E-6. Experimental system

E-6-1. Experimental system

Experimental system of MRI/JMA used in the 2008 experiment was almost the same as that of the 2007 experiment except the following points;

- Run a 6-hour control forecast in PREMF step with the change of breeding cycle in the MBD method.
- Because RA was replaced with the new global 4D-Var analysis (GA), we improved the system to use GA as initial conditions. After some verification, we applied the Meso 4D-Var analysis to the B08RDP domain to produce initial conditions for the control forecast (see E-3-1).
- To reduce the spin-up problem caused by the difference of the models, the incremental forecast in the outer-loop of MA was carried out using JMANHM.
- Initial perturbations were calculated under the same horizontal resolution in the MSV, MBD and LETKF methods for the easy comparison.
- To save the elapse time, the POST processes were parallelized.
- The wind components at 700hPa level were involved in the grib2 data transferred to CMA (see E-6-3).

Because of the introduction of MA, which took about 2 hours, total computation time became longer compared with the 2007 experiment. Nevertheless thanks to collaboration of the Numerical Prediction Division of JMA, timing of the data transfer from JMA to MRI was advanced so that we could send the data required by CMA on time.

E-6-2. Web visualization

To monitor the results of ensemble predictions and their verifications, we developed a visualization system on the website. During the period of the 2008 experiments, the system was run automatically just after the ensemble forecast, then we could check the results every day without any further work.

Figure E-6-1 shows the top page of the system. We can select the initial date, initial and boundary perturbation methods. When submit button is clicked after the selection, we move the individual result page (Fig. E-6-2). In the page, we can select the forecast time, element, level and the area (whole or common verification area). If the element ‘verif’ is selected, we can see the verification results such as ensemble spreads and RMSE/ME of the 24-hour forecast compared with the analysis fields at each level (Fig. E-6-3).

Because this visualization system was designed for the portable use, we can run the system on any Linux PC as long as the FORTRAN and C compilers were installed. After the B08RDP, we added the display of the probability distributions of 3-hour accumulated precipitation.

E-6-3. GRIB-2 transformation

The output of each forecast was the NuSDaS format, which was developed and widely used in the JMA. The advantages of using NuSDaS format were compact data size and the capability of utilizing some useful tools developed at JMA. Because the ensemble forecast data were required to be converted to the GRIB-2 format, we utilized one of the tools ‘nus2grib2’, which convert from the

NuSDaS format into the GRIB-2 format.

To remove ambiguity, a filename and template of the dataset were defined strictly, such as "F_FOR_FMT-*yyyymmddhh*-JMA-*member*.grd". In the 2006 experiment, each dataset included only surface variables, and further upper variables were added in 2007 and 2008 experiments (denoted by the symbols '*' or '**' in Table E-6-1).

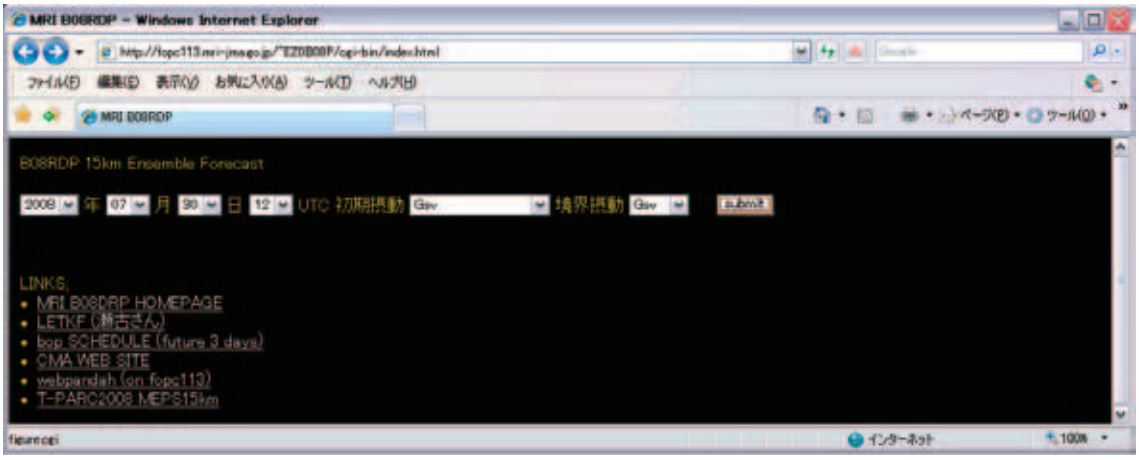


Fig. E-6-1. Top page of the B08RDP web site at MRI.

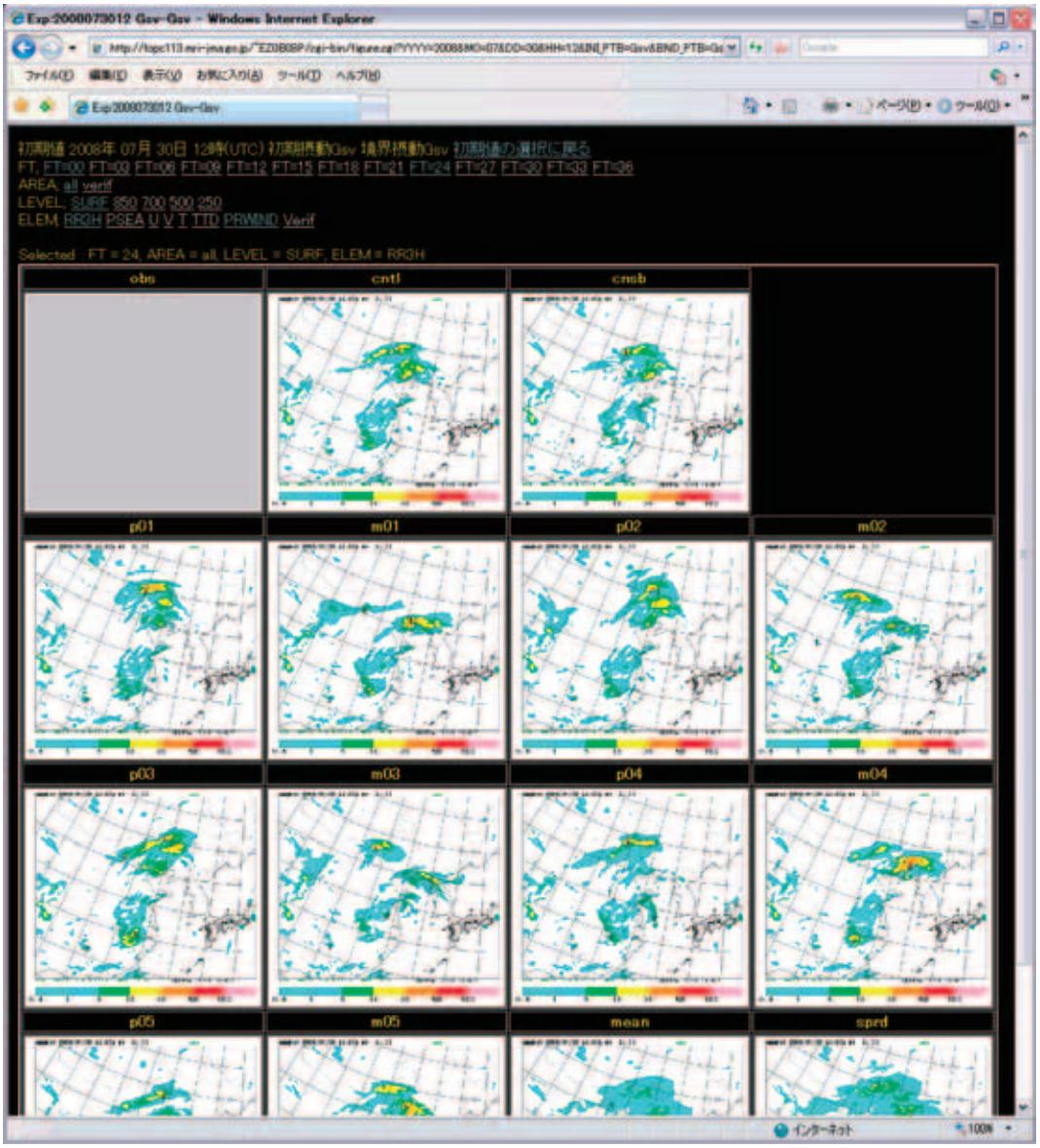


Fig. E-6-2. Horizontal distribution of 3-hour accumulated precipitation predicted by all members. The 'cntl' is a control forecast and the 'cnsb' is a no-perturbed forecast using the GSM analysis for its initial conditions.

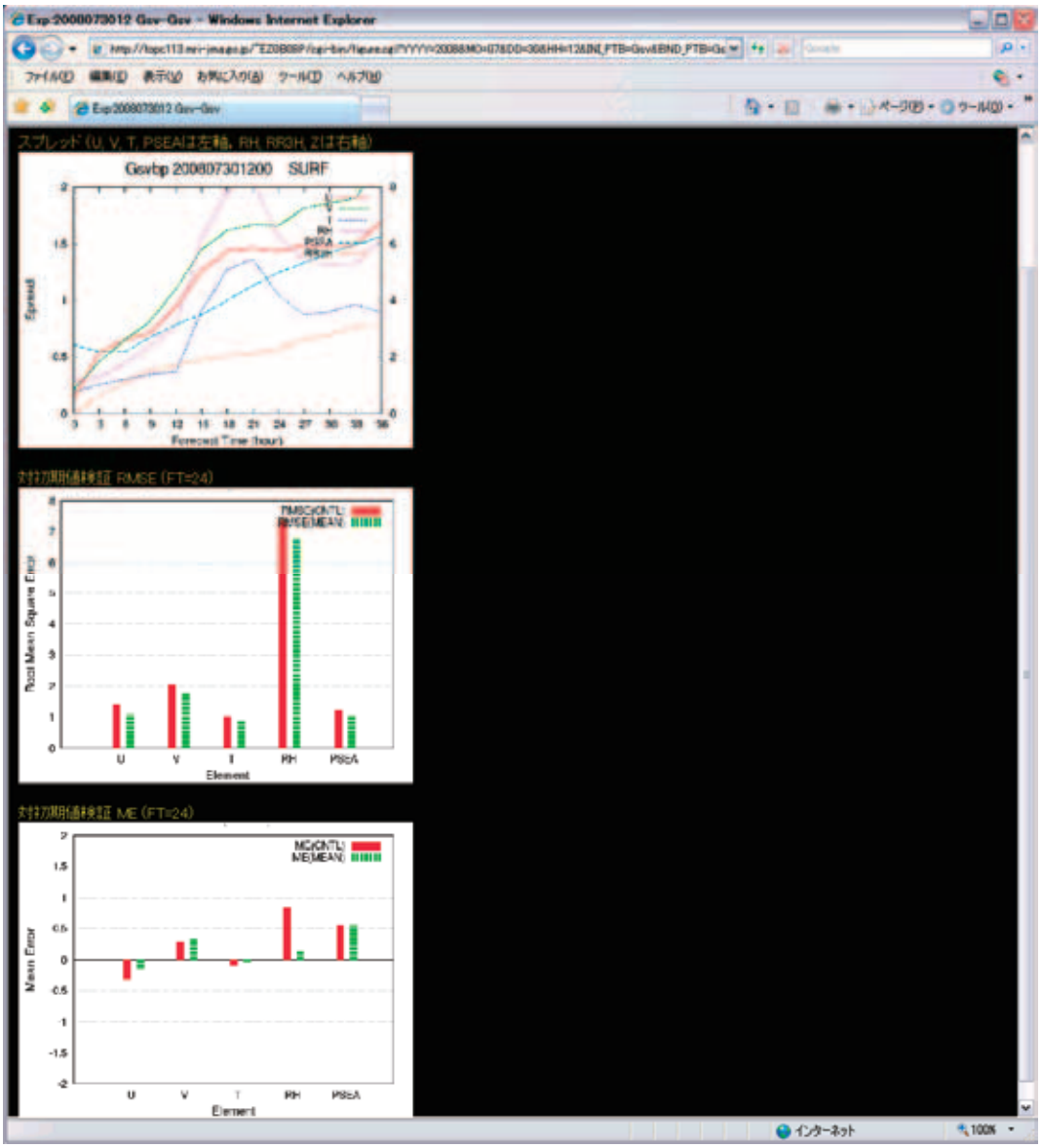


Fig. E-6-3. Verification result of a day. The RMSE and ME of the 24-hour forecast were calculated by comparing with the analysis fields, MA.

Table E-6-1. Data transfer list in B08RDP experiment. The symbols '*' and '**' denote the data included from 2007 experiment and 2008 experiment respectively.

LEVEL	SURF	850 hPa	700 hPa	500 hPa	250 hPa
U	○(10 m)	○*	○**	—	○*
V	○(10 m)	○*	○**	—	○*
T	○(2 m)	○*	—	○*	○*
Z	—	—	—	○*	—
RH	○(2 m)	○*	—	—	—
PSEA	○	—	—	—	—
RAIN	○	—	—	—	—
CAPE			○*		
CIN			○*		
SAUI			○*		

E-7. Verification

It is important to understand the characteristics of the MEP system through the verifications and intercomparisons in order to refine each MEP system in future. Verification is an indispensable part of meteorological research and operational forecasting activities (Casati et al 2008). In this section, the verification results as deterministic and probabilistic forecasts are presented. For more details, see Kunii et al. (2010c).

E-7-1. Verification of control run

Motivated by the interest for the performances of no-perturbed control forecasts of each participant, an intercomparison was performed among them for the 2008 B08RDP experiment period over the common verification area. The analysis system used for preparing initial condition and forecast model of each participant are listed in Table E-7.1. MSC applied ensemble Kalman filter technique (EnKF) for data assimilation by their global model and the others used variational assimilation system, 3D-Var or 4D-Var. In the following intercomparison, for MSC, we treated a forecast initialized by the ensemble mean at an initial time as the control forecast. NCEP employed three different forecast models for ensemble prediction whose ensemble members were five respectively. We tentatively regarded a forecast by WRF-NMM from no-perturbed initial fields as the control run.

The intercomparison of control forecasts was carried out on the common verification domain through the period of B08RDP experiment conducted in 2008. We verified surface parameters such as 6-hour accumulated precipitation and 2 m temperature by comparing the surface observation obtained from CMA and the mean value of four model grid points surrounding the observation point.

Figure E-7-1 shows the threat scores for 6-hour precipitation computed at 3-hour intervals between 6 and 36 hours. For weak and moderate rains, the scores of MRI/JMA are substantially better than the others while those of MSC are the best for intense rains. Figure E-7-2 presents the bias scores for 6-hour precipitation. The scores of MSC and MRI/JMA are comparatively flat for almost all thresholds. We can see that the forecast results of some participants overestimated weak rain frequency and underestimated intense rains. This characteristic is obvious in the results of NCEP and CAMS. Figures E-7-2b, E-7-2c and E-7-2d indicate that the diurnal change of bias score is apparent in the control forecast of ZAMG, especially for weak rains. Concerning MRI/JMA's results, the scores are comparatively flat and closed to one during the forecast period.

The verification results of surface temperature are indicated in Fig. E-7-3. With regard to the threat and bias scores at 14 LT (Figs. E-7-3a and E-7-3b), scores of MRI/JMA are the best of all. MRI/JMA also has lowest scores on mean error and root mean square error (Figs. E-7-3c and E-7-3d). These results of MRI/JMA indicate that tunings of model parameters for physical processes (K-F scheme and surface wetness) performed well, which were done in order to ameliorate underestimations of convective rains and maximum temperatures in abnormally hot days found in the 2007 experiment (Saito et al., 2008). Implementation of the Meso 4D-Var analysis (E-3-1) was also indispensable to improve the performance of the control run.

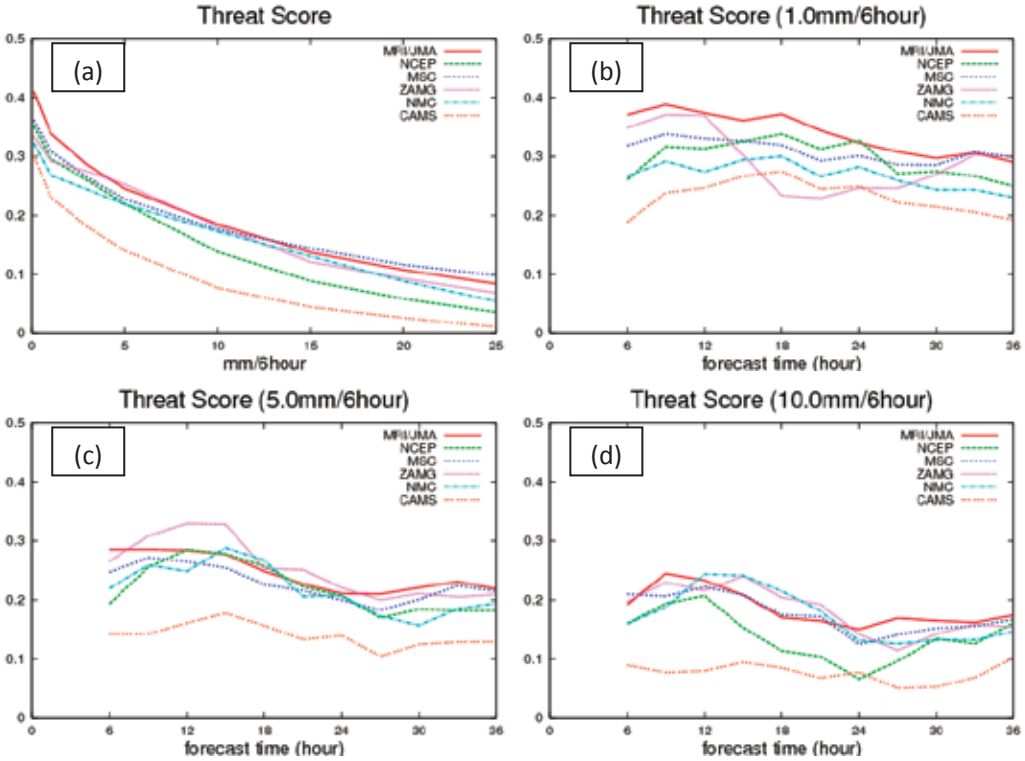


Fig. E-7-1. Threat scores of all participants for 6-hour accumulated precipitation averaged between 25 July and 23 August 2008, (a) as a function of threshold, (b) at a threshold of 1mm, (c) at a threshold of 5mm, (d) at a threshold of 10mm.

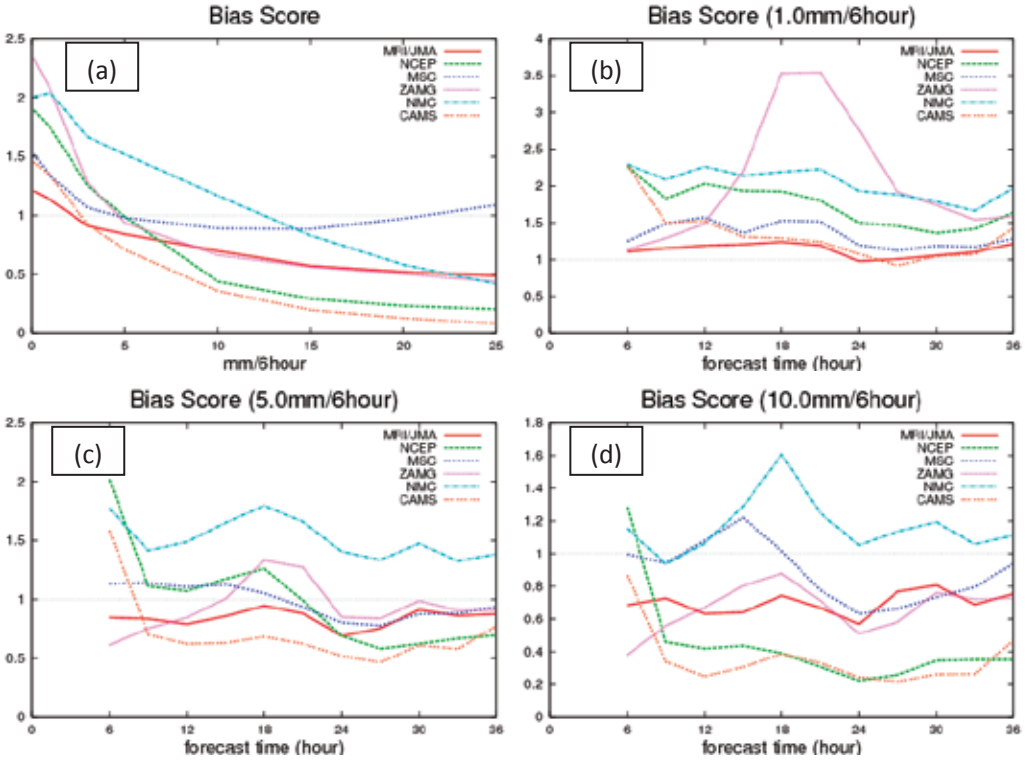


Fig. E-7-2. Same as Figure 7.1 but for bias scores.

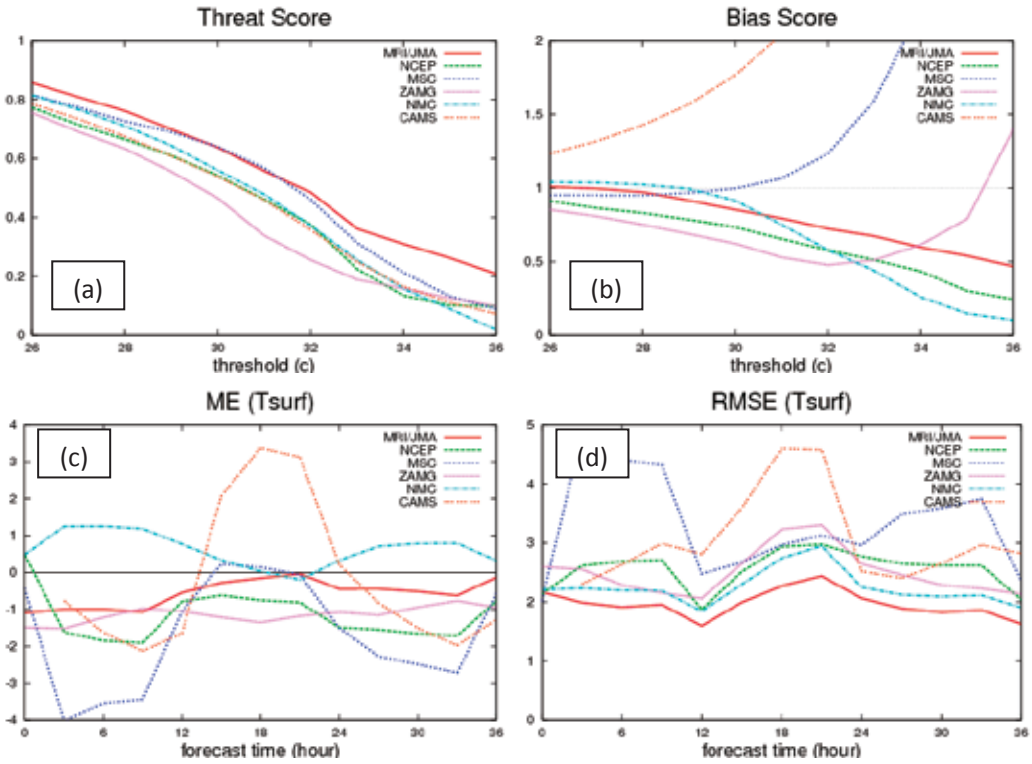


Fig. E-7-3. (a) Threat scores of all participants for surface (2 m) temperature at 14 LT averaged between 25 July and 23 August 2008, (b) same as (a) but for bias scores, (c) same as (a) but for mean error for 36 hours forecasts, (d) same as (c) but for root mean square error.

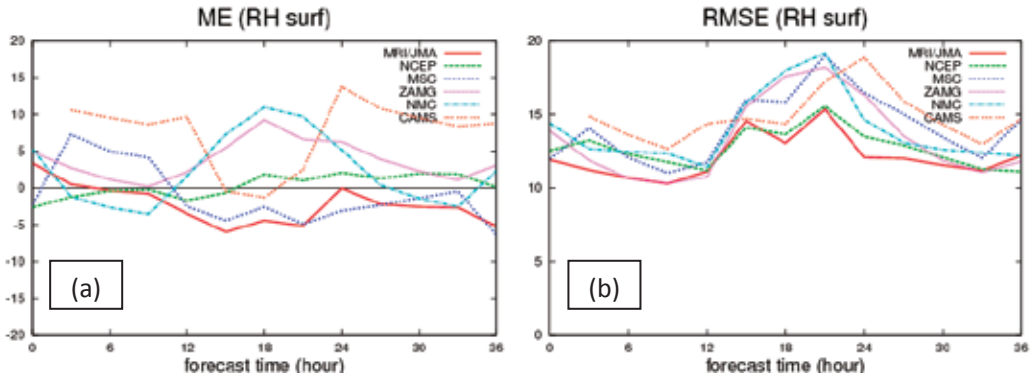


Fig. E-7-4. (a) Mean error of surface (2 m) relative humidity for 36 hours forecasts, (b) same as (a) but for root mean square error.

Table E-7-1. Specifications of control forecasts of each participant in B08RDP experiment.

Participants	Initial Condition	Forecast Model
MRI/JMA	JMA Meso-4DVar (+NHM 3hr forecast)	NHM (L40)
MSC	MSC Global EnKF	GEM (L28)
ZAMG & Meteo-Fr.	ECMWF Global 4DVar	ALADIN (L37)
NCEP	NCEP Global 3DVar	WRF-ARW, WRF-NMM, GEFS-Downscaled (L60)
NMC/CMA	WRF-3DVar	WRF-ARW (L31)
CAMS/CMA	GRAPES-3DVar	GRAPES (L31)

E-7-2. Verification as the ensemble prediction system

In this section, verification results of the ensemble forecasts performed by MRI/JMA are presented. The specification of the Tier-1 ensemble prediction system of each participant is listed in Table E-1-1. The number of ensemble members varies from 9 of CAMS to 20 of MSC. All participants applied lateral boundary perturbations. While the physical perturbation method was not implemented in the MRI/JMA system, others adopted the multi-model or multi-physics method to represent the uncertainty in the forecast model.

The verification shown in this subsection is performed for the period from 25th July to 23rd August in 2008 and for the surface and upper variables in the common verification region.

a. Evolution of ensemble spreads

Figure E-7-5 shows the ensemble spreads for surface variables. For all participants, we can see the steady growth of ensemble spreads through the forecast period. MRI/JMA, MSC, ZAMG, and CAMS ensembles have maxima in spreads of temperature and relative humidity at around FT=18 (1400 local time), whereas the others, NCEP and NMC, have another peak around FT=06 to 09 (0200 to 0500 local time). The peak around FT=18 was caused by the diurnal cycle. The first peak of NCEP may be due to the use of two (global and mesoscale) models in their ensemble system; the perturbation growth patterns varied in the different models. It seems that NMC's first peak was associated with the initial imbalance because of the large amplitude of initial perturbations[1]. This early imbalance can also be seen clearly in the spread of wind components of MSC and relative humidity of ZAMG. The most notable characteristic in Fig. E-7-5 is that the spread of precipitation of MSC system grows most rapidly in first 6 hours. That spread at FT=06 is more than 3 times as large as the second largest one (NCEP). This suggests that initial perturbations of moist fields had large amplitudes or some ensemble members predicted intense precipitation because of their physical perturbations.

The ensemble spreads for upper variables are shown in Fig. E-7-6. The spread of these variables evolves differently from the surface variables because they were not directory influenced by the diurnal change. For almost all variables, the spreads of MSC and NCEP are larger than the others'. As with the growth rate of ensemble spreads, those of MRI/JMA have the steepest gradient. This is partly because of MRI/JMA' utilization of singular vectors for both initial and boundary perturbations, which have characteristic that maximizes the growth rate of forecast perturbations.

b. Forecast errors of ensemble mean

To investigate the advantage of the ensemble forecast against a single forecast, we carried out forecast verifications comparing the forecast error of the individual control forecast with that of ensemble mean. Moreover, comparison of forecast errors of ensemble mean and ensemble spreads were performed to assess the validity of the magnitude of ensemble spread. In this study, the 24 hour forecast by each EPS was verified by evaluating the RMSE against each participant's analysis (the initial condition of the control run) at the same valid time. The verification of precipitation forecasts is done against the rain-gauge observations. Figure E-7-7 shows forecast verifications for surface

[1] Since the transferred data by NMC and CAMS didn't include surface variables at initial times, ensemble spreads of those centers were set to 0 at FT=0.

temperature (T2m)[2] at FT=24. One can see that ensemble mean provides a better forecast than the control one; in particular the improvement of NCEP is the most obvious. However, for all participants, ensemble spreads are smaller to some extent than corresponding forecast errors of ensemble mean. The same comparison for the sea surface pressure is shown in Fig. E-7-8. Except MRI/JMA and NCEP, the ensemble spreads substantially smaller than the forecast errors, and this characteristic can be seen in verifications of other surface elements, such as relative humidity and wind components (not shown).

For T2m and surface relative humidity (RHs), more detailed verifications against observations are performed. Figure E-7-9 shows the time series of the RMSE and the ME of the ensemble averages as well as the individual control forecast for T2m. One can see that the CAMS system has a strong diurnal cycle even if the ensemble average is used, which seems to be caused by the characteristic of its forecast model to have a tendency to overestimate T2m in daytime. The most distinctive feature is that, compared with the control forecast, the ensemble mean reduces the variation of mean error originated from its average, but it doesn't always reduce the mean bias itself. This trend is relatively obvious in results of MSC and NMC. As with the RMSE, NCEP and MSC ensemble systems clearly improve the scores, but for the others, we cannot identify the differences between them.

For the quantitative evaluation, the improvement (ensemble average minus individual control forecast) of the RMSE, the ME and the square of the standard deviation of the error (SDE) due to the usage of ensemble averages is plotted in Fig. E-7-10. The RMSE can be decomposed into the ME and the SDE (Murphy 1988; Hou *et al.*, 2001);

$$\mathrm{RMSE}^2 = \mathrm{ME}^2 + \mathrm{SDE}^2 \tag{E-7-1}$$

$$\mathrm{SDE}^2 = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}(\mathrm{x_i} - \mathrm{a_i} - \mathrm{ME})^2 \ , \tag{E-7-2}$$

where N is the number of samples, x is a model forecast, and a is a observation. These figures indicate that NCEP and MSC improve the RMSE in this case, and for MSC, this improvement is particularly attributed to the large improvement of SDE. The SDE can be interpreted as the standard deviation of random errors. It seems that the usage of multi models or multi physics which influenced on the surface perturbation effectively contributed to ameliorate the score. On the other hand, for MRI/JMA and CAMS, their ensemble systems showed little improvement in the RMSEs reduction. This is probably due to the lack of physical perturbation and insufficiency of initial perturbations in the surface processes. In Fig. E-7-10b, it is found that the ME of the ensemble average was not always ameliorated compared with the individual control forecast. This result suggests that the difficulty of the introduction and tuning of physical perturbation methods.

Regarding to the RHs, the same figures are presented (Figs. E-7-11 and E-7-12). For the mean error, although the bias of NCEP is neutral in the control verification, it turns negative in the ensemble verification. Since NCEP used three different models for ensemble forecasting, this negative bias was probably caused by the different model from the control forecast. On the contrary, the ensemble average of MSC and NMC improved the negative bias observed in daytime of its control forecast. That subsequently leads to more amelioration of the RMSE in NMC case, whereas the contribution of

[2] The ensemble forecast data of CAMS did not include the surface temperature and relative humidity at initial time.

SDE is larger in MSC.

c. Verification scores for the precipitation

Next we focus on the comparison between the control forecast and ensemble mean with the verification scores for 6-hour accumulated precipitation. Figure E-7-13 shows the threat and bias scores for the precipitation against observation. For weak and moderate rains, ensemble means of some participants improved the threat scores, while the scores of ensemble mean become worse for intense rains. This is not a surprising result since the forecast fields become smooth by the averaging, and the reproducibility of heavy rain in that fields worsens. As with the bias score, one can see the overestimation for weak rains and the underestimation for intense rains, the border of which is around 5mm / 6hour.

We now introduce the probabilistic verification scores, such as Brier score and relative operating characteristic (ROC) diagram (Harvey *et al.*, 1992). Figure E-7-14 shows the Brier scores for individual and ensemble forecast. Futhermore, the improvement rate of Brier score defined as follows;

$$\text{improvement rate} = \left(1 - \frac{\text{Brier Score (ensemble)}}{\text{Brier Score (cntl)}}\right) \times 100 \ , \qquad \text{(E-7-3)}$$

is also illustrated. One can see that ensemble forecasts improved Brier scores clearly compared with the individual control forecast. The improvement rates of MSC and NMC are relatively higher than other systems. For intense rains, the improvement rates of NCEP and CAMS systems are rather small, which are caused by the underestimation of intense rains (Fig. E-7-13d). The rate of MRI/JMA is almost constant even for intense rains, which is similar to the result of MSC, whereas the amelioration for weak rains is relatively poor. One possible reason is that regarding to the individual control forecast, MRI/JMA's system showed the best threat and bias score especially for weak to moderate rains (See E-7-1) therefore to ameliorate the score by using ensemble forecasting further might be relatively difficult.

To assess the relationship between the improvement rate of Brier score and the number of EPS members, the rates of individual systems are plotted as a function of the number of EPS members with different thresholds (Fig. E-7-15). One can see the strong positive relationship between the improvement due to the ensemble forecasting and the number of members. For weak rains, the improvement rate of MRI/JMA is slightly inferior to the others for the number of members. MSC ensemble indicates the superiority of its probabilistic forecast for the intense rains.

ROC diagrams and ROC area skill score (ROCASS) are shown in Fig. E-7-16. ROCASS is defined by the area enclosed by the curve and x-y axis (denoted by 'S');

$$\text{ROCASS} \equiv 2(\text{S} - 0.5) \quad (-1 \leq \text{ROCASS} \leq 1) \ , \qquad \text{(E-7-4)}$$

where ROCASS equals to 1 if the probabilistic forecast is perfect. These figures indicate that the detection rate of precipitation of MRI system was insufficient, which suggests that importance of number of EPS members and physical perturbations. For intense rains, MSC system shows its higher performance of the probabilistic forecast of precipitation. These features can also be seen in Fig. E-7-17.

Not only verification scores but also actual fields need to be investigated. Figure E-7-18 illustrates an example of distributions of ensemble spread of 3-hour precipitation for a heavy rainfall

case occurred in the experiment period. A glance at these figures will reveal that the spread of MSC is by far the largest of all, and its horizontal scale is unnaturally small for the resolution of the forecast model. In order to examine the characteristics of individual precipitation forecast, we plotted the frequency distribution of 6-hour accumulated precipitation averaged by the number of EPS members for all participants and observations in Fig. E-7-19. Note that the scale of the vertical axis is logarithmic. This figure indicates that the frequency of MRI/JMA, NCEP and NMC are relatively close to the observations, while the frequency of very intense rains predicted by MSC system is higher than the observation. It can be seen from this result that some of the MSC members tend to predict intense rainfalls even when such severe rainfalls are not observed, which is consistent with the features of MSC's precipitation spread shown in Fig. E-7-5.

d. Summary

We performed the verification and intercomparison of ensemble prediction systems in B08RDP. Through the results obtained here, it can be summarized that;

- For all systems, the ensemble spreads grew as the forecast time increased, and the ensemble mean improved the forecast errors compared with individual control forecast in the verification against the analysis fields.
- For almost all systems, the ensemble spreads were somewhat smaller than forecast errors of the ensemble mean (weak under-dispersive).
- Ensemble forecasts improved Brier scores clearly compared with the individual control forecast. For intense rains, the improvement rates of NCEP and CAMS systems were rather small, which were caused by the underestimation of intense rains.
- Some problems with individual ensemble systems have been identified.
  - In the ROC curves, the hit rates of precipitation for moderate rains of the MRI system were insufficient, probably due to the small number of ensemble members and the lack of physical perturbations.
  - The frequency of very intense rains predicted by the MSC system was higher than the observations
  - For surface conditions (T and RH at 2m), the MRI (and CAMS) ensemble system showed little improvement in the RMSEs reduction, probably due to the lack of physical perturbation and insufficiency of initial perturbations in the surface processes.
  - It was found that mean error of the ensemble forecast was sometimes worse than that of the individual control forecast. This suggests that the careful attention should be paid for the physical perturbations.
- As a whole, concerning the control forecast, the performance of the MRI system was the best, whereas the remarkable improvement of using ensemble forecast was seen in the MSC system. It seems that the improvement of the MSC ensemble system is due to the fact that the MSC system had a large number of EPS members, large ensemble spreads and that initial perturbation and physics perturbation functioned effectively in the ensemble forecast.

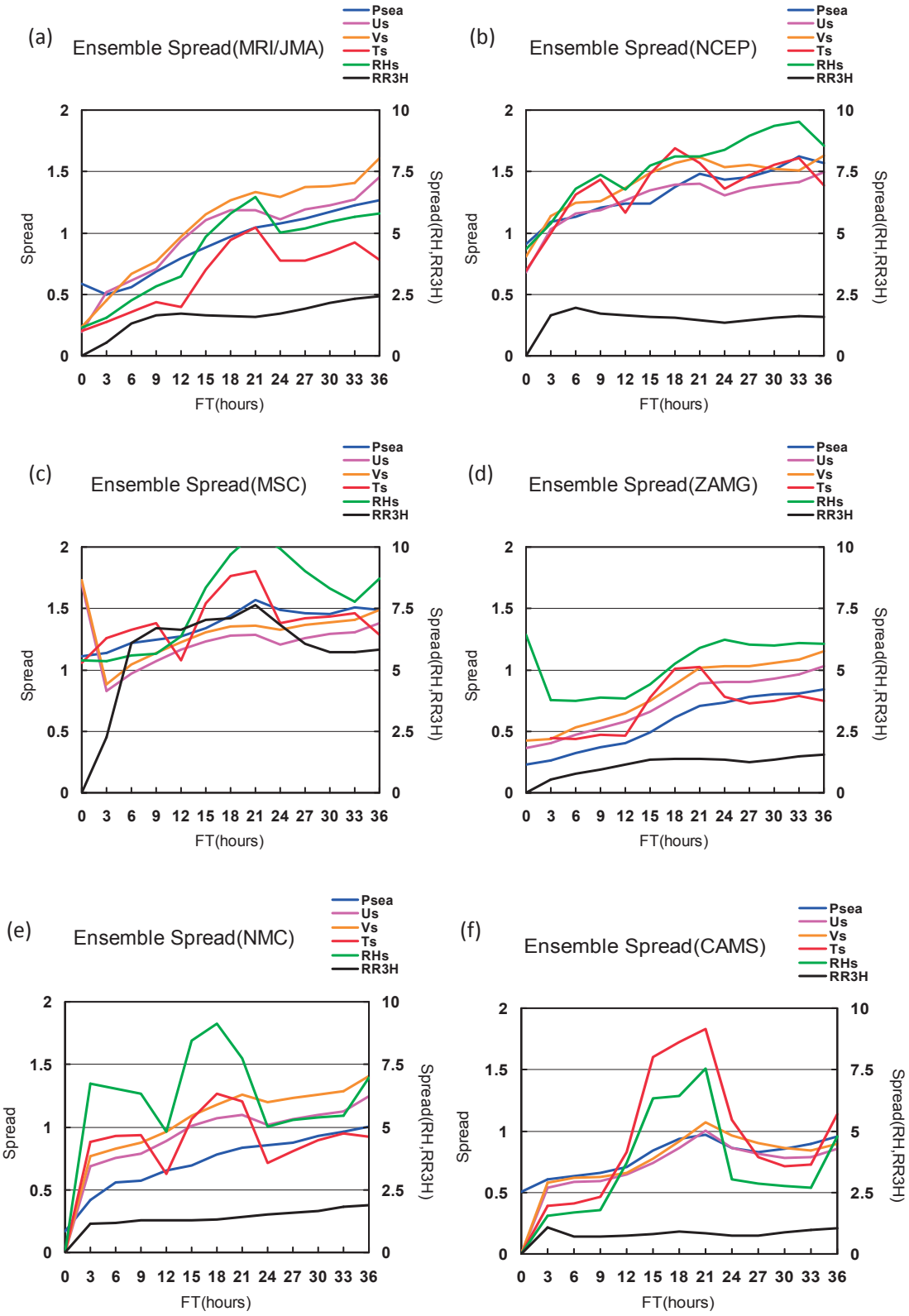


Fig. E-7-5. Ensemble spreads of surface variables of (a) MRI/JMA, (b) NCEP, (c) MSC, (d) ZAMG, (e) NMC, (f) CAMS. The right axis indicates the spread of 3-hour accumulated precipitation (RR3H), and the left axis indicates the others' spreads.

