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

Tellus (2011), 63A, 445–467
Printed in Singapore. All rights reserved

TELLUS

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)

By KAZUO SAITO*, MASAHIRO HARA, MASARU KUNII, HIROMU SEKO and MUNEHIKO YAMAGUCHI, *Meteorological Research Institute, 1-1 Nagamine, Tsukuba, Ibaraki 305-0052, Japan*

(Manuscript received 17 May 2010; in final form 21 December 2010)

ABSTRACT

Different initial perturbation methods for the mesoscale ensemble prediction were compared by the Meteorological Research Institute (MRI) as a part of the intercomparison of mesoscale ensemble prediction systems (EPSs) of the World Weather Research Programme (WWRP) Beijing 2008 Olympics Research and Development Project (B08RDP).

Five initial perturbation methods for mesoscale ensemble prediction were developed for B08RDP and compared at MRI: (1) a downscaling method of the Japan Meteorological Agency (JMA)'s operational one-week EPS (WEP), (2) a targeted global model singular vector (GSV) method, (3) a mesoscale model singular vector (MSV) method based on the adjoint model of the JMA non-hydrostatic model (NHM), (4) a mesoscale breeding growing mode (MBD) method based on the NHM forecast and (5) a local ensemble transform (LET) method based on the local ensemble transform Kalman filter (LETKF) using NHM. These perturbation methods were applied to the preliminary experiments of the B08RDP Tier-1 mesoscale ensemble prediction with a horizontal resolution of 15 km. To make the comparison easier, the same horizontal resolution (40 km) was employed for the three mesoscale model-based initial perturbation methods (MSV, MBD and LET).

The GSV method completely outperformed the WEP method, confirming the advantage of targeting in mesoscale EPS. The GSV method generally performed well with regard to root mean square errors of the ensemble mean, large growth rates of ensemble spreads throughout the 36-h forecast period, and high detection rates and high Brier skill scores (BSSs) for weak rains. On the other hand, the mesoscale model-based initial perturbation methods showed good detection rates and BSSs for intense rains. The MSV method showed a rapid growth in the ensemble spread of precipitation up to a forecast time of 6 h, which suggests suitability of the mesoscale SV for short-range EPSs, but the initial large growth of the perturbation did not last long. The performance of the MBD method was good for ensemble prediction of intense rain with a relatively small computing cost. The LET method showed similar characteristics to the MBD method, but the spread and growth rate were slightly smaller and the relative operating characteristic area skill score and BSS did not surpass those of MBD. These characteristic features of the five methods were confirmed by checking the evolution of the total energy norms and their growth rates.

Characteristics of the initial perturbations obtained by four methods (GSV, MSV, MBD and LET) were examined for the case of a synoptic low-pressure system passing over eastern China. With GSV and MSV, the regions of large spread were near the low-pressure system, but with MSV, the distribution was more concentrated on the mesoscale disturbance. On the other hand, large-spread areas were observed southwest of the disturbance in MBD and LET. The horizontal pattern of LET perturbation was similar to that of MBD, but the amplitude of the LET perturbation reflected the observation density.

*Corresponding author.
e-mail: ksaito@mri-jma.go.jp
DOI: 10.1111/j.1600-0870.2010.00509.x

1. Introduction

Rapid progress in computer power in recent years has enabled us to develop higher resolution models for numerical weather

prediction (NWP). The accuracy of operational NWP in the world's major forecast centres has been considerably improved by advances in numerical modelling and data assimilation techniques, and meso-β scale convective systems are now becoming the target of short-range forecasting (Shapiro and Thorpe, 2004). However, there are still many difficulties in producing predictions of severe mesoscale phenomena that specify their intensity, location and timing. The performance of quantitative precipitation forecasting is totally unsatisfactory if we focus on the statistical scores of intense rain (e.g. Ebert et al., 2003; Bousquet et al., 2006). One reason for this low performance of current operational NWP of intense rains is the lack of available observations for high-resolution real time data assimilation (e.g. Nuret et al., 2005). Another reason is the inherent low predictability of severe small-scale phenomena, which occur under convectively unstable atmospheric conditions. Without an orographic trigger or synoptic-scale forcing, initiation of such phenomena is very sensitive to small differences in the initial conditions. In some cases, trivial initial perturbations with magnitudes less than those of the direct in-situ observation errors result in a completely different evolution of deep convection (e.g. Hohenegger and Schar, 2007; Yoden, 2007).

To consider forecast errors due to uncertainties in initial conditions, ensemble prediction systems (EPSs) are widely used for medium range NWP. These systems, however, are designed for global forecasts and mainly focus on the uncertainties due to synoptic-scale baroclinic instability, whereas many natural hazards, such as local heavy rainfall or wind gusts, are caused by mesoscale disturbances rather than by synoptic-scale systems. To cope with the significant forecast uncertainty of severe mesoscale phenomena, there is now increasing interest in the development of mesoscale EPSs. Several forecast centres have begun to use short-range regional EPSs: the Short-Range Ensemble Forecast (SREF) system of the National Center for Environment Prediction (NCEP) (Du et al., 2003; Stensrud and Yussouf, 2007); the Met Office Global and Regional Ensemble Prediction System (MOGREPS) of the UK Met Office (Bowler et al., 2008); the Limited Area Ensemble Prediction System within the COnsortium for Small-scale Modeling consortium (COSMO-LEPS) of Agenzia Regionale Prevenzione e Ambiennte Romagna (ARPA-SMR) of Italy (Marsigli et al., 2005); and the Limited Area Ensemble Forecasting system using the ALADIN model (ALADIN-LAEF) of the Central Institute for Meteorology and Geodynamics (Zentralanstalt für Meteorologie und Geodynamik, ZAMG) of Austria (Wang et al., 2011).

Initial perturbation is a key problem in ensemble forecasting. Hoffman and Kalnay (1983) proposed lagged average forecasting (LAF) and showed superiority to Monte Carlo forecasting. Toth and Kalnay (1993, 1997), at NCEP, developed the first operational ensemble forecasting method, the breeding growing mode (BGM) method, in which Lyapunov vectors are obtained by time integration of the non-linear forecast model. The singular vector (SV) method, developed by Buizza and Palmer (1995), was implemented at the European Centre for Medium-Range Weather Forecasts (ECMWF; Molteni et al., 1996). The SV method computes the leading eigenvectors using a linear tangent model and its adjoint. JMA's operational global EPS has also implemented the SV method for 1-week forecasting since 2006 and for typhoon prediction since 2007 (Yamaguchi et al., 2009). Yamaguchi and Majumdar (2010) compared global ensemble initial perturbations by ECMWF, NCEP and JMA using The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE) data, and investigated the dynamical mechanisms of perturbation growth associated with the tropical cyclone.

Recently, the ensemble transform of the ETKF (Bishop et al., 2001) is becoming more popular as a generator of initial perturbations for global EPSs. Wang and Bishop (2003) compared BGM and ETKF for global ensemble prediction and showed that ETKF has the advantage that the amplitude of the initial perturbation reflects the magnitude of the local analysis error. Global EPS of the UK Met Office (Bowler et al., 2008) uses ETKF, while the Meteorological Service of Canada (MSC; Houtekamer and Mitchell, 2005) uses an incremental EnKF system.

The features of regional EPSs differ from those of global EPSs. One choice for the initial perturbation method of a limited model is dynamical downscaling of the global EPS. Several current operational regional EPSs use downscaling of perturbations processed by a global EPS (e.g. Marsigli et al., 2005; Bowler et al., 2008; Houtekamer et al., 2009). On the other hand, some regional EPSs generate initial perturbations using regional model-based perturbation methods such as BGM (Du et al., 2003; Wang et al., 2011), SV (Li et al., 2008) and ETKF (Bowler and Mylne, 2009).

Another important factor in regional EPSs is the existence of lateral boundary conditions, which influence ensemble solutions. Nutter et al. (2004) first examined the effect of lateral boundary perturbations on limited-area ensemble forecasts. Torn et al. (2006) tested ensemble boundary conditions on limited-area ensemble Kalman filters (EnKF), but a coarse horizontal resolution of 100 km was used in their study. Since mesoscale EPSs are relatively new applications in the NWP field, their perturbation methods have not been fully investigated. For example, Wang and Bishop (2003) compared BGM and ETKF, but employed a low-resolution (T42 and 18 levels) global climate model. Bowler (2006) compared SV, BGM and ETKF in ensemble prediction, but with a simple Lorenz model. Bowler and Mylne (2009) tested ETKF for the regional MOGREPS with a horizontal resolution of 24 km, but they failed to show any relative advantage to using a regional ETKF compared with downscaling a global ETKF. No comparisons between a regional ETKF and a regional BGM have been made. The regional SV method is a relatively new application, and its performance in mesoscale EPS has not been investigated so far.

In 2008, the World Weather Research Programme (WWRP) Beijing 2008 Olympics Forecast Demonstration

Project/Research and Development Project (B08FDP/RDP), an international research project of the WWRP of the World Meteorological Organization (WMO) for short-range weather forecasting, was conducted in conjunction with the Beijing Olympic Games. The main part of B08RDP, called Tier-1, was an intercomparison of mesoscale EPSs with a horizontal resolution of 15 km. The Meteorological Research Institute (MRI) of the Japan Meteorological Agency (JMA) participated in this project, applying the JMA non-hydrostatic model (NHM). Prior to the 2008 intercomparison period (one month from 25 July to 23 August 2008), MRI developed five initial perturbation methods: (1) a downscaling method of JMA's operational one-week EPS (hereafter, WEP), (2) a targeted GSV method, (3) a mesoscale singular vector (MSV) method based on the adjoint model of NHM, (4) a mesoscale breeding growing mode (MBD) method based on the NHM forecast and (5) a local ensemble transform (LET) method based on the local ETKF using NHM (NHM-LETKF). Results of the ensemble forecasts made with these five methods were compared objectively by evaluating the evolution of the ensemble spreads, the root mean square error (RMSE) of the ensemble mean against the four-dimensional variational data assimilation (4D-Var) analysis, and two quantitative precipitation scores, the relative operating characteristics area skill score (ROCSS) and the Brier skill score (BSS). In addition, two lateral boundary perturbation methods were developed, and their impacts on the ensemble forecast were also examined.

In this paper we compare the use of these five initial perturbation methods for the mesoscale ensemble prediction as the first part of the MRI's studies for the B08RDP project. To the best of our knowledge, this is the first comprehensive comparison of initial perturbation methods for a full-scale mesoscale EPS that includes global and mesoscale SVs, BGM and LETKF.

This paper is organized as follows. Section 2 describes the WWRP B08FDP/RDP project and MRI's participation in that project. Section 3 presents the five initial perturbation methods for mesoscale ensembles developed at MRI for the B08RDP Tier-1 ensemble prediction. Section 4 presents the design of the experiments. Section 5 shows the results of the comparison of the five perturbation methods. Section 6 discusses the vertical structures of the energy norms and their time evolutions in the ensemble prediction. Characteristics of horizontal and vertical distributions of the initial perturbations by the five methods are also discussed. Summary and concluding remarks are given in Section 7.

2. The WWRP B08FDP/RDP project

The B08FDP/RDP was an international research project of WWRP conducted in conjunction with the Beijing 2008 Olympic Games. The plan was first proposed by the China Meteorological Administration (CMA) in September 2003 as a research project which succeeded the Sydney 2000 Forecast Demonstration Project (Sydney 2000FDP; Keenan et al., 2003). In October 2004 the project was endorsed at the seventh session of the WWRP Science Steering Committee. The B08FDP/RDP was divided into two components: a FDP component for very short-range forecasting of up to 6 h based on nowcasting, and a RDP component for short-range forecasting of up to 36 h based on mesoscale EPSs. MRI participated in the RDP component in collaboration with the Numerical Prediction Division (NPD) of JMA. The RDP component was further divided into two parts: Tier-1, consisting of mesoscale ensemble prediction by regional models with a horizontal resolution of 15 km, and Tier-2, case studies using cloud-resolving models with horizontal resolutions of less than 3 km. Prior to the intercomparison performed in summer 2008, preliminary experiments testing the data transfer and the EPSs of the project participants were carried out in 2006 and 2007.

The 2008 B08FDP/RDP experiment was conducted over about one month in summer 2008 to coincide with the period of the Beijing Olympic Games, which took place from 8 to 24 August 2008. In the Tier-1 ensemble experiment, the six participants [MRI/JMA, NCEP, MSC, ZAMG, the National Meteorological Center (NMC) of CMA and the Chinese Academy of Meteorological Sciences (CAMS)] were requested to run their ensemble predictions for a forecast time (FT) of up to 36 h, starting every day at 12 UTC. The results were interpolated into verification grids with a resolution of 0.15° over a common verification domain (105–125°E, 30–45°N). Duan et al. (2009) provide a detailed overview of B08RDP, and Kunii et al. (2011) have reported the results of the international EPS intercomparison. Details of MRI and JMA's activities in B08FDP/RDP have been published as an MRI Technical Report (Saito et al., 2010a).

3. Initial perturbation methods

Five initial perturbation methods were developed at MRI in collaboration with NPD/JMA and applied in the B08RDP Tier-1 mesoscale EPS. Considering different features of horizontal and vertical distributions of initial perturbations by the five methods (see section 6.2), we did not apply the same scaling factor but employed different scaling manners in the five methods. The details are as follows.

3.1. WEP method

The WEP method employs perturbations of the forecast of JMA's operational one-week EPS, which uses a singular vector method in a global model. This perturbation method was regarded as a reference method for the comparison of initial perturbation methods and was also used to supply lateral boundary perturbations to the mesoscale ensemble predictions, except for that employing the GSV method. We employed 6-hourly pressure plane (12 levels) forecast grid point values (GPVs) of JMA's

one-week EPS, which covers the Regional Specialized Meteorological Center (RSMC) Tokyo responsible area (90–180°E, 0–71.5°N) with a horizontal grid interval of 1.25°. The data (hereafter, JMA EPS GPV) are processed at NPD/JMA in Tokyo once a day at the initial time of 12 UTC and transferred to MRI in Tsukuba through an exclusive line.

