Tellus
SERIES A DYNAMIC METEOROLOGY AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tellus (2011), 63A, 402–414

TELLUS

Data assimilation experiments of precipitable water vapour using the LETKF system: intense rainfall event over Japan 28 July 2008

By HIROMU SEKO[1]*, TAKEMASA MIYOSHI[2], YOSHINORI SHOJI[1] and KAZUO SAITO[1],
[1]*Meteorological Research Institute, 1-1 Nagamine, Tsukuba, Ibaraki 305-0052, Japan;* [2]*University of Maryland, College Park, MD 20742, USA*

(Manuscript received 26 April 2010; in final form 21 December 2010)

ABSTRACT

We performed data assimilation experiments using data from an intense rainfall event in western Japan on 28 July 2008 with the Local Ensemble Transform Kalman Filter (LETKF) and analysed fields of the mesoscale four-dimensional variational assimilation (4DVAR) system of the Japan Meteorological Agency (JMA). We supplemented conventional observation data with precipitable water vapour (PWV) data derived from the Global Positioning System (GPS) Earth Observation Network of the Geospatial Information Authority of Japan. Because the LETKF system assimilates fewer data than the 4DVAR system, ensemble mean fields of the LETKF cycle experiment were replaced with analysed fields from the 4DVAR system each day at 12 UTC. PWV values were converted to relative humidity profiles for assimilation by the LETKF. The addition of PWV data tended to increase low-level water vapour and improve the precipitation forecast. We attempted to reproduce the intense rainfall band using downscale forecast experiments with the JMA non-hydrostatic model (JMANHM) with grid spacings of 5 and 1.6 km. The experiment with 5-km resolution generated a rainfall band in western Japan that was not reproduced using conventional data, although the rainfall was smaller than observations. The experiment with 1.6-km resolution faithfully reproduced the observed band of intense rainfall.

1. Introduction

Heavy rainfall events are more likely when the Baiu front stays over Japan or when a typhoon exists in southwestern Japan. Previous studies have attempted deterministic forecasts to reproduce heavy rainfall (e.g. Seko and Nakamura, 2005), but efforts have suffered owing to insufficient accuracy of the initial conditions, even with excellent numerical models and model resolutions that were fine enough for modelling severe convective cells. To improve knowledge of initial conditions for reproducing heavy rainfalls, several previous researchers have conducted experiments using variational data assimilation systems. For example, Seko et al. (2004) and Kawabata et al. (2007) used four-dimensional variational data assimilation systems to assimilate radial wind velocities from Doppler radar and water vapour data [precipitable water vapour (PWV) and slant water vapour] from the Global Positioning System (GPS), and they noted that improvements in low-level wind and water vapour fields were essential for accurate forecasts of heavy rainfalls. The usefulness of ground-based GPS data has been demonstrated in European case studies. For instance, Yan et al. (2009) showed the benefit of assimilating high-resolution GPS zenith tropospheric delay data within a high-resolution numerical weather prediction (NWP) system by using data from Intensive Observing Period (IOP) 9 of the Convective and Orographically induced Precipitation Study (COPS) campaign. Bauer et al. (2011) also showed the usefulness of slant total delay data using the MM5 4-D data assimilation system.

Other studies have used the ensemble Kalman filter (EnKF) for mesoscale applications and obtained promising results (e.g. Snyder and Zhang 2003; Dowell et al., 2004; Zhang et al., 2004, 2006; Tong and Xue 2005; Xue et al., 2006; Meng and Zhang 2008). EnKF is advantageous in its relative ease of implementation and the direct representation of analysis and forecast uncertainties by ensembles. In the case of heavy rainfall applications, some ensemble members may reproduce heavy rainfall events even if a deterministic forecast fails to reproduce them. Moreover, probabilities of heavy rainfalls can be obtained using ensemble members.

Water vapour plays an important role in convective initiation and development, as well as dynamical fields obtained by Doppler radar. Assimilation experiments of GPS-derived

*Corresponding author.
e-mail: hseko@mri-jma.go.jp
DOI: 10.1111/j.1600-0870.2010.00508.x

PWV data have been performed with variational data assimilation methods (e.g. Seko et al., 2004; Kawabata et al., 2007; M. Zhang et al., 2007; Shoji et al., 2009). However, to the best of our knowledge, there have been no studies published on the assimilation of PWV data with EnKF systems.

In this study, the data assimilation system using the Local Ensemble Transform Kalman Filter (LETKF, Hunt et al., 2007) based on the Japan Meteorological Agency (JMA) Non-hydrostatic Model (JMANHM, Saito et al., 2006), known as the NHM-LETKF (Miyoshi and Aranami 2006), was modified to assimilate PWV data from the intense rainfall event of 28 July 2008 in western Japan. PWV data used in this study were provided by the GPS Earth Observation Network (GEONET) of the Geospatial Information Authority of Japan (GSI), which has a mean distance between GPS receivers of about 25 km, among the densest in the world.

In this paper, Section 2 introduces the NHM-LETKF and Section 3 discusses the intense rainfall event of 28 July 2008. Sections 4 and 5 describe the experimental design and the assimilation method of PWV, respectively. Section 6 shows the analysis results of NHM-LETKF, including the impact of PWV on the intense rainfall and Section 7 is conclusion of this study.

2. NHM-LETKF

Hunt et al. (2007) derived the LETKF equations as a kind of ensemble of square-root filters (EnSRF, Tippett et al., 2003). Following the notation of Miyoshi (2010), the analytical equations for the LETKF are written as

$$\tilde{\mathbf{P}}^a = [(m-1)\mathbf{I}/\rho + (\delta\mathbf{Y})^T\mathbf{R}^{-1}\delta\mathbf{Y}]^{-1} = \mathbf{U}\mathbf{D}^{-1}\mathbf{U}^T, \quad (1)$$

$$\mathbf{X}^a = \bar{x}^f e + \delta\mathbf{X}^f\left(\tilde{\mathbf{P}}^a(\delta\mathbf{Y})^T\mathbf{R}^{-1}(y^o - \overline{H(x^f)})e + \sqrt{m-1}\mathbf{U}\mathbf{D}^{-1/2}\mathbf{U}^T\right), \quad (2)$$

where ρ is the covariance inflation factor and e is an m-dimensional row vector $(1, \ldots, 1)$. A multiplicative spread inflation of 10% was employed in this study. $\mathbf{U}$ and $\mathbf{D}$ are obtained by the following eigenvalue decomposition:

$$(m-1)\mathbf{I}/\rho + (\delta\mathbf{Y})^T\mathbf{R}^{-1}\delta\mathbf{Y} = \mathbf{U}\mathbf{D}\mathbf{U}^T. \quad (3)$$

More detailed derivation of equations has been provided by Miyoshi (2010).

The NHM-LETKF system used in this study is based on the JMA's operational non-hydrostatic mesoscale model JMANHM. It contains the following variables: horizontal wind components (u, v), vertical wind component (w), temperature (T), water vapour (q_v), surface pressure (p_s), 3-D atmospheric pressure (p) and various water quantities such as rain (q_r), cloud liquid water (q_c), cloud ice (q_{ci}), snow (q_s) and graupel (q_g). Miyoshi and Aranami (2006) investigated the performance of mesoscale ensemble forecasts with microphysical processes by using JMANHM with a grid interval of 5 km. As the first test of the NHM-LETKF system, they performed a perfect model observing system simulation experiment using synthetic observation data every 10 min at grid points 10 km apart. Although they obtained excellent results, the simulated observations were much denser than reality.

The NHM-LETKF was modified for ensemble experiments of the World Weather Research Program Beijing 2008 Olympics Forecast Demonstration Project/Research and Development Project (WWRP B08FDP/RDP), which was the first international intercomparison experiment of mesoscale ensemble forecasts. The Meteorological Research Institute (MRI) participated in the B08RDP project (Saito et al., 2010) and tested NHM-LETKF as one of the methods to generate initial perturbations. The implementations and amendments of the NHM-LETKF system for the B08RDP project included the following (Seko, 2010): (1) accurate treatment of air density, (2) implementation of vertical hybrid coordinates, (3) saturation adjustment of the initial field, (4) removal of local patches (Miyoshi et al., 2007), (5) cluster analysis procedure for choosing 10 members out of 20 members and (6) lateral boundary perturbations in forecast-analysis cycles.

An important advantage of the LETKF is its parallel efficiency, solving eq. (2) for each grid point independently; the idea is based on the local ensemble Kalman filter (Ott et al., 2004). The error covariance is localized when a subset of the observational data is chosen to be assimilated. However, this localization is the Heaviside function and causes discontinuities in the analysis. To avoid these discontinuities, smooth localization weights are applied (Miyoshi et al., 2007). In this process of 'observation localization' (Hunt et al., 2007), the observation error variance σ is inflated according to the distance from the analysed grid point. This modification was implemented for NHM-LETKF (Fujita et al., 2009).

In B08FDP/RDP, 6-h forecast-analysis cycles of 20 ensemble members were performed with JMANHM (40 km L40) and NHM-LETKF. Surface and upper sounding data that were used in JMA's operational global analysis and mesoscale analysis were assimilated by NHM-LETKF. Assimilation results were recentred around the analysed fields obtained by a mesoscale four-dimensional variational assimilation system of JMA (Meso-4DVAR) every day at 12 UTC; that is, the ensemble mean was replaced with Meso-4DVAR analysis results once a day. The covariance estimated by LETKF was not used in the estimation of analysis fields in the Meso-4DVAR system. The boundary-condition perturbations of NHM-LETKF were produced by normalizing the perturbation of the operational 1-week ensemble forecast. 10 members out of the 20-member analyses were chosen by cluster analysis for downscale forecasts. A detailed comparison of NHM-LETKF and other methods as the initial perturbation generator in B08RDP has been given by Saito et al. (2011).

3. Intense rainfall event of 28 July 2008

This study focuses on an intense rainfall event over western Japan on 28 July 2008. Figure 1a shows sea-level pressure at 00 UTC on 28 July 2008, produced from the JMA operational mesoscale analysis. The position of the front in Fig. 1a was traced from the weather map at the same time. The Baiu front, a stationary rainy front that stays over Japan in early summer, extended along the north side of western Japan. There was a typhoon near Taiwan and a high-pressure system off the south coast of western Japan. Southeasterly flow, enhanced by the relatively large pressure gradient between the typhoon and the high-pressure system, was dominant over the East China Sea. This low-level airflow moved northwards over the East China Sea, then turned eastwards near Korea (indicated by the thick arrow in Fig. 1b). This airflow supplied high equivalent potential temperature air to the rainfall systems that developed along the Baiu front. At the 500-hPa level, a cold air mass with a temperature lower than –6 °C (shaded region in Fig. 1c), extended over western Japan. This combination of low-level high equivalent potential temperature and mid-level low temperature indicates that the atmosphere over western Japan was very unstable.

Figure 2 shows rainfall intensities observed by JMA operational radar. A rainfall region associated with the Baiu front (indicated by **A** in Fig. 2a) was generated along the northern part of western Japan by 01 UTC 28 July. This rainfall region organized into a linear system that remained stationary through 05 UTC. South of this rainfall region, several convective cells were generated at 03 UTC (indicated by **B** in Fig. 2b). Warm southwesterly flow prevailed on the southern side of region **B**, and the cold air pool produced by the rainfall in region **A** did not reach the convective cells. Although the western part of the convective cells was generated over mountain slopes, the cells in the eastern part were not related to mountains (Fig. 2e).

Figure 3 depicts the increase of PWV after 22 UTC 27 July. The region of increased PWV expanded from the western side of the convective cells of region **B**. At 03 UTC 28 July, the humid region had reached where the convective cells of intense rainfall were being generated. These observations suggest that the convective cells in region **B** were generated mainly by the convergence of warm humid air caused by southwesterly flow. By 05 UTC these convective cells were organized into an intense convective band, which caused 1-h rainfall amounts of 23.5 mm at 06 UTC at Kobe (black circle in Fig. 2) and 53.0 mm at Sanda 20 km north of Kobe. The rainfall system, including these convective bands, was the focus of the data assimilation experiments in this study.

This intense rainfall system has been the subject of previous studies. In the most notable of these, Shoji et al. (2009) assimilated PWV data from Japan and East Asia using the Meso-4DVAR system of JMA, to show their impacts on rainfall forecasts. They found that the rainfall band was not reproduced by assimilation of operational data such as upper sounding, surface, wind profiler and Doppler radar radial wind (see Fig. 5 in Shoji et al., 2009). When PWV data were added to these, water vapour over the Sea of Japan was increased in the model, which helped reproduce the rainfall system in region **A**. However, the water vapour over western Japan was not increased enough to reproduce the convective band in region **B**. Their results indicate that water vapour is a key factor for the intense rainfall band in this case.

4. Design of the experiments

The LETKF system used in this study is nearly the same as the one used in the B08FDP/RDP project (Seko, 2010), differing from it in the horizontal resolution, boundary perturbations in

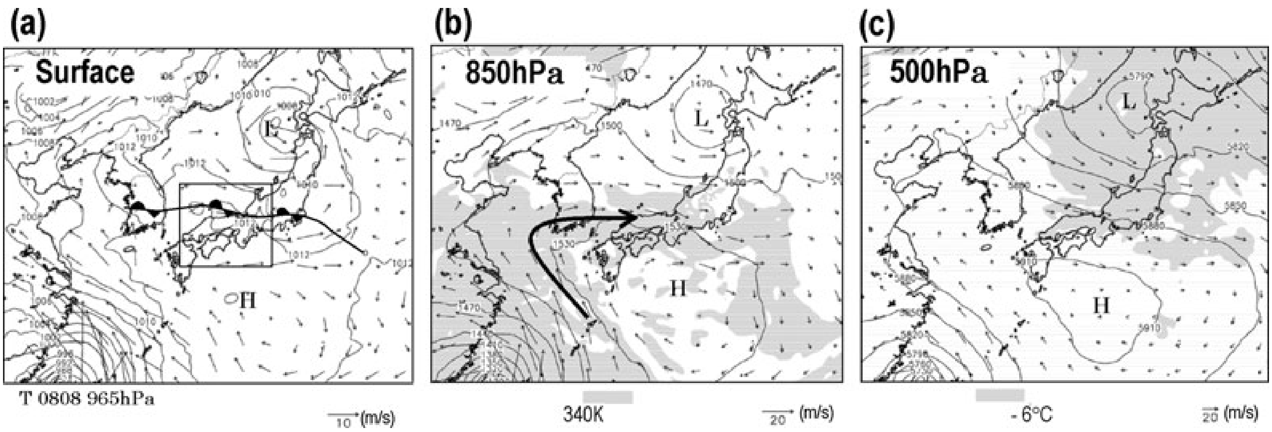

Fig. 1. Synoptic conditions over Japanese Islands and adjacent Asia at 00 UTC on 28 July 2008. (a) Sea level pressure (contours in hPa) and surface wind vectors, produced from the operational mesoscale analysis data of JMA. Rectangle shows outline of Fig. 2. (b) Height (contours, m), equivalent potential temperature (shade, >340 K), and wind vectors at 850 hPa. (c) Height (contours, m), temperature (shade, <–6 °C), and wind vectors at 500 hPa. Thick arrow in (b) indicates the airflow into the rainfall system.

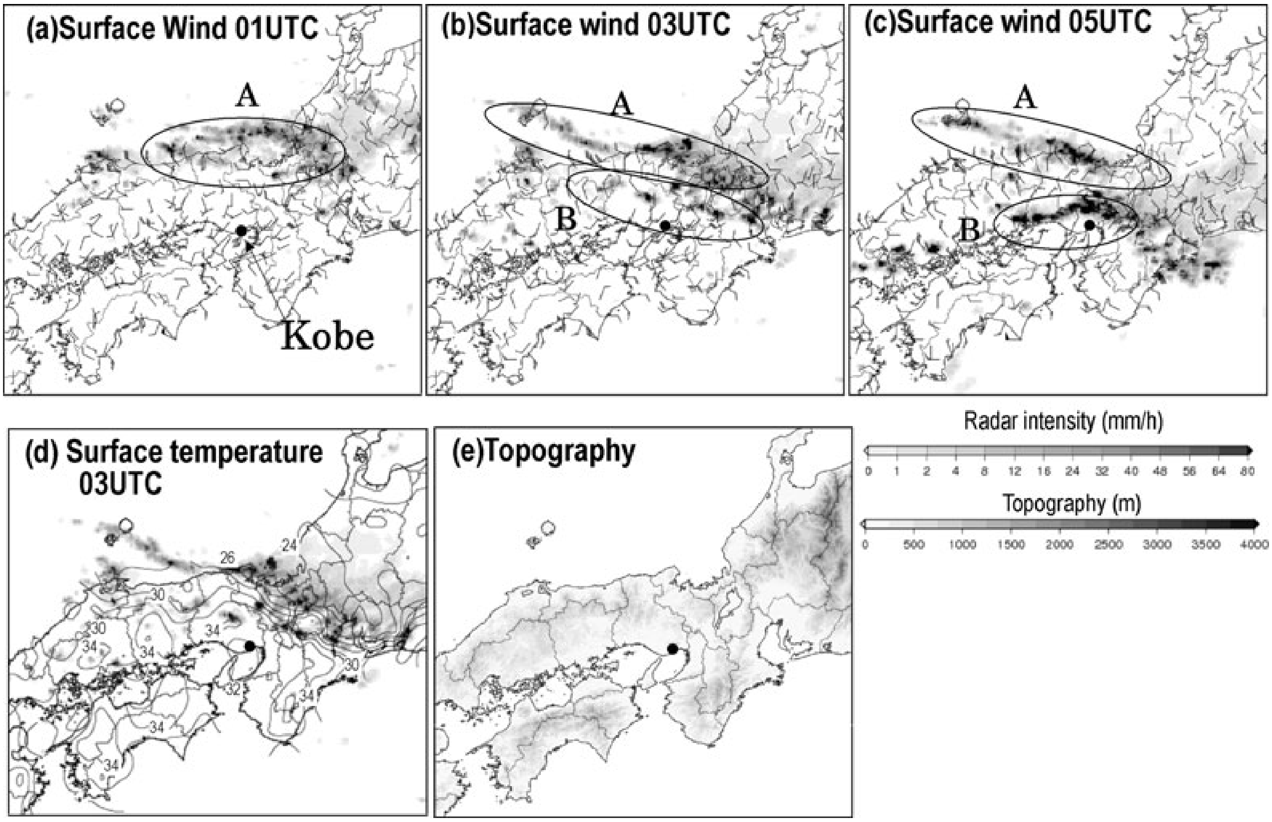

Fig. 2. Maps of western Japan on 28 July 2008. (a)–(c) Surface wind and rainfall intensity observed by JMA operational radar at 01 UTC, 03 UTC and 05 UTC. Ellipse **A** outlines a rainfall region in the northern part of western Japan, and ellipse **B** outlines a band of intense rainfall to its south. Position of Kobe is indicated by black circle. (d) Surface temperature and rainfall intensity at 03 UTC 28 July. Influence of altitude on temperature has been corrected. (e) Topography of western Japan.

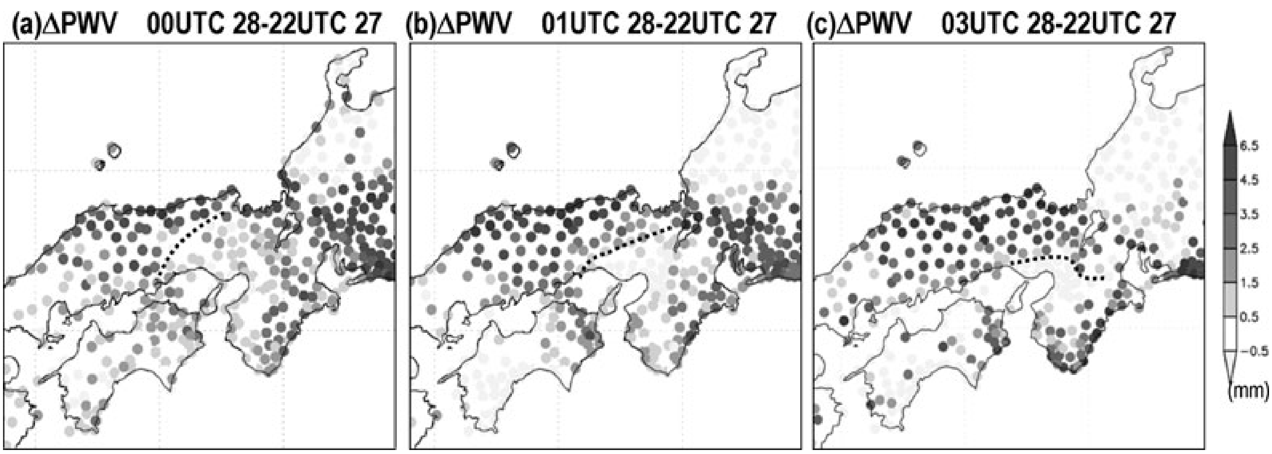

Fig. 3. Maps of western Japan showing incremental change of PWV from that of 22 UTC 27 July (dots) at (a) 00 UTC 28 July, (b) 01 UTC 28 July and (c) 03 UTC 28 July. Dotted lines indicate the eastern edge of the region of increased PWV.

the forecast-analysis cycles and ensemble size of the downscale forecasts.

Figure 4 is a flowchart of our forecast-analysis cycles and subsequent forecasts. We performed two kinds of assimilation experiments to show the impact of PWV data. The control experiment (CNTL) assimilated only conventional data, and the test experiment (PWV) assimilated conventional data and PWV data. 6-h forecast-analysis cycles of 20 members were performed with JMANHM and NHM-LETKF beginning at 00 UTC on 26 July 2008. The horizontal resolution of the LETKF and forecast

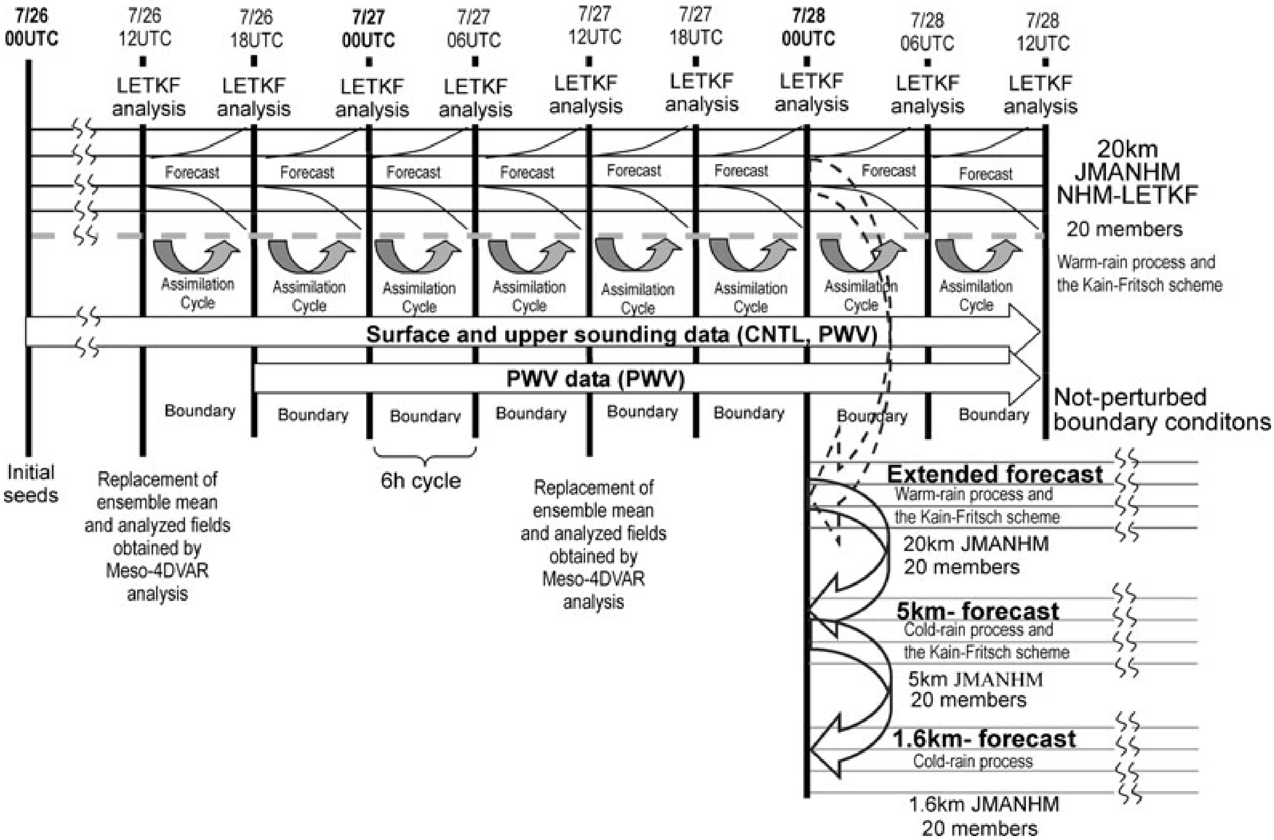

Fig. 4. Flowchart of the experiments in this study. See text for explanation.

model in the forecast-analysis cycles was set at 20 km, which is smaller than that of the B08FDP/RDP project. The number of vertical layers was 40. The lowest level was 20 m above ground level, and the depth of the layers increased from 40 to 1180 m as their height increased. The domain of the NHM-LETKF system is the same as the operational mesoscale model of JMA, which covers the region from Taiwan to the Aleutian Islands. To accelerate computations of the NHM-LETKF system, a warm-rain process and the Kain-Fritsch convective parameterization scheme were adopted in the forecast model of NHM-LETKF, because microphysical processes are not very important at 20-km resolution.

Surface and upper sounding data that were used in the JMA operational global analysis and mesoscale analysis were assimilated by NHM-LETKF. However, these data were not sufficient to produce accurate analyses owing to the lack of satellite data and rainfall analysis data used in the mesoscale analysis. To retain the accuracy of NHM-LETKF during the forecast-analysis cycles, the analysis ensemble mean was replaced with the result of the Meso-4DVAR analysis each day at 12 UTC. As mentioned in the explanation of B08FDP/RDP, the covariance estimated by LETKF was not used in the Meso-4DVAR system.

Initial conditions were taken from those at 00 UTC 7 July, 20 d before the initial time of the forecast-analysis cycle. To check the convergence of the analysed fields, we examined the time-series of ensemble spread in the inner domain, at least 30 grid spaces inside the lateral boundary (Fig. 5a). The temporal change of spread became smaller 1.5 d after the initial time of the forecast-analysis cycles. Therefore, we used the analysed fields at 00 UTC 28 July, 2 d into the cycles, as the initial fields of our subsequent extended forecasts and downscale experiments.

Although the importance of the lateral boundary perturbations in LETKF was noted by Saito et al. (2010), for simplicity in this study the boundary conditions of the forecast-analysis cycles were not perturbed. Figure 5b shows the spread distribution of surface pressure and rainfall at 00 UTC 28 July after convergence. The ensemble spread near the boundary was small because of the non-perturbed boundary condition. However, spread was recovered over the Yellow Sea, Sea of Japan and Pacific Ocean where the number of observation data were small, although the region of small spread was expanded from the southern boundary to the East China Sea because of the intense southerly flow. Over the Japanese Islands, located well within the domain, spreads were reduced by the assimilation of observation data. We found areas of large spread associated with low-pressure systems such as the one over western Hokkaido, showing the flow-dependent errors of the day.

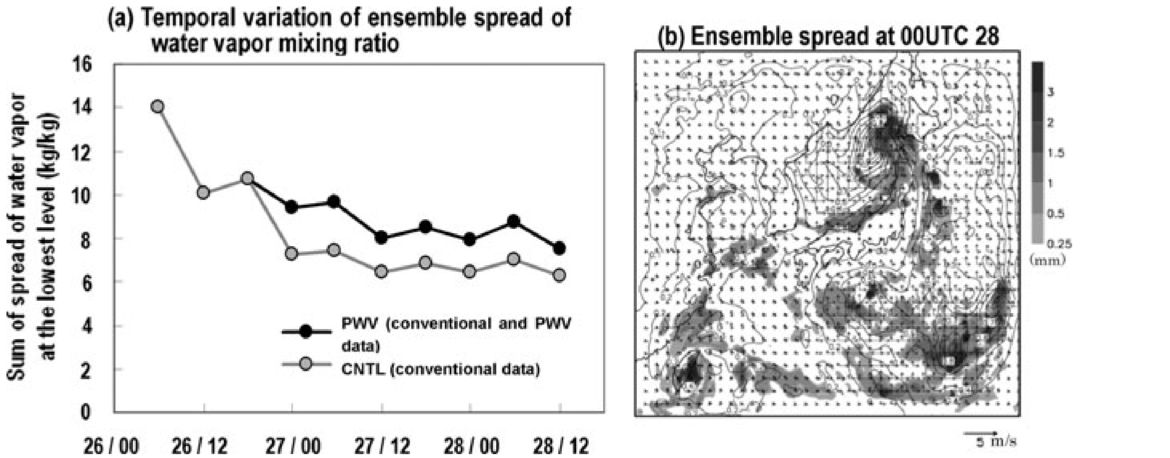

Fig. 5. (a) Time-series of ensemble spread of water vapour mixing ratio at the lowest level of NHM-LETKF, plotting the horizontal sum over the inner part of the model domain 30 grids inside the lateral boundary. Assimilation of PWV began at 00 UTC 27 July 2008. (b) Map of Japan region showing ensemble spread distributions of 6-h rainfall amount (shaded region), sea surface pressure (contours) and horizontal wind of the lowest level of NHM-LETKF (vectors) at 00 UTC 28 July when conventional data (surface and upper sounding data in mesoscale and global analysis) were assimilated.

After completing the forecast-analysis cycles, we performed an extended forecast and downscale forecasts with all ensemble members (Fig. 4). The initial time of the extended forecast was 00 UTC 28 July. The ensemble mean of the analysis fields on 00 UTC 28 July was not replaced with the analysed initial condition of Meso-4DVAR, although the ensemble mean on 12 UTC 27 July was replaced. Again, we employed a warm-rain process and the Kain–Fritsch convective parameterization scheme, and the boundary conditions were obtained from the JMA operational regional model.

Downscale forecasts were performed with JMANHM using horizontal grid intervals of 5 and 1.6 km and the same vertical resolution as in the forecast-analysis cycles of NHM-LETKF. The initial and boundary conditions of the 5-km downscale forecast were produced by the extended forecast, and the initial and boundary conditions of the 1.6 km forecast were given by the 5 km forecast. The domain of the 5 km forecast covers most of Japan except for Hokkaido, and that of the 1.6 km forecast covers the area where the heavy rainfall occurred. The initial times of both forecasts were 00 UTC 28 July. The three-ice bulk cloud microphysics scheme and Kain–Fritch scheme were applied to the 5 km forecast, but only the three-ice bulk cloud microphysics scheme was employed in the 1.6 km forecast.

5. Assimilation method of PWV

5.1. Estimation of PWV data

GPS-derived PWV data were estimated from GEONET data of GSI and the surface meteorological data of JMA using the following procedure (Shoji et al., 2004):

(1) Total delay was obtained by GPS software GIPSY/OASIS II from RINEX data that were provided by GEONET.

(2) Total delay is composed of hydrostatic delay and wet delay. Because hydrostatic delay is proportional to the surface pressure, surface pressure at GPS sites was estimated by interpolation of surface meteorological data, then hydrostatic delay was estimated in turn.

(3) Wet delay was obtained by subtracting hydrostatic delay from total delay.

(4) Wet delay is proportional to PWV, and the coefficient between them is a function of surface temperature. This coefficient was estimated using surface temperature at GPS sites from the JMA surface meteorological data, then PWV was estimated in turn.

Because the altitudes in the model topography using 20-km grid cells differ from the actual altitudes of GPS sites, we accounted for the influence of this difference on the PWV data by adding or subtracting the products of water vapour mixing ratio, air density and altitude difference; however, the error was only partially eliminated. To further reduce the error caused by this difference of altitudes, only GPS data from stations whose altitude differed from the model topography by less than 100 m were used. This step reduced the number of qualifying PWV data, especially in the mountainous area of central Japan. For western Japan, where the mountains are lower, many GPS data were acceptable (Fig. 6a).

5.2. Assimilation of PWV data

For this study, we produced intermediate profiles of relative humidity from observations and statistical data from outputs of

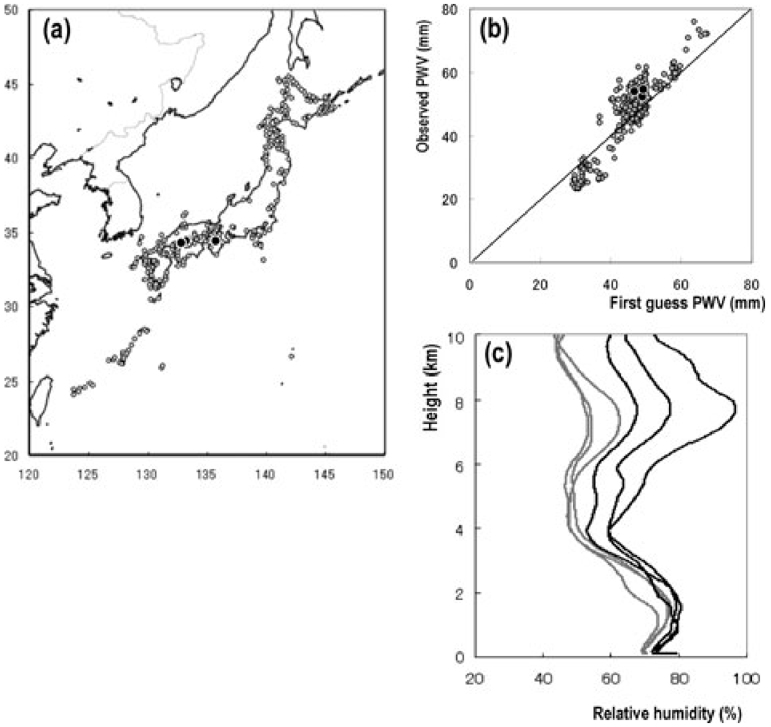

Fig. 6. (a) Map showing positions of PWV data assimilated in this study. (b) Scatter diagram of observed GPS-derived PWV versus first-guess PWV at 00 UTC 28 July 2008. (c) Examples of the intermediate relative humidity profile produced from the observed PWV (black lines) and first guess of the ensemble forecast (grey lines) for three localities marked by black circles in (a) and (b).

the ensemble forecast, and these profiles were assimilated by NHM-LETKF. We took this step because NHM-LETKF had an observation operator for relative humidity. We began with the PWV data described in Section 5.1 and proceeded in two steps as follows:

(1) Vertical profiles of the ensemble means and spreads of water vapour, temperature and pressure were obtained from the forecasts in the forecast-analysis cycles by using the equations

$$\mathrm{mean\_}X_{\mathrm{fg\_GPS}}(k) = \sum_{i=i_{\mathrm{GPS}}-5}^{i_{\mathrm{GPS}}+5} \sum_{j=j_{\mathrm{GPS}}-5}^{j_{\mathrm{GPS}}+5} \mathrm{mean\_}X_{\mathrm{fg\_GPV}}(k)/121, \quad (4)$$

$$\mathrm{mean\_RH}_{\mathrm{fg\_GPS}}(k) = \mathrm{RH}\left(\mathrm{mean\_}Qv_{\mathrm{fg\_GPV}}(k),\ \mathrm{mean\_}T_{\mathrm{fg\_GPV}}(k),\ \mathrm{mean\_}p_{\mathrm{fg\_GPV}}(k)\right), \quad (5)$$

$$\mathrm{spread\_RH}_{\mathrm{fg\_GPS}}(k) = \mathrm{RH}\left(\mathrm{spread\_}Qv_{\mathrm{fg\_GPV}}(k),\ \mathrm{mean\_}T_{\mathrm{fg\_GPV}}(k),\ \mathrm{mean\_}p_{\mathrm{fg\_GPV}}(k)\right), \quad (6)$$

$$\mathrm{spread\_}Qv_{\mathrm{fg\_GPS}}(k) = \max_{i=i_{\mathrm{GPS}}-5,\, i_{\mathrm{GPS}}+5;\, j=j_{\mathrm{GPS}}-5,\, j_{\mathrm{GPS}}+5} \left(\mathrm{spread\_}Qv_{\mathrm{fg\_gpv}}(k)\right), \quad (7)$$

where X represents water vapour mixing ratio (Qv), temperature (T), pressure (p) and air density (ρ). Errors in horizontal positions were considered by averaging the profiles of the ensemble mean over the region within 100 km (±5 grid cells) of GPS sites, and by obtaining the maximum of ensemble spread over the same region, respectively. Profiles of the ensemble means were used as the first-guess profiles. The spread profiles indicate the degree to which relative humidity can be perturbed at the positions of GPS receivers.

(2) We assumed that the difference between the intermediate profiles and first-guess profiles is proportional to the maximum of the ensemble spread. In other words, the intermediate profiles were produced by distributing the difference between the observed PWV data and the first-guess PWV (the vertical sum of water vapour in the first-guess profile) to the values of each model layer on the basis of this assumption. The equations of this procedure are

$$\mathrm{PWV}_{\mathrm{fg\_GPS}} = \sum_{l=1}^{L} Qv_{\mathrm{fg\_GPS}}(k)\,\rho_{\mathrm{fg\_GPS}}(k)\mathrm{d}z, \quad (8)$$

$$\mathrm{PWV}_{\mathrm{intermediate\_GPS}} = \mathrm{PWV}_{\mathrm{fg\_GPS}} + \alpha \sum_{l=1}^{L} \frac{\mathrm{spread\_RH}_{\mathrm{fg\_GPS}}(k)}{\mathrm{mean\_RH}_{\mathrm{fg\_GPS}}(k)} \times \mathrm{mean\_}Qv_{\mathrm{fg\_GPS}}(k)\,\mathrm{mean\_}\rho(k)\,\mathrm{d}z, \quad (9)$$

$$\mathrm{RH}_{\mathrm{intermediate\_GPS}} = \mathrm{mean\_RH}_{\mathrm{fg\_GPS}}(k) + \alpha\,\mathrm{spread\_RH}_{\mathrm{fg\_GPS}}(k), \quad (10)$$

where $\mathrm{PWV}_{\mathrm{fg\_GPS}}$ and dz are the first-guess PWV and thickness of the layer, and α was determined so that $\mathrm{PWV}_{\mathrm{intermediate\_GPS}}$ was equal to the observed PWV. Then, intermediate profiles of relative humidity $\mathrm{RH}_{\mathrm{intermediate\_GPS}}$ were expressed as $\mathrm{mean\_RH}_{\mathrm{fg\_GPS}}(k) + \alpha\,\mathrm{spread\_RH}_{\mathrm{fg\_GPS}}(k)$.

Profiles of maximum spreads were used in order to consider any possibility in the region within 100 km of the profiles. This approach might result in excessive changes in the profiles; however, the increment of relative humidity becomes small if the first-guess PWV is similar to the observed PWV.

Figure 6b shows a scatter diagram of the observed and first-guess PWV at 00 UTC 28 July. At many GPS sites where the observed PWV exceeded 40 mm, the observed PWV were larger than the first-guess PWV. Because these humid points were mainly in western Japan, assimilation of the observed PWV data should intensify the modelled rainfall over western Japan. Figure 6c shows examples of these intermediate and first-guess profiles from three GPS sites indicated by black circles in Figs. 6a and b. At these sites, the observed PWV was larger than the first-guess PWV (Fig. 6b). The relative humidities in the intermediate profiles exceeded those of the first-guess profiles at all altitudes (Fig. 6c). The differences between profiles in the water vapour amount are very small in the higher layers, though the differences in relative humidity are large.

6. Results of the LETKF analysis and impacts of PWV data

6.1. Results of the forecast-analysis cycles

Figure 7 shows the analysed fields of the rainfall region obtained from the forecast-analysis cycles. These distributions contain the assimilation impacts of conventional data after re-centring of the ensemble mean. Figures 7a and b show results of the control experiments, which assimilated only conventional data. Rainfall along the north side of western Japan was reproduced at 00 UTC 28 July, though the rainfall amount was smaller and the region of rainfall was larger. These differences from observations were partly caused by the averaging process of ensemble forecasts. At 06 UTC 28 July (Fig. 7b), a small rainfall region near Kobe was reproduced. Figures 7c and d show results from the assimilation of PWV data in addition to conventional data. The analysed fields of this experiment had already converged by 00 UTC

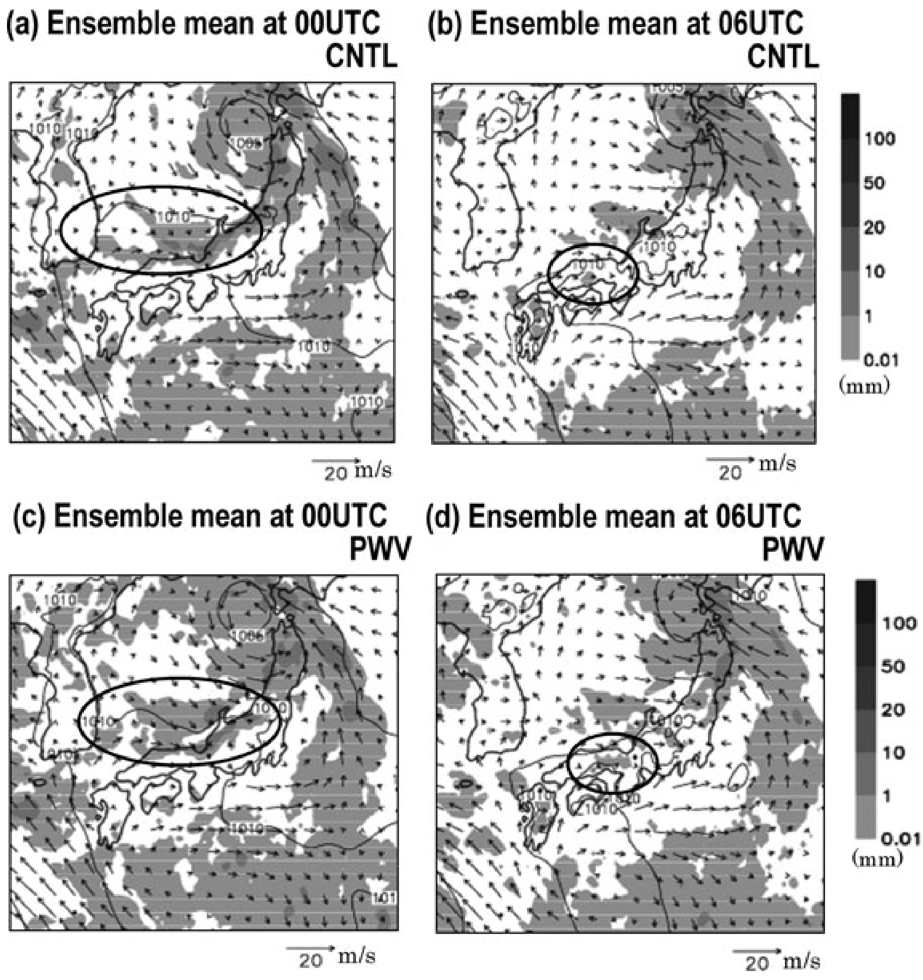

Fig. 7. Maps showing ensemble mean distributions of 6-h rainfall amount (shading), sea surface pressure (contours) and horizontal wind at the lowest level of NHM-LETKF (vectors) at (a) 00 UTC 28 July and (b) 06 UTC 28 July when conventional data were assimilated, and (c) 00 UTC 28 July and (d) 06 UTC 28 July when conventional data and GPS-derived PWV data were assimilated. Ellipses indicate the positions of intense rainfall regions.

28 July (Fig. 5a). Despite the addition of PWV data, pressure patterns and horizontal wind and rainfall distributions were not greatly changed. However, the rainfall region indicated by the ellipse in Fig. 7d, which corresponds to the intense convective band in region **B**, became larger with the addition of PWV data.

To demonstrate the impact of PWV data more clearly, the difference in water vapour fields between the control and PWV experiments are shown in Fig. 8. When PWV data were assimilated, the ensemble mean of water vapour at the lowest altitude increased over western Japan (ellipses in Figs 8a and b). This difference was consistent with the increment we expected (see Section 4). Other regions with a relatively small increment appeared in Korea and northeastern China. The ensemble spread of water vapour became smaller over western Japan (indicated by broken ellipses in Figs 8c and d) because of the assimilation of PWV data.

As shown in Fig. 8, water vapour at the lowest level of the NHM-LETKF near Japan increased with the assimilation of PWV data. Thus, the reinforcement of convection was expected during numerical simulations by JMANHM. Because the grid interval of 20 km should be too coarse to simulate convective cells, whose horizontal scale is only a few kilometres, the results of downscale experiments are explained in the next subsection.

6.2. *Results of downscale experiments*

The initial and boundary conditions of the 5-km downscale forecast were produced by spatial interpolation of analysis fields at 00 UTC 28 July and the outputs of the extended forecasts (Fig. 3). When the 5 km forecast using PWV data was performed, the band of rainfall in region **B** was more distinct over western Japan in the ensemble mean field (Fig. 9a), but the rainfall amount was still smaller than the observed amount. Figure 10 shows that the probability of the 1-h rainfall amount in the convective band exceeding 1 mm and 5 mm were 70% and 20%, respectively, among members of the ensemble. These probabilities mean that the rainfall intensity of the convective band in region **B** was reproduced in only a few ensemble members, although most members reproduced the rainfall band. The rainfall band in region **A** was not reproduced by any ensemble

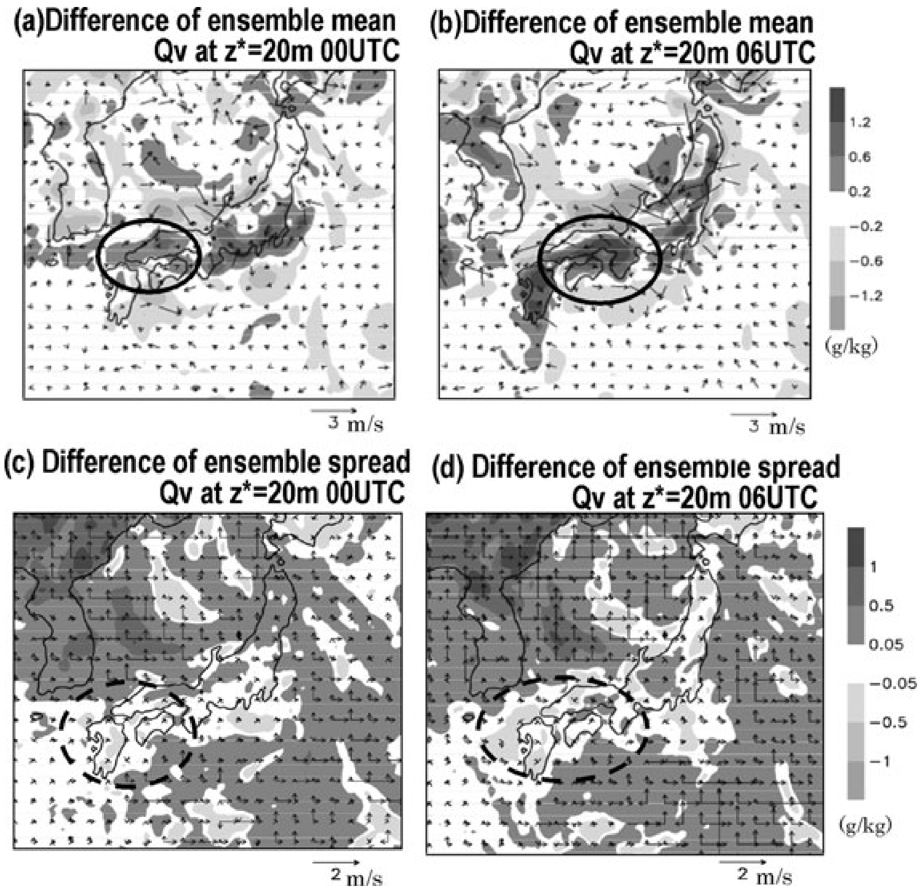

Fig. 8. Maps showing (a) differences of ensemble means (PWV minus CNTL) of the water vapour mixing ratio (shading) and horizontal wind at the lowest level of NHM-LETKF (vectors) at 00 UTC 28 July and (b) 06 UTC July, and (c) differences of ensemble spreads at 00 UTC 28 July and (d) 06 UTC July.

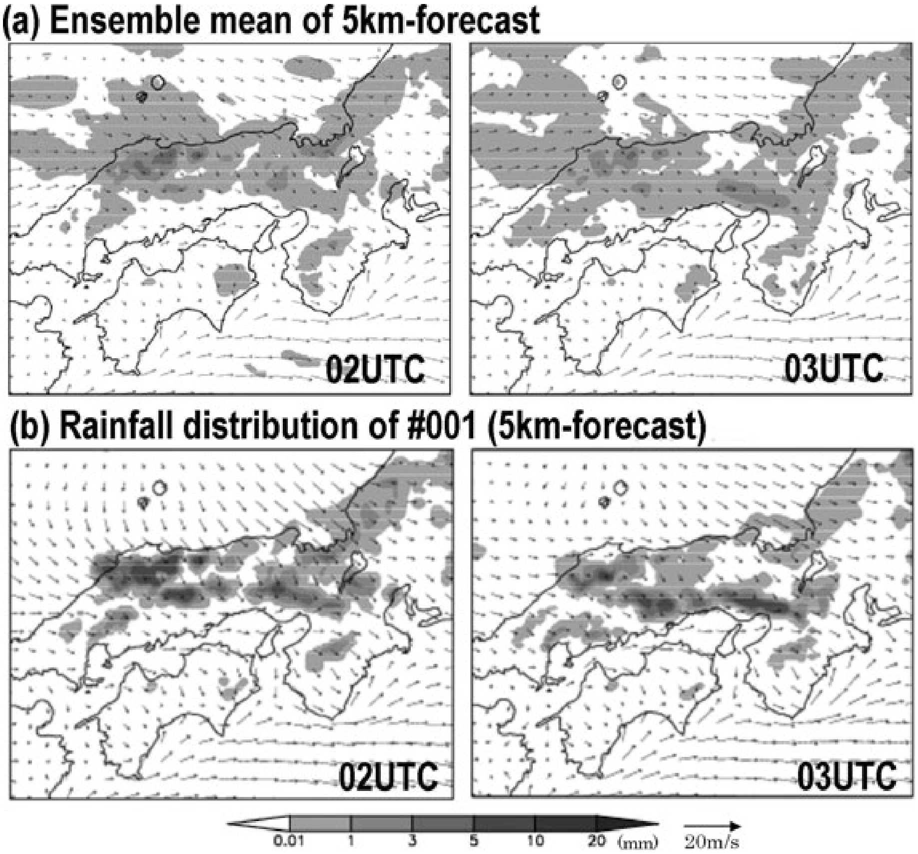

Fig. 9. Maps of 5 km forecast region showing (a) ensemble mean distributions of rainfall (shaded region) and surface wind (vectors) and (b) rainfall and surface wind distributions of ensemble member #001.

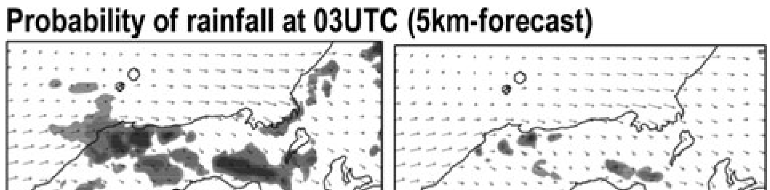

Fig. 10. Maps of 5 km forecast region showing probability distributions of rainfall in which 1-h rainfall amount exceeded 1 mm and 5 mm (shading). Vectors indicate surface wind of ensemble mean.

members. Comparison of Figs 8a and b shows that water vapour over the sea north of western Japan was not increased. This increment distribution is one of the reasons why the rainfall in region **A** was not reproduced.

We compared the result of our 5 km forecast with that of Shoji et al. (2009), in which the PWV data were assimilated by the Meso-4DVAR system. In their results, water vapour over the sea north of western Japan was increased by the assimilation of PWV data of East Asia and developed the convective band in region **A**, whereas water vapour over western Japan was increased in our NHM-LETKF experiment. Besides this difference, the impact of PWV assimilation on our 5 km forecasts was much smaller, because we had fewer assimilation data and the assimilation period was shorter than those of Shoji et al. (2009). However, there were a few ensemble members that reproduced the convective band. This result also shows that the rainfall band can be reproduced by giving the ensemble perturbation, even if the improvement of the analysis fields was modest.

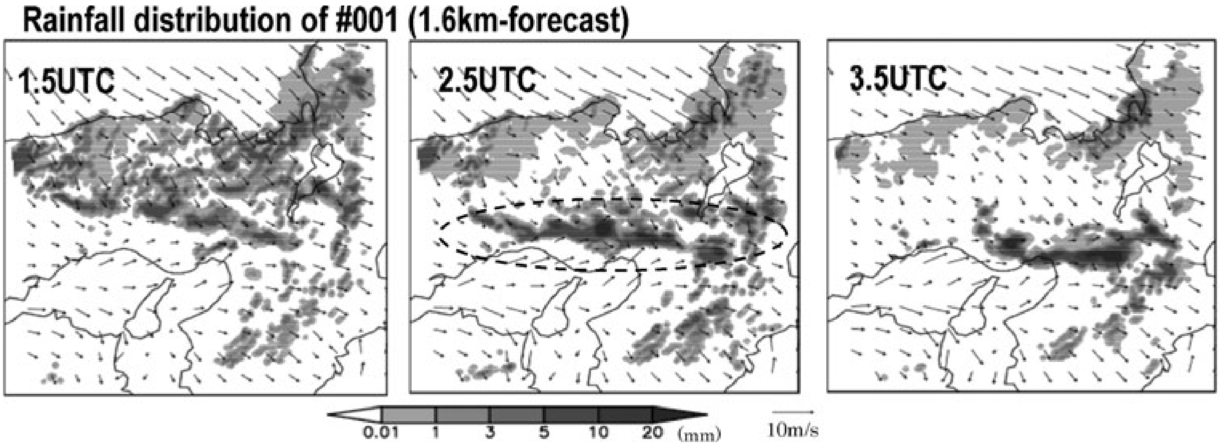

Fig. 11. Maps of 1.6 km forecast showing rainfall and surface wind distributions obtained by downscale experiments of ensemble member #001 for 28 July. Rainfall is 1-h rainfall amount extrapolated from 10-min rainfall amount. Dashed ellipse at 2.5 UTC indicates the rainfall band of region **B**.

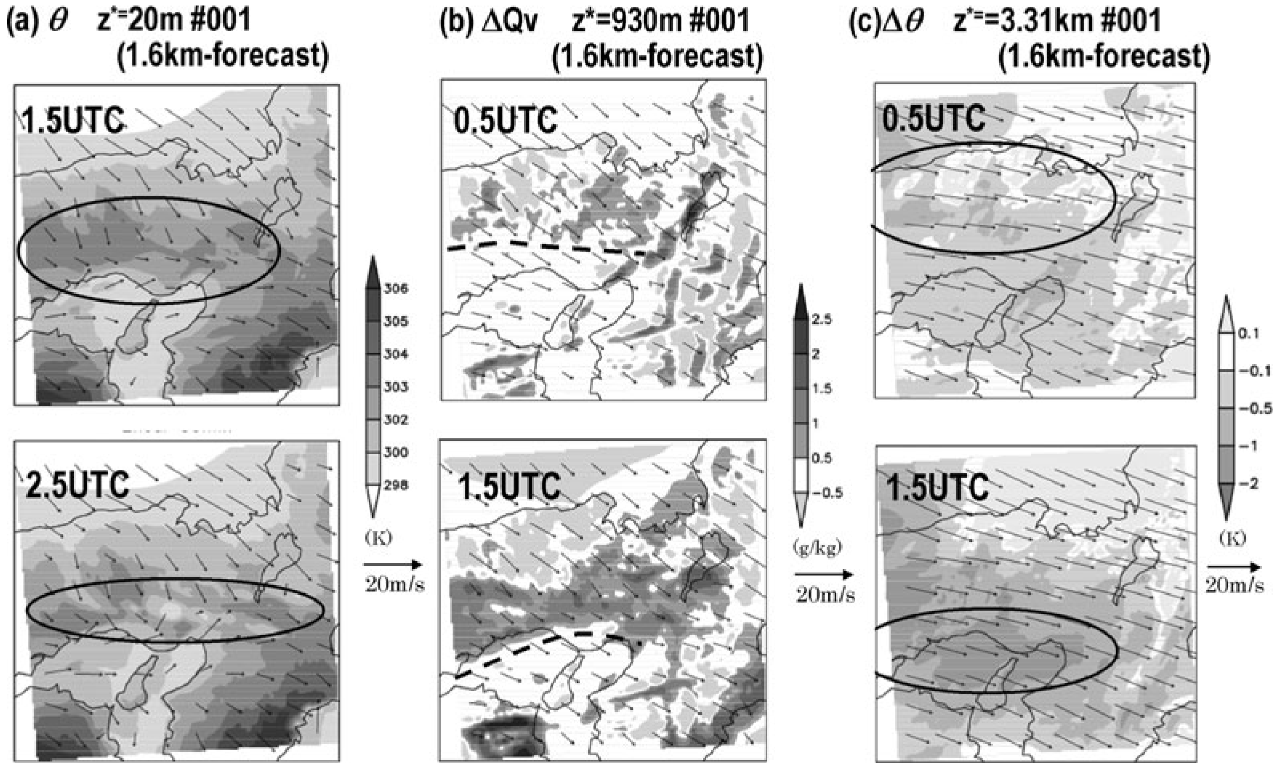

Fig. 12. Maps of 1.6 km forecast. (a) Distribution of potential temperature and horizontal wind at the height of 20 m at 1.5 and 2.5 UTC 28 July. Ellipse at 1.5 UTC denotes a warm region where potential temperature exceeds 303 K. Ellipse at 2.5 UTC denotes position of convective band in region **B**. (b) Increments of water vapour mixing ratio from that of 00 UTC 28 July at the height of 930 m. Dashed lines denote southern edge of region of elevated water vapour mixing ratio. (c) Increments of potential temperature from that of 00 UTC 28 July at the height of 3.31 km. Ellipses denote intrusion of relatively cold air from the northwest.

Rainfall distribution in ensemble member #001, in which the rainfall amount was largest, is shown in Fig. 9b. Figure 11 shows the downscale experiment of this same member at 1.6-km resolution. Rainfall is 1-h rainfall amount extrapolated from 10-min rainfall amount. Low-level convergence generated convective cells over western Japan that were then organized into the rainfall band, which moved southeastwards. This evolution of convective cells was similar to observations (Fig. 2). The potential temperatures and water vapour mixing ratios around the time that the convective band in region **B** was generated are shown in Fig. 12. Convective cells were generated in the warm region where the potential temperature exceeded 303 K (ellipse in Fig. 12a at 1.5 UTC). Regions of increased water vapour at 930 m altitude are shown in Fig. 12b. The region of increased water vapour moved southwards, generating the intense rainfall in region **B**. This evolution of the humid region was consistent with that of the observed PWV (Fig. 3). The region of increased water vapour at 1.5 UTC is partly the result of the low-level convergence shown in Fig. 12a. Relatively cold air, shown as the ellipse for 1.5 UTC in Fig. 12c, intruded from the northwest just before the generation of the convective cells in region **B**. These changes indicate that favourable conditions for convection, including low-level convergence of warm moist air and cold air intrusion at middle levels, caused the intense rainfall band of region **B**. Beside these factors, the chosen grid interval and parameterization scheme are important for reproducing the intense rainfall band, because it was reproduced more realistically when the grid interval of 1.6 km was used and a microphysical scheme was adopted.

7. Summary

GPS-derived PWV data were assimilated by the NHM-LETKF system in an attempt to reproduce the intense rainfall system of 28 July 2008 in Japan. This system consisted of two bands of intense rainfall, one in region **A** along the north side of western Japan and one in region **B** to the south of region **A**, that were not reproduced by the operational weather forecast model. When PWV data from 18 UTC 26 July to 06 UTC 28 July were assimilated, low-level water vapour was increased and the rainfall region in western Japan became larger, closer to the actual observations.

Downscale forecast experiments were conducted from the analysed fields of 00 UTC 28 July using JMANHM and grid intervals of 5 km and 1.6 km. In the 5 km forecast with convective parameterization, the rainfall band was reproduced in a few ensemble members, but their rainfall amounts were smaller than observations. In the 1.6 km forecast, the rainfall intensity of the band in region **B** became larger. When the intense convective band was generated, low-level humid air converged from the northwest and was overlaid by cold air at 3 km altitude. These results show that low-level convergence of warm humid air in unstable atmospheric conditions generated and developed the convective cells in the intense rainfall band of region **B**.

In this study, we produced intermediate profiles of relative humidity from observed PWV data and statistical data obtained from outputs of the ensemble forecast, based on the first-step approximation that the increment of relative humidity is proportional to the ensemble spread, and used these profiles as the pseudo-observation data. The intermediate profiles are not necessarily similar to the observed ones because our assumptions were crude. For better assimilation of PWV, it would be desirable to treat PWV more directly as the integrated value of water vapour.

The results of this study also suggest that the detection of severe events would be improved by using large ensemble forecasts when the assimilation is performed. Even when intense rainfall is not reproduced by the assimilation of conventional data, assimilation of additional data, such as PWV data, improves the rainfall forecast.

In the experiments of this study, the ensemble mean was replaced by the analysis from the Meso-4DVAR system every day at 12 UTC during the forecast-analysis cycles. The assimilation period of this study was as short as 3 d, and number of the ensemble member was only 20. The slant water vapour data such as SWV or slant wet delay, which have the information of anisotropic distribution of water vapour around GPS receivers, is expected to improve rainfall forecasts by intensifying the horizontal gradient of water vapour if SWV is used instead of PWV (Bauer et al., 2011; Seko et al., 2010). Doppler radar data (e.g. reflectivity and radial wind) are also expected to improve the rainfall forecast, if the high-resolution LETKF is used. To improve the accuracy of analyses, more assimilation data such as slant water vapour data and Doppler radar data, the longer assimilation period and the increases of ensemble size and of the horizontal resolution of analysis system are needed.

8. Acknowledgments

We thank two anonymous referees for valuable comments that significantly improved the manuscript. Initial and boundary conditions for the numerical models and conventional data were provided by the Numerical Prediction Division (NPD) of JMA. GPS data were provided by GSI. Dr. Tadashi Fujita of NPD provided the source code of the nonlocal scheme of LETKF. The 'Kasaneru 3D tool' developed by Tokyo District Meteorological Observatory of JMA was used to depict the rainfall distribution observed by operational radars. We also thank Dr. Tetsuya Iwabuchi of University Corporation for Atmospheric Research, Mr. Masaru Kunii of MRI and Mr. Hiroshi Kusabiraki of NPD who provided useful comments.

References

Bauer, H.-S., Wulfmeyer, V., Schwitalla, T., Zus, F. and Grzeschik, M. 2011. Operational assimilation of GPS slant path delay measurements into the MM5 4DVAR system, *Tellus A*, this issue, doi:10.1111/j.1600–0870.2010.00489.x

Dowell, D. C., Zhang, F., Wicker, L. J., Snyder, C. and Crook, N. A. 2004. Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, Supercell: ensemble Kalman filter experiments. *Mon. Wea. Rev.* **132**, 1982–2005.

Fujita, T., Tsuguchi, H., Miyoshi, T., Seko, H. and Saito, K. 2009. Development of a mesoscale ensemble data assimilation system at JMA. Report of the Grant-in-Aid for Scientific Research (B) (2005–2008), No. 17110035, 232–235. (in Japanese)

Hunt, B. R., Kostelich, E. J. and Szunyogh, I. 2007. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D*, **230**, 112–126.

Kawabata, T., Seko, H., Saito, K., Kuroda, T., Tamiya, K. and co-authors. 2007. An assimilation and forecasting experiment of the Nerima heavy rainfall with a cloud-resolving nonhydrostatic 4-dimensional variational data assimilation system. *J. Meteorol. Soc. Japan*, **85**, 255–276.

Meng, Z. and Zhang, F. 2008. Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVar in a real-data case study. *Mon. Wea. Rev.* **136**, 522–540.

Miyoshi, T. 2010. NHM-LETKF, *Tech. Rep. MRI*, **62**, 159–163.

Miyoshi, T. and Aranami, K. 2006. Applying a four-dimensional local ensemble transform Kalman filter (4D-LETKF) to the JMA nonhydrostatic model (NHM). *SOLA*, **2**, 128–131.

Miyoshi, T., Yamane, S. and Enomoto, T. 2007. Localizing the error covariance by physical distances within a local ensemble transform Kalman filter (LETKF). *SOLA*, **3**, 89–92.

Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J. and co-authors. 2004. A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428.

Saito, K., Fujita, T., Yamada, Y., Ishida, J., Kumagai, Y. and co-authors. 2006. The operational JMA nonhydrostatic mesoscale model. *Mon. Wea. Rev.* **134**, 1266–1298.

Saito, K., Kunii, M., Hara, M., Seko, H., Hara, T. and co-authors. 2010. WWRP Beijing 2008 Olympics Forecast Demonstration/Research and Development Project (B08FDP/RDP). *Tech. Rep. MRI*, **62**, 210 pp.

Saito, K., Hara, M., Kunii, M., Seko, H. and Yamaguchi, M. 2011. Comparison of initial perturbation methods for the mesoscale ensemble prediction system of the Meteorological Research Institute for the WWRP Beijing 2008 Olympics Research and Development Project (B08RDP). *Tellus*, **63A**, this issue.

Seko, H. 2010. Local ensemble transform Kalman filter (LET) method. *Tech. Rep. MRI*, **62**, 80–84.

Seko, H. and Nakamura, H. 2005. Analytical and numerical studies on meso-β scale precipitation bands observed over southern Kyushu on 7 July 1996. *Papers in Meteorology and Geophysics*, **55**, 55–74.

Seko, H., Kawabata, T., Tsuyuki, T., Nakamura, H., Koizumi, K. and co-authors. 2004. Impacts of GPS-derived water vapor and radial wind measured by Doppler radar on numerical prediction of precipitation. *J. Meteorol. Soc. Japan*, **82**, 473–489.

Seko, H., Kunii, M., Shoji, Y. and Saito, K. 2010. Improvement of rainfall forecast by assimilations of ground-based GPS data and radio occultation data. *SOLA*, **6**, 81–84.

Shoji, Y., Nakamura, H., Iwabuchi, T., Aonashi, K., Seko, H. and co-authors. 2004, Tsukuba GPS dense net campaign observation: improvement in GPS analysis of Slant Path delay by stacking one-way postfit phase residuals. *J. Meteorol. Soc. Japan*, **82**, 301–314.

Shoji, Y., Kunii, M. and Saito, K. 2009. Assimilation of nationwide and global GPS PWV data for heavy rain event on 28 July 2008 in Hokuriku and Kinki, Japan. *SOLA*, **5**, 45–48.

Snyder, C. and Zhang, F. 2003. Assimilation of simulated doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.* **131**, 1663–1677.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2003. Ensemble square root filters. *Mon. Wea. Rev.* **131**, 1485–1490.

Tong, M. and Xue, M. 2005. Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSSE experiments. *Mon. Wea. Rev.* **133**, 1789–1807.

Xue, M., Tong, M. and Droegemeier, K. K. 2006. An OSSE framework based on the ensemble square root Kalman filter for evaluating the impact of data from radar networks on thunderstorm analysis and forecasting. *J. Atmos. Oceanic Tech.* **23**, 46–66.

Yan, X., Ducrocq, V., Jaubert, G., Brousseau, P., Poli, P. and co-authors. 2009. The benefit of GPS zenith delay assimilation to high-resolution quantitative precipitation forecasts: a case-study from COPS IOP 9. *Q. J. R. Meteorol. Soc.* **135**, 1788–1800,

Zhang, F., Snyder, C. and Sun, J. 2004. Impacts of initial estimate and observation availability with an ensemble Kalman filter. *Mon. Wea. Rev.* **132**, 1238–1253.

Zhang, F., Meng, Z. and Aksoy, A. 2006. Test of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: perfect model experiments. *Mon. Wea. Rev.* **134**, 722–736.

Zhang, M., Ni, Y. and Zhang, F. 2007. Variational assimilation of GPS precipitable water vapor and hourly rainfall observations for a meso-β-scale heavy precipitation event during the 2002 Mei-Yu season. *Adv. Atmos. Sci.* **24**, 509–526.