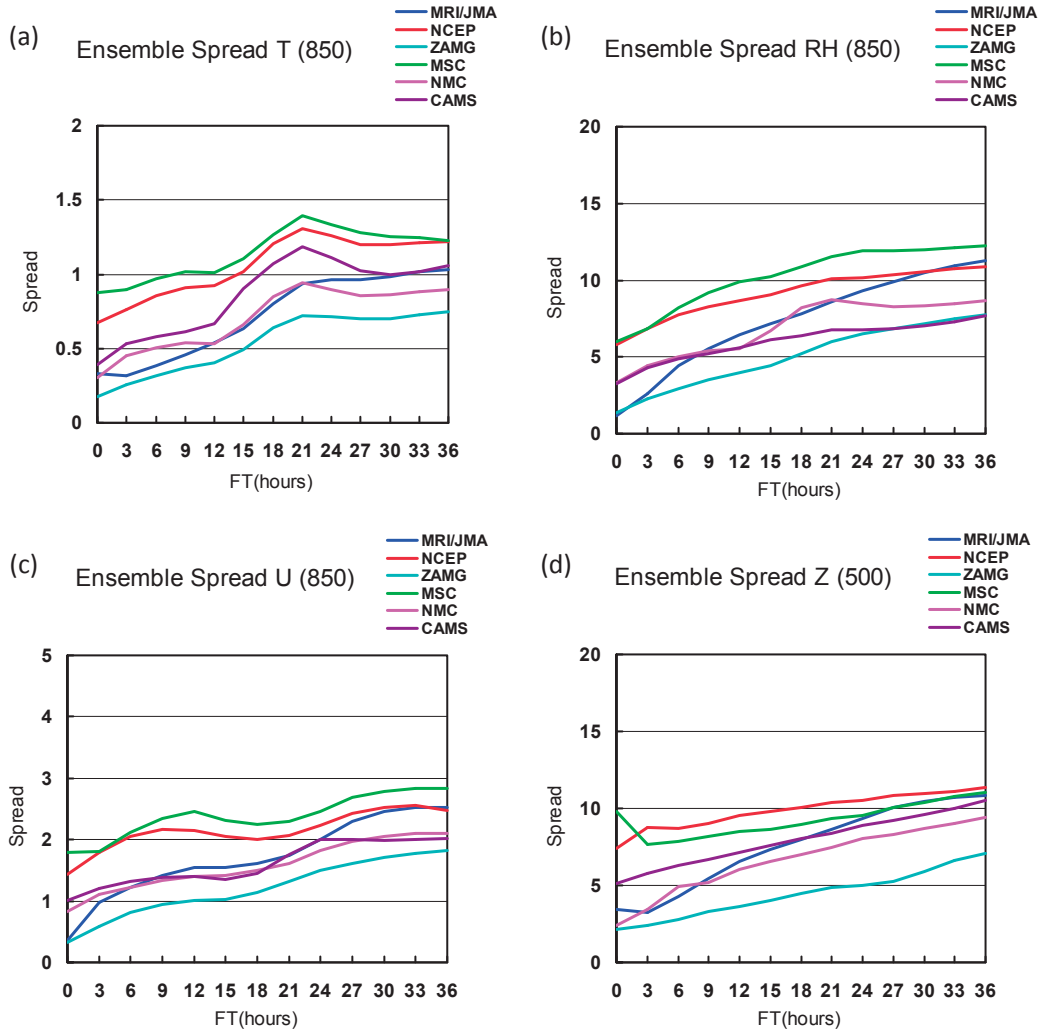


Fig. E-7-6. Ensemble spreads of upper variables. (a) temperature at 850 hPa, (b) relative humidity at 850 hPa, (c) zonal wind at 850 hPa, (d) geopotential height at 500 hPa.

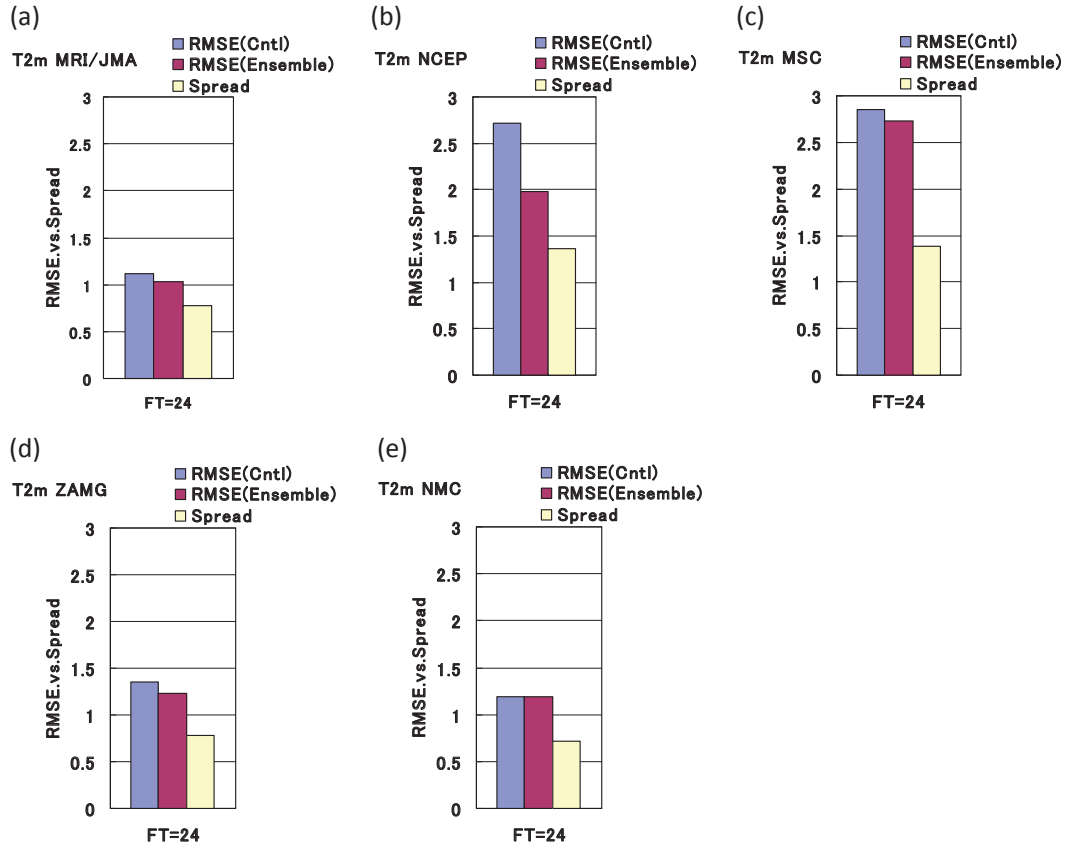
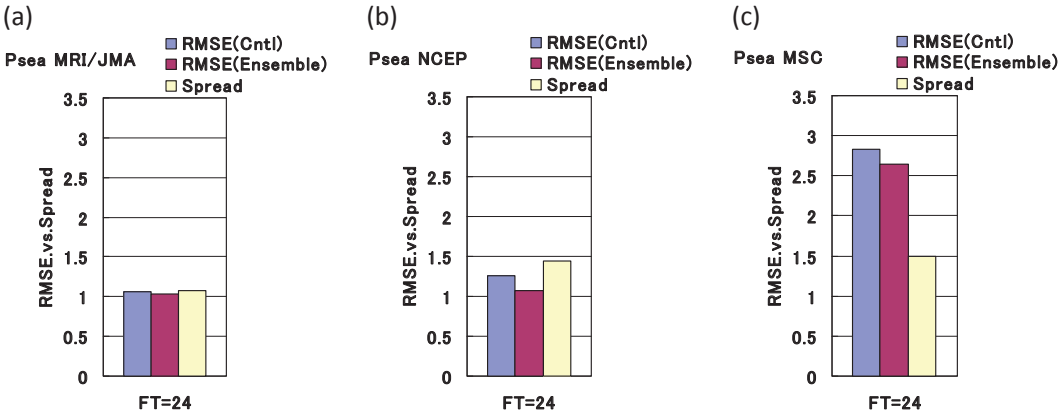


Fig. E-7-7. The comparison among the RMSE of a control forecast, ensemble mean, and ensemble spread for surface temperature (T2m). (a) MRI/JMA, (b) NCEP, (c) MSC, (d) ZAMG, (e) NMC. T2m data were not included in the grib2 data of CAMS. For MSC, member 21 was assumed as the control forecast.



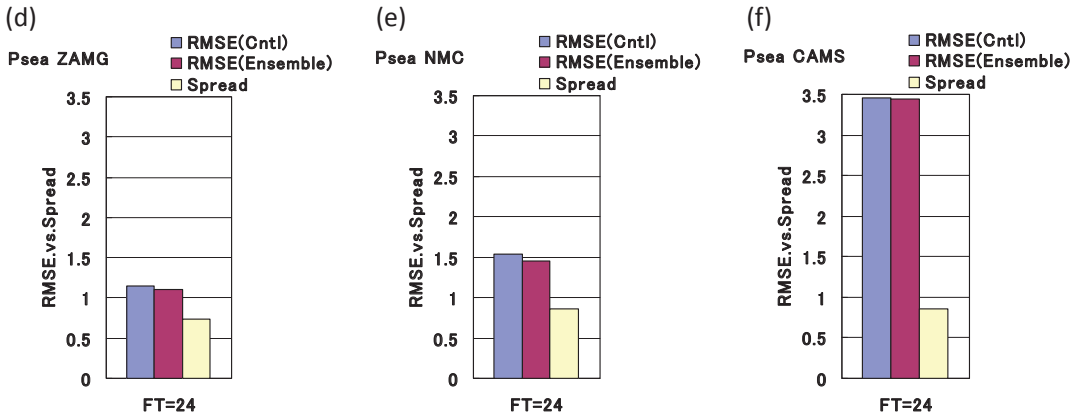


Fig. E-7-8. Same as Fig. E-7-7, but for sea surface pressure. The result of CAMS (f) is included.

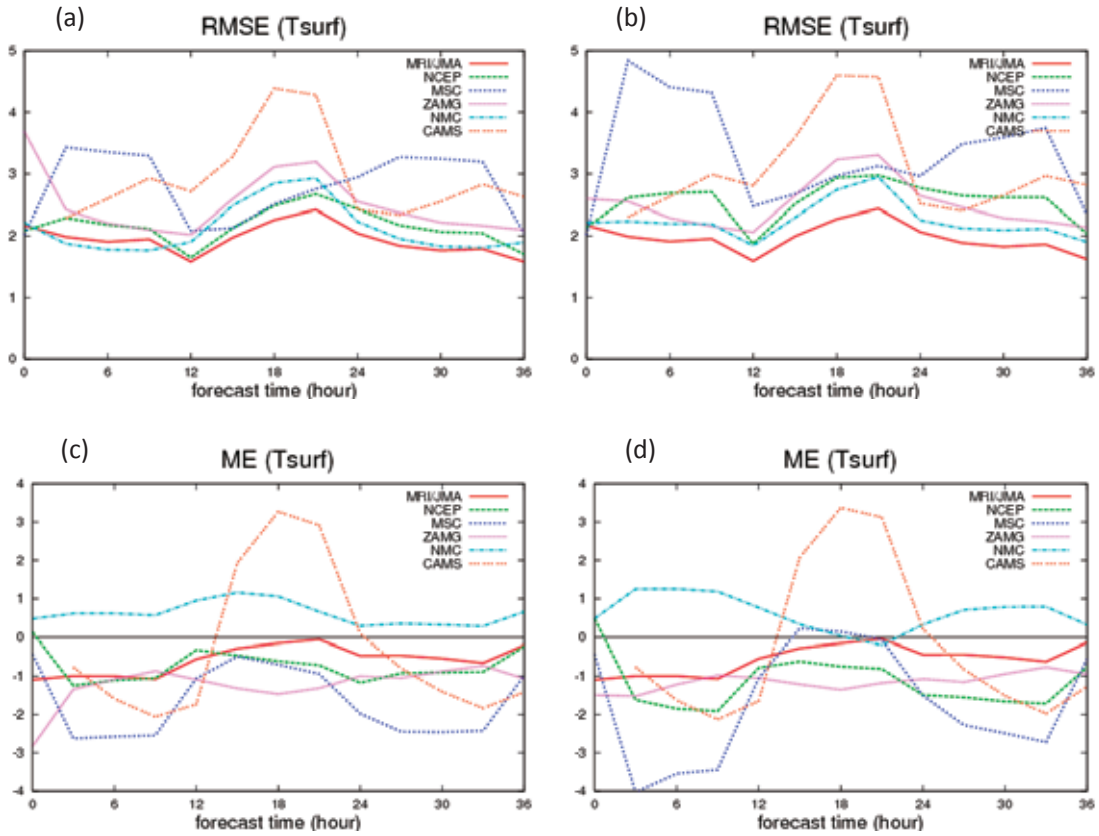


Fig. E-7-9. Time series of verifications of surface temperature against observations. (a) RMSE of ensemble mean, (b) RMSE of the control forecast, (c) ME of ensemble mean, (d) ME of the control forecast.

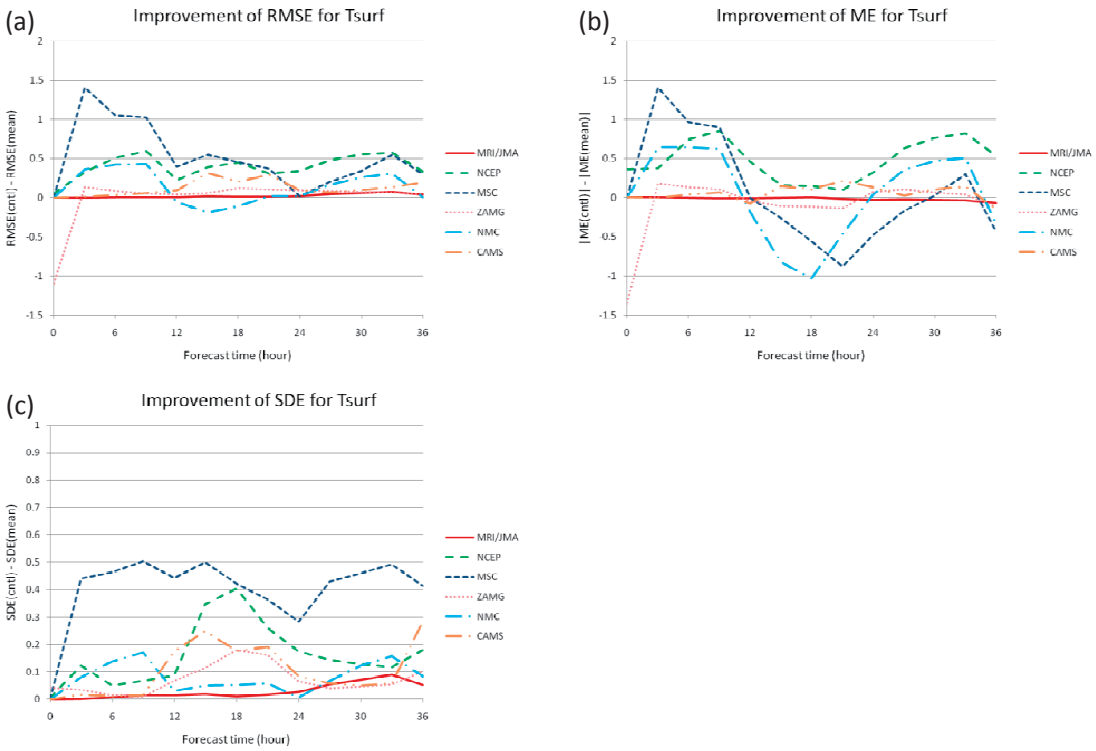


Fig. E-7-10. The improvement of the (a) RMSE, (b) ME and (c) SDE due to the usage of ensemble averages for T2m.

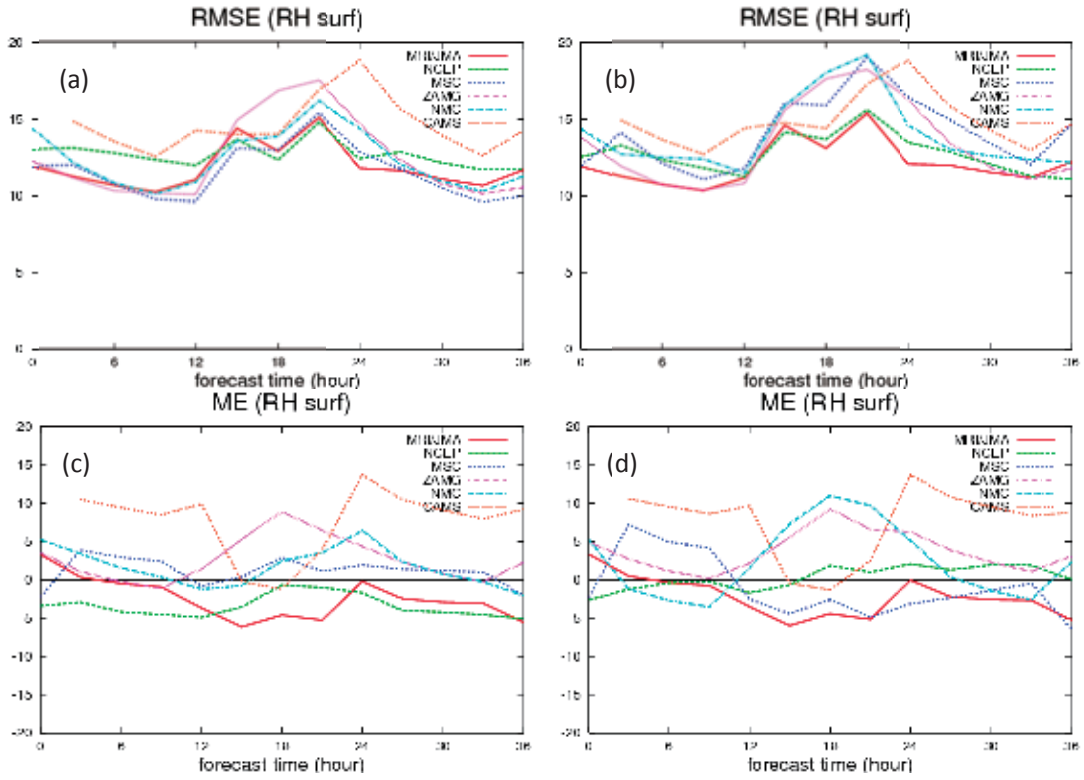


Fig. E-7-11. Same as Fig. E-7-9, but for RHs.

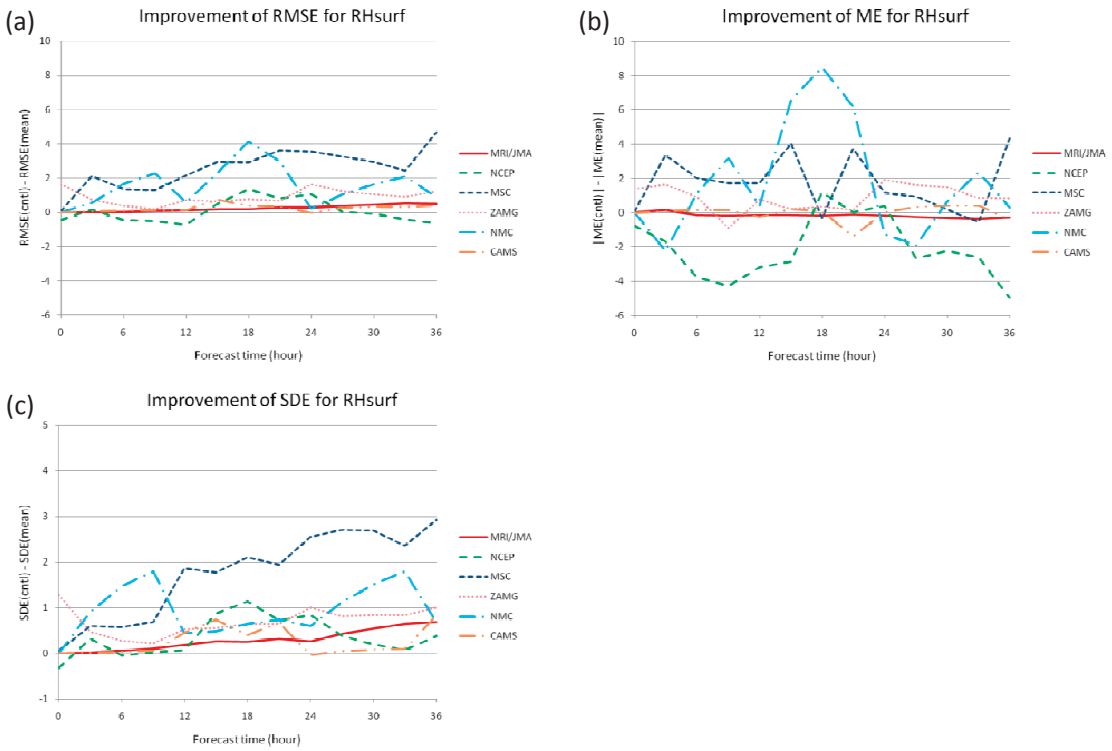


Fig. E-7-12. Same as Fig. E-7-10, but for RHs.

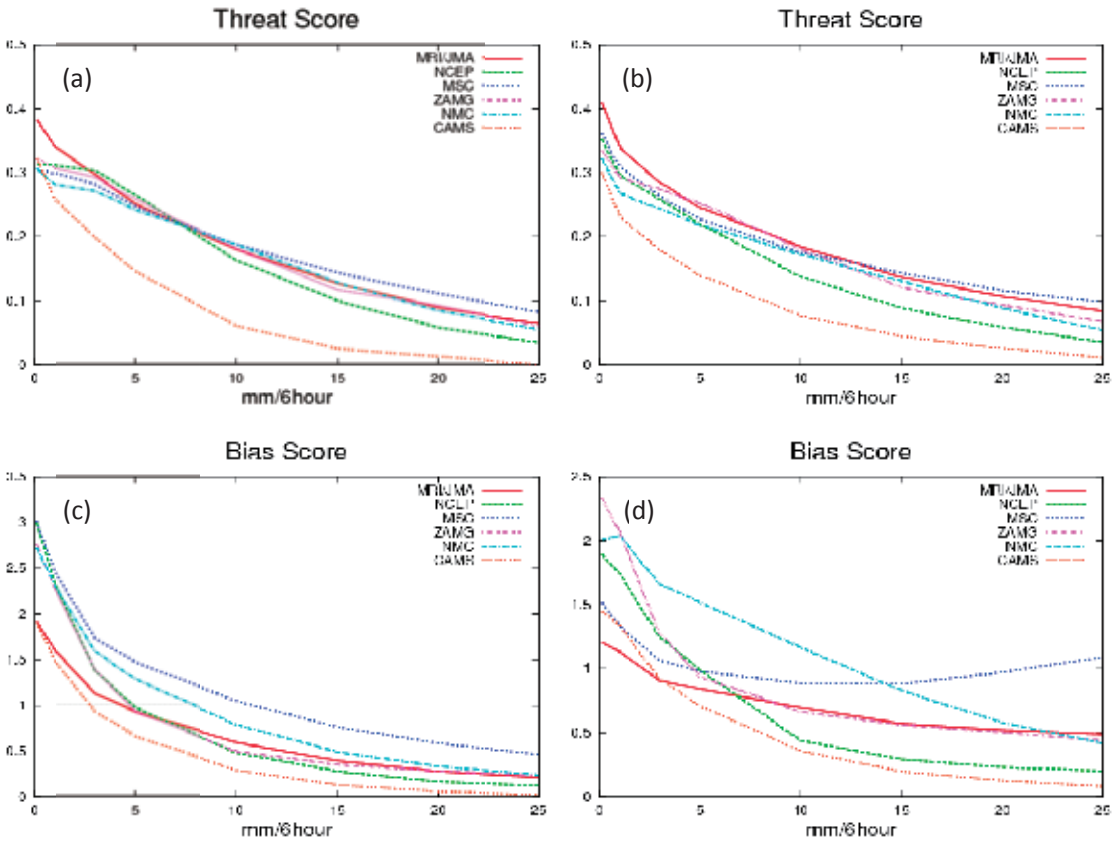


Fig. E-7-13. Threat and bias score for 6-hour accumulated precipitation of all participants: (a) threat score for ensemble mean, (b) threat score for a control forecast, (c), (d) same as (a), (b) but for bias scores.