As described later in Section 4, an NHM with a horizontal resolution of 15 km and 40 vertical levels (NHM L40) is employed in the ensemble prediction, but the highest JMA EPS GPV level is 100 hPa, which is lower than the top of NHM L40 (22.1 km, ∼40 hPa). The procedure to prepare the initial perturbation is as follows. First, JMA EPS GPV data are interpolated to model planes of a 32-level NHM (NHM L32; the model top is at 13.8 km, ∼160 hPa), and perturbations are extracted by subtracting the interpolated field of the control run from those of the perturbed runs. Then, the perturbations are normalized and added to the initial conditions of the NHM L40 control run. Next, the perturbation at the top of NHM L32 is vertically extrapolated to the highest eight levels of NHM L40. At the kz-th level, the amplitude of perturbation is multiplied by the coefficient $c1$:

$$c1 = \frac{1 - \cos\{\pi^*(kz - 32)/8\}}{2}, \tag{1}$$

so that the perturbation amplitude becomes zero at the top of NHM L40.

Because the timing of the file transfer of JMA EPS GPV from NPD/JMA to MRI is too late to conduct a near-real-time EPS run of B08RDP, instead of the initial conditions of the JMA EPS, 24-h forecast of the JMA EPS from 12 UTC 1 d before was used for the experiment. The amplitudes of JMA EPS GPV were adjusted to take into account the analysis errors of the initial conditions of NHM L40 given by JMA's hydrostatic 4D-VAR system (Meso 4D-VAR; Koizumi et al., 2005). Normalization coefficients were determined so that the root mean square values of the perturbations at each model level did not exceed the prescribed upper limits of the standard errors of the Meso 4D-VAR analysis [0.7 hPa for mean sea level pressure, 1.8 m/s for horizontal winds (U and V), 0.7 K for potential temperature, and 15% for relative humidity]. These values are about 70% of the magnitudes of JMA's operational Meso 4D-VAR's statistical background errors. Although this 'capping' may deform the structures of the 24-h forecasted perturbations of JMA's global EPS, the resulting perturbations do not badly degrade the dynamical balance of the NHM initial conditions (as shown the RMSEs of the ensemble mean in Fig. 7). A similar method has been used in mesoscale ensemble experiments for a heavy rainfall event in northern Japan (Saito et al., 2006b). Seko et al. (2009) also applied this method to estimate potential parameters for Tornado formation in Japan and obtained a good result.

Figure 1 depicts the procedures used to prepare the initial conditions by the WEP method. To prepare the lateral boundary

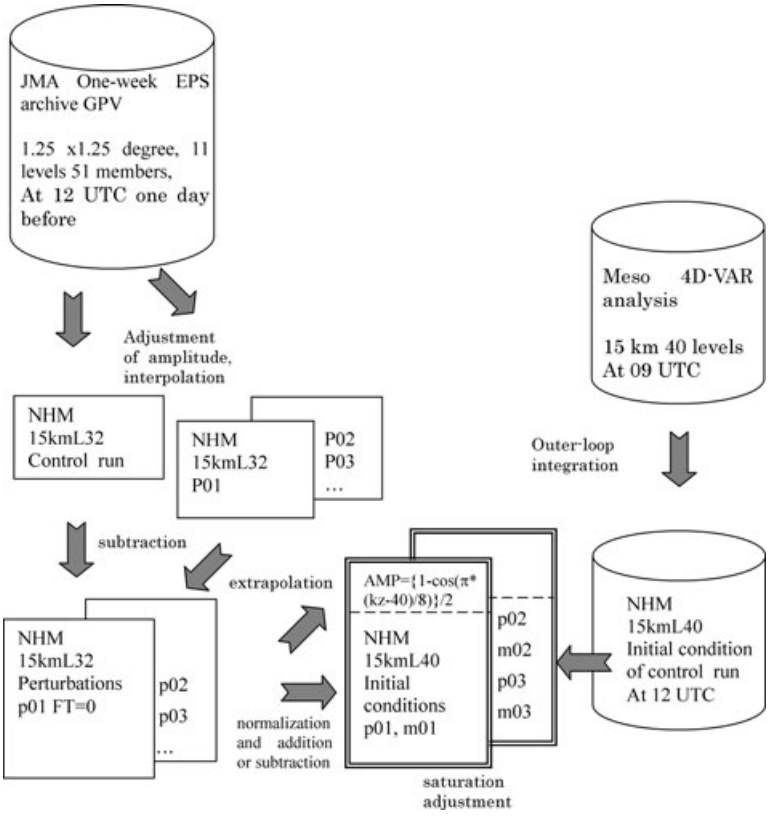

Fig. 1. Schematic chart of the procedures used for preparing initial conditions by the WEP method.

perturbations, a similar procedure was applied to the 6-hourly JMA EPS GPV data of 1 d previously, but instead of applying the same limiters in the initial perturbations, the amplitudes of the global EPS forecast perturbations were adjusted by taking into account the difference in initial times between the global and mesoscale EPSs. Details of the procedure are given in Section E-5 of Saito et al. (2010a). Similar procedures were employed in the ensemble prediction experiment for the Myanmar cyclone Nargis (Saito et al., 2010b), but in that case, instead of the 6-hourly RSMC Tokyo EPS GPV data of 1 d previously, other 12-hourly global EPS GPV data of JMA for the same day were used.[1]

3.2. GSV method

The GSV method is based on the targeted GSV method of the typhoon EPS at JMA (Yamaguchi et al., 2009). To calculate the SVs, a tangent linear model (TLM) and its ADjoint Model (ADM) of the JMA global 4D-VAR system (Kadowaki, 2005) were used. In the B08RDP application, the horizontal resolutions and vertical levels of TLM and ADM were T63 (about 200 km) and L40. The models consist of a full dynamical core and physical processes, including vertical diffusion, gravity wave drag, large-scale condensation, long-wave radiation, and deep cumulus convection. The chosen optimization time of the SVs was 24 h. The moist total energy (*TE*) norm (Barkmeijer et al., 2001) was used:

$$TE = \frac{1}{2}\int_0^1 \int_s \left(\nabla\Delta^{-1}\zeta_x \cdot \nabla\Delta^{-1}\zeta_y + \nabla\Delta^{-1}D_x\nabla\Delta^{-1}D_y + \frac{c_p}{T_r}T_xT_y + w_q\frac{L^2}{c_pT_r}q_xq_y \right) \mathrm{d}S \left(\frac{\partial p}{\partial \eta}\right) \mathrm{d}\eta + \frac{1}{2}\int_S R_dT_rP_r \ln\pi_x \cdot \ln\pi_y \mathrm{d}S, \quad (2)$$

where ζ, D, T, q and $\ln\pi$ are the vorticity, divergence, temperature, specific humidity and logarithm of the surface pressure, respectively. c_p *is* the specific heat at constant pressure, R_d the gas constant for dry air, and L the latent heat of condensation. $T_r = 300$ K is the reference temperature, $P_r = 800$ hPa is the reference pressure, and w_q is a weight constant for moist energy. $w_q = 1.0$ was the value employed in JMA's operational typhoon EPS,[2] whereas $w_q = 0.6$ was employed in the preliminary experiment in 2007 to increase the initial perturbation of specific humidity. On the basis of the results of the 2007 experiment, $w_q = 0.3$ was chosen in the 2008 B08RDP experiment (this study) to further increase the initial perturbation of specific humidity.

As a tuning in B08RDP, the target area of the GSV was shifted southeastward from the common verification area (from 105–125°E, 30–45°N to 110–130°E, 27–42°N) to capture the northward water vapour flux at lower levels in the southern part of the model domain and to improve the ensemble spreads in the early stage of the forecast period.

To obtain consistent initial and lateral boundary perturbations for NHM, both perturbations were given by initial conditions and the forecast of the global spectral model (GSM) of JMA. The first five SVs were added to the initial conditions (JMA global analysis) of GSM (T63L40), and time integration of GSM was performed for 36 h in the control run and in the five ensemble members, using only positive perturbations. The SV amplitudes were adjusted so that none of the maximum values of the five variables (temperature, specific humidity, x and y components of the horizontal velocity, and surface pressure) exceeded the typical GSM forecast error estimated from the statistics of the 2007 experiment.[3] Then, forecasted 3-hourly GSM model-plane data were interpolated into the NHM model planes. The initial and lateral boundary perturbations for the mesoscale EPS were obtained from the differences between the control run and the perturbed runs of the GSM EPS. Then, the perturbations were added to (or subtracted from) the initial and boundary conditions of the NHM control run. Super-saturation was removed from the perturbed fields. For more details of the GSV method and its modifications in B08RDP, see Hara (2010a).

3.3. MSV method

In the MSV method, the SVs are calculated using the TLM and ADM of the JMA non-hydrostatic mesoscale 4D-VAR system (Honda et al., 2005). In TLM and ADM, some parts of the non-linear model are simplified; only the large-scale condensation and the moisture convective adjustment are used in moisture processes. To solve the eigenvalue problem, the Lanczos algorithm (Simon and Parlett, 1980) with Gram-Schmidt re-orthogonalization was adopted. The TE norm including the moisture term (Ehrendorfer et al., 1999) is

$$TE = \frac{1}{S}\int_{z_1}^{z_2}\int_s \frac{1}{2}\rho\left[u'^2 + v'^2 + w'^2 + w_t\frac{c_p(\theta)'^2}{\Theta} + RT_r\left(\frac{p'}{p_r}\right)^2 + w_q\frac{L^2}{c_pT_r}q'^2\right]\mathrm{d}S\mathrm{d}z, \quad (3)$$

where ρ is density; u, v and w are wind components; θ, p and q are potential temperature, pressure, and the mixing ratio of water vapour, respectively; and C_p is the specific heat at constant

[1] Saito et al. (2010b) did not use the RSMC Tokyo EPS data set for the Nargis EPS because its coverage area (90–180°E, 0–71.5°N) does not include all of the Bay of Bengal. In the case of the Nargis simulation, we did not have to finish the computation within a specific time; thus, we could use the data at the initial time of the mesoscale model simulation without considering the timing of the file transfer from NPD/JMA to MRI.

[2] $w_q = 0.04$ is used in JMA's operational 1-week EPS.

[3] In the 2007 preliminary experiment, the amplitude of the initial perturbation was fixed at the mean value of the estimated standard error of the Meso 4D-VAR analysis, but the ensemble spread was insufficient (see section D-4–2 of Saito *et al.* (2010a)).

pressure. $\Theta = 300$ K, $Tr = 300$ K and $Pr = 10^5$ hPa are the reference values of potential temperature, temperature, and pressure. R is the gas constant, L the latent heat constant, and $w_t = 3.0$ and $w_q = 0.5$ are the weights of the potential temperature and water vapour mixing ratio terms, respectively. These weights were determined so that the composition ratio of individual terms in the initial norm would be equivalent to that calculated using standard analysis errors.

Because mesoscale model-based SVs sometimes show an unnatural 'peak' value for a particular parameter such as qv, a simple normalization based on the maximum value only tends to rescale the perturbation amplitude unreasonably small. To overcome the problem, a two-step approach to the normalization of the initial perturbations was used. In this approach, first, a normalization factor to the initial perturbations is determined so that one of their maximum values becomes any of $p = 1.0$ hPa, $(u, v) = 6.0\,\mathrm{m\,s^{-1}}$, $\theta = 3.0$ K, and $qv = 3.0\,\mathrm{g\,kg^{-1}}$. Then, the maximum peak size limit was set to three times the analysis error. In other words, if the absolute value of the initial perturbation exceeded three times the standard deviation of the analysis error, the amplitude was reduced to equal three times the analysis error. The analysis errors used here were functions of height, based on the background errors of the JMA Meso 4D-VAR, which were calculated statistically for a summer season. These procedures may deform the structure of each SV, but we confirmed that the modified initial perturbations could grow in the non-linear model as steadily as in the linear model. Furthermore, the variance minimum method (Yamaguchi et al., 2009) for reorganizing initial perturbations was utilized. For more details of the MSV method and its modifications in B08RDP, see Kunii (2010).

3.5. MBD method

The MBD method employs a self-breeding cycle with NHM. To evaluate the magnitude of the bred perturbations, the moist TE norm by Barkmeijer et al. (2001), similar to eq. (3), is employed. In B08RDP, referring to the JMA global EPS, the values of $\Theta = 300$ K, $P_r = 800$ hPa and $w_q = 0.1$ were used, and the norm was computed below the height of 5.3 km above ground level.

Six-hourly self-breeding cycles with a horizontal resolution of 40 km were conducted (Fig. 2a). In the 2007 preliminary experiment, self-breeding cycles were conducted with a horizontal resolution of 15 km as in the extended ensemble run, but bred vectors contained small-scale (high wavenumber) perturbations, which tended to shrink in a short period (see section D-4 of Saito et al., 2010a). In the 2008 intercomparison, we unified the horizontal resolutions of the three mesoscale model-based initial perturbation methods (MSV, MBD and LET) to 40 km. This unification made implementation of the lateral boundary perturbations easier, and contributed to the elimination of such small-scale perturbations.

The moist TE norms are computed from the differences between the control runs and perturbed runs, and the bred perturbations of all prognostic variables except four-levels soil temperatures are normalized every 6 h. The normalization coefficients are determined by the square root of the ratios of the TE norms of perturbed runs to the standard norm, which is computed by using the prescribed values of the model variables (0.35 hPa for mean sea level pressure, $1.0\,\mathrm{m\,s^{-1}}$ for U and V, 0.4 K for θ and 5% for relative humidity). These values are about 50% of the magnitudes of JMA's operational Meso 4D-VAR's statistical background errors (Koizumi et al., 2005).

Perturbations from the operational global one-week EPS of JMA at 12 UTC 2 July 2008 were used as the initial seed of the ensemble perturbation, and 6-h breeding cycles were performed sequentially throughout the experimental period using JMA's operational global 4D-VAR analyses without orthogonalization. A sequential breeding cycle without orthogonalization may cause conversion of bred vectors to a single Liapunov vector, but this problem was avoided by inclusion of the lateral boundary perturbations in the breeding cycles (Saito et al., 2009). Water vapour supersaturation was removed for each perturbed member at the initial times of all breeding cycles. To save computational cost, forecasts in breeding cycles were performed by using the warm rain process.

Five bred vectors were interpolated and added to the initial condition of the control run as the incremental perturbation to make five positive ensemble members with a horizontal resolution of 15 km. Additionally, the interpolated bred vectors were subtracted from the initial condition of the control run to make five negative ensemble members. These negative members are almost symmetric to positive members, but the ensemble mean is slightly modified owing to the saturation adjustment.

