset can be used to perform budget studies and to drive single-column models.

The paper is organized as follows. Section 2 describes the analysis strategy and its numerical implementation. Section 3 discusses the input data and preprocessing procedure. Experimental design is described in section 4. The results are presented in section 5 in the form of adjustments to the state variables with different constraints. Section 6 presents sensitivities of diagnostic variables to the implementation of conservation constraints. Section 7 discusses restrictions and possible improvements of the analysis scheme. The last section summarizes the results.

2. The method

a. Theoretical formulation

The governing equations of the *large-scale* atmospheric fields are

$$\frac{\partial \mathbf{V}}{\partial t} + \mathbf{V}\cdot\nabla\mathbf{V} + \omega\frac{\partial \mathbf{V}}{\partial p} + f\,\mathbf{k}\times\mathbf{V} + \nabla\phi = -\overline{\nabla\cdot(\mathbf{V}'\mathbf{V}')} - \frac{\overline{\partial\omega'\mathbf{V}'}}{\partial p}, \tag{1}$$

$$\frac{\partial s}{\partial t} + \mathbf{V}\cdot\nabla s + \omega\frac{\partial s}{\partial p} = Q_{\text{rad}} + L(C - E) - \overline{\nabla\cdot(\mathbf{V}'s')} - \frac{\overline{\partial\omega's'}}{\partial p} + L\frac{\partial q_l}{\partial t}, \tag{2}$$

$$\frac{\partial q}{\partial t} + \mathbf{V}\cdot\nabla q + \omega\frac{\partial q}{\partial p} = E - C - \overline{\nabla\cdot(\mathbf{V}'q')} - \frac{\overline{\partial\omega'q'}}{\partial p} - \frac{\partial q_l}{\partial t}, \tag{3}$$

$$\frac{\partial \omega}{\partial p} + \nabla\cdot\mathbf{V} = 0, \tag{4}$$

with boundary conditions

$$\omega\big|_{p=p_s} = \frac{\partial p_s}{\partial t} + \mathbf{V}_s\cdot\nabla p_s \tag{5}$$

and

$$\omega\big|_{p=p_0} = 0, \tag{6}$$

where $\mathbf{V}$ is the wind, $s = C_pT + gz$ is the dry static energy, q is the mixing ratio of water vapor, and p_s is the surface pressure. Large scale is also referred to as grid scale, defined as the size of a sounding array comprising several stations. Prime denotes unresolvable motions of the observational network. Here, Q_{rad} is the net

[1] After this paper was accepted, the authors acquired ARM data for other IOPs in which there was one more sounding station than in April 1994. Processed data for these IOPs can be accessed through anonymous FTP to ftp://atmsci.msrc.sunysb.edu.

radiative heating related with clouds, C is the condensation of water vapor to rainwater, and E is the evaporation of rainwater. Phase changes associated with ice can be included but are neglected for simplicity. Here q_l is the cloud liquid water content.

All variables on the left-hand side of (1)–(4) can be calculated from coordinated sounding measurements of winds, temperature, and humidity over a network of a few stations. Terms on the right-hand side of (1)–(4) constitute the unknowns. They are the subject of parameterizations in a large-scale model. Their validation often requires the diagnostics of the left-hand-side terms. With no prior knowledge of the unknowns, vertical integration of the above equations yields the following constraints:

$$\langle \nabla\cdot\mathbf{V}\rangle = -\frac{1}{g}\frac{dp_s}{dt}, \tag{7}$$

$$\frac{\partial\langle q\rangle}{\partial t} + \langle \nabla\cdot\mathbf{V}q\rangle = E_s - P_{\text{rec}} - \frac{\partial\langle q_l\rangle}{\partial t}, \tag{8}$$

$$\frac{\partial\langle s\rangle}{\partial t} + \langle \nabla\cdot\mathbf{V}s\rangle = R_{\text{TOA}} - R_{\text{SRF}} + LP_{\text{rec}} + \text{SH} + L\frac{\partial\langle q_l\rangle}{\partial t}, \tag{9}$$

$$\frac{\partial\langle \mathbf{V}\rangle}{\partial t} + \langle \nabla\cdot\mathbf{V}\mathbf{V}\rangle + f\,\mathbf{k}\times\langle\mathbf{V}\rangle + \nabla\langle\phi\rangle = \boldsymbol{\tau}_s, \tag{10}$$

where

$$\langle X\rangle \equiv \frac{1}{g}\int_{p_T}^{p_s}(X)\,dp.$$

These are column-integrated conservations of mass, water vapor, dry static energy, and momentum. Here, R is the net downward radiative flux at the TOA and at the surface (SRF), $\boldsymbol{\tau}_s$ is the surface wind stress, P_{rec} is precipitation, SH is the sensible heat flux, and E_s is the surface evaporation. Horizontal advection of hygrometers and all horizontal eddy covariance terms have been neglected due to insufficient knowledge. These omissions are not expected to seriously affect the present study if the scale of the phenomena to be studied is much larger than the scale of the observational network.

Terms on the right-hand side of (7)–(10) are available from surface and satellite measurements. They are the area-averaged fluxes within the observational network. The strategy of objective analysis in this study is to constrain the atmospheric variables (**v**, s, q) to satisfy (7)–(10) with minimum adjustments to direct sounding measurements. The adjustment is justifiable after consideration of instrument and measurement uncertainties, errors from handling of missing sounding data, and aliasing of small-scale features to large-scale fields in the instantaneous soundings.

The analyzed product, denoted as $\mathbf{v}^*$ (u^* and v^*), s^*, and q^*, is derived by minimizing the cost function

$$I(t) = \iiint_{p,x,y} [\alpha_u(\mathbf{u}^* - \mathbf{u}_o)^2 + \alpha_v(\mathbf{v}^* - \mathbf{v}_o)^2 + \alpha_s(s^* - s_o)^2 + \alpha_q(q^* - q_o)^2]\,dx\,dy\,dp, \tag{11}$$

with (7)–(10) as strong constraints, where subscript "o" denotes direct measurements and α is the weighting function (discussed later). The integration will be replaced by summation over the stations and on vertical layers. It should be pointed out that, alternatively, we can also minimize the time integration of the above cost function over the whole measurement period to perform data analysis. Such a choice, however, is not well justified if the data uncertainties remain constant with time, and thus it is not used here.

As always, the analyzed data are typically neither measurements nor true values of the variables. An important aspect of the analysis procedure is, therefore, to justify the magnitude of adjustments made to the direct measurements.

Since the general constraints of (7)–(10) apply to all models, this analysis strategy is not particular to any one model. It should be pointed out that most available assimilated products are derived using (1)–(4) with particular parameterization of the unknowns. The technique of the adjoint model is often used to minimize the difference between model-generated variables and observations (e.g., Ghil and Malanotte-Rizzoli 1991; Zou et al. 1993; Xu 1996). Another approach is to include a nudge term in (1)–(4) to relax the model variables to observations (Daley 1991). While these approaches have been very useful in filling observational gaps, the utility of the assimilated data, however, is inherently limited for validating the parameterization of the unknowns of a model in (1)–(4). The current approach

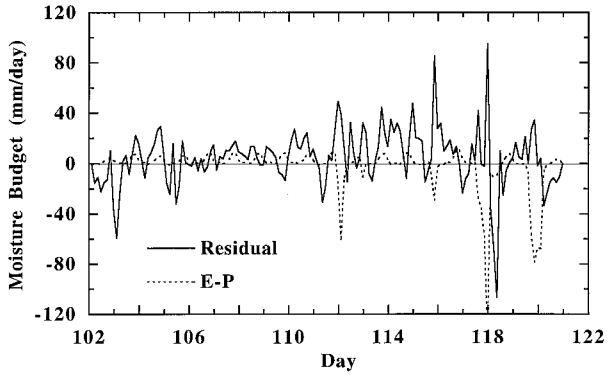

FIG. 1. Moisture residual diagnosed from column moisture budget (solid line) around the ARM CART facility in April 1994. Dashed line is the observed surface evaporation minus precipitation. The winds have been already adjusted to conserve column-integrated mass. The abscissa is time in Julian days.

differs from these assimilation practices in that it relies entirely on measurements.

b. Numerical implementation

For I stations in the sounding network, each with K layers, we use x_{ik} to denote a state variable at station i and layer k, and use column vector $\mathbf{X}$ to denote this variable (u, v, s, q) at all grids,

$$\mathbf{X}^{\mathbf{T}} = (x_{11}, x_{12}, \ldots, x_{1K}, x_{21}, \ldots, x_{ik}, \ldots x_{IK}), \quad (12)$$

where superscript T means transpose. The cost function of (11) can be written as

$$I(t) = (\mathbf{u}^* - \mathbf{u}_o)^{\mathbf{T}}\mathbf{Q}_u(\mathbf{u}^* - \mathbf{u}_o) + (\mathbf{v}^* - \mathbf{v}_o)^{\mathbf{T}}\mathbf{Q}_v(\mathbf{v}^* - \mathbf{v}_o) + (\mathbf{s}^* - \mathbf{s}_o)^{\mathbf{T}}\mathbf{Q}_T(\mathbf{s}^* - \mathbf{s}_o) + (\mathbf{q}^* - \mathbf{q}_o)^{\mathbf{T}}\mathbf{Q}_q(\mathbf{q}^* - \mathbf{q}_o), \quad (13)$$

where $\mathbf{Q}$ is the weighting matrix related with error covariances of a variable. The analyzed data are subject to the strong constraints of (7)–(10). They can be written in the discrete forms

$$A_{p_s}(\mathbf{V}^*_{ik}) = \left(\frac{\partial p_s}{\partial t}\right)_m + \sum_{k=1}^{k=K} (\boldsymbol{\nabla}\cdot\mathbf{V}^*_k)_m \Delta p_k = 0, \quad (14)$$

$$A_q(\mathbf{V}^*_{i,k}, q^*_{ik}) = \left\langle\left(\frac{\partial q^*}{\partial t}\right)_m\right\rangle_k + \langle(\boldsymbol{\nabla}\cdot\mathbf{V}^* q^*)_m\rangle_k - E_s + P_{\text{rec}} + \left\langle\left(\frac{\partial q_l}{\partial t}\right)_m\right\rangle_k = 0, \quad (15)$$

$$A_s(\mathbf{V}^*_{i,k}, s^*_{ik}) = \left\langle\left(\frac{\partial s^*}{\partial t}\right)_m\right\rangle_k + \langle(\boldsymbol{\nabla}\cdot\mathbf{V}^* s^*)_m\rangle_k - R_{\text{TOA}} + R_{\text{srf}} - LP_{\text{rec}} - \text{SH} + L\left\langle\left(\frac{\partial q_l}{\partial t}\right)_m\right\rangle_k = 0, \quad (16)$$

$$\mathbf{A}_{\mathbf{V}}(\mathbf{V}^*_{i,k}, \phi^*_{ik}) = \left\langle\left(\frac{\partial \mathbf{V}^*}{\partial t}\right)_m\right\rangle_k + \langle(\boldsymbol{\nabla}\cdot\mathbf{V}^*\mathbf{V}^*)_m\rangle_k + f\mathbf{k} \times \langle(\mathbf{V}^*)_m\rangle_k + \langle(\boldsymbol{\nabla}\phi^*)_m\rangle_k - \boldsymbol{\tau}_s = 0, \quad (17)$$

where

$$\langle X\rangle_k = \sum_{k=1}^{k=K} (X_k)\Delta p_k$$

and subscript m represents average over the area covered by the I stations. Geopotential height can be derived from the virtual temperature analysis using the hydrostatic balance

$$\frac{\partial \phi^*}{\partial p} = -\frac{RT^*_v}{p}. \quad (18)$$

Surface pressure p_s is not included in the cost function for adjustment, nor are surface evaporation, sensible heat flux, surface wind stress, and cloud liquid content, although their inclusion for adjustment is possible.