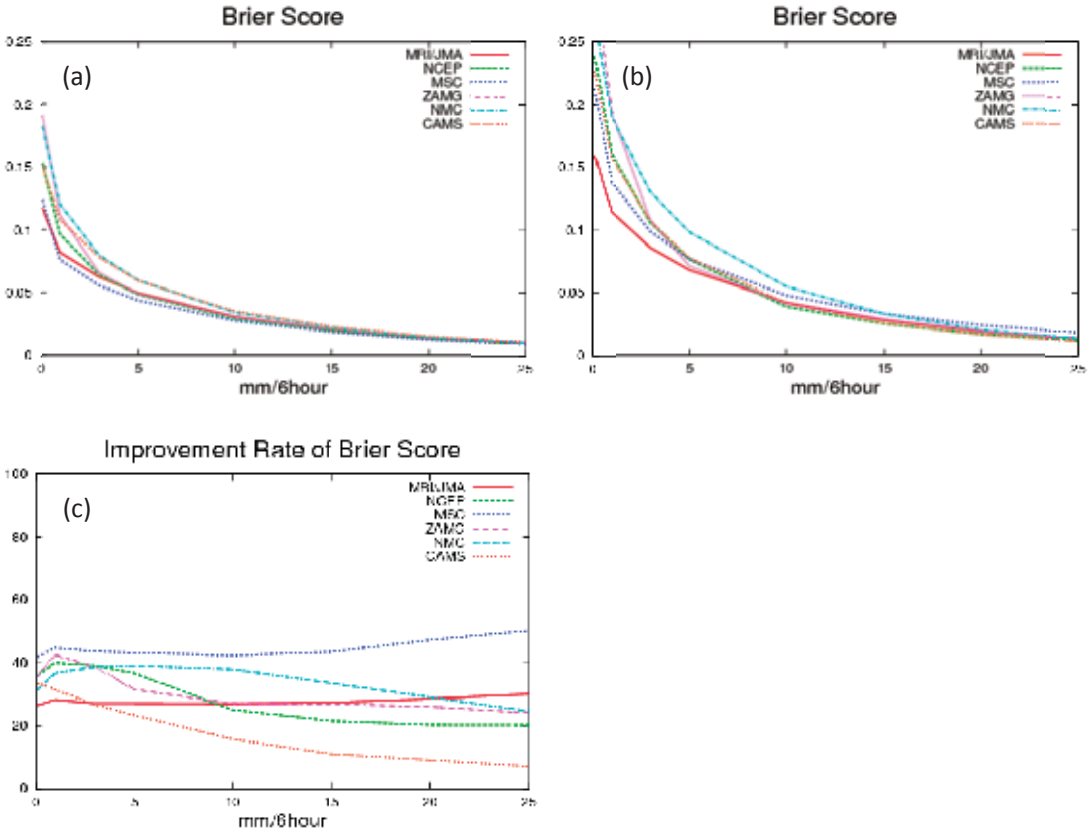


Fig. E-7-14. Brier score of 6-hour accumulated precipitation for (a) ensemble forecast, (b) a control forecast. (c) The improvement rate of the Brier score.

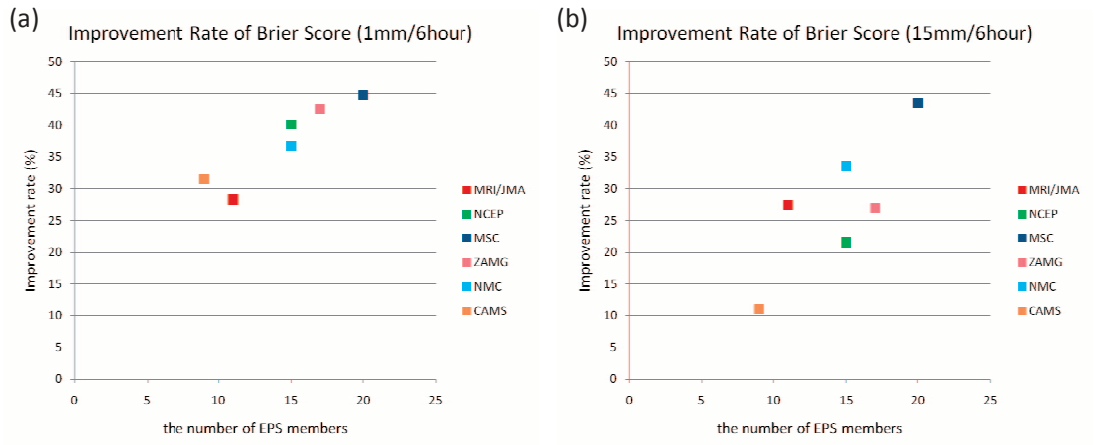


Fig. E-7-15. The relationship between the improvement rate of Brier score (6-hour precipitation) and the number of EPS members. The thresholds are (a) 5mm / 6hour, (b) 15mm / 6hour.

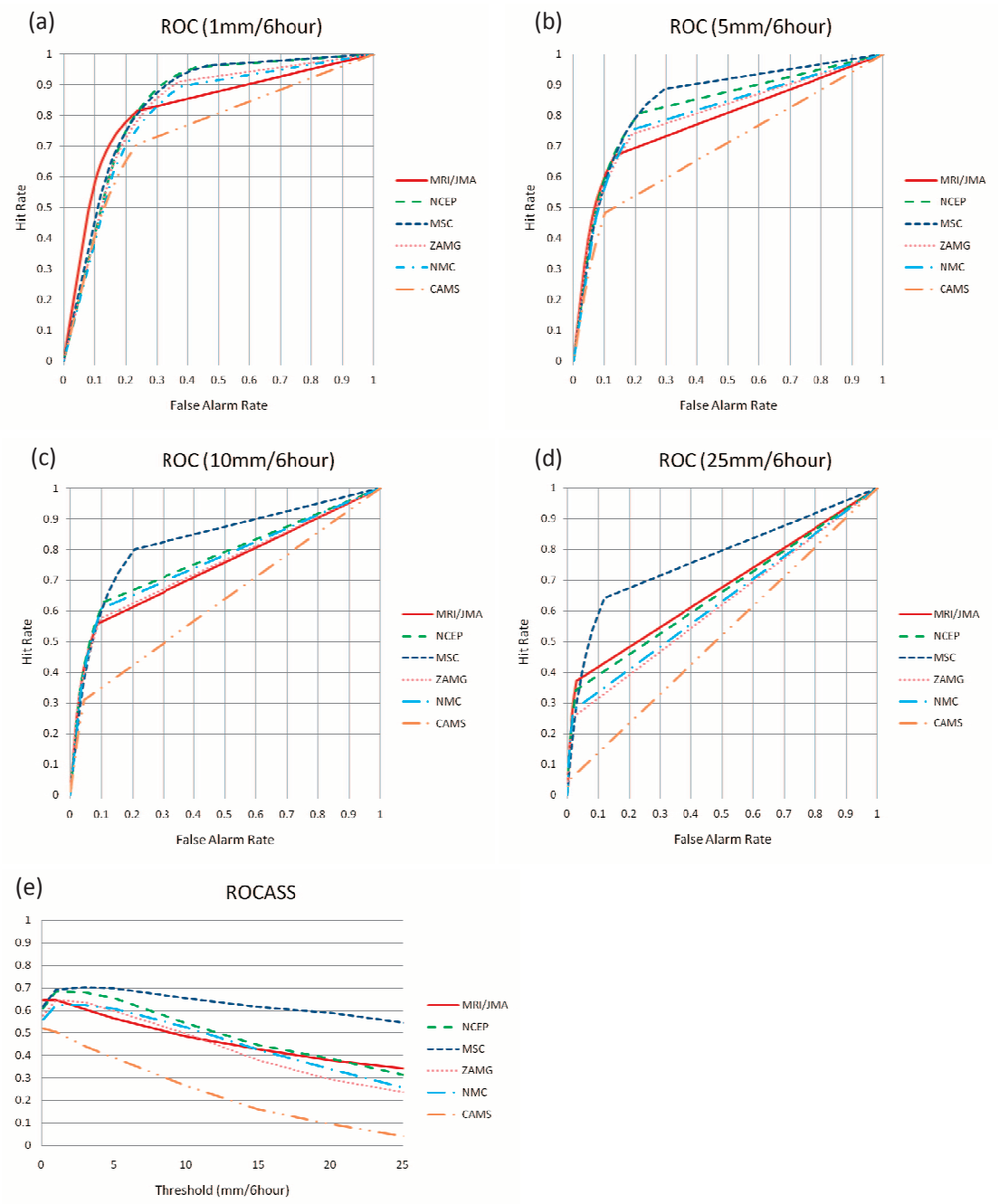


Fig. E-7-16. ROC diagrams for 6-hour accumulated precipitation, the thresholds of which are (a) 1mm, (b) 5mm, (c) 10mm, (d) 25mm. (e) ROC area skill score.

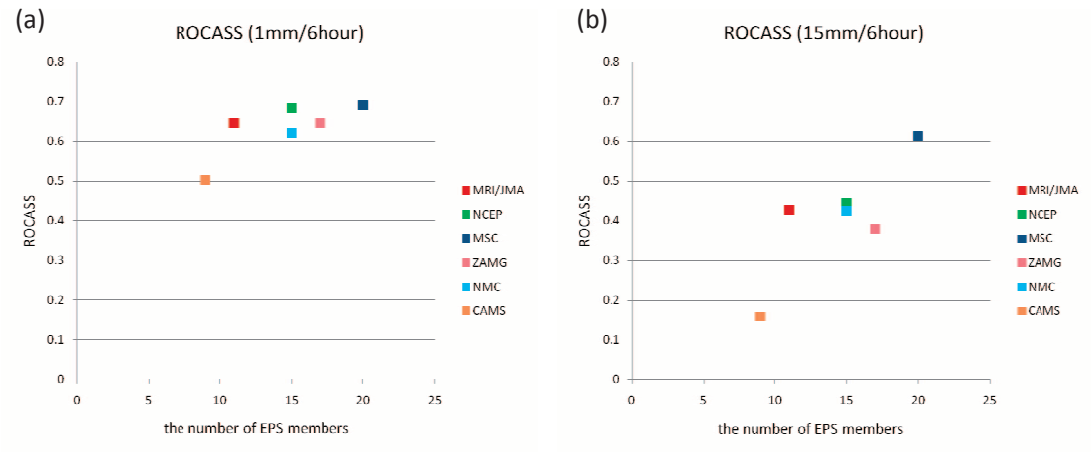


Fig. E-7-17. Same as Fig. E-7-15, but for ROC area skill score.

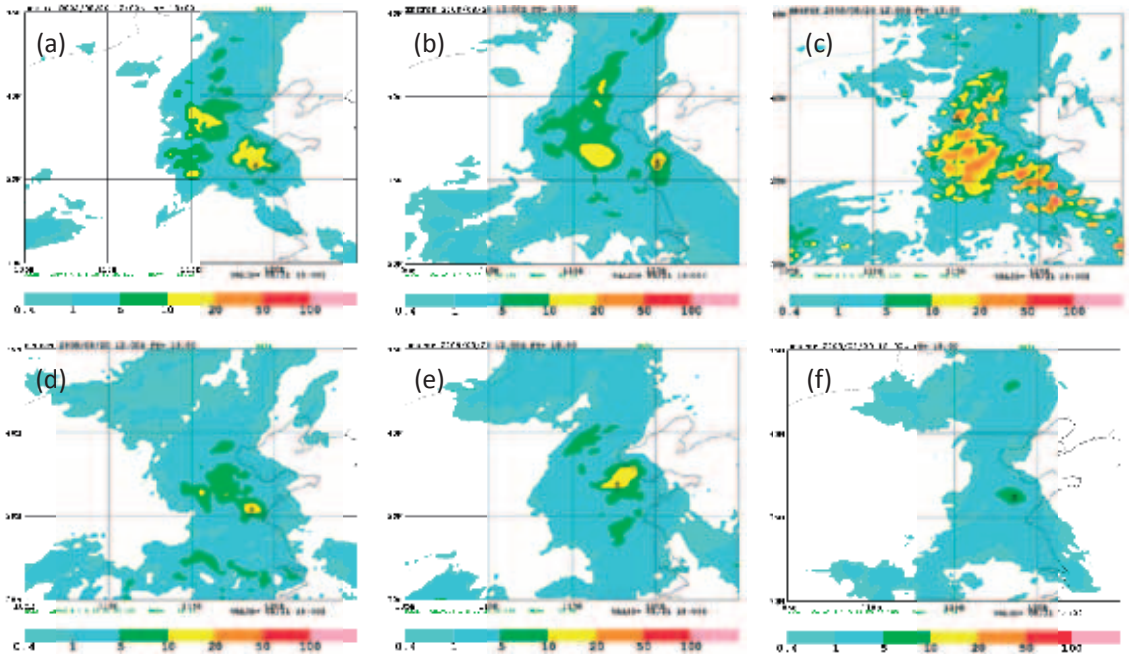


Fig. E-7-18. Ensemble spread of 3-hour accumulated precipitation. Initial time is 12 UTC 20 August 2008 (FT=18). (a) MRI/JMA, (b) NCEP, (c) MSC, (d) ZAMG, (e) NMC, (f) CAMS.

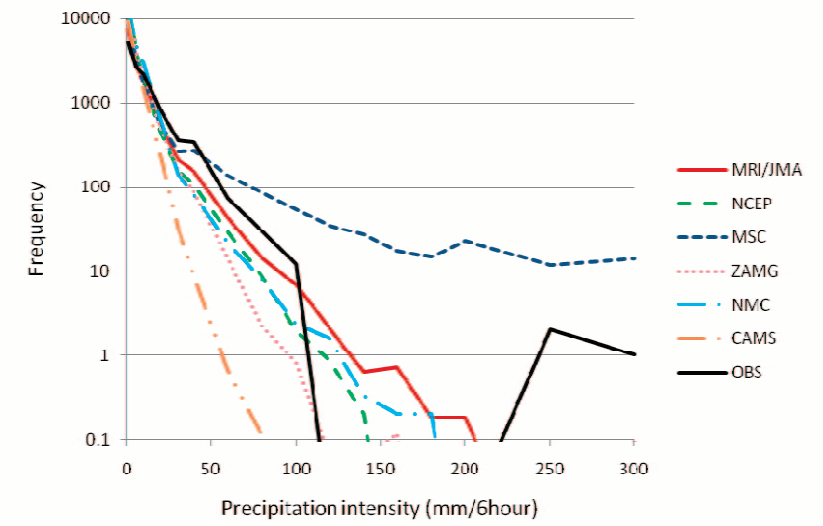


Fig. E-7-19. Histograms for 6-hour precipitation intensity. The scale of the vertical axis is logarithmic.

E-7-3. Verification by CMA

In this section, verification results of the 2008 experiment presented in the 4th B08FDP/RDP workshop by CMA (Li Li, 2009) are partly introduced. As in the 2007 experiment (D-7-2), forecast results in the common verification grids with a resolution of 0.15° (105~125°E, 30~45°N) were compared with 400 synoptic observation stations and 722 dense auto observation stations. Since 2007, upper air variables have also been verified with 28 radio sonde stations (Fig. D-7-1).

Figure E-7-20 shows the ensemble spread and RMSE of temperatures at 850 hPa level predicted by the MRI's EPS for 55 days from 1 July to 24 August 2009. Note that this period is longer than the formal intercomparisons period (30 days from 24 July to 23 August), thus MRI's EPS was not well tuned in the pre-formal period (1 July to 23 July), where the ensemble spread was small. Despite this flaw of EPS, compared with the 2006 and 2007 verifications (Figs. C-6-1 and D-7-2), RMSE against sondes became smaller while the ensemble spread at FT=36 became larger. These improvements seem attributable to the implementation of the Meso 4Dvar analysis in the control run and lateral boundary perturbations. Magnitude of RMSE is larger than the MRI's verification for the formal intercomparison period (E-7-2) because MRI's verification was against initial conditions but CMA's verification was performed against sondes. Improvements of RMSE and spread are more obvious in the forecast of 500 m height (Fig. E-7-21).

Talagrand diagram for 2m temperature and 10m wind at FT=36 is shown in Fig. E-7-22. The temperature (Fig. E-7-22a) still has the U-shaped pattern, but the histogram was ameliorated compared with the 2007 experiment (Fig. D-7-5). However, amelioration in the 10m wind speed was not clear and strong bias remained for weak wind observations (Fig. E-7-22b). This histogram bias may be explained by the following reason; in the anemometer observation, calm is dominated in the cases of weak wind situations but weak wind speeds remain in the NWP output.

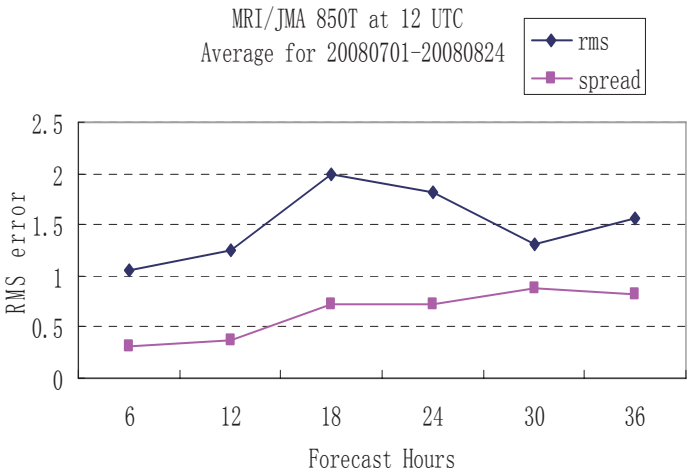


Fig. E-7-20. Ensemble spread and RMSE of temperatures at 850 hPa level. Period is 57 days from 1 July to 31 August 2008. After Li Li (2007). Courtesy of NMC/CMA.

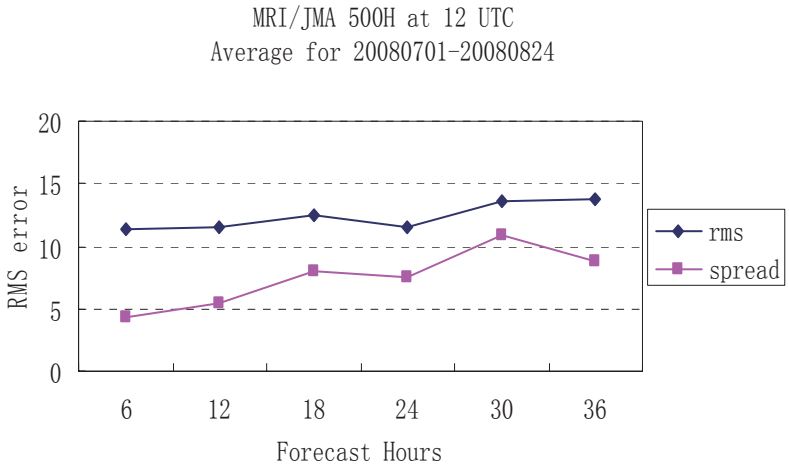


Fig. E-7-21. Same as in Fig. E-7-1, but for 500 hPa height. After Li Li (2008). Courtesy of NMC/CMA.

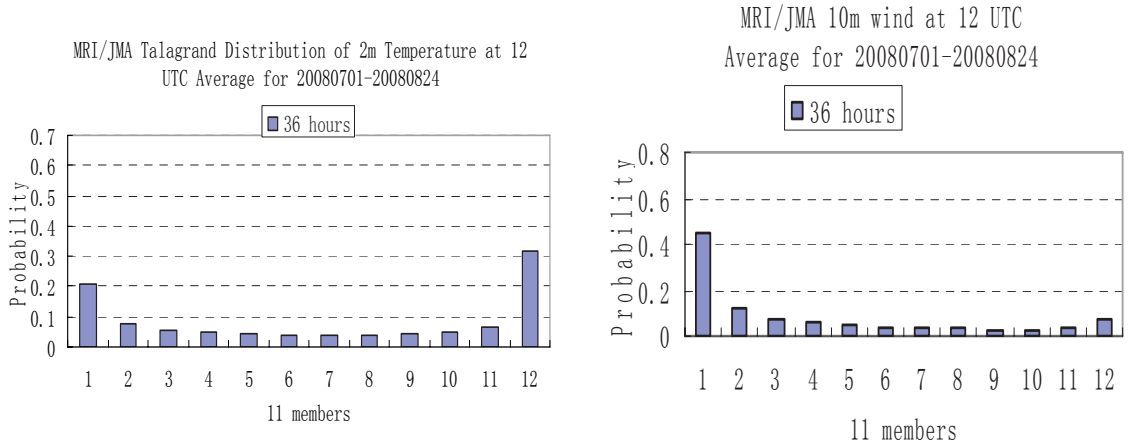


Fig. E-7-22. Talagrand diagram on the surface (2m) temperature (left) and the 10m wind speed (right) at FT=36 of the MRI/JMA ensemble forecast. Period is 57 days from 1 July to 31 August 2008. After Li Li (2009). Courtesy of NMC/CMA.

E-8. Contribution to the FDP project

E-8-1. NWP Support to the Hong Kong Observatory

Analysis and forecast data of the JMA Global Spectral Model (GSM) were provided in full model resolution by MRI from May to September 2008. GSM data were used as the initial and boundary conditions of the NHM (version NHM-0712) in support to the operation of: (1) HKO nowcasting system – SWIRLS (Short-range Warning of Intense Rainstorms in Localized Systems) in B08FDP (see E-8-2); and (2) weather service provided to the Hong Kong Windsurfing Team for the Olympics windsurfing races in Qingdao.

GSM forecast data from JMA were sent to MRI after several hours delay for the B08RDP experiment and process of boundary data for NHM operated by HKO. To save the transmission time, the spatial coverage of boundary file was reduced to just cover the above two forecast domains of NHM before the boundary data were upload to the MRI's ftp site. These data were prepared four times per day with initial time of GSM at 00, 06, 12 and 18 UTC.

As the computational domain of NHM for B08FDP was not fully covered with the mesoscale analysis of the JMA, the global land surface analysis was provided for the operation of NHM. Special treatment was applied to use the soil temperatures in the global analysis due to the different land process and configuration adopted in the mesoscale model (E-3-3).

Figures E-8-1 and E-8-2 show examples of surface forecasts and the full spatial coverage of NHM in the Beijing and Qingdao domains. Details on model configurations are summarized in Table 1. NHM in the Beijing domain was operated on an hourly basis at 5-km horizontal resolution up to a forecast range of 9 hours to provide model precipitation forecasts for blending with the SWIRLS rainfall nowcast product. In Qingdao domain, NHM was operated once per day at a horizontal resolution of 3 km and a forecast range of 18 hours (from 18 UTC to 12 UTC the following day). Graphical products were made available on a dedicated web site of HKO for reference by the windsurfing team. Besides, NHM output on wind speed, wind direction and temperature at the windsurfing venue were combined with site-specific prognoses from WRF model (version 2.2.1, also at a horizontal resolution of 3 km) to generate forecast guidance for the windsurfers.

E-8-2. SWIRLS in B08FDP

The second generation of the HKO nowcasting system SWIRLS (Short-range Warning of Intense Rainstorms in Localized Systems), or SWIRLS-2 in short (Yeung *et al.* 2009), was deployed to support the nowcast service during the Beijing Olympic Games. Apart from functionalities on radar-based quantitative precipitation estimation and forecast in the original system, new techniques were successfully implemented in SWIRLS-2, namely: (a) blending and combined use of radar-based nowcast and high-resolution NWP model analysis and forecast based on 5-km NHM, also known as RAPIDS (Rainstorm Analysis and Prediction Integrated Data-processing System) (Wong and Lai, 2006, Wong *et al.* 2009,); (b) conceptual model-based detection and nowcasting of high-impact weathers including lightning, severe squalls and hail; (c) a grid-based, multi-scale and more robust storm-tracking method; and (d) probabilistic representation of nowcast uncertainties arising from storm tracking, growth and decay. Figures E-8-3 (a)-(d) show examples of SWIRLS-2 nowcast products during the B08FDP operation period. More details on the design and technological

improvements in SWIRLS-2 and RAPIDS can be found in Yeung *et al.* (2009) and Wong *et al.* (2009). In particular, the provision of GSM data at 20-km full resolution facilitated a much simpler set-up of the 5-km NHM by removing the need of an intermediate step of 10-km NHM used in the two trial operations in 2006 and 2007. This also allowed the 5-km NHM to be operated on an hourly basis. In addition, GSM forecasts were utilized as the first guess field in the JNoVA-3DVAR data assimilation system with horizontal resolution of 5 km to produce the initial condition of NHM. The observations assimilated in the 3DVAR included local data like automatic weather stations and radiosonde, as well as the derived vertical profiles of wind, temperature and moisture over rainy areas from the LAPS (Local Analysis and Prediction System). In LAPS, radar reflectivity, Doppler radial velocity, satellite cloud imagery, surface and radiosonde data were analysed by different modules to generate a three dimensional analysis of wind, temperature and moisture (Albers *et al.* 1996). The overall effect of the first guess from GSM and assimilated observations in the JNoVA analysis resulted in an improved initial field of 5-km NHM and the model precipitation in the first few hours of forecasts (Fig. E-8-4).

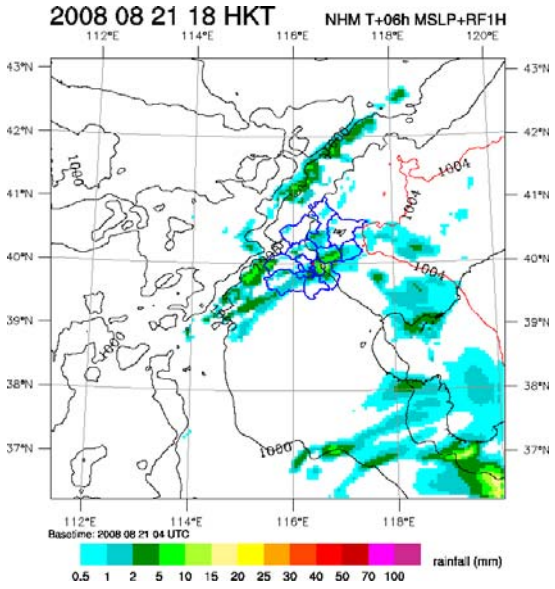


Fig. E-8-1. Map showing the forecast domain of NHM in B08FDP operation with blue lines drawing the administrative boundaries of Beijing city and red dot the location of Olympic Park. Also overlaid are the 6-hour NHM forecasts for mean-sea-level pressure (isobars as dark lines) and 1-hour accumulated rainfall (colour pixels) initialized at 04 UTC (or 12 noon), 21 August 2008, indicating significant precipitation might occur in Beijing in that afternoon.

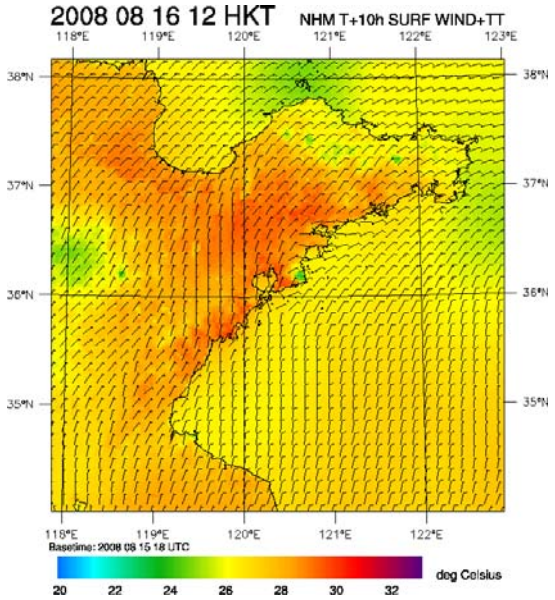


Fig. E-8-2. Domain of 3-km NHM in Qingdao forecast support. 10-hour NHM forecasts for winds and temperature (colour pixels) initialized at 18 UTC 15 August 2008 are shown which indicate the convergence of winds in the vicinity of the windsurfing venue as the sea-breeze began to establish around noon on 16 August.

Table E-8-1. Configuration of NHM for the support of B08FDP and Qingdao forecasts.

	Beijing Domain	**Qingdao Domain**
System (version no.)	**JMA-NHM (0712)**	
Horizontal resolution	5 km	3 km
Number of horizontal grid points	151x151	151x151
Number of vertical levels	50	45
Forecast range	9 hours	18 hours
Update cycle	Hourly	Once per day with initial time at 18 UTC
Horizontal grid	Arakawa C	
Vertical coordinates	Terrain following height coordinates (z* coordinates) on Lorenz grid	
Initial condition	- JNoVA-3DVAR - Specific humidity of water vapour, cloud liquid water, cloud ice, snow and graupel can be initialized by LAPS cloud moisture analysis - Sub-surface temperatures (4 layers) initialized by blending GSM land surface analysis with previous NHM forecast	- 6 hour forecast from full resolution 20 km GSM at 12 UTC - Sub-surface temperatures (4 layers) initialized by blending GSM land surface analysis with 6-hr NHM forecast initialized at 12 UTC
Boundary condition	JMA GSM at 20 km resolution	
Nesting configuration	One-way nesting	
Map projection	Lambert conformal	
Topography and land-use	USGS GTOPO30 and GLCC2	
Dynamics	Fully compressible non-hydrostatic governing equations, solved by time-splitting horizontal-explicit-vertical-implicit (HEVI) scheme and 4-order spatial centred finite differencing in flux form	Same as 5 km NHM but without splitting of gravity wave and momentum terms.
Model physics	- Kain-Fritsch convective parameterization scheme - Three ice bulk cloud microphysics scheme - Flux and bulk coefficients based on Beljaars and Holtslag (1991) - Mellor-Yamada-Nakanishi-Niino Level 3 (MYNN-3) turbulence closure model with partial condensation scheme (PCS) (see F-5) - Long wave radiation process follows Kitagawa (2000) - Short wave radiation process using Yabu *et al.* (2005) (as in JMA-GSM-0412)	

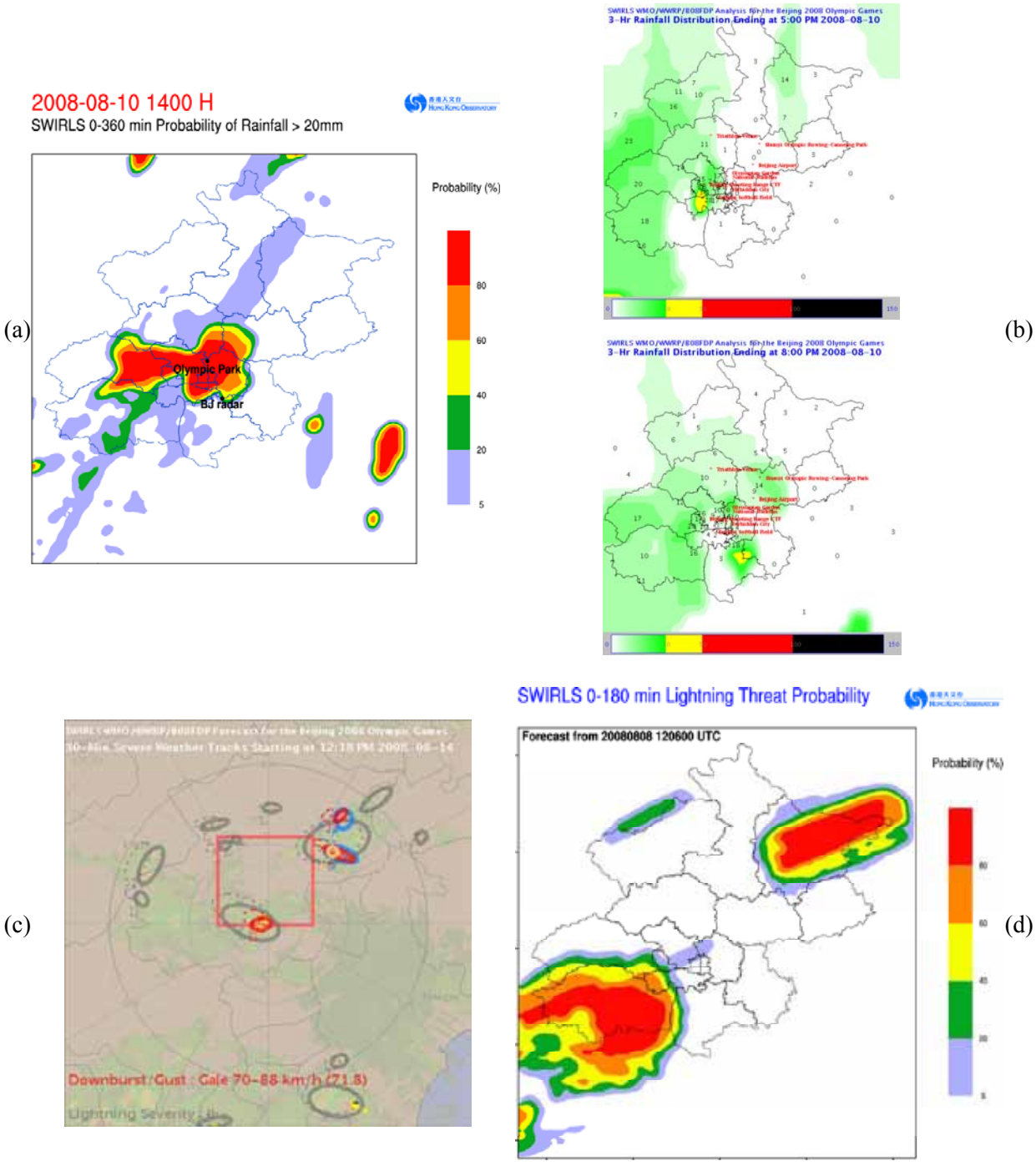


Fig. E-8-3. Real-time examples of SWIRLS-2 products: (a) Probability of precipitation (PoP) forecast from RAPIDS generated at 2 p.m. on 10 August 2008 with threshold greater than 20 mm in the next 6 hours; (b) the actual 3-hour rainfall ending at 5 p.m. and 8 p.m.; (c) Severe weather map generated at 12:18 pm on 14 August 2008. Four types of high-impact weather were shown as ellipses in different colours, namely rainstorm in blue, CG lightning in grey, downburst/squalls in red and hail in orange. Solid and dashed lines represent respectively their analyzed and predicted 30-min locations; (d) Probability of lightning (PoL) threat forecast for the next 3 hours generated at 8:06 p.m. on 8 August 2008 for opening ceremony of the Beijing Olympic Games.

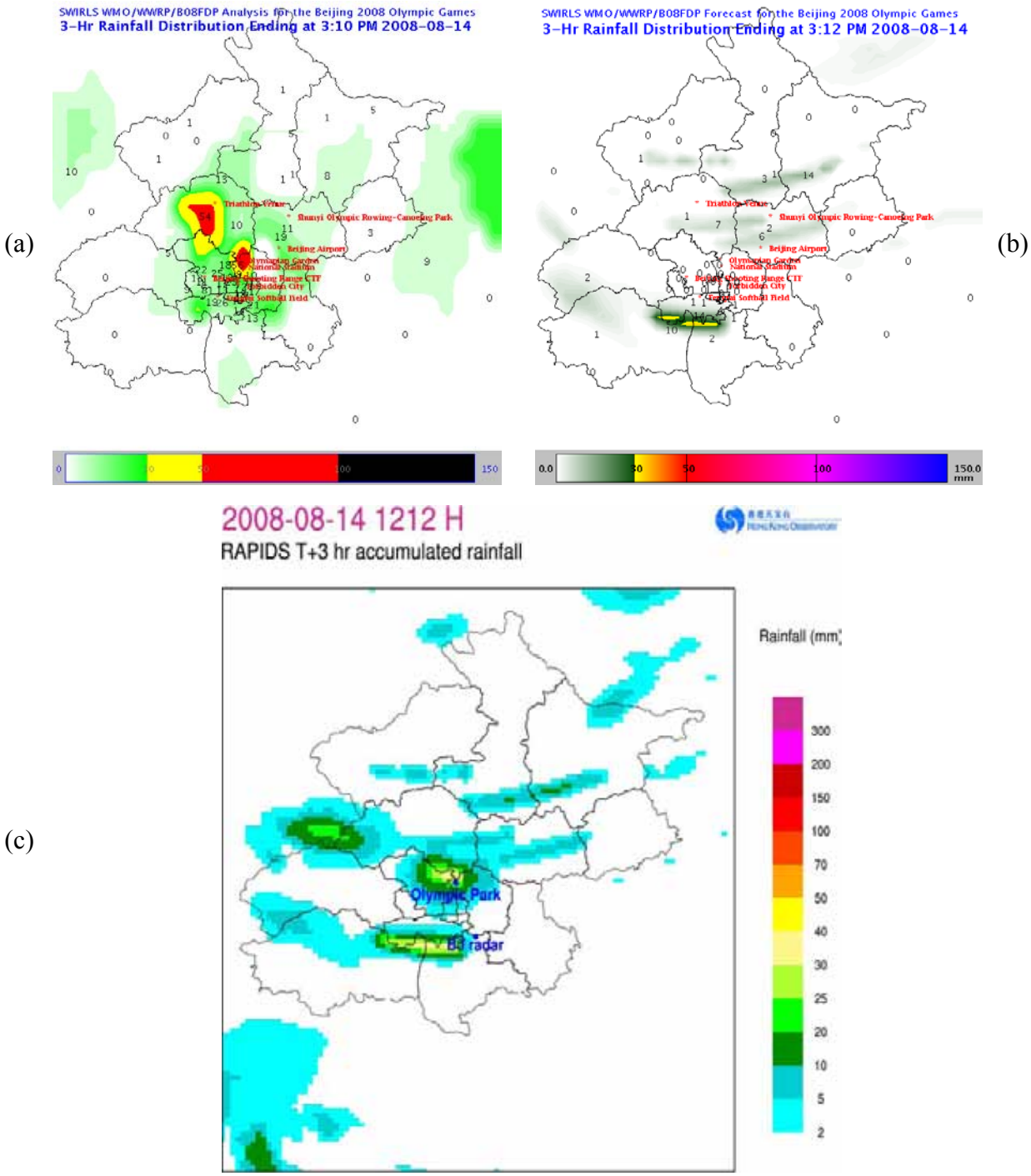


Fig. E-8-4. Example of RAPIDS forecast for 14 August 2008: (a) actual 3-hour rainfall accumulation ending at 3:10 p.m. local time, estimated from raingauge data; (b) SWIRLS radar-based 3-hour rainfall nowcast for nearly the same period (ending at 3:12 p.m. local time) with an underestimation trend generally seen; (c) the corresponding rainfall forecast from RAPIDS, demonstrating an improvement over heavy rain regions (e.g. near the Olympic Park) made possible by blending NHM forecast initialized with JNoVA-3DVAR.