Four-dimensional Variational Assimilation for the Tropical Atmosphere Using Precipitation Data

降水量データを用いた熱帯大気の4次元変分法によるデータ同化

1999

露木 毅

Four-dimensional Variational Assimilation for the Tropical Atmosphere Using Precipitation Data

Tadashi Tsuyuki

Japan Meteorological Agency

1-3-4 Ote-machi, Chiyoda-ku, Tokyo 100-8122, Japan

Doctoral dissertation submitted to Graduate School of Science

University of Tokyo, in February 1999

Abstract

Four-dimensional variational assimilation (4D-Var) is one of the advanced data assimilation methods to which considerable attention has been paid in recent years. However, the application of 4D-Var to the tropical atmosphere has not been much examined due to the difficulty associated with moist processes that is essential for the tropical atmosphere. This is because parameterization schemes of the moist processes in atmospheric models are highly nonlinear and often include several discontinuities, which degrade the efficiency of commonly used minimization algorithms.

This paper investigates the feasibility of 4D-Var for the tropical atmosphere by removing several discontinuities from the moist processes and by including a penalty term for controlling gravity wave level in the analysis. The impact of assimilating precipitation data, which is closely related to the moist processes, with 4D-Var on the tropical analysis is also examined. The cost function consists of a discrepancy term between model and observations and the penalty term, and does not include a background term for simplicity. Physics of the adjoint model that is backward integrated for calculating the gradient of the cost function includes moist processes, horizontal diffusion, and simplified surface friction only.

A column model assimilation experiment is carried out using simulated data to examine the impact of the removal of discontinuities from the moist processes. Then a global primitive-equation model with the smoothed moist processes is used to investigate the effect of the penalty term for accelerating convergence of 4D-Var with the moist processes included. In this experiment, the true state and observations are provided by the full-physics model, while the assimilation model (forward model) and the corresponding adjoint model use the above-mentioned reduced physics. An idealized observation network that is sparse in the tropics and the Southern Hemisphere is adopted to roughly simulate the real observation network. Finally, in order to evaluate the overall feasibility of 4D-Var in the tropics, real observational data are assimilated with the global model. The radiosonde data over the globe and Special Sensor Microwave Imager (SSM/I) precipitation rate data over the tropical oceans are assimilated for the period 0 - 12 UTC 22 August 1992. The assimilation model is the full-physics model and the adjoint model uses the reduced physics.

It is found that 4D-Var for the tropical atmosphere is feasible if the following three procedures are adopted: the appropriate control of gravity wave level, the removal of zeroth-order discontinuities from parameterization schemes of the moist processes, and the use of a higher-order interpolation

operator for precipitation field when assimilating precipitation data. These procedures are helpful to improve the convergence performance of 4D-Var in which the adjoint model includes moist processes. 4D-Var using the adjoint model which lacks the moist processes produces a poor analysis in the tropics even if the full-physics model is used as the assimilation model. Inclusion of the moist processes in the adjoint model leads to a better precipitation analysis even without assimilating precipitation data, although convergence is slightly decelerated by including the moist processes.

The impact of assimilating SSM/I precipitation rates on the precipitation analysis is not confined to near SSM/I observation times, but spreads over the whole assimilation window. Its impact on the precipitable water analysis over the tropical oceans is positive but marginal, suggesting the necessity of assimilating SSM/I precipitable water data for a better moisture analysis. Assimilation of the SSM/I precipitation rates improves the precipitation forecast over the tropical oceans.

Contents

1. Introduction

The concept of four-dimensional data assimilation was first introduced to numerical weather prediction (NWP) by Charney et al. (1969) in order to utilize incomplete historical data to infer the present state of the atmosphere. The information contained in the past observations that are distributed inhomogeniously in space and time is transferred to the present time by using a numerical model, and it is blended with the present observations to yield an optimal estimate of state variables of the present atmosphere. The product of four-dimensional data assimilation is called the analysis. The accuracy of analysis in data-void or sparse regions strongly depends on assimilation methods and performance of the model. Four-dimensional data assimilation is hereafter simply referred to as data assimilation to avoid confusion with the four-dimensional variational assimilation (4D-Var) method that is the subject of the present study.

Several methods for data assimilation have been proposed and investigated (Daley, 1991). In the direct insertion method, the model is integrated forward in time as observations are inserted at observation times. In the nudging method, nudging terms that force the model trajectory to get close to observations are added to the right-hand side of model prediction equations and time integration of the model is carried out. These methods have a couple of deficiencies. First, observation errors are not appropriately taken into account. This means that these methods are incapable of providing analysis error, which is necessary for constructing initial perturbations for ensemble prediction. Second, observational data for a variable different from model state variables can not be directly assimilated, and conversion of the observational data into model state variables by using limited information is required prior to assimilation. The assimilation method that had been widely used in NWP centers until recently is intermittent data assimilation with the multivariate optimal interpolation (OI) method. The OI method yields analysis by taking a weighting average of observations and short-range forecast with weights based on observation error and short-range forecast error. Since this method is based on the estimation theory, the first deficiency of the previous two methods is alleviated, but it is still not free from the second deficiency due to its methodology. This method has another difficulty in appropriately assimilating asynoptic observations such as polar-orbiting satellite data, because it assimilates observations intermittently such as at 6-hour intervals. Recently, NWP centers are adopting advanced data assimilation methods such as the three-dimensional variational assimilation (3D-Var) method and the 4D-Var method. 3D-Var is capable of directly

assimilating observations for a variable different from model state variables, and 4D-Var, in addition to that, can appropriately assimilates asynoptic observations.

Considerable attention has been paid to 4D-Var in the NWP community in recent years. It is a comprehensive multivariate analysis technique using model dynamics, and imposes no limitation on the type of data to be assimilated. This procedure gives the maximum likelihood estimate of the atmospheric state under certain conditions. Although the potential usefulness of variational approach for data assimilation was pointed out very early by Sasaki (1970), its heavy demands upon computer resources had prevented 4D-Var from being widely investigated until recently. Theoretical aspects of 4D-Var were presented by Le Dimet and Talagrand (1986) and Talagrand and Courtier (1987). 4D-Var seeks the model trajectory in phase space that minimizes the difference between model and observations over a certain time interval. This difference is measured by a cost function, and an adjoint model that is the adjoint of the tangent linear version of the model is needed to search for the minimum of the cost function. Several 4D-Var experiments with dry-adiabatic models have shown its advantages against the conventional OI method. For instance, experiments excluding data over an area with a strong baroclinic development showed that the information contained in the dynamics of the model is used successfully in the analysis over this area (Thépaut et al., 1993a). The impact of surface data, such as ERS-1 C-band measurements of backscatter, in 4D-Var is not confined to the surface and balanced analysis increments are obtained (Thépaut et al., 1993b). Andersson et al. (1994) showed that 4D-Var can extract wind information from the TOVS humidity channel that is one of the cloud-cleared radiance data. Their study also demonstrated the ability of 4D-Var to generate flow dependent and baroclinic structure functions in meteorological data analysis.

4D-Var is especially suitable for data analysis for the tropical atmosphere. The major advantage of 4D-Var in the tropics is the use of full model dynamics to assimilate observational data. Geostrophic balance is not a reasonable approximation for the flow in synoptic-scale tropical systems. As a result, it is difficult to improve the wind analysis from temperature observations and vice versa in conventional OI procedures. The geopotential and wind fields are usually analyzed univariately in the tropics. A more appropriate balance relationship for tropical multivariate data analysis can be derived under certain conditions. A scale analysis of synoptic-scale circulations in which the vertical scale is comparable to the scale height of the atmosphere shows that in the absence of condensation heating or precipitating systems the flow in the tropics is nearly nondivergent (Holton, 1992).

Absolute vorticity is approximately conserved following the nondivergent wind, and the following diagnostic relationship between the geopotential Φ and streamfunction Ψ can be derived from the divergence equation:

$$\nabla^2[\Phi + \frac{1}{2}(\nabla\Psi)^2] = \nabla \cdot [(f + \nabla^2\Psi)\,\nabla\Psi]\,. \qquad (1)$$

where f is the Coriolis parameter, ∇ the horizontal gradient operator, and ∇^2 the horizontal Laplacian operator. This nonlinear balance equation can be introduced as a weak constraint in 3D-Var procedures. For synoptic-scale precipitating systems in the tropics, however, the average vertical motion is an order of magnitude larger than the vertical motion in non-precipitating areas. Consequently, the flow in these systems has a relatively large divergent component so that the barotropic vorticity equation is no longer a good approximation. The primitive equations must be applied to analyze the flow. OI and 3D-Var procedures face an intrinsic limitation for this purpose. 4D-Var uses the primitive equations to assimilate observational data and, therefore, is quite appropriate for tropical data assimilation.

In order to improve NWP in the tropics, Krishnamurti et al. (1984) introduced the concept of physical initialization. Physical initialization modifies an analysis by using precipitation data which are provided by rain gauges and meteorological radars, or estimated from Special Sensor Microwave/Imager (SSM/I) and infrared (IR) observations by satellites. In their revised physical initialization procedure (Krishnamurti et al., 1991, 1993, 1994) outgoing longwave radiation (OLR) measured by polar-orbiting satellites is also used to restructure the moisture field. Furthermore, the temperature and moisture at the top of the constant flux layer are modified so that the parameterized surface fluxes are consistent with observed precipitation data. They demonstrated that the procedure leads to a significant reduction in the spin-up time and improves the skill of short-range forecasts. A variety of algorithms for physical initialization have been proposed by several other authors (e.g., Puri and Miller, 1990; Heckley et al., 1990; Davidson and Puri, 1992; Aonashi, 1993; Chang and Holt, 1994; Manobianco et al., 1994; Kasahara et al., 1994, 1996, Kasahara and Mizzi, 1996; Peng and Chang, 1996, 1997; Aonashi et al., 1997). The success of physical initialization provides promise for assimilating precipitation data in the tropics.

4D-Var is capable of assimilating precipitation data in a straightforward way, provided that

the assimilation model and the corresponding adjoint model include a treatment of moist processes. Zupanski and Mesinger (1995) carried out 4D-Var of precipitation data for the first time using a midlatitude regional model. Their 4D-Var procedure outperformed an OI method with a considerable improvement not only in the precipitation forecast but also in the forecast of other variables. A rainfall assimilation with 4D-Var using a limited-area mesoscale model was also tried by Zou and Kuo (1996). One of the major advantages of 4D-Var over physical initialization is that 4D-Var uses the full model dynamics to adjust the model variables to the observed precipitation. This is the most appropriate constraint for data assimilation. On the other hand, physical initialization directly modifies some of the model variables in a column model framework using reverse parameterization schemes or minimization procedures subject to certain constraints, and then applies a nudging technique or diabatic nonlinear normal mode initialization to adjust the other variables. Some physical initialization procedures use even a prescribed vertical heating profile to adjust model variables to observed precipitation rates. The second advantage of 4D-Var over physical initialization is that 4D-Var assimilate precipitation data with the other conventional data simultaneously. It allows interactions with other observations, leading to an optimal usage of observational data. Precipitation data errors also can be correctly handled in 4D-Var. Knowledge of observation errors is necessary for optimizing data assimilation procedures. Physical initialization assimilates precipitation data independently of conventional data, thus making it difficult to utilize precipitation data errors. The third advantage is that 4D-Var does not require synthesizing precipitation data from other observations. Precipitation rates retrieved from the SSM/I instrument are more accurate than those retrieved from IR radiometers, but the temporal resolution of the SSM/I from polar-orbiting satellites is much coarser (twice a day) than that of the IR radiometers on geostationary satellites. Manobianco et al. (1994) synthesized SSM/I and IR satellite data to yield precipitation estimates with realistic spatial and temporal structure for use in their physical initialization procedure. This type of synthesizing is unnecessary in 4D-Var procedures.

Application of 4D-Var to the tropical atmosphere has not been well examined in spite of the above mentioned advantages. This is because there is a difficulty in applying 4D-Var to tropical data analysis. In contrast to the extratropical atmosphere, cumulus convection plays a essential role in determining large-scale atmospheric circulation in the tropics. This property makes it important to include the parameterization of cumulus convection in 4D-Var. Since physical processes are far

more nonlinear than dynamical processes, the tangent linear approximation is less valid for a model with physical processes than for the same model without physical processes. In particular, threshold processes in parameterization schemes make a cost function discontinuous (zeroth-order discontinuity) or nondifferentiable (first-order discontinuity). A zeroth (first)-order discontinuity in prediction equations introduces a zeroth (first)-order discontinuity in the cost function. Although one of the subgradients of a nonsmooth cost function can be calculated using the adjoint model, smooth minimization methods may fail to reach the minimum of the cost function. In a smooth optimization algorithm, one may replace the cost function at a certain point by a quadratic or cubic model and minimize these models. Obviously these models no longer provide an efficient approximation of the function at a nondifferentiable or discontinuous point. Another problem is that an implementable termination rule of iteration is not easily applied in nonsmooth optimization. Furthermore, if the minimum of the cost function is located at a point of discontinuity, the accuracy of result depends upon whether the algorithm terminated with a point on the same branch as that the minimum belongs to. Several efficient algorithms for nonsmooth optimization have been proposed, such as the bundle methods, the SDG algorithms, and the r-algorithms (Lemarechal and Mifflin, 1978; Shor, 1985). But their computational costs are still much higher than several methods for smooth optimization.

From a practical point of view, the effects of discontinuities in cumulus convection schemes on variational data assimilation might be small compared to other effects because of local character of convection. In fact, Zou et al. (1993b) showed negligible effects of discontinuity in their variational data assimilation using a global model that includes a modified Kuo cumulus convection scheme. However, since cumulus convection plays a major role in the tropical circulation, discontinuities in moist processes may result in serious difficulties in variational data assimilation in the tropics. According to Vukicevic and Errico (1993), significant linearization errors may be expected in regions where transitions between convective and nonconvective conditions are frequent. They showed using a two-dimensional kinematic microphysical model that a smooth optimization algorithm did not converge when a zeroth-order discontinuity was included in the prediction equations. They demonstrated that a spline interpolation over a wide interval spanning the discontinuity works well to alleviate this difficulty. D. Zupanski (1993) found that discontinuities in the Betts-Miller cumulus convection scheme have the most serious effect in lower layers, and modified the scheme to render

it more continuous for variational data assimilation. Although the those authors have shown the feasibility of 4D-Var with moist processes included, their assimilation experiments were performed primarily for the extratropics.

This paper investigates the feasibility of 4D-Var for the tropical atmosphere by removing several types of zeroth-order discontinuity from the moist processes and by including a penalty term for controlling gravity wave level in the analysis. The impact of assimilating precipitation data, which is closely related to the moist processes, with 4D-Var on the tropical analysis is also examined. The cost function consists of a discrepancy term between model and observations and the penalty term, and does not include the background term that measures the difference from first guess. Physics of the adjoint model includes moist processes, horizontal diffusion, and simplified surface friction only. This package of physics is hereafter referred to as the reduced physics. Preconditioning and model systematic error correction are desirable for faster convergence and better quality of analysis, but these procedures are not introduced in the present study for the sake of simplicity.

Three variational data assimilation experiments are carried out. Prior to perform assimilation experiments with a three-dimensional model, a variational data assimilation experiment with a column model is carried out using simulated data to examine the impact of the removal of discontinuities from the parameterizations of moist processes. This experiment is hereafter referred to as the column model experiment. In this column model experiment, results obtained by using the continuous parameterization package are compared with those using the discontinuous one to assess the influence of discontinuities in parameterization schemes. The experiment will also give insight into the impact of assimilating precipitation data on tropical analysis.

Then a global primitive-equation model with the smoothed moist processes is used to investigate the effect of the penalty term for accelerating convergence of 4D-Var. The model used is a low-resolution version of the Florida State University global spectral model (FSU-GSM). In this experiment, the true state and observations are provided by the full-physics model, while the assimilation model (forward model) and the corresponding adjoint model use the above mentioned reduced physics. Since overly optimistic results will be obtained if the same model is used to both generate and assimilate the simulated data, a different model is used for assimilating the simulated data. An idealized observation network that is sparse in the tropics and the Southern Hemisphere is

adopted to roughly simulate the real observation network. This experiment is hereafter referred to as the simulated data experiment. It will demonstrate the importance of moist processes in the tropical data assimilation and the effect of the penalty term for increasing the efficiency of 4D-Var with moist processes included. Since a reliable estimate of precipitation in the global tropics is difficult, such a simulated data approach is suitable as a first step. A serious problem in this approach is that the parameterization schemes of moist processes are regarded as perfect. This assumption is unrealistic and, so, care should be taken in interpreting results from the experiment.

Finally, in order to evaluate the overall feasibility of 4D-Var in the tropics, real observational data are assimilated with the FSU-GSM. Radiosonde data are major observational data even in the tropics. For the purpose of demonstrating the advantages of 4D-Var, assimilation of asynoptic satellite data such as the SSM/I estimates for precipitation rates is more appropriate than assimilation of synoptic satellite data such as IR estimates from geostationary satellites. Therefore, in this experiment, the radiosonde data over the globe and SSM/I precipitation rate data over the tropical oceans are assimilated for the period 0 - 12 UTC 22 August 1992. Since real data is assimilated, the full-physics models is used as the assimilation model, while the adjoint model includes the reduced physics only. So the two models are not mathematically consistent. Development of an adjoint model with full physics, which is consistent with the assimilation model, is left for the future work. This experiment is hereafter referred to as the real data experiment, in which sensitivities of the convergence performance of 4D-Var to the size of the penalty term, the smoothness of moist processes, and horizontal interpolation methods for model precipitation are examined. Results from the real data experiment confirm the conclusions of the previous two experiments. Note that direct assimilation of SSM/I measurements of microwave brightness temperature may be better than assimilating precipitation rates retrieved from SSM/I brightness temperatures. However, the direct assimilation requires prediction of cloud water and rainwater as well as a microwave radiative transfer model. This approach is not taken in the present study.

Section 2 describes a basic formulation of 4D-Var, the method to assimilate precipitation data with 4D-Var, and an outline of smooth optimization algorithms. Section 3 presents an outline of the FSU-GSM and modifications to the parameterization schemes of moist processes to eliminate zeroth-order discontinuities. A brief description of a method for writing the computer code of the adjoint model is also included in this section. Sections 4, 5, and 6 describes in detail the method of

and results from the above three assimilation experiments: the column model experiment, simulated data experiment and real data experiment. Conclusions and future work are described in Section 7.

2. Formulation of 4D-Var

2.1 Basic formulation

4D-Var seeks the model trajectory in phase space that minimizes the difference between model and observations over a time interval considered, which is called an assimilation window. The difference is measured by a cost function. The column vector consisting of model state variables at all time levels in the assimilation window, $\boldsymbol{x}$, is determined by the model and the initial condition of the model state variables at the beginning of the assimilation windows, $\boldsymbol{x}_0$:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_N \end{pmatrix} = F(\boldsymbol{x}_0), \tag{2}$$

where $\boldsymbol{x}_n$ is the column vector consisting of the model state variables at time level n, and $F(\cdot)$ is the prediction equation of the model as a function of the initial condition. The total number of time levels in the assimilation window is N+1. The most general form of the cost function is given by

$$J(\boldsymbol{x}_0) = \frac{1}{2}(\boldsymbol{x}_0 - \boldsymbol{x}_0^b)^{\mathrm{T}} \mathbf{B}^{-1}(\boldsymbol{x}_0 - \boldsymbol{x}_0^b) + \frac{1}{2}[H(\boldsymbol{x}) - \boldsymbol{y}]^{\mathrm{T}} \mathbf{O}^{-1}[H(\boldsymbol{x}) - \boldsymbol{y}] + J_c, \tag{3}$$

where the superscript T indicates the transpose of vectors or matrices, $\boldsymbol{x}_0^b$ denotes the first guess for the initial condition $\boldsymbol{x}_0$, $\boldsymbol{y}$ the column vector consisting of observational data available in the assimilation window, $H(\cdot)$ the observation operator that converts the model state variables to observed variables and interpolates from model grid points to observation points, $\mathbf{B}$ and $\mathbf{O}$ the

error covariance matrices for $\boldsymbol{x}_0^b$ and $\boldsymbol{y}$, called the background error covariance matrix and the observation error covariance matrix, respectively, and J_c the penalty term for suppressing gravity wave noise. The value of the cost function is calculated by a forward intergration of the model.

The cost function is a function of the initial condition $\boldsymbol{x}_0$ only, which is the control variable for the minimization. If the model is perfect, background and observation errors are normally distributed and the penalty term is neglected, then the minimum of the cost function yields the maximum likelihood estimate of the model state variables over the assimilation window. The first term on the right-hand side of (3) is called the background term and measures misfit to the first guess, which contains the information of observational data prior to the beginning of the assimilation window. The second term is called the observation term and measures misfit to the observational data in the assimilation windows. If the penalty term J_c is not included, numerous noisy structures consisting mainly of gravity waves may appear in low latitudes (Courtier and Talagrand, 1990). This is due to the fact that the minimization process uses all the degrees of freedom of the model to minimize the cost function. Consequently, if observational data have errors or the model is not perfect, the minimization process introduces into the analysis as much gravity wave noise as necessary to reach the minimum. This problem may become more serious when precipitation data in the tropics is assimilated.

The cost function for the present study does not include the background term for simplicity, and the penalty term consists of the area mean of the squared magnitude of horizontal divergence tendency at all model grid points and all time levels in the assimilation window. The cost function is written as

$$J(\boldsymbol{x}_0) = \frac{1}{2}[H(\boldsymbol{x}) - \boldsymbol{y}]^{\mathrm{T}} \mathbf{O}^{-1}[H(\boldsymbol{x}) - \boldsymbol{y}] + \frac{1}{2} r \left(\frac{\partial D}{\partial t}\right)^{\mathrm{T}} \mathbf{W} \frac{\partial D}{\partial t}, \tag{4}$$

where r denotes the penalty parameter, $\frac{\partial D}{\partial t}$ the column vector consisting of time tendency of horizontal divergence at all model grid points and all time levels, and $\mathbf{W}$ a positive-definite block-diagonal symmetric matrix consisting of weights to calculate the area mean. The time tendency is approximated by a two-time level finite difference scheme. An explicit form of the penalty term is given in the Subsection 4.1 for the column model experiment. This formulation of the penalty

term may not be the best and a more sophisticated approach is desirable in the future work. A couple of formulations of the penalty term have been proposed. Courtier and Talagrand (1990) and Thépaut and Courtier (1991) showed that gravity wave noise can be efficiently eliminated by adding a penalty term which controls the magnitude of the time tendency of the gravity wave component of the flow and by introducing a nonlinear normal mode initialization algorithm in the variational process. On the other hand, Zou et al. (1992, 1993a) demonstrated that inclusion of simple quadratic penalty terms consisting of the time tendency of surface pressure and divergence efficiently damps high-frequency gravity waves. A digital filter technique proposed by Lynch and Huang (1992) as an efficient initialization method is easily applicable to 4D-Var, and Gustafsson (1992) reported a preliminary study in which the digital filter is applied to 4D-Var as a weak constraint.

Minimization of the above cost function using smooth optimization algorithms requires the gradient of the cost function with respect to the initial condition $\boldsymbol{x}_0$. It is written as

$$\nabla_{\boldsymbol{x}_0} J = \left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}_0}\right)^{\mathrm{T}} \left\{ \left(\frac{\partial \boldsymbol{H}}{\partial \boldsymbol{x}}\right)^{\mathrm{T}} \mathbf{O}^{-1} [H(\boldsymbol{x}) - \boldsymbol{y}] + r \left(\frac{\partial}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{D}}{\partial t}\right)^{\mathrm{T}} \mathbf{W} \frac{\partial \boldsymbol{D}}{\partial t} \right\} . \tag{5}$$

The transpose of the Jacobian matrix of the model on the right-hand side of (5), $\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}_0}\right)^{\mathrm{T}}$, is the propagator matrix of the adjoint model. The adjoint model is the adjoint of the tangent linear model. Since application of the adjoint operator to a multiplication of matrices reverses the order of multiplication, the adjoint model is integrated backward in time to calculate the gradient of the cost function. The basic state at each time level, which is necessary for the backward integration of the adjoint model, is provided by a forward integration of the original model (2). The latter model is usually referred to as the assimilation model or forward model. Another transposed matrix in (5), $\left(\frac{\partial \boldsymbol{H}}{\partial \boldsymbol{x}}\right)^{\mathrm{T}}$, is the adjoint of the linearized observation operator. Since the time tendency of divergence is approximated by a finite difference scheme, the calculation of the third transposed matrix $\left(\frac{\partial}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{D}}{\partial t}\right)^{\mathrm{T}}$ is not difficult. A method of writing the computer program for those adjoint operations, which is called the adjoint code, will be briefly described in Subsection 3.3.

One of the criticisms of variational data assimilation is that it requires the solution to exactly satisfy the model prediction equation. Wergen (1992) showed that variational data assimilation with an imperfect model distributes the model errors over the entire analysis domain, in contrast to the conventional OI method which projects the model errors only in the data-void areas. The model error can be classified into systematic error and random error. Derber (1989) developed a variational data assimilation method which takes the model systematic error into account. In his technique the systematic error is also included in the control variables and is estimated during the minimization process. M. Zupanski (1993a) applied this method to the variational data assimilation using a sophisticated regional model and obtained significantly better results. Since the largest part of forecast error in the tropics is due to the model systematic error (Kanamitsu, 1985), the inclusion of systematic error correction into 4D-Var is expected to have an even greater positive impact in the tropics. The model random error may be partially incorporated into observation error covariance matrices. Recently, however, D. Zupanski (1997) proposed a technique to correct random error during the minimization process by extending the Derber's method. In the present study, the model error is not considered for the sake of simplicity.

The above mathematical formulation of 4D-Var seems to require differentiability of the model with respect to the initial condition. From a practical point of view, however, differentiability of the model is not strictly required. Application of (5) to a nonsmooth but continuous cost function yields one of the subgradients at non-differentiable points of the cost function. This information can be efficiently utilized to seek the minimum of the cost function using methods for nonsmooth optimization. The problem is that nonsmooth optimization methods are much less efficient than smooth optimization methods. If the cost function has just a few discontinuities in its first derivative, and these discontinuities do not occur in the neighborhood of the minimum, then methods for smooth optimization are likely to be more efficient than methods for nonsmooth optimization (Gill et al., 1981). If, however, the cost function itself has many discontinuities in its value, it is much more likely that smooth optimization methods will fail to reach the minimum. Unfortunately, physical parameterizations used in most of the current NWP models have several discontinuities which introduce zeroth-order discontinuities in the cost function. Therefore, removal of zeroth-order discontinuities from parameterization schemes may be necessary for successful development of 4D-Var with physics included. Modifications to the parameterizations of moist processes to eliminate discontinuities are

described in Subsection 3.2.

The removal of discontinuities, however, tends to introduce locally large gradients in the cost function around discontinuous points. On the other hand, although mathematically the gradient of]the cost function is not defined at a discontinuous point, numerical application of (5) yields a finite gradient even at a discontinuous point. This is one of the subgradient of the cost function. Fitting a smoothing function over the discontinuity makes the cost function continuous and differentiable, but the magnitude of the gradient of the smoothed cost function around the discontinuous point is likely to be much larger than that of the subgradient of the discontinuous cost function computed by (5). Since physical processes are highly nonlinear, it may be inevitable that the cost function has a locally large gradient. Still, it is desirable to avoid encountering such a locally large gradient during the minimization process because it may considerably degrade convergence performance. It is shown in the present study that the penalty term is helpful to accelerate the convergence of 4D-Var with moist processes included.

2.2 Assimilation of precipitation data

As mentioned in the introduction, precipitation data are straightforwardly assimilated with 4D-Var. Assume that observational data including precipitation data are available at all time levels in the assimilation window and that the observed variables except for precipitation are calculated from model state variables at observed time levels only. Note that precipitation is usually not a model state variable in large-scale models and derived from model state variables at more than one time level. It is also assumed that observation errors at different time levels are uncorrelated. If the penalty term is neglected, then the cost function (4) is reduced to

$$J(\boldsymbol{x}_0)=\frac{1}{2}\sum_{n=0}^{N}\left[H_{y_n}(\boldsymbol{x}_n)-\boldsymbol{y}_n\right]^{\mathrm{T}}\mathbf{O}_{y_n}^{-1}\left[H_{y_n}(\boldsymbol{x}_n)-\boldsymbol{y}_n\right]+\frac{1}{2}\sum_{n=1}^{N}\left[H_{Q_n}(P_n)-Q_n\right]^{\mathrm{T}}\mathbf{O}_{Q_n}^{-1}\left[H_{Q_n}(P_n)-Q_n\right]\ . \tag{6}$$

where $\boldsymbol{x}_n$ denotes the column vector of model state variables at time level n, P_n the column vector of model precipitation accumulated between time levels (n-1) and n, $\boldsymbol{y}_n$ the column vector of observational data at time level n except precipitation, Q_n the column vector of observed precipitation accumulated between time levels (n-1) and n, $\mathbf{O}_{y_n}$ and $\mathbf{O}_{Q_n}$ the observation error covariance

matrix for $\boldsymbol{y}_n$ and Q_n, and $H_{y_n}(\cdot)$ and $H_{Q_n}(\cdot)$ the observation operators for $\boldsymbol{y}_n$ and Q_n, respectively. The observation operator $H_{Q_n}(\cdot)$ consists only of horizontal interpolation. If observations are not available at some time levels, the observation terms for these time levels are absent. When precipitation data accumulated over more than a one-time step interval are assimilated, only independent precipitation data should be included in the cost function. For example, to assimilate precipitation data accumulated over the whole the assimilation window, the second term on the right-hand side of (6) consists only of the term for the last time level, and P_N and Q_N are model and observed precipitations accumulated over the assimilation window. Note that it may be better to convert precipitation into, for instance, the cubic root of precipitation in order to ensure that precipitation observation error is more normally distributed, but this approach was not taken here.

As mentioned above, precipitation is not a model state variable, but is diagnosed from model state variables in most of large-scale atmospheric models including the FSU-GSM. A simple method to assimilate precipitation data with 4D-Var is to regard precipitation as one of the model state variables. Let $\tilde{\boldsymbol{x}}_n$ denote the model state variables at time level n thus extended. If a three-time level scheme is applied with a time filter for time integration of the model, as in the FSU-GSM, the model prediction equation can be written as

$$\tilde{\boldsymbol{x}}_{n+1}=F_n(\tilde{\boldsymbol{x}}_n)\qquad(n=0,\cdots,N-1)\ , \tag{7}$$

where

$$\tilde{\boldsymbol{x}}_0=\boldsymbol{x}_0\ , \tag{8}$$

$$\tilde{\boldsymbol{x}}_n=\begin{pmatrix}\boldsymbol{x}_n\\ \boldsymbol{x}_{n-1}^{f}\\ P_n\end{pmatrix}\qquad(n=1,\cdots,N)\ . \tag{9}$$

The superscript f indicates a time filtered value. Note that precipitation P_n is a diagnostic variable; the value of P_n is not necessary to compute the model state variables at the next time level, $\boldsymbol{x}_{n+1}$. However, precipitation at the previous time level is necessary for computing precipitation accumulated over more than a one-time step interval, so it is included in the model prediction equation in (7) -

(9).

The gradient of the cost function with respect to the initial condition x_0 can be directly derived by utilizing (5), but application of the Lagrange multiplier method is more straightforward. The problem to solve is to minimize the cost function (6) with the constraint of (7). The following Lagrangian function is introduced.

$$L(\tilde{x}_0, \cdots, \tilde{x}_N; \lambda_1, \cdots, \lambda_N) = J(x_0) + \sum_{n=0}^{N-1} [F_n(\tilde{x}_n) - \tilde{x}_{n+1}]^T \lambda_{n+1} , \tag{10}$$

where $\lambda_1, \cdots, \lambda_N$ are the column vectors of Lagrange multipliers of the same size as the extended model state variables $\tilde{x}_1, \cdots, \tilde{x}_N$. By substituting (6) into (10) and taking the first variation of (10) with respect to all the arguments of the Lagrangian function, the following algorithm for computing the gradient of the cost function, in addition to (7), is obtained.

$$\lambda_{N+1} = 0 , \tag{11}$$

$$\lambda_n = \left(\frac{\partial F_n}{\partial \tilde{x}_n}\right)^T \lambda_{n+1} + \begin{pmatrix} \left(\frac{\partial H_{y_n}}{\partial x_n}\right)^T \mathbf{O}_{y_n}^{-1} [H_{y_n}(x_n) - y_n] \\ 0 \\ \left(\frac{\partial H_{Q_n}}{\partial P_n}\right)^T \mathbf{O}_{Q_n}^{-1} [H_{Q_n}(P_n) - Q_n] \end{pmatrix} \qquad (n = N, \cdots, 1) , \tag{12}$$

$$\nabla_{x_0} J = \nabla_{\tilde{x}_0} J = \left(\frac{\partial F_0}{\partial \tilde{x}_0}\right)^T \lambda_1 + \left(\frac{\partial H_{y_0}}{\partial x_0}\right)^T \mathbf{O}_{y_0}^{-1} [H_{y_0}(x_0) - y_0] . \tag{13}$$

where $\left(\frac{\partial F_n}{\partial \tilde{x}_n}\right)^T$ is the propagator matrix of the adjoint model corresponding to the original model (7). The time integration of the adjoint model is carried out backward in time. Note that although $\left(\frac{\partial F_N}{\partial \tilde{x}_N}\right)^T$ in (12) is not defined, it is actually not used due to (11), and that filtered model state variables are not used to calculate observed variables.