The variational equations (Euler–Lagrange equations) for the analyzed variables of u^*_{ik}, v^*_{ik}, q^*_{ik}, s^*_{ik} are

$$\frac{\partial I(t)}{\partial x^*_{ik}} + \lambda_1(t)\frac{\partial A_{p_s}}{\partial x^*_{ik}} + \lambda_2(t)\frac{\partial A_q}{\partial x^*_{ik}} + \lambda_3(t)\frac{\partial A_s}{\partial x^*_{ik}} + \lambda_4(t)\frac{\partial A_u}{\partial x^*_{ik}} + \lambda_5(t)\frac{\partial A_v}{\partial x^*_{ik}} = 0, \quad (19)$$

where x^*_{ik} stands for any one of the variables among u^*_{ik}, v^*_{ik}, q^*_{ik}, s^*_{ik}. Here λ_l are the Lagrange multipliers. Each variable has $I \times K$ grids. With a total of four variables and five Lagrange multipliers, the total number of variables to calculate at any given time is $4 \times I \times K + 5$. They are solved from the $4 \times I \times K$ equations in (19) and the five equations in (14)–(17). It is noted here that the Lagrange multipliers are only functions of time because the constraints are vertically integrated budgets.

We assume measurement errors at different locations and for different variables uncorrelated. The covariance matrix is then diagonal. The diagonal elements are the reciprocal of error variances $\sigma^2_{x_{ik}}$. Thus, (19) becomes

$$2\sigma^{-2}_{x_{ik}}(x^*_{ik} - x_{o,ik}) + \lambda_1(t)\frac{\partial A_{p_s}}{\partial x^*_{ik}} + \lambda_2(t)\frac{\partial A_q}{\partial x^*_{ik}} + \lambda_3(t)\frac{\partial A_s}{\partial x^*_{ik}} + \lambda_4(t)\frac{\partial A_u}{\partial x^*_{ik}} + \lambda_5(t)\frac{\partial A_v}{\partial x^*_{ik}} = 0 \quad (20)$$

or

$$x^*_{ik} = x_{o,ik} - \frac{\sigma^2_{x_{ik}}}{2}\left[\lambda_1(t)\frac{\partial A_{p_s}}{\partial x^*_{ik}} + \lambda_2(t)\frac{\partial A_q}{\partial x^*_{ik}} + \lambda_3(t)\frac{\partial A_s}{\partial x^*_{ik}} + \lambda_4(t)\frac{\partial A_u}{\partial x^*_{ik}} + \lambda_5(t)\frac{\partial A_v}{\partial x^*_{ik}}\right]. \quad (21)$$

The analyzed product is equal to the corresponding measurement plus an adjustment term that is associated with the sensitivity of the column-integrated mass, water vapor, moist static energy, and momentum.

Numerical calculation of (21) and (14)–(17) is carried out in an iterative mode. The iteration is performed for the whole measurement period. To simplify the description, we describe the iteration procedure for a single time level. The iteration for the whole measurement period can be carried out by executing an iteration step to all successive time levels of measurements and then proceeding to the next iteration step up to the completion of all steps. The iteration, when described to a single time level, contains three steps. First, the previous estimate or original measurements are used to calculate each partial derivative to x_{ik} on the right-hand side of (21) using the formulas in (14)–(17). Actual numerical calculation is simplified because the partial derivative does not require vertical summation. The total number of partial derivatives at any given time is $5 \times 4 \times I \times K$, with five constraints (A_{ps}, A_q, A_s, A_u, A_v) taking partial derivatives to four variables (u_{ik}^*, v_{ik}^*, q_{ik}^*, s_{ik}^*) at each station and layer.

A general form of constraint in (14)–(17) can be written as

$$A_Y(x_{ik}^*, y_{ik}^*) = 0. \tag{22}$$

In an iteration, it can be written as

$$A_Y(x_{ik}^{(n)}, y_{ik}^{(n-1)}) = 0, \tag{23}$$

where n denotes the iteration index. When (21) is written as

$$x_{ik}^{(n)} = x_{o,ik} - \frac{\sigma_{x_{ik}}^2}{2} \sum_{l=1}^{5} \lambda_l(t) B_{l,x_{ik}}^{(n-1)}, \tag{24}$$

where $B_{l,x_{ik}}^{(n-1)}$ are the partial derivatives from the first step, substitution of (24) into (23) yields a linearized set of equations for λ_l. More specifically, the substitution gives

$$A_{Y_l}\left(x_{o,ik} - \frac{\sigma_{x_{ik}}^2}{2} \sum_{l=1}^{5} \lambda_l(t) B_{l,x_{ik}}^{(n-1)}, Y_{ik}^{(n-1)}\right) = 0. \tag{25}$$

Because of the linearity of the above operator, it can be further written as

$$A_Y(x_{o,ik}, Y_{ik}^{(n-1)}) - \sum_{l'=1}^{5} \frac{\sigma_{X_{ik}}^2}{2} A_Y(B_{l,X_{ik}}^{(n-1)}, Y_{ik}^{(n-1)}) \lambda_l(t) = 0. \tag{26}$$

This set of five equations for the five constraints (A_{ps}, A_q, A_s, A_u, A_v) is used to solve for λ_l at any given time. This constitutes the second step in the iteration.

In the third step, the adjustments are calculated by using the newly obtained λ_l in (24). After that, the next iteration is performed. In our calculation, good convergence of iteration occurs within 20 cycles of iterations for the whole measurement period, which is judged by column-integrated balance of mass, moisture, heat, and momentum to the order of better than 0.1 (with units Pa day^{-1}, W m^{-2} and N m^{-2}) in (14)–(17) at all time levels.

Note that time derivatives appear in (14)–(17); therefore, more than one measurement time is involved each time. The time derivatives in (14)–(17) are calculated using a central difference scheme with values from the previous iteration cycle. That is why the actual iteration needs to be done for the whole measurement period. It is also noted that we do not adjust the first and last time levels of measurements because of the central time differencing, and budgets are not derived for these two time levels. It is possible to include them in the analysis if forward and backward time difference schemes are used for these two end time levels.

We also note that there might be other algorithms available to obtain the minimization solution of (13) with constraints of (14)–(17), such as using a maximum gradient descent approach. The adjoint method cannot be easily applied here because (14)–(17) are not models that can be integrated in time. The iteration procedure described above only involves solving linear equations and seems to be quite efficient for our purpose. Nevertheless, we welcome the reader to suggest other more efficient algorithms.

The divergence terms are calculated by first assuming that field vectors vary linearly along the sides of the outmost domain formed by connecting the stations and then taking the convergent fluxes perpendicular to the sides divided by the area of the domain. This represents an area-averaged divergence. Gradient is calculated by first taking a least squares fit of a multilinear field to the data and then calculating the derivatives. Detailed procedures are the same as the flux method described in Davies-Jones (1993). As pointed out in Davies-Jones (1993), this procedure, when used for a triangle defined by three stations, is equivalent to several methods of divergence and gradient calculation, including Bellamy's graphical method (Bellamy 1949) and spatial linear fitting. It should be pointed out that even though the analysis is independent of any particular model, there is the possibility that the analysis depends on the numerical approximations. A sensitivity study is performed by using various existing numerical schemes (Davies-Jones 1993) to calculate (14)–(17). It is found that these numerical approximations do not affect the analysis to a degree with which to be concerned.

3. Data and preprocessing

a. Input data

Input measurements can be categorized into two groups. The first is measurements of variables to be adjusted. The second is measurements of variables to form the constraints. In this study, adjustment variables are the state variables only, namely, winds, temperature, and humidity. The constraint variables include all terms on the right-hand side of (7)–(10), namely, surface pres-

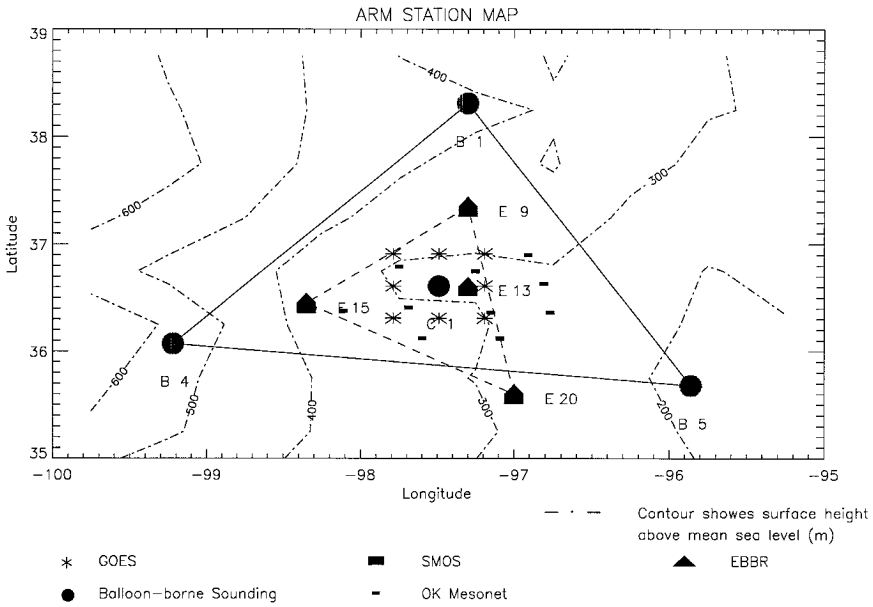

FIG. 2. April 1994 ARM IOP measurement network. Dashed lines are topography heights in meters.

sure, surface latent and sensible heat fluxes, wind stress, precipitation, net radiation at the surface and at the TOA, and the variability of column total cloud water content. It is possible to treat some of the variables in the latter category as adjustment variables. For example, when winds near the surface are adjusted, latent and sensible heat fluxes can be recalculated.

Balloonborne soundings of winds, temperature, and dewpoint temperature, available every 3 h in the April 1994 IOP from 12 April to 30 April, are used as input for the first category. These include sounding data from the CART central facility and three boundary stations (Fig. 2). It is noted here that the observational network shown in Fig. 2 is not an optimal one. In later IOPs another boundary facility has been added. Therefore, results shown in this study are mainly intended to illustrate our method and its impact on the analysis results. For applications using the processed data, we suggest the reader use data processed for other IOPs, which will become available shortly. There are profiler measurements of winds every hour at several stations. These have not been used in the current study. The original data has a vertical resolution of 2 s of balloon ascending time. The data contains the actual time and location of the balloon when measurements are taken.

Constraint variables are derived from measurements of surface pressure and precipitation at four Surface Meteorological Observation System (SMOS) stations and 11 Oklahoma mesonet stations around the CART site within the triangle formed by the three boundary facilities (Fig. 2). Surface latent and sensible heat fluxes are obtained from the four Energy Balance Bowen Ratio (EBBR) stations. They were calculated from a Bowen ratio method. Surface net radiative heat fluxes are obtained from measurements at the four EBBR stations. Net radiation at the TOA is taken from Minnis et al. (1995) over an 3×3 array of grids as shown in Fig. 2. Shortwave and longwave radiative fluxes were derived from the GOES narrowband brightness temperatures through calibration schemes. Only daytime fluxes are available for this IOP; nighttime longwave fluxes have been linearly interpolated. The column total cloud liquid water is taken from the Microwave Water Radiometer (MWR) at the four sounding stations. The surface wind stress is calculated from the SMOS winds, temperature, and pressure following Peixoto and Oort (1992). All these measurements, except the Oklahoma Mesonet, are available every half-hour covering the same period as do the sounding measurements. The Oklahoma Mesonet data are available every 5 min.

b. Preprocessing

The original data are first preprocessed for suitability of automated analysis. The preprocessing consists of quality control of raw data, averaging the data to form large-scale quantities, filling in missing measurements, and interpolation to consistent observation times.

For the sounding data, data entries flagged as suspicious by the ARM data center are treated as missing. The available data are then averaged vertically in 20-mb intervals to get the mean quantities in each layer. They are then used to calculate the mean of all available stations. If more than half the data are available, temperature values that deviate more than 15° from the mean value are considered outliers and treated as missing. The same procedure was applied for height, using a deviation

criterion of 100 m. The data rejected by this procedure are several complete soundings.

The majority of missing data are due to missing original balloon soundings. All missing data are filled and interpolated to the standard launching time using the following method.

1) If missing data last no more than two time steps, they are filled with linear interpolation in time.
2) If missing data last more than three time steps, values averaged over the other three stations are used for filling. If values at other stations are also missing (a very small percentage of grids), linear interpolation in time is used.

Because the free balloons move horizontally with the winds, their position differs from the stations. Displacement of the balloons at higher levels may be large. For the ARM dataset, the displacement is up to 0.7° in the zonal direction and 0.1° in the meridional direction. In our calculation, all divergence and gradient terms on the left-hand side of (7)–(10) are calculated according to the actual locations of balloons. They represent the area average in the triangle formed by the balloons rather than the sounding stations. Since the constraint variables are calculated at the fixed surface stations, the variational procedure tends to adjust the area-averaged tendencies to a fixed domain.

