Verification of high resolution ensemble forecasts using the 4-dimensional Fractions Skill Score

Le Duc[12], Kazuo Saito[21], and Hiromu Seko[21]

[1]Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan; leduc@jamstec.go.jp

[2]Meteorological Research Institute, Tsukuba, Japan

We conducted two 11-member ensemble prediction systems MF10km and MF2km with the resolutions of 10 km and 2 km, the later nested inside the former with a 6-hour lag, for 15 days in the summer of 2010. The two ensemble systems were verified to examine the value of cloud-resolving ensemble forecast in predicting small spatiotemporal-scale precipitation. Therefore we focused on short time (1-hour) rainfall with a small scale (5 km) verification grid. The domains of the two systems and the verification area are shown in Fig. 1.

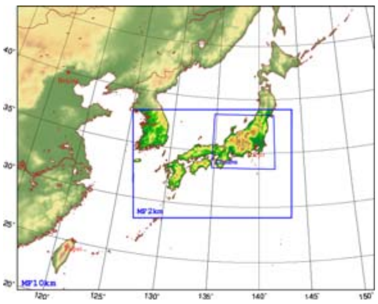


Fig. 1. Domains of MF10km and MF2km and verification area.

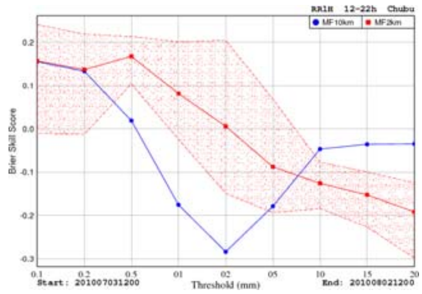


Fig. 2. BSS of hourly precipitation forecasts from MF10km and MF2km.

Since the verification was performed on short-term precipitation at high resolution, uncertainties from small-scale processes in space and time caused the traditional verification methods inconsistent with the subjective evaluation. This is illustrated in Fig. 2 and Fig. 3 in which the Brier Skill Scores indicate that two systems had no skills with respect to moderate and heavy rain, whereas the “eyeball” verification did not.

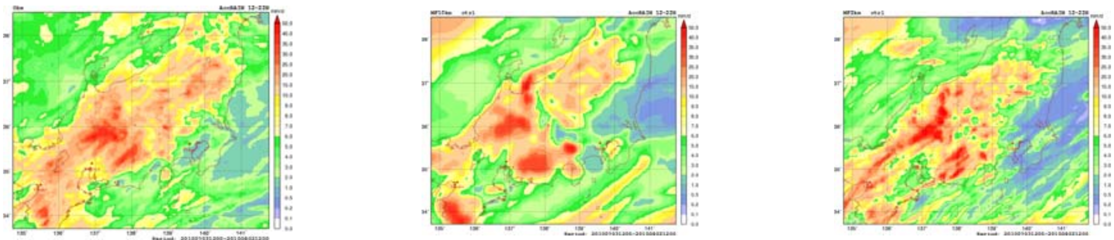


Fig. 3. Rainfall analysis (left) and corresponding forecasts by MF10km (center) and MF2km (right) controls (the average rainrates, unit mm d^{-1}, are shown in the plot).

The fall of the traditional verification methods with high resolution precipitation forecasts is attributed to the fact that the short predictability of small-scale processes in space and time was not taken into account in these methods. This problem remains even in the ensemble forecast if the verification grid size is small. In recent year, a number of new verification methods were proposed to account for spatial mismatches between deterministic forecasts and observations. However, not many studies were performed on the time lag issue. The other issue is how to apply the spatial verification methods into ensemble forecast. In this study, we introduced an extended verification method based on the Fractions Skill Score (FSS) by Roberts and Lean (2008) to account for these uncertainties. The main idea is to extend the concept of spatial neighborhood in FSS to the time and ensemble dimension.