The observation error covariance matrix for conventional data consists of measurement error

and interpolation error, while the observation error covariance matrix for precipitation data derived from remote sensing, such as satellite and meteorological radar observations, contains additional errors: retrieval algorithm error and parameterization error for moist processes in the model. To assimilate precipitation data with 4D-Var, a parametrization scheme of moist processes is used as in physical initialization techniques. However, the parameterization gives a statistical estimate of the influence of subgrid-scale phenomena on grid-scale fields: an ensemble mean over all possible realizations under the same grid-scale forcing. On the other hand, what is really observed and assimilated is only one of the realizations. Numerical simulations with a cumulus ensemble model under periodic large-scale forcing showed that domain-averaged precipitations for each period are considerably different from each other (Xu et al., 1992). The absence of such a fluctuation in the parameterization scheme introduces random error, even if the scheme does not have systematic error. The problem is that random error in a parameterization scheme is not easy to estimate. Numerical studies with a cumulus ensemble model might provide a reliable estimate of the random error in a specific cumulus parameterization scheme, but this is beyond the scope of the present study. Note that accumulated precipitation over a longer period computed by a parameterization scheme may be expected to have less relative magnitude of random error. 4D-Var is capable of assimilating the accumulated precipitation in a straightforward way, different from physical initialization techniques.

2.3 Optimization algorithm

A large-scale smooth optimization algorithm is used to seek the minimum of the cost function in the present study. Popular large-scale smooth optimization methods are of Newton type. Zou et al. (1993c) conducted comparative tests of several limited-memory quasi-Newton methods and truncated Newton methods for unconstrained nonlinear optimization. They concluded that among the tested limited-memory quasi-Newton methods the L-BFGS method (Nocedal, 1980; Liu and Nocedal, 1989) had the best overall performance for several test problems including large-scale problems in oceanography and meteorology, and that the numerical performance of truncated Newton methods such as TNPACK (Schlick and Fogelson, 1992a, b) is competitive with that of L-BFGS. Wang (1993) used the L-BFGS method in his 4D-Var experiments with the dry-adiabatic version of the FSU-GSM. The L-BFGS method is taken for the 4D-Var experiments of the present

study. This method requires as input data the values of the cost function and its gradient with respect to the control variables. It continuously updates the quasi-Newton approximation of the inverse Hessian matrix by using information from the last m quasi-Newton iterations. The number of updates m is supplied by the user (generally, $3 \leq m \leq 7$).

Minimization algorithms work best when all the control variables are of order unity, and convergence rates strongly depend on how to scale the control variables. The optimum way of the scaling needs computation of the Hessian matrix of the cost function (Thacker, 1989); however, in general it is prohibitively expensive. If the cost function contains the background term, the inverse background covariance error matrix can be used as an approximation to the Hessian matrix, but this is not the case in the present study. Although a few sophisticated preconditioning techniques have been proposed by M. Zupanski (1993b, 1996), Courtier et al. (1994), and Yang et al. (1996), a simple scaling method is used in the present study; each state variable is scaled by its representative constant value. Since it is not easy to determine the optimal values for these scaling constants, the scaling constants used are not well optimized to yield faster convergence.

3. Description of model

3.1 Outline of FSU-GSM

The model used is a low-resolution version of the FSU-GSM. This model was developed at FSU for NWP in the tropics, and has been used for numerical study of tropical meteorology. Although several modifications have been introduced so far, the basic dynamical formulation of the FSU-GSM is quite similar to that described by Daley et al. (1976). Since detailed descriptions of the model are given by Pasch (1983) and Wang (1993), this subsection describes a brief outline of the model.

a. Governing equations

The horizontal coordinates are the longitude λ and the latitude θ, and the vertical coordinate is sigma, defined as

$$\sigma = \frac{p}{p_s}, \tag{14}$$

where p is the pressure, and p_s is the surface pressure. The model state variables are the vertical component of relative vorticity ζ, the horizontal divergence D, the virtual temperature T_v, the dewpoint depression $T - T_d$, and the natural logarithm of surface pressure $\ln p_s$. The virtual temperature T_v and the dewpoint depression T_d are related to the temperature T and the specific humidity q as

$$T_v = \left[1 + \left(\frac{R_v}{R_d} - 1\right) q\right] T, \tag{15}$$

$$q = \frac{\varepsilon}{p} e_s(T_0) \exp\left[\frac{\varepsilon L}{R_d}\left(\frac{1}{T_0} - \frac{1}{T_d}\right)\right], \tag{16}$$

where R_d and R_v denote the gas constant of dry air and water vapor, respectively, ε the ratio of the molecular weight of water vapor to that of dry air, L the specific latent heat of condensation or sublimation of water vapor, and $e_s(T_0)$ the saturation water vapor pressure at a reference temperature T_0.

The governing equations of the model are the primitive equations written as

$$\frac{\partial \zeta}{\partial t} = -\alpha(A, B), \tag{17}$$

$$\frac{\partial D}{\partial t} + \nabla^2(\Phi + R_d T_v^* \ln p_s) = \alpha(B, -A) - a^2 \nabla^2 \frac{(U^2 + V^2)}{2\cos^2\theta}, \tag{18}$$

$$\frac{\partial T_v}{\partial t} - \gamma^* \dot{\sigma} + \frac{R_d T_v^*}{C_{pd}} \int_0^1 (G + D)\, d\sigma = -\alpha(U T_v', V T_v') + Q, \tag{19}$$

$$\frac{\partial (T - T_d)}{\partial t} = -\alpha[U (T - T_d), V (T - T_d)] + H, \tag{20}$$

$$\frac{\partial \ln p_s}{\partial t} + \int_0^1 (G + D)\, d\sigma = 0, \tag{21}$$

where

$$(U, V) = \frac{\cos\theta}{a}(u, v), \tag{22}$$

$$\gamma = \frac{R_d T_v}{C_{pd}} - \frac{\partial T_v}{\partial \sigma}, \tag{23}$$

$$\dot{\sigma} = \sigma \int_0^1 (G + D)\, d\sigma - \int_0^{\sigma} (G + D)\, d\sigma, \tag{24}$$

$$G = \frac{1}{\cos^2\theta}\left(U \frac{\partial \ln p_s}{\partial \lambda} + V \cos\theta \frac{\partial \ln p_s}{\partial \theta}\right), \tag{25}$$

$$\alpha(A, B) = \frac{1}{\cos^2\theta}\left(\frac{\partial A}{\partial \lambda} + \cos\theta \frac{\partial B}{\partial \theta}\right), \tag{26}$$

$$A = (f + \zeta) U + \dot{\sigma} \frac{\partial V}{\partial \sigma} + \frac{R_d T_v'}{a^2 \cos\theta} \frac{\partial \ln p_s}{\partial \theta} - \frac{\cos\theta}{a}\left(\frac{dv}{dt}\right)_{phys}, \tag{27}$$

$$B = (f + \zeta) V - \dot{\sigma} \frac{\partial U}{\partial \sigma} - \frac{R_d T_v'}{a^2} \frac{\partial \ln p_s}{\partial \lambda} - \frac{\cos\theta}{a}\left(\frac{du}{dt}\right)_{phys}, \tag{28}$$

$$\begin{aligned} Q = {} & T_v' D + \gamma' \dot{\sigma} - \frac{R_d T_v'}{C_{pd}} \int_0^1 (G + D)\, d\sigma + \frac{R_d T_v}{C_{pd}} G + \left(\frac{dT}{dt}\right)_{phys} \\ & + \left(\frac{R_v}{R_d} - 1\right)\left[q\left(\frac{dT}{dt}\right)_{phys} + T\left(\frac{dq}{dt}\right)_{phys}\right], \end{aligned} \tag{29}$$

$$\begin{aligned} H = {} & (T - T_d) D - \dot{\sigma} \frac{\partial (T - T_d)}{\partial \sigma} - \left(\frac{R_d T}{C_{pd}} - \frac{R_d T_d^2}{\varepsilon L}\right)\left|\frac{\sigma}{\sigma} + G - \int_0^1 (G + D)\, d\sigma\right| \\ & + \left(\frac{dT}{dt}\right)_{phys} - \frac{R_d T_d^2}{\varepsilon L q}\left(\frac{dq}{dt}\right)_{phys}. \end{aligned} \tag{30}$$

The standard meteorological notation is used: t is time, u and v the zonal and meridional velocity component, Φ the geopotential height, a the radius of the earth, f the Coriolis parameter, and

C_{pd} the heat capacity at constant pressure of dry air. The virtual temperature T_v and the static stability γ are decomposed into their initial horizontal means, denoted by superscript *, and deviations from them, denoted by superscript $'$. The terms with subscript *phys* on the right-hand side of (27) - (30) represent the time tendencies due to subgrid-scale physical processes.

b. Dynamical processes

The model is a global spectral model with a horizontal resolution of triangular truncation at total wavenumber 42 and it has 12 levels in the vertical (T42L12). The transform method (Orszag, 1970) is utilized for calculating the spectral coefficients of nonlinear terms. The number of Gaussian grid points is 128 (longitude) x 64 (latitude) with an approximate equivalent resolution of 2.8125 ° x 2.8125 °. The schematic diagram of vertical configuration of model levels are depicted in Fig. 1. The vertical levels where vorticity and divergence are defined are σ = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.85, 0.9, 0.95, 0.99. The vorticity, divergence, and virtual temperature are defined at all model levels, while the dewpoint depression is defined only at the lowest 10 levels.

To save computational time, a trapezoidal semi-implicit scheme is applied to linearized gravity wave terms. The time integration of nonlinear terms is made with a leap-frog scheme and a time filter for eliminating computational modes (Asselin, 1972). Therefore, the extended representation of the model prediction equation of (7) - (9) is directly applicable. The time step is set to 1200 s. The first-order upstream scheme is applied to discretize the vertical advection of dewpoint depression, while the second-order symmetric difference scheme is applied for the vertical advection of the other variables. Note that the cost function for the FSU-GSM is not smooth due to the use of the upstream scheme, even if the parameterization schemes are not included. The original FSU-GSM uses an envelope orography to represent the surface topography. It was found that noise was sometimes produced around the Andes due to the steep orography, and that this noise was detrimental to 4D-Var experiments. In order to reduce the noise, the envelope orography was replaced by a mean orography for the present study.

c. Physical processes

The physics of the model includes radiation, surface fluxes, planetary boundary layer, vertical

diffusion, horizontal diffusion, dry convective adjustment, large-scale condensation, evaporation from falling precipitation, and deep cumulus convection. The latter four physical processes are referred to as moist processes in this paper. Note that the dry convective adjustment is included as a moist process for the sake of convenience. Shallow cumulus convection is not taken into account, because it is not directly associated with precipitation process.

The radiative process is essentially the same as that of Chang (1979). This scheme includes the diurnal change in the incoming shortwave radiation and a parameterization of clouds. Three basic cloud types are considered: low, middle, and high cloud, and clouds are specified from the relative humidity distribution. The emissivity method is utilized to compute the net longwave radiation. Calculation of the surface fluxes follows the formulation of Businger et al. (1971), which makes use of the similarity theory. The fluxes of heat, moisture and momentum in the planetary boundary layer are specified by the mixing length theory, as in Smagorinsky et al. (1965). The vertical diffusion is introduced following the formulation of Louis (1979), and the horizontal diffusion is a Laplacian-type. A modified Kuo scheme (Krishnamurti et al., 1983) is used for the parameterization of deep cumulus convection. Details of the moist processes are described in the next subsection.

3.2 Modifications to moist processes

Several zeroth-order discontinuity are removed from the original parameterization schemes of the moist processes in order to improve the convergence performance of 4D-Var. A couple of additional modifications are also made for convenience in writing the adjoint code.

a. Dry Convective adjustment and large-scale condensation

The dry convective adjustment and large-scale condensation are, in principle, continuous parameterization schemes, though not differentiable. However, in the FSU-GSM, the relative humidity after the execution of large-scale condensation is not equal to the threshold relative humidity for the large-scale condensation, but it is equated to the relative humidity before the dry convective adjustment is executed. Such a formulation introduces a discontinuity into the moist processes. Thus, the large-scale condensation scheme is altered such that the relative humidity after the large-scale condensation is executed is equal to the threshold relative humidity. This is a more usual formulation

in most of the current atmospheric models.

b. Evaporation of falling precipitation

The incremental change in specific humidity and temperature at the kth vertical level due to the evaporation of falling precipitation, and the precipitation at the (k+1)th level, which is the adjacent lower level, are given by

$$\delta q_k = \min\left[\alpha P_k (q_{sk}^* - q_k^*), \frac{\rho_w g P_k}{p_s \Delta\sigma_k}\right], \tag{31}$$

$$\delta T_k = -\frac{L}{C_{pd}^*}\delta q_k , \tag{32}$$

$$P_{k+1} = P_k - \frac{p_s \Delta\sigma_k}{\rho_w g}\delta q_k , \tag{33}$$

respectively, where α is a positive constant (100 m^{-1}), P_k the precipitation at the kth level, q_{sk}^* and q_k^* the saturation specific humidity and specific humidity at the kth level before the execution of the scheme, respectively, ρ_w the density of liquid water, g the acceleration of gravity, and $\Delta\sigma_k$ the thickness between the kth and (k+1)th levels.

This process is invoked, if the following two conditions are satisfied.

$$P_k > 0 , \tag{34}$$

$$q_{sk}^* - q_k^* > 0.1\, q_{sk}^* . \tag{35}$$

The second condition introduces a discontinuity into the moist processes. To remove this discontinuity, the term on the right-hand side of (31) is multiplied by the following threshold function instead of applying the condition (35).

$$\Theta(q_s^* - q^*, 0.1\, q_s^*) = 1 - \left(\frac{1}{2}\right)^{\left(\frac{q_s^* - q^*}{0.1\, q_s^*}\right)^2} . \tag{36}$$

Since multiplying a smooth function to eliminate discontinuity does not require on/off switches, it may be superior to a spline interpolation used by Verlinde and Cotton (1993).

c. Deep cumulus convection

A partial description of the modified Kuo scheme used in the FSU-GSM is given by Krishnamurti et al. (1983). The following is a full description of the scheme. Although a more recent version of the Kuo scheme is adopted in the FSU-GSM, the one described below is used in the present study.

The time tendencies of temperature, specific humidity, and precipitation due to the Kuo scheme are given by

$$\left(\frac{dT}{dt}\right)_{Kuo} = \frac{(1-b)(1+\eta) I_L}{Q_\theta} \max\left\{\frac{T_c - T}{\Delta\tau} + \left(\frac{p}{p_0}\right)^{\frac{R_d}{C_{pd}}} \omega \frac{\partial}{\partial p}\left[\left(\frac{p_0}{p}\right)^{\frac{R_d}{C_{pd}}} T\right], 0\right\}, \tag{37}$$

$$\left(\frac{dq}{dt}\right)_{Kuo} = \frac{b(1+\eta) I_L}{Q_q} \frac{q_c - q}{\Delta\tau} + \omega \frac{\partial q}{\partial p}, \tag{38}$$

$$\left(\frac{dP}{dt}\right)_{Kuo} = \frac{(1-b)(1+\eta) I_L}{\rho_W}, \tag{39}$$

where

$$\omega = p\left(-D - \frac{\partial\dot{\sigma}}{\partial\sigma} + \frac{\dot{\sigma}}{\sigma}\right), \tag{40}$$

$$I_L = -\frac{p_s}{g}\int_{\sigma_T}^{\sigma_B} \omega \frac{\partial q}{\partial p} d\sigma, \tag{41}$$

$$Q_\theta = \frac{p_s}{g}\int_{\sigma_T}^{\sigma_B} \frac{C_{pd}}{L} \max\left\{\frac{T_c - T}{\Delta\tau} + \left(\frac{p}{p_0}\right)^{\frac{R_d}{C_{pd}}} \omega \frac{\partial}{\partial p}\left[\left(\frac{p_0}{p}\right)^{\frac{R_d}{C_{pd}}} T\right], 0\right\} d\sigma, \tag{42}$$

$$Q_q = \frac{p_s}{g}\int_{\sigma_T}^{\sigma_B} \frac{q_c - q}{\Delta\tau} d\sigma. \tag{43}$$

In the above equations, ω is the vertical p-velocity, T_c and q_c the temperature and specific humidity of the moist adiabat starting from the cloud base, respectively, p_0 a reference pressure,

$\Delta\tau$ the cloud formation time scale (1800 s), and σ_B and σ_T the height of cloud base and cloud top, respectively. The moistening parameter b and the mesoscale convergence parameter η are assumed to be a function of relative vorticity at 700 hPa, ζ_{700}, and vertically averaged p-velocity, ϖ:

$$b = \begin{cases} 0 & (b^* < 0,\ \eta_{10} < 0), \\ b^* & (0 \le b^* \le 0.1,\ b^* \ge \eta_{10}), \\ \eta_{10} & (0 \le \eta_{10} \le 0.1,\ \eta_{10} > b^*), \\ 0.1 & (b^* > 0.1 \text{ or } \eta_{10} > 0.1), \end{cases} \tag{44}$$

$$\eta = \begin{cases} -0.1 & (\eta^* < -0.1), \\ \eta^* & (-0.1 \le \eta^* \le 0.5), \\ 0.5 & (\eta^* > 0.5), \end{cases} \tag{45}$$

where

$$b^* = \frac{a_1 \zeta_{700} + b_1 \varpi + c_1}{1 + \eta^*}, \tag{46}$$

$$\eta^* = (a_1 + a_2)\zeta_{700} + (b_1 + b_2)\varpi + (c_1 + c_2) - 1, \tag{47}$$

$$\eta_{10} = \frac{\eta}{1+\eta}. \tag{48}$$

The numerical values of the constant parameters in (46) and (47) are $a_1 = 1.58 \times 10^5$ s, $a_2 = 1.07 \times 10^5$ s, $b_1 = 3.04 \times 10^3$ hPa^{-1} s, $b_2 = 1.07 \times 10^3$ hPa^{-1} s, $c_1 = 0.476$, and $c_2 = 0.870$, as in Krishnamurti et al. (1983).

The cloud base σ_B is set to the lowest level satisfying the following four conditions.

$$\frac{\partial\theta_c^*}{\partial p} > 0, \tag{49}$$

$$\omega \frac{\partial q}{\partial p} < 0 , \tag{50}$$

$$\omega < 0 , \tag{51}$$

$$RH \geq RH_c , \tag{52}$$

where θ_e^* is the saturation equivalent potential temperature, RH the relative humidity, RH_c a threshold relative humidity, set to 0.7. Equation (49) implies a conditionally unstable stratification. The temperature of an air parcel lifted adiabatically from the cloud base is then compared with the surrounding air temperature until a vertical level of neutral or negative buoyancy is found. This level is set to the cloud top σ_T. The deep cumulus convection is invoked if the following two conditions are met in addition to the existence of σ_B.

$$I_L > 0 , \tag{53}$$

$$\sigma_B - \sigma_T > d_c , \tag{54}$$

where d_c is a threshold cloud thickness in sigma, set to 0.125.

The modified Kuo scheme described above has several discontinuities. First, the cloud base and cloud top are assigned to one of the model vertical levels as in most of the state-of-the-art atmospheric models. This may cause a serious problem in variational data assimilation as pointed out by D. Zupanski (1993). She multiplied the tendencies of temperature and specific humidity due to the Betts-Miller cumulus convection scheme by a smooth function of vertical coordinate to reduce the abrupt jumps from nonconvective to convective regimes. However, her method does not perfectly remove the discontinuity. In this study, the heights of cloud base and cloud top are allowed to have continuous values by linearly interpolating vertical temperature profiles calculated by the model.

Second, it is easily seen that the specification of the cloud base using the conditions (49)-(52) introduces a discontinuity. Since it is difficult to smooth this specification, the cloud base conditions are replaced by the condition of the lower bound of a conditionally unstable layer above the lifting condensation level, and a threshold function of the relative humidity at the cloud base RH_B is introduced. Therefore, the original conditions (49) and (52) are taken into account in the modified scheme, but the other two conditions are not. Third, the discontinuity associated with the condition

(54) is eliminated by introducing a smooth threshold function of cloud thickness. The terms on the right-hand side of (37)-(39) are multiplied by the following function combining the above two threshold functions, in place of applying the conditions (52) and (54).

$$\tilde{\Theta} = \begin{cases} 0 & (0 \leq RH_B \leq RH_0) , \\ \Theta(\sigma_B - \sigma_T, d_c)\, \Theta(RH_B - RH_0, RH_c - RH_0) & (RH_0 < RH_B \leq 1) , \end{cases} \tag{55}$$

where RH_0 is a positive constant less than RH_c, and the explicit form of the function $\Theta(x, y)$ is given by (36). The value of RH_0 is set to 0.6.

There are other types of zeroth-order discontinuities in the modified Kuo scheme. One of them is associated with the method for specifying the height of cloud top. As shown in Fig. 2, the cloud top height may change abruptly when there is an inversion layer. This discontinuity is in common with other deep cumulus parameterization schemes where a similar method is used to determine the cloud top height. The removal of this discontinuity causes the scheme to be more complicated, so that it is left unchanged. Incidentally, it is worthwhile to note that the moist convective adjustment scheme does not have such a discontinuity; it is a continuous parameterization scheme of cumulus convection in spite of the application of vertical discretization, although it has several drawbacks (Frank and Molinari, 1993).

There are a couple of discontinuities which are not present in the original Kuo scheme (Kuo, 1974) but are introduced by the unique formulation of the modified Kuo scheme in the FSU-GSM. The time tendency of specific humidity has a discontinuity at $I_L = 0$ due to the presence of the second term on the right-hand side of (38). The moistening parameter b has a discontinuity at $\eta^* = -1$, as is seen from (44) and (46). Although it is not difficult to remove these two discontinuities, they are left as they are in this study for the sake of simplicity.

d. Other modifications

In the FSU-GSM, the dry convective adjustment, the large-scale condensation and the evaporation of falling precipitation are executed in this order for every vertical level. This calculation are iterated until absolutely unstable stratification disappears at all levels with the evaporation of falling precipitation executed at the first iteration only. Then the Kuo scheme is carried out. In order

to avoid following-up this iteration process in the adjoint code, the dry convective adjustment is separated from the large-scale condensation and the evaporation of falling precipitation, and is carried out first with an iteration algorithm before the latter two processes are executed. In this case, the adjoint code of the dry convective adjustment scheme is reduced to a code for solving several sets of simultaneous linear equations. However, by so doing a possible unstable stratification is allowed before the execution of the Kuo scheme, because the large-scale condensation and evaporation of falling precipitation may render a neutral stratification produced by the dry convective adjustment unstable.

Another modification is the removal of ice phase from the moist processes. In the FSU-GSM, the latent heat of condensation or sublimation and the saturation vapor pressure of water vapor are discontinuous functions of temperature with a discontinuity at -10 °C. In order to eliminate this discontinuity the function forms defined above -10 °C are simply applied even for temperatures below this value. Since the major interest is in the impact of precipitation data in the tropics, this modification may not give rise to a serious problem.

e. Performance of the modifications

The parameterization package for moist processes modified as mentioned above is referred to as the continuous version, although it still has a few discontinuities. A parameterization package with less modifications is constructed for comparison, in which the modification to the evaporation of falling precipitation and all the modifications to the Kuo scheme except the replacement of the method for specifying the cloud base are removed from the continuous version. This package is referred to as the discontinuous version. It is to be noted that the discontinuous version is more continuous than the original version.

The changes induced by the above modifications to the parameterization schemes have to be tested to ensure that they do not degrade the model performance. Since a detailed examination is beyond the scope of the present study, only a comparison of 6-hour precipitation simulated by the FSU-GSM with the original version, continuous version, and discontinuous version of moist processes is presented in Fig. 3. The initial condition is an initialized analysis at 12 UTC 27 July 1979 prepared by the European Centre for Medium-Range Forecasts (ECMWF). It is found that overall precipitation patterns are almost identical. In the tropics the difference between the continuous

version and the discontinuous version is smaller than the difference between the original version and the continuous or discontinuous version. It may be concluded from Fig. 3 that the modifications introduced above do not significantly degrade the performance of the model.

3.3 Adjoint model

The adjoint model of a dry-adiabatic version of the T42L12 FSU-GSM was developed by Zhi Wang (Wang, 1993). The model did not predict the dewpoint depression, and the only physics included was horizontal diffusion. For the purpose of the present study, the prediction equation for dewpoint depression and the parameterization of the moist processes and a simplified surface friction are included in the adjoint model. The simplified surface friction is taken from Thépaut et al. (1993b), and the following frictional force is added to the momentum equation at the lowest two model levels:

$$F_{PBL} = -\frac{u_*}{h}\frac{k}{\ln\left(\frac{z_K}{z_0}\right)}\mathbf{v}_K , \qquad (56)$$

where h is the thickness of the planetary boundary layer (1 km), u_* the surface friction velocity (0.5 m s^{-1} over land; 0.2 m s^{-1} over sea), k the von Kalman's constant (0.4), z_0 the roughness length (0.05 m over land; 0.0005 m over sea), z_K the height of the lowest model level, and $\mathbf{v}_K$ the horizontal velocity vector at the lowest model level. For simplicity, z_K is set to a constant (100 m). The package of physics that consists of the moist processes, the simplified surface friction, and the horizontal diffusion is referred to as the reduced physics. Replacement of the simplified surface friction with the original planetary boundary layer parameterization scheme and inclusion of the remaining physical processes in the adjoint model are desirable to obtain a more optimal solution to 4D-Var. Development of the adjoint code for these processes is left for future work.

In order to write the adjoint code of the above additional processes, the tangent linear code of them was first written by linearizing the original nonlinear code around an arbitrary basic state. A problem encountered at this stage is the treatment of conditional statements that are contained in the original code. In the tangent linear code, the if-branch to be taken in a conditional statement was determined solely by the basic state, and a possible effect of perturbations on the selection of the

if-branch was entirely neglected, as was done by Zou et al. (1993b). By doing so, the resulting tangent linear code is less valid near a branching point. A precise treatment of conditional statements in the tangent linear code has recently been proposed by Xu (1996a, b), but his procedure is too complicated to be applied to large-scale atmospheric models like the FSU-GSM. The correctness of the tangent linear code was numerically verified by using the Taylor expansion:

$$\frac{|A(x+\alpha\,\delta x)-A(x)|}{|\alpha|\,|\mathbf{A}\,\delta x|}=1+O(\alpha)\,, \tag{57}$$

where $A(\cdot)$ represent a process in the original code, $\mathbf{A}$ the matrix that represents the corresponding process in the tangent linear code, x the column vector of the input variables, δx the perturbation to x, α a scaler sufficiently smaller than unity. The above equation indicates that as the value of α goes to zero, the value of the left-hand side goes to unity linearly if the tangent linear code is correctly written. Note that (57) holds if continuous partial derivatives of order two of $A(\cdot)$ exist in the neighborhood of x.

The adjoint code corresponding to the tangent linear code was written by converting the latter code following the definition of the adjoint operator. First, the adjoint of a multiplication of matrices is the reversed-order multiplication of adjoint matrices, so the order of the processes are all reversed and the input and output variables are interchanged in the adjoint code. Second, most statements in the tangent linear code are assignment statements like this:

```
Z = A*X + B*Y,
```

where X, Y, and Z are perturbation variables, and A and B depend on the basic state only. This is a simple example. If X and Y are no longer used later in the tangent linear code, the corresponding adjoint code is

```
X = A*Z,
Y = B*Z.
```

On the other hand, if X and Y are used again later in the tangent linear code, the corresponding adjoint code is

```
X = X + A*Z,
Y = Y + B*Z.
```


These conversion rules from the tangent linear code to the adjoint code are easily verified by representing the tangent linear code in a matrix form and applying the adjoint operation to the matrix. The correctness of the adjoint code was numerically verified by using the following identity, as was done by Thépaut and Courtier (1991) and Navon et al. (1992):

$$(\mathbf{A}\,x)^{\mathrm{T}}\,\mathbf{A}\,x=x^{\mathrm{T}}\,\mathbf{A}^{\mathrm{T}}(\mathbf{A}\,x)\,. \tag{58}$$

where $\mathbf{A}$ is a matrix that represents a process in the tangent linear code and x is a column vector of the input perturbation variables to the process. The left-hand side of (58) is computed by using the tangent linear code only, while the adjoint code is used for calculating the column vector $\mathbf{A}^{\mathrm{T}}(\mathbf{A}\,x)$ on the right hand side.

4. Column model experiment

4.1 Method

a. Column Model

The governing equations of the column model are the same as those of the FSU-GSM (17) - (30), except that the prediction equations of vorticity and divergence are omitted and that all the horizontal advection terms are neglected. They are written as

$$\frac{\partial T_v}{\partial t}+\dot{\sigma}\frac{\partial T_v}{\partial\sigma}=-\frac{R_d\,T_v}{C_{pd}\,\sigma}\int_0^{\sigma}D\,d\sigma+\left(\frac{dT}{dt}\right)_{phys}+\left(\frac{R_v}{R_d}-1\right)\left[q\left(\frac{dT}{dt}\right)_{phys}+T\left(\frac{dq}{dt}\right)_{phys}\right], \tag{59}$$

$$\frac{\partial(T-T_d)}{\partial t}+\dot{\sigma}\frac{\partial(T-T_d)}{\partial\sigma}=-\left(\frac{R_d\,T}{C_{pd}}-\frac{R_d\,T_d^{\,2}}{\varepsilon\,L}\right)\frac{1}{\sigma}\int_0^{\sigma}D\,d\sigma+\left(\frac{dT}{dt}\right)_{phys}-\frac{R_d\,T_d^{\,2}}{\varepsilon\,L\,q}\left(\frac{dq}{dt}\right)_{phys}, \tag{60}$$

$$\frac{\partial\ln p_s}{\partial t}=-\int_0^1 D\,d\sigma\,, \tag{61}$$

$$\dot{\sigma}=\sigma\int_0^1 D\,d\sigma-\int_0^{\sigma}D\,d\sigma\,. \tag{62}$$

Divergence and vorticity at 700 hPa, which is used in the modified Kuo scheme are to be prescribed.

The physical processes employed in computing $\left(\frac{dT}{dt}\right)_{phys}$ and $\left(\frac{dq}{dt}\right)_{phys}$ in the column model are the moist processes only: dry convective adjustment, large-scale condensation, evaporation of falling precipitation, and deep cumulus convection. The same FORTRAN code as that used in the T42L12 FSU-GSM is applied to discretize (59) - (62) except that the semi-implicit scheme for the time integration is replaced with an explicit one using the leap-frog scheme. This is because divergence field is prescribed and, therefore, gravity waves are absent in the column model.

Precipitation data is expected to have a large impact on the analysis of divergence field in the tropics, but divergence is prescribed in the column model. It is possible, however, to include the divergence variable in the control variables for minimization and to estimate its optimal value during the minimization process if enough observational data are available. By doing so the impact of precipitation data on divergence analysis is partly examined, although its exact estimation is difficult in the column model experiment. For this purpose, only the vertical structure of divergence is prescribed as follows:

$$D_{n,k} = D_n(1 - 2\sigma_k) \qquad (n = 0, \cdots, N-1;\ k = 1, \cdots, K-1), \qquad (63)$$

where D_0 , ···, D_{N-1} are divergence parameters to be estimated, the subscript n the index of time levels, and the subscript k the index of vertical levels, increasing downward. The total number of time levels in the assimilation window and that of vertical levels are N+1 and K, respectively. In this study K is equal to 12. The value of divergence at the lowest level, $D_{n,K}$, is determined so that the vertical integral of divergence over the vertical coordinate vanishes. Therefore, the right-hand side of (61) and the first term on the right-hand side of (62) are absent. Since the vertical levels are not placed at equal intervals as shown in Fig. 1, if (63) is applied to all vertical levels, then the vertical integral of divergence does not vanish.

b. Experimental design

A simulation with the column model starting from a reference state with a prescribed wind field is used as the truth, against which results from assimilation experiments are verified. The reference vertical profiles of virtual temperature and dewpoint depression are derived from Table 1 of Yamasaki (1983). They are similar to the mean undisturbed state of the tropical atmosphere in

summer. The reference surface pressure is set to 1013.25 hPa. The relative vorticity at 700 hPa, which is used in the modified Kuo scheme, is specified to be $1 \times 10^{-5}\ s^{-1}$.

Figure 4 shows the prescribed divergence at the model top and the one-time step precipitation simulated by the column model with the continuous version of moist processes for a 6-hour interval. Since the time step is set to 1200 s as in the T42L12 FSU-GSM, the number of time levels is 19 with $N = 18$. The divergence and precipitation at time level 0 are set to zeros. The discontinuous version yields almost the same simulation of precipitation, as may be expected from Fig. 3. Most of the precipitation is brought about by deep cumulus convection, and the amount of total precipitation during the 6-hour assimilation window is 17.0 mm. An inspection of Fig. 3 reveals that this value of precipitation corresponds to a heavy precipitation case. A close correspondence between the divergence and precipitation in Fig. 4 suggests the usefulness of precipitation data for estimating divergence in the tropics.

This 6-hour period is taken as an assimilation window for assimilation experiments. Two types of experiment are carried out with the column model using model-generated data: inversion experiments and analysis experiments. In the former experiments, the true state at the beginning of the assimilation window is retrieved using accurate observations with divergence field specified to the truth, so that the divergence parameters are not included in the control variables and the penalty term is omitted. The purpose of this experiment is to investigate the impact of the removal of discontinuities from the parameterization schemes of the moist processes on the convergence performance of variational data assimilation. In the latter experiments, small random errors are added to observations. The analysis errors at the end of the assimilation window and the estimated divergence field are examined. The purpose of this experiment is to examine the impact of precipitation data on the analysis. Since the physics of the column model includes the moist processes only, the adjoint model is mathematically consistent with the assimilation model. Results obtained by using the continuous version of moist processes are compared with results obtained by using the discontinuous version. Note that the truth is not the same for the continuous version and the discontinuous version of moist processes.