Because sounding data are available every 3 h, all surface measurements are averaged to a 3-h period centered around the sounding measurements. Furthermore, Eqs. (1)–(5) are intended to describe large-scale resolvable fields. At the surface and TOA, measurements from all stations are averaged to form an area mean of the triangle domain. Because of the surface inhomogeneity in this domain (see Fig. 2 for distribution of topography height), an appropriate averaging procedure would use fractional area to weight each station, the fraction being determined by the surface characteristics. In this study, we simply take the arithmetic mean of all available stations in Fig. 2 to compute the average. Eleven mesonet stations and four SMOS stations contributed to the time- and space-averaged precipitation.

c. Effect of topography

To calculate the horizontal divergence and gradient terms, we have to use four stations with different topography heights that vary as much as 400 m among the stations (Fig. 2). Calculations of convergence and gradient at layers above the highest station are straightforward. Calculation at the surface pressure level of the central facility was performed using surface observations at the four SMOS stations, while assuming that measurements can be extrapolated to below the ground at high-elevation stations. Quantities in layers between the central facility height and the highest station are linearly interpolated. Data at the central facility is used only to form the area-mean quantities, not in the calculation of the divergences and gradients. The top of the atmosphere is set at 50 mb in this study, where vertical velocity is assumed to be zero.

TABLE 1. Measurement uncertainties in the sounding data.

Temperature	±0.2 K
Winds	±0.5 m s^{-1}
Mixing ratio	2% of the climatological vertical profile when relative humidity is less than 90% and 3% when relative humidity is larger than 90%

d. Error specification

Errors in the preprocessed variables include instrument and measurement uncertainties, errors from interpolation of missing data, and aliasing errors of small-scale features. Manufacturers' specifications of instrument errors in balloonborne soundings are listed in Table 1 (M. Wesely 1996, personal communication). The values are tentative in nature. Note that the uncertainty associated with winds is not treated as height dependent in this study because the supplied error specifications are height independent. The largest uncertainty is probably associated with the humidity. Humidity sensors tend to be slow in reaching relative humidities above 90% in clouds and recover slowly when exiting clouds. At temperatures less than −20°C, the humidity sensor may be contaminated by ice. In addition, radiative heating of sensors increases the uncertainties.

Errors of interpolated data at missing grids should carry a large uncertainty. This uncertainty is estimated by using the standard deviations of all available soundings. As expected, the largest variability is associated with winds. Error variances are therefore specified at the missing grids using standard deviations. This is not applied to the temperature and humidity fields in the current analysis. At all other grids, instrument errors listed in Table 1 are used as error variance. It should be kept in mind that actual errors in the data could be larger than manufacturers' instrument errors. We also point out that what matters to the analysis scheme is the relative magnitudes of error specifications among winds, temperature, and moisture rather than the absolute magnitudes of the specified error uncertainties. Therefore, the above specifications of uncertainties can be viewed as specifications of relative magnitudes of errors.

4. Experimental design

To assess the impact of different conservation constraints on the analysis of state variables and on the vertical velocity and advective tendencies, we describe results from three analysis schemes. They are listed as AS1, AS3, and AS5 in Table 2. Also listed in the table is the preprocessed data AS0. They differ in the number of state variables for adjustments and in the number of

TABLE 2. Experimental design of data analysis schemes.

ID	Input	Variable acted	Action
AS0	Raw sounding data	Winds, temperature, humidity	Missing soundings are quality controlled and objectively interpolated
AS1	AS0 product	Winds	Constraint of mass conservation only
AS3	AS0 product	Winds, temperature, humidity	Constraints of mass, moisture, and static energy conservations
AS5	AS0 product	Winds, temperature, humidity	Constraints of mass, moisture, static energy, and momentum conservations

constraints imposed. In the numerical calculation, the number of adjustment variables is reflected in (13), and the number of constraints is reflected in the choice of (14)–(17).

Figure 3 gives an example of the time–height cross sections of the state variables at the boundary facility B1 (see Fig. 2) in the preprocessed AS0 product. The shaded area represents missing sounding data. In AS1, we adjust winds to impose mass conservation, which is widely used in other objective analysis schemes. In AS3, we constrain winds, temperature, and humidity to conserve mass, moisture, and dry static energy. In AS5, we impose further constraints of momentum conservation. The momentum constraints are included for possible use of the analysis data in studying momentum transport of subgrid motions.

5. Results

a. AS1 scheme

Figure 4 shows the time–height distributions of adjustments made to winds to conserve column integrated mass [Eq. (14)] at the three boundary facilities (B1, B4, B5; see Fig. 2). The adjustment magnitude is typically less than 1 m s^{-1}. At some points, it exceeds 2 m s^{-1}. These are associated with interpolated data corresponding to missing soundings. It is seen that wind adjustments are made simultaneously at the three boundary stations, for example, around day 105 and day 116. This is consistent with the minimum adjustment requirement.

Since weighting coefficients vary only between missing and nonmissing grids, the wind adjustments are almost uniform vertically. To the divergence field, this is

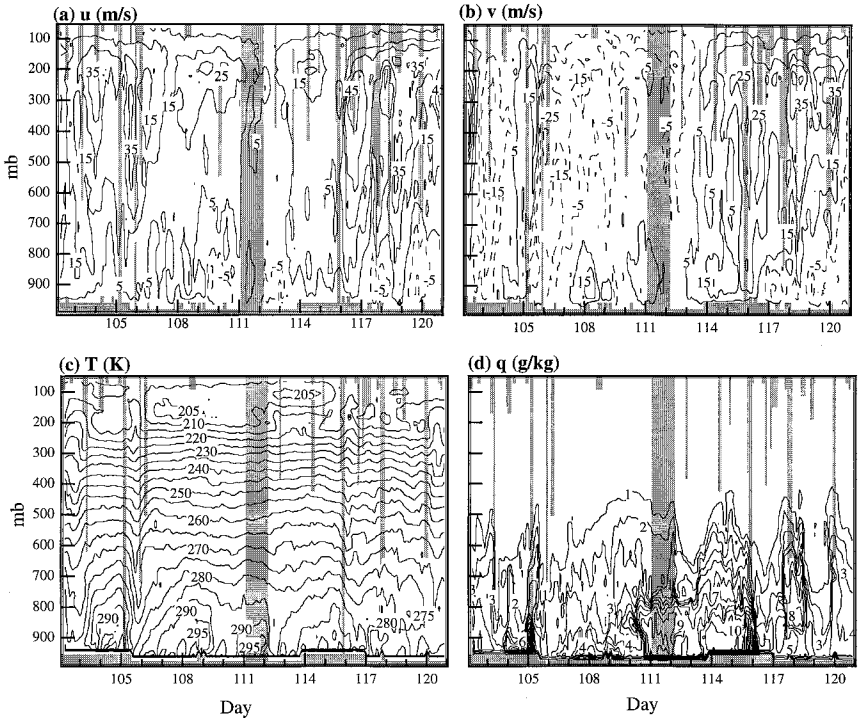

FIG. 3. Time–height distribution of winds (u, v components), temperature, and water vapor mixing ratio in the April IOP at boundary facility B1 from the preprocessed data (AS0). Shaded region represents missing sounding data.

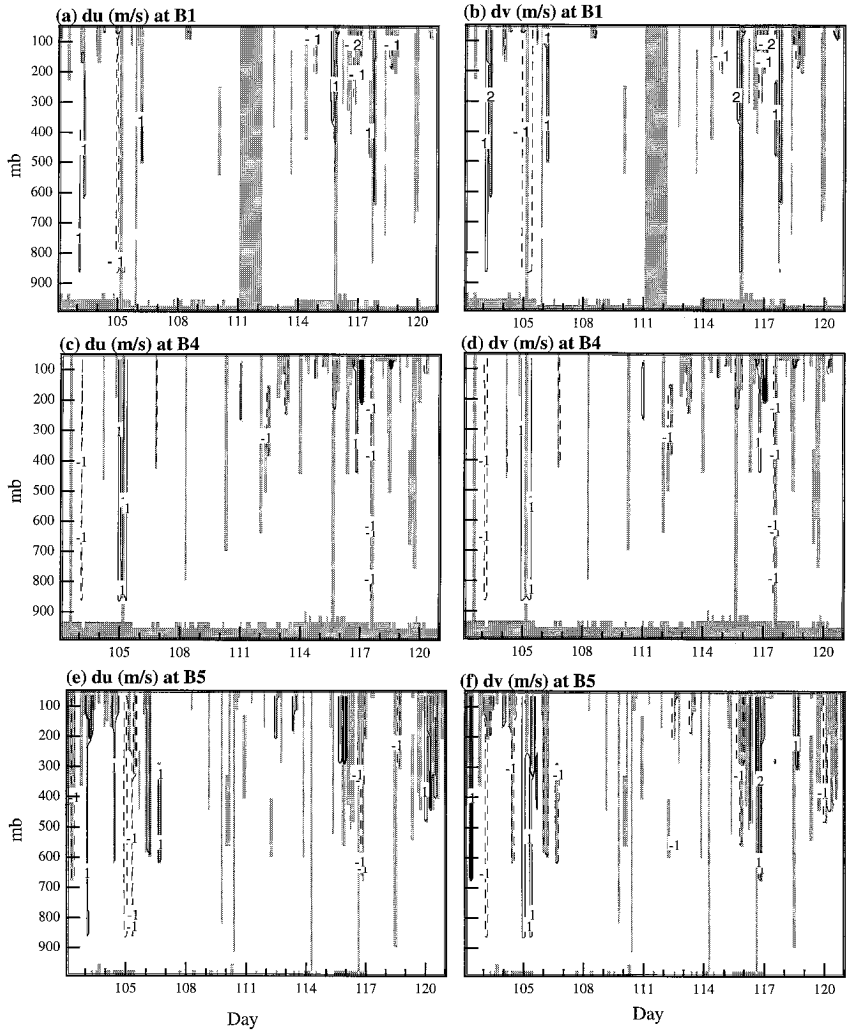

FIG. 4. Distributions of wind adjustments (u, v) at the three boundary facilities (B1, B4, B5) to conserve column integrated mass in the AS1 product. Shaded region represents missing sounding data.

equivalent to a uniform adjustment at all heights. It should be noted, however, that the traditional way of adjusting the divergence field does not give the wind adjustments.

Mass conservation in AS1 is expected to result in better conservation of other quantities. This is found to be true for the column-integrated dry static energy. Figure 5a compares the spurious residual heat source in the AS0 and AS1 products. The AS1 product greatly improves the conservation of dry static energy, although the actual magnitude of the residual term is still very large in comparison with other surface energy fluxes when they are plotted on a different scale.

Figure 5b shows the spurious residual moisture source in the AS0 and AS1 products. It is seen that the mass conservation in AS1 produces little improvement in the moisture conservation. The impact of mass conservation on other budgets can be understood by examining the horizontal advection term. If a quantity is vertically uniform, then mass flux compensation at different layers results in a compensation of the quantity flux. Since dry static energy is relatively more uniform in the vertical direction than is the water vapor mixing ratio, mass conservation better improves the budget of dry static energy than water vapor.

b. AS3 scheme

Winds, dry static energy, and moisture at the three boundary stations are all adjusted to conserve the column-integrated mass, water vapor, and dry static energy. Figure 6 shows the distribution of adjustments made to the u and v components at the three stations. Features to be noted are as follows. First, the magnitude of adjustment is generally within 1 m s^{-1} at nonmissing grids but could be as large as 3 m s^{-1} at missing grids with

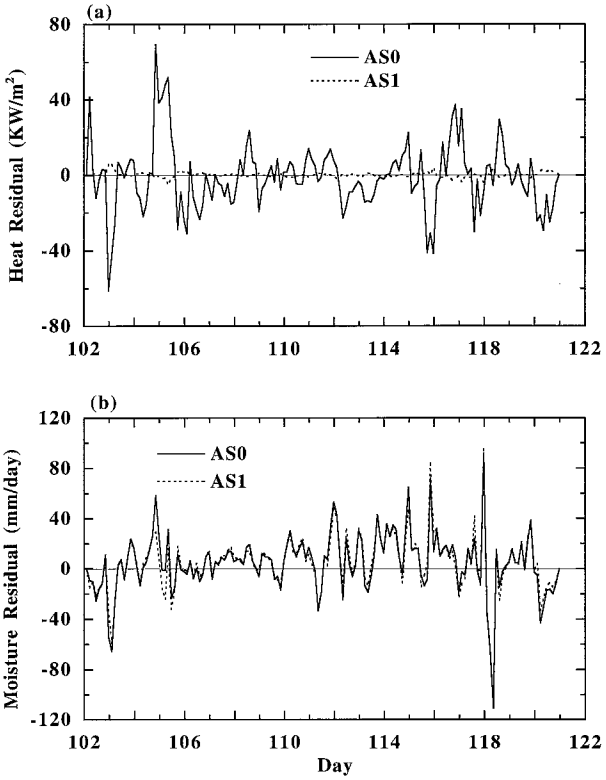

FIG. 5. Residuals diagnosed from column-integrated budget in the AS0 and AS1 products. (a) Residuals of the dry static energy, (b) residuals in the moisture budget. Column-integrated mass is conserved in AS1, not in AS0.

interpolated data. Second, the adjustment varies vertically because of the vertical distribution of moisture and dry static energy. Third, wind adjustment appears in the lower troposphere because relatively small adjustments to winds here can make effective adjustments to the moisture budget.