3.6. Local ensemble transform Kalman filter (LET) method

The LET method uses the ensemble transform of LETKF (Hunt et al., 2007) as the initial perturbation generator. NHM-LETKF (Miyoshi and Aranami, 2006) was modified for B08RDP. The procedures of the LET method are outlined in Fig. 2b. As in the MBD method, 6-hourly forecast-analysis cycles with NHM (40 km, L40) are performed, with 20 ensemble members. Also as in the MBD method, the warm rain process was adopted in the EnKF forecast-analysis cycles to reduce the computation time. An inflation factor of 10% (i.e. 21% covariance inflation) is employed. Localization is performed with the observation localization scales of five grid points in the horizontal and three vertical levels, in which the influence of an observation is cut off when the distance from the observation is larger than the product of the localization scale and 2.0*sqrt(10/3). Surface and upper-level sounding data that have passed JMA's operational quality control procedures are assimilated by LETKF. The

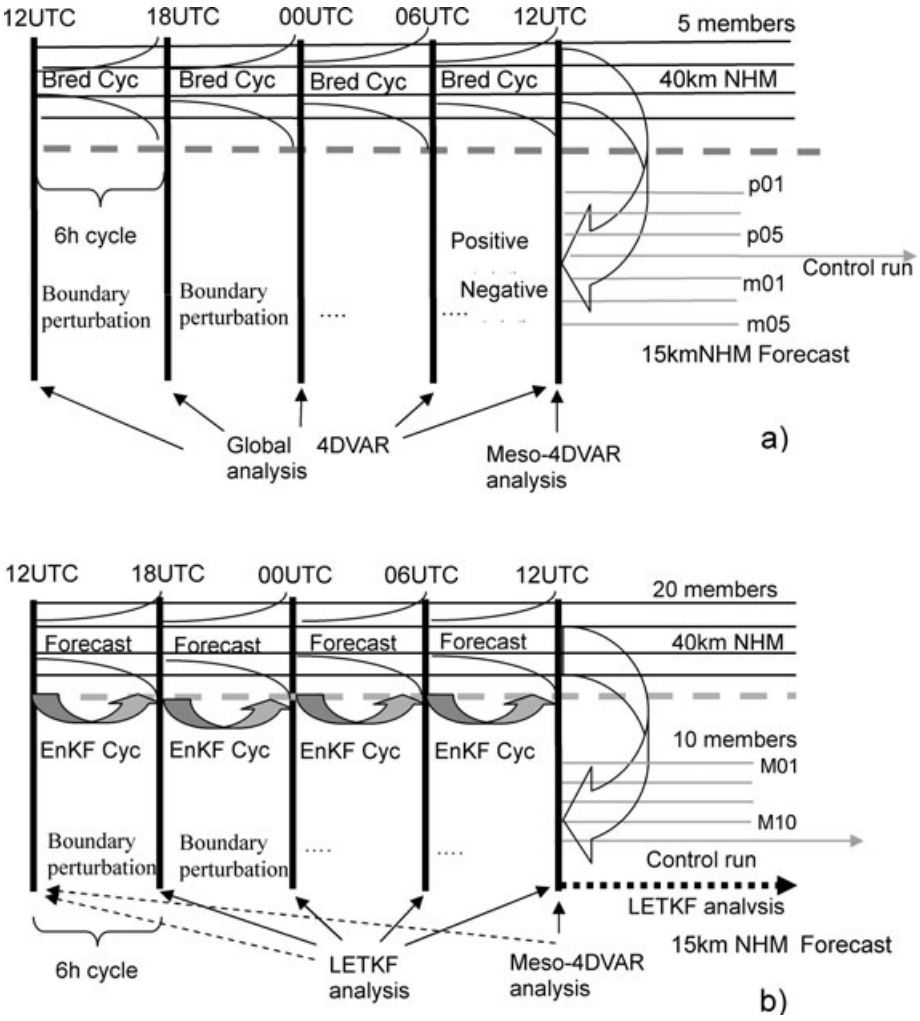

Fig. 2. Schematic diagrams of (a) the MBD method and (b) the LETKF method.

ensemble mean of the assimilation results is re-centred around the analysed fields obtained by the Meso 4D-Var system at 12 UTC in each EnKF forecast-analysis cycle. Then, 10 members are selected by cluster analysis, and perturbations produced by LETKF are extracted. Initial conditions of ensemble members are produced by adding the perturbations of the selected 10 members to the analysed fields obtained by the Meso 4D-Var system.

The following modifications were made to Miyoshi and Aranami's (2006) NHM-LETKF: (1) rigorous treatments of momentum and pressure (e.g. accurate treatment of air density); (2) implementation of vertical hybrid coordinates; (3) saturation adjustment of the initial field; (4) removal of local patches and (5) implementation of the lateral boundary perturbation in forecast-analysis cycles. Local patches that originated from the local EnKF (Ott et al., 2004) were removed by the method of Miyoshi et al. (2007), which was implemented for NHM-LETKF by Fujita (2009). With the above modifications, the ensemble spread was increased by about 5% and the analysis fields became closer to the next day's initial condition (Seko, 2010).

A cluster analysis procedure was also applied in the LETKF in B08RDP to choose the 10 members from the 20-member ensemble perturbations. The variation of normalized energy in the lower atmosphere was obtained by the following equation:

$$E = \int_{20\,\mathrm{m}}^{1.46\,\mathrm{km}} \int\int \left\{ \frac{u^2 + v^2 - (u^2 + v^2)_{\min}}{(u^2 + v^2)_{\max} - (u^2 + v^2)_{\min}} + \frac{\theta^2 - \theta^2_{\min}}{\theta^2_{\max} - \theta^2_{\min}} + \frac{q_v^2 - q_{v\,\min}^2}{q_{v\,\max}^2 - q_{v\,\min}^2} \right\} \mathrm{d}S\,\mathrm{d}z, \tag{4}$$

where u, v, θ and q_v are horizontal wind components, potential temperature, and the mixing ratio of water vapour, respectively. $()_{\min}$ and $()_{\max}$ are the maximum and minimum values among all ensemble members. The integration was performed over the whole horizontal domain and from 20 m to 1.46 km in the vertical domain. Distances between each cluster and the ensemble mean were evaluated in terms of the variation of normalized energy

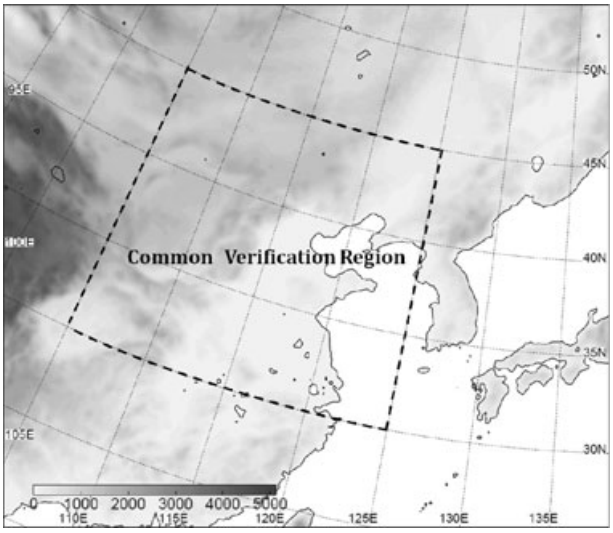

Fig. 3. Domain of mesoscale ensemble prediction. The fan-shaped area over eastern China indicates the common verification area in B08RDP.

in the cluster analysis. Selection of ensemble members started with the one in the largest group that was the farthest from the ensemble mean.[4]

A detailed comparison of MBD and LET methods with and without the lateral boundary perturbations has been given by Saito et al (2009). Seko et al. (2011) described mesoscale and cloud-resolving ensemble forecast experiments for a heavy rainfall event in western Japan using NHM-LETKF.

4. Design of experiments

Performances of the five initial perturbation methods (WEP, GSV, MSV, MBD and LET) were compared by applying their perturbations to the MRI/JMA's mesoscale EPS in B08RDP. The model domain (Fig. 3) is covered by 232 × 200 gridpoints with a horizontal resolution of 15 km. The number of vertical levels is 40, with the lowest level at 20 m above ground level, and the depth of the layers stretches from 40 to 1180 m with increasing height. The initial conditions and lateral boundary conditions of the control run were prepared by using the Meso 4D-VAR analysis over the Beijing area (Kunii et al., 2010) and the forecast of JMA's high-resolution operational global model (Japan Meteorological Agency, 2007), respectively.

A three-ice bulk cloud microphysics scheme that predicts cloud water, rain, cloud ice, snow and graupel and a Kain-Fritsch convective parametrization scheme were included as the moist processes. The Mellor-Yamada-Nakanishi-Niino's level 3 turbulent closure model (MYNN3; Nakanishi and Niino, 2004) as implemented by Hara (2010b) was used. Surface momentum, heat, and moisture fluxes over the sea were computed using Beljaars and Holtslag's (1991) scheme. These physical processes, including the atmospheric radiation scheme, are basically the same as those in JMA's operational mesoscale model (MSM; Saito et al., 2006a, 2007; Japan Meteorological Agency, 2007). Small perturbations in initial soil temperatures corresponding to the lowest level atmospheric temperatures were applied to the ensemble members, but they had almost no impact on the ensemble spread (see section D-5 in Saito et al., 2010a).

A few points in the operational MSM were modified for the extended forecast in B08RDP with a horizontal resolution of 15 km: first, the lower limit of soil humidity was modified to attenuate the overestimation of the surface sensible heat on very hot days; second, the parameters in the Kain-Fritch scheme such as the autoconversion threshold from condensed water to rain (kf_thresh) and convection life times (cu_lifetime_min and shallow_lifetime) were modified. Details of the modifications of NHM in B08RDP are given by Saito and Hara (2010).

The 36-h extended ensemble forecast with 11 members was assessed for two initial times, 12 UTC on 3 July and 12 UTC on 4 July 2008, and RMSEs, ensemble spreads, and precipitation were verified using average values of the two 36-h EPSs. Although the period of the comparison (2.5 d, from 12 UTC on 3 July to 00 UTC on 6 July) was possibly not long enough for full-scale validation of the initial perturbation period, given the forecast range (36 h), this kind of comparison using five different initial perturbation methods is expensive, and we therefore could test only two cases. Specifications of the mesoscale EPS of MRI/JMA are listed in Table 1, along with comparisons with the operational mesoscale NWP at JMA.

During the period of the comparison, a synoptic low-pressure system passed over eastern China. Figure 4 shows the forecast of the control run of the extended EPS at initial time 12 UTC, 4 July 2008. The low-pressure system was located southwest of Beijing at 15 UTC, 4 July (Fig. 4a; FT = 3), and moved northeastward bringing rainfall to the area northeast of Beijing on 5 July (Fig. 4b; FT = 24).

5. Comparison of the five methods

5.1. Ensemble spread and RMSE

Performances of the five initial perturbation methods (WEP, GSV, MSV, MBD and LET) were verified by checking their ensemble spreads and the RMSEs of the ensemble means against the Meso 4D-VAR analysis at the valid time.

Figure 5 shows the time sequence of the ensemble spreads of surface variables (Psea, sea level pressure; Vs, surface (10 m) meridional wind; Ts, surface (2 m) temperature; RHs, surface (2 m) relative humidity; and RR3H, 3-h accumulated rainfall amount) averaged for two EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008. Spreads of surface variables

[4] In fact, this clustering procedure was not decisive for the performance of EPS using LETKF. The ensemble transform makes each ensemble equal-likely in the model coordinates, though the distances between the ensemble members are slightly modified in the energy norm space.

Table 1. Specifications of the B08RDP MEP system of MRI/JMA

Forecast model	Operational MSM at JMA	B08RDP
Horizontal grid	721 × 577 ($\Delta x = 5$ km)	232 × 200 ($\Delta x = 15$ km)
Time step	24 s	60 s
Forecast time	33 h (03, 19, 15, 21 UTC) 15 h (00, 06, 12, 18 UTC)	36 h (12 UTC)
Vertical grid	Generalized hybrid coordinates, 50 levels	Generalized hybrid coordinates, 40 levels
Number of members	Only deterministic run	11 members
Initial condition of the control run	JMA non-hydrostatic 4DVAR analysis over Japan area	Meso 4DVAR analysis for Beijing area
Lateral boundary condition of the control run	JMA GSM forecast (TL959 L60), 1 hourly	JMA GSM forecast (TL959 L60), 3 hourly

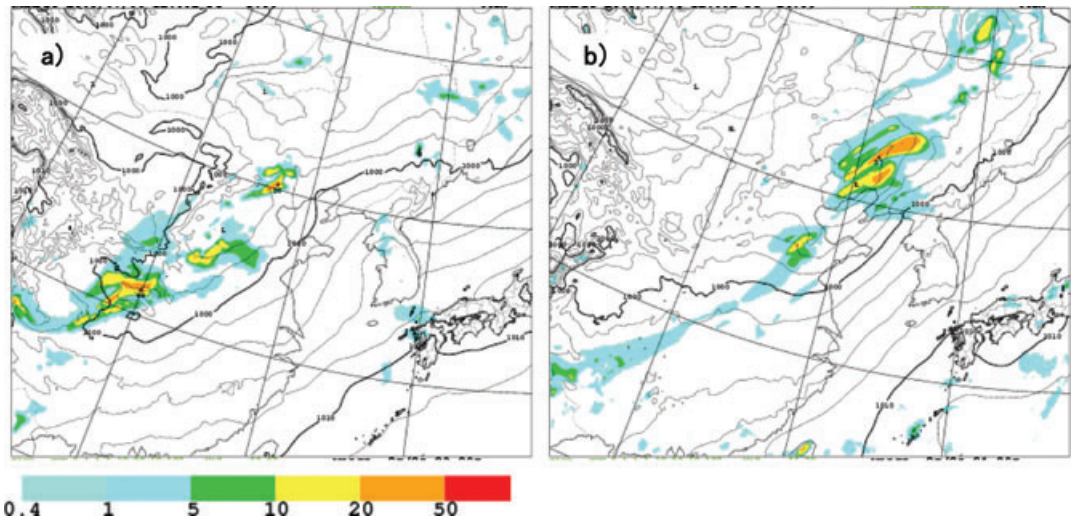

Fig. 4. Sea level pressure (contours) and accumulated 3-h precipitation (colour scale) predicted by the control run. Initial time is 12 UTC, 4 July 2008. The colour bar indicates precipitation intensity in mm. (a) FT = 3. (b) FT = 24.

except Psea and RR3H show distinct diurnal changes. Since the initial time 12 UTC corresponds to 8 pm local standard time in Beijing, spreads of surface wind, Ts, and RHs show daytime peaks, from FT = 12 to FT = 24, whereas spreads of Psea generally increase steadily with time.