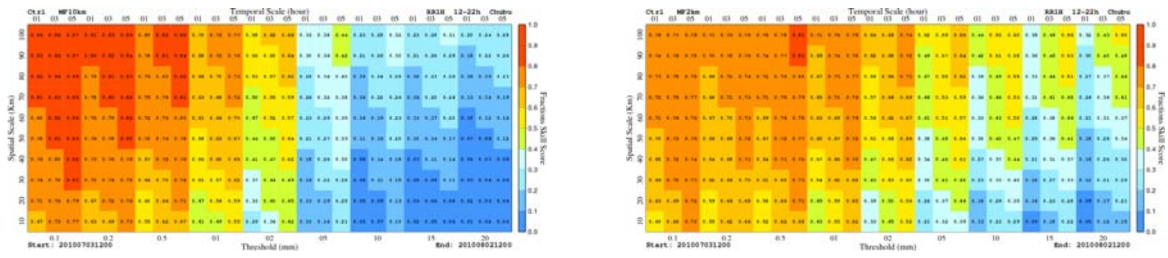
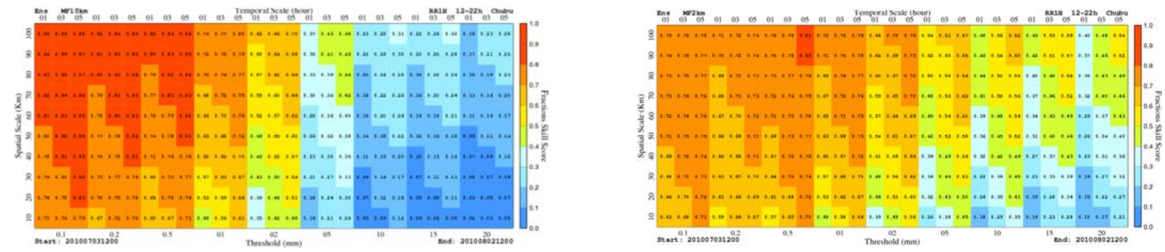


Fig. 4. Intensity-scale diagrams with temporal scales incorporated.

The incorporation of the time dimension into FSS requires a new form for intensity-scale diagrams. For each intensity value in the intensity-scale diagram, an additional horizontal temporal scale axis will be embedded. This new intensity-scale diagram is given in Fig. 4. In this figure, FSS-constant lines with a slope of -20 km / 2 hour can be identified in each spatial-temporal sub-diagram, indicating that the FSS values at small spatial and long temporal scales are equal to the ones at large spatial and short temporal scales. This reveals the important role of temporal scales in short-term precipitation verification at small spatial scales.



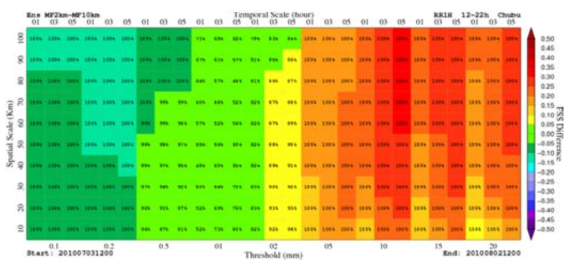


Fig. 5. Extended intensity-scale diagrams showing FSSs from MF10km (top left) and MF2km (top right) hourly precipitation forecasts and their differences (bottom left).

The performances of two systems as measured by the extended FSS are shown in Fig. 5. MF2km clearly outperforms MF10km in predicting heavy rain. In contrast, MF10km is slightly better than MF2km with respect to light rain. Two systems have the same performance with respect to moderate rain. The fact that the lower-resolution forecasts MF10km were better than the high-resolution forecasts MF2km for light rains suggests that the horizontal resolution of 2 km is not necessarily fine enough to completely remove the convective parameterization.

In addition to the FSS, the neighborhood concept was also incorporated to reliability diagrams and relative operating characteristics to verify the reliability and resolution of two systems. In verification of MF2km, the reliability and resolution with spatial lag of 20 km and the temporal scale of 1 hour was almost the same that with temporal lag of 2 hours and the spatial scale of 10km (not shown here).

Reference

Roberts, N. M., and Lean H. W., 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97.