At boundary station B1, there is a period of continuous missing data in day 111. The adjustments of winds in Figs. 6a and 6b have coherent structures with amplitude of about 3 m s^{-1}. Though these changes in u and v have no impact on the mass conservation, they affect the moisture budget. Moisture varies greatly around that time.

Figure 7 presents the adjustments to dry static energy and the water vapor mixing ratio. For the dry static energy, the magnitude is mostly less than 0.2 K, within instrument uncertainty. For the moisture, the maximum centers are generally within 0.6 g kg^{-1} in the low layers of the troposphere. Considering the large variability of moisture shown in Fig. 3, this magnitude could be considered as an acceptable measurement uncertainty. Adjustments in Figs. 6 and 7 appear to distribute more uniformly in the vertical direction. While there is no specific reason why this should be the case, there are plausible causes: errors due to missing soundings, aliasing of subscale features in the grid-scale fields, and time delay of balloon launches, all of which could be more uniform in the vertical direction.

Since moisture and dry static energy are conserved in the AS3 product, it is meaningful to examine relationships among the budget components. For comparison, we also show the corresponding relationships in the AS1 product. Figure 8 plots the local variability of column moisture against the integrated horizontal moisture transport. The time variability term is large, comparable in magnitude with the moisture convergence term. This is consistent with findings of several early studies using other field data (e.g., Thompson et al. 1979; Lin and Johnson 1996). In the AS1 product, there is no systematic relationship between moisture transport and local moisture variability. But in AS3, a good relationship exists between moisture divergence and reduction of local moisture. Deviation of local moisture variability from the moisture convergence is primarily related to precipitation and, to a much smaller degree, to surface evaporation. It is seen that when moisture convergence exceeds 40 mm day^{-1}, local moisture variability never exceeds 40 mm day^{-1}.

Figure 9 shows the plot of local variability of integrated dry static energy against its horizontal convergence. Again, there is no clear relationship in AS1. But there is a relationship in the AS3 product. Convergence of air in upper layers, accompanied by downward motion, typically increases the column dry static energy because of larger dry static energy in upper layers. This corresponds to a convergent increase of the local dry static energy in the figure. Divergence in upper layers, accompanied by upward motion, typically decreases the column dry static energy; this can be compensated for, however, by latent heat released in the upward motion, as seen in Fig. 9b. The differences between the AS1 and AS3 products shown in Figs. 8 and 9 suggest limited utility of the AS1 product for interpreting the relationships among the budget components.

c. AS5 scheme

Adjustments are made to the same variables as in AS3, but two more constraints are imposed to conserve momentum in the u and v directions. The motivation for enforcing such conservation is twofold. The first is to understand the sensitivity of the adjustments to these additional constraints. The second is to produce a dataset that can be used to evaluate the subgrid-scale momentum forcing.

Figure 10 shows the distribution of wind adjustments to the AS0 product in AS5 at the three boundary stations. Comparing it with Fig. 6, we see that there is hardly any difference in the wind fields between the AS3 and AS5 products. Figure 11 shows the adjustments of dry static energy and water vapor mixing ratio to the AS0 product. Comparing it with Fig. 7 of the AS3 product, we find that, as expected, there is little difference in the moisture field. The main difference appears in the

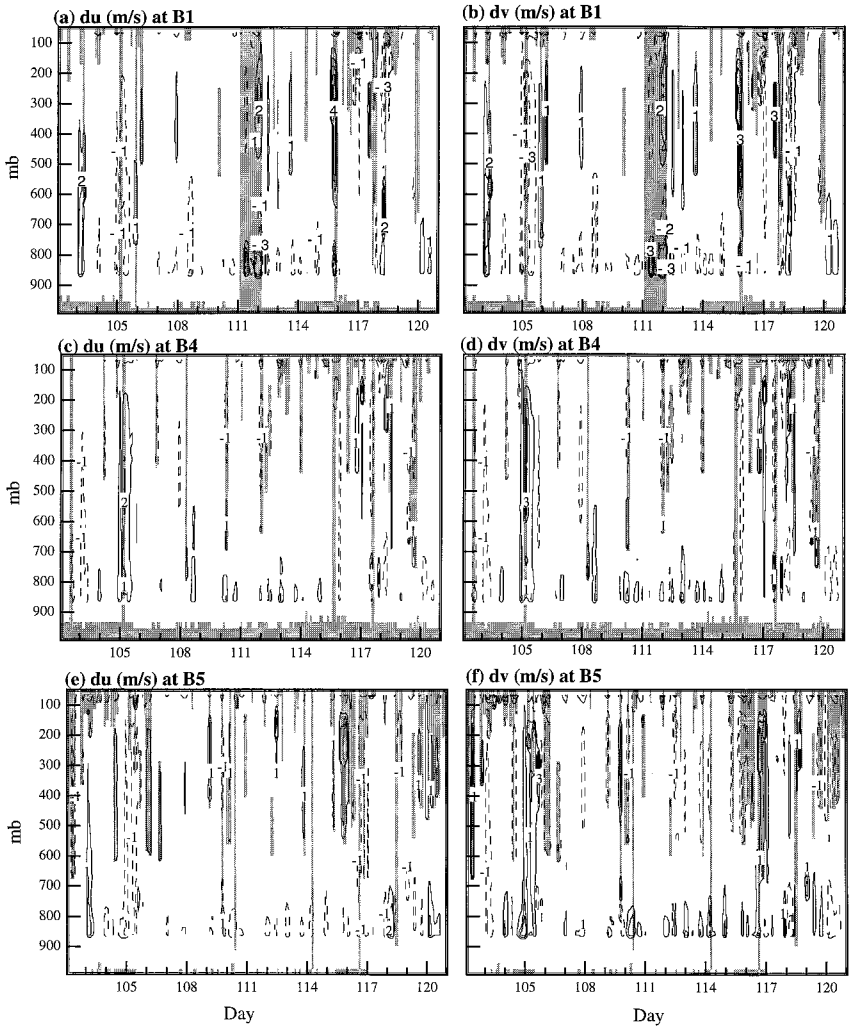

FIG. 6. Distributions of wind adjustments (u, v) at the three boundary facilities (B1, B4, B5) to conserve column-integrated mass, moisture, and dry static energy in the AS3 product. Shaded region represents missing sounding data.

temperature field. Dry static energy, a good approximation of the potential temperature, is mostly adjusted by a magnitude of about 0.2 K. At some time this reaches 0.6 K. This can be considered as within measurement uncertainty. It may be noted that temperature adjustments at different stations differ in magnitude and sign, which constitutes an effective adjustment to the pressure gradient force.

The spurious momentum source in the AS0 and AS3 products is shown in Fig. 12. It is seen that conservations of mass, moisture, and dry static energy do not have a significant impact on the momentum budget. The spurious residual could be as large as 10 m s^{-1} day^{-1}. The temperature adjustment required to balance the budget is less than 0.2 K at 65% of the grids, less than 0.6 K at 93% of the grids, and less than 1 K at 99.4% of the grids. Maximum temperature adjustment is 1.44 K. This sensitivity of the momentum budget to temperature has an important implication to the coupling of dynamical process and physical parameterization in a large-scale model. If spatially inhomogeneous errors are produced in the thermodynamic fields from the physical parameterization, they could induce large errors in the pressure gradient force affecting the dynamical circulation that further influence the physical parameterization.

Time–height distributions of winds, dry static energy, and water vapor mixing ratio in the AS5 product are shown at station B1 in Fig. 13. This is to be compared with the preprocessed distribution in Fig. 3. There is little visible variation in the fields before and after the variational analysis. Yet the AS5 product satisfies all conservation requirements.

The time series of moisture budget components in the AS5 product is presented in Fig. 14. Storage of liquid water in clouds is not plotted because it is much smaller

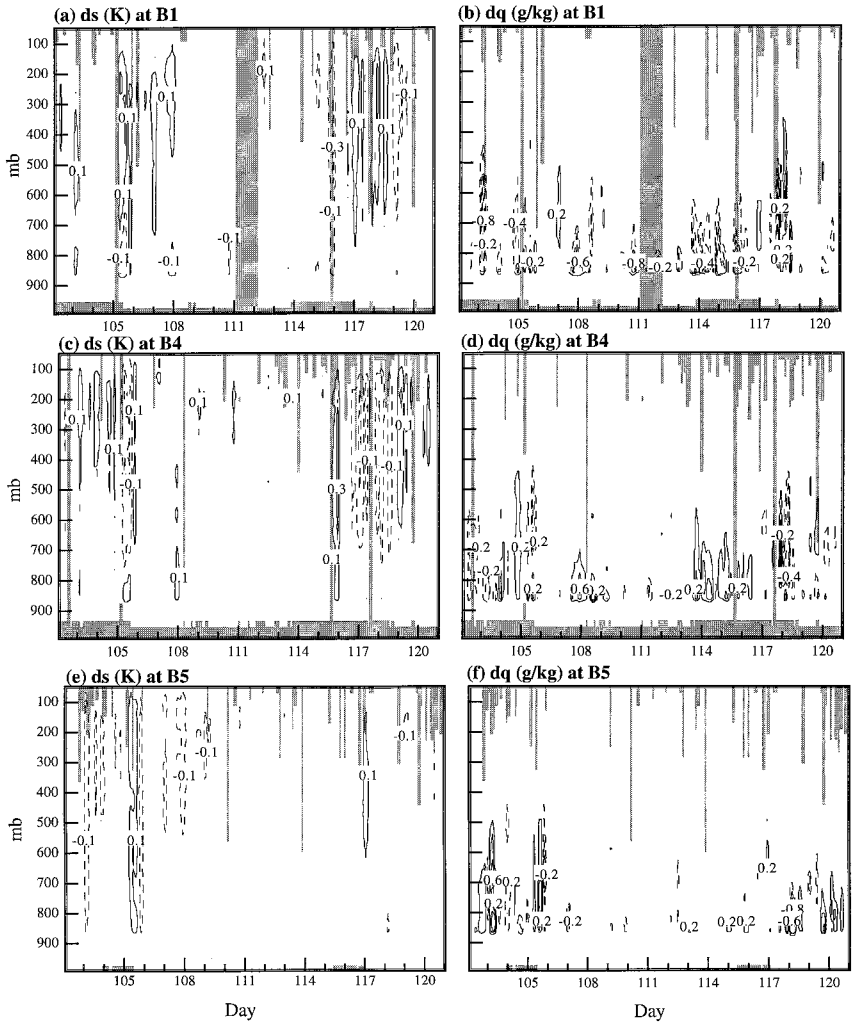

FIG. 7. Adjustments of dry static energy and water vapor mixing ratio at the three boundary facilities (B1, B4, B5) in the AS3 product. Shaded region represents missing sounding data.

than all other terms. The two largest terms are precipitation and horizontal moisture convergence. They balance each other to the first order during the peak events. But moisture variability is also large. There are times when moisture transport is primarily balanced by its time variability, such as from day 103 to day 106. On day 106, moisture convergence does not correspond to precipitation. Instead it corresponds to a large increase of moisture storage, the evidence for which can also be found in Fig. 3d. Another unusual feature appears on day 115, when precipitation corresponds to column moisture divergence and a large reduction in moisture storage.

Figure 15 shows the components of the dry static energy budget. Leading terms are the latent heat and the horizontal energy transport. The local storage term is also much larger than the surface sensible heat flux and net radiation. There are strong diurnal cycles in the latter two terms. Precipitation tends to suppress the surface sensible heat flux.

Components of column momentum budget are shown in Figs. 16 and 17. The two leading terms are the pressure gradient force and the Coriolis force, which are in balance to the first order. The horizontal transport and the local variability may also be large. They balance the ageostrophic component. Since surface wind stress is one order of magnitude smaller than these components, it is not plotted here.

6. Impact on vertical velocity and other diagnostic variables

Impacts of the analysis scheme on several diagnostic variables are now examined. These are important variables for the description of physical processes within the atmospheric column. The first is the large-scale ver-

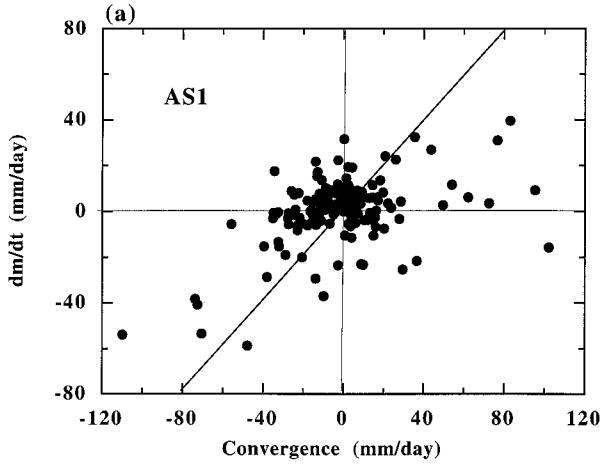

FIG. 8. Relationships between variation of column moisture storage and horizontal moisture transport, (a) in the AS1 product and (b) in the AS3 product.

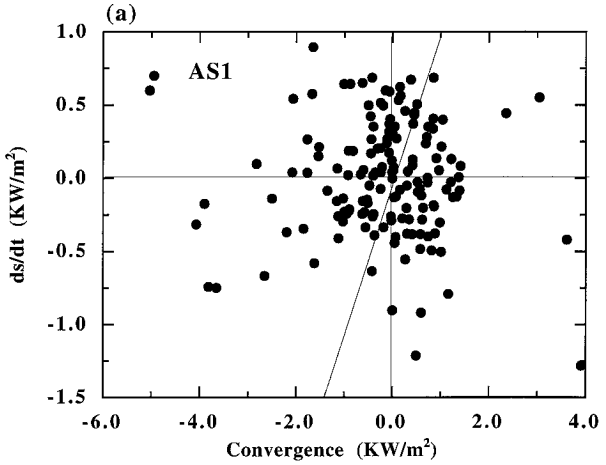

FIG. 9. Relationships between variation of column storage of dry static and its horizontal transport, (a) in the AS1 product and (b) in the AS3 product.

tical velocity calculated from the divergence field using surface boundary conditions. Figure 18 compares the vertical velocity in the preprocessed product AS0 with those in AS1, AS3, and AS5. In AS0 (Fig. 18a), integration of the horizontal divergence produces a vertical velocity that does not vanish at the upper boundary. Constraint of mass conservation (Fig. 18b) greatly modifies the structure of the vertical velocity. Maximum values are seen in the middle troposphere. Unlike findings from GATE (Yanai and Johnson 1993) for the tropical Atlantic, both upward and downward motions appear to peak at one altitude in the middle troposphere in Oklahoma. The magnitude of the maximum upward velocity exceeds 1 Pa s^{-1} (36 mb h^{-1}); it appears to be larger than the magnitude of maximum downward velocity. Furthermore, moderately strong upward motions are concentrated in a few short time periods, but moderately strong downward motions occur very often. Around day 117, upward motion exists in the lower troposphere below 450 mb, but downward motion exists above it; thus the maximum divergence is in the middle troposphere.

The vertical velocity in AS3 (Fig. 18c) shows many features that are different from those in AS1. Downward motions are significantly weaker in AS3. Strong upward motion appears in day 112 and day 116. Structures of upward motion from day 117 to day 120 are also changed. Vertical motion in day 112 is associated with the observed precipitation (Fig. 1). Since original soundings are missing in that day at one of the boundary facilities (see Fig. 4), the interpolated data may not describe the actual divergence field. There are also missing data on day 116 when precipitation occurred. In the AS1 product, vertical velocity relies heavily on the interpolation scheme when there are missing data. On the other hand, the AS3 product has used precipitation and other surface and TOA measurements, and thus the corresponding vertical velocity is not as sensitive to the interpolation scheme as it is in the AS1 product. Vertical velocity in the AS5 product is nearly the same as that in AS3. This is consistent with the notion that few wind

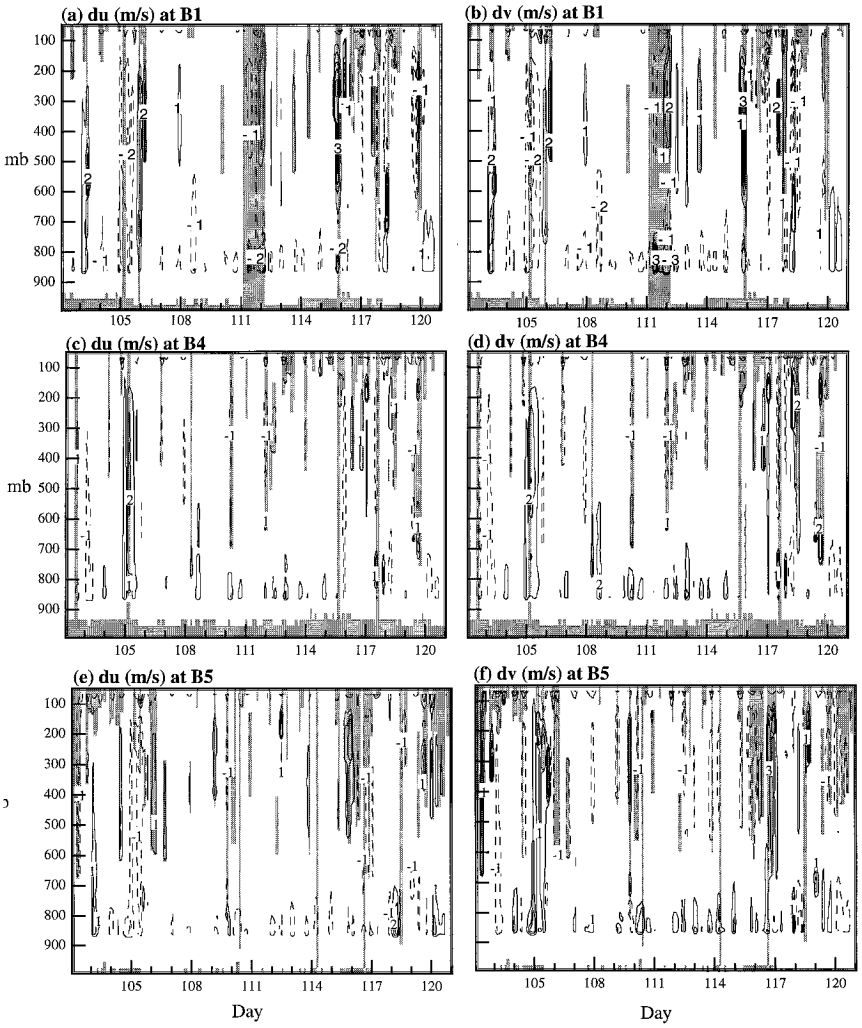

FIG. 10. Distributions of wind adjustments (u, v) at the three boundary facilities (B1, B4, B5) to conserve column-integrated mass, moisture, and dry static energy and momentum in the AS5 product. Shaded region represents missing sounding data.

adjustments are required to preserve the momentum budgets.

Next we examine the apparent heat source and moisture sinks Q_1 and Q_2 as introduced in Yanai et al. (1973). They are defined as

$$Q_1 = \frac{1}{C_p}\left(\frac{\partial s}{\partial t} + \mathbf{V}\cdot\nabla s + \omega\frac{\partial s}{\partial p}\right), \tag{27}$$

$$Q_2 = -\frac{L}{C_p}\left(\frac{\partial q}{\partial t} + \mathbf{V}\cdot\nabla q + \omega\frac{\partial q}{\partial p}\right). \tag{28}$$

Many previous studies have diagnosed these variables using sounding measurements from field experiments (e.g., Nitta and Esbensen 1974; Esbensen 1978; Frank 1979; Johnson and Young 1983; Kuo and Anthes 1984; Gallus and Johnson 1991). Figure 19 shows Q_1 in the four data products in this study. Mass conservation in AS1 does not strongly modify the Q_1 distribution in the middle and lower troposphere, although it improves the integrated balance of the dry static energy as shown in Fig. 5. This is because mass conservation primarily modifies Q_1 near the top of the troposphere where large vertical velocity in AS0 has been modified to zero in AS1. Diurnal variation can be seen below 850 mb, presumably reflecting vertical diffusion associated with diurnal variability of surface heat fluxes. Maximum heating centers appear around day 102, day 106, day 116, day 118, and day 120, with the last three events coincident with precipitation. Maximum value of apparent heating exceeds 1.5° per hour.

In AS3, Q_1 (Fig. 19c) has several noteworthy features that differ from those in AS1. Some scattered features in Fig. 19b, such as the cooling pattern from 600 to 800 mb in the middle of the IOP, have disappeared in AS3. The most important change from Q_1 in AS1 to that in

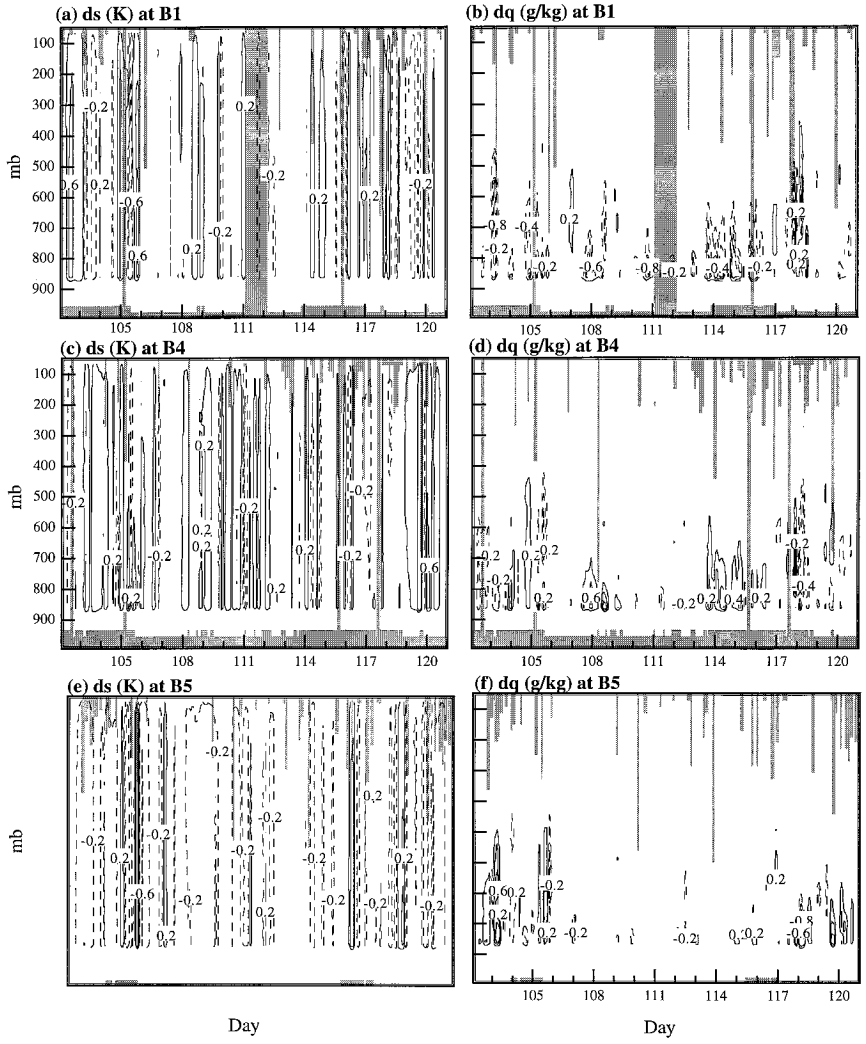

FIG. 11. Adjustments of dry static energy and water vapor mixing ratio at the three boundary facilities (B1, B4, B5) in the AS5 product. Shaded region represents missing sounding data.

AS3 is probably the significant reduction of heating in day 103 and day 106 along with the appearance of a new heating center in day 112. We note that there is no precipitation observed on days 103 and 106, but precipitation is observed on day 112. There are also differences in the heating pattern from day 117 to day 120. It is noted that in most of the events, Q_1 peaks at 500 mb; there is one event on day 118 when Q_1 peaks at around 400 and 750 mb. The distribution of Q_1 in AS5 is similar to that in AS3. This implies that constraining the momentum does not affect the diagnostics of the apparent heat source.

Apparent moisture sink Q_2 in the four data products is shown in Fig. 20. Consistent with the results shown in Fig. 5, mass conservation in AS1 has little impact on the moisture budget. Large moisture sinks (positive area) are seen with the precipitation events, while moisture sources (negative areas) are seen in the lower troposphere. Figure 20c shows that Q_2 in AS3 differs from Q_2 in AS1. Several moisture source centers in AS1 become much weaker in AS3. Moisture sinks on day 103 and day 116 are reduced, while those in day 112 and day 118 are amplified. In AS5, Q_2 is basically the same as in AS3.

Following Yanai et al. (1973), subgrid vertical transport of most static energy can be diagnosed from

$$Q_1 - Q_2 - Q_{\text{rad}} = -\frac{1}{C_p}\frac{\overline{\partial \omega'(s' + Lq')}}{\partial p}. \quad (29)$$

Since we do not have the radiative heating rate, we examined $Q_1 - Q_2$, keeping in mind the contribution of radiative heating (not shown). The diurnal cycle is more pronounced near the surface, indicating the subgrid-scale nature of the vertical eddy transport near the surface. Positive heating centers are always associated with negative cooling centers below. Furthermore, there are many days when this quantity appears to be very

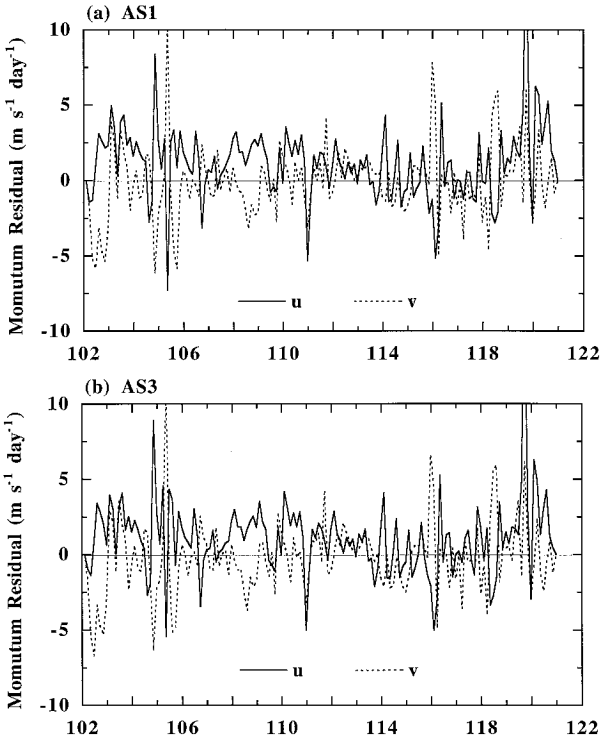

FIG. 12. Momentum residuals diagnosed from column-integrated u and v momentum budgets, (a) in the AS0 product and (b) in the AS3 product.

weak and smooth in the AS3 and AS5 products, indicating relatively good data quality.

The last variable we present here is the apparent momentum source. We show it only for the u component. It is defined as

$$F_x = \frac{\partial u}{\partial t} + \mathbf{V}\cdot\nabla u + \omega\frac{\partial u}{\partial p} - f v + \frac{\partial \phi}{\partial x}. \qquad (30)$$

Previous studies of the subgrid momentum forcing include LeMone et al. (1984), LeMone and Moncrieff (1994), Sui and Yanai (1986), Gallus and Johnson (1992), and Wu and Yanai (1994). Figure 21 shows the time–height distribution of F_x in the four data products. A very pronounced feature is that F_x does not change much among AS0, AS1, and AS3. Thus conservations of mass, moisture, and momentum have little impact on the momentum forcing. Westward forcing near the surface up to 900 mb appears for periods when the westerly reaches ground (see Fig. 3a); eastward forcing occurs when the easterly reaches ground, with magnitude larger than 1 m s^{-1} h^{-1}. When the momentum budget is enforced, Fig. 21d shows that the distribution of subgrid forcing is changed. In the precipitation events, there appear an eastward acceleration below 400 mb and a westward acceleration above that level, with magnitude exceeding 4 m s^{-1} h^{-1}. It is not clear why there is a westward acceleration around day 111. The magnitude of subgrid-scale momentum forcing is consistent with

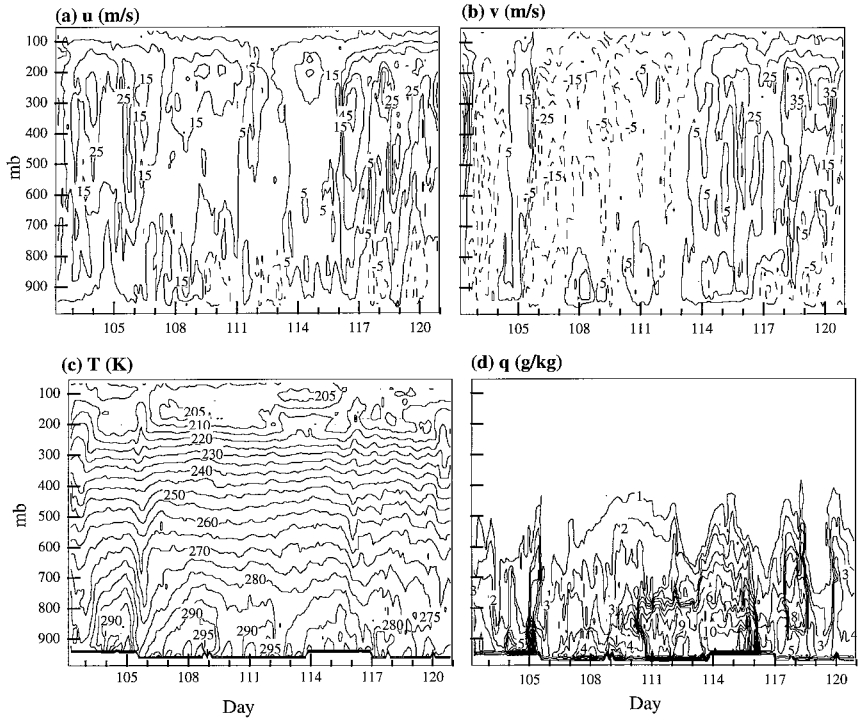

FIG. 13. Time–height distribution of winds (u, v components), temperature, and water vapor mixing ratio in the April IOP at boundary facility B1 in the AS5 product.

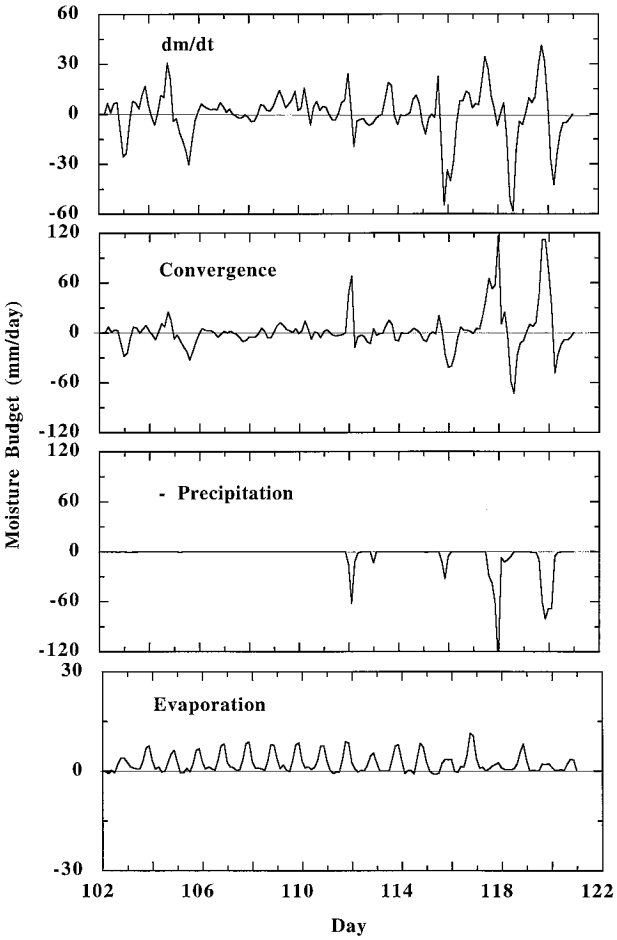

FIG. 14. Time variation of the budget components of column-integrated moisture. (a) Variation of storage, (b) horizontal transport, (c) precipitation, and (d) evaporation.

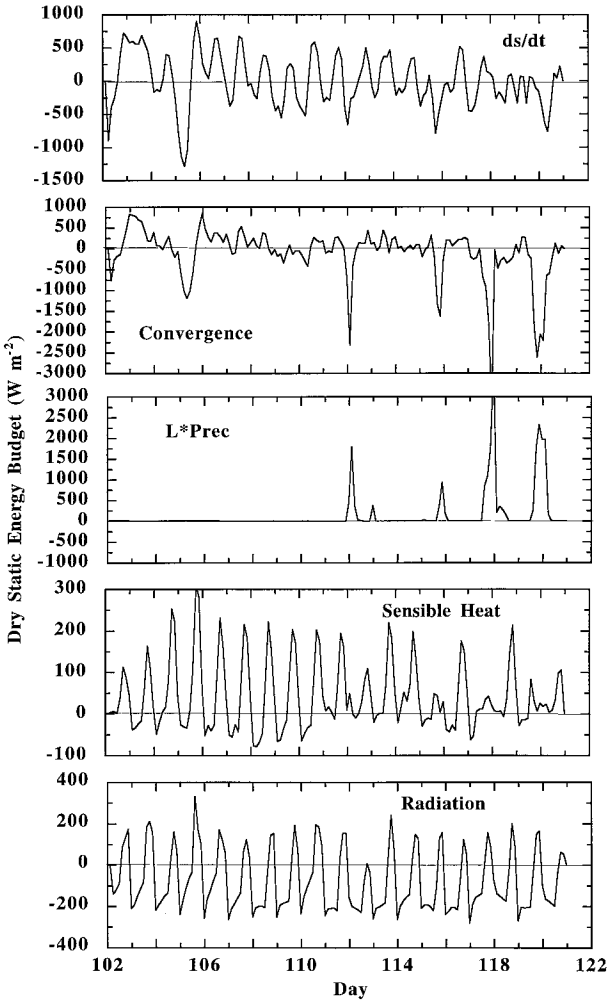

FIG. 15. Time variation of the budget components of column-integrated dry static energy. (a) Variation of storage, (b) horizontal transport, (c) latent heat, (d) sensible heat, and (e) net radiation.

previous reports from aircraft measurements (LeMone et al. 1984).

7. Discussion of the schemes

As with all objective analysis methods, there are numerous ways by which the method described in the present study can be possibly improved. The first aspect that warrants further study is the specification of error uncertainties. These uncertainties become the weighting factors in the cost function of (11). As mentioned in section 3d, what matters to the objective analysis is the relative magnitude of errors among winds, temperature, and moisture and their spatial distributions. The perfect knowledge of this information never exists; otherwise there would be no need to perform the objective analysis. Another difficulty of the problem is the error covariances in time and space for a single variable and covariances between different variables. This is especially true when missing soundings are interpolated or data are preprocessed through a statistical interpolation procedure. It is conjectured that these problems associated with error specifications have larger impact on the adjustments of atmospheric state variables (which are small anyway) than on the diagnosed vertical velocity and advective tendencies, since the latter are constrained by the surface and TOA measurements of fluxes. A comprehensive sensitivity study of the final data product using various forms of weighting factors in (13) may help to shed more light on this issue.

The second aspect concerns the smoothness of the analyzed fields. The idealized goal here is to extract the true information of the grid-scale fields without measurement and instrument errors. One of the formidable tasks is then to de-alias small-scale features from the undersampled sounding network (Ooyama 1987). We note that the method described in this study has a unique advantage of automatically de-aliasing small-scale features from the instantaneous soundings, since area-averaged constraint variables are used to construct the

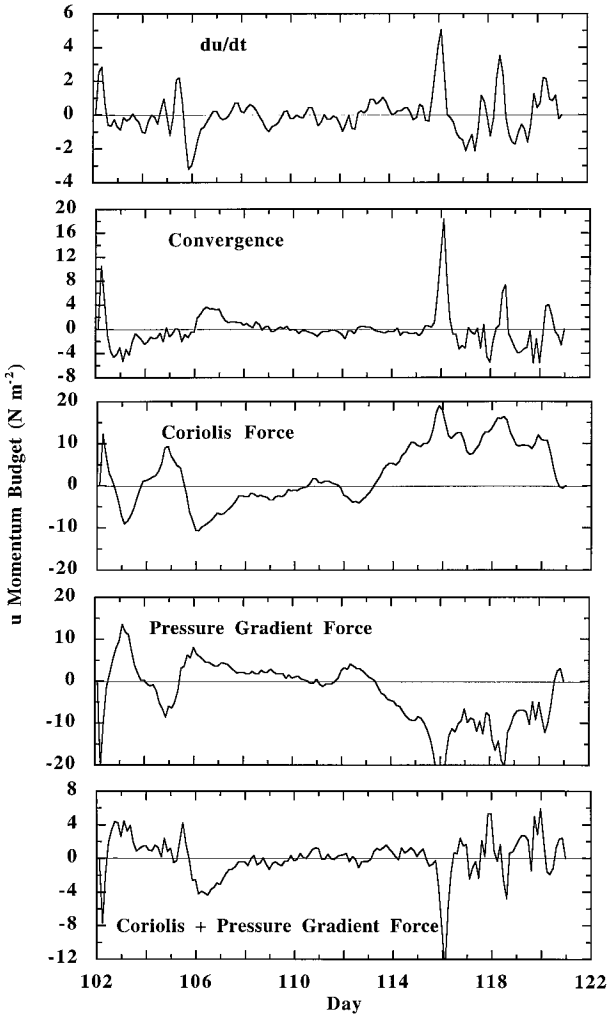

FIG. 16. Time variation of the budget components of column-integrated u momentum. (a) Storage variation, (b) horizontal transport, (c) Coriolis force, (d) pressure gradient force, and (e) sum of pressure gradient force and Coriolis force.

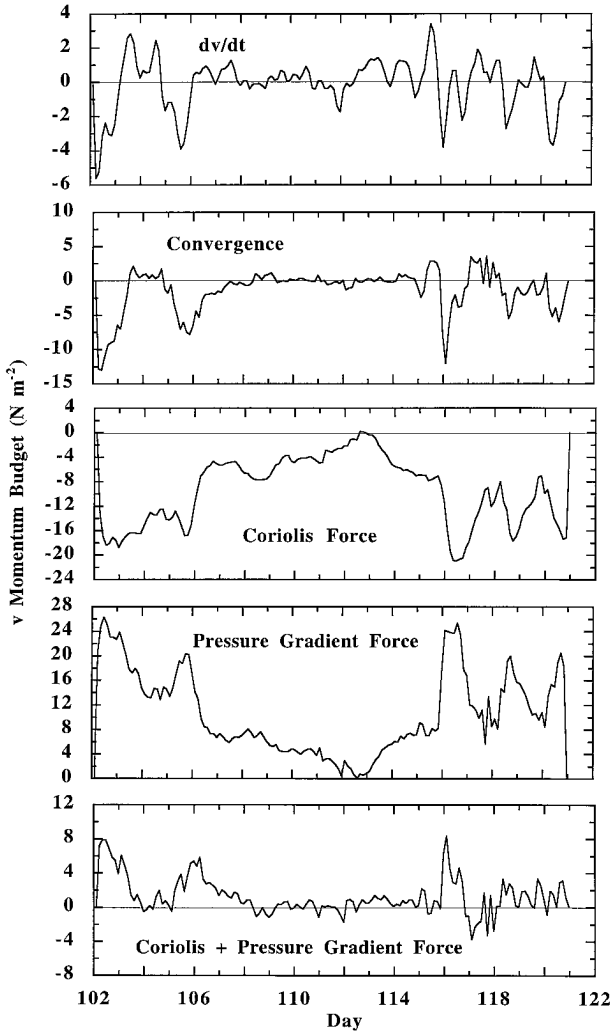

FIG. 17. Same as Fig. 16 except for the v component.

area-averaged large-scale diagnostics. The product should be more reliable when sufficient surface and TOA measurements are available, which is the case in ARM. Yet the method does not dealias data in time and in the vertical direction. Possible improvements include vertical smoothing and time filtering of data in the preprocessing step.

The approach we discussed can be also technically extended. One way is to include more variables for adjustment, such as surface pressure, temperature, heat fluxes, and cloud liquid water content. This can be achieved by including their adjustments in the cost function of (13). Analysis variables can be used to replace prescribed variables in the constraints. Another aspect is to include more observational data such as those from wind profilers. The procedure would include added adjustments to the cost function with constraints remaining the same. Additional constraints, such as narrowband TOA radiative fluxes, can be introduced as either strong constraints or weak constraints. These extensions have the potential to make the maximum use of available observational data. Water substance in the form of hydrometeors should be considered when measurements become available.

It is possible to relax the strong constraints in the current study to weak constraints in order to consider numerical errors. Minimization of cost function of (11) integrated over the whole measurement period instead of at each time level can also be considered. Furthermore, a more rigorous preprocessing procedure can possibly further improve the quality of the derived data. In cases where more than a few sounding stations are available, objective analysis such as those described in Barnes (1964) and Ooyama (1987) can be used in the preprocessing step.

Finally, further sensitivity studies of the data products to various input data and processing procedures are necessary to give an estimate of the accuracy limit in the

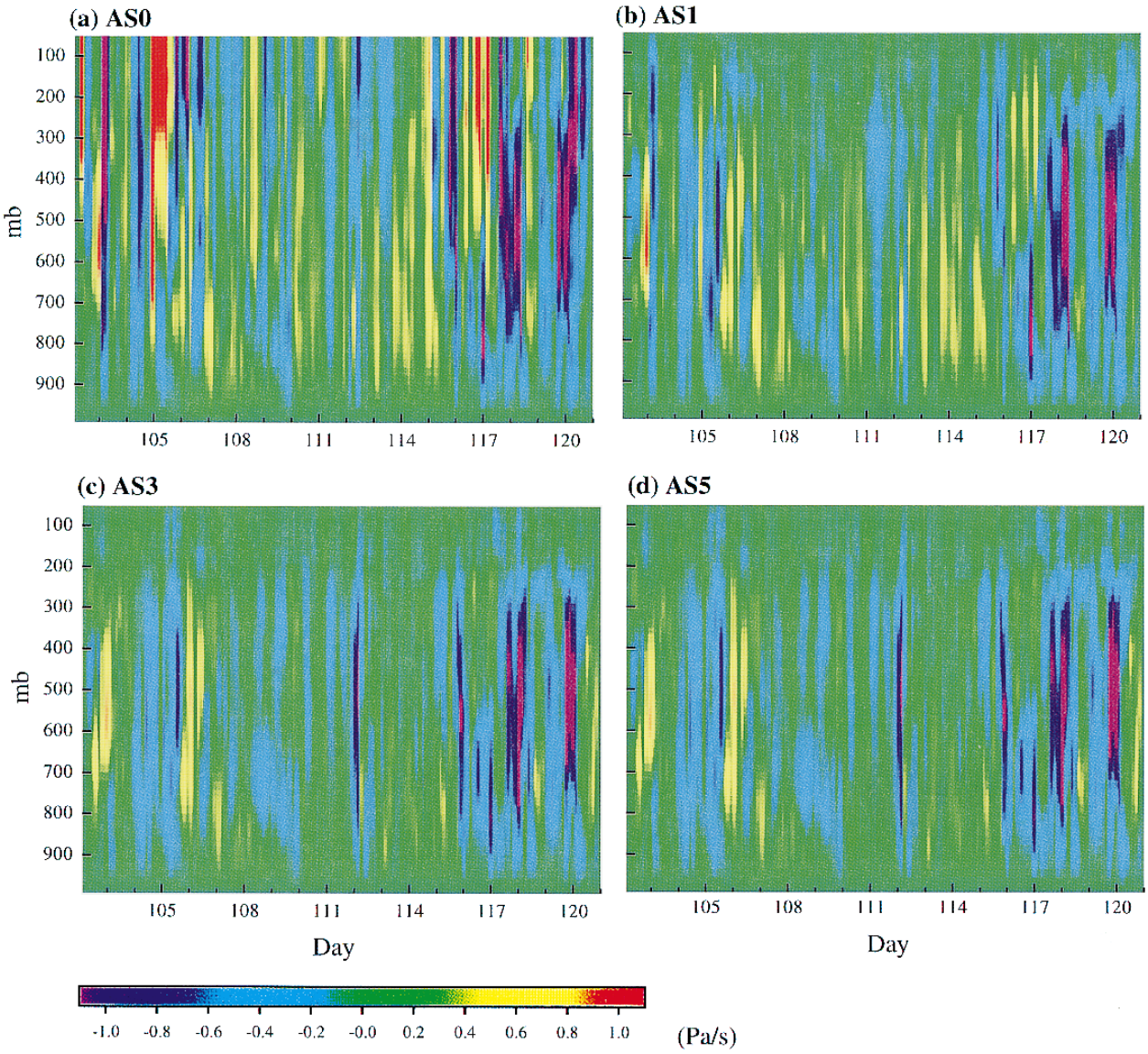

FIG. 18. Time–height distribution of vertical velocity in the four data products.

large-scale diagnostic fields. These will be the subject of future research.

8. Conclusions

We have reported an objective method to analyze sounding measurements over a network of a small number of stations. The scheme has been implemented to process the ARM IOP measurements at the boundary facilities around the central facility in Northern Oklahoma. The essence of the method is, accounting for uncertainties in sounding measurements, to make the smallest adjustments to original soundings of winds, temperature, and water vapor mixing ratio so that the adjusted data conserve column-integrated mass, moisture, energy, and momentum. Input data include not only measurements of the adjustment variables, but also measurements of constraint variables over a large number of surface stations and at the TOA. As a result, the scheme tends to make full use of available measurements and is independent of any particular model.

Our results are summarized as follows.

1) Column-integrated conservations of mass, moisture, static energy, and momentum can be achieved in the analysis when state variables are adjusted by an amount comparable to their measurement uncertainty. Temperature adjustment to the preprocessed field, including original measurement and interpolated data, is less than 0.4 K at 85% of the grids and less

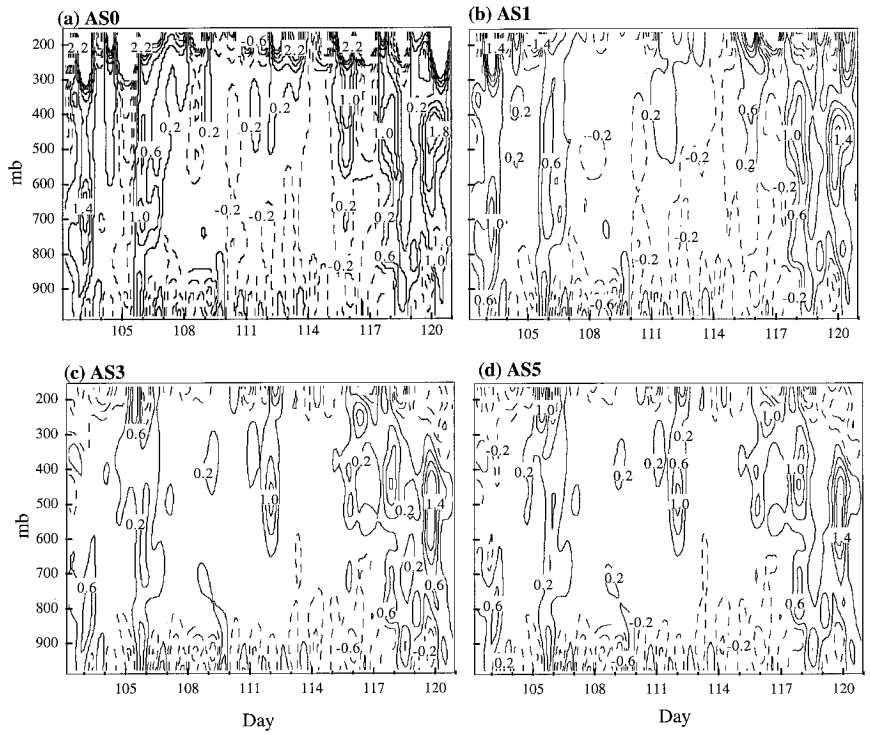

FIG. 19. Time–height distribution of the apparent heat source Q_1 in the four data products. The units are kelvin per hour.