It is assumed that errors in the observational data are uncorrelated and that the statistics of observation errors are homogeneous in the vertical. The observation errors are assumed to be normally distributed with no bias. The following values of the standard deviation of observation

error are used: 1 K for virtual temperature and dewpoint depression, 10^{-4} for the logarithm of surface pressure, and 0.25 mm or 1 mm for one-time step precipitation. Since the true one-time step precipitations are around 1 mm (see Fig. 4), the observation error for one-time step precipitation of 1 mm corresponds to a case of poor quality of precipitation data. The observation error of accumulated precipitation is computed by assuming that observation errors of one-time step precipitation are uncorrelated. Negative values of precipitation data perturbed by random error are replaced by zeros.

Two inversion experiments are carried out. In Experiment A, observations of all the model state variables are available at the beginning and end of the assimilation window. In Experiment B, only observations of one-time step precipitation are available at every time level. The standard deviation of observation error of one-time step precipitation is set to 0.25 mm. The first guess is the reference state perturbed by normally-distributed random error. It is the same as the perturbed observations at the beginning of the assimilation window that are used in the analysis experiments. The first guess is used in the minimization algorithm to compute the cost function and its gradient at iteration 0. Each experiment consists of 20 cases using different perturbations to the reference state.

Three analysis experiments are carried out. In Experiment C, observations of all the model state variables are available at the beginning and end of the assimilation window. In Experiment D, observation of 6-hour accumulated precipitation is available in addition to the observations used in Experiment C. In Experiment E, observations of one-time step precipitation are available at every time level in addition to the observations used in Experiment C. Each analysis experiment is performed for two standard deviations of observation error of one-time step precipitation: 0.25 mm and 1 mm. The first guesses for the model state variables are the observations for them at the beginning of the assimilation window. The first guess for the divergence parameters is set to 1 x 10^{-5} s^{-1} for all time levels. Each experiment is carried out for 20 cases using different perturbations to the observations.

c. Cost function and minimization

In the column model experiment, it is assumed that direct observations of the model state variables, virtual temperature, dewpoint depression and the logarithm of surface pressure, are available at the model vertical levels and that observation errors at different time levels are uncorrelated.

Hence, the observation term in the cost function is the same as the right-hand side of (6) with the observation operators $H_{y_n}(\cdot)$ and $H_{Q_n}(\cdot)$ reduced to identity operators. For later convenience, the observation term is divided into the term for the observational data except for precipitation, J_o, and that for precipitation, J_r, as follows:

$$J_o = \frac{1}{2}\sum_{n=0}^{N} (\boldsymbol{x}_n - \boldsymbol{y}_n)^{\mathrm{T}} \mathbf{O}_{y_n}^{-1} (\boldsymbol{x}_n - \boldsymbol{y}_n) , \tag{64}$$

$$J_r = \frac{1}{2}\sum_{n=1}^{N} (P_n - Q_n)^{\mathrm{T}} \mathbf{O}_{Q_n}^{-1} (P_n - Q_n) , \tag{65}$$

Since the observations are assumed to be uncorrelated with each other, the observation error covariance matrices, $\mathbf{O}_{y_n}$ and $\mathbf{O}_{Q_n}$, are diagonal. These matrices are formally set to infinity for time levels at which observational data are not available.

In the analysis experiments, the divergence parameters in (63) are to be estimated during the minimization process, and the cost function is a function of the divergence parameters, $D_0, \cdots, D_{N-1}$, as well as the model state variables at time level 0, $\boldsymbol{x}_0$. They are the control variables for the minimization. Then, it is necessary to control the time tendency of divergence even in the column model where gravity waves are totally absent, so a penalty term is added to the cost function as in (4). The explicit expression of the penalty term J_c used is

$$J_c = \frac{1}{2} r \sum_{n=0}^{N-2} \left(\frac{D_{n+1} - D_n}{\Delta t} \right)^2 , \tag{66}$$

where Δt denotes the time step for time integration, which is set to 1200 s. The value of the penalty parameter r in (66) is set to 2 x 10^{15} s^4. The cost function is given by the sum of the above three terms:

$$J(\boldsymbol{x}_0, D_0, \cdots, D_{N-1}) = J_o + J_r + J_c . \tag{67}$$

An algorithm for calculating the gradient of the cost function with respect to the control variables is easily obtained by applying the Lagrange multiplier method as in Subsection 2.2. The gradient with respect to $\boldsymbol{x}_0$ is given by (11) - (13) except that the observation operators $H_{y_n}(\cdot)$ and $H_{Q_n}(\cdot)$ are replaced by identity operators. The gradient with respect to the divergence parameters

are given by

$$\frac{\partial J}{\partial D_n} = \left(\frac{\partial F_n}{\partial D_n}\right)^T \lambda_{n+1} + \frac{r}{\Delta t^2} \sum_{m=0}^{N-2} (\delta_{m+1,n} - \delta_{m,n}) (D_{m+1} - D_m) \qquad (n = 0, \cdots, N-1) \qquad (68)$$

where $\delta_{m,n}$ is the Kronecker delta. Note that the model $F_n(\cdot)$ depends on the divergence parameters.

The convergence performance of the minimization process will be presented in terms of number of function calls rather than number of iterations in this section. One function call consists of one forward integration of the assimilation model and one backward integration of the adjoint model. One iteration of the unconstrained minimization algorithm sometimes requires more than one function call due to the line search. Since in variational data assimilation the computational effort for evaluating the cost function and its gradient dominates the cost of the minimization iteration itself, number of function calls may provide a better assessment of convergence performance. The scaling factors for minimization are set equal to the standard deviations of observation errors for the model state variables. The scaling factor for the divergence parameters is set to 5 x 10^{-5} s^{-1}, the same order of the prescribed divergence field (see Fig. 4). The number of updates m in the L-BFGS method is set to 4.

4.2 Results

a. Inversion experiments

Figure 5 shows the values of the cost function for the 20 cases of Experiment A plotted as a function of the number of function calls for the continuous version and the discontinuous version of moist processes. The values are plotted on a logarithmic scale. In this experiment, observations of all the model state variables are available without error at the beginning and end of the assimilation window. It is seen that the minimization procedure converges for all cases even for the discontinuous version, where convergence rates are slightly slower and more case dependent. This result is consistent with that of Zou et al. (1993b). They did not find any difficulties in their 4D-Var experiment using discontinuous parameterization schemes of moist processes. Their result, however, showed that the error of retrieved divergence was increased by including the moist processes, contrary to expectation.

The corresponding result from Experiment B is displayed in Fig. 6. The standard deviation of

observation error of one-time step precipitation is set to 0.25 mm. In this experiment, only observations of one-time step precipitation without error are available at every time level. It is to be noted that the number of observations is 18, less than the number of the control variables, 23. This underdeterminacy implies the existence of multiple minima of the cost function (Zou et al., 1992). In fact, none of the cases do not seem to converge to the truth with several cases exhibiting large fluctuations during the minimization process. Some cases for the discontinuous version do not decrease at all (Fig. 6a). This result indicates that discontinuous parameterization schemes of moist processes cause serious difficulties in variational data assimilation if only precipitation data is assimilated. The convergence performance for the continuous version is much better, with only two cases exhibiting large fluctuations: Case 3 and Case 20 (Fig. 6b). The cost function for these two cases are plotted by a thick line. In Case 3 the cost function oscillates between orders of 10^{-5} and 10^{-3} after the 35th function call, while in Case 20 the oscillation occurs between orders of 10^{-1} and 10^{-3} after the 6th function call.

These oscillations are associated with the discontinuous change in the cloud top height shown in Fig. 2. Figure 7 shows the time sequences of cloud top height and one-time step precipitation at the 13th and 14th function calls for Case 20. The corresponding time sequence at the 15th function call is almost the same as that at the 13th function call. In this case, an abrupt change in cloud top height occurs early in the assimilation window during the minimization process. On the other hand, such a change occurs late in the assimilation windows for Case 3 (not shown). The discontinuity in the specification method for cloud top height seems detrimental to 4D-Var and it is desirable to remove it. It is to be noted that first-order discontinuities in the parameterization schemes, which are left as they are, do not seem to cause serious difficulties with the smooth optimization method. From a practical point of view, elimination of zeroth-order discontinuities might be sufficient for application of 4D-Var.

b. Analysis experiments

In two cases out of the 20 cases, an iteration algorithm that converts virtual temperature and dewpoint depression into temperature and specific humidity failed to converge at a certain stage during the minimization process for the experiments with the continuous version of moist processes. This problem is caused by unacceptable values of virtual temperature or dewpoint depression for

the iteration algorithm to converge. Fitting a smoothing function over the discontinuity may introduce a large gradient to the cost function, which might lead to physically unacceptable values of the control variables during the minimization process. This difficulty indicates that eliminating zeroth-order discontinuity by smoothing does not necessarily work well unconditionally, and some care should be taken. One method to alleviate this difficulty is to broaden the smoothing interval. Another method is to add to the minimization algorithm a subroutine which shortens the step length taken along the gradient of the cost function when unacceptable values occurs for some of the control variables. Results from the assimilation experiments with the FSU-GSM, which will be described in Sections 5 and 6, show that the appropriate control of gravity wave level is helpful to avoid encountering a locally large gradient of the cost function during the minimization process. In the following, results from the analysis experiments will be presented only for the remaining 18 cases.

Figure 8 shows the analysis errors in virtual temperature and dewpoint depression for Experiments C, D and E for the precipitation observation error of 0.25 mm. Analysis error is defined as a difference between the analysis and the truth. The plotted values are the root mean square errors (RMSEs) averaged over the vertical levels and over the 18 cases. It is seen that the analysis errors for the discontinuous version exhibits larger fluctuations, slower convergence, and larger analysis errors than those for the continuous version, even if precipitation data are not assimilated. Assimilation of precipitation data introduces large fluctuations of the analysis errors during the minimization process. This result is consistent with that of the inversion experiments. The impact of precipitation data on the analysis of temperature and moisture is not clear. For example, for the continuous version the analysis error of virtual temperature is slightly decreased with an increase of the number of precipitation data, while that of dewpoint depression is slightly increased. The reason for the negative impact of precipitation data on dewpoint depression analysis is not clear. A clear positive impact of precipitation data was also not obtained from analysis experiments in which the size of observation errors of virtual temperature and dewpoint depression are doubled or the divergence parameters are set to the truth during the minimization process (not shown).

The divergence parameters estimated at the 20th function call are displayed in Fig. 9. They are the averages over the 18 cases. It is found that precipitation data is very useful to evaluate

divergence, provided that it is available at every time level (Experiment E). The continuous version yields a slightly better result than the discontinuous version. However, the estimated divergence by using the 6-hour precipitation (Experiment D) is not better than that obtained by using no precipitation data (Experiment C). It is to be noted that two-time step oscillations of divergence are evident in the results for Experiments C and D, especially for the discontinuous version. If the penalty term is not added to the cost function, the amplitude of these oscillations becomes one order of magnitude larger (not shown).

Figures 10 and 11 show the analysis errors in virtual temperature and dewpoint depression and the divergence parameters, respectively, for the case of the precipitation observation error of 1 mm. Experiment C in these figures are the same as in Figs. 8 and 9. In this case, precipitation data does not have any positive impacts on the analysis of temperature and moisture. However, the estimated divergence using observations of one-time step precipitation at every time levels is still much better than that for the other two experiments in spite of the larger precipitation observation error. The estimated divergence using the 6-hour accumulated precipitation is slightly better than that using no precipitation data in this case, but the impact of 6-hour accumulated precipitation data is not large.

4.3 Discussion

The column model experiment was carried out primarily to examine the impact of removal of zeroth-order discontinuities from the parameterization schemes of moist processes. The results from the experiment are summarized as follows:

- Variational data assimilation with discontinuous parameterization schemes of moist processes exhibits slow convergence and large analysis errors. The convergence performance becomes further deteriorated when precipitation data is assimilated.
- The removal of zeroth-order discontinuities from the moist processes accelerates the convergence of variational data assimilation, and yields a better analysis. The existence of first-order discontinuities in the moist processes does not seem to be very harmful for variational data assimilation.
- Precipitation data is useful to estimate divergence with variational data assimilation, provided that the temporal resolution of the data is sufficiently high.
- The impact of precipitation data on the analysis of temperature and moisture is small in the

column model experiment in which the horizontal advections of temperature and moisture are neglected.

The column model experiment demonstrates adverse effects of discontinuities in the parametrization schemes of moist processes on variational data assimilation. The results shows that it is necessary to eliminate zeroth-order discontinuities from the moist processes for successful application of 4D-Var to the tropical data analysis. A similar conclusion was recently presented by Zou (1997). The existence of zeroth-order discontinuities in parameterization schemes seems to be inappropriate from a physical point of view. If prediction equations are not continuous, the uniqueness of the solution to initial value problems is not ensured. This implies that the solution may strongly depend upon discretization methods or the length of time step in a certain situation, although the likelihood that we encounter such a situation is very small. Another problem is that the solutions to initial value problems of discontinuous prediction equations are discontinuous functions of the initial condition. Such an excessive sensitivity to initial condition is unlikely, because parameterization represents an ensemble mean of the influence of subgrid-scale phenomena on grid-scale fields and, therefore, a possible abrupt change in each realization of subgrid-scale phenomena should be smoothed away in parameterization schemes. Discontinuities in a deep cumulus convection scheme may degrade the predictability of the tropical atmosphere and the midlatitude mesoscale disturbances, but this problem has not been investigated. If this is demonstrated to be the case, we would need continuous parameterization schemes for the prediction model itself, and we should carefully formulate parameterization schemes to retain zeroth-order continuities. Such a change may provide a more auspicious circumstance for successfully developing variational data assimilation. The only justifiable reason to use discontinuous parameterization schemes in atmospheric models may be due to their requiring less computer resources.

Since divergence is not a model state variable in the column model, the third result on the impact on the analysis of divergence may not be directly applied to 4D-Var with a three-dimensional model. However, previous studies by several authors mentioned in the Introduction have shown that satellite imagery data or precipitation data improve the analysis of divergence in the tropics by using physical initialization techniques. Therefore, a similar large impact from precipitation data on the analysis of divergence is expected to be obtained in 4D-Var. The negligible impact on the analysis of temperature and moisture does not imply that precipitation data is useless to estimate

these variables with 4D-Var. The moisture field is closely tied to divergence in the tropics; low level convergence is associated with much moisture due to the horizontal advection of moisture. The column model does not include the horizontal advection. In three-dimensional models it is expected that precipitation data will improve the analysis of moisture field through the horizontal and vertical advection of moisture. This will be demonstrated to be the case in the simulated data experiment.

5. Simulated data experiment

5.1 Method

a. Experimental design

The design of the simulated data experiment is illustrated in Fig. 12. The truth is determined by a simulation with the T42L12 FSU-GSM with the full physics starting from an initialized analysis at 12 UTC 27 July 1979 prepared by the ECMWF. In order to avoid problems associated with model spin-up (Krishnamurti et al., 1988), the simulated atmospheric state 48 hours after the initial time is taken as the target against which results from assimilation experiments are verified. That time is set to t = 0 h. The assimilation window is the 6-hour period from t = -6 h to t = 0 h, and the simulation from t = -18 h to t = -12 h is used as the first guess. The first guess is used in the minimization process to compute the cost function and its gradient at iteration 0. Observations are generated by adding random error to the truth. Forward integration of the assimilation model and backward integration of the adjoint model are iterated to minimize the cost function. The physics of the assimilation model is the reduced physics as well as in the adjoint model. Thus, the assimilation model and the adjoint model are consistent with each other, although some authors used an inconsistent adjoint model (e.g., M. Zupanski, 1993a; Zupanski and Mesinger, 1995). The reason why this procedure was adopted here is because overly optimistic results are obtained if the same model was used to both generate and assimilate the simulated data. Since the assimilation model and the adjoint mode do not include the full physics, they are not perfect models to simulate the truth. Note, however, that the analyses and forecasts are obtained by integrating the full-physics model from the initial condition at t = -6 h determined by the minimization.

Three assimilation methods are compared. The first method uses the assimilation model and the adjoint model with no moist processes included, and assimilates conventional data only. Since

the moist processes are not included, supersaturation of water vapor is not removed during the forward integration. Yet, since the analysis is prepared by using the full-physics model, this method can yield analyses of precipitation and other fields relating to the moist processes. The second method uses the assimilation model and the adjoint model with the moist processes included, and assimilates conventional data only. The third method uses the same models as the second method, and assimilates both conventional data and precipitation data. These three methods are hereafter referred to as the NOMP, MP and MP+RR method, respectively. A comparison of results from the three methods will reveal the importance of moist processes for 4D-Var in the tropics and the benefits of assimilating precipitation data.

b. Simulated observations

An idealized observation network is used in which the observations are provided at model Gaussian grid points and model vertical levels. The observation network is sparse in the tropics and the Southern Hemisphere. North of 30 °N it is assumed that conventional observations of wind, temperature, relative humidity, and surface pressure are available at all model levels at the beginning (t = -6 h) and end (t = 0 h) of the assimilation window. These observations are intended to correspond to radiosonde observations in the real observation network. South of 30 °N it is assumed that there are no radiosonde stations. Wind observations are available only at σ = 0.2 and σ = 0.85, corresponding to cloud drift winds measured by geostationary satellites. Temperature observations at all model levels and surface pressure observations are assumed to be available as in the north of 30 °N region. The former observations are intended to correspond to vertical temperature soundings provided by polar-orbiting satellites; the latter to ship, buoy, and other surface observations. The above three observations are provided at the beginning and end of the assimilation window. Relative humidity observations are assumed not to be available. Instead, one-hour accumulated precipitation retrieved from satellite remote sensing is assumed available every hour in the south of 30 °N region.

Observation error statistics are assumed to be homogeneous and independent of each other. The observation errors are assumed to be normally distributed with no bias. Negative values of observed relative humidity and 1-hour precipitation are replaced by zeros. The standard deviations of observation error for wind, temperature, relative humidity, and surface pressure are set to 3 m s^{-1}, 1 K, 10 %, and 1 hPa, respectively. These values may be too small for the observations south of 30 °N where radiosonde observations are assumed not to be available. For example, the operational 3D-Var system at the National Centers for Environmental Prediction (NCEP) uses observation errors of 6.1 m s^{-1}, 3.9 m s^{-1}, and 4 - 5 K for high cloud drift wind, low cloud drift wind, and satellite retrieved temperature, respectively (Parrish and Derber, 1992). These smaller values are used in the present study to ensure that the observation error is not much larger than error in the first guess field (see Fig. 19).

The standard deviation of observation error for 1-hour accumulated precipitation is set to 0.5 mm for the area average over one Gaussian grid box (approximately 2.8125 ° x 2.8125 °). The observation error is added to the true precipitation only if the latter is non-zero. This value of precipitation data error is not very different from estimated errors for several satellite-based precipitation retrieval algorithms (e.g., Adler and Negri, 1988; Liu and Curry, 1992; Kummerow, 1994a, b). For example, the RMSE of SSM/I precipitation rates determined by an algorithm proposed by Liu and Curry (1992) is 0.62 mm h^{-1} over the ocean for an area average over a 1.25 ° x 1.25 ° domain. Precipitation rates retrieved from IR observations are likely to be less accurate. The use of the relatively small value for the precipitation error may be partly justified by the small magnitude of errors in the other observations specified in the tropics and the Southern Hemisphere. This small value of the precipitation error maintains the relative importance of precipitation data to the other observations in these regions.

c. Cost function and minimization

The cost function is given by (4) with the observation term, the first term on the right-hand side of this equation, replaced by a more specific form given by (6). To cancel the horizontal inhomogeniety of the Gaussian grid distribution, the summation over the Gaussian grid points in the observation term is replaced by the area mean over the globe as in the penalty term. Since the observations are assumed to be uncorrelated with each other, the observation error covariance matrices are diagonal. Since the observations are available at model grid points, the observation operators in the cost function consists only of inverse spectral transform to go from spectral to grid-point representation and conversions from model state variables to observed variables. The gradient of the cost function is given by (11) - (13) with additional contributions from the penalty

term. These additional terms due to the penalty term are easily computed, because the time tendency of divergence is approximated by a finite difference scheme and the divergence is one of the model state variables of the FSU-GSM.

The initial condition for the FSU-GSM are originally provided on the Gaussian grid, so the minimization is performed in physical space, not spectral space for simplicity. Therefore, the control variables are the model state variables, vorticity, divergence, virtual temperature, dewpoint depression and the logarithm of surface pressure, at the Gaussian grid points. It may be desirable to change the control variables from the grid point values on the Gaussian grid to the spectral coefficients since the FSU-GSM is a spectral model. Such an approach is taken in the real data experiment described in the next section. It is expected that this modification will considerably reduce memory requirements and that the efficiency of the minimization will be improved. The maximum number of iterations in the minimization is set to 30. Scaling factors for the gradient of the cost function with respect to vorticity, divergence, virtual temperature, dewpoint depression, and the logarithm of surface pressure are set to 10^{5} s^{-1}, 10^{5} s^{-1}, 1 K, 2 K, and 10^{3} divided by the square root of the Gaussian weight, respectively. The number of updates m in the L-BFGS method is set to 5.

5.2 Results

a. Verification of gradient calculation

Prior to discussing results from the experiment, the correctness of the gradient calculation using (5) is verified. If the cost function has continuous partial derivatives of order two in the neighborhood of $\boldsymbol{x}_0$, application of the Taylor's theorem to the cost function yields

$$\varphi(\alpha) = \frac{J(\boldsymbol{x}_0) - J(\boldsymbol{x}_0 - \alpha \mathbf{S}^2 \nabla_{\boldsymbol{x}_0} J)}{\alpha (\nabla_{\boldsymbol{x}_0} J)^{\mathrm{T}} \mathbf{S}^2 \nabla_{\boldsymbol{x}_0} J} = 1 + O(\alpha) , \tag{69}$$

where α is a small nondimensional parameter and $\mathbf{S}$ is a scaling matrix for the gradient of the cost function. The scaling matrix is a diagonal matrix, its diagonal elements consisting of the scaling factors introduced at the end of the previous subsection. For sufficiently small α but not too close to the machine accuracy, the value of $\phi(\alpha)$ should be close to unity if the gradient $\nabla_{\boldsymbol{x}_0} J$ is correctly calculated.

The variation of $\phi(\alpha)$ with log α for the NOMP and MP methods are plotted in Fig. 13. In this example, $\boldsymbol{x}_0$ is taken as the simulated atmospheric state at t = -48 h in Fig. 12. The observations available for the north of 30 °N region are assumed to be available over the globe for the verification purpose. The observation errors are not added, nor is the penalty term included. The verifications are carried out on a CRAY-YMP computer which has intrinsic double precision. Although it may be better to verify that $\phi(\alpha) - 1$ is of order α, this is not easy for 4D-Var with the moist processes included due to the fact that tangent linear approximation is less valid in this case. It is seen from Fig. 13 that the value of $\phi(\alpha)$ for the NOMP method is almost equal to unity for α less than 10^{-1}, while for the MP method $\phi(\alpha)$ is close to unity for α less than 10^{-5}. Larger values of $\phi(\alpha)$ for α less than 10^{-10} result from numerical cancellation error, so they are to be neglected. Since the gradients of the cost function for the two cases are not very different (e.g., $|\mathbf{S}\nabla_{\boldsymbol{x}_0} J|$ = 293.0 for the NOMP method, 278.9 for the MP method, and the cosine of angle between the two gradient vectors is 0.958), the difference in $\phi(\alpha)$ may primarily reflect the difference in the shape of the cost function. In particular, the existence of a local minimum around $\alpha = 10^{-4}$ for the MP method indicates that the cost function in the direction of the scaled gradient is not convex. Given the highly nonlinear character of physical parameterizations, the cost function including moist processes has a more complicated shape with, possibly, multiple minima. These factors render the minimization for 4D-Var with the moist processes included more difficult than for 4D-Var without the moist processes.

b. Convergence performance

Figure 14 shows the variation of the cost function as a function of the number of iterations in the minimization process up to 30 iterations for the three assimilation methods: the NOMP, MP and MP+RR methods. Convergence performance is evaluated for two values of the penalty parameter r. The upper panel is for $r = 4 \times 10^{17}$ s^4; the lower panel is for $r = 1 \times 10^{18}$ s^4. The cost function is divided into three terms as in (67): the observation term for observational data except for precipitation data, J_o, the observation term for precipitation data, J_r, and the penalty term, J_c. The second term

J_r is absent for the NOMP and MP methods. In each panel the upper three lines and the middle three lines are the values of J_o and J_c for the three methods, respectively, and the lower line is the values of J_r for the MP+RR method. Figure 15 displays the corresponding variation of the scaled gradient of the cost function. The parenthetical values in the legend of each panel indicate the number of function calls required for 30 iterations. One function call consists of one forward integration of the assimilation model and one backward integration of the adjoint model. One iteration of the unconstrained minimization algorithm sometimes requires more than one function call due to the line search.

The slopes of J_o and the norms of the scaled gradient in Figs. 14 and 15 indicate that convergence has not been reached for the three methods. The norms of the gradient do not show large decreases even after 30 iterations. These slow convergence rates may be primarily due to the use of the suboptimal values as the scaling constants. The inclusion of the penalty term in the cost function also decelerates the convergence (Courtier and Talagrand, 1990; Thépaut and Courtier, 1991). Since the observations have errors and are not available over the globe, exclusion of the penalty term or use of very small penalty parameters leads to numerous noisy structures, consisting mostly of gravity waves, in the analysis. A point to note here is that a decreasing cost function does not necessarily mean that the analysis is still being improved. In contrast, a continually decreasing cost function can, at times, correspond to an increasing level of gravity wave noise (Thépaut et al., 1993a). Although the iteration process does not converge in the results presented here, the minimization with 30 iterations yields analyses that are sufficiently close to the truth, as will be described later.

The smaller penalty parameter makes the value of J_c after 30 iterations almost the same as the value before the iteration (iteration 0) for the three methods, with a slight increase of J_c after around 10 iterations (Fig. 14a). J_r is only a small fraction of the total cost function. The inclusion of the moist processes considerably slows down the convergence. The MP method requires up to 49 function calls for 30 iterations and, still, the value of J_o after 30 iterations is much larger than that for the NOMP and MP+RR methods. The step-like variation in J_o for the MP method results from the occasional occurrence of small step lengths in the line search. These instances are often associated with large gradients in the cost function such as at the 5th and 13th iterations (Fig. 15a). The MP+RR method also encounters a large gradient in the cost function at the 4th iteration. Comparison

of results from the MP and MP+RR methods reveals that assimilating precipitation data accelerates the convergence of 4D-Var with the moist processes included. However, this is not always the case when real precipitation data are assimilated, as will be shown in the real data experiment. In the simulated data framework, observed precipitation data are fairly consistent with the model. Thus, the acceleration of convergence by assimilating precipitation data may be an expected result.

Increasing the penalty parameter decelerates the convergence for the NOMP method, while it considerably improves the convergence performance for the MP and MP+RR methods (Fig. 14b). Although the number of function calls for the MP method is greatly reduced, it is still larger than that for the MP+RR method, still indicating again the benefit of assimilating precipitation data. The values of J_o for the three methods are almost the same, with J_o for the MP and MP+RR methods a little larger than that for the NOMP method. This is due to acceleration of convergence for the MP method and deceleration for the NOMP method. There are no large differences in the values of J_o for the MP+RR method for the two penalty parameters. The values of J_c decrease during the first several iterations for the three methods and increase after around 10 iterations. As will be shown later, the larger penalty parameter yields a better analysis in the tropics after 30 iterations even for the NOMP method. Figure 15b shows that large gradients in the cost function are not encountered during the minimization processes for the MP and MP+RR methods when using the larger penalty parameter. This result suggests that the penalty term helps to avoid encountering locally large gradients and, as a result, increases the efficiency of 4D-Var with the moist processes included. Locally large gradients in the cost function arise from the parameterized moist processes, and the moist processes tend to generate a large time tendency of divergence. Therefore, locally large gradients can be avoided by including a penalty term which suppresses a large time tendency of divergence. This result will be confirmed by the real data experiment described in Section 6.

c. Analyzed fields

Analysis errors in the surface pressure, 200 hPa divergence, 6-hour accumulated precipitation, and total precipitable water at t = 0 h in the tropics are summarized in Table 1. Analysis error is defined as a difference between the analysis and the truth. The RMSEs over the latitudinal belt 30 °N to 30 °S are listed for the two penalty parameters, $r = 4 \times 10^{17}$ s^4 and $r = 1 \times 10^{18}$ s^4.

Comparison of the two cases reveals that the larger penalty parameter tends to yield a better analysis in the tropics even for the NOMP method. Recall that the NOMP method gives a slower convergence rate with the larger penalty parameter. It is also seen that 4D-Var without the moist processes yields poorer analyses of precipitation and total precipitable water than the first guess. On the other hand, 4D-Var with the moist processes included yields better analyses of those fields with both the penalty parameters, despite the degraded convergence performance. This result strongly indicates the importance of moist processes for 4D-Var in the tropics. Although moisture observations are not available in the tropics, analysis errors in total precipitable water for the MP and MP+RR methods are smaller than the first guess error. This is due to the fact that 4D-Var uses all available observations and the model dynamics to retrieve the moisture field. It is also seen that assimilating precipitation data further improves the moisture analysis as well as the analysis of 200 hPa divergence and precipitation. In the following, results are presented only for the case of $r = 1 \times 10^{18}$ s^4 which yields better analyses.

Figures 16 - 18 compare the analyses of 6-hour accumulated precipitation, 200 hPa divergence, and mean sea level pressure, respectively, at $t = 0$ h after 30 iterations. The top panel in each figure shows the truth, as defined from the full-physics simulation. A developing tropical cyclone east of the Philippines (Fig. 18a) brings heavy precipitation (Fig. 16a) and an accompanying pattern of strong divergence at 200 hPa (Fig. 17a). This simulated cyclone corresponds to Super Typhoon Hope (27 July - 3 August 1979) in the real atmosphere, so that it is referred to as a typhoon hereafter. Another tropical disturbance with heavy precipitation and strong divergence is found in the northern Arabian Sea. 4D-Var without the moist processes considerably underestimates precipitation in the tropics and almost completely fails to retrieve the areas of heavier precipitation (Fig. 16b). The strong divergence associated with the typhoon and the other tropical disturbance in the northern Arabian Sea is also considerably underestimated as well as strong divergence in the eastern tropical Atlantic (Fig. 17b). The sharp structure of the typhoon that is seen on the mean sea level pressure map is not well reproduced (Fig. 18b).

The inclusion of the moist processes greatly improves the analysis of 6-hour accumulated precipitation (Fig. 16c). Regions of heavy precipitation are well retrieved, although they are still underestimated in several areas and the locations of maximum precipitation east of the Philippines and in the Bay of Bengal are misplaced. The analysis of strong divergence in the northern Arabian

Sea and eastern tropical Atlantic is improved (Fig. 17c). The retrieved divergence in the region from the Bay of Bengal to the western tropical Pacific is still not satisfactory. The center of the divergence associated with the typhoon is erroneously shifted eastward, corresponding to the misplaced maximum of 6-hour accumulated precipitation. The central pressure of the simulated typhoon is slightly underestimated (Fig. 18c). Assimilation of precipitation data with 4D-Var works very well. Taking into account the size of observation error in precipitation data, agreement between the analyzed precipitation and the truth is remarkable (Fig. 16d). The divergence field in the region from the Bay of Bengal to the western tropical Pacific is well retrieved (Fig. 17d), although the divergence associated with the typhoon is slightly underestimated. The central pressure of the retrieved typhoon is almost the same as the truth (Fig. 18d).

The vertical distributions of analysis error in vorticity, divergence, temperature, and relative humidity in the tropics at $t = 0$ h are displayed in Fig. 19. Since wind observations are available at $\sigma = 0.20$ and $\sigma = 0.85$ in the tropics, it is reasonable that the analysis errors in vorticity and divergence at these two levels are considerably decreased from the first guess errors. Note that analysis errors in wind field at other levels are also decreased from the corresponding first guess errors without using specific analysis structure functions. This is a nice property of 4D-Var. A similar advantage of 4D-Var is seen in the relative humidity analysis error. The errors for the MP and MP+RR methods are smaller than the first guess error in the lower troposphere despite the fact that no moisture observations are available in the tropics. This advantage has been discussed earlier when referring to the RMSE of total precipitable water in Table 1. Positive impacts from including the moist processes and assimilating precipitation data are evident in the analyses of divergence, moisture, and lower tropospheric vorticity. In particular, the winds in the tropical planetary boundary layer are better analyzed. This result may reflect the fact that circulation in the tropical planetary boundary layer is closely related to cumulus convection. The impacts, however, are not clear for the temperature analysis. This is partly because temperature observations with relatively small errors are assumed to be available at all model levels and, as a result, the impact of model dynamics and the other observations on the temperature analysis is negligible. Such insensitivity of the temperature analysis to precipitation data was also seen in the column model experiment (see Figs. 8 and 10).

d. Forecast experiment

It is expected that a better analysis will produce a better forecast. To confirm this statement for the present study, a 72-hour forecast experiment was carried out by using the initial conditions at $t = 0$ h prepared by the simulated data experiment with $r = 1 \times 10^{08}$ s^4. The full-physics model was used for the forecast experiment. Initialization is not necessary, because undesirable gravity wave noise is already suppressed by inclusion of the penalty term in the cost function.

Results of the forecast experiment are presented only for the MP and MP+RR methods. The forecast starting from the NOMP analysis did not run successfully due to unrealistic values of dewpoint depression at lower levels around the Andes. As mentioned in Subsection 3.1, the FSU-GSM tends to produce noise around the Andes. The unrealistic values prevented an iterative algorithm which converts virtual temperature and dewpoint depression into temperature and specific humidity from converging after several iterations and, as a result, the time integration was stopped before reaching forecast hour 24. It is likely that the larger value of the penalty parameter is not appropriate for the NOMP method, since the convergence performance is poor compared to that for the NOMP method with the smaller penalty parameter (see Fig. 14). It is also to be noted that the value of J_r after 30 iterations for the NOMP method is only a third of the value before the iteration is started. The combination of the penalty term constraint being too strong and the absence of the moist processes may distort some part of the analyzed atmospheric state and lead to the failure of the forecast. In fact, a 72-hour forecast starting from the NOMP analysis with the smaller penalty parameter did run successfully. The exact cause of the forecast failure, however, is not known.

The skill of predicted 24-hour precipitation totals in the tropics is shown in Fig. 20. The benefit of assimilating precipitation data into the subsequent forecast is evident. The forecast starting from the MP+RR analysis is more skillful during the entire forecast period. The large difference in the errors for the 48 - 72 h period is primarily due to differences in precipitation caused by the typhoon which was located to the east of the Philippines at $t = 0$ h (see Fig. 18). Figure 21 compares the true and forecasted 24-hour precipitation totals in the Asian monsoon region for $t = 48$ - 72 h. Heavy precipitation around the East China Sea is brought about by the typhoon. Another area of heavy precipitation around southern Pakistan is due to the tropical disturbance which was located in the northern Arabian Sea at $t = 0$ h. The forecast staring from the MP analysis erroneously shifts the heavy precipitation area around the East China Sea slightly southwestward and overestimates the maximum value of precipitation. The forecast also fails to

predict the heavy precipitation around southern Pakistan. The assimilation of precipitation data improves both deficiencies, although the heavy precipitation due to the typhoon is still misplaced southwestward and the other area of heavy precipitation is underestimated. Prediction of precipitation in eastern India and the eastern Indochina Peninsula, however, is slightly degraded. Note that there are almost no differences in areas of weak precipitation. Positive impacts from assimilating precipitation data are also found in the prediction of other variables such as mean sea level pressure (not shown).

5.3 Discussion

The FSU-GSM with the smoothed moist processes is used primarily to investigate the effect of the penalty term for accelerating convergence of 4D-Var with the moist processes included. The results from the experiment are summarized as follows:

- The penalty term for suppressing gravity wave noise helps to increase the efficiency of 4D-Var with the moist processes included by avoiding locally large gradients in the cost function during the minimization process.

- Despite the fact that wind observations in the tropics are assumed to be available only at two model levels, the wind analysis error of 4D-Var is decreased from the first guess error at all model levels without using specific vertical structure functions. The moisture analysis error in the lower troposphere is also decreased from the first guess error without assimilating moisture observations.

- 4D-Var with the moist processes included yields a much better analysis in the tropics despite a slower convergence rate than 4D-Var without the moist processes. Inclusion of the moist processes improves analyses of divergence, moisture, and lower tropospheric vorticity. The wind field in the planetary boundary layer is better analyzed in the tropics, and the structure of a tropical cyclone is well retrieved.

- Further improvements result from assimilating precipitation data. The 72-hour forecast experiment shows a benefit of assimilating precipitation data in the tropics. However, major improvements are generally confined to areas of heavy precipitation.

Although the above results are derived from assimilation experiments using simulated data for one synoptic situation only, the first result strongly suggests that control of gravity wave level is generally helpful in accelerating the convergence of 4D-Var with moist processes included. This

result makes a contrast to results from 4D-Var experiments with dry-adiabatic models by previous authors. For instance, Courtier and Talagrand (1990) reported that when a penalty term to eliminate gravity wave noise was included in the cost function, the convergence turned out to be extremely slow. To alleviate this difficulty they devised a method of introducing a nonlinear normal mode initialization in 4D-Var in addition to the penalty term. The reason why adding the penalty term decelerates the convergence of 4D-Var is that it makes more difficult for the model trajectory to fit observations due to the additional constraint. The present result, however, shows that this is not the case if moist processes are included and it is applied to the tropics, where deep cumulus convection plays an essential role. Deep cumulus convection is a highly nonlinear process, so unrealistically strong convective activities that may be encountered during the minimization process may significantly decelerate convergence. The penalty term is helpful to suppress unrealistically strong convection through the control of the time tendency of divergence and, therefore, to increase the efficiency of 4D-Var. However, too strong constraint will deteriorate the convergence due to the same reason as in 4D-Var experiments with dry models, as will be demonstrated in the real data experiment.

The present study demonstrates the benefits of including moist processes for 4D-Var in the tropics, but difficulties still arise with the strong nonlinearities and threshold processes associated with physical parameterizations, particularly the parameterization of deep cumulus convection. The minimization failed to converge in some cases for the MP and MP+RR methods in which the penalty term was not included or the 6-hour simulation starting from t = -48 h was used as the first guess. The failures were primarily caused by the occurrence of unrealistic dewpoint depression values. Similar minimization failures were also experienced in the column model experiment using the smoothed moist processes, but were not experienced when using the discontinuous moist processes. As discussed at the end of Subsection 2.1, removal of discontinuities from parameterization schemes may introduce a locally large gradient in the cost function which might lead to physically unacceptable values of some control variables during the minimization processes. Further smoothing is desirable for successfully developing 4D-Var in the tropics, but care must be taken not to degrade the performance of the physical parameterizations. It is also desirable to devise an effective method for constraining the value of control variables, especially the moisture variable, to realistic values, but such a approach usually requires further demands upon computer resources and may not be feasible.

6. Real data experiment

6.1 Method

a. Experimental design

A schematic diagram of experimental design for the real data experiment is presented in Fig. 22. The 12-hour period of 0 - 12 UTC 22 August 1992 (t = -12 - 0 h) is chosen as the assimilation window. The observations assimilated are radiosonde data over the globe and precipitation data over the tropical oceans. The precipitation data are instantaneous precipitation rates retrieved from the SSM/I observation by the Defense Meteorological Satellite Program (DMSP) satellites. Details of the observational data will be described later. The SSM/I precipitation data do not entirely cover the tropical oceans during this 12-hour period. Increasing the length of the assimilation window is possible for the precipitation data to cover more area, but the adverse effects of model error on 4D-Var will also increase. The nonlinear normal mode initialization with physics (Kitade, 1983) is applied to the operational analysis prepared by the NCEP at the beginning of the assimilation window (t = -12 h). This initialized analysis serves as the first guess for the model state at t = -12 h.

Different from the simulated data experiment, the FSU-GSM with the full physics is used as the assimilation model. Since the 4D-Var formulation usually assumes that the model is perfect, the full-physics model is more appropriate as the assimilation model when assimilating real observations. In this case, however, the adjoint model with the reduced physics, which uses the basic states produced by a forward integration of the assimilation model with the full physics, does not produce exact gradients of the cost function. Such an approximate procedure has already been adopted by M. Zupanski (1993) and Zupanski and Mesinger (1995). They showed that this procedure produced better analyses than another procedure in which the reduced-physics model consistent with the adjoint model is used as the assimilation model. As in the simulated data experiment, three assimilation methods are compared to assess the importance of moist processes for 4D-Var in the tropics and to evaluate the benefits of assimilating precipitation data. The NOMP method uses the adjoint model that lacks the moist processes. The MP method uses the adjoint model with the moist processes included but does not assimilate the SSM/I precipitation rates. The MP+RR method uses the same adjoint model as in the MP method but assimilates the SSM/I precipitation rates.

Since the background term is not included in the cost function as in the previous two experiments,

a comparison with operational analysis at an NWP center is not intended here although real data are assimilated. In an operational analysis, information obtained from the previous analysis cycle is effectively utilized through the background term. A 4D-Var formulation without the background term makes it difficult to fully utilize this information except for using it as the first guess. As a result, the quality of analysis produced from the assimilation experiments performed here may be worse than the quality of operational analysis.

b. Radiosonde data

Radiosonde data in the assimilation window of 0 - 12 UTC 22 August 1992 are taken from the NCEP upper air data subsets archived at the National Center for Atmospheric Research (NCAR). These data represent what NCEP collected before their cutoff time when they commenced the operational analysis. Assimilated data are geopotential height at 1000 hPa, wind and temperature at ten standard levels below 100 hPa and relative humidity at six standard levels below 300 hPa. Geopotential height data at the other levels are not assimilated because they are not independent of temperature, moisture, and 1000 hPa geopotential height data. Since the model top is placed at σ = 0.1, upper air data above 100 hPa are not assimilated.

The observation error standard deviations used are given in Table 2. They are almost the same as those used in the NCEP operational 3D-Var system. The observation error standard deviations at 1000, 700, 500, 300, and 100 hPa except for the geopotential height are taken from Table 1 of Parrish and Derber (1992), and those at the other levels are calculated by applying a spline interpolation in the logarithm of pressure. The observation error standard deviation of geopotential height at 1000 hPa is taken from Table 5 of Dey and Morone (1985). Observation errors in the radiosonde data are assumed to be independent of each other, so that the observation error covariance matrix for these data is diagonal.

A simple quality control procedure is applied to the radiosonde data by using the NCEP operational analyses at the beginning and end of the assimilation window. The operational analyses are prepared on a 2.5 ° x 2.5 ° grid at the fifteen standard pressure levels. The analyses are linearly interpolated to observation times, and then bicubic horizontal interpolation is performed from analysis grid points to observation points. If the magnitude of the difference between an observational datum and interpolated NCEP analysis is greater than twice the observation error, the observational

datum is rejected. Table 3 summarizes the numbers of reported data and rejected data. Zonal and meridional components of the wind are rejected separately in the present study, although simultaneous rejection or acceptance of both wind components would be more appropriate for quality control of radiosonde wind data. Rejection rates for wind and temperature data are about 5 %. The small rejection rate for relative humidity data is due to the use of a relatively large observation error standard deviation. The total number of radiosonde data that are assimilated during the 12-hour assimilation window is 32681. Since this number is much less than the degrees of freedom in the model, 86879, the minimization problem is underdetermined. The inclusion of the penalty term in the cost function restricts possible solutions and alleviates this underdeterminacy. Figure 23 displays the horizontal distribution of radiosonde stations from which data are assimilated in the experiments. The radiosonde stations are densely distributed over land in the Northern Hemisphere, while they are very sparse over the tropical oceans except for the western tropical Pacific.

c. SSM/I precipitation rates

The SSM/I is a passive microwave radiometer operating on several DMSP satellites since 1987. The DMSP satellites are in sun-synchronous orbit at about 860 km elevation with a 98.8 ° inclination. Their passes cover each earth location twice per day. Two DMSP satellites, F10 and F11, were in orbit during the assimilation window of 0 - 12 UTC 22 August 1992. The F10 satellite was launched into a slightly elliptical orbit in November 1990. Its equator crossing time drifted slowly from about 0755 local time (LT) in May 1991 to about 0830 LT in May 1992. The F11 satellite was launched into a near-circular orbit in December 1991 with equator crossing times of about 0600 and 1800 LT. The SSM/I has four channels with frequencies of 19.35, 22.235, 37.0, and 85.5 GHz with nadir horizontal resolution of 15 - 60 km depending on the frequencies. All measurements except for the 22-GHz channels are obtained with dual polarization. The addition of the 85-GHz channel makes the SSM/I unique in comparison to previous passive microwave instruments. This channel provides the highest sensitivity for measuring precipitation over land. A conical electrical scan sweeps out a 1400 km swath aft of the satellite. The details of the SSM/I are described in Hollinger et al. (1987).

The SSM/I precipitation rates over the tropical oceans are obtained from a regression formula developed by Olson et al. (1990). This algorithm was used by Krishnamurti et al. (1993, 1995) for

their physical initialization procedure. The algorithm for precipitation over ocean is as follows:

if

$$-11.7939 - 0.02727\, T_B^{37V} + 0.09920\, T_B^{37H} > 0\,, \tag{70}$$

then

$$R = \max\,[\,\exp(3.06231 - 0.0056036\, T_B^{85V} + 0.0029478\, T_B^{85H} - 0.0018119\, T_B^{37V} - 0.00750\, T_B^{22V} + 0.0097550\, T_B^{19V}) - 8.0,\, 0\,]\,, \tag{71}$$

else

$$R = 0\,, \tag{72}$$

where R is the precipitation rate in mm h^{-1}, T_B the brightness temperature in Kelvin. The superscripts 85, 37, 22, and 19 refer to frequencies in GHz, and the superscripts H and V refer to horizontal and vertical polarizations. A total of 1365 collocated SSM/I and tropical radar measurements over ocean were used to obtain the above regression formula. The bias and RMSE of the precipitation rate estimate are 0.13 mm h^{-1} and 0.59 mm h^{-1}, respectively. The correlation coefficient between the SSM/I measurement and the radar measurement is 0.772. Gairola and Krishnamurti (1992) found a close agreement between rain gauge based precipitation at island stations and SSM/I based oceanic precipitation derived from the above algorithm. Olson et al. (1990) also proposed regression formulae for retrieval of precipitation rates over land. Gairola and Krishnamurti (1992), however, found that over land areas the SSM/I precipitation derived from the Olson algorithm showed large departures from rain gauge measurements. Therefore, SSM/I precipitation rates over land are not assimilated in the present study.

The SSM/I precipitation rates are gridded to a 1 ° x 1 ° mesh over ocean between 35 °N and 35 °S by taking an area average. Since model precipitation is computed on the Gaussian grid with an approximate horizontal resolution of 2.8125 ° x 2.8125 °, an area average over a Gaussian grid box appears more appropriate. However, data coverage of the SSM/I precipitation rates is so limited that an area average over a larger domain may dilute the information carried by the original estimates from the Olson algorithm. The 12-hour assimilation window is divided into 12 one-hour intervals. The SSM/I estimates of instantaneous precipitation rates belonging to one of the 1-hour intervals are regarded as hourly mean precipitation rates over this 1-hour interval. Bicubic interpolation

is applied to interpolate model precipitation rates from model grid points to observation points. In view of the fact that model precipitation rates are computed in physical space not spectral space, bilinear interpolation might appear more appropriate. Results from an assimilation experiment using bilinear interpolation, which will be described later, indicate the opposite. A quality control procedure is not applied, nor is the bias in the Olson algorithm corrected.

As mentioned in Introduction, one of the advantages of 4D-Var over physical initialization in assimilating precipitation data is the proper treatment of precipitation data error. However, assigning appropriate observation error standard deviation for precipitation rates is not easy, primarily because of the skewed probability distribution of precipitation rates. The necessity of transforming the precipitation rates to make the observation error more normally distributed was pointed out in Subsection 2.2. In the present study such a transformation are not applied and a constant observation error standard deviation is used for the sake of simplicity. Taking the area average of the SSM/I precipitation rates over the 1 ° x 1 ° grid box reduces the RMSE of the original SSM/I estimates from the Olson algorithm. On the other hand, representativeness error should be added to the observation error because the horizontal resolution of model precipitation rates is much larger than that of the SSM/I precipitation rates. There is another error that results from regarding instantaneous precipitation rates as hourly mean precipitation rates. In the present study the observation error standard deviation of the SSM/I precipitation rates is set to 1 mm h^{-1} as a crude estimate. Observation errors in the SSM/I precipitation rates are inherently mutually correlated because it is obtained by the same sensing device. However, the observation errors are assumed to be horizontally uncorrelated for simplicity. Thus, the observation error covariance matrix for the precipitation data is also diagonal. The total number of the SSM/I precipitation rates available over the assimilation window is about 2 x 10^4, so the minimization problem is still underdetermined even if the precipitation data are assimilated. Since the radiosonde data and the SSM/I precipitation rates are independent of each other, the entire observation error correlation matrix is also diagonal.

Figure 24 shows the SSM/I precipitation rates over the assimilation window at 1-hour intervals. It is seen that almost the same regions are observed twice by the two DMSP satellites. For example, a region of heavy precipitation in the equatorial Indian ocean was observed during the period -12 h to -11 h (0 - 1 UTC 22 August 1992) and the period -8 h to -7 h (4 - 5 UTC). Regions of heavy precipitation east of New Guinea were observed during the period -6 h to -5 h (6 - 7 UTC) and the

period -2 h to -1 h (10 - 11 UTC). Figure 25 shows the composite of the SSM/I precipitation rates over the assimilation window. Mean values are plotted in areas where more than one observation is available. For later convenience, only the latitudinal belt between 30 °N and 30 °S is displayed. It is seen that the SSM/I observations do not fully cover the tropical oceans despite the fact that two DMSP satellites are in orbit and that a 12-hour assimilation window is taken. About a third of the tropical oceans remain unobserved. In this period there are no strong tropical disturbances over the tropical oceans except for Hurricane Andrew in the western Atlantic. A small region of precipitation at 25 °N, 70 °W adjacent to a data-void region is associated with this compact tropical cyclone. Unfortunately the center of the hurricane is not observed by the DMSP satellites during the assimilation window. However, even if the SSM/I observations covered the entire region, the horizontal scale of Hurricane Andrew is too small to be properly resolved by the T42L12 global spectral model. A region of heavy precipitation around 5 °N, 165 °E is associated with a tropical depression that is developing into Super Typhoon Omar.

OLR data obtained from the infrared radiometer aboard NOAA polar-orbiting satellites have more data coverage in the tropics, although the OLR data have a poorer physical basis for estimating precipitation rates. Figure 26 displays the night pass and day pass composites of OLR for 22 August 1992. These composites are prepared on a 2.5 ° x 2.5 ° grid by the National Environmental Satellite, Data, and Information Service (NESDIS) from measurements by the NOAA 11 and NOAA 12 satellites. Equator crossings of the two satellites are at 0230 LT and 0730 LT for night pass and at 1430 LT and 1930 LT for day pass, respectively. Thus, OLR data used for the night pass composite in the Atlantic region between 70 °W and 40 °E and those used for the day pass composite in the western Pacific region between 110 °E and 180 °E fall into the assimilation window of 0 - 12 UTC 22 August 1992. Therefore, the OLR composites in these two regions can be compared with the SSM/I composite in Fig. 25. A close agreement between the SSM/I composite and the OLR composites in the regions suggests that the OLR composites can be used to qualitatively verify the precipitation analysis in SSM/I data-void areas.

d. Cost function and minimization

The cost function is given by (4) with the observation term replaced by a more specific form of (6) as in the simulated data experiment. Since the observations are assumed to be uncorrelated

with each other, the observation error covariance matrices are diagonal. The observation operators consist of inverse spectral transforms to go from spectral to grid-point representation, conversions from model state variables to observed variables, bicubic horizontal interpolation from model Gaussian grid points to observation points, and cubic vertical interpolations from model sigma levels to observation levels. The gradient of the cost function is given by (11) - (13) with additional contributions from the penalty term.

The minimization is performed in spectral space in contrast to the simulated data experiment, where the control variables were defined in physical space. Minimization in spectral space may be more appropriate because the model is a spectral model and this procedure requires much less memory. A preliminary experiment using a 1-hour assimilation window showed that there was no significant difference in convergence performance between minimizations in spectral space and physical space. The maximum numbers of iterations and function calls are set to 30 and 40, respectively. One function call consists of one forward integration of the assimilation model and one backward integration of the adjoint model. One iteration of the unconstrained minimization algorithm sometimes requires more than one function call due to the line search.

Scaling factors for the control variables are 10^{-5} s^{-1} for the vorticity and divergence, 1 K for the virtual temperature, 2 K for the dewpoint depression, and 10^{-3} for the logarithm of surface pressure. These values are the same as those used in the simulated data experiment except that the inverse square roots of the Gaussian weights are not multiplied. The multiplication of the inverse square roots was introduced in the previous experiment because the observations were provided on the model Gaussian grid and the cost function was defined in terms of the global integral rather than a summation over the Gaussian grid points. The number of updates m in the L-BFGS method is set to 5, as in the simulated data experiment.

6.2 Results

a. Gradient calculation

Since the assimilation model and the adjoint model in the present experiment are not consistent with each other, the adjoint model does not provide the exact gradient of the cost function but yields an approximate gradient. The precision of the approximation can be examined by using the following function similar to (69):

$$\varphi(\alpha)=\frac{J(x_0)-J(x_0-\alpha S^2\nabla_{x_0}J)}{\alpha(\nabla_{x_0}J)^T S^2\nabla_{x_0}J}, \tag{73}$$

where α is a small nondimensional parameter and S is a scaling matrix. The scaling matrix is a diagonal matrix, its diagonal elements consisting of the scaling factors introduced in the previous subsection. $\nabla_{x_0}J$ is an approximate gradient obtained from backward integration of the inconsistent adjoint model. If the approximate gradient is exact, the value of $\varphi(\alpha)$ should be close to unity for sufficiently small α.

The variations of $\varphi(\alpha)$ with $\log\alpha$ for the NOMP and MP methods are plotted in Fig. 27. The first guess used in the experiment, the initialized NCEP analysis at t = -12 h, is taken as x_0 in (73). Thus, this figure shows the precision of the gradient calculation at iteration 0 in the minimization process. The penalty parameter r is set to 8 x 10^{20} s^4. Large values of $\varphi(\alpha)$ for α = 10^{-12} result from numerical cancellation error, so they are to be neglected. It is seen that $\varphi(\alpha)$ is sufficiently close to unity when α is less than 10^{-3} for both the NOMP and MP methods and that, therefore, approximate gradients computed from the inconsistent adjoint model appear to have enough precision. Note that $\varphi(\alpha)$ is less than unity for almost all values of α for both methods, indicating that the calculated gradients tend to be slightly overestimated. Unexpectedly, $\varphi(\alpha)$ for the NOMP method is slightly closer to unity than $\varphi(\alpha)$ for the MP method. Since the directions of the gradients computed from the two methods are not the same, this does not necessarily mean that the adjoint model without the moist processes generates a more precise gradient than the adjoint model with the moist processes included. However, this result may be partly due to less validity of tangent linear approximation for the MP method. Figure 13 for the simulated data experiment showed that the range of α in which tangent linear approximation is valid for the MP method is much smaller than the range of α for the NOMP method.

b. Convergence performance

The cost function and the norm of the scaled approximate gradient are plotted in Fig. 28 as a

function of the number of iterations for the NOMP, MP and MP+RR methods. The bicubic interpolation is used to interpolate model precipitation rates from model grid points to observation points. The cost function is divided into three terms as in the previous experiment: the observation term for observational data except for precipitation, J_o, the observation term for precipitation data, J_r, and the penalty term, J_c. The second term is absent for the NOMP and MP methods. In the upper panel the upper three lines and the middle three lines are the values of J_o and J_c for the three methods, respectively, and the lower line is the values of J_r for the MP+RR method. The parenthetical values in the legend of each panel indicate the number of function calls required for 30 iterations. It is seen that the NOMP method shows slightly better convergence performance than the MP and MP+RR methods as in the simulated data experiment. If the cost function has locally large gradients resulting from highly nonlinear nature of physical processes, a better approximation to the gradient of the cost function does not always leads to better convergence performance because of less validity of the tangent linear approximation. Although the MP and MP+RR methods yield slightly worse fits to the radiosonde data than the NOMP method, the MP and MP+RR methods produce better precipitation analyses in the tropics as will be shown later. Note that assimilation of precipitation data does not improve the convergence performance, different from the results from the simulated data experiment. The values of the penalty term remain almost constant during the minimization process for the three methods. Since the first guess used is the initialized operational analysis, the gravity wave level at iteration 0 may already be at a reasonable level. Therefore, the almost constant penalty terms indicate that the resulting analyses after 30 iterations also have acceptable gravity wave levels. It is to be noted that the approximate gradients of the cost function are not reduced by an order of magnitude even after 30 iterations. Convergence performance will be improved by introducing a sophisticated preconditioning procedure instead of the simple scaling adopted here.

One of the results from the simulated data experiment was that the penalty term for suppressing gravity wave noise increased the convergence rate of 4D-Var with the moist processes included by avoiding locally large gradients in the cost function during the minimization process. The sensitivity of convergence performance to the penalty parameter r is examined in Fig. 29 for the MP+RR method. The cost function and the norm of the scaled approximate gradient are plotted as a function

of the number of iterations for the following three penalty parameters: $r = 5 \times 10^{20}$ s^4, 8×10^{20} s^4, and 15×10^{20} s^4. The curves for $r = 8 \times 10^{20}$ s^4 are identical to the curves shown in Fig. 28. Minus signs in the parentheses in the legends indicate that the number of function calls reaches the maximum value, 40, before reaching 30 iterations. It is seen that there is an optimal value of the penalty parameter, $r = 8 \times 10^{20}$ s^4, which gives the best convergence performance. This value also keeps the magnitude of the penalty term almost constant during the minimization process. When $r = 5 \times 10^{20}$ s^4 is used, the convergence is considerably decelerated and the maximum number of function calls is reached before 28 iterations. The lower panel suggests that this deceleration may be primarily caused by excessively large gradients computed by the adjoint model at the 1st and 10th iterations. This result agrees with the result from the simulated data experiment: too small penalty parameters lead to poor convergence performance of 4D-Var with the moist processes included. The penalty term gradually increases during the minimization process, indicating that gravity wave level is too high towards the end of iteration. When $r = 15 \times 10^{20}$ s^4 is used, the convergence performance is also degraded, although large gradients in the cost function are not computed. A decreasing tendency of the penalty term indicates that the gravity wave level is gradually reduced. The excessive suppression of gravity wave level makes it difficult to fit the model trajectory to the observations and, therefore, decelerates the convergence.

So far in the present experiment, bicubic interpolation has been used to interpolate model precipitation rates from model grid points to observation points. In view of the fact that model precipitation is calculated in physical space not spectral space and that convective rainfall, which predominates in the tropics, is a rather localized phenomenon, bilinear interpolation might appear more appropriate. Figure 30 compares convergence performance between bicubic and bilinear interpolations for the MP+RR method. The penalty parameter r is set to 8×10^{20} s^4. It is seen that the convergence rate is considerably decelerated when bilinear interpolation is used; the maximum number of function calls are reached before 27 iterations. Excessively large gradients are calculated at the 2nd and 6th iterations and may contribute to degradation of the convergence performance. A higher-order interpolator such as bicubic interpolation has a smoothing effect on the cost function and, as a result, improves convergence performance.

The results from the column model experiment showed that variational data assimilation with discontinuous moist processes had poor convergence properties and large analysis errors. However,

this conclusion was obtained by using the column model and simulated data. It is desirable to examine whether this conclusion holds when using the global primitive-equation model and real observations. An assimilation experiment is performed with the MP+RR method with $r = 8 \times 10^{20}$ s^4 using the discontinuous version of the moist processes and the bicubic interpolation for precipitation rates. The discontinuous version is the same as that used in the column model experiment. In this version several types of discontinuity such as discontinuous changes in cloud depth and a step-function condition for invoking cumulus convection are not removed from the original parameterization schemes of the FSU-GSM. Figure 31 compares the convergence performance between the discontinuous moist processes and the continuous moist processes that has been used so far in the present experiment. It is seen that the discontinuous moist processes deteriorate the convergence performance of 4D-Var. The discontinuous version produces a larger misfit in radiosonde data, and requires three more function calls for 30 iterations than the continuous version. The misfit in precipitation data is almost the same, but the discontinuous version yields a slightly larger misfit. This result agrees with that from the column model experiment, although differences in convergence performance between the continuous and discontinuous versions are much smaller than differences which may be expected from the result of the column model experiment. This is because the column model experiment was performed for a case of deep cumulus convection and an area covered by deep cumulus convection account for only a small portion over the globe.

c. Analyzed fields

Precipitation and divergence fields in the tropics after 30 iterations are examined for the assimilation experiments presented in Fig. 28. Impacts of the SSM/I precipitation rates on other fields such as vorticity are rather small and, therefore, are not described here. Figure 32a compares misfits in hourly mean precipitation rates for the first guess and the three assimilation methods over the tropical oceans between 30 °N and 30 °S where SSM/I data are available. The magnitude of the misfit for the NOMP method is almost the same as the misfit for the first guess, indicating that the precipitation analysis is hardly improved when using the adjoint model without the moist processes despite the fact that the full-physics model is used as the assimilation model. When using the adjoint model with the moist processes included, the misfit is reduced from the first guess by 9 % even without assimilating SSM/I precipitation rates. This result indicates a benefit of including the

moist processes in the adjoint model for the tropical data analysis. Assimilation of the SSM/I precipitation rates considerably reduces the misfit, as expected.

Figure 32b shows misfits in precipitable water over the assimilation window. The observed precipitable water is obtained from the SSM/I brightness temperatures using an algorithm developed by Lojou et al. (1994):

$$\begin{aligned} PW = {} & 20.75 - 2.582 \log (280 - T_B^{19H}) \\ & - 0.3919 \log (280 - T_B^{19V}) - 3.610 \log (280 - T_B^{22H}) \\ & + 2.729 \log (280 - T_B^{37H}) - 0.5118 \log (280 - T_B^{37V}) , \end{aligned} \tag{74}$$

where PW is the precipitable water in g cm^{-2}. The brightness temperatures are corrected to eliminate the weak biases observed between the brightness temperature measured by the SSM/I and those computed by the ECMWF global NWP model following the method proposed by Lojou et al. (1994). The SSM/I precipitable water is only calculated over the tropical oceans. Similar to the SSM/I precipitation rates used in the present experiment, the precipitable water data are prepared on a 1 ° x 1 ° grid with a temporal resolution of one hour. Bicubic interpolation is used to interpolate from model grid points to observation points. It is found that the impact of including the moist processes in the adjoint model and assimilating the SSM/I precipitation rates on the precipitable water analysis is very small; the misfit for MP+RR is decreased from the first guess only by 1.5 %. A larger impact of assimilating precipitation data on precipitable water analysis was obtained in the simulated data experiment where observations of hourly mean precipitation rates were assumed to be available over the entire tropics at every hour. This assumption was too optimistic. Figure 32b strongly suggests the necessity of assimilating the SSM/I precipitable water data retrieved from SSM/I observations to improve the tropical moisture analysis.

Figure 33 compares accumulated precipitations over the 12-hour assimilation window produced by the three methods. The corresponding accumulated precipitation obtained from the first guess is also displayed for comparison. The first guess precipitation is calculated by making a 12-hour prediction with the assimilation model from the first guess. In order to evaluate the quality of these precipitation analyses we need observations of 12-hour accumulated precipitation over the tropics, but precipitation observations with sufficient spatial coverage and precision are not available. Here

we assume that the spatial pattern of the true accumulated precipitation distribution is not very different from the spatial pattern of the composite of the SSM/I precipitation rates shown in Fig. 25. When using the SSM/I composite for verification purposes, only a qualitative comparison is possible. Since the SSM/I composite is not available over land and data-void regions over ocean, the night and day pass composites of OLR shown in Fig. 26 are used as complementary data for the qualitative verification.

The spatial distribution of the first guess precipitation (Fig. 33a) exhibits several discrepancies with the spatial distribution of the SSM/I precipitation composite. For example, the first guess generates heavy precipitation in the eastern Arabian Sea, while the SSM/I composite does not show large precipitation rates in this region. Regions of heavy precipitation over the equatorial Indian Ocean in the SSM/I composite are not seen in the first guess precipitation. In the tropical Pacific a region of precipitation observed around 10 °N, 135 °E is misplaced southward by about 10 °. A region of precipitation south of the Hawaii islands is missing in the first guess precipitation. The spatial pattern of the precipitation analysis by the NOMP method is almost the same as the spatial pattern of the first guess precipitation, although precipitation amounts are different in several regions (Fig. 33b). Precipitation around 5 °N, 135 °E, 5 °N, 170 °E, and 25 °N, 115 °W is intensified, while precipitation around the Indian subcontinent is reduced. This result suggests that the adjoint model without moist processes is not adequate to retrieve precipitation in the tropics even if the full-physics model is used as the assimilation model.

When using the adjoint model with the moist processes included, significant changes in precipitation analysis are introduced in the western tropical Pacific (Fig. 33c). The region of heavy precipitation north of New Guinea is shifted northward by about 10 °, and the center of heavy precipitation just west of the date line is shifted westward by about 5 °. A comparison with the SSM/I precipitation composite in Fig. 25 reveals that the former change makes the spatial pattern of analyzed accumulated precipitation closer to the SSM/I composite. The existence of a SSM/I data-void region around the date line makes it difficult to locate the observed center of the latter heavy precipitation region. This heavy precipitation is associated with a developing tropical depression. The day pass composite of OLR in Fig. 26 is useful for this purpose. As mentioned earlier, the day-pass composite of OLR in the western tropical Pacific is produced by using OLR observations distributed in the assimilation window. Consequently, the spatial pattern of the OLR composite in

this region is expected to resemble that of the true accumulated precipitation over the assimilation window. The center of low OLR just west of the date line in the day pass OLR composite is located at 5 °N, 165 °E (Fig. 26b). Thus, the position of the heavy precipitation produced by the MP method is closer to the observed position than that produced by the NOMP method. There are several radiosonde stations in the western tropical Pacific in contrast to the other tropical oceans (see Fig. 23). Observational data provided by these stations may contribute to the improved precipitation analysis in the western tropical Pacific. On the other hand, regions of heavy precipitation observed in the equatorial Indian Ocean and south of the Hawaii islands are not well retrieved by the MP method possibly because no radiosonde data are available in those areas. A point to note here is that if we use the adjoint model without the moist processes radiosonde data hardly contribute to improvement of precipitation analysis. This result demonstrates the necessity of including the moist processes in the adjoint model for 4D-Var in the tropics. A similar result was also obtained from the simulated data experiment.

A comparison between Figs. 25 and 33d reveals that assimilation of the SSM/I precipitation rates with 4D-Var works quite well. Regions of heavy precipitation in the equatorial Indian Ocean and south of the Hawaii islands are well retrieved. Erroneous areas of precipitation over the ocean around western India, northwest of Madagascar, and northwest of Australia are removed. The spatial distribution of the precipitation analysis by the MP+RR method exhibits a closer agreement in the western tropical Pacific with the day pass OLR composite (Fig. 26b) than that of the precipitation analysis by the MP method even in SSM/I data-void areas. Since wind speeds in the tropics are generally smaller than in the extratropics, significant horizontal advection of information from data-available areas to data-void areas is not expected when using a 12-hour assimilation window. The improvement in the precipitation analysis in SSM/I data-void areas may primarily result from assimilating other observational data by using the model dynamics. This is one of the advantages of 4D-Var over physical initialization techniques. In physical initialization some of the model state variables are directly modified so that model precipitation rates are close to observed precipitation rates, but such direct modifications are not performed in regions where precipitation observations are not available. For this reason the existence of data-void areas are undesirable for physical initialization procedures.

The time sequences of hourly mean precipitation rates at two Gaussian grid points at 7 °N,

166 °E and 4 °S, 65 °E are shown in Fig. 34. These two points are near the centers of heavy precipitation regions in the precipitation analysis by the MP+RR method (Fig. 33d). One-hour intervals in which the SSM/I precipitation rates are available are shaded. The horizontal distribution of the SSM/I precipitation rates at these 1-hour intervals is displayed in Fig. 24. At the grid point in the western tropical Pacific (Fig. 34a) the precipitation rates for the first guess and the NOMP analysis are much smaller than those for the MP and MP+RR analyses. The precipitation rate produced by the MP+RR method is slightly larger than the precipitation rate produced by the MP method. In other words, the use of the adjoint model with the moist processes included has much more impact on the precipitation analysis at this point than assimilation of the SSM/I precipitation rates. On the other hand, at the grid point in the equatorial Indian Ocean the precipitation rates for the first guess, NOMP analysis and MP analysis are much smaller than that for the MP+RR analysis (Fig. 34b). Since radiosonde data are not available around this point, assimilation of the SSM/I precipitation rates is necessary to significantly improve the precipitation analysis. Note that the precipitation analysis by the MP+RR method exhibits a decreasing tendency in the assimilation window. An inspection of Figs. 24 reveals that the SSM/I precipitation rates around this point actually show such a decreasing tendency. This time tendency in the observed precipitation rates is reflected in the analyzed precipitation rates. It is also to be noted in Fig. 34 that the impact of assimilating the SSM/I precipitation rates with 4D-Var spreads over the whole assimilation window, not confined around the observation times.

Figure 35 compares analyzed 200 hPa divergence fields at the end of the assimilation window (t = 0 h). Divergence in the tropics is closely related to convective activity. Centers of strong divergence north of New Guinea obtained from the first guess and the NOMP method are associated with the misplaced heavy precipitation in this region (see Figs. 33a and b). Strong divergence west of India obtained from the first guess is gradually replaced by convergence moving through the NOMP, MP and MP+RR analyses. In a similar fashion the upper level divergence associated with the developing tropical depression around 5 °N, 165 °E becomes stronger and more compact. A comparison between Figs. 35c and d reveals that the impact of assimilating the SSM/I precipitation rates on divergence analysis is not very remarkable. The magnitude of the impact may be highly case dependent. If there is a strong tropical disturbance in a region where conventional observations are not available, much larger impacts may be obtained as in the simulated data experiment. In the

western Atlantic, possible strong divergence associated with Hurricane Andrew is not retrieved by any of the methods. This may be due to the fact that the spatial resolution of the model is not high enough to resolve this compact tropical cyclone. Note that the divergence fields produced by the three methods are smoother than the divergence field obtained from the first guess. This may be primarily due to the inclusion of the penalty term in the cost function which suppresses large time tendencies of divergence.

d. Precipitation forecast

The quality of the analyses can also be evaluated by examining the skill of forecasts starting from the analyses. Figure 36 compares forecast skills for 24-hour precipitation totals over the tropical oceans between 30 °N and 30 °S for the first guess, NOMP, MP and MP+RR analyses. The problem here is that reliable verification data for the precipitation totals are not available over the verification area. It is assumed, as before, that the spatial pattern of the composites of the SSM/I precipitation rates for 24-hour periods are not very different from that of the true 24-hour precipitation totals. The correlation coefficient is used to score the forecasts. The SSM/I precipitation rates measured by the two DMSP satellites cover almost all the tropical oceans over 24-hour periods. Since the composites are produced from instantaneous precipitation rates, small-scale features that are unlikely in the 24-hour precipitation totals are present in the composites. In order to remove these undesirable small-scale features a moving average over a 5 ° x 5 ° domain is applied to the SSM/I precipitation composites.

It is seen from Fig. 36 that the forecast skill up to 72 hours tends to increase by including the moist processes in the adjoint model or by assimilating the SSM/I precipitation rates. However, differences in the forecast skill between the first guess and the three assimilation methods are rather small. Inclusion of the background term that possesses the observational information prior to the assimilation window might further increase the forecast skill of the three assimilation methods. An unexpected result in Fig. 36 is the better skill of the forecast starting from the NOMP analysis compared to the forecast starting from the MP analysis over the first 24-hour interval. As discussed earlier, 4D-Var with the moist processes included has a slightly slower convergence rate than 4D-Var without the moist processes. The discrepancy term for radiosonde data for the MP method is larger than that for the NOMP method (see Fig. 28a). Although the MP method produces better

precipitation and divergence analyses in the tropics, it is not clear whether the quality of the MP analysis in the tropics after 30 iterations is better as a whole than the quality of the NOMP analysis after the same number of iterations.

The small impact of assimilating the SSM/I precipitation rates on the precipitation forecast skill even over the first 24-hour interval may be an expected result in view of the small impact on the divergence analysis (see Fig. 35). Assimilation of SSM/I precipitable water data will improve the moisture analysis in the tropics and will be helpful to increase the precipitation forecast skill. Another factor to consider is the general inability of the model to well simulate synoptic-scale tropical precipitating systems. The simulated data experiment demonstrated that the T42L12 FSU-GSM is capable of simulating a super typhoon quite well at least up to 72 hours. However, such a large-scale strong tropical precipitating system is absent in the present case, so that a high forecast skill may be hard to obtain even if a good initial condition is prepared. A forecast experiment for the same case was carried out with a T213 version of the FSU-GSM by Krishnamurti et al. (1995) using their physical initialization procedure. Their results showed that the high resolution model succeeded in simulating the developing stage of the tropical depression located around 5 °N, 165 °E. In contrast, this weak tropical disturbance gradually decayed in the forecast experiment performed here with the T42 version of the FSU-GSM (not shown).

6.3 Discussion

The real data experiment was carried out to investigate the overall feasibility of 4D-Var in the tropics in a realistic situation and to evaluate the impact of assimilating the SSM/I precipitation data for the tropical analysis. The results from the experiment are summarized as follows:

- There is an optimal value of the penalty parameter for the fastest convergence of 4D-Var with moist processes included. Too small penalty parameters deteriorate the convergence performance due to strong nonlinearities in the moist processes, and too large ones deteriorate it due to excessive constraints on the model trajectory.

- When bilinear interpolation is used instead of bicubic one for interpolating precipitation rates from model grid points to observation points, the convergence performance is considerably deteriorated.

- The discontinuous moist processes result in poor convergence properties of 4D-Var, but

difference in convergence performance between the continuous and discontinuous versions are much smaller than difference which may be expected from the results from the column model experiment.

• 4D-Var using the adjoint model which lacks the moist processes produces a poor analysis in the tropics despite the fact that the full-physics model is used as the assimilation model.

• Inclusion of the moist processes in the adjoint model yields a better precipitation analysis even without assimilating the SSM/I precipitation rates, especially in areas where several radiosonde observations are available. However, the convergence rate is slightly decelerated.

• The impact of assimilating the SSM/I precipitation rates on the precipitation analysis is not confined to around SSM/I observation times, but spreads over the whole assimilation window. It improves 72-hour precipitation forecast over the tropical oceans.

• The impact of assimilating the SSM/I precipitation rates on the precipitable water analysis over the tropical oceans is positive but marginal, suggesting the necessity of assimilating SSM/I precipitable water data for a better tropical analysis.

The above results are generally in agreement with those from the column model experiment and the simulated data experiment. They strongly suggest that 4D-Var in the tropics is feasible despite the existence of strong nonlinearities in the moist processes, if zeroth-order discontinuities are removed from the parameterization schemes of the moist processes and an appropriate penalty term for suppressing gravity wave noise is included in the cost function. It should be emphasized that inclusion of the moist processes in the adjoint model is necessary for application of 4D-Var to the tropical data analysis, even if precipitation data are not assimilated. The use of the full-physics model as the assimilation model only is not enough. A problem introduced by including the moist processes in the adjoint model is the slight degradation in the convergence performance due to less validity of tangent linear approximation. Additional smoothing of the moist processes, especially parameterization schemes for deep cumulus convection, will probably improve the convergence performance.

The real data experiment shows the benefits of assimilating SSM/I precipitation data with 4D-Var for a better tropical analysis. As mentioned in the Introduction, 4D-Var has several advantages over conventional physical initialization techniques in assimilating precipitation data. One of the consequences is that, in contrast to physical initialization, 4D-Var does not yields a sharp difference

in the quality of analysis between precipitation data-void areas and the other areas, because it uses full model dynamics including horizontal advection in assimilating precipitation data and, therefore, spreads information contained in precipitation data to precipitation data-void areas. 4D-Var with the moist processes included is also capable of extracting information on precipitation from other observational data such as radiosonde data. On the other hand, assimilation of precipitation data with physical initialization is performed only in a column model framework and independently of assimilation of other observational data. The extent of improvement in the 72-hour precipitation forecast due to the assimilation of the SSM/I precipitation data with 4D-Var was not so large in the present experiment compared with some of the previous results on physical initialization. The use of a higher resolution model is expected to increase the impact of the SSM/I precipitation data on precipitation forecast through better ability of the model to simulate tropical precipitating disturbances. It should also be noted that results from observation impact study is highly case-dependent.

7. Conclusions

The feasibility of 4D-Var for the tropical atmosphere was investigated by carrying out the three assimilation experiments: the column model experiment, the simulated data experiment and the real data experiment. Emphases were placed on impacts of including the moist processes in the adjoint model on the quality of 4D-Var tropical data analysis, effects of removing zeroth-order discontinuities from the parameterization schemes of the moist processes, and effects of the penalty term for accelerating the convergence performance of 4D-Var with moist processes included.

The following is the major conclusions which are deduced from the results of the three assimilation experiments.

• 4D-Var for the tropical atmosphere is feasible if the following three procedures are adopted: the appropriate control of gravity wave level, the removal of zeroth-order discontinuities from parameterization schemes of the moist processes, and the use of a higher-order horizontal interpolation operator for precipitation when assimilating precipitation data. These procedures are helpful to improve the convergence performance of 4D-Var using the adjoint model with the moist processes included.

• 4D-Var using the adjoint model which lacks the moist processes produces a poor analysis in the tropics even if the full-physics model is used as the assimilation model.

• Inclusion of the moist processes in the adjoint model yields a better precipitation analysis even without assimilating precipitation observations, especially in areas where conventional observations are available, although the convergence rate is slightly decelerated. It also improves the analyses of divergence, moisture, and lower tropospheric vorticity in the tropics. Further improvements in the quality of analyzed fields are obtained by assimilating precipitation data.

• The impact of assimilating the SSM/I precipitation rates on the precipitation analysis is not confined to around SSM/I observation times, but spreads over the whole assimilation window. It improves precipitation forecast over the tropical oceans. Its impact on the precipitable water analysis over the tropical oceans is positive but marginal, suggesting the necessity of assimilating SSM/I precipitable water data for a better tropical data analysis.

Since the conclusions mentioned above were derived from a few case studies, assimilation experiments using real data for other cases including a strong tropical cyclone case are desirable to confirm the conclusions. Further studies are also needed on the following issues. One of the deficiencies of the present study is that the background term is not included in the cost function. Inclusion of the background term may possibly improve the impact of SSM/I precipitation rates on precipitation forecast, since much more information is available in data assimilation. The first conclusion indicates that control of gravity waves is generally helpful in accelerating the convergence of 4D-Var with moist processes included. It is worth noting that in 3D-Var the multivariate balance formulation of the background term effectively controls the amount of gravity waves in the analysis (Andersson et al., 1994). Parrish and Derber (1992) showed that a separate initialization step is almost unnecessary in 3D-Var. In light of their results, it may be expected that inclusion of a background term in the cost function also improves the convergence performance of 4D-Var with moist processes included. Interactions of precipitation data with prior information contained in the background term are also remained to be explored. Recently, Fillion and Errico (1997) investigated these interactions in a column model framework. The second issue is to introduce a sophisticated preconditioning technique to attain faster convergence rates. The simple scaling adopted here does not work satisfactorily; for instance, the norm of the gradient of the cost function in Fig. 15 does not show a significant decrease with iterations. Third, since the largest part of forecast error in the tropics is due to the model systematic error, the introduction of a systematic error correction may be more important for 4D-Var in the tropics than it is in the extratropics. A method for correcting the

model bias was proposed by Derber (1989). The sparsity of observations in the tropics, however, complicates efforts to obtain reliable estimates of the model bias. The degrees of freedom in the estimated systematic error will have to be decreased to a number much smaller than the degrees of freedom in the model.

The results of this study suggest that 4D-Var can provide a reliable estimate of precipitation over the global tropics. No single technique exists for reliably estimating global precipitation. Rain gauges provide relatively accurate measurements at a specific point, but area averages derived from gauge observations suffer from severe limitations due to sampling problems. Rain radars observe a relatively large area of 200 - 300 km in radius, but the spatial distribution of radar data is generally limited to populated land areas. Satellite IR estimates from geosynchronous platforms have a poorer physical basis for estimating precipitation. Satellite microwave estimates have a better physical basis, but they are only available a few times a day from polar-orbiting platforms. Short-range prediction by NWP models can provide precipitation estimates with complete global coverage by using physical laws and observed initial conditions, but the quality of model precipitation is limited by errors in models and initial conditions. Objective techniques that combine the useful portions of available estimates are desirable to improve estimates of global precipitation. Huffman et al. (1995) proposed a technique to combine precipitation estimates from satellite microwave data, satellite IR data, rain gauge data, and short-range prediction from a NWP model. 4D-Var utilizes the model dynamics and observations in the assimilation window in an optimal way to produce global precipitation estimates. Although precipitation estimates derived from 4D-Var are still not free from model error, the present study demonstrates the ability of 4D-Var to provide a reliable estimate of accumulated precipitation over the tropics.

In the Tropical Rainfall Measuring Mission (TRMM) project (Theon and Fugono, 1988), a rain radar aboard a satellite in a non-sunsynchronous orbit are providing more accurate precipitation estimates over the global tropics than microwave estimates. 4D-Var can easily assimilate TRMM precipitation data along with other available observations including other satellite precipitation and precipitable water estimates. It can also produce dynamically-consistent estimates of related fields such as condensational heating rates by using model dynamics. Thus, application of 4D-Var to assimilate TRMM data may be the most promising approach to accomplishing one of the goals of the TRMM project, that is, a reliable estimate of precipitation and latent heating over the global

tropics.

Acknowledgments

This research was done during the author's visit to the Florida State University. The author is sincerely thankful to Prof. T. N. Krishnamurti of the Florida State University for providing the opportunity for the visit and for his encouragement through the work. It is a pleasure to thank Profs. I. M. Navon and X. Zou of the Florida State University, Dr. A. Kasahara of the National Center for Atmospheric Research, Drs. D. Zupanski, and M. Zupanski of the National Centers of Environmental Prediction, and Dr. A. da Silva of the Goddard Space Flight Center for helpful discussions and comments. Special thanks go to Profs. R. Kimura and H. Niino of the University of Tokyo for encouraging the author to submit the work as the doctoral dissertation. Part of this doctoral dissertation was published as Tsuyuki (1996a), Tsuyuki (1996b) and Tsuyuki (1997). This research was supported by NSF Grant ATM 9312537, NASA Grants NAG 8-914, and NAG 5-1595. Most of the computations for this work were carried out on the CRAY-YMP at the Florida State University and the CRAY-C90 at the San Diego Supercomputer Center.

References

Adler, R. F., and A. J. Negri, 1988: A satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteor.*, **27**, 30-51.

Andersson, E., J. Pailleux, J.-N. Thépaut, J. R. Eyre, A. P. McNally, G. A. Kelly, and P. Courtier, 1994: Use of cloud-clear radiances in three/four-dimensional variational data assimilation. *Quart. J. Roy. Meteor. Soc.*, **120**, 627-653.

Aonashi, K., 1993: An initialization method to incorporate precipitation data into a mesoscale numerical weather prediction model. *J. Meteor. Soc. Japan*, **71**, 393-406.

Aonashi, K., K. Kuma, and Y. Matsushita, 1997: A physical initialization method for the economical prognostic Arakawa-Schubert scheme. *J. Meteor. Soc. Japan*, **75**, 597-618.

Asselin, R., 1972: Frequency filter for time integrations. *Mon. Wea. Rev.*, **100**, 487-490.

Businger, J. A., J. C. Wyngard, Y. Izumi, and E. F. Bradley, 1971: Flux profile relationship in the atmospheric surface layer. *J. Atmos. Sci.*, **28**, 181-189.

Chang, C. B., 1979: *On the Influences of Solar Radiation and Diurnal Variation of Surface Temperatures on African Disturbances*. FSU Report No. 79-3, Department of Meteorology, Florida State University, Tallahassee, FL 32306, 157pp.

Chang, S. W., and T. R. Holt, 1994: Impact of SSM/I rainfall rate on numerical prediction of cyclones. *Mon. Wea. Rev.*, **122**, 151-164.

Charney, J., M. Halem, and R. Jastrow, 1969: Use of incomplete historical data to infer the present state of the atmosphere. *J. Atmos. Sci.*, **26**, 1160-1163.

Courtier, P., and O. Talagrand, 1990: Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, **42A**, 531-549.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367-1387.

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

Daley, R., C. Girard, J. Henderson, and I. Simmonds, 1976: Short-term forecasting with a multi-level spectral primitive model. Part I-Model formulation. *Atmosphere*, **14**, 98-116.

Davidson, N. E., and K. Puri, 1992: Tropical prediction using dynamical nudging, satellite-defined convective heat sources, and a cyclone bogus. *Mon. Wea. Rev.*, **120**, 2501-2522.

Derber, J. C., 1989: A variational continuous assimilation technique. *Mon. Wea. Rev.*, **117**, 2437-2446.

Dey, C. H., and L. L. Morone, 1985: Evolution of the National Meteorological Center global data assimilation system: January 1982 - December 1983. *Mon. Wea. Rev.*, **113**, 304-318.

Fillion, L., and R. Errico, 1997: Variational assimilation of precipitation data using moist convective parameterization schemes: A 1D-Var study. *Mon. Wea. Rev.*, **125**, 2917-2942.

Frank, W. M., and J. Molinari, 1993: Convective adjustment. In *The Representation of Cumulus Convection in Numerical Models* (eds. Emanuel, K. A., and D. J. Raymond). American Meteorological Society, 101-105.

Gairola, R. K., and T. N. Krishnamurti, 1992: Rain rates based on SSM/I, OLR and raingauge data sets. *Meteorol. Atmos. Phys.*, **50**, 165-174.

Gill, P. E., W. Murray, and M. H. Wright, 1981: *Practical Optimization*, Academic Press, London, 401pp.

Gustafsson, N., 1992: Use of a digital filter as weak constraint in variational data assimilation. *Proc. of ECMWF Workshop on Variational Assimilation with Special Emphasis on Three-dimensional Aspects*, Reading, U.K., ECMWF, 327-338.

Heckley, W. A., A. G. Kelly, and M. Tiedtke, 1990: On the use of satellite-derived heating rates for data assimilation within the tropics. *Mon. Wea. Rev.*, **118**, 1743-1757.

Hollinger, J., R. Lo, G. Poe, R. Savage and J. Pierce, 1987: *Special Sensor Microwave/Imager User's Guide*. Naval Research Laboratory, 120pp.

Holton, J. R., 1992: *An Introduction to Dynamic Meteorology, Third Edition*. Academic Press, San Diego, 511pp.

Huffman, G. J., R. F. Adler, B. Rudolf, and U. Schneider, 1995: Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. *J. Climate*, **8**, 1284-1295.

Kanamitsu, M., 1985: A study of the predictability of the ECMWF operational forecast model in the tropics. *J. Meteor. Soc. Japan*, **63**, 779-804.

Kasahara, A., A. P. Mizzi, and L. J. Donner, 1994: Diabatic initialization for improvement in the tropical analysis of divergence and moisture using satellite radiometric imagery data. *Tellus*, **46A**, 242-264.

Kasahara, A., and A. P. Mizzi, 1996: Use of precipitation data for diabatic initialization to

improve the tropical analysis of divergence and moisture. *Meteor. Atmos. Phys.*, **60**, 143-156.

Kasahara, A., J. Tsutsui, and H. Hirakuchi, 1996: Inversion methods of three cumulus parameterizations for diabatic initialization of a tropical cyclone model. *Mon. Wea. Rev.*, **124**, 2304-2321.

Kitade, T., 1983: Nonlinear normal mode initialization with physics. *Mon. Wea. Rev.*, **111**, 2194-2213.

Krishnamurti, T. N, K. Ingles, S. Cocke, R. Pasch, and T. Kitade, 1984: Details of low latitude medium-range numerical weather prediction using a global spectral model II. Effect of orography and physical initialization. *J. Meteor. Soc. Japan*, **62**, 613-649.

Krishnamurti, T. N., H. S. Bedi, and K. Ingles, 1993: Physical Initialization using SSM/I rain rates. *Tellus*, **45A**, 247-269.

Krishnamurti, T. N., H. S. Bedi, W. Heckley, and K. Ingles, 1988: Reduction of the spinup time for evaporation and precipitation in a spectral model. *Mon. Wea. Rev.*, **116**, 907-920.

Krishnamurti, T. N., J. Xue, H. S. Bedi, K. Ingles, and D. Oosterhof, 1991: Physical initialization for numerical weather prediction over the tropics. *Tellus*, **43AB**, 53-81.

Krishnamurti, T. N., S. K. R. Bhowmik, D. Oosterhof, G. Rohaly, and N. Surgi, 1995: Mesoscale signatures within the tropics generated by physical initialization. *Mon. Wea. Rev.*, **123**, 2771-2790.

Krishnamurti, T. N., S. Low-Nam, and R. Pasch, 1983: Cumulus parameterization and rainfall rates II. *Mon. Wea. Rev.*, **111**, 815-828.

Kummerow, C., and L. Giglio, 1994a: A passive technique for estimating rainfall and vertical structure information from space. Part I: Algorithm description. *J. Appl. Meteor.*, **33**, 3-18.

Kummerow, C., and L. Giglio, 1994b: A passive technique for estimating rainfall and vertical structure information from space. Part II: Applications to SSM/I data. *J. Appl. Meteor.*, **33**, 19-34.

Kuo, H. L., 1974: Further studies of the parameterization of the influence of cumulus convection on large-scale flow. *J. Atmos. Sci.*, **31**, 1232-1240.

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, **38A**, 97-110.

Lemarechal, C., and R. Mifflin, 1978: *Nonsmooth Optimization* , IIASA Proceedings 3, Pergamon,

Oxford, 186pp.

Liu, D. C., and J. Nocedal, 1989: On the limited memory BFGS method for large scale optimization, *Mathematical Programming*, **45**, 503-528.

Liu, G., and J. A. Curry, 1992: Retrieval of precipitation from satellite microwave measurement using both emission and scattering. *J. Geophys. Res.*, **97**, 9959-9974.

Lojou, J.-Y., R. Bernard, and L. Eymard, 1994: A simple method for testing brightness temperatures from satellite microwave radiometers. *J. Atmos. Oceanic Technol.*, **11**, 387-400.

Louis, J. F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteor.*, **17**, 187-202.

Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019-1034.

Manobianco, J., S. Koch, V. M. Karyampudi, and A. J. Negri, 1994: The impact of assimilating satellite-derived precipitation rates on numerical simulations of the ERICA IOP 4 cyclone. *Mon. Wea. Rev.*, **122**, 341-365.

Navon, I. M., X. Zou, J. C. Derber, and J. Sela, 1992: Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon. Wea. Rev.*, **120**, 1433-1446.

Nocedal, J., 1980: Updating quasi-Newton matrices with limited storage. *Math. Comp.*, **35**, 773-782.

Olson, W. S., F. J. LaFontaine, W. L. Smith, and T. H. Achtor, 1990: *Recommended algorithms for the retrieval of rainfall rates in the tropics using the SSM/I (DMSP-8).* CIMSS/Space Science and Engineering Center, Univ. of Wisconsin, Madison, WI 53706, 10pp.

Orszag, L, 1970: Transform method for calculation of vector coupled sums: application to the spectral form of the vorticity equation. *J. Atmos. Sci.*, **27**, 890-895.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Pasch, P. J., 1983: *On the Onset of the Planetary Scale Monsoon.* FSU Report No. 83-9, Department of Meteorology, Florida State University, Tallahassee, FL 32306, 220pp.

Peng, M. S., and S. W. Chang, 1996: Impacts of SSM/I retrieved rainfall rates on numerical prediction of a tropical cyclone. *Mon. Wea. Rev.*, **124**, 1181-1198.

Peng, M. S., and S. W. Chang, 1997: Prediction of typhoon Flo of 1990 with assimilation of

rainfall rates inferred from infrared brightness temperatures. *J. Meteor. Soc. Japan*, **75**, 1073-1089.

Puri, K., and M. J. Miller, 1990: The use of satellite data in the specification of convective heating for diabatic initialization and moisture adjustment in numerical weather prediction. *Mon. Wea. Rev.*, **118**, 67-93.

Sasaki, Y., 1970: Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.*, **98**, 875-883.

Schlick, T., and A. Fogelson, 1992: TNPACK --- a truncated Newton minimization package for large-scale problems: I. Algorithm and usage. *ACM Trans. Math. Software*, **18**, 46-70.

Schlick, T., and A. Fogelson, 1992: TNPACK --- a truncated Newton minimization package for large-scale problems: II. Implementation examples. *ACM Trans. Math. Software*, **18**, 71-111.

Shor, N. Z., 1985: *Minimization Methods for Non-differentiable Functions*. Springer Series in Computational Mathematics 3, Springer-Verlag, Berlin, 162pp.

Smagorinsky, J., S. Manabe, and J. L. Holloway, 1965: Numerical results from a nine-level general circulation model of the atmosphere. *Mon. Wea. Rev.*, **93**, 727-768.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311-1328.

Thacker, W. C., 1989: The role of the Hessian matrix in fitting models to measurements. *J. Geophy. Res.*, **94**, 6177-6196.

Theon, J., and N. Fugono, 1988: *Tropical Rainfall Measurements*. A. Deepak Publishers, Hampton, Virginia, 523pp.

Thépaut, J.-N., and P. Courtier, 1991: Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model. *Quart. J. Roy. Meteor. Soc.*, **117**, 1225-1254.

Thépaut, J.-N., D. Vasiljevic, P. Courtier, and J. Pailleux, 1993a: Variational assimilation of conventional meteorological observations with a multilevel primitive-equation model. *Quart. J. Roy. Meteor. Soc.*, **119**, 153-186.

Thépaut, J.-N., R. N. Hoffman, and P. Courtier, 1993b: Interactions of dynamics and observations in a four-dimensional variational assimilation. *Mon. Wea. Rev.*, **121**, 3393-3414.

Tsuyuki, T., 1996a: Variational data assimilation in the tropics using precipitation data. Part I: Column model. Meteor. Atmos. Phys., 60, 87-104.

Tsuyuki, T., 1996b: Variational data assimilation in the tropics using precipitation data. Part II: 3D model. Mon. Wea. Rev., 124, 2545-2561.

Tsuyuki, T., 1997: Variational data assimilation in the tropics using precipitation data. Part III: Assimilation of SSM/I precipitation rates. Mon. Wea. Rev., 125, 1447-1464.

Verlinde, J., and W. R. Cotton, 1993: Fitting microphysical observations of nonsteady convective clouds to a numerical model: an application of adjoint technique of data assimilation to a kinematic model. *Mon. Wea. Rev.*, **121**, 2776-2793.

Vukicevic, T., and R. M. Errico, 1993: Linearization and adjoint of parameterized moist diabatic processes. *Tellus*, **45A**, 493-510.

Wang, Z., 1993: *Variational data assimilation with 2-D shallow water equations and 3-D FSU global spectral models*. Ph. D. dissertation, Department of Mathematics, Florida State University, Tallahassee, FL 32306, 232pp.

Wergen, W., 1992: The effect of model errors in variational assimilation. *Tellus*, **44A**, 297-313.

Xu, K.-M., A. Arakawa, and K. Krueger, 1992: The macroscopic behavior of cumulus ensembles simulated by a cumulus ensemble model. *J. Atmos. Sci.*, **49**, 2402-2426.

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

Xu, Q., 1996b: Generalized adjoint for physical processes with parameterized discontinuities. Part II: Vector formulations and matching conditions. *J. Atmos. Sci.*, **53**, 1143-1155.

Yamasaki, M., 1983: A further study of the tropical cyclone without parameterizing the effects of cumulus convection. *Papers Meteor. Geophys.*, **34**, 221-260.

Yang, W., I. M. Navon, and P. Courtier, 1996: A new Hessian preconditioning method applied to variational data assimilation experiments using NASA general circulation models. *Mon. Wea. Rev.*, **124**, 1000-1017.

Zou, X., 1997: Tangent linear and adjoint of "on-off" processes and their feasibility for use in 4-dimensional variational data assimilation. *Tellus*, **49A**, 3-31.

Zou, X., and Y.-H. Kuo, 1996: Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.*, **124**, 2859-2882.

Zou, X., I. M. Navon, and F. X. Le Dimet, 1992: Incomplete observations and control of gravity

waves in variational data assimilation. *Tellus*, **44A**, 273-296.

Zou, X., I. M. Navon, and J. Sela, 1993a: Control of gravitational oscillations in variational data assimilation. *Mon. Wea. Rev.*, **121**, 272-289.

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

Zou, X., I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F.-X. Le Dimet, 1993c: Numerical experience with limited-memory quasi-Newton and truncated Newton methods. *SIAM J. Optimization*, **3**, 582-608.

Zupanski, D., 1993: The effects of discontinuities in the Betts-Miller cumulus convection scheme on four-dimensional variational data assimilation. *Tellus*, **45A**, 511-524.

Zupanski, D., 1997: A general weak constraint applicable to operational 4DVAR data assimilation systems. *Mon. Wea. Rev.*, **125**, 2274-2292.

Zupanski, D., and F. Mesinger, 1995: Four-dimensional variational assimilation of precipitation data. *Mon. Wea. Rev.*, **123**, 1112-1127.

Zupanski, M., 1993a: Regional four-dimensional variational data assimilation in a quasi-operational forecasting environment. *Mon. Wea. Rev.*, **121**, 2396-2408.

Zupanski, M., 1993b: A preconditioning algorithm for large-scale minimization problems. *Tellus*, **45A**, 478-492.

Zupanski, M., 1996: A preconditioning algorithm for four-dimensional variational data assimilation. *Mon. Wea. Rev.*, **124**, 2562-2573.

Tables

Table 1. RMSE in surface pressure, divergence at 200 hPa, 6-hour accumulated precipitation, and total precipitable water in the tropics (30 °N - 30 °S) at $t = 0$ h after 30 iterations.

(a) $r = 4 \times 10^{17}$ s^4

	Surface pressure (hPa)	Divergence at 200 hPa (10^{-6} s^{-1})	Precipitation (mm)	Precipitable water (mm)
first guess	1.20	7.00	1.78	3.17
NOMP	0.80	5.00	2.03	3.27
MP	1.18	5.15	1.48	3.08
MP+RR	1.22	4.20	0.82	2.87

(b) $r = 1 \times 10^{18}$ s^4

	Surface pressure (hPa)	Divergence at 200 hPa (10^{-6} s^{-1})	Precipitation (mm)	Precipitable water (mm)
first guess	1.20	7.00	1.78	3.17
NOMP	0.66	4.91	1.89	3.24
MP	0.92	4.67	1.36	3.08
MP+RR	0.56	4.31	0.89	2.90

Table 2. Observation error standard deviations for assimilated radiosonde data.

Level (hPa)	Geopotential height (gpm)	Wind ($m\ s^{-1}$)	Temperature (K)	Relative humidity (%)
100	-	2.5	3.1	-
150	-	3.0	2.8	-
200	-	3.3	2.5	-
250	-	3.4	2.2	-
300	-	3.4	2.0	20
400	-	3.1	1.6	20
500	-	2.8	1.3	20
700	-	2.4	1.3	20
850	-	2.0	1.5	20
1000	7.0	1.4	1.8	20

Table 3. Numbers of reported radiosonde data with the numbers of rejected data in parentheses.

Time	Geopotential height at 1000 hPa	Zonal wind	Meridional wind	Temperature	Relative humidity
0 - 3 UTC	403 (60)	3546 (214)	3546 (157)	3679 (166)	2097 (39)
3 - 9 UTC	28 (5)	225 (21)	225 (19)	228 (12)	133 (2)
9 - 12 UTC	604 (131)	5429 (342)	5429 (293)	5648 (224)	3200 (54)
Total	1035 (196)	9200 (577)	9200 (469)	9555 (402)	5430 (95)
Rejection rate	18.9 %	6.3 %	5.1 %	4.2 %	1.7 %

Figures

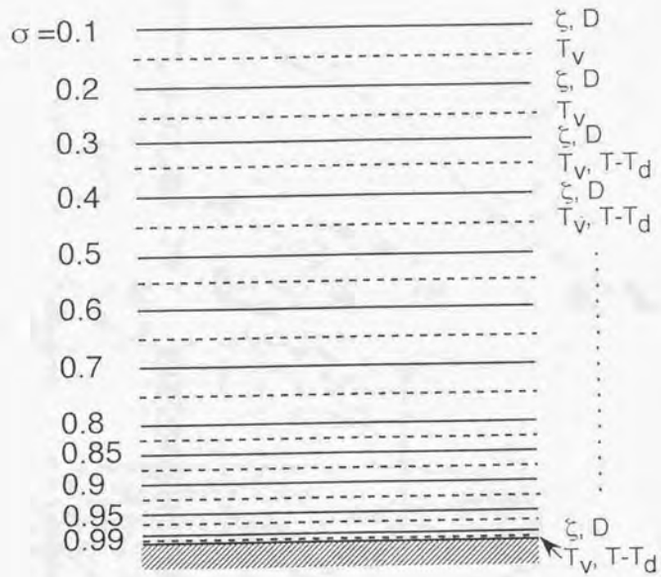

Fig. 1. Configuration of vertical levels where model state variables are defined in the T42L12 FSU-GSM.

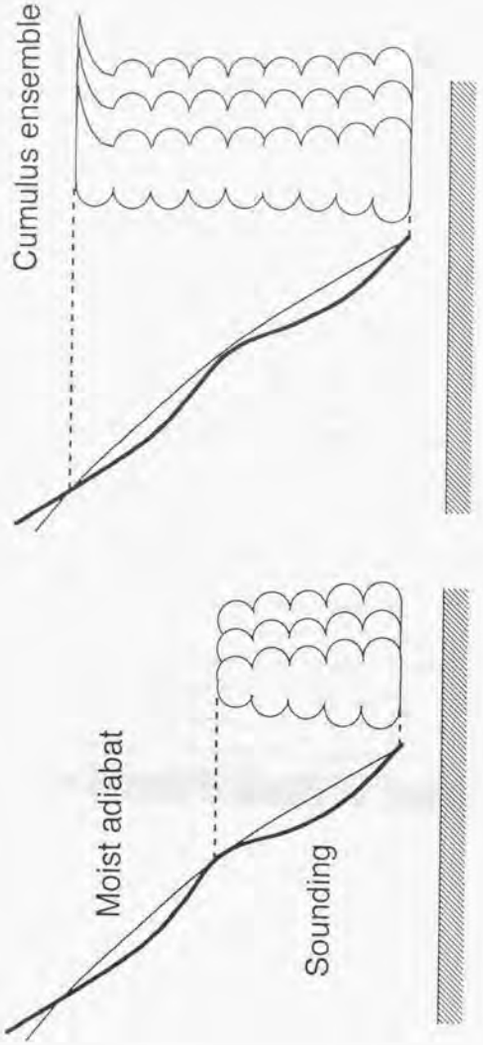

Fig. 2. A discontinuity of the Kuo scheme associated with an inversion layer.

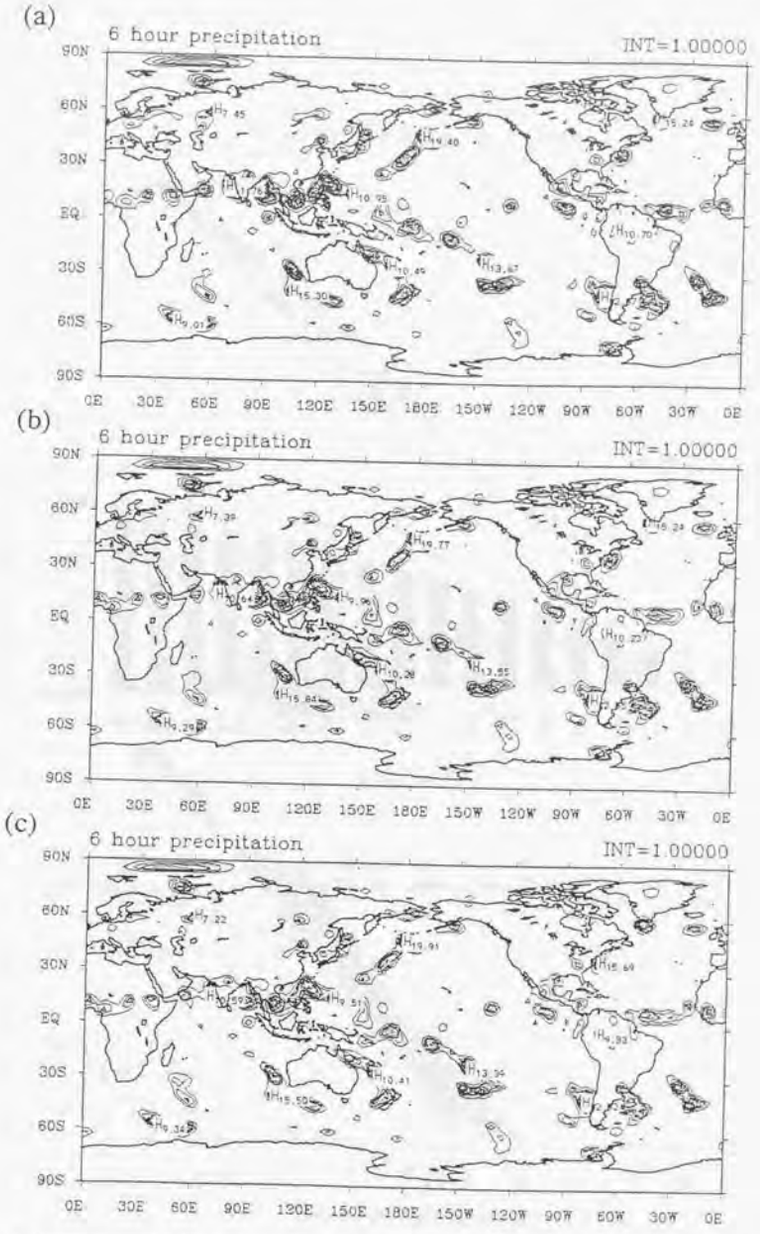

Fig. 3. Comparison of 6-hour accumulated precipitations predicted by the T42L12 FSU-GSM using (a) the original version, (b) the continuous version, and (c) the discontinuous version of moist processes. The contour interval is 1 mm.

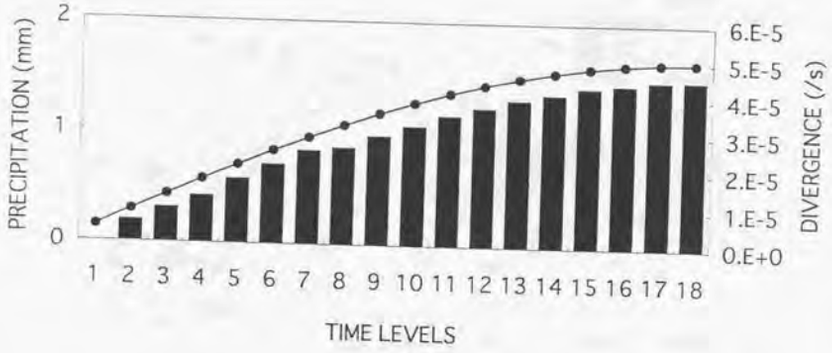

Fig. 4. Time sequences of one-time step precipitation (bar) and the prescribed divergence at the model top (line). The precipitation is generated by the column model with the continuous version of moist processes.

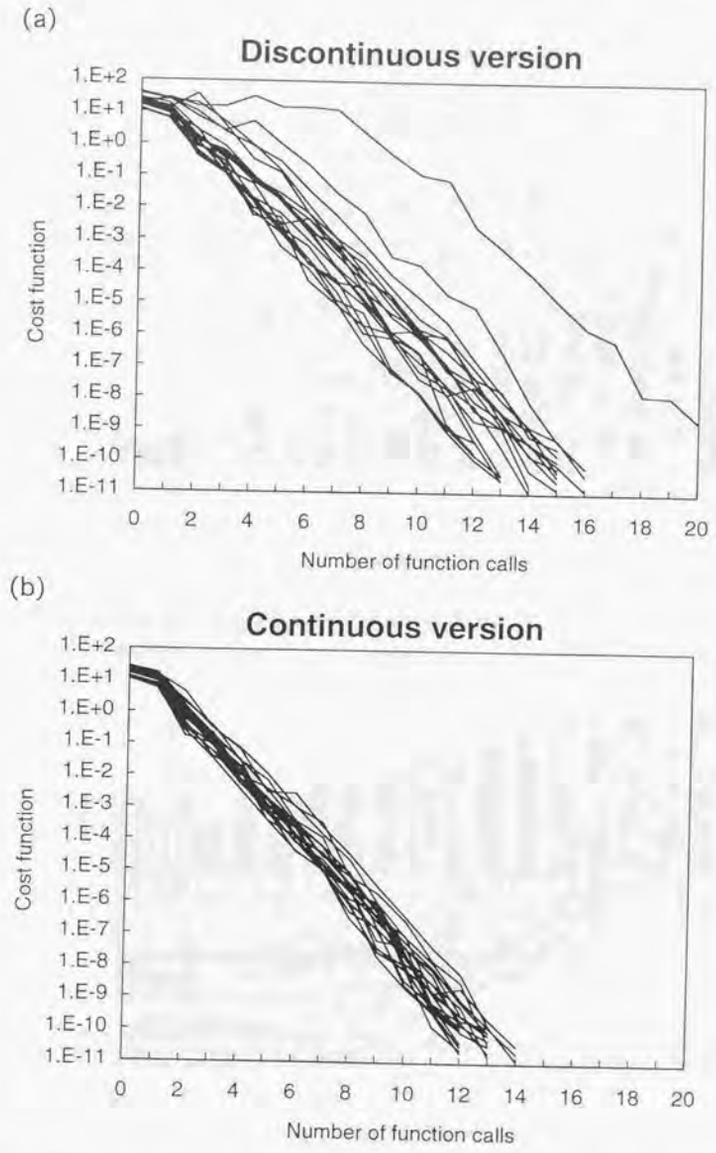

Fig. 5. Variations of the logarithm of the cost function with the number of function calls for Experiment A using (a) the discontinuous version and (b) the continuous version of moist processes.

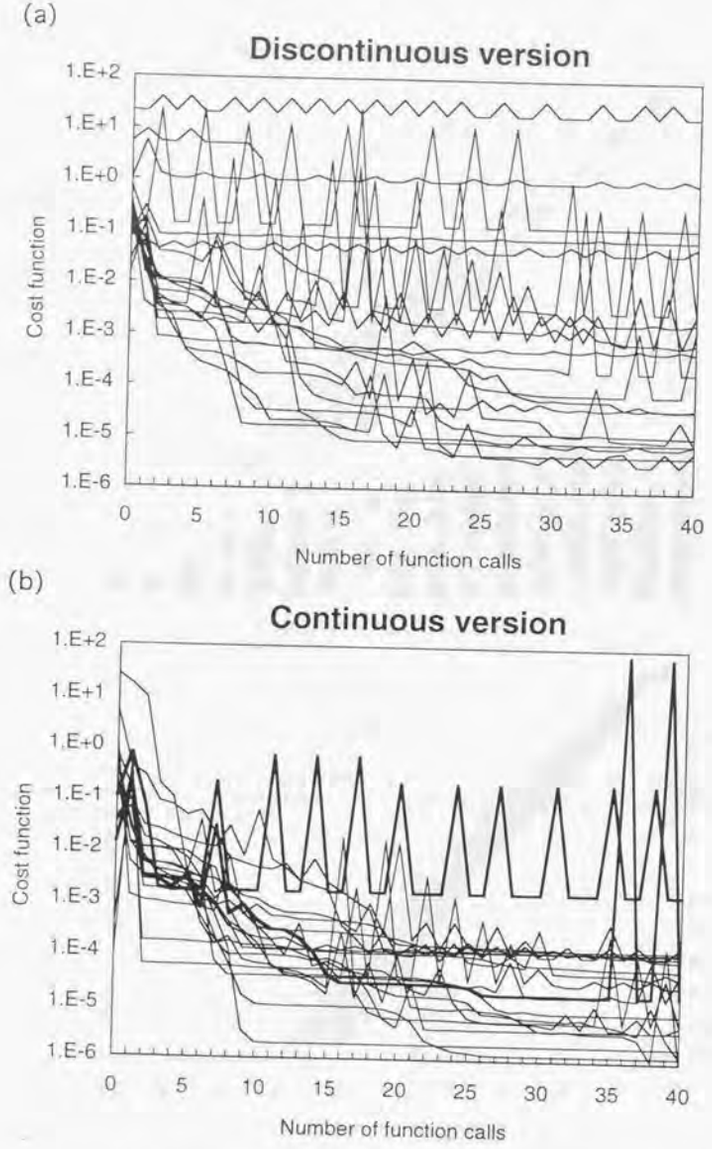

Fig. 6. Same as Fig. 5 except for Experiment B. Thick lines in (b) are for Case 3 and Case 20.

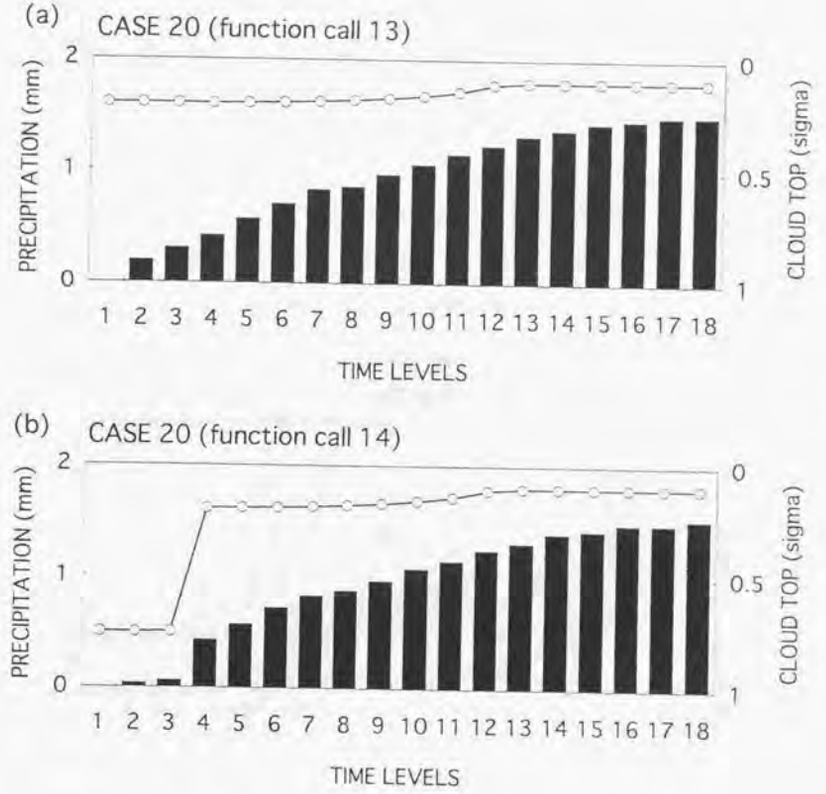

Fig. 7. Time sequences of one-time step precipitation (bar) and cloud top height (line) for Case 20 at (a) the 13th and (b) the 14th function calls.

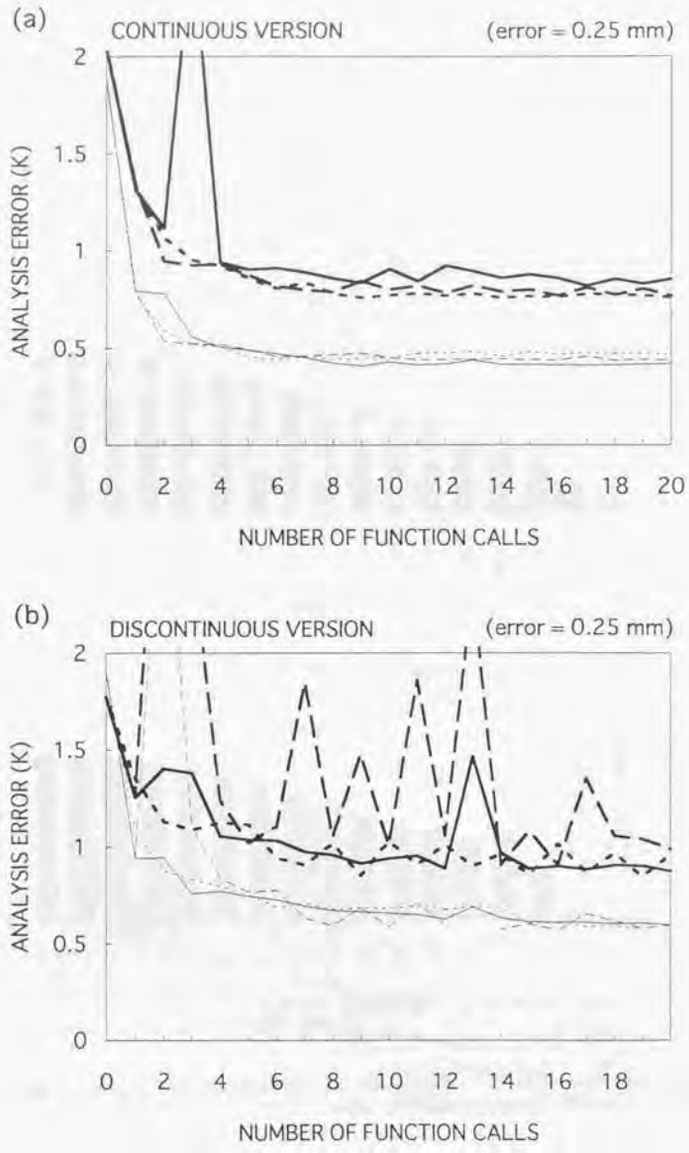

Fig. 8. Variations of mean analysis errors in virtual temperature (thin line) and dewpoint depression (thick line) with the number of function calls for (a) the continuous version and (b) the discontinuous version of moist processes. Dotted line is for Experiment C, dashed line for Experiment D, and solid line for Experiment E. Observation error of one-time step precipitation is 0.25 mm.

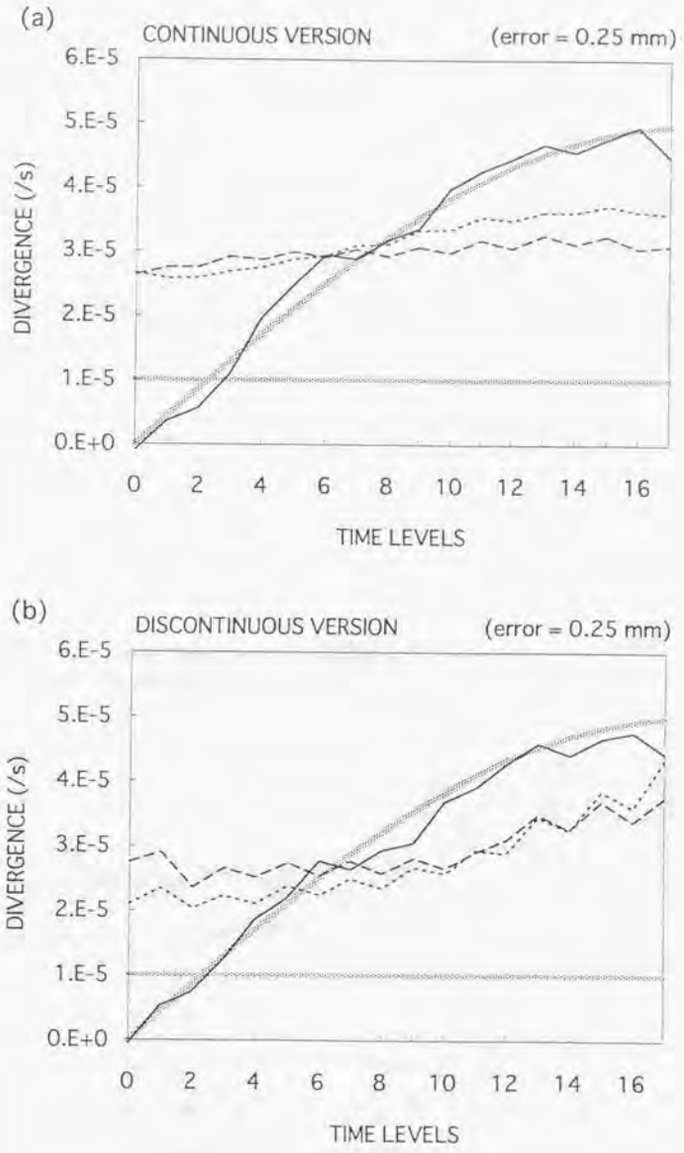

Fig. 9. Mean divergence at the model top at every time level at the 20th function call for (a) the continuous version and (b) the discontinuous version of moist processes. Horizontal grey line is for the first guess, other grey line for the truth, dotted line for Experiment C, dashed line for Experiment D, and solid line for Experiment E. Observation error of one-time step precipitation is 0.25 mm.

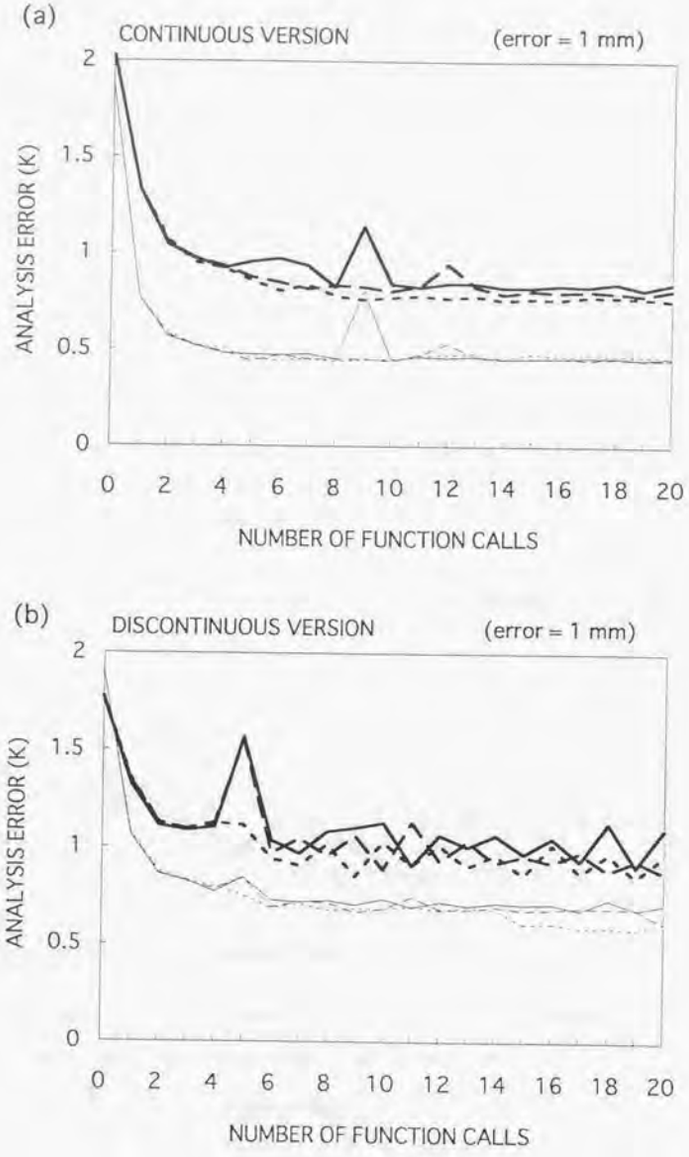

Fig. 10. Same as Fig. 8 except that observational error of one-time step precipitation is 1 mm.

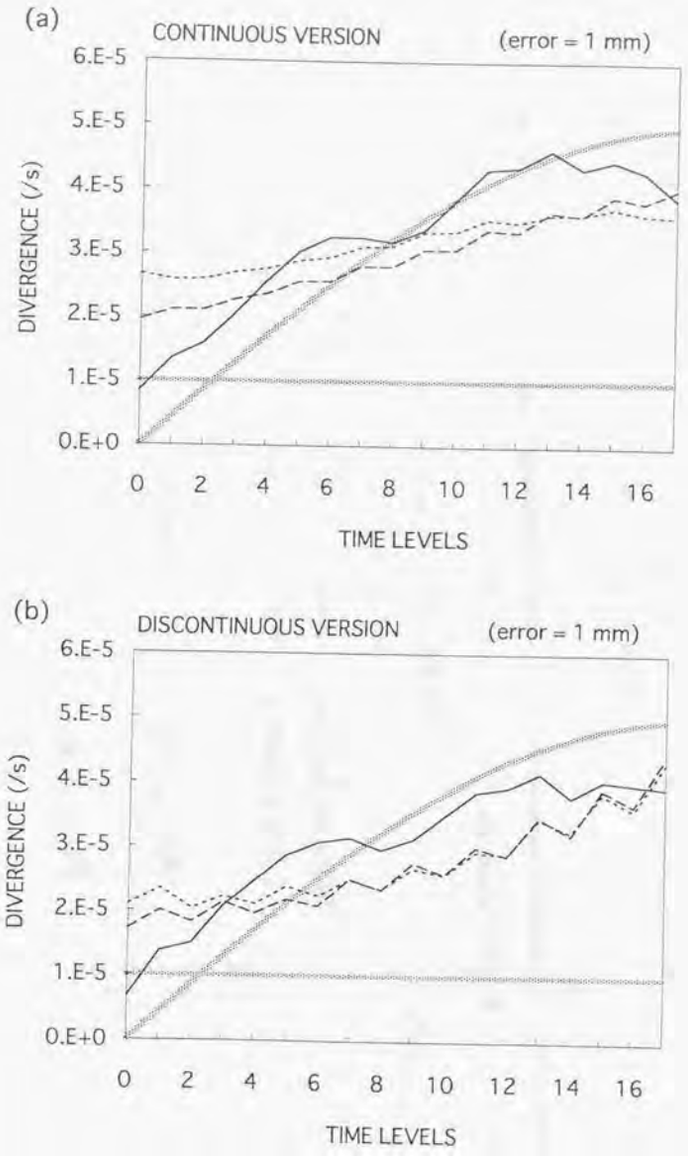

Fig. 11. Same as Fig. 9 except that observational error of one-time step precipitation is 1 mm.

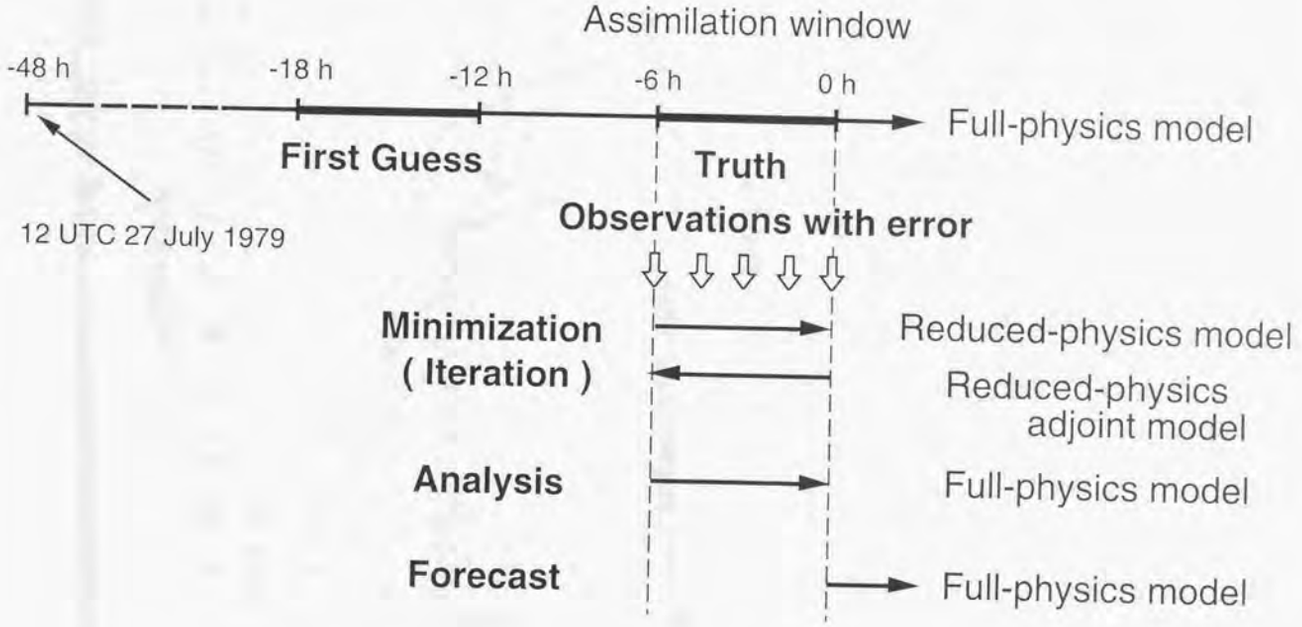

Fig. 12. Schematic illustration of experimental design for the simulated data experiment.

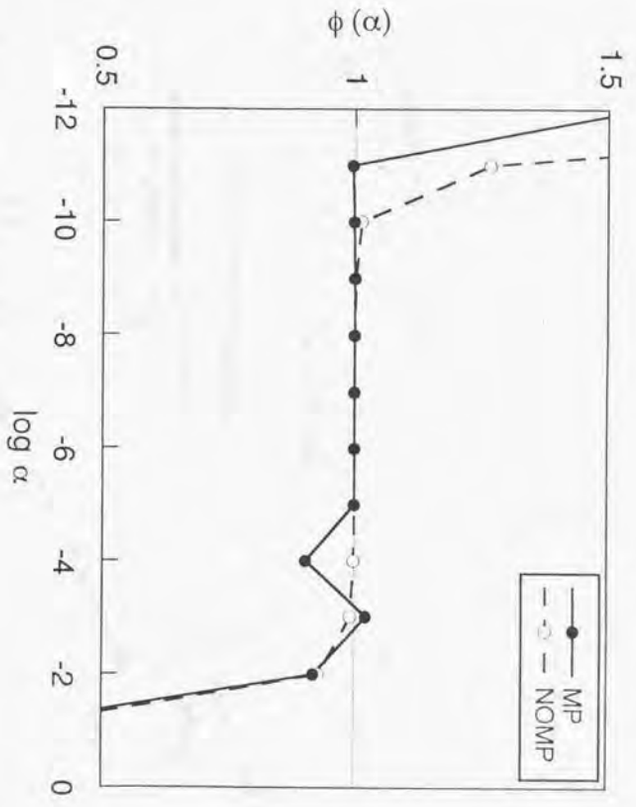

Fig. 13. Variations of the function $\phi(\alpha)$ with $\log \alpha$ for the NOMP and MP methods. The atmospheric state at t = -48 h is taken as x_0 in (69). The observation errors and the penalty term are not included.

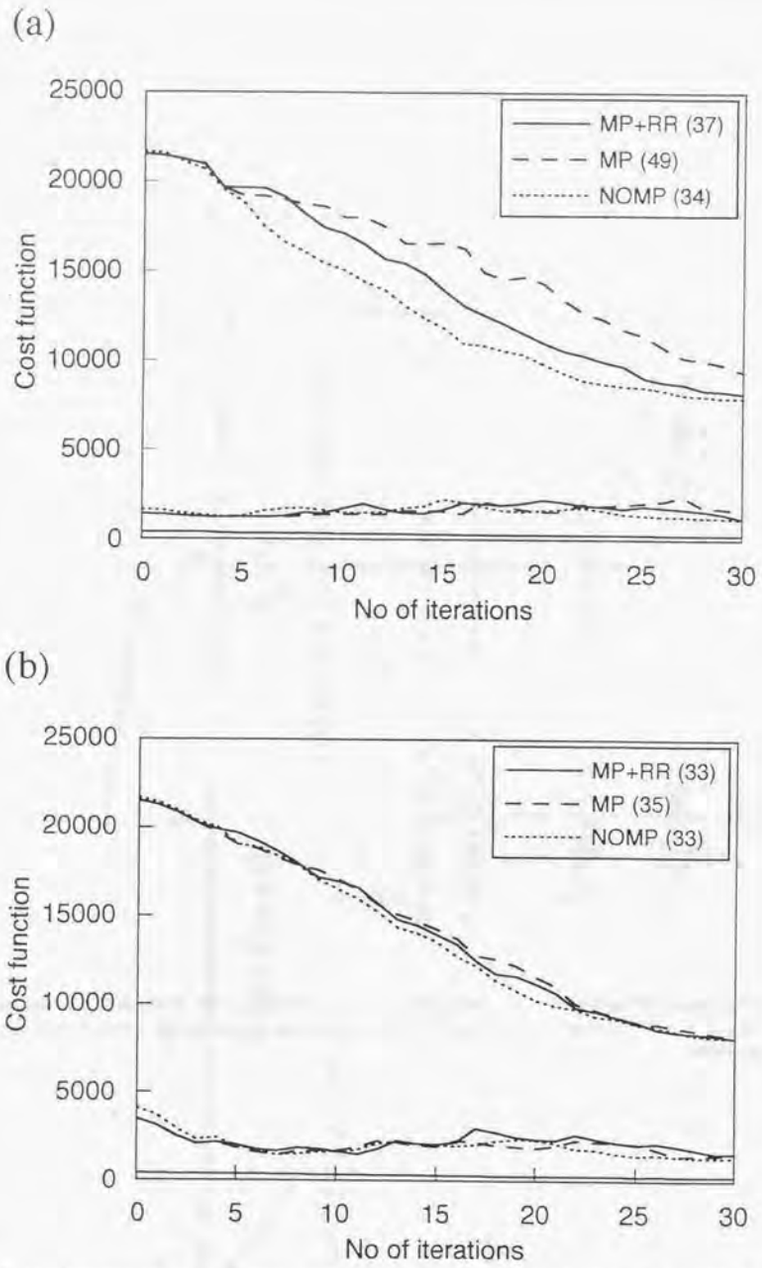

Fig. 14. Variation of the cost function with the number of iterations for (a) $r = 4 \times 10^{17}$ s^4 and (b) $r = 1 \times 10^{18}$ s^4. In each panel the upper three lines are J_o, the middle three lines are J_c, and the lower line is J_r for the MP+RR method. Parenthetical values in the legend indicate the number of function calls required for 30 iterations.

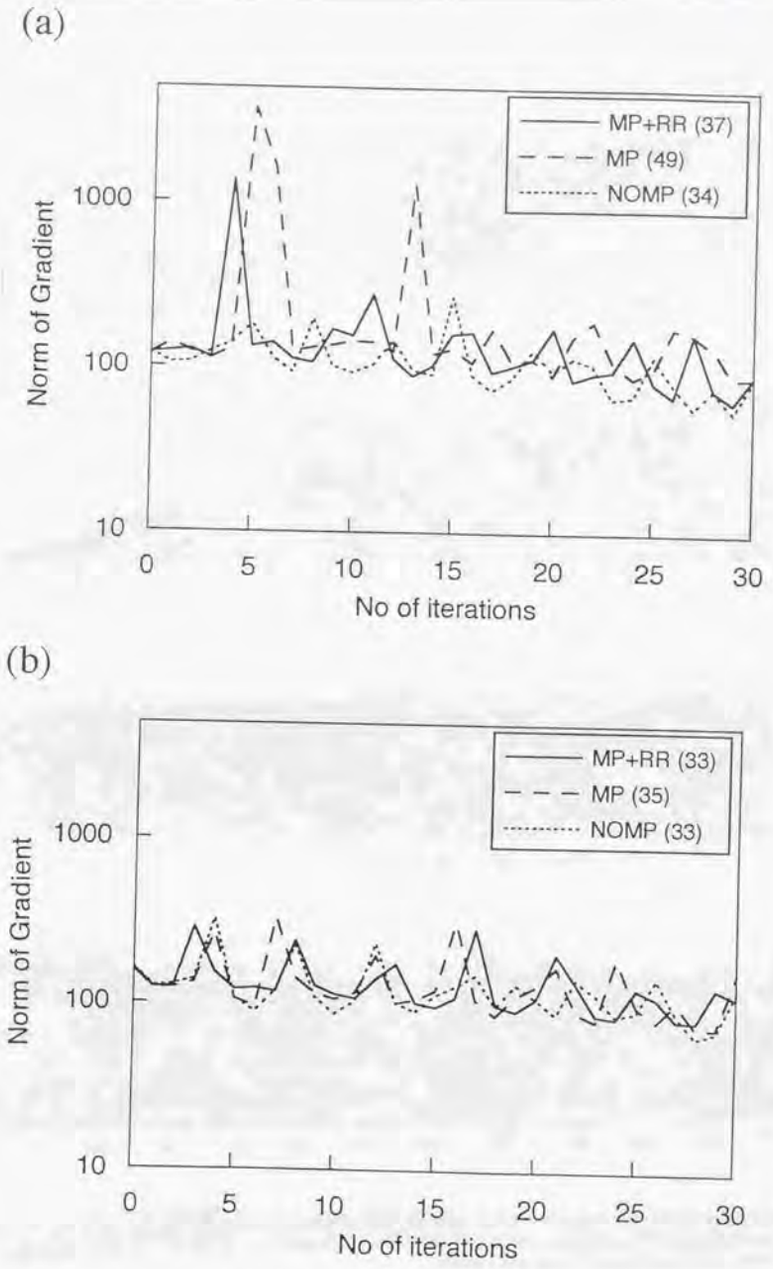

Fig. 15. Variation in the scaled gradient of the cost function with the number of iterations for (a) $r = 4 \times 10^{17}$ s^4 and (b) $r = 1 \times 10^{18}$ s^4. Parenthetical values in the legend indicate the number of function calls required for 30 iterations. The ordinate is on a logarithmic scale.

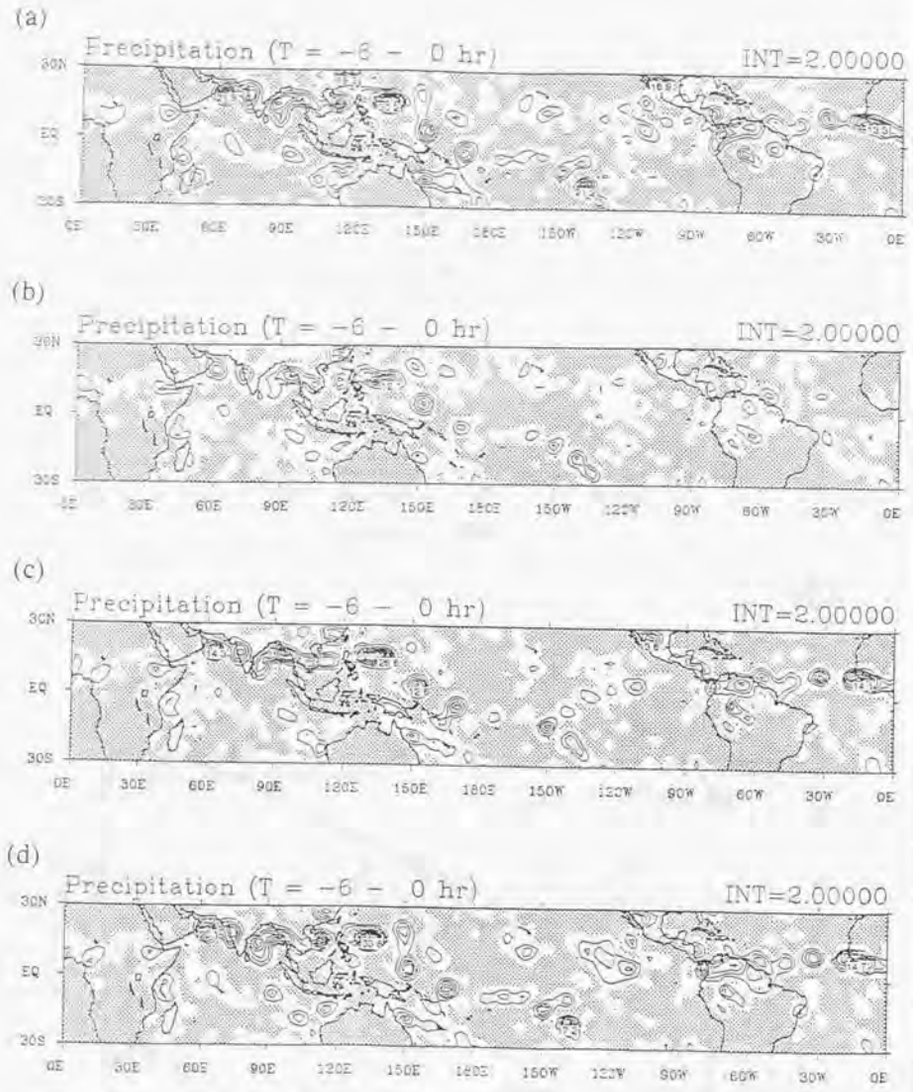

Fig. 16. 6-hour precipitation accumulated over the assimilation window for (a) the truth, (b) NOMP analysis, (c) MP analysis, and (d) MP+RR analysis with $r = 1 \times 10^{18}$ s^4. Contour interval is 2 mm. Areas less than 0.2 mm are shaded.

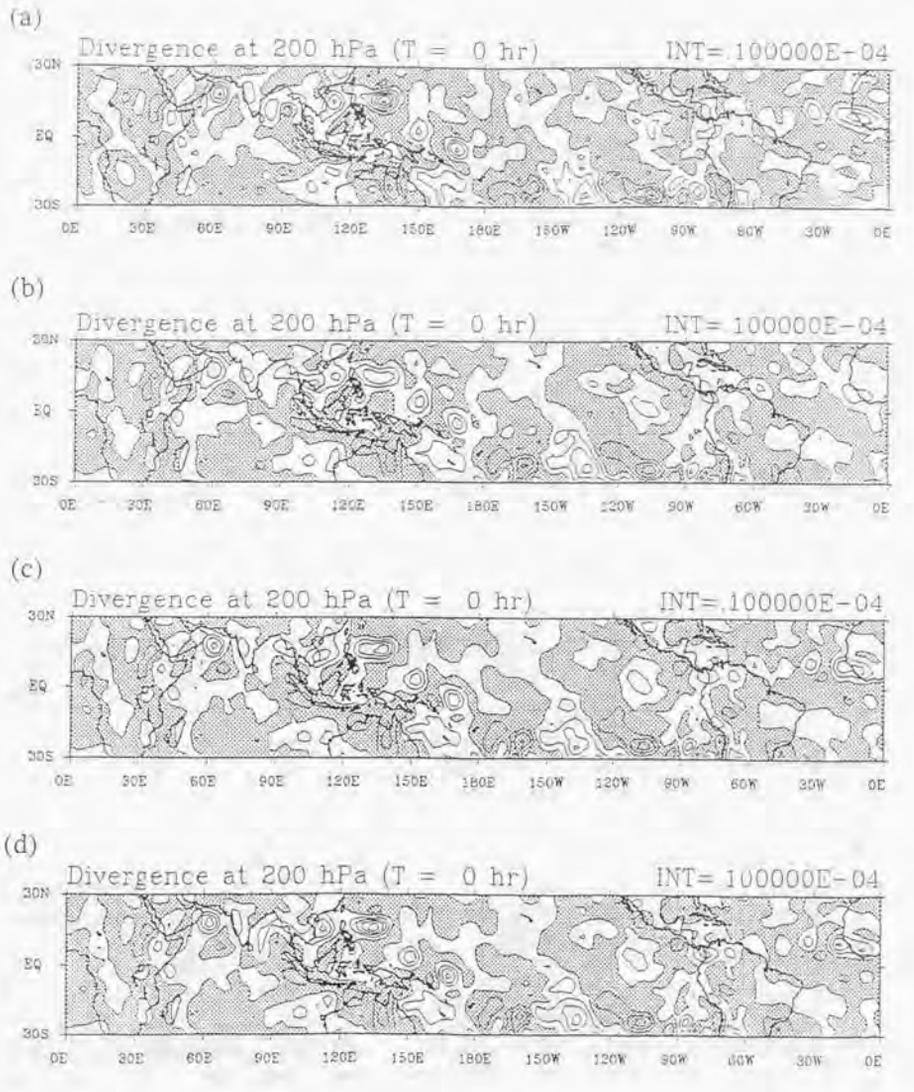

Fig. 17. Divergence at 200 hPa at $t = 0$ h for (a) the truth, (b) NOMP analysis, (c) MP analysis, and (d) MP+RR analysis with $r = 1 \times 10^{18}$ s^4. Contour interval is 1×10^{-5} s^{-1}. Negative areas are shaded.

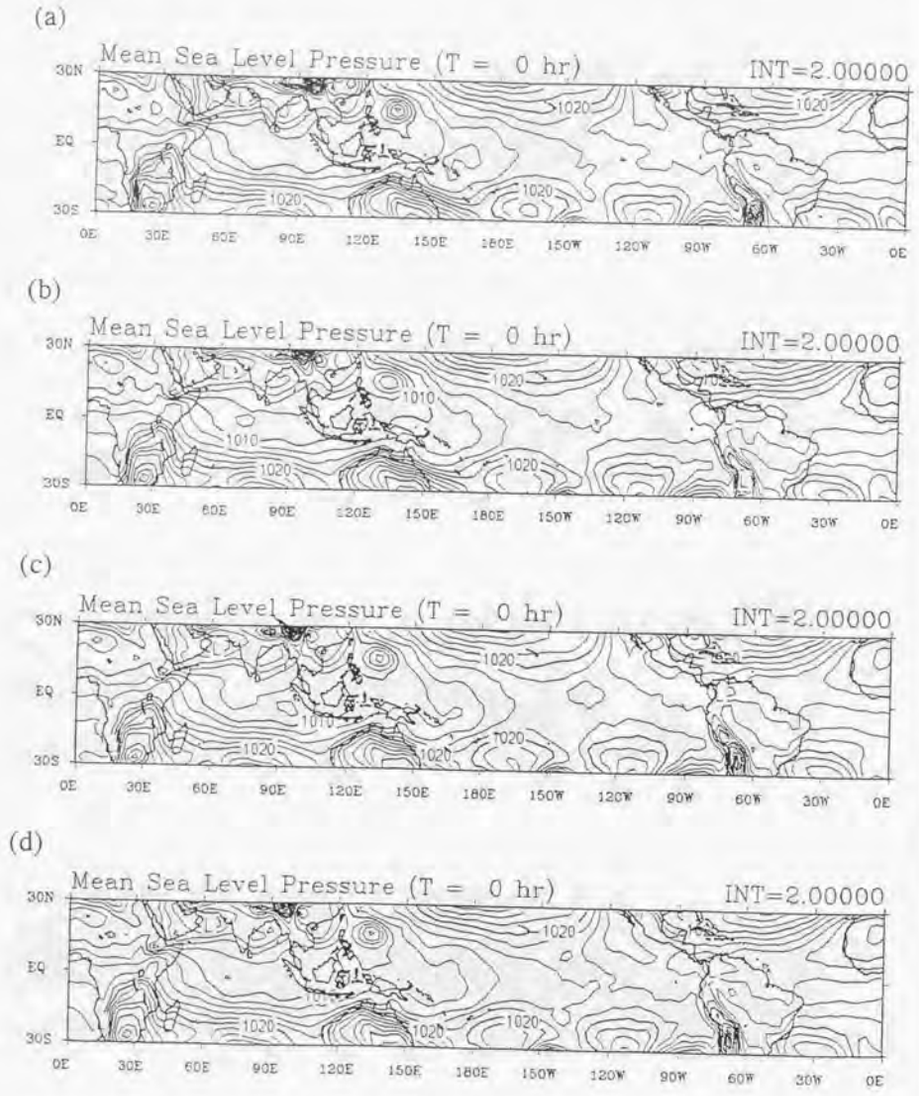

Fig. 18. Mean sea level pressure at $t = 0$ h for (a) the truth, (b) NOMP analysis, (c) MP analysis, and (d) MP+RR analysis with $r = 1 \times 10^{18}$ s^4. Contour interval is 2 hPa.

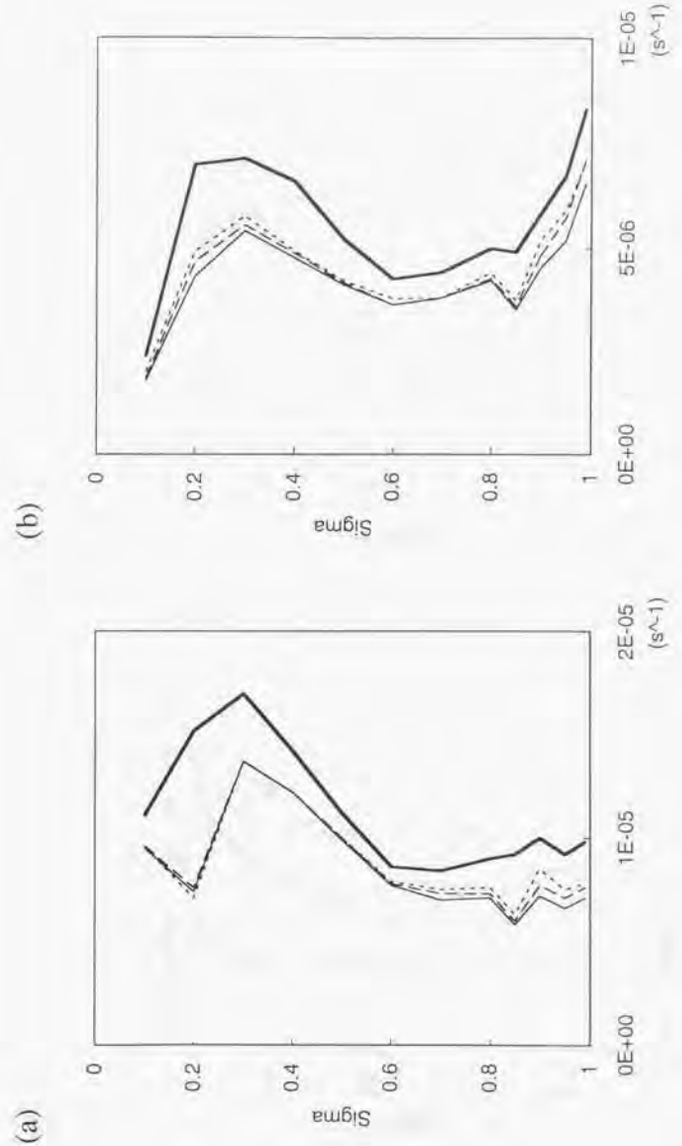

Fig. 19. Vertical distribution of RMSEs in (a) vorticity, (b) divergence, (c) temperature, and (d) relative humidity in the tropics (30 °N - 30 °S) at $t = 0$ h with $r = 1 \times 10^{18}$ s^4. Thick solid line is the first guess, dotted line is the NOMP analysis, dashed line is the MP analysis, and thin solid line is the MP+RR analysis.

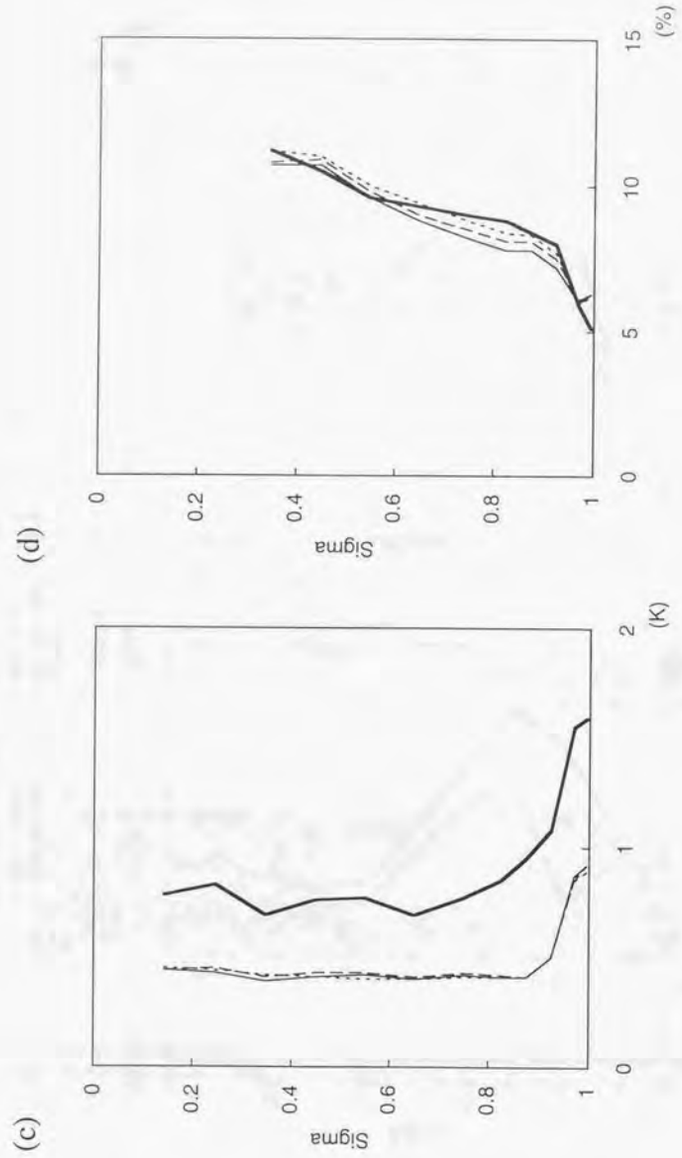

Fig. 19. *(Continued)*

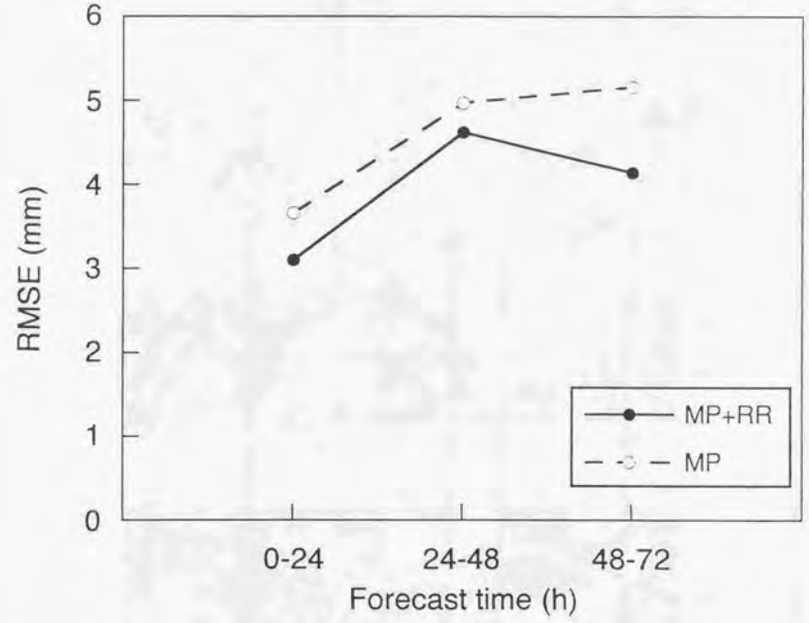

Fig. 20. RMSEs of 24-hour accumulated precipitation in the tropics (30 °N - 30 °S) for 72-hour forecasts starting from the initial conditions produced by the MP and MP+RR methods with $r = 1 \times 10^{18}$ s^4.

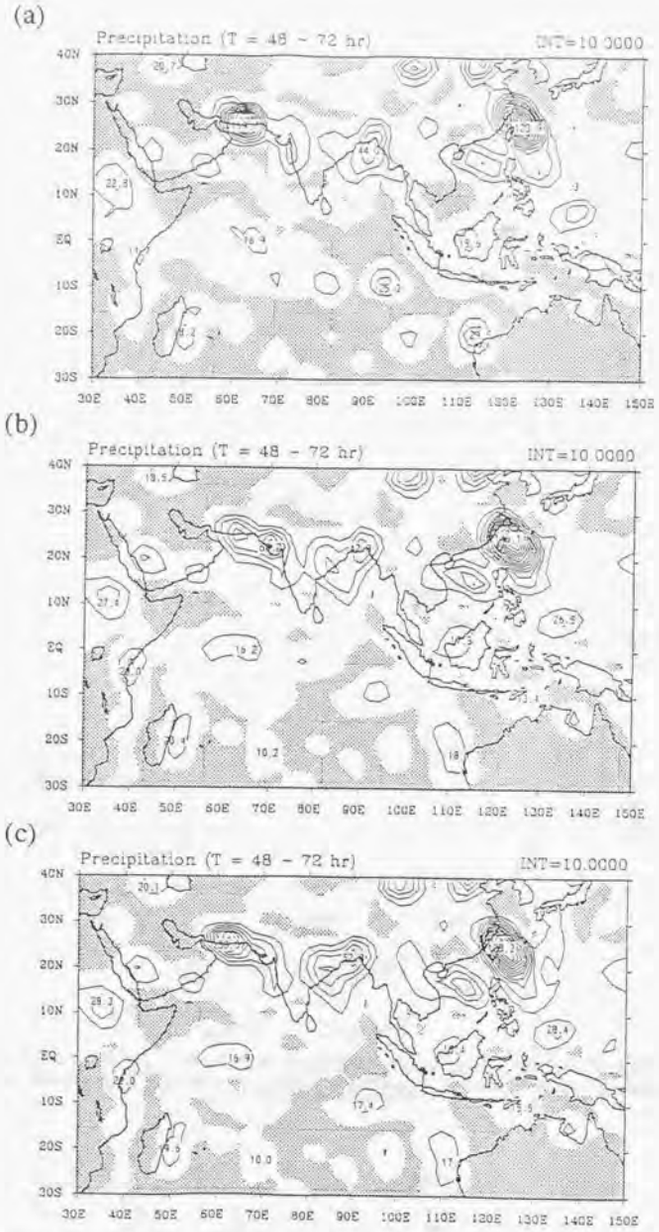

Fig. 21. 24-hour precipitation accumulated over t = 48 - 72 h for (a) the truth, (b) forecast starting from the MP analysis, and (c) forecast starting from the MP+RR analysis with $r = 1 \times 10^{28}$ s^4. Contour interval is 10 mm. Areas less than 1 mm are shaded.

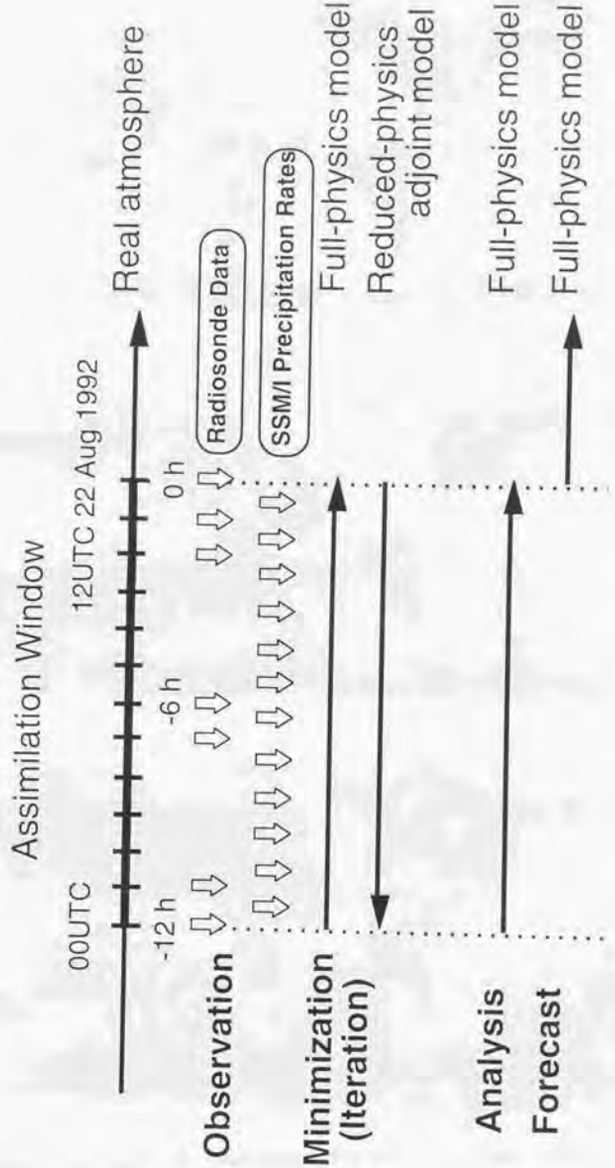

Fig. 22. Schematic illustration of experimental design for the real data experiment.

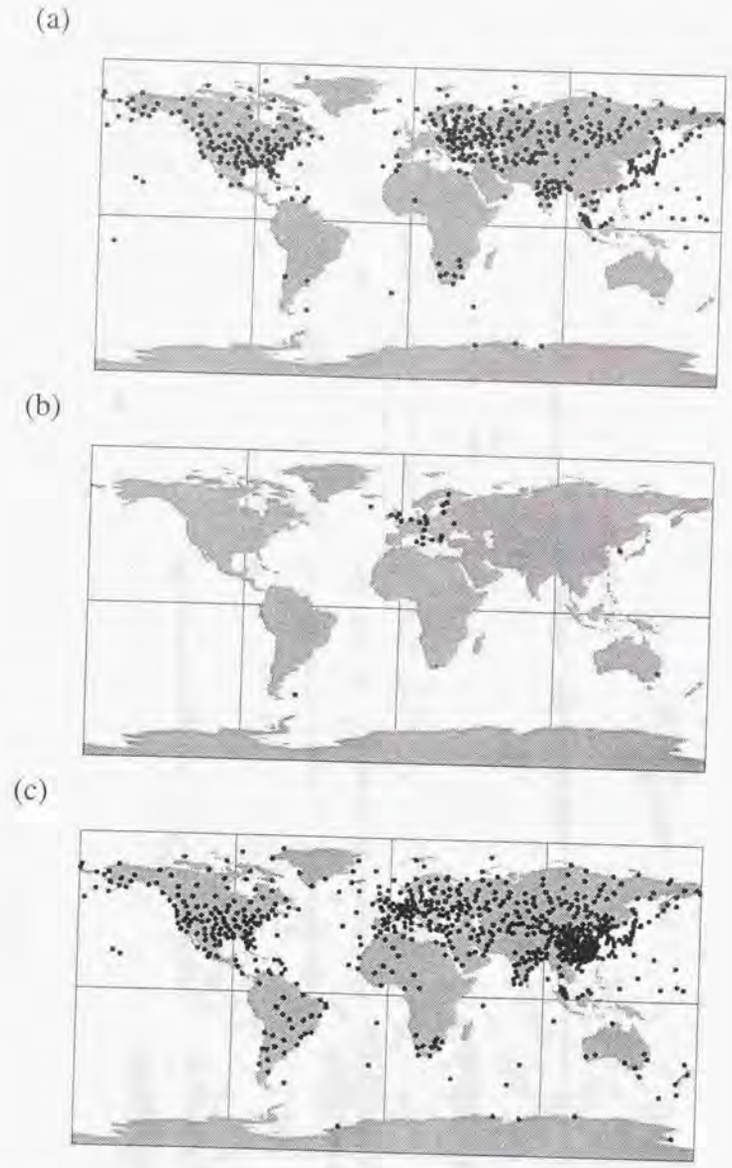

Fig. 23. Horizontal distribution of radiosonde stations for (a) 0 - 3 UTC, (b) 3 - 9 UTC, and (c) 9 - 12 UTC 22 August 1992.

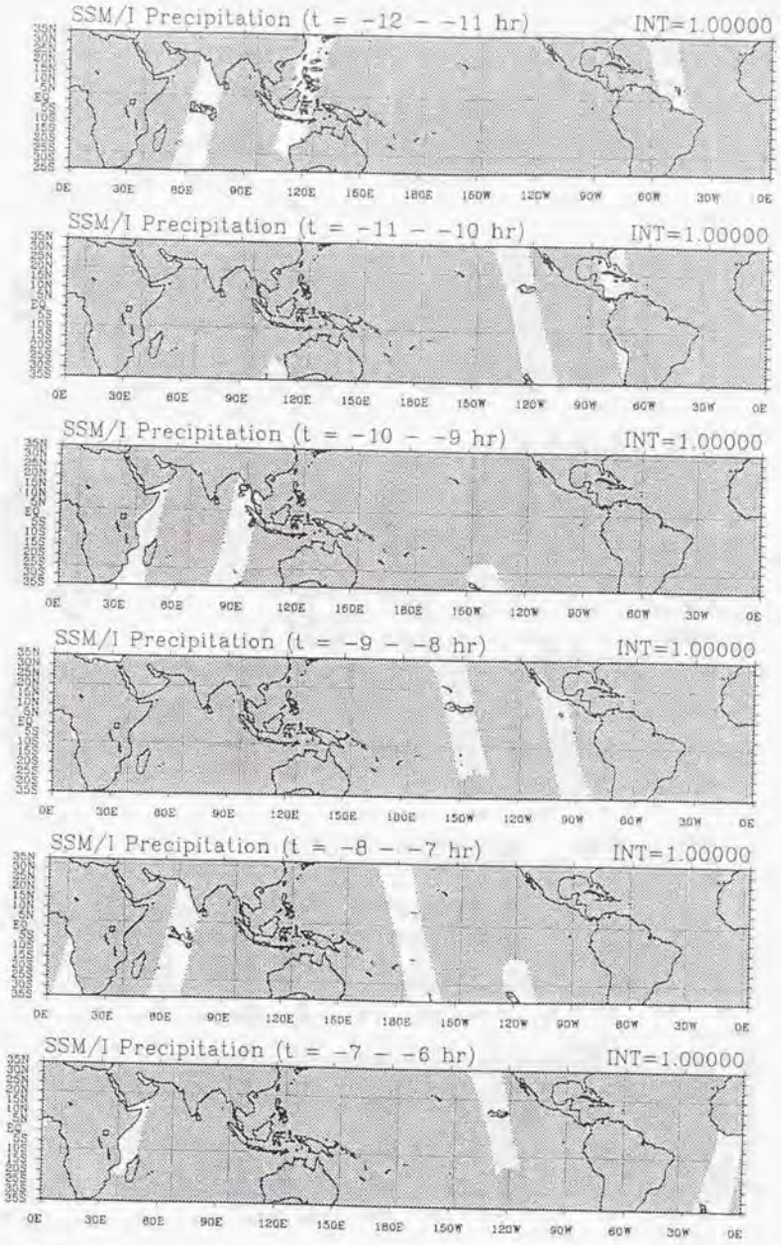

Fig. 24. SSM/I precipitation rates for 0 - 12 UTC (t = -12 - 0 h) 22 August 1992 at a 1-hour interval. Contour intervals are 0.25, 0.5, 1, 2, and 4 mm h^{-1}. Data void areas are shaded.

