Research on fuzzy verification in high-resolution precipitation forecasts

Yasutaka Ikuta[1], Yuki Honda[2]
Numerical Prediction Division, Japan Meteorological Agency,
1-3-4, Otemachi, Chiyoda-ku, Tokyo 100-8122, Japan

A high-resolution local forecast model (LFM) with a grid spacing of 2 km is being developed by the Japan Meteorological Agency. The LFM is nested into the operational mesoscale model (MSM), which is another limited-area model with a grid spacing of 5km. To compare the performance of the LFM with that of the MSM, verification of quantitative precipitation forecasts (QPFs) is required. However, traditional verification methods such as bias score and threat score are not adequate because these methods cannot handle forecast displacement error appropriately. This kind of issue regarding verification scores is often referred to as the 'double penalty' problem.

For the last several years, the verification for high-resolution model forecasts has been discussed enthusiastically in research fields, and several new verification techniques have been proposed. One of the main characteristics of these new techniques is the introduction of the concept of spatial scale into the verification. The new verification techniques tolerate the difference of the spatial structure pattern within the spatial scale, and are referred to as *spatial verification* methods. One of them is the fuzzy verification methods, which considers the precipitation of neighborhood grids. In this report, several fuzzy verification methods are tested to evaluate the QPFs of the LFM and the MSM.

We tested the following fuzzy verification methods: 1) Upscaling (Yates et al., 2006); 2) Fuzzy Logic (Damrath, 2004; Ebert, 2002); 3) Intensity Scale (Casati et al. 2004); 4) Fractions Skill Score (Roberts and Lean 2008). More fuzzy verification methods in addition to the ones given here are listed in Ebert (2008). We ascertained the characteristics of these methods as described below.

The upscaling method has the problem of providing no computed values for large spatial scales and high thresholds. Fuzzy logic, on other hand, gives uncertain scores for large spatial scales. As an example, in the perfect forecast of an idealized experiment, the ETS becomes a value other than 1. The Intensity scale skill score (ISS) can detect displacement error, but is insensitive to the frequency bias of forecasts. The Fractions Skill Score (FSS) provides a target skill giving a general criterion for a skillful forecast. This target skill is defined as half the value of the scores between a random forecast and a perfect forecast, and a poor forecast with typical displacement larger than twice as long as neighborhood length gives score under the target skill. In addition, the rates exceeding the target skill provide an estimate of the minimum useful scale, but this scale becomes useless in forecasts with large biases. The forecast frequency bias is represented by a conventional bias f_M/f_O, where f_O is the observed frequency and f_M is the model-forecast frequency. This conventional bias is equal to the bias score from a contingency table in the grid scale. If the neighborhood length covers the whole verification area, the FSS becomes $2f_Of_M/\left(f_O^2+f_M^2\right)$. This skill score is called the *asymptotic fractions skill score* (AFSS). In particular, the AFSS of a no-bias forecast is equal to the value of 1.

We have reached the conclusion that upscaling and the fractions skill score are appropriate for verification of high-resolution QPFs. In particular, the FSS provides new information on displacement error as the target scale. The following discussion of a real case study is focused on the FSS approach. Figure 1 shows FSSs with two different spatial scales of QPFs of the MSM and the LFM for April 2008. Figure 2 shows a comparison with the averaged conventional bias of these values. In the grid scale (Fig. 1 (a)), the LFM had a lower value for the skill score and the over-target rate than the MSM at almost all thresholds, and the AFSS of the LFM was larger than that of the MSM for the 15 mm/3 h and the 20 mm/3 h thresholds. For thresholds lower than 10 mm/3 h the AFSS of the LFM was close to that of the MSM, when the LFM conventional bias was smaller than that of the MSM. In contrