A Cloud-Resolving 4DVAR Assimilation Experiment for a Local Heavy Rainfall Event in the Tokyo Metropolitan Area

Takuya Kawabata, Tohru Kuroda, Hiromu Seko, and Kazuo Saito

Meteorological Research Institute/Japan Meteorological Agency, Tsukuba, Ibaraki, Japan

(Manuscript received 23 March 2010, in final form 9 December 2010)

ABSTRACT

A cloud-resolving nonhydrostatic four-dimensional variational data assimilation system (NHM-4DVAR) was modified to directly assimilate radar reflectivity and applied to a data assimilation experiment using actual observations of a heavy rainfall event. Modifications included development of an adjoint model of the warm rain process, extension of control variables, and development of an observation operator for radar reflectivity.

The responses of the modified NHM-4DVAR were confirmed by single-observation assimilation experiments for an isolated deep convection, using pseudo-observations of rainwater at the initial and end times of the data assimilation window. The results showed that the intensity of convection could be adjusted by assimilating appropriate observations of rainwater near the convection and that undesirable convection could be suppressed by assimilating small or no reflectivity.

An assimilation experiment using actual observations of a local heavy rainfall in the Tokyo, Japan, metropolitan area was conducted with a horizontal resolution of 2 km. Precipitable water vapor derived from global positioning system data was assimilated at 5-min intervals within 30-min assimilation windows, and surface and wind profiler data were assimilated at 10-min intervals. Doppler radial wind and radar-reflectivity data below the elevation angle of 5.4° were assimilated at 1-min intervals.

The 4DVAR assimilation reproduced a line-shaped rainband with a shape and intensity consistent with the observation. Assimilation of radar-reflectivity data intensified the rainband and suppressed false convection. The simulated rainband lasted for 1 h in the extended forecast and then gradually decayed. Sustaining the low-level convergence produced by northerly winds in the western part of the rainband was key to prolonging the predictability of the convective system.

1. Introduction

Heavy rainfalls are extreme meteorological phenomena and often cause disasters with loss of human life. Recent progress in numerical modeling and assimilation techniques has made it possible to predict to some extent the occurrence of heavy rainfalls induced by orographic or synoptic forcing. However, predicting small-scale convective rainfalls with weak forcing is still a numerical weather prediction (NWP) challenge. In Japan, such local heavy rainfalls are sometimes called “guerrilla” heavy rainfalls because of their unforeseeability. At present, local heavy rainfall is primarily forecast by kinematical methods based on extrapolation, but such methods have limited accuracy that degrades in a short time. Thus, the dynamical forecasting of local heavy rainfalls using a numerical model is one of the most important challenges in meteorology today.

A primary reason for the difficulty in predicting local heavy rainfalls is their small size. Conventional observations (e.g., aerological soundings) used in operational NWP systems are unable to capture phenomena at the typical spatial and temporal scales of deep convection of 1 h and 10 km, respectively. Another reason is the chaotic characteristics of deep convection, which is initiated in an unstable atmosphere and whose evolution is very sensitive to small perturbations in the initial conditions. To reproduce and predict such phenomena with numerical models, it is necessary to spatially and temporally...
assimilate dense observations corresponding to the scale of the deep convection and to precisely determine the initial conditions.

Data assimilation techniques connect observations with numerical models. One sophisticated assimilation method is the four-dimensional variational data assimilation (4DVAR) method, which can consider the model trajectory over the assimilation window. In a pioneering work, Zou and Kuo (1996) implemented 4DVAR with a mesoscale numerical model. They developed their 4DVAR based on the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) Mesoscale Model (MM5; Dudhia 1993) and investigated the impact of precipitable water vapor (PWV) derived from global positioning system (GPS) data on predictions of mesoscale convective systems (MCSs). However, in their system, the horizontal grid spacing was 40 km and cumulus convection was parameterized. Thus, local heavy rainfall was not targeted. The Japan Meteorological Agency (JMA) initiated its mesoscale 4DVAR system (Meso-4DVAR; Koizumi et al. 2005) in 2002, making it the first operational regional 4DVAR system in the world. By assimilating 1-h accumulated rainfall amounts derived from radar-reflectivity data, they improved the accuracy of the JMA operational mesoscale forecasts. However, the horizontal grid spacing of Meso-4DVAR was 20 km, and the precipitation scheme in the adjoint model adopted only large-scale condensation and convective adjustment. In 2009, Meso-4DVAR was replaced with a different 4DVAR system [i.e., the JMA Nonhydrostatic Model-based Variational Data Assimilation System (JNoVA); Honda et al. 2005]. Although JNoVA is based on nonhydrostatic dynamics, its horizontal grid spacing is 15 km and only large-scale condensation is considered in the adjoint model. The Met Office has also been operating a mesoscale 4DVAR system, but with a horizontal grid spacing of 24 km. Huang et al. (2009) developed and obtained preliminary results from an enhanced version of MM5-4DVAR called the Weather Research and Forecasting 4DVAR (WRF-4DVAR) with a horizontal resolution of 45 km.

Radar observations provide the most useful data for high-resolution assimilation systems because radars are deployed at many observation sites and can capture MCSs with high temporal and spatial resolution. One of the most important radar observational elements is Doppler radial winds (RW). Most MCSs are induced by the low-level convergence of water vapor. Therefore, detecting low-level convergence is key to successfully predicting MCSs. The assimilation of RW contributes to the reproduction of low-level wind fields. Another important observational element of radars is reflectivity. Various methods of using reflectivity for cloud-scale assimilation have been examined.

The simplest method for assimilating reflectivity may be to nudge the rainwater produced by the model toward that retrieved from the observed reflectivity. Sokol and Rezacova (2009) assimilated radar reflectivity to local model Consortium for Small-scale Modeling (COSMO; Doms and Schaeftler 1999) using a nudging method that converted the reflectivity to latent heating, and succeeded in simulating MCSs for several hours. However, with the nudging method, the impact of assimilation usually does not continue for a long time, and the assimilated rainwater sometimes vanishes too soon because dynamical and thermodynamical balances are neglected.