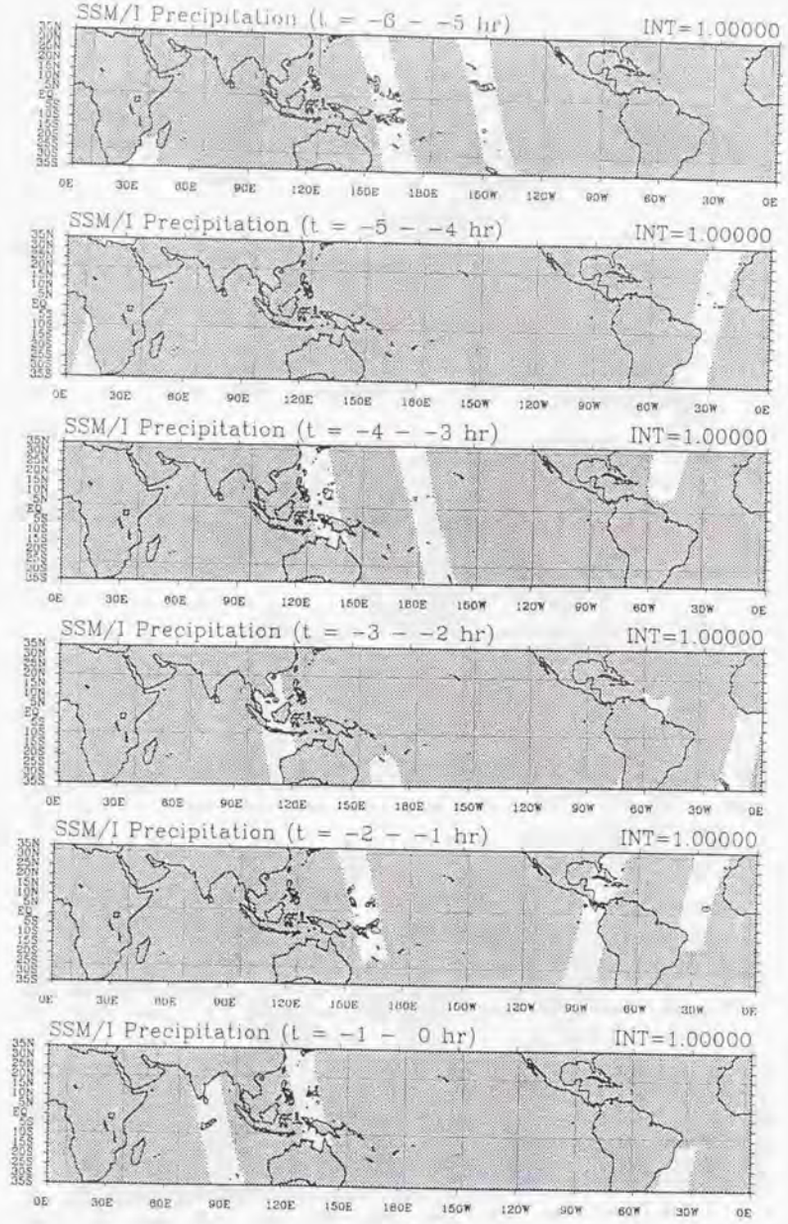

Fig. 24. (*Continued*)

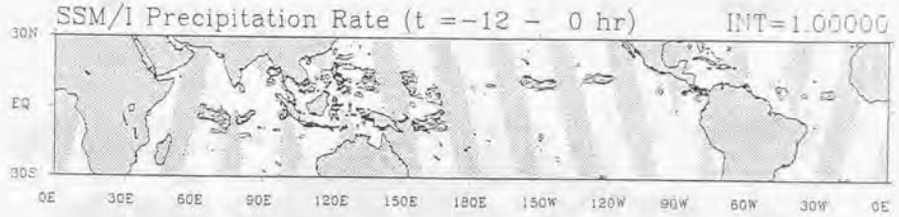

Fig. 25. Composite of SSM/I precipitation rates for 0 - 12 UTC 22 August 1992. Contour intervals are 0.25, 0.5, 1, 2, and 4 mm h^{-1}. Data void areas are shaded.

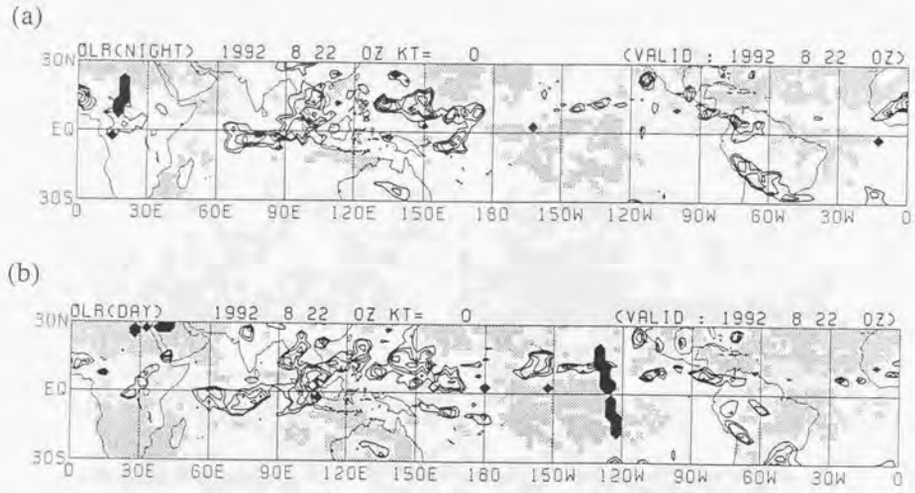

Fig. 26. Composites of outgoing longwave radiation for (a) the night pass and (b) the day pass of NOAA satellites for 22 August 1992. Contour interval is 25 W m^{-2} and contours greater than 200 W m^{-2} are omitted. Areas of greater than 275 W m^{-2} are lightly shaded. Data void areas are densely shaded.

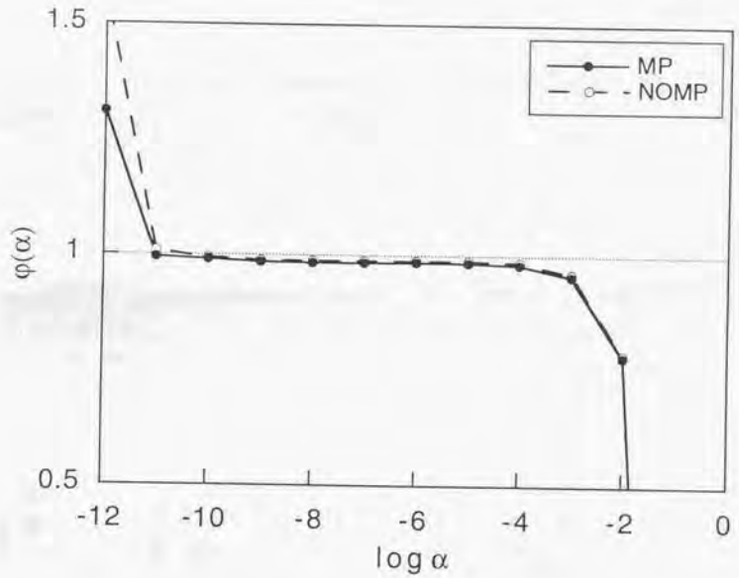

Fig. 27. Variations of the function $\varphi(\alpha)$ with $\log \alpha$ for the NOMP and MP methods with $r = 8 \times 10^{20}$ s^4. The first guess is taken as x_0 in (73).

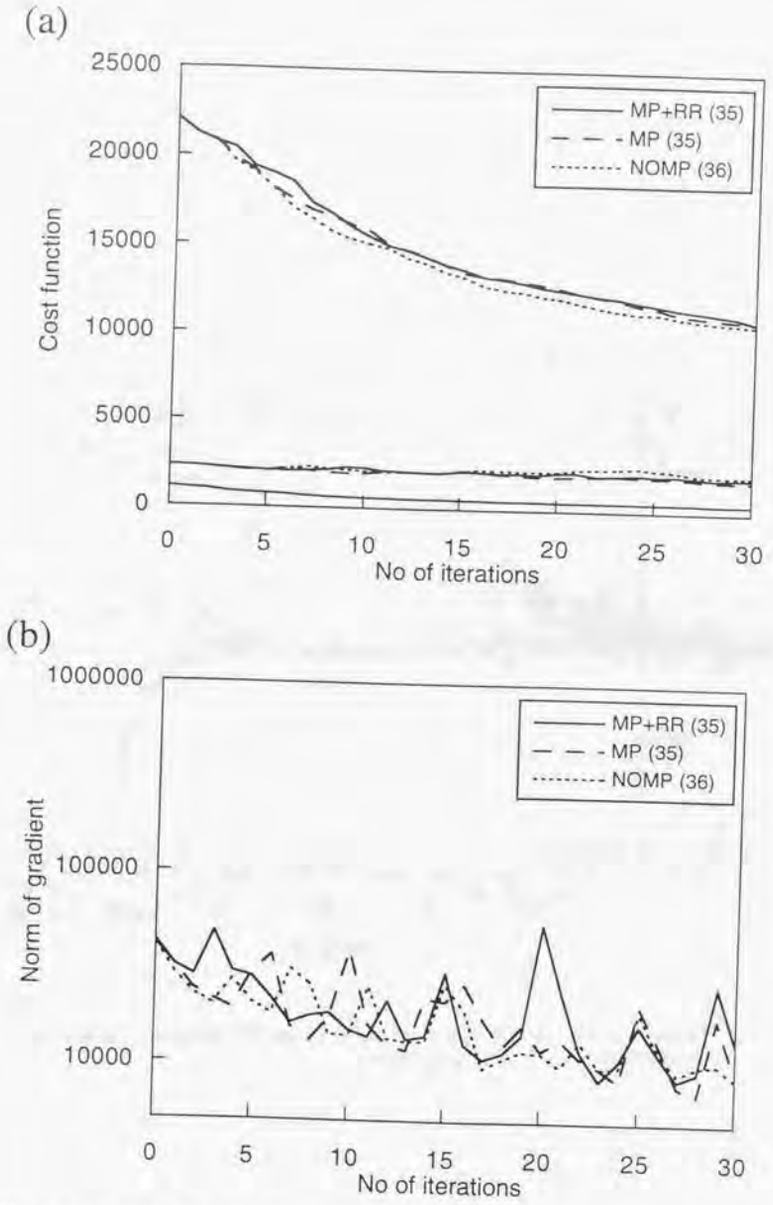

Fig. 28. Variations of (a) the cost function and (b) the scaled approximate gradient of the cost function with the number of iterations for the NOMP, MP, and MP+RR methods with $r = 8 \times 10^{20}$ s^4. In (a) the upper three lines are the discrepancy terms for radiosonde data, the middle three lines are the penalty terms, and the lower line is the discrepancy term for precipitation data for the MP+RR method. In (b) the ordinate is on a logarithmic scale. Parenthetical values in the legend of each panel indicate the number of function calls required for 30 iterations.

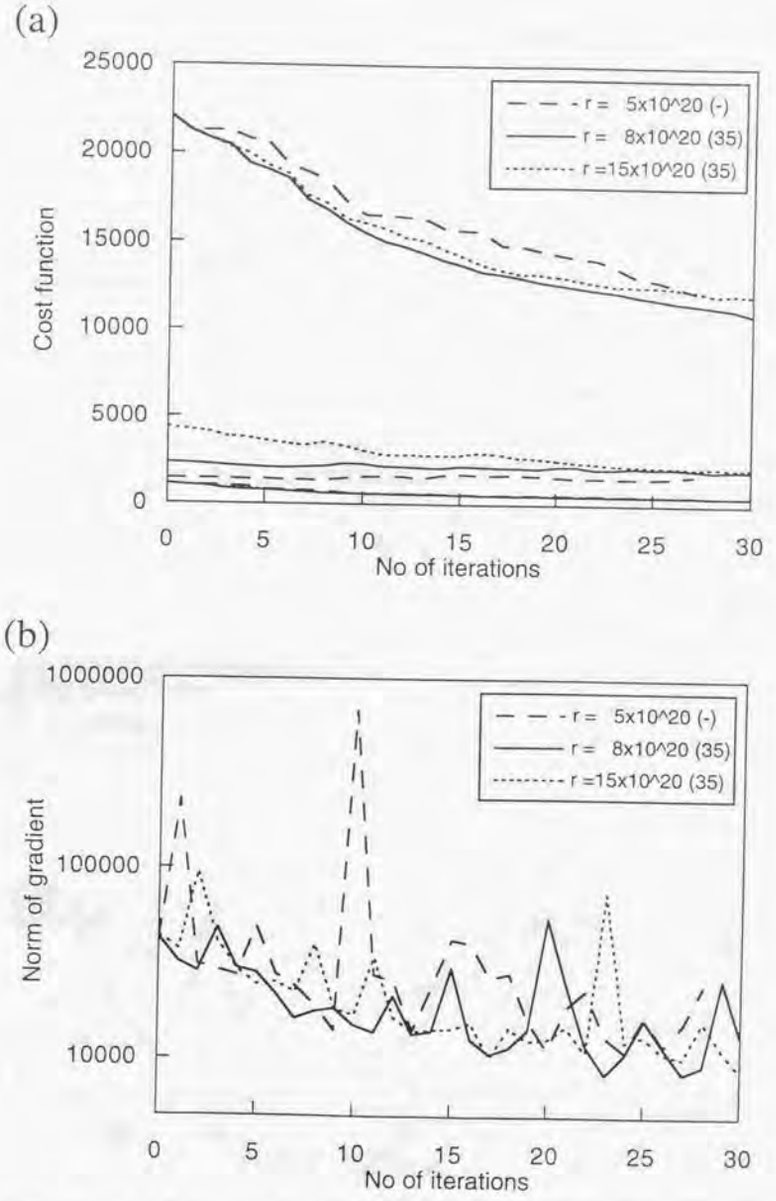

Fig. 29. Same as Fig. 28, except for the MP+RR method with $r = 5 \times 10^{20}$ s^4, 8×10^{20} s^4, and 15×10^{20} s^4. In (a) the upper three lines are the discrepancy terms for radiosonde data, the middle three lines are the penalty terms, and the lower three lines are the discrepancy terms for precipitation data.

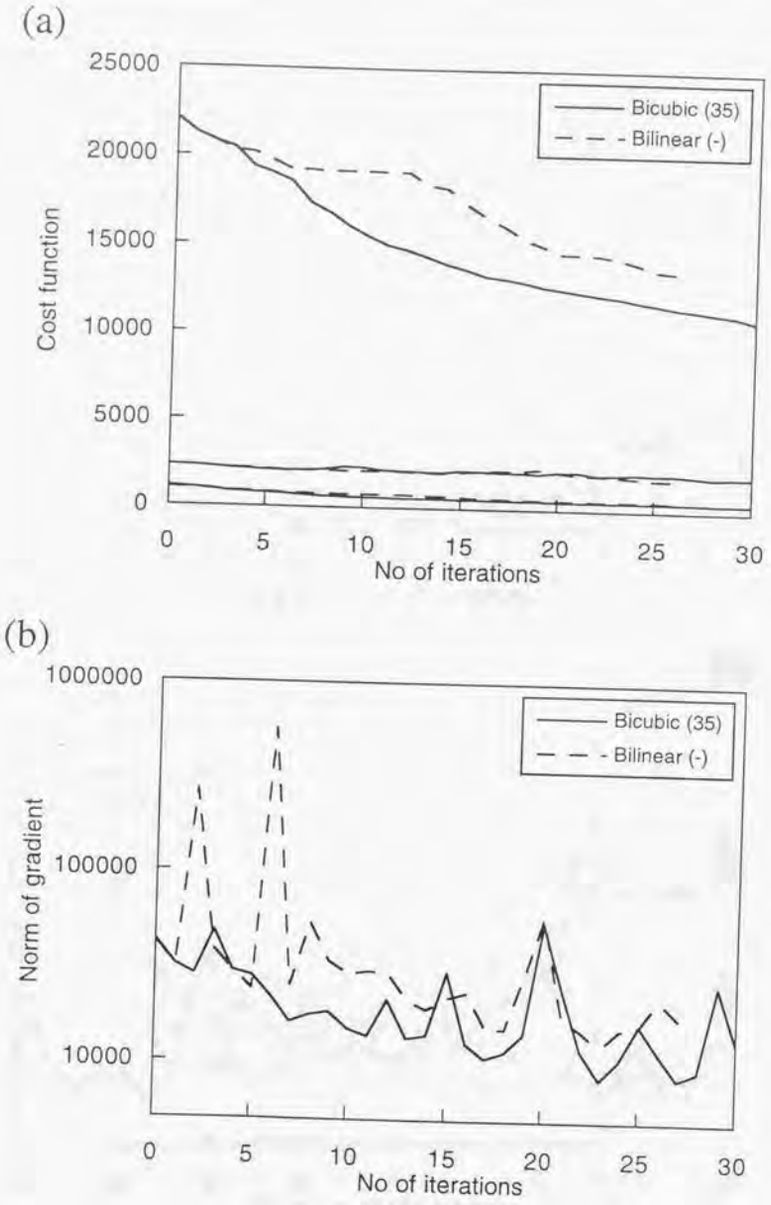

Fig. 30. Same as Fig. 28, except for the MP+RR method with bicubic and bilinear interpolations for precipitation data. In (a) the upper two lines are the discrepancy terms for radiosonde data, the middle two lines are the penalty terms, and the lower two lines are the discrepancy terms for precipitation data.

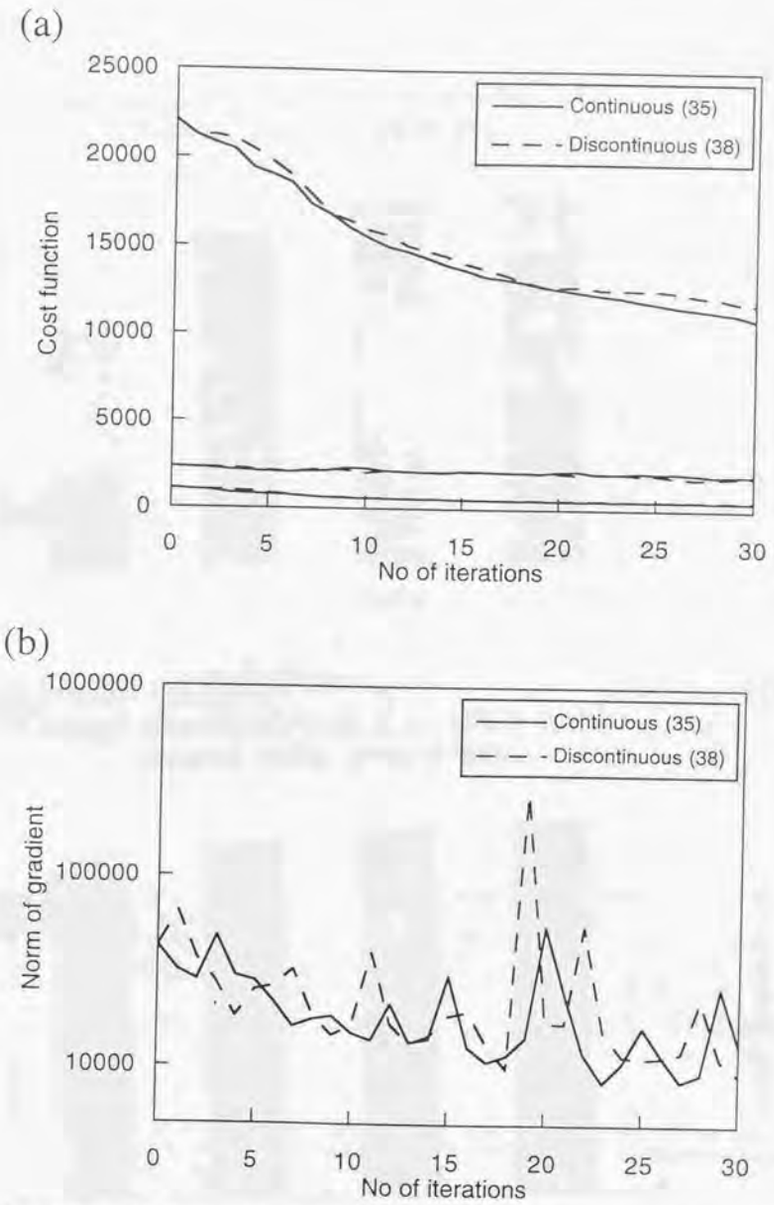

Fig. 31. Same as Fig. 28, except for the MP+RR method with the continuous and discontinuous moist processes. In (a) the upper two lines are the discrepancy terms for radiosonde data, the middle two lines are the penalty terms, and the lower two lines are the discrepancy terms for precipitation data.

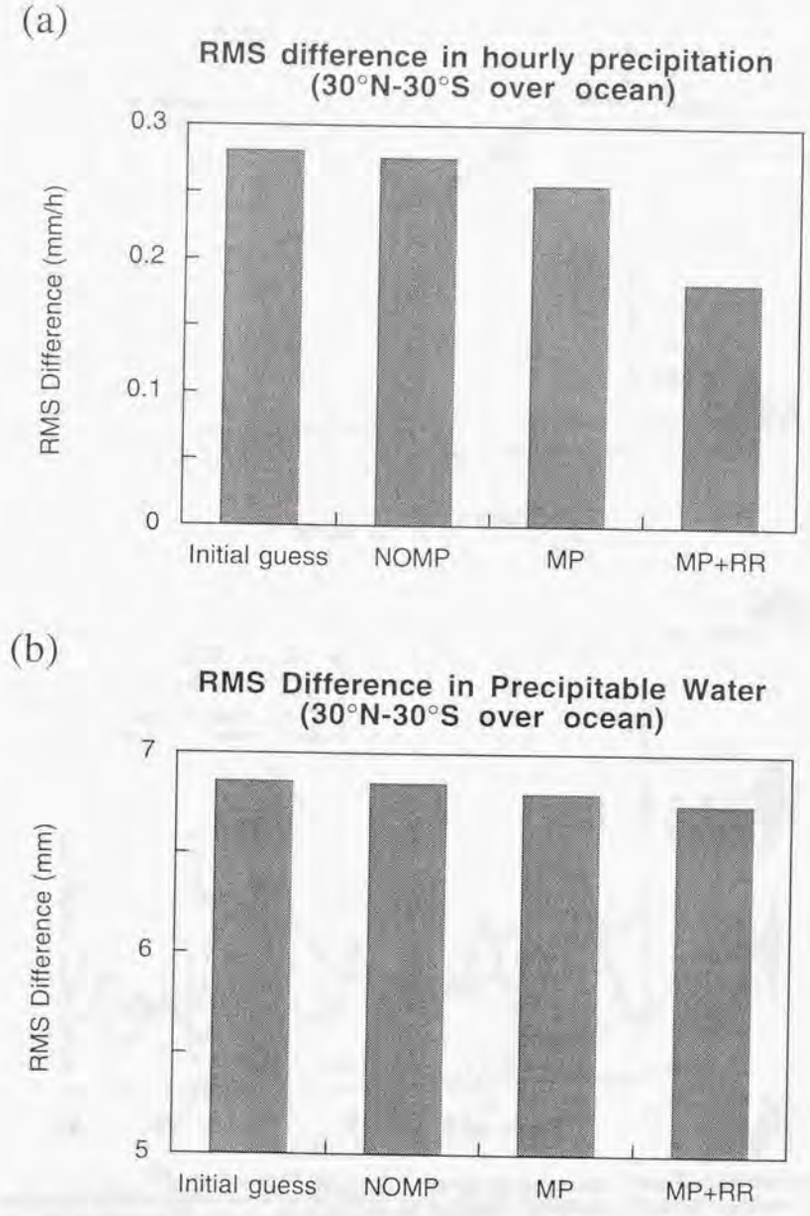

Fig. 32. Root mean square differences in (a) hourly mean precipitation rates and (b) precipitable water between the analyses and SSM/I observations over the tropical oceans (30 °N - 30 °S) over the assimilation window.

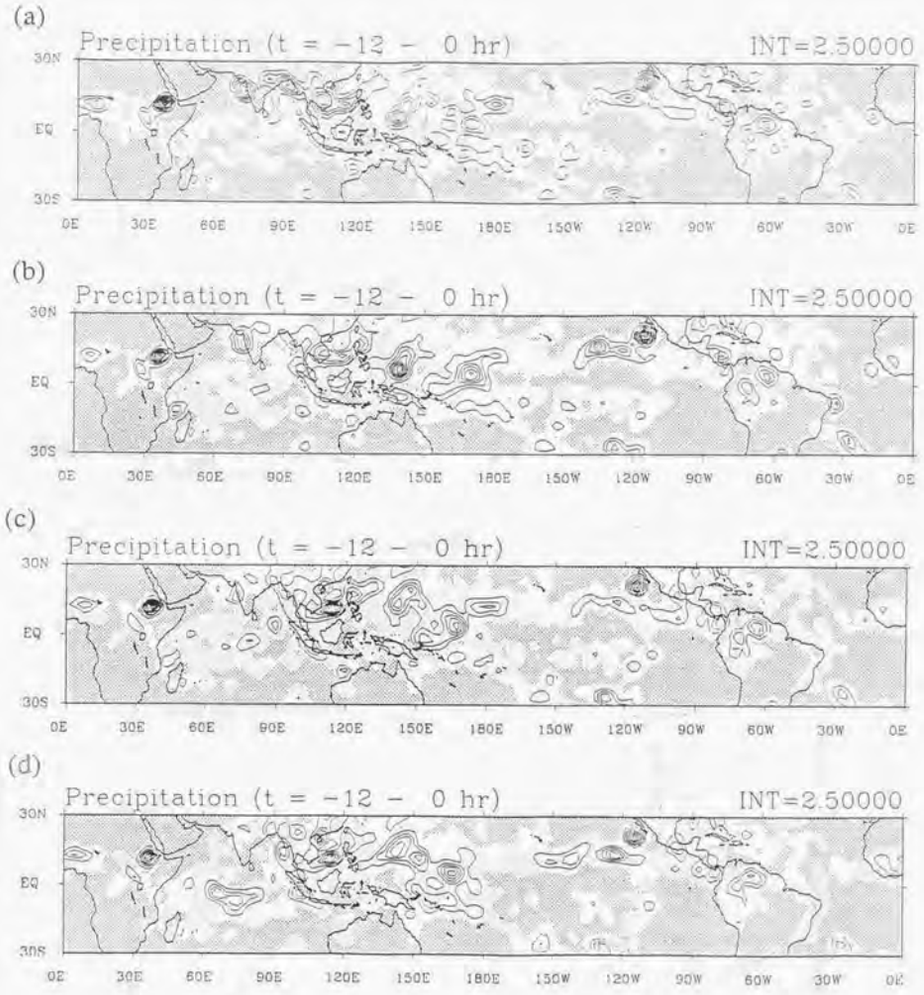

Fig. 33. 12-hour accumulated precipitation over the assimilation window for (a) the first guess, (b) NOMP analysis, (c) MP analysis, and (d) MP+RR analysis. Contour interval is 2.5 mm. Areas less than 0.25 mm are shaded.

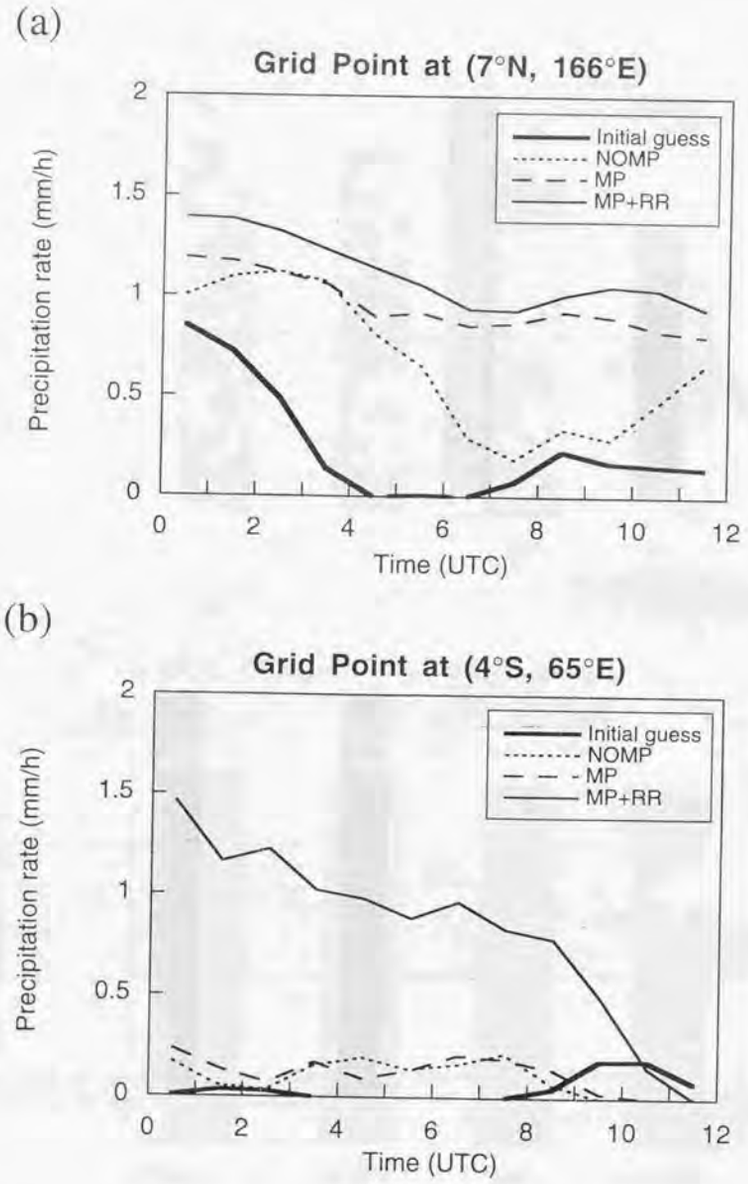

Fig. 34. Time sequences of precipitation rates for the assimilation window at (a) 7 °N, 166 °E and (b) 4 °S, 65 °E. One-hour intervals when SSM/I observations are available are shaded.

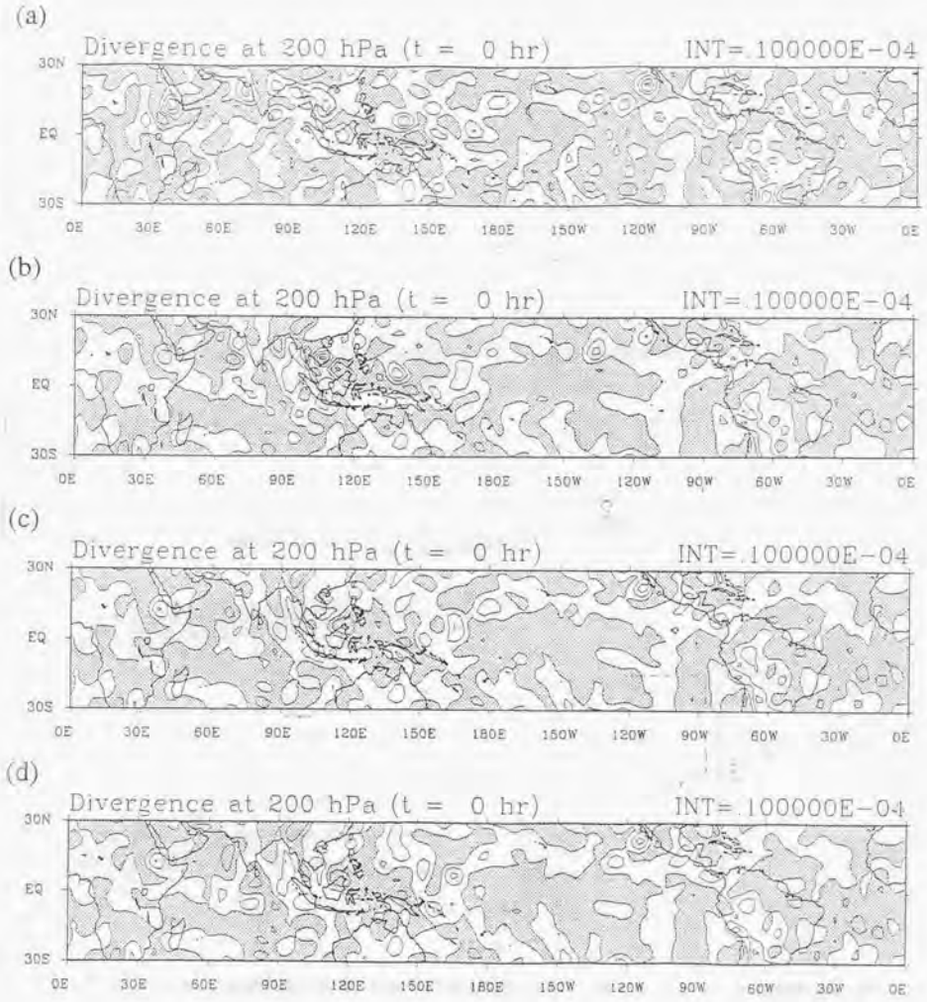

Fig. 35. Divergence at 200 hPa at the end of assimilation window for (a) the first guess, (b) NOMP analysis, (c) MP analysis, and (d) MP+RR analysis. Contour interval is 1 x 10^{-5} s^{-1}. Negative areas are shaded.

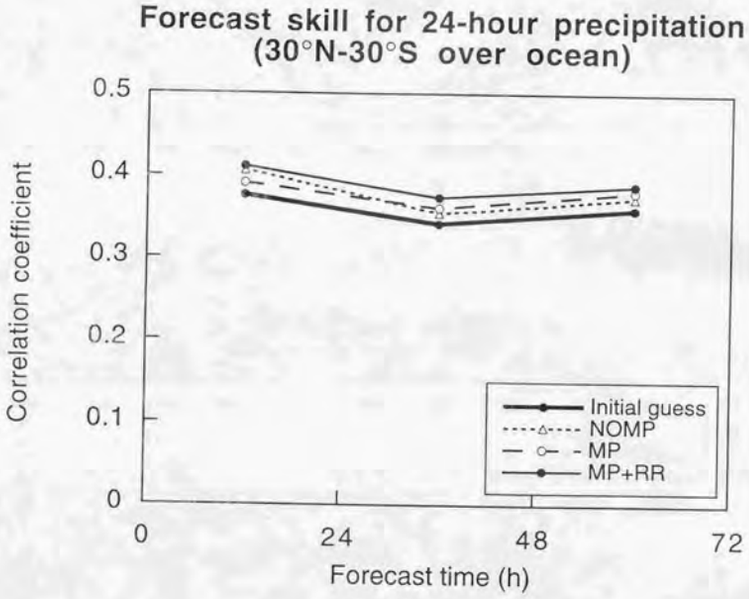

Fig. 36. Forecast skill for 24-hour accumulated precipitation over the tropical oceans (30 °N - 30 °S) up to 72 hours.