WEP spreads have relatively large amplitudes at the initial time (FT = 0) and increase slowly. Spreads of surface wind and RHs decrease at FT = 3, after the initial spin-up. GSV spreads are relatively smaller than those of WEP, MBD, or LET at the initial time (FT = 0) and then grow steadily throughout the 36-h FT. However, the increase in the spread of surface precipitation is slow during the first 3 h. By FT = 18, GSV spreads are largest among the five methods.

MSV spreads at FT = 0 are the smallest among the five methods but grow rapidly after the spin-up. Surface precipitation (RR3H) grows particularly rapidly during the first 6 h, reaching 3.75 mm at FT = 6, the largest value among the five methods. However, the spread of RR3H decreases subsequently and becomes comparable to those obtained with the other methods after FT = 12. This behaviour of the precipitation spread after the model spin-up in MSV may be partially come from the adjustment process of the atmospheric unstableness in the initial condition. Nevertheless, as shown in subsection 6.1, the energy norm of the MSV method increases steadily and keeps relatively large growth rate even after FT = 3. The tendencies of the MBD and LET spreads are similar to each other: values at FT = 0 are relatively large and grow slowly. These differences between SV and BGM perturbations are consistent with Yamaguchi and Majumdar (2010), where the global ensemble spreads by ECMWF SV and NCEP BGM were compared. Generally, growths of spreads in LET are slightly more sluggish than those in MBD.

Figure 6 shows the evolution of the ensemble spreads of the variables at 500 hPa (Z, geopotential height; V, meridional wind; T, temperature; RH, relative humidity). No distinct diurnal changes are seen in these variables. The characteristics of the spreads at 500 hPa obtained with the five initial perturbation methods are similar to those at the surface (Fig. 5). WEP spreads have relatively large amplitude at the initial time (FT = 0) and increase slowly. Generally, initial spreads for SV methods (GSV

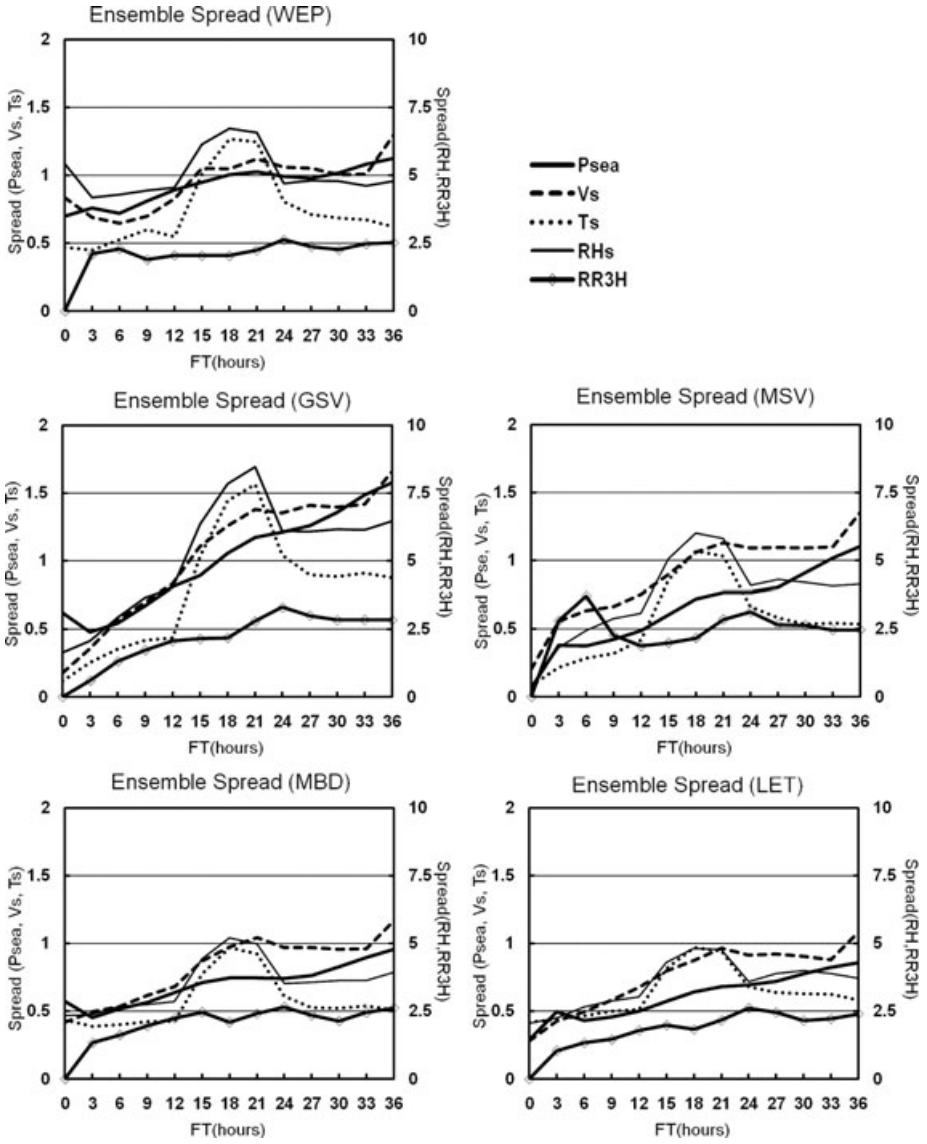

Fig. 5. Ensemble spreads of surface variables for the five initial perturbation methods (WEP, GAV, MSV, MBD and LET) in the verification area averaged over two EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008. The units on the left vertical axis are hPa for sea level pressure (Psea), m s^{-1} for surface (10 m) meridional winds (Vs) and K for 2 m temperature (Ts). The units on the right vertical axis are % for surface (2 m) relative humidity (RH) and mm for 3-h precipitation.

and MSV) are smaller than those for other methods but their amplitudes grow rapidly. GSV spreads grow steadily throughout the 36-h forecast range. The large differences at FT = 36 between the GSV spreads and those of the other four methods probably reflect differences in the lateral boundary perturbation methods; lateral boundary perturbations in WEP, MSV, MBD and LET were given by JMA's 1-week EPS, whereas GSV lateral boundary perturbations were given by the GSM forecast perturbed by the targeted SV. In the second half of the simulation period (after FT = 18), the effect of the lateral boundary perturbation dominates that of the initial perturbation (Saito et al., 2009). The MBD and LET methods show similar tendencies, but the growth of the LET spreads is slightly more sluggish than that of the MBD spreads.

Figure 7a shows RMSEs of the surface variables at FT = 24 against the initial conditions (4D-Var analysis) at the same valid times for the 2 d of 3–4 July 2008. RMSEs of the control run and those of the ensemble means obtained with the five methods are shown. Error bars show standard deviations of improvements by the ensemble means against the control runs. RMSEs of the ensemble means are smaller than those of the control run for all variables except surface pressure by WEP, which indicates

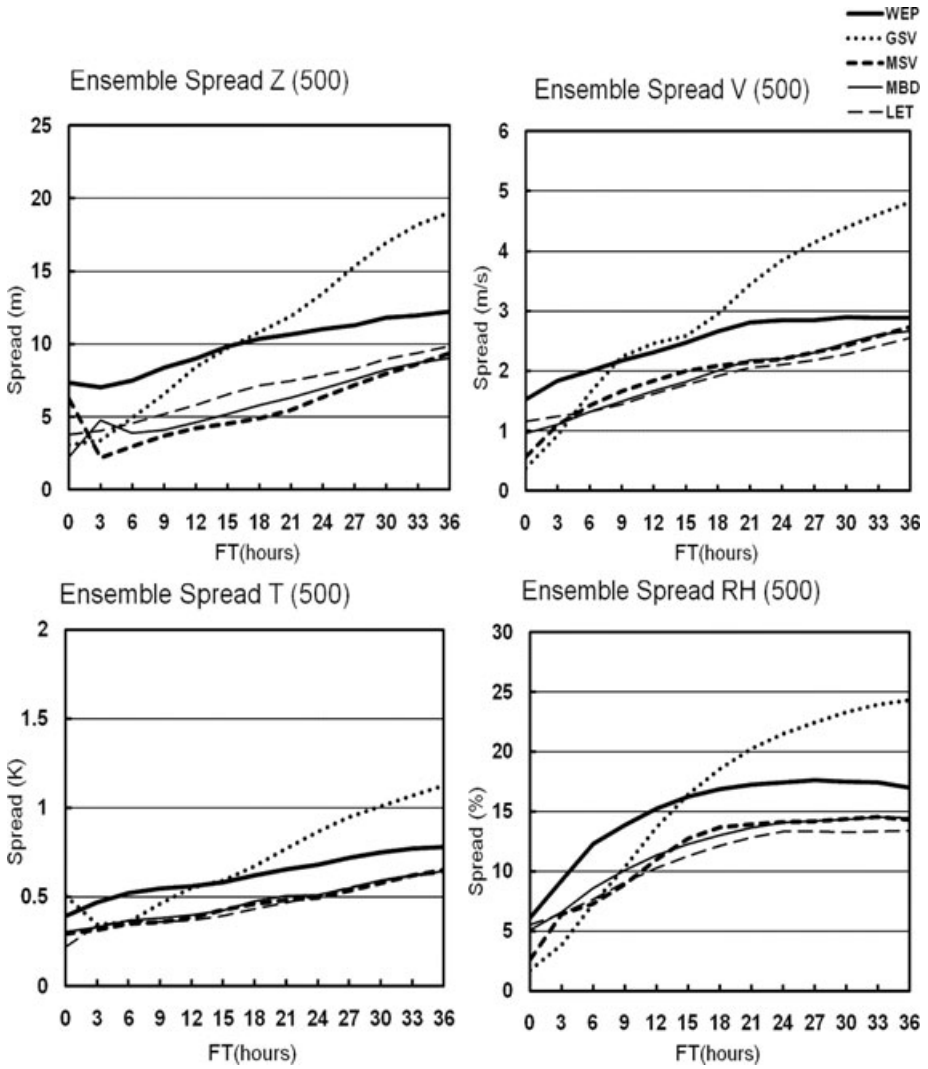

Fig. 6. Ensemble spreads of the 500-hPa level variables for the five initial perturbation methods (WEP, GSV, MSV, MBD and LET) in the verification area averaged over two EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008. Unit on vertical axis is m for height (Z), m s^{-1} for meridional wind (V), K for temperature (T) and % for relative humidity (RH), respectively.

that the ensemble predictions by all initial perturbation methods were conducted properly.[5] WEP RMSEs are larger than those of the other methods in the case of Psea, Vs and Ts. GSV RMSEs are the smallest for all variables. MSV RMSEs are larger than those of MBD and LET, but still smaller than those of WEP for Psea, Vs and Ts. MBD and LET show similar tendencies; their RMSEs are almost the same for Ts, but the RMSEs of LET are marginally smaller than those of MBD for Psea and RHs.

Figure 7b shows the RMSEs for the variables at 500 hPa. WEP does not show a clear advantage against the control run, but the other four methods show an obvious advantage. The GSV RMSE is smallest for V but largest for Z. MSV RMSEs are again slightly larger than those of MBD and LET. MBD and LET RMSEs are fairly comparable. LET RMSEs were minimum for T and RH, but as shown in the error bars, the differences are marginal. The automasking effect of initial perturbation amplitude in the EnKF which reflects the accuracy of the analysis is regarded as the relative advantage of the EnKF against BGM in EPS (Wang and Bishop, 2003), but the advantage of LET against MBD is not obvious in the RMSEs in this figure.

Figure 8 compares the RMSEs and ensemble spreads obtained with the five methods for Psea (Fig. 8a) and Ts (Fig. 8b). In this figure, error bars show their standard deviations. RMSEs of the ensemble means are smaller than those of the control runs except for Psea by the WEP method. Ensemble spreads are

[5] RMSE can be decomposed into the mean error (ME) and the standard deviation error (SDE) (Murphy, 1988), and ME does not depend on the initial perturbation method. Besides, in the B08RDP experiment, model bias for the NHM surface condition was very small (Kunii et al., 2010b).

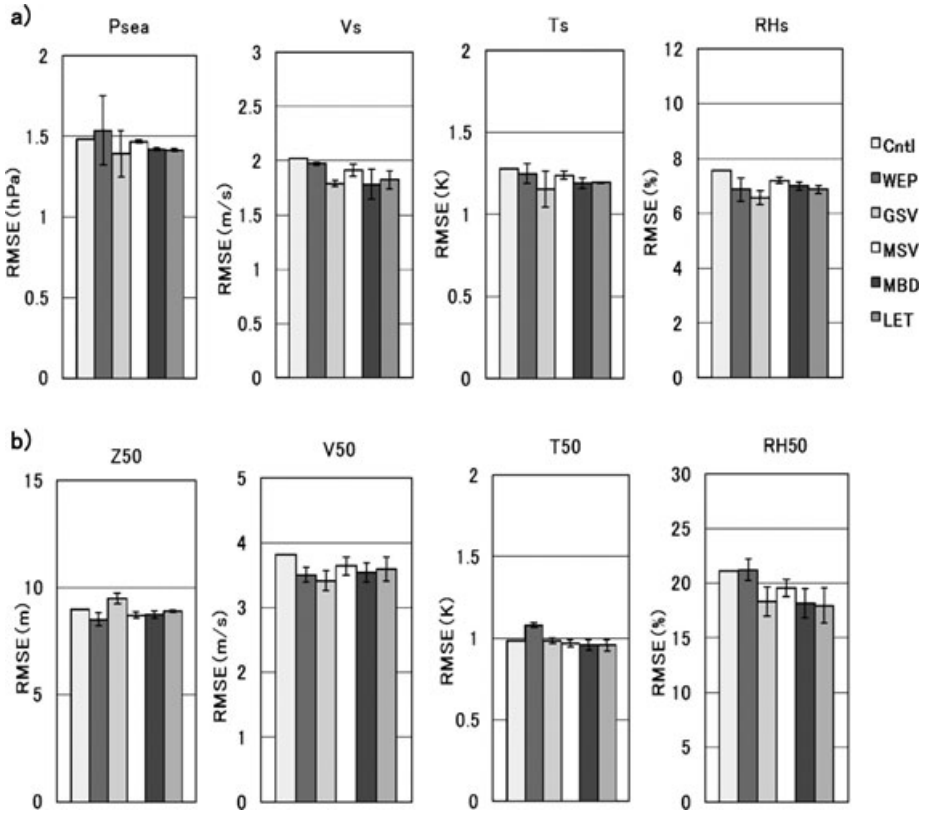

Fig. 7. (a) RMSEs of surface variables (Psea, Us, Vs, Ts and RHs) against the 4D-Var analysis in the verification area at FT = 24 averaged over two EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008. From left to right, control run (Cntl), WEP, GSV, MSV, MBD and LET. Error bars show standard deviations of improvements by the ensemble means against the control runs. (b) Same as (a) but for the variable at 500 hPa (Z, U, V, T and RH).

smaller than the RMSEs for all ensembles. These results indicate that the ensemble forecasts are still underdispersive compared with the model errors. If we examine the ratio of the ensemble spread of the ensemble mean to the RMSE, the GSV method give the best results of 87 and 90% for surface pressure and surface temperature, respectively. Underdispersion of ensemble spreads in surface variables were commonly seen in the other participants' mesoscale EPSs as well, and the results obtained by MRI/JMA's EPS with the GSV method were relatively better than the EPS results obtained by the other participants (Kunii et al., 2011).