Xiao et al. (2007) developed a three-dimensional variational data assimilation system (3DVAR) to directly assimilate radar reflectivity. By applying cloud microphysics in the conversion process between the model prognostic variables and control variables in MM5-3DVAR (Barker et al. 2004) with 10-km grid spacing, they were able to improve the forecast of a typhoon. 3DVAR is more sophisticated than the nudging method and can directly assimilate radar reflectivity. However, 3DVAR cannot consider model dynamics in the assimilation procedure.

The pioneers of radar data assimilation in the cloud scale are Sun and Crook (1997, 1998). Sun and Crook (1997) developed a 4DVAR system called the variational Doppler radar analysis system (VDRAS) to assimilate radar reflectivity and RW with an anelastic nonhydrostatic model. They demonstrated the impact of the radar data assimilation using an Observing System Simulation Experiment (OSSE). Their system reproduced winds, thermodynamics, and cloud microphysics fields well. Sun and Crook (1998) applied their system to an actual convective storm, and reported good agreement between the simulation and aircraft observations. However, in their experiment, they used a uniform horizontal field as the initial condition, and convection was first initiated by an artificial warm bubble. Only 5- and 7-min assimilation windows were adopted in a narrow experimental domain of 11.2 km², and orography was not considered in the model. Their studies were the first trial of radar data assimilation in a cloud-scale 4DVAR, but the experimental configurations were unable to support actual full-scale short-range forecasts of local convective rainfall.

Snyder and Zhang (2003) reported the potential of an ensemble Kalman filter (EnKF) for use in radar data assimilation. They used 50 members of the numerical model of VDRAS with 2-km horizontal grid spacing and assimilated radial wind and reflectivity observations with 5-min intervals. The EnKF reduced the analysis error during assimilation cycles and reproduced
unobserved variable fields. Xue et al. (2006) also used the EnKF technique to assimilate radar reflectivity data directly with a horizontal resolution of 1.5 km as an OSSE. Aksoy et al. (2009) attempted to assimilate radial wind and reflectivity data obtained by actual radar observations from the Weather Surveillance Radar-1988 Doppler (WSR-88D) into an EnKF, which consisted of 50 members of WRF model with 2-km horizontal grid spacing, and obtained reasonable analysis results for their experiment cases.

Recently, Sun and Zhang (2008) assimilated radar reflectivity from multiple Doppler radars using VDRAS improved in a number of ways with 4-km grid spacing. In their experiment, the forecast of a squall line 4 h after the initial time was improved compared with the background fields, but the reproduced squall line had a horizontal scale of 300 km and was thus more predictable than a local heavy rainfall event. Kawabata et al. (2007) developed a cloud-resolving 4DVAR based on the JMA (nonhydrostatic model) NHM-4DVAR. Although the adjoint model included only dry dynamics and advection of water vapor, they succeeded in reproducing observed cumulonimbi by assimilating RW and GPS-PWV data. Their study was the first to demonstrate the feasibility of short-range forecasting of local heavy rainfall brought about by deep convection, using a full-blown numerical model and a dense observation network. They were able to successfully reproduce deep convection because the convection was initiated by the low-level convergence of a sea-breeze front, and low-level winds were observed by Doppler radars as clear-air echoes around the initiating point of the cumulonimbi. However, it should be noted that low-level features are not always observed by radar instruments.

To prevent disasters caused by heavy rainfall events, JMA issues 33-h mesoscale model (MSM) predictions 8 times a day, and 6-h kinematic very-short-range forecasts every 30 min. A merging method that integrates extrapolations of radar observations and forecasts by MSM with a weighting function is used in this kinematic forecast. It is very difficult, however, to predict convection initiation and decay by such methods. The aim of this study is to issue 2–6-h forecasts, because a kinematic forecast is useful only within 1–2 h.

On 4–5 September 2005, a local heavy rainfall event occurred in the Tokyo, Japan, metropolitan area. A maximum total rainfall of over 200 mm was recorded in the western part of Tokyo. This event was caused by a small-scale line-shaped convective system about 100 km long and 15 km wide. As is discussed in section 4, the Tokyo metropolitan area has a very high density of operational observation networks, including three Doppler radars, more than 30 GPS observation sites, over 30 automated surface-observation sites, and four wind profilers in an area of 150 km². We chose this event as a test case for an assimilation experiment of deep convection with a cloud-resolving 4DVAR system.

In this study, we first modified NHM-4DVAR to assimilate radar-reflectivity data directly, newly considering perturbations in rainwater and cloud water, and implementing an adjoint model of cloud microphysics. Then an observation operator for radar reflectivity was developed.

Here, we demonstrate that radar-reflectivity assimilation with a cloud-resolving 4DVAR improves the reproduction of observed MCSs. Because Kawabata et al. (2007) demonstrated the impact of GPS-PWV and RW, and these data have already been used in JNoVA as well, we focus on the impact of radar reflectivity in this paper. We describe the NHM-4DVAR modifications in section 2, and single-observation experiments to observe the 4DVAR responses to reflectivity data in section 3. In 4, we describe the heavy rainfall event and observations, and the results of the 4DVAR experiment in section 5. In section 6, we discuss the effect of assimilation of 0-dBZ information and the sustainment of the convective system. Section 7 consists of a summary and the conclusions.

2. NHM-4DVAR

2a. System

NHM-4DVAR is a cloud-resolving nonhydrostatic four-dimensional variational assimilation system based on the JMA NHM (Saito et al. 2007) and is designed to reproduce and predict MCSs with cloud-resolving resolutions. A full model of NHM is adopted as a forward model, which includes in particular three-ice bulk cloud physics. The adjoint model used previously (Kawabata et al. 2007) considered dry dynamics and advection of water vapor, but this study implements a warm rain process (see section 2b). NHM adopts a terrain-following vertical coordinate (z* coordinate; Gal-Chen and Somerville 1975).

Perturbations of water substances (the mixing ratios of rainwater and cloud water) are considered, and new control variables associated with water substances are installed (see section 2c). The main prognostic variables and processes of the NHM and perturbations of NHM-4DVAR are listed in Table 1. The cost function is formulated as follows:

\[
J(x^0, x_{lbc}) = \frac{1}{2}(x^0 - x_b^0)^T B^{-1}(x^0 - x_b^0) \\
+ \frac{1}{2}(x_{lbc} - x_{lbc}^b)^T B_{lbc}^{-1}(x_{lbc} - x_{lbc}^b) \\
+ \frac{1}{2}(Hx - y^o)^T R^{-1}(Hx - y^o),
\]

(1)
where $x_0$ comprises the model prognostic variables and lateral boundary conditions at the beginning time of the assimilation window, $x_{lbc}$ is the lateral boundary conditions at the end time, and $x_0^b$ and $x_{lbc}^b$ are first-guess fields of $x_0$ and $x_{lbc}$ respectively. Here $B$ ($B_{lbc}$) represents the background covariance matrix associated with the initial (lateral) boundary conditions. Lateral boundary conditions, except at the initial and end times, are obtained by linear interpolation. Here $H$ is the observation operator, $y^o$ comprises the observations, and $R$ is the observation error covariance matrix.

In a precise sense, $B$ and $B_{lbc}$ should be different, because initial and lateral boundary conditions are generally given by different models. In this study, because first guesses of the initial and lateral boundary conditions are derived from the same simulation results, the same matrix is used for both. In NHM-4DVAR, perturbations of water substances are considered in the initial field but not in the lateral boundary conditions. Therefore, $B_{lbc}$ does not include the covariance associated with water substances, but uses instead pseudorrelative humidity.

Because the forward model of NHM-4DVAR is a full-blown nonlinear model whereas the adjoint model is based on a “simplified” tangent linear model, there is an inconsistency between the forward and the adjoint models. Sometimes this inconsistency may cause solution divergence in the minimization procedures. In this study, we adopted short-term assimilation windows and a spinup process to improve first-guess fields and prevent solution divergence (see section 4c). Since the full nonlinear model can represent a more realistic field than the simplified tangent linear model, especially in a highly nonlinear situation, the full nonlinear model can be used to evaluate a precise cost function $J$ for assimilation of such a highly nonlinear situation. This point is important for convergence of minimization. In our discussion of this problem, we will describe an incremental 4DVAR system that adopts the tangent linear model as a forward model (section 6b).

### Warm rain process

We developed an adjoint model of the warm rain microphysical process that includes a Kessler-type parameterization scheme. A technique developed by T. Kuroda (2007, personal communication) was employed to detect nonlinear branches in this microphysical scheme. In this method, the results of all branches in particular subroutines are stored in memory and their influences are investigated. Using this method, an “IF” branch in FORTRAN program codes with strong nonlinearity was found in the warm rain process of the original NHM code. This branch was added to prohibit the evaporation of rainwater at a grid point where cloud water exists.

With careful treatment, a nonlinear IF branch can exist in a tangent linear model (Xu 1996), but such treatment has a high development cost. Tsuyuki (1996a,b) determined that removing the discontinuity from a nonlinear process can improve the linearity of a tangent linear model. This nonlinear IF was removed from the forward, tangent linear, and adjoint models of NHM-4DVAR after confirmation that its removal led to only very trivial differences in the simulation results.

### Control variables

In an assimilation system in which water substances are not treated as control variables, cloud water and rainwater are absent at the initial time and are produced from water vapor during the model time integration. The time lag that is needed to create water substances is called the spinup time. Assimilation of the radar reflectivity at the start time of an assimilation window thus allows the spinup time to be minimized. For this reason, we added control variables associated with water substances to NHM-4DVAR.

One challenge in including water substances in the control variables is treatment of their background errors. We assumed that the background errors had a horizontally uniform distribution. However, water substances in the model often exist very locally, thus, the shapes of their background errors may not follow a Gaussian distribution.

In MM5-3DVAR and VDRAS, the control variable associated with water substances is total water, which is the sum of the mixing ratios of water vapor $q_v$, cloud water $q_c$, and rainwater $q_r$. Adopting total water requires separating it into $q_v$, $q_c$, and $q_r$. In VDRAS, total water is used as a prognostic variable in the forward nonhydrostatic model, so separation is handled during the model integration. The prognostic variables of water substances in MM5 are $q_v$, $q_c$, and $q_r$ as in NHM. In MM5 and NHM, it is necessary to employ a separation process so that the control variables can be converted into the
model prognostic variables. Separation of \( q_v \) and \( q_c \) is a linear process that uses a saturation adjustment. However, separation of \( q_r \) is a nonlinear process because of cloud microphysics and it is inconvenient to perform the separation for the assimilation.

To implement microphysical processes in NHM-4DVAR, we examined several candidate control variables considering the aspects of background errors and nonlinearity. Finally, we chose two variables: total water (\( q_r + q_c \)) and the relative mixing ratio of rainwater (\( q_r/q_{\text{vap}} \)), where \( q_{\text{vap}} \) is the saturation mixing ratio of water vapor given by the background). A probability density function (pdf) of the background error of total water (\( q_r + q_c \)) is presented in Fig. 1. The shape is not Gaussian but it is much closer to Gaussian than pdfs of the individual variables (\( q_r, q_c \)). The shape of the \( q_r \) error is similar to that of total water, but the \( q_c \) error is spiked shaped (not shown). The individual variable \( q_r \) is also spike shaped, whereas the relative mixing ratio of rainwater becomes Gaussian, similar to relative humidity. We confirmed that the correlations between the two chosen variables and other variables were not large. Specifically, the correlation between potential temperature and total water is about 0.2, that between total water and the relative mixing ratio of rainwater is less than 0.1, and that between potential temperature and the relative mixing ratio of rainwater is -0.4.

Background errors were determined by the NMC method (Parrish and Derber 1992). One-hour forecast errors of NHM with 2-km grid spacing were calculated from the difference between two forecasts initiated at different times. These forecasts were initiated using the downscaled initial fields from the relatively coarse contemporary JMA operational regional model. As shown by Kawabata et al. (2007), their Fig. 1), the horizontal scales of the errors were suitable for the cloud-resolving assimilation system, but their standard deviations were too small. Therefore, the standard deviations of the errors were inflated by a factor of 2.0 for dynamical variables and by a factor of 3.0 for variables related to water vapor and water substances. Kawabata et al. (2007) present detailed information and figures regarding this topic.

Table 2 lists the new control variables. Pseudorelative humidity (Dee and Da Silva 2003, \( q_r/q_{\text{vap}} \)), which was employed in the former version of NHM-4DVAR, is used as a control variable only for lateral boundary conditions, while total water is adopted as the control variable related to water vapor for the initial condition.

### Observation operator

Only the observation operator for radar reflectivity is described here; other operators have been described by Kawabata et al. (2007).

We adopted the \( Z-Q_r \) relationship from Sun and Crook (1997) as the observation operator for assimilating radar reflectivity:

\[
10^{[Z-43.1]/17.5]-3.0 = \rho q_r. \tag{2}
\]

Here \( Z \) is the observed radar reflectivity (\( Z \)), \( \rho \) is the density of the model atmosphere (kg m\(^{-3}\)), and \( q_r \) is the mixing ratio of rainwater (kg kg\(^{-1}\)).

An alternative approach is to use a \( Z-R \) relation to relate the reflectivity to the precipitation accumulation \( R \), but this involves weak nonlinearity. In contrast, Eq. (2) is perfectly linear because the lhs is a model variable and can be calculated prior to the assimilation process.

To avoid the uncertainty of weak reflectivity (i.e., the uncertainty in discriminating rainwater from other possible targets), observations below 10 dBZ were given.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind</td>
<td>( u, v, w )</td>
</tr>
<tr>
<td>Potential temperature, surface pressure</td>
<td>( \theta, P_s )</td>
</tr>
<tr>
<td>Nonhydrostatic pressure</td>
<td>( \pi - \pi_B )</td>
</tr>
<tr>
<td>Total water (for ICs)</td>
<td>( q_v + q_c )</td>
</tr>
<tr>
<td>Relative mixing ratio of rainwater</td>
<td>( q_r/q_{\text{vap}} )</td>
</tr>
<tr>
<td>Pseudo–relative humidity (for LBCs)</td>
<td>( q_r/q_{\text{vap}} )</td>
</tr>
</tbody>
</table>

---

**Fig. 1.** Probability density function of the background error of total water (\( q_r + q_c \)).
special treatment, as described in section 5d. The observational error includes a measurement error and a representativeness error. The instrument error, which is provided by the manufacturer, is about 1 dBZ, but the error in actual observations will be larger than the instrument error. Moreover, the observational error will be larger than the measurement error alone because of the representativeness error and the use of a simplified observation operator [Eq. (2)] and an adjoint model in the data assimilation system. To confirm the typical magnitude of the observational error of our system, the departure (observation vs first guess) was determined, and the standard deviation was found to be about 15 dBZ. The specific value of 15 dBZ is used in NHM-4DVAR via Eq. (2).

In addition, in a preliminary study, we performed impact tests (results not shown) in order to determine the specific value of the observational error, testing values from 5 to 20 dBZ. In the impact test using the value of 5 dBZ, the rainband did not last as long as predicted by the result of this paper, whereas in the impact test using the value of 20 dBZ, the reproduced rainband was very weak. In light of these results, we set the observational error to 10 dBZ. Observational data are transformed into “superobservations” by interpolating the data to model grids for each elevation angle. A similar treatment was used in the assimilation of Doppler velocity observations.

3. Single-observation experiments

Before our experiment using real data, we conducted single-observation data assimilation experiments without using the observation operator described in section 2d to verify whether the responses of our assimilation system were reasonable. In these experiments, the result of the numerical simulation of an isolated deep convection reported by Kawabata et al. (2007) was adopted as the first-guess field. The length of the assimilation window was set to 5 min, and we assumed that the mixing ratio of rainwater was observed at the start and end times of the assimilation window, separated by 5 min. We used two observations, one at the start and the second at the end of the assimilation window, because this configuration was determined to incorporate the tendency of the observed convection into the analysis. The case of one observation carried out only at the start is equivalent to the 3DVAR case, a configuration not suitable for our study. Furthermore, if a single observation is carried out only at the end, then the data is not sufficient for determining whether the convection is in the development or decay stage. We examined three cases by changing the amount and location of the pseudo-observation data. The assimilation was carried out on a 24 × 24 × 40 grid. The horizontal resolution was 2 km.

Figure 2 illustrates the horizontal distribution of the mixing ratio of rainwater, divergence, and horizontal wind at the end time of the assimilation window. Here, Fig. 2a shows the first-guess field, and Figs. 2b–d present the results for cases 1, 2, and 3, respectively, as explained below. The pseudo-observation data are introduced at the crisscross mark (×) in each figure, and the height $z^*$ of each figure (3.17 km in Figs. 2b,c; 1.7 km in Fig. 2d) corresponds to the height of the observation data.

Figures 3 and 4 present vertical cross-sections of the mixing ratios of $q_r$ and $q_c$ and the wind along the line A–B in Fig. 2a at the start and end times, respectively, of the assimilation window. In the first-guess field (Figs. 3a and 4a), a convective updraft is seen at the center of the model domain, and cloud water is being generated by the updraft. The surrounding low-level water vapor is also being lifted up, producing more cloud water.

a. Case 1: Enhancement

In the enhancement case, 10 g kg$^{-1}$ of rainwater was introduced at the center of the convection cell. This experiment was conducted to determine the response of the assimilation system when the observed rainwater exceeded the first guess. The result showed that the horizontal wind convergence is intensified and the convective cloud becomes bigger (Fig. 2b). The maximum value of the mixing ratio of cloud water at the start time is located around the assimilation point (Fig. 3b).

When a large amount of rainwater must exist in order to fit the observation, water vapor increases until saturation pressure causes excess water vapor to be converted to cloud water in the separation process. The large amount of cloud water seen in Fig. 3b is provided by this process. The cloud water is then converted to rainwater by a cloud microphysical process. At the end of the assimilation window (Fig. 4b), rainwater is descending and a downdraft is seen in the lower part of the convection cell.

From this result, we can expect that the underestimation of convection is adjustable by data assimilation if proper observations of rainwater are obtained near the convection.

b. Case 2: Decay

In the decay case, 0.01 g kg$^{-1}$ of rainwater was introduced at the center of the convection. This experiment was conducted to see the response of the assimilation system when the observed rainwater was less than the first guess. In this case, the convection is weakened (Fig. 2c). Not only is the rainwater directly reduced, but also the updraft around the convection is weakened (Figs. 3c and 4c). The cloud water does not decrease as much as would be expected from the small amount of rainwater.

This result shows that data assimilation of a small amount of rainwater weakens convection. We can expect
that undesirable convection in a first guess can be weakened by assimilation of small or no reflectivity.

c. Case 3: Generation

In this case, 10 g kg\(^{-1}\) of rainwater was introduced at a point away from the convection. This experiment was conducted to see the response of the assimilation system when the observed rainwater is in a convection-free area or has a large positional error. Figure 2d depicts the generation of rainwater around the observation point. At the start time of the assimilation window, rainwater is produced above and below the observation point (Fig. 3d). A large amount of cloud water is seen above the observation point, but its height is lower than in the enhancement case and little updraft is observed. At the end of the assimilation window (Fig. 4d), rainwater is

---

**Fig. 2.** Mixing ratio of rainwater (shade), divergence (contours), and wind (vectors) at the end of the assimilation window. (b)–(d) Crisscross marks (×) show the grid points where \( q_r \) observation data are put. Each figure’s height corresponds to the level of observation data. (a) First-guess field (\( z^* = 3.17 \) km), (b) enhancement case (\( z^* = 3.17 \) km), (c) decay case (\( z^* = 3.17 \) km), and (d) generation case (\( z^* = 1.70 \) km). Line A–B in (a) shows the line of the vertical cross section in Figs. 3, 4, and 5.
descending below the observation point. A downdraft has developed, and the cloud water has disappeared.

The shading in Fig. 5 illustrates the differences of mixing ratio of water vapor, the vertical wind speed, and the potential temperature between case 3 and the first guess at the end of the assimilation window, except the vectors in Fig. 5 show the wind field of the generation case. Water vapor (Fig. 5a) increases in the upper levels above the observation point, and potential temperature (Fig. 5c) increases around the observation point. These features are consistent with typical environmental fields upon initiation of convection. However, a downdraft is generated at the end of the assimilation window. Since potential temperature around the observation point is increased compared with the first-guess field, it can be inferred that this downdraft is caused by the effect of rainwater drag. It is difficult to create and maintain new convection with just the assimilation of single observations of rainwater.

The case 1 and case 2 experiments clarified that the assimilation of rainwater has a positive impact in enhancing and decaying convection. However, case 3 is more difficult to interpret because the assimilation of rainwater alone cannot maintain convection for a long time. However, the reproduced environmental field seems to be harmonious with natural convection. The assimilation of radar reflectivity would thus have a positive impact on modifying the first guess.

**Fig. 3.** Vertical cross sections of mixing ratios of $q_r$ (shade), $q_c$ (contours), and winds (vectors) along the line A–B in Fig. 2 at the start time of the assimilation window. (a)–(d) Experiment cases are the same as in Fig. 2.
4. Real observation experiment

a. Heavy rainfall event on 4–5 September 2005

An experiment based on actual observations was conducted for a heavy rainfall event that occurred around Tokyo on 4–5 September 2005. The assimilation-experiment domain, shown by the rectangle on a surface weather map (Fig. 6), covered the Kanto plain. A stationary front extended from north of Kyushu to the northern part of Japan. Typhoon 0514 (NABI) can be seen over the sea east of Okinawa, but the Kanto plain is 1500 km away and was not directly affected by the typhoon. Since there were no distinct disturbances around the Kanto plain, meteorological forcing larger than meso- \( \alpha - \beta \) scale was weak.

Convective rainfall started in the Kanto plain around noon on 4 September. A line-shaped rainband (hereafter L1) developed after 1800 Japan Standard Time (JST) on 4 September 2005 in the western part of Tokyo. The 1-h accumulated rainfall amounts from 2000 to 2300 JST are shown in Fig. 7. The rainband, which was oriented north to south-southwest, gradually increased in intensity during this period.

Although, L1 was very small, about 100 km long and 15 km wide, it caused heavy rainfall in a narrow area in the southwestern part of the Tokyo metropolitan area. The maximum total rainfall [2000–2400 JST; observed by Automated Meteorological Data Acquisition System (AMeDAS) of JMA; average resolution about 17 km] among rain gauge stations was 113 mm at Nerima,
whereas 10 observation points of local municipalities recorded over 200 mm. A surprising 264 mm was observed at Shimoigusa over 9 h from 2000 to 0500 JST on 4–5 September (Fig. 8). The Myoshoji River, a branch of the Arakawa River, overflowed and more than 5000 houses were flooded up to their floorboards.

b. Observational data

Observations used in the experiment consisted of RW and reflectivity observed by the HANEDA and NARITA airport radars, GPS-PWV observed by GPS Earth Observation Network System (GEONET) of the Geographical Survey Institute of Japan, vertical profiles of horizontal wind observed by wind profilers of the JMA, and surface winds and temperature observed by AMeDAS and at UMIHOTARU (an observation site on Tokyo Bay monitored by the Ministry of the Environment of Japan). The locations of these sites are plotted in Fig. 9c.

The assimilation methods used in these experiments, except for radar reflectivity, are the same as those used by Kawabata et al. (2007). Radar data (RW and reflectivity) were assimilated at 1-min intervals for data of each level of elevation. GPS-PWV, processed by Shoji (2009), was assimilated at 5-min intervals, and surface and wind profiler data were assimilated at 10-min intervals.

The RW can be expressed with three-dimensional wind velocity components and the mean terminal velocity of water substances, (i.e., hail, graupel, snow, and rain droplets) in the numerical model. Although reflectivity was observed, there are uncertainties in the conversion from reflectivity to the mixing ratio of water substances. The melting of snowflakes in the melting layer also causes uncertainty. We assumed that the uncertainty in vertical velocity was up to 10 m $s^{-1}$ and that the accuracy of the radial wind was around 1 m $s^{-1}$. When the uncertainty in the radial wind component of vertical velocity ($10 \text{ m s}^{-1}$) is less than the accuracy of radar beam velocity ($1 \text{ m s}^{-1}$), the elevation angle is $5.7^\circ$. Therefore, the RW data for an elevation angle of less than $5.4^\circ$ were used and was treated as horizontal winds instead of as radial winds (see Seko et al. 2004) in this study. Data for a higher elevation angle will be used in our future studies.
Likewise, the reflectivity observations at high elevations of more than 5.4° were ignored because the high-elevation data are confined to small areas near radar sites. In a preliminary study, we confirmed that the impact of high-elevation-angle data on the forecast results is small.

c. Design of the assimilation experiment

First, a simulation by NHM with 5-km grid spacing (5km-NHM; Fig. 9a) was conducted from 0900 to 2400 JST on 4 September 2005 using the Meso-4DVAR analysis for the initial and boundary conditions. Second, a simulation by NHM with 2-km grid spacing (2km-NHM-L) was performed in a domain of about 400 km² (Fig. 9b) from 0900 to 2400 JST using the 5km-NHM forecast result for the initial and boundary conditions. Finally, an NHM simulation with 2-km grid spacing (2km-NHM-S) was conducted in the assimilation domain (244 km²; Fig. 9c) from 1000 to 1200 JST using the 2km-NHM-L result for the initial and boundary conditions.

The 2km-NHM-S result was then used as the first-guess field at 1200 JST in the first assimilation window. A first guess of the lateral boundary conditions of NHM-4DVAR was given by the 2km-NHM-L forecast. As mentioned in section 2a, the same background covariance matrices can be used for $B$ and $B_{lbc}$ with the same simulations in first guesses of the initial and lateral boundary conditions.

To minimize the effects of a coarse model, a forecast based on downscaled initial conditions should be started several hours before the short-range forecast experiment and forecast–analysis cycles have been made. Since a forecast based on downscaled initial conditions failed badly in the forecast of convection (see Figs. 12a,d), it is better to set the start time of the cycle during a period of calm weather.

For the above reasons, 8 analysis–forecast cycles were performed from 1200 to 2000 JST with 1-h assimilation windows using the NHM-4DVAR. In these spinup windows, considering computation costs, the number of iterations for minimizing $J$ was limited to 10. Since there were few convection areas during the first half of this period, only observations of RW, GPS-PWV, surface observations, and wind profiler data were assimilated until 1800 JST. Radar reflectivity was added between 1800 and 2000 JST.

After the spinup process, two assimilations with 30-min windows were performed, from 2000 to 2030 JST and from 2030 to 2100 JST. In these windows, owing to computational restrictions, about 50 iterations for minimization were performed. A 1-h free forecast was conducted from 2030 to 2130 JST, using the NHM-4DVAR analysis at 2030 JST provided by the second window as the initial conditions over the same domain with the analysis. In fact, in this configuration, the pure forecast, unconstrained by observation data, is 30-min...
Fig. 9. Nesting conditions of the assimilation experiment. The domain of (a) 5km-NHM and the model topography, (b) 2km-NHM, and (c) NHM-4DVAR and distribution of observation stations used in the experiment. Crisscrosses denote AMeDAS and the UMIHOTARU surface observation stations, triangles denote GPS observations, black circles denote Doppler radars, and rectangles denote wind profilers. Black shade denotes orography.
5. Results

a. Statistical results

Time series of the threat scores of the reflectivity field on a 0.7° elevation plane are shown in Fig. 11. In the ctl case, the scores for both the 15- and 30-dBZ thresholds were very low, below 0.1, at every forecast time. In the wo-ref case, the 15-dBZ score was high, above 0.25, at the first 15 min, but low, below 0.2, at forecast time after 15 min. The 30-dBZ score was low at the forecast start time, but became relatively high, about 2.0, by the forecast end time. In the ref case, the 15-dBZ score was very high, over 0.6, at the forecast start time, but became relatively low, about 0.3, by the forecast end time. The 30-dBZ score was also notably high, about 0.5, during the first 30 min, but became low, about 0.25, by the forecast end time. These results indicate that the assimilation of indirect observations (e.g., GPS-PWV, RW, and surface observations) can modify the reflectivity field, and that these observations substantially affect weak convection, but they do not affect strong convection to a large extent. Furthermore, the direct assimilation of reflectivity can substantially modify the forecast of strong convection, especially at early forecast times.

b. Reflectivity

The distribution of reflectivity observed by the Haneda airport radar and the corresponding simulated reflectivity from the ref, wo-ref, and ctl experiments are shown in Fig. 12. The ctl experiment could not simulate the line-shaped rainband, convective cells are only sparsely distributed along the foot of the mountainous region west of the Kanto plain (Fig. 12d).

In wo-ref (Fig. 12c), L1 is reproduced, though with less intensity than the observation. This result was mainly attained by the assimilation of GPS-PWV, RW, and the wind profiler data. Surface observations include information about surface-wind circulation, but in this case their impact on the reproduction of L1 was small. In the experiment, weak convective cells, not found in the observation, also developed around Tokyo bay.

The wo-ref reflectivity field at 2030 JST seems noisy in (Fig. 12c), possibly because NHM-4DVAR adopts a control variable related to rainwater. Since there was no rainwater observation in wo-ref, the minimizer cannot determine the optimal direction of the rainwater state.

In ref (Fig. 12b), L1 is well reproduced, with a shape and intensity consistent with the observation (Fig. 12a). At 2130 JST, false convective cells appear over Tokyo bay, but their intensity is weak. Compared with wo-ref, assimilation of the reflectivity intensified L1 and assimilation of 0 dBZ (see section 6d) suppressed the false convection.

Vertical cross sections of the differences between ref and wo-ref are presented in Fig. 13. In the ref experiment, the convective updraft was intensified around L1, and water vapor and potential temperature also increased beneath L1. Thus, the dynamical and thermodynamical fields were modified by the assimilation of observed radar reflectivity, such that the circumstances became suitable for intensification of the rainband.
FIG. 12. Radar reflectivity on 0.7° elevation plane from 2030 to 2130 JST. (a) Observation from the Hanaeda airport radar. (b) Assimilation and forecast results with radar reflectivity. (c) As in (b), but without the assimilation of reflectivity. (d) First-guess field. Black shows orography.
c. Evolution of the rainband

In this section, we compare the evolution of L1 as simulated in the ref experiment with the observation. At 2030 JST, the start time of the assimilation window, the three convection, L1, L2, and L3, recognized in the observation (Fig. 12a) are roughly reproduced in the ref experiment (Fig. 12b). Areas of weak reflectivity are more widespread in ref, however, than in the observation. Rain water in these areas is produced by horizontal correlation of the background errors. Since these are neither dynamically nor thermodynamically related to the environmental fields, they soon evaporate and vanish.

At 2050 JST, L1 is well reproduced in the ref experiment in terms of intensity, size, and location; L2 has shifted northwestward; and L3 and other precipitation areas have disappeared. Around this time, the observed L1 becomes enhanced (Fig. 12a), and a similar tendency is reproduced in the experiment.

At 2111 JST, L1 has moved northward while maintaining its intensity. Convection in the ref experiment is slightly more intense than the observed convection. In the experiment, a few false convective cells appear over Tokyo Bay, but they are weak.

At 2130 JST, L1 remains in the same region because new convection is generated at its southern tip, although each convective cell in L1 has moved northward. Thus, L1 seems to have a back-building formation mechanism. The intensity of L1 in ref is similar to the observation, but in ref, L1 is entering a decaying stage. After 2130 JST, the intensity and shape of L1 could not be maintained in the extended forecast.

The horizontal distributions of the mixing ratio of $q_r$ and the winds at $z^* = 225$ m in ref are shown in Fig. 14, vertical cross sections of the mixing ratio of $q_r$, winds, and the equivalent potential temperature of ref below 5 km along line A–B (see Fig. 14d) are shown in Fig. 15. From 2050 to 2110 JST, easterly winds blowing toward L1 are seen in the eastern part of L1 in the assimilation window, and northerly winds west of L1 weaken. As a result, the low-level convergence around L1 decreases, ultimately vanishing at 2130 JST (Fig. 14d). However, the distribution of equivalent potential temperature does not change during this period. Thus, the decay of L1 is caused by a change in the dynamical rather than the thermodynamical environment. Hence, we infer that a northerly wind would have to be reproduced to maintain the low-level convergence, and the rainband, after 2130 JST.

Regarding the thermodynamical field, a warm, wet easterly inflow with an equivalent potential temperature of more than 350 K can be recognized in the lower part of the rainband (at about $z^* = 1$ km) at 2030 JST (Fig. 15, top left). A small cold pool with an equivalent potential temperature of less than 336 K exists at the surface to the west of the rainband. This cold pool contributes to the maintenance of the rainband. At 2050 JST (Fig. 15, top right), a strong updraft and associated heating are identified in the rainband between 0.5 and 3 km. This inflow and the updraft are important for the rainband maintenance, but they are not sufficient to sustain it for a long time. The updraft weakens after 2110 JST (Fig. 15, bottom).

d. Assimilation of 0-dBZ information

In this study, we regarded only reflectivity values greater than 10 dBZ as proper observations. Initially, we neglected weak reflectivity of less than 10 dBZ in order to avoid inappropriate reflectivity caused by materials
other than rainwater. However, as discussed in section 3, such very weak reflectivity data convey the information that rainwater is absent or that the amount is very small in the observed area.

To incorporate this information, we assimilated weak reflectivity of less than 10 dBZ as an observation of 0 dBZ. The observational error was set to 30 dBZ for this observation (3 times the normal error) because 0 dBZ is not genuine observational data. In addition, we assimilated 0 dBZ only in regions where the reflectivity in the first-guess field exceeded 10 dBZ.

The impact test results are presented in Fig. 16. Reflectivity observed by the Haneda airport radar is presented in Fig. 16a, and the first-guess field, the assimilation result of ref with 0 dBZ, and the result without 0 dBZ are illustrated in Figs. 16b–d), respectively.

There is a false convective area on the east side of Tokyo bay in the first-guess field (circled in Fig. 16b), but not in the observation (Fig. 16a). After the assimilation of reflectivity, the reproduction of L1 is improved (Figs. 16c,d), but the false cells are still seen in the assimilation result without 0 dBZ. It is clear that the assimilation of
0 dBZ suppressed the false convection. This suppression is similar to the case 2 result described in section 3. The area of 0 dBZ is very much wider than other areas, and the cost function tends to be minimized in a direction without convection. Therefore, use of 0 dBZ should be limited, as pointed out by Koizumi et al. (2005).

6. Discussion

a. Sustaining low-level convergence

As described in section 4, the line-shaped rainband L1 is reproduced in the assimilation and in the corresponding extended forecast. Horizontal wind convergence along the reproduced rainband was seen below a height of 2000 m. However, this convergence weakened at 2110 JST and vanished after 2140 JST. This convergence appears to have maintained the rainband. Figure 17 illustrates the differences in horizontal wind and the mixing ratio of rainwater between ref and wo-ref at 225-m height at the end time of the first assimilation window. The mixing ratio of rainwater in the main rainband L1 increases, and the easterly wind is strengthened in the east of L1. The southerly wind is also strengthened in the inflow region of the rainband. This change in the wind field is caused by the assimilation of reflectivity. Therefore, it is clear that the assimilation of reflectivity changes the low-level wind circulation and contributes to producing the line-shaped rainband.

However, the northerly winds in west of L1 did not continue for a long time, and L1 began to weaken after 2110 JST. This problem was likely caused by insufficient retrieval of the thermodynamical field, as mentioned in section 5c. To improve this situation, more observations [e.g., wind profiler, radiosonde, and the Aircraft Communications Addressing and Reporting System (ACARS)] are needed in these areas.

To extend the influence of the assimilation of reflectivity and RW, it seems necessary to lengthen the assimilation windows, but because of the strong nonlinearity we could not make them longer than 30 min. A test case with a 1-h assimilation window did not reproduce the intensity of L1 even in the assimilation window (not shown).

b. Incremental method with a tangent linear model

As described in section 2a, there is an inconsistency between the forward and adjoint models in NHM-4DVAR. Since this inconsistency caused problems, as described in section 6a, we attempted to use an incremental method (Courtier et al. 1994), which uses tangent linear and adjoint models in an inner loop and the nonlinear model in an outer loop. In this system, the
outer loop trajectory is updated every 10 inner loop iterations. An advantage of this method is that the convergence of the minimization process is improved by the consistency between the forward and adjoint models, and the nonlinearity is taken into consideration by the frequent updating of the outer loop trajectory.

An experiment with 10-min windows was carried out. Figure 18 compares the reflectivity fields obtained using the incremental system and the original system. Reflectivity of more than 25 dBZ can be recognized widely in the reproduced rainband in the original system, whereas, reflectivity is less than 25 dBZ in the incremental system. The incremental system using the simplified tangent linear model thus could not reproduce the strong rainband. A possible reason for this failure is that the strong nonlinearity of the mesoscale convective rainband cannot be reproduced by the tangent linear system. Figure 19 shows the time sequence of cost function $J$ in the two systems. The cost
function obtained by the incremental system is sluggish during the iteration process and shows a jump when the outer loop trajectory is updated. Most of this jump is in the observation term, because that term is recalculated with a new trajectory after the updating. Given these results, we did not adopt the incremental system in our study.

7. Summary and conclusions

The cloud-resolving nonhydrostatic four-dimensional variational assimilation system (NHM-4DVAR; Kawabata et al. 2007) was evolved in order to directly assimilate radar reflectivity. Modifications included the development of an adjoint model of the warm rain process, the extension of control variables, and the development of an observation operator for radar reflectivity using the $Z-Q_r$ relationship.

Single-observation assimilation experiments were conducted to observe the responses of the modified NHM-4DVAR. A case of isolated deep convection was adopted as a first-guess field. Three experiments were performed using pseudo-observations of rainwater at the start and end times of a 5-min assimilation window for specific enhancement, decay, and generation situations. When $10 \text{ g kg}^{-1}$ of rainwater was introduced at the center of the convection, the convergence of horizontal wind intensified and the convective cloud became larger. When $0.01 \text{ g kg}^{-1}$ of rainwater was introduced at the center of the convection, the convection was weakened. This result demonstrates that the underestimation of convection can be corrected by data assimilation if appropriate rainwater observations are obtained near the convection, and that assimilating a small amount of reflectivity or none at all suppresses undesirable convection in the first guess. When $10 \text{ g kg}^{-1}$ of rainwater was introduced at a distance from the convection, rainwater was produced above and below the observation point at the start.
time of the assimilation window, but by the end of the assimilation window, downdrafts had developed and the rainwater descended below the observation point. It is likely to be difficult to create and maintain new convection by assimilation of only single observations of rainwater.

We then conducted an assimilation experiment with a horizontal resolution of 2 km based on actual observations of a local heavy rainfall in the Tokyo metropolitan area. Precipitable water vapor derived from GPS data were assimilated at 5-min intervals during 30-min assimilation windows, and surface data and wind profiler data were assimilated at 10-min intervals. The Doppler radial wind and radar reflectivity were assimilated at 1-min intervals for elevation angles of less than 5.4°.

A line-shaped rainband was well reproduced by the assimilation of reflectivity, with a shape and intensity consistent with the observation. Compared with the woref experiment, the ref experiment intensified an originally weak rainband and suppressed false convection. False convection was controlled by assimilating 0 dBZ, and the results were consistent with the results of the single-observation assimilation experiments.

In the extended forecast, the simulated line-shaped rainband gradually decayed after 1 h. Sustaining the low-level convergence produced by northerly winds in the western part of the rainband seems to be key to maintaining the convective system. The RW and reflectivity data can be obtained only in areas where rainwater actually exists. The environmental field should be further modified to enhance the predictability of the observed convective system. This problem corresponds to the inverse problem, in which the environmental field is estimated from the observations. Since this relationship is nonlinear, solving the problem is not easy.

To predict rainbands, it is important to estimate their environmental field. However, we have not carried out sufficient observations outside the rainband, although we have carried out many observations within the rainband. Therefore, we have to estimate the environmental field (thermodynamical and dynamical fields) from the observation data from within the rainband. Reflectivity indicates the existence of rain drops, but the relationships between raindrops in clouds and environmental thermodynamical and dynamical fields are not linear. Hence, it is difficult to solve the inverse problem.

More information is needed to improve short-range forecasts of local heavy rain events. We are currently conducting data assimilation studies using GPS slant-delay data and radioacoustic sounding-system data. To enhance the predictability of heavy rainfall, it will be necessary to develop techniques to overcome the nonlinearity. In addition, the introduction of a flow-dependent background error and the consideration of model error are future subjects for study.

Acknowledgments. This study was partly supported by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) through a Grant-in-Aid for Scientific Research (21244074) “Study of Advanced Data Assimilation and Cloud Resolving Ensemble Technique for Prediction of Local Heavy Rainfall.” The authors would like to express their appreciation to Mr. Yoshinori Shoji, senior researcher of MRI, who provided GPS-PWV data, and the Ministry of the Environment which provided the surface observation at UMIHOTARU. We also thank Dr. Tadashi Tsuyuki of MRI for giving us useful comments and Mr. Hisaki Eito, senior researcher of MRI, for advising us about the observation operator of reflectivity. Finally, we are grateful to the reviewers and editors.

REFERENCES