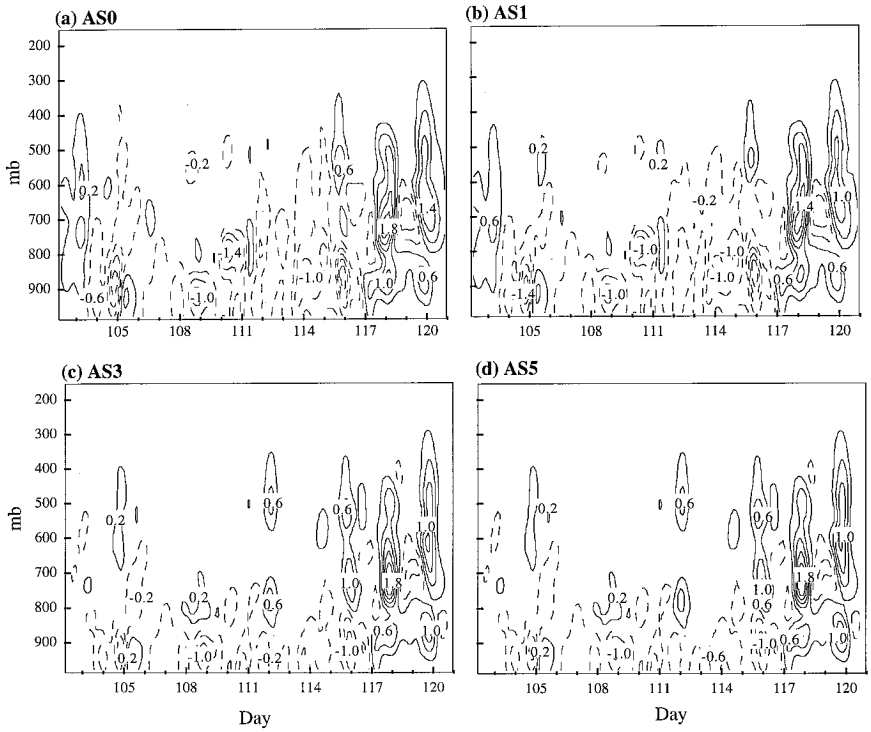

FIG. 20. Time–height distribution of the apparent moisture sink Q_2 in the four data products. The units are kelvin per hour.

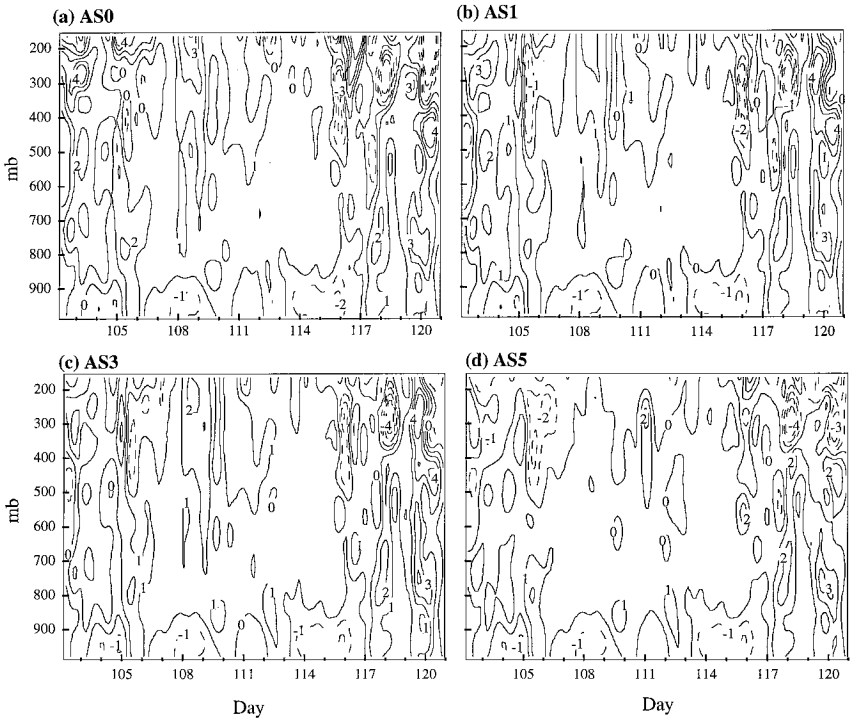

FIG. 21. Time–height distribution of the apparent momentum forcing F_x in the four data products. The units are meters per second per hour.

than 0.8 K at 99% of the grids. Adjustment of mixing ratio is less than 0.6 g kg^{-1} in the lower troposphere at 99% of the grids, and wind adjustments are less than 2 m s^{-1} at 99% of the nonmissing grids. Without such adjustments, the spurious residual source in the budget of each quantity is large and comparable with other leading components.

2) Vertical velocity, apparent heat source, and moisture source, diagnosed from the large-scale fields, are sensitive to the constraints of mass, moisture, and dry static energy conservation. There is evidence that atmospheric diagnostic fields are more realistic with the constraints. Large changes are coincident with periods when some of the sounding data are interpolated. Philosophically, inevitable uncertainties in the sounding data prevent accurate derivation of some atmospheric diagnostics solely from sounding measurements. Additional measurements at surface and at the top-of-the-atmosphere can be effectively used to improve the accuracy of these diagnostics.
3) Constraint of mass conservation alone does not improve the conservation of moisture in the dataset used in this study, although it improves the conservation of dry static energy to a certain degree. This result may be different in the Tropics where state variables are horizontally more homogeneous. Constraints of mass, water vapor, and heat conservation do not improve the momentum budget. Conservation of momentum is realized by small adjustments to the temperature field (less than 0.6 K at 93% of the grids). Adjustments to winds and water vapor are not sensitive to the momentum budget constraint.
4) Analysis of ARM data shows that there are occasions when precipitation corresponds to large reductions of column precipitable water and column-moisture divergence, even though moisture convergence often corresponds to precipitation. Subgrid momentum transport generally appears with precipitation events, and westward acceleration aloft is typically accompanied by eastward acceleration below. The magnitude could be more than 2 m s^{-1} h^{-1}.

Finally, it is noted here that there is no simple way to validate the analyzed data products other than through general discussion of data features and by examining the adjustment magnitudes. Comparisons with assimilation products, such as those from numerical weather prediction models, will also be useful. The data can be tested in physical parameterizations where additional independent observations may be available to serve as indirect validation. A comprehensive sensitivity study can also give insights into the robustness of the final data products. From the results of this study, we advocate the use of additional constraint measurements to improve the accuracy of large-scale diagnostic fields from balloonborne soundings.

Acknowledgments. The authors wish to thank Joyce Tichler at Brookhaven National Laboratory for her help in acquiring several ARM datasets and making a work-

station available to us. We are also grateful to Patrick Minnis at NASA LARC for sending us the document on the GOES cloud product and to Marvin Wesely at ANL for providing us information on the radiosonde and EBBR instruments. Dr. Iqbal Pittalwala made valuable comments on the paper. We thank two of the three anonymous reviewers whose comments have led improvements to the original manuscript. This research has been partly supported by DOE under Grant DEFG0285ER60314 and partly by NASA under Grant NAG3517 to the State University of New York at Stony Brook.

REFERENCES

Barnes, S. L., 1964: A technique for maximizing details in numerical map analysis. *J. Appl. Meteor.,* **3,** 396–409.

Bellamy, J. C., 1949: Objective calculations of divergence, vertical velocity and vorticity. *Bull. Amer. Meteor. Soc.,* **30,** 45–49.

Cotton, W. R., and R. A. Anthes, 1989: *Storm and Cloud Dynamics.* Academic Press, 880 pp.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

Davies-Jones, R. P., 1993: Useful formulas for computing divergence, vorticity, and their errors from three or more stations. *Mon. Wea. Rev.,* **121,** 713–725.

Emanuel, K. A., 1994: *Atmospheric Convection.* Oxford University Press, 580 pp.

Esbensen, S. K., 1978: Bulk thermodynamic effects and properties of small tropical cumuli. *J. Atmos. Sci.,* **35,** 826–837.

Frank, W. M., 1979: Individual time period analysis over the GATE ship array. *Mon. Wea. Rev.,* **107,** 1600–1616.

Gallus, W. A., and R. H. Johnson, 1991: Heat and moisture budgets of an intense midlatitude squall line. *J. Atmos. Sci.,* **48,** 122–146.

——, and ——, 1992: The momentum budget of an intense midlatitude squall line. *J. Atmos. Sci.,* **49,** 422–450.

Ghil, M., and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Advances in Geophysics,* Vol. 33, Academic Press, 141–266.

Johnson, R. H., and G. S. Young, 1983: Heat and moisture budgets of tropical mesoscale anvil clouds. *J. Atmos. Sci.,* **40,** 2138–2147.

Kuo, Y.-H., and R. A. Anthes, 1984: Mesoscale budgets of heat and moisture in a convective system over the central United States. *Mon. Wea. Rev.,* **112,** 1482–1497.

LeMone, M. A., and M. W. Moncrieff, 1994: Momentum and mass transport by convective bands: Comparisons of highly idealized dynamical models to observations. *J. Atmos. Sci.,* **51,** 281–305.

——, G. B. Barnes, and E. J. Zipser, 1984: Momentum flux by lines of cumulonimbus over the tropical oceans. *J. Atmos. Sci.,* **41,** 1914–1932.

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.

Lord, S. J., 1982: Interaction of cumulus cloud ensemble with the large-scale environment. Part III: Semi-prognostic test of the Arakawa–Schubert cumulus parameterization. *J. Atmos. Sci.,* **39,** 88–103.

Minnis, P., W. L. Smith, D. P. Garber, J. K. Ayers, and D. R. Doeling, 1995: Cloud properties derived from GOES-7 for spring 1994 ARM Intensive Observing Period using version 1.0.0 of ARM satellite data analysis program. NASA Ref. Publ. 1366, 59 pp. [Available from NASA/Langley Research Center, Technical Library, MS 185, Hampton, VA 23655-5225.]

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

O'Brien, J. J., 1970: Alternative solutions to the classical vertical velocity problem. *J. Appl. Meteor.,* **9,** 197–203.

Ooyama, K., 1987: Scale-controlled objective analysis. *Mon. Wea. Rev.,* **115,** 2476–2506.

Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate.* American Institute of Physics.

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

Stokes, G. M., and S. E. Schwartz, 1994: The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed. *Bull. Amer. Meteor. Soc.,* **75,** 1201–1221.

Sui, C.-H., and M. Yanai, 1986: Cumulus ensemble effects on the large-scale vorticity and momentum fields of GATE. Part I: Observational evidence. *J. Atmos. Sci.,* **43,** 1618–1642.

Thompson, R. M., S. W. Payne, E. E. Recker, and R. J. Reed, 1979: Structure and properties of synoptic-scale wave disturbances in the intertropical convergence zone of the eastern Atlantic. *J. Atmos. Sci.,* **36,** 53–72.

Wakimoto, R. M., and N. T. Atkins, 1994: Observations of the seabreeze front during CaPE. Part I: Single Doppler, satellite, and cloud photogram analysis. *Mon. Wea. Rev.,* **122,** 1092–1114.

Wang, J., and D. A. Randall, 1996: A cumulus parameterization based on the generalized convective available potential energy. *J. Atmos. Sci.,* **53,** 712–727.

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

Wu, X., and M. Yanai, 1994: Effects of vertical wind shear on the cumulus transport of momentum: Observations and parameterization. *J. Atmos. Sci.,* **51,** 1640–1660.

Xu, Q., 1996: Generalized adjoint for physical processes with parameterized discontinuities. Part I: Basic issues and heuristic examples. *J. Atmos. Sci.,* **53,** 1123–1142.

Yanai, M., and R. H. Johnson, 1993: Impact of cumulus convection on thermodynamic fields. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.,* No. 46, Amer. Meteor. Soc., 36–62.

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

Zou, X., I. M. Navon, and J. Sela, 1993: Variational data assimilation with moist threshold processes using the NMC spectral model. *Tellus,* **45A,** 370–387.