5.2. *Performance of precipitation forecast*

We examined the performances of the five initial perturbation methods with regard to the ensemble forecast on precipitation. Figure 9 shows the relative operating characteristics (ROC; Mason, 1982) curves for the comparison period (2.5 d from 12 UTC, 3 July, to 00 UTC, 6 July 2008). For verification, forecast results in the common domain were interpolated to verification grids with a resolution of 0.15° and compared with CMA's surface rain gauge network data (400 synoptic observation stations and 722 automated observation stations (fig. 1 of Kunii et al., 2011).

The five methods exhibit interesting characteristics. For weak to moderate rains, such as 1 mm/6 h and 5 mm/6 h (Figs 9a and b), 'hit rates' (probability of detection, i.e. the proportion of occurrences that were correctly forecasted; Jolliffe and Stephenson, 2003) of the global model-based methods (GSV and WEP) were better (higher) than those of the mesoscale model-based methods (MSV, MBD and LET), whereas 'false alarm rates' (probability of false detection, that is, the proportion of non-occurrences that were incorrectly forecasted) were slightly worse (larger) for high probabilistic thresholds. Maximum detection values (hit rates) by the GSV method were about 90% and 75% for 1 mm and 5 mm per 6 h, respectively. On the other hand, maximum detection values by the LET method were about 75% and 65% for 1 mm and 5 mm per 6 h, respectively. This relatively poor precipitation detection performance of LET was likely due to underestimation of the ensemble spread. For moderate to intense rains, such as 10 mm/6 h and 25 mm/6 h, the situation is reversed. For moderate rains (Fig. 9c), MBD and LET showed a maximum detection rate of 61% and WEP showed a rate of 60%. For intense rains (Fig. 9d), the MBD and MSV methods

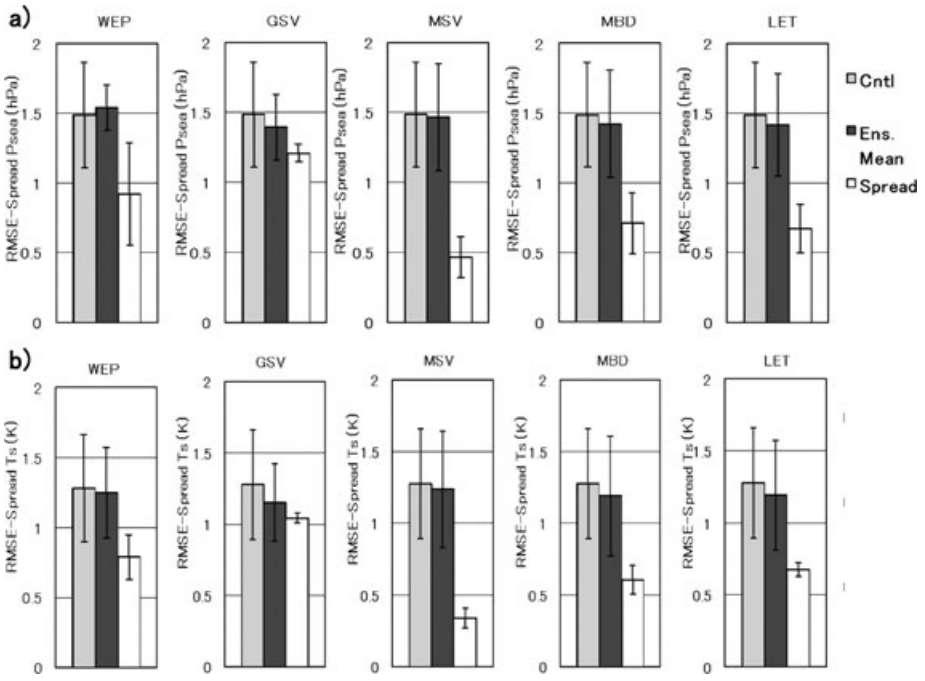

Fig. 8. (a) RMSEs of the control run and the ensemble mean against the 4D-Var analysis, and ensemble spreads of mean sea level pressure (Psea) in the verification area at FT = 24 averaged over two EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008. Error bars show standard deviations. From left to right, WEP, GSV, MSV, MBD and LET. (b) Same as (a), but for surface (2 m) temperature (Ts).

yielded maximum rates of 45% and 41%, respectively, whereas WEP yielded the lowest value of 35%. These results indicate that mesoscale model-based methods (MBD, MSV and LET) tend to be more suitable for detecting intense rains than global model-based initial perturbation methods (WEP and GSV).

This tendency is clearly confirmed by the ROC area skill scores (ROCSSs). Figure 10 shows ROCSSs with thresholds of 0.1, 1, 3, 5, 10, 15, 20 and 25 mm per 6 h. GSV was best (highest) for weak rains of less than 5 mm/6 h, whereas MBD and MSV were more suitable for moderate or intense rain (>10 mm/6 h). A likely reason for this result is that GSV (and WEP) tends to perturb synoptic-scale disturbances, whereas MBD, LET, and MSV tend to perturb mesoscale disturbances, which has more effect on local intense rains.

Figure 11 shows time series of Brier skill scores (BSSs) for precipitation of 1 mm/6 h by the five initial perturbation methods. Averages of two EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008, are depicted. For this weak rainfall, the global model-based methods, WEP and GSV, show relatively better scores than the mesoscale model-based methods (MSV, MBD and LET). WEP shows the best performance up to FT = 15, but its skill score decreases after FT = 18. GSV shows relatively good scores from FT = 18 to 30. The three mesoscale-based methods show similar score evolutions, peaking at FT = 12, then decreasing gradually, but showing second and third peaks at FT = 18 and 33. Roughly speaking, MBD performance among the three is better up to FT = 18, but it shows no clear advantage after that. Since we compared only two cases, the detailed structures of the time series are likely unimportant.

Figure 12 compares BSSs for precipitation thresholds of 0.1, 1, 3, 5, 10, 15, 20 and 25 mm per 6 h. These scores show the same basic tendencies as the ROCSSs (Fig. 10). GSV was best for weak rains of less than 3 mm/6 h, whereas MBD and MSV were more suitable for moderate or intense rains of more than 5 mm/6 h. However, the performance of GSV was good for intense rains of more than 20 mm/6 h, too. It is not clear why the BSS of GSV is relatively better than the ROCSS for intense rains, but its relatively smaller RMSEs than other initial perturbation methods (Fig. 7) may allow rainfall areas to be predicted more precisely through the dynamical balance. On the other hand, WEP and LET loose their skills for intense rains of more than 20 mm/6 h.

Taking into account each method's performance as indicated by the ensemble spread and the RMSEs as well as the precipitation performance, MRI/JMA selected GSV as the initial perturbation method for the 2008 B08RDP experiment.[6]

[6] The results presented in this paper are based on the re-computation for the MBD and LET methods after the near-real-time MEP intercomparisons of B08RDP. In the preliminary test conducted at MRI before B08RDP in summer 2008, there was a bug in the initial perturbation interpolation procedure in the MBD and LET methods, causing their performances to be poor. The flaw was found after the B08RDP experiment, and the two methods were recomputed (see section E-4–6 of MRI technical report; Saito et al., 2010a).

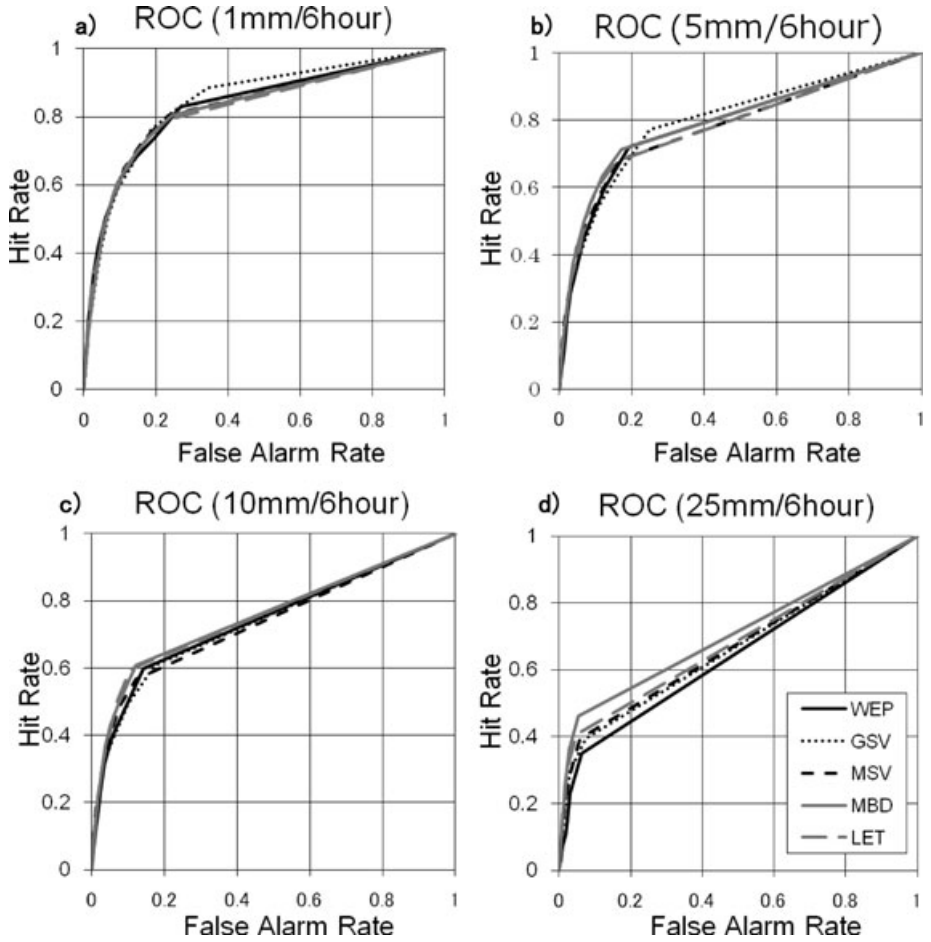

Fig. 9. Relative operating characteristics (ROC) curves for five initial perturbation methods (WEP, GSV, MSV, MBD and LET) averaged over two 36-h EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008. For (a) 1 mm/6 h; (b) 5 mm/6 h; (c) 10 mm/6 h and (d) 20 mm/6 h.

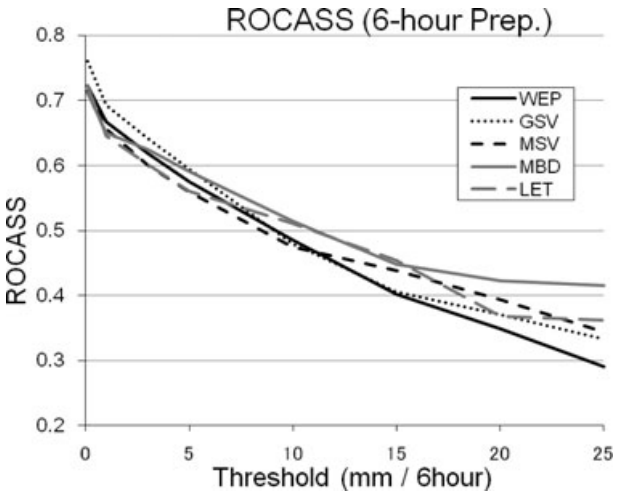

Fig. 10. ROC area skill scores (ROCSS) against different 6-h precipitation intensity thresholds for the five initial perturbation methods over two 36-h EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008.

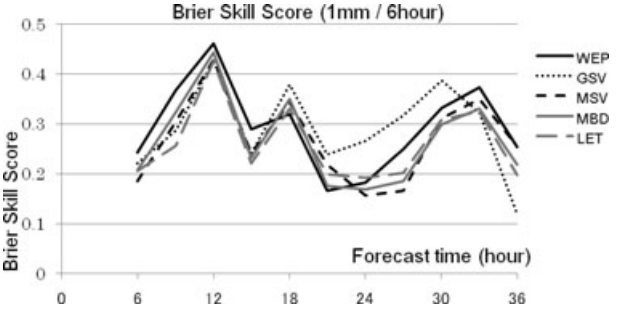

Fig. 11. Time series of Brier skill scores (BSSs) for precipitation of 1 mm/6 h by the five initial perturbation methods. Averages of two EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008.

5.3. *Computational costs*

Another important factor influencing the choice of the initial perturbation method for a practical mesoscale EPS is the computational cost of producing the perturbation. To obtain the initial

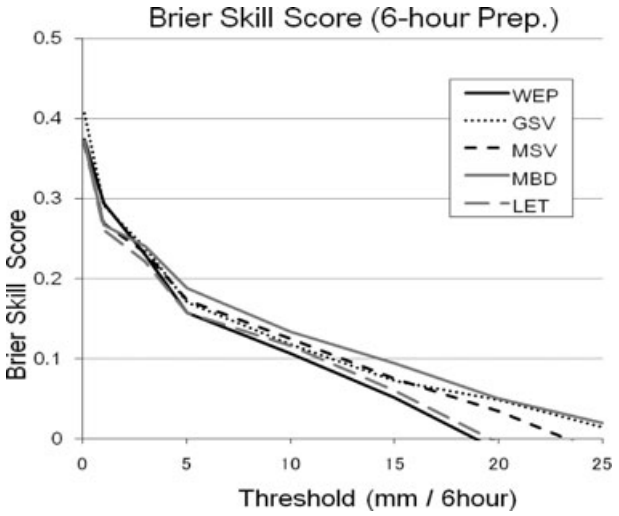

Fig. 12. Same as Fig. 10 but Brier skill scores (BSSs).

perturbation, SV methods use the Lanczos method, which integrates TLM and ADM iteratively, while the breeding method and LETKF integrate the non-linear model in the breeding/EnKF cycles. On the other hand, for perturbations given by downscaling the operational global EPS, the computational cost is negligible.

Table 2 summarizes the characteristics and computational costs of the five initial perturbation methods. To produce five SVs with a horizontal scale of 200 km, the GSV method needed 45 min of elapsed time using 4-node (32-CPU) MPI parallel computing on the NEC SX-6 supercomputer at MRI. The MSV method required 120 min to perform 20 iterations to obtain five SVs with a horizontal scale of 200 km. In the case of the MBD (LET) method, each of the five (20) members in the breeding (EnKF) cycles was integrated using a single CPU of the SX-6, and the total CPU time was converted to the elapsed time if 4 nodes had been used. For four 6-h breeding cycles, the MBD method needed 5 min, and the LETKF method required about 20 min.

These results indicate that SV methods are relatively expensive in computational cost, especially methods based on the mesoscale model (MSV). Note that if the number of ensemble members in the extended EPS becomes larger, the relative computational overhead of LETKF decreases from four times the overhead of MBD to double the overhead, because it creates the same number of independent initial perturbations as the number of ensemble members in the EnKF cycles. In our experiment, only half of the 20 LET perturbations were used for the 10-member EPS and the remaining perturbations were discarded.

6. Discussion

6.1. Evolution of energy norm

In Section 5.1, we verified ensemble spreads of the model variables at surface and 500 hPa in the ensemble predictions from the five initial perturbation methods. Ensemble spread is a basic index to see the properness of perturbations in EPS, but its magnitude depends on the way of perturbation scaling. As shown in section 3, we employed different rescaling manners in the five methods. To check the initial ensemble variance and the growth of perturbations, vertical profiles of the energy norm and their time evolution in the ensemble prediction were examined. The energy norms by Barkmeijer et al. (2001) with $w_q = 0.2$ in the common verification area were computed for five vertical levels (surface, 850, 700, 500 and 250 hPa) 3-hourly until FT = 24 for the two cases with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008.

Figure 13a shows vertical profiles of the energy norms for the five EPSs at the initial time. WEP had large amplitude throughout the troposphere because it used the 24-h forecast of JMA's one-week global EPS. MBD and LET profiles showed similar tendencies, but the MBD norm is relatively larger than that of LET in lower levels, whereas the LET norm is larger at upper levels. Initial energy norms by GSV and MSV were small, especially at surface and at 250 hPa.

At FT = 24, the energy norms at all levels increased for all EPSs (Fig. 13b). The GSV norm became the largest for all levels, showing a remarkable growth at upper levels. MBD and LET profiles showed similar tendencies, but the MBD norm became larger than that of LET for all levels except for 250 hPa. The MSV norm is still the smallest among the five methods. Generally speaking, perturbations in the mesoscale ensemble prediction

Table 2. Specification of the five perturbation methods and computational time

Method	Resolution	Target area	Optimization time Window	Vector numbers	Elapsed time[a]
WEP	T63 (200 km)	Northern Hemisphere (20–90°N)	48 h/24 h	51	–
GSV	T63 (200 km)	SE of verification area 27–42°N, 110–130°E)	24 h	5	45 min
MSV	40 km	Common verification area 30–45°N, 105–125°E)	18 h	5	120 min
MBD	40 km	–	6 h	5	5 min
LET	40 km	–	6 h	20	20 min

[a]NEC SX-6 4 nodes (32 CPU) MBD and LET are estimation from 1 CPU run.

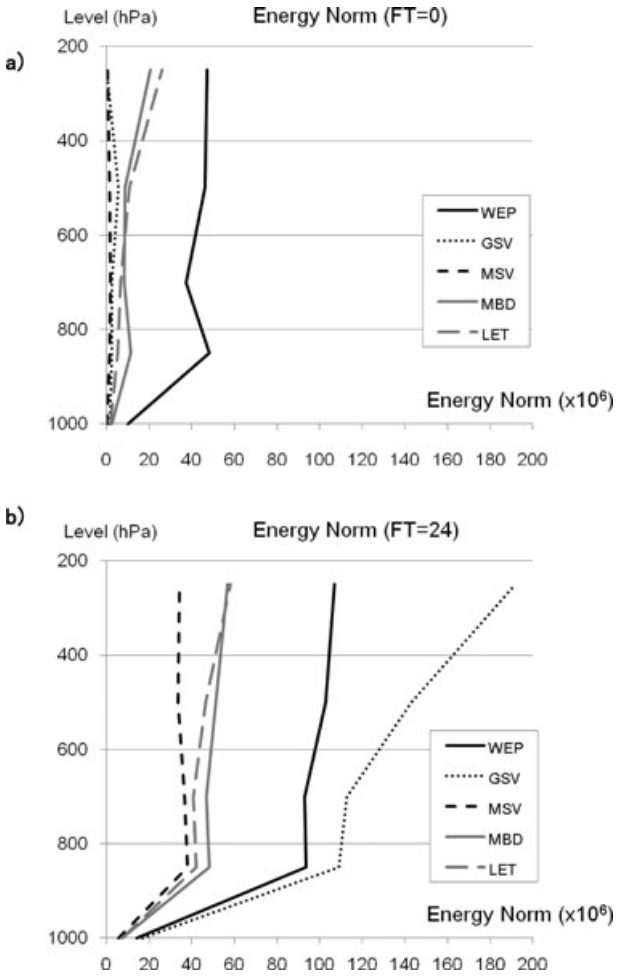

Fig. 13. Vertical profile of the energy norm for the five EPSs. (a) FT = 0. (b) FT = 24.

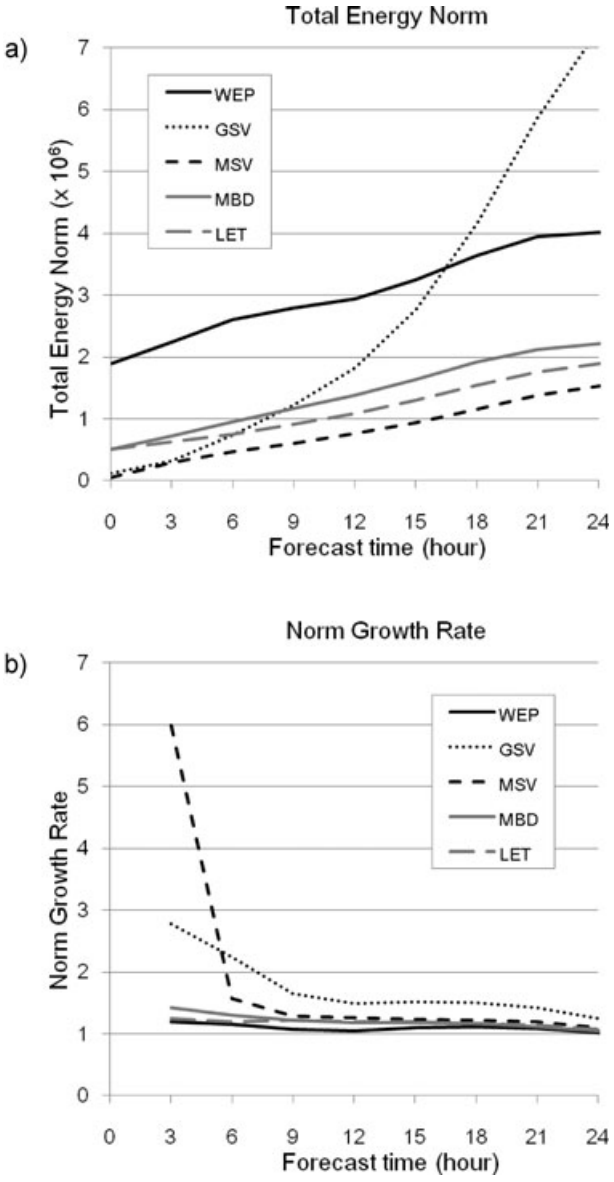

Fig. 14. (a) Time series of the total energy (TE) norms. Averages of two EPSs with initial times of 12 UTC, 3 July and 12 UTC, 4 July 2008. (b) Time series of growth rates of TE norms.

are influenced by the lateral boundary conditions, but at FT = 24, the influence of the lateral boundary perturbations seems to be not so dominant; despite the fact that the four methods (WEP, MBD, LET, and MSV) used the same lateral boundary perturbations, large differences of norm magnitudes were seen in Fig. 13b. Therefore, the large growth of the GSV energy norm in upper levels is predominantly brought not by the lateral boundary perturbations, but by the synoptic scale dynamical growing mode.

Figure 14a shows evolution of the TE norms by the five methods. All methods showed steady growth of the TE norms, but the GSV norm remarkably increased throughout this period. Magnitudes of the initial norms of MBD and LET were almost the same, but the MBD norm was larger than that of LET after FT = 3.

Figure 14b depicts the time series of the norm growth rates. The MSV TE norm showed the minimum magnitude among the five methods (Fig. 14a), but its growth rate was the largest at FT = 3, which means that the MSV method produces the perturbations which grow rapidly. Although the very large growth rate of the MSV norm at FT = 3 decreased soon, the MSV norm growth rate was larger than those of other methods except for GSV (WEP, MBD and LET) even after FT = 6. The GSV norm kept its relatively large growth rate throughout the period, supporting the large increase of the TE norm shown in Fig. 14a. Growth rates of MBD and LET profiles showed similar tendencies, but the MBD growth rate was relatively larger than that of LET until FT = 6. This result suggests that the BGM produces the mesoscale growing modes properly whereas the auto-masking effect in the EnKF and the non-diagonal components in the ensemble transform may reduce the growth rate

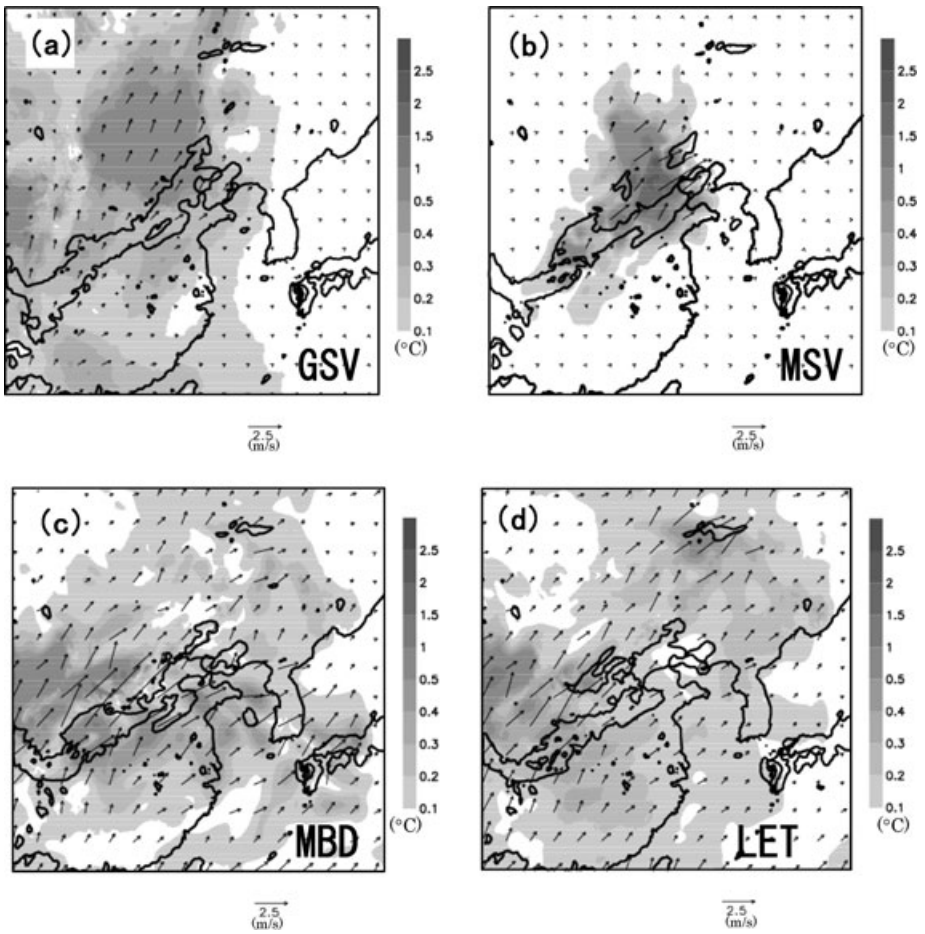

Fig. 15. Initial ensemble spreads of temperature and horizontal wind at the height of 700 hPa at 12 UTC, 4 July 2008. (a) GSV, (b) MSV, (c) MBD and (d) LET. Thick contours indicate the regions where 6-h rainfall exceeded 1 mm at 00 UTC, 5 July 2008.

of the initial perturbations. Although the magnitude of the TE norm of WEP was the largest until FT = 15 (Fig. 14a), the WEP growth rate was the smallest among the five methods throughout the period of Fig. 14b.

6.2. *Characteristics of the initial perturbation*

In this subsection, we examine the characteristics of the initial perturbations obtained by four initial perturbation methods (GSV, MSV, MBD and LET) for the case of initial time 12 UTC, 4 July 2008. As shown in Fig. 4, there was a low-pressure system near Beijing, and a rainfall band extended from the low-pressure system southwestward. Southwesterly airflow prevailed on the southern side of the rain band, supplying warm, humid air. A relatively cold northerly flow on the north side converged with the southwesterly flow along the rain band.

Figure 15 shows the distributions of the initial ensemble spreads of temperature at the height of 700 hPa. The four methods can be divided into two groups: an SV group (GSV and MSV) and a breeding group (LET and MBD). In GSV and MSV, the regions of large spread are near the low-pressure system around Beijing, whereas in MBD and LET the large-spread areas are over southwestern China. These differences between the two groups may be inferred by the difference in the concepts of the methods. In the SV methods, perturbations that achieve maximum growth in the evaluation period are obtained by referring to the analysis field at the initial time. On the other hand, perturbations of the MBD and LET methods are given by time integration of the non-linear model in the 6-hourly breeding (EnKF) cycles before the initial time; the perturbations by the breeding group may tend to be produced on the more upwind (i.e. past) side than those by the SV group. The horizontal pattern of GSV perturbation is much broader than that of MSV, reflecting the difference in horizontal resolution between the two methods. The horizontal pattern of LET perturbation is similar to that of MBD but the amplitude of the spread is smaller over Japan and around Beijing, reflecting the auto-masking effect in LETKF corresponding to the observation density.

To examine the characteristics of the initial perturbations, we selected the member that showed the maximum rainfall amount for each initial perturbation method. Table 3 shows the maximum 6-h rainfall amounts obtained with the four methods by the

Table 3. Maximum 6-h rainfall and their horizontal position (I_max, J_max) and number of the member

All	Rain_max	I_max	J_max	Member
CNTL	121.29	58	75	–
GSV	140.85	58	75	5
MSV	134.79	58	75	5
MBD	155.22	57	76	1
LET	142.22	58	74	5

selected members. Because the rainfall amounts obtained by the chosen members were larger than the control run rainfall (CNTL), the characteristic features of the perturbations of each chosen member are those that explain the rainfall enhancement.

Figure 16 shows the horizontal distributions for 1000 hPa temperature perturbation by the members that predicted the maximum rainfalls. As mentioned before, the horizontal extension of the perturbation in MSV was smaller than that of the other methods. Black circles indicate the positions of the maximum 6-h rainfall amount in each of the four methods; all were in the rainfall band in southwestern China, though the low-pressure system was around Beijing. These maximum rainfall locations coincide with the positive perturbation regions of temperature in GSV, MBD and LET. Positive low-level temperature perturbations favour the intensification of precipitation, because they cause the atmosphere to become unstable. In LET and MBD, contrasting perturbations are seen along the rain band. For example, in LET the perturbation on the southern side of the rainfall band is positive and that on the northern side is negative. The horizontal pattern of perturbation in MBD is similar to that of LET. A similar tendency was also seen in the equivalent potential temperature (figure not shown).

Figure 17 displays vertical cross sections of the ensemble spreads of the equivalent potential temperature (θe) in the north–south (y)-direction through the point of the maximum 6-h rainfall. In GSV (Fig. 17a), the regions of large perturbation spread are below the height of 300 hPa. In the case of the maximum rainfall member, the positive perturbation of θe in the lower atmosphere is overlaid by negative perturbation layers, causing atmospheric conditions to become unstable (Fig. 18a). In MSV (Fig. 17b), perturbations are more localized near the

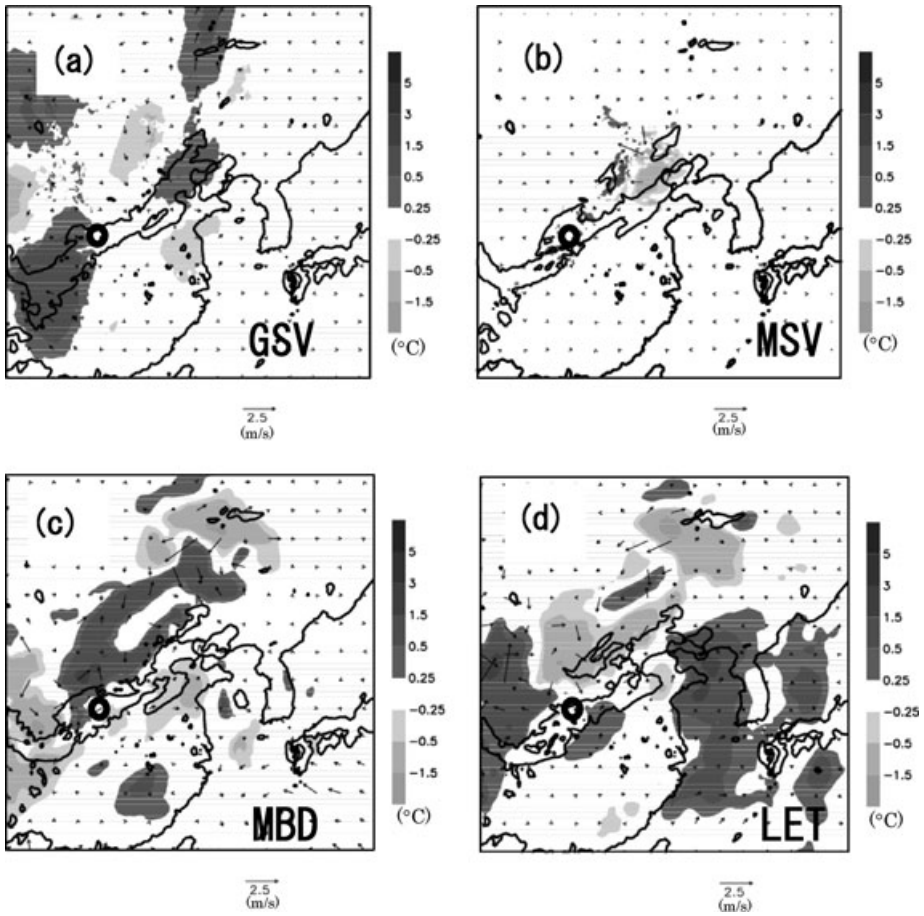

Fig. 16. Horizontal distributions of the 1000-hPa temperature perturbation at 12 UTC, 4 July 2008. Thick contours indicate the regions where 6-h rainfall exceeded 1 mm at 00 UTC, 5 July 2008. (a) GSV, (b) MSV, (c) MBD and (d) LET.

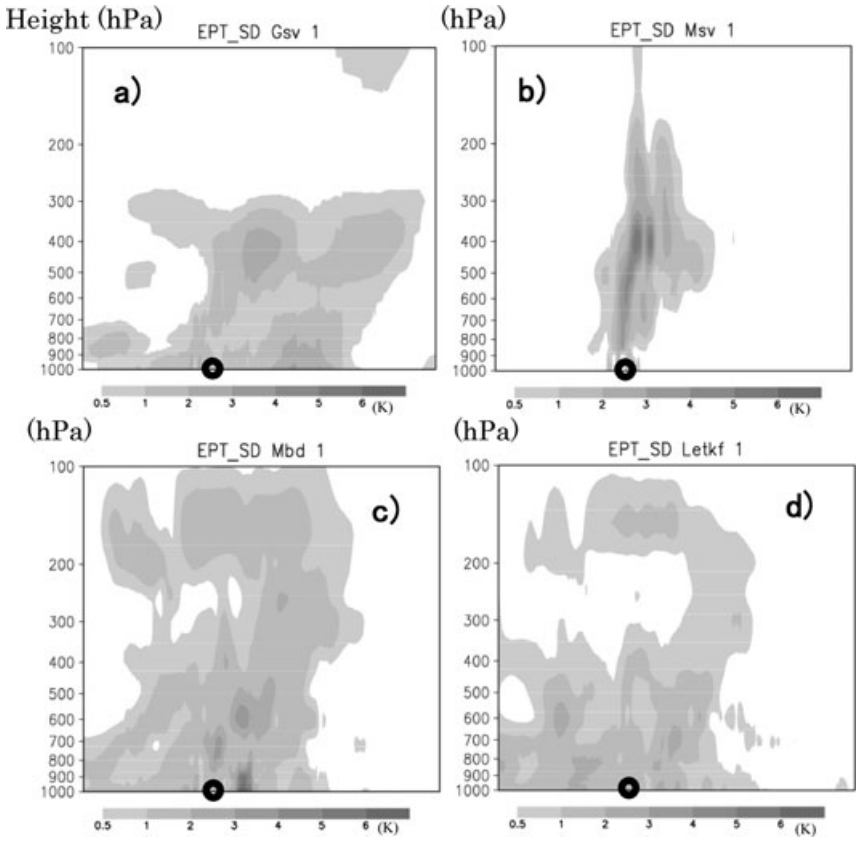

Fig. 17. Vertical cross sections of ensemble spreads of equivalent potential temperature (θe) at 12 UTC, 4 July 2008. (a) GSV, (b) MSV, (c) MBD, (d) LET. The sections are along the north–south direction thought the points of maximum 6-h rainfall at 00 UTC 5 July 2008. Black circles at the surface indicate the maximum rainfall locations.

rainfall band. Regions of large negative perturbations of θe are on the northern side of the positive perturbations, favouring intensification of the rainfall. In LET and MBD, the perturbation spreads extend to levels above 300 hPa (Figs 17c and d). In the maximum rainfall member, positive and negative perturbations are seen at lower levels on the southern and northern sides of the rainfall band, respectively, and a warmer, more humid air mass on the southern side of the rainfall band enhanced convection (Figs 18a and d). As for airflow, in MSV areas of large perturbations correspond to the updraft and rear inflow regions (Fig. 18b), but the low-level airflow was not modified. On the other hand, in the breeding group methods, the low-level inflow is modified (figures not shown). This difference between MSV and the breeding group methods probably affects the steadiness of the growth rate of the ensemble spread of the surface precipitation after the spin-up period.

7. Summary and concluding remarks

Comparison of different initial perturbation methods for mesoscale ensemble prediction was performed as part of MRI's intercomparison studies of the mesoscale EPSs of B08RDP. Five initial perturbation methods (WEP, GSV, MSV, MBD and LET) were developed and applied in preliminary experiments for the B08RDP Tier-1 mesoscale ensemble prediction with a horizontal resolution of 15 km. The same horizontal resolution (40 km) was employed for the three mesoscale model-based initial perturbation methods (MSV, MBD and LET) to make the comparison easier and so that the same lateral boundary perturbations could be applied in the breeding/EnKF cycles.

The evolution of ensemble spreads was compared. WEP spreads had relatively large amplitude at the initial time and increased slowly. GSV spreads were smaller than those of WEP, MBD and LET at the initial time, but grew steadily throughout the 36-h forecast period. After FT = 18, GSV spreads were the largest among the five methods. MSV spreads at the initial time were the smallest among the five methods but grew rapidly after the spin-up. In MSV, the surface precipitation spread increased particularly rapidly in the first 6 h. However, it decreased after that and became comparable to those of the other methods after FT = 12. MBD and LET spreads showed similar tendencies: relatively large values at FT = 0 and slow growth. Growths of spreads in LET were slightly slower than those in MBD.

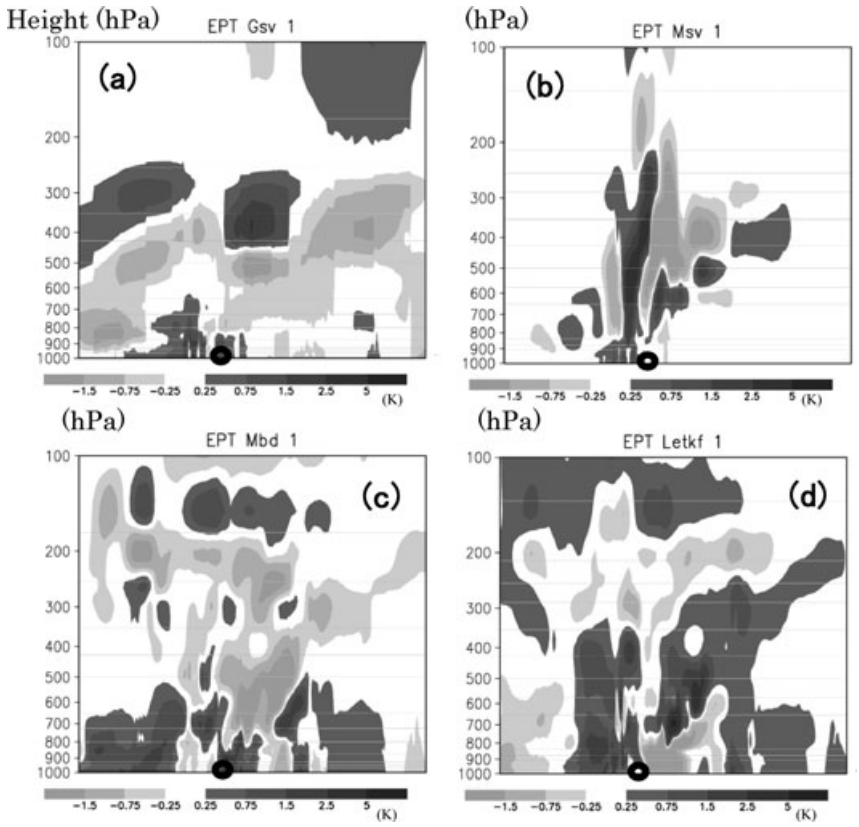

Fig. 18. Same as Fig. 17 except for perturbations of θe by the members that predicted the maximum rainfalls.

We examined RMSEs of the variables at FT = 24 against the initial condition at the same valid time. RMSEs of the ensemble means were smaller than those of the control run in almost all cases, suggesting the properness as the mesoscale ensemble forecast. WEP RMSEs were larger than those of the other methods for several variables, whereas RMSEs of GSV were the smallest among the methods for most surface variables. RMSEs of MSV were larger than those of MBD and LET, but still smaller than the WEP RMSEs for several variables. MBD and LET showed similar RMSE values. Comparison of the RMSEs and ensemble spreads of the five methods indicated that the ensemble forecasts were underdispersive compared with the model errors; the GSV method showed the best ratio of the ensemble spread to the RMSE.

Performances of the precipitation forecasts using the five methods were examined. For weak to moderate rains, the hit rates (detection rate) by the global model-based methods (GSV and WEP) were higher than those by the mesoscale model-based methods (MSV, MBD and LET). On the other hand, for moderate to intense rains, MSV, MBD and LET showed better detection rates than GSV or WEP. GSV had the best ROCSS for weak rains, whereas MBD and MSV had better scores for moderate or intense rain. A likely reason for this result is that GSV (and WEP) tends to perturb synoptic-scale disturbances whereas MBD, LET and MSV tend to perturb mesoscale disturbances, thus affecting local intense rains more. A similar tendency was confirmed by examining the probabilistic forecasts by BSS. GSV was best for weak rains of less than 3 mm/6 h, whereas MBD and MSV had better scores for moderate or intense rain larger than 5 mm/6 h. However, the BSS of GSV indicated good performance for intense rains of more than 20 mm/6 h, too.

To check the initial ensemble variance and the growth of perturbations, vertical profiles of the energy norm and their time evolution in the ensemble prediction were examined. All methods showed steady growth of the TE norms, but the GSV norm remarkably increased throughout the forecast period. Magnitudes of the initial norms of MBD and LET were almost the same, but the MBD norm became larger than that of LET after the model spi-up, showing relatively larger growth rate than that of LET. The MSV TE norm showed the minimum magnitude among the five methods, but its growth rate was the largest at the initial stage, demonstrating that the MSV method produces the perturbations which grow rapidly.

Characteristics of the five initial perturbation methods were also examined in the same case study. The regions of large spread were near the low-pressure system in GSV and MSV, whereas the large spread areas were over southwestern China, the upstream region, in MBD and LET. The spread in MSV

was concentrated near the mesoscale disturbance, and showed a narrow vertical structure.

In this study, the GSV method showed relatively good performance in terms of steady growth of the ensemble spread and small RMSEs of the ensemble mean, and a good ROCSS and BSS for weak to moderate rains. The GSV method outperformed the WEP method, confirming the advantage of targeting a domain of EPS to compute the SVs. However, for intense rains, mesoscale model-based initial perturbation methods (MSV, MBD and LET) performed better than GSV, which may be more important for mesoscale EPSs used for disaster prevention and risk assessment. Recently, Ono et al. (2010) developed a hybrid method merging GSV and MSV perturbations that performs well for both weak and intense rains. From the viewpoint of computational cost, SV methods are more expensive than MBD or LET.

The horizontal pattern of LET perturbations was similar to that of MBD but the distribution of the initial perturbations reflected the observation density by the automasking effect in LETKF. However, by objective measures such as RMSE and ROCSS, LET did not outperform MBD in our study, and its relative advantage as an initial perturbation generator is unclear. More detailed comparisons of MBD and LET with and without lateral boundary perturbation are given in section E-5 of Saito et al. (2010a), and will be reported in detail at a later date.

Several issues need further research for a full-scale validation of initial perturbation methods. The relatively good performance of the GSV method in our comparison may depend on the lateral boundary perturbations, especially in the latter half of the simulation period. We previously tested the impact of lateral boundary perturbations on results obtained with the MBD and LET methods (Saito et al., 2009), but not on those obtained with GSV. The spread of precipitation depends on the choice of the weight factor w_q in the TE norm of Barkmeijer et al. (2001), but the values are not necessarily optimized. As for the LET method, several factors such as the number of ensemble members, the inflation factor, and localization affect its performance (e.g. Miyoshi, 2011), and our LETKF has ample room for improvement. Model error is a large source of forecast uncertainty even in a short-range EPS (Eckel and Mass, 2005), and physical process perturbations should be considered in the non-linear models in MBD and LET. Since MBD and LET performances improve remarkably by implementation of lateral boundary perturbations (Saito et al., 2010a), the introduction of model uncertainty should further improve their performances. The period of the comparison in this study (2.5 d by two 36-h EPSs) was not adequate for full-scale verification of the initial perturbation period as suggested by error bars in Figs 7 and 8. A validation exercise requires a sample size much larger than only two cases, and it must cover a range of different synoptic patterns. These topics are beyond the scope of this paper, but they remain important subjects for future studies.

8. 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 thank Dr. Takemasa Miyoshi of the University of Maryland for allowing us to use NHM-LETKF and for valuable comments on this study. Thanks are also extended to Tabito Hara, Masayuki Kyouda, Tadashi Fujita, Kosuke Ono, Yuki Honda and other staff members of NPD/JMA for their help and support. We also express our sincere gratitude to Dr. Harald Rejenas, editor of *Tellus*, and two anonymous reviewers, whose comments significantly improved the manuscript. We also thank members of the International Science Steering Committees and the International Technical Supporting Team of B08RDP/FDP, especially Dr. Gong Jiandong of CMA. Surface observation data for verification in this study were provided by CMA.

References

Barkmeijer, J., Buizza, R., Palmer, T. N., Puri, K. and Mahfouf, J. 2001. Tropical singular vectors computed with linearized diabatic physics. *Q. J. R. Meteorol. Soc.* **127**, 658–708.

Beljaars, A. C. M. and Holtslag, A. A. M. 1991. Flux parameterization and land surfaces in atmospheric models. *J. Appl. Meteor.* **30**, 327–341.

Bishop, C. H., Etherton, B. J. and Majumdar, S. J. 2001. Adaptive Sampling with Ensemble Transform Kalman Filter. Part I: theoretical aspects. *Mon. Wea. Rev.* **129**, 420–436.

Bousquet, O., Lin, C.A. and Zawadzki, I. 2006. Analysis of scale dependence of quantitative precipitation forecast verification: a case-study over the Mackenzie river basin. *Q. J. R. Meteorol. Soc.* **132**, 2107–2125.

Bowler, N. E. 2006. Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model. *Tellus* **58A**, 538–548.

Bowler N. E., Arribas, A., Mylne, K. R., Robertson, K. B. and Beare S. E. 2008. The MOGREPS short-range ensemble prediction system. *Q. J. R. Meteorol. Soc.* **134**, 703–722.

Bowler N. E. and Mylne, K. R. 2009. Ensemble transform Kalman filter perturbations for a regional ensemble prediction system. *Q. J. R. Meteorol. Soc.* **135**, 757–766.

Buizza, R. and Palmer, T. N. 1995. The singular vector structure of the atmospheric global circulation. *J. Atmos. Sci.* **52**, 1434–1456.

Du J., DiMego, G., Tracton, M. S. and Zhou, B. 2003. NCEP short-range ensemble forecasting (SREF) system: multi-IC, multi-model and multi-physics approach. *CAS/JSC WGNE Res. Act. Atmos. Ocea. Modell.* **33**, 5.09–5.10.

Duan, Y., Gong, J., Chen, D., DiMego, G., Kuo, B. and co-authors. 2009. B08RDP—a WWRP Research and Development Project: Beijing 2008 Olympics Meso-scale Ensemble Prediction Research & Development Project (B08RDP).*A Report on the WWRP Research and Development Project B08RDP to the WWRP Joint Scientific Com-*

mittee. Available at: http://www.wmo.int/pages/prog/arep/wwrp/new/documents/Doc3_2_3_B08RDP.doc.

Ebert, E. E., Damrath, U., Wergen, W. and Baldwin, M. E. 2003. The WGNE assessment of short-term quantitative precipitation forecasts. *Bull. Amer. Meteor. Soc.* **84**, 481–492.

Eckel, F. A. and Mass, C. F. 2005. Aspects of effective short-range ensemble forecasting. *Wea. Forecast.* **20**, 328–350.

Ehrendorfer, M. and Errico, R. M. and Raeder, K. D. 1999. Singular-vector perturbation growth in a primitive equation model with moist physics.*J. Atmos. Sci.* **56**, 1627–1648.

Fujita, T., Tsuguti, 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)

Hara, M. 2010a. Global singular vector method. *Tech. Rep. MRI* **62**, 61–72.

Hara, T. 2010b. Turbulent processes. *Tech. Rep. MRI* **62**, 168–176.

Hoffman, R. N. and Kalnay, E. 1983. Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus* **35A**, 100–118.

Hohenegger, C. and Schar, C. 2007. Predictability and error growth dynamics in cloud-resolving models. *J. Atmos. Sci.* **56**, 4467–4478.

Honda, Y., Nishijima, M. Koizumi, K. Ohta, Y., Tamiya, K. and co-authors. 2005. A pre-operational variational data assimilation system for a nonhydrostatic model at Japan Meteorological Agency: formulation and preliminary results. *Q. J. R. Meteorol. Soc.* **131**, 3465–3475.

Houtekamer, P. L. and Mitchell, H. L. 2005. Ensemble Kalman filtering. *Q. J. R. Meteorol. Soc.* **131**, 3269–3289.

Houtekamer, P. L., Mitchell, H. L. and Deng, X. 2009. Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.* **137**, 2126–2143.

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.

Japan Meteorological Agency 2007. Outline of the operational numerical weather prediction at the Japan Meteorological Agency. *Appendix to WMO Numerical Weather Prediction Progress Report. Japan Meteorological Agency, Tokyo, Japan*, 194 pp. Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline-nwp/index.htm

Jolliffe, I. T. and Stephenson, D. B. 2003. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, John Wiley. Chichester, 245 pp.

Kadowaki, T. 2005. A 4-dimensional variational assimilation system for the JMA Global Spectrum Model. *CAS/JSC WGNE Res. Act. Atmos. Ocea. Modell.* **34**, 117–118.

Keenan, T., Joe, P., Wilson, J., Collier, C., Golding, B. and co-authors. 2003. The Sydney 2000 World Weather Research Programme Forecast Demonstration Project. *Bull. Amer. Meteor. Soc.* **84**, 1041–1054.

Koizumi, K., Ishikawa, Y. and Tsuyuki, T. 2005: Assimilation of Precipitation Data to the JMA Mesoscale Model with a Four-dimensional Variational Method and its Impact on Precipitation Forecasts. *SOLA* **1**, 45–48.

Kunii, M. 2010. MSV method. *Tech. Rep. MRI* **62**, 73–77.

Kunii, M., Saito, K. and Seko, H. 2010. Mesoscale data assimilation experiment in the WWRP B08RDP. *SOLA* **6**, 33–36.

Kunii, M., Saito, K., Seko, H., Hara, M., Hara, T. and co-authors. 2011. Verification and intercomparison of mesoscale ensemble prediction systems in the Beijing 2008 Olympics Research and Development Project. *Tellus* **63A**, this issue.

Li, X., Charron, M., Spacek, L. and Candille, G. 2008. A regional ensemble prediction system based on moist targeted singular vectors and stochastic parameter perturbations. *Mon. Wea. Rev.* **136**, 443–462.

Marsigli, C., Boccanera, F., Montani, A. and Paccagnella, T. 2005. The COSMO-LEPS mesoscale ensemble system: validation of the methodology and verification. *Nonlin. Processes Geophys.* **12**, 527–536.

Mason, I. 1982. A model for assessment of weather forecasts. *Aust. Meteor. Mag.* **30**, 291–303.

Miyoshi, T. 2011. The Gaussian approach to adaptive covariance inflation and its implementation with the Local Ensemble Transform Kalman Filter. *Mon. Wea. Rev.* **139**. doi: 10.1175/2010MWR3570.1.

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.

Molteni, F., Buizza, R., Palmer, T. N. and Petroliagis, T. 1996. The ECMWF ensemble prediction system: meteorology and validation. *Q. J. R. Meteorol. Soc.* **122**, 73–120.

Murphy, A. H. 1988. Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.* **116**, 2417–2424.

Nakanishi, M. and Niino, H. 2004. An improved Mellor-Yamada level 3 model with condensation physics : its design and verification. *Bound.-Layer Meteor.* **112**, 1–31.

Nuret, M., Lafore, J. P., Gouget, V. and Ducrocq, V. 2005. Mesoscale analysis and impact on simulation of IOP14 of the MAP experiment. *Q. J. R. Meteorol. Soc.* **131**, 2769–2793.

Nutter, P., Stensrud, D. and Xue, M. 2004. Application of lateral boundary condition perturbations to help restore dispersion in limited-area ensemble forecasts. *Mon. Wea. Rev.* **132**, 2378–2390.

Ono, K., Honda, Y. and Kunii, M. 2010. Development of a mesoscale ensemble prediction system using a singular vector method. *CAS/JSC WGNE Res. Act. Atmos. Ocea. Modell*. **40**, 5.17–5.18.

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. 2006a. The operational JMA nonhydrostatic mesoscale model. *Mon. Wea. Rev.* **134**, 1266–1298.

Saito, K., Kyouda, M. and Yamaguchi, M. 2006b. Mesoscale Ensemble Prediction Experiment of a Heavy Rain Event with the JMA Mesoscale Model. *CAS/JSC WGNE Res. Act. Atmos. Ocea. Modell.* **36**, 5.49–5.50.

Saito, K. and Hara, T. 2010. Numerical model for the 2008 experiment. *Tech. Rep. MRI* **62**, 40–50.

Saito, K., Ishida, J., Aranami, K., Hara, T., Segawa, T. and co-authors. 2007. Nonhydrostatic atmospheric models and operational development at JMA.*J. Meteor. Soc. Japan* **85B**, 271–304.

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

MRI **62**, 210 pp. Available at: http://www.mri-jma.go.jp/Publish/Technical/DATA/VOL_62/62_en.html.

Saito, K., Kuroda, T., Kunii, M. and Kohno, N. 2010b. Numerical simulations of Myanmar Cyclone Nargis and the associated storm surge. Part 2: ensemble prediction. *J. Meteor. Soc. Japan* **88**, 547–570.

Saito, K., Seko, H., Kunii, M., Hara, M. and Miyoshi, T. 2009. Influence of lateral boundary perturbations on the mesoscale EPS using BGM and LETKF. *CAS/JSC WGNE Res. Activities Atmos. Oceanic Modell.* **39**, 5.21–5.22.

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

Seko, H., Miyoshi, T., Shoji, Y. and Saito, K. 2011. Data assimilation experiments of precipitable water vapor using the LETKF system: intense rainfall event over Japan 28 July 2008. *Tellus* **63A**, this issue.

Seko, H., Saito, K., Kunii, M. and Kyouda, M. 2009. Mesoscale ensemble experiments on potential parameters for Tornado formation. *SOLA* **5**, 57–60.

Shapiro, M.A. and Thorpe, A.J. 2004. THORPEX International Science Plan. Version 3. WMO/TD-No. 1246, WWRP/THORPEX No. 2, 51 pp.

Simon, H. and Parlett, B. 1980. Solving Ax = b using the Lanczos algorithm with selective orthogonalization. University of California Berkeley.

Stensrud, D. J. and Yussouf, N. 2007. Reliable probabilistic quantitative precipitation forecasts from a short-range ensemble forecasting system. *Wea. Forecast.* **22**, 3–17.

Torn, R. D., Hakim, G. J. and Snyder, C. 2006. Boundary conditions for limited-area ensemble Kalman filters. *Mon. Wea. Rev.* **134**, 2490–2502.

Toth, Z. and Kalnay, E. 1993. Ensemble forecasting at NMC: the generation of perturbations. *Bull. Amer. Meteor. Soc.* **74**, 2317–2330.

Toth, Z. and Kalnay, E. 1997. Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.* **125**, 3297–3319.

Wang, X. and Bishop, C. H. 2003. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.* **60**, 1140–1158.

Wang, Y., Bellus, M., Wittmann, C., Steinheimer, M., Weidle, F. and co-authors. 2011. The Central European limited area ensemble forecasting system: ALADIN-LAEF. *Quart. J. Roy. Meteor. Soc.* **137**. doi: 10.1002/qj.751.

Yamaguchi, M. and Majumdar, S. J. 2010. Using TIGGE data to diagnose initial perturbations and their growth for tropical cyclone ensemble forecasts. *Mon. Wea. Rev.* **138**, 3634–3655.

Yamaguchi, M., Sakai, R., Kyoda, M., Komori, T. and Kadowaki, T. 2009. Typhoon ensemble prediction system developed at the Japan Meteorological Agency. *Mon. Wea. Rev.* **137**, 2592–2604.

Yoden, S. 2007. Atmospheric predictability. *J. Meteor. Soc. Jpn.* **85B**, 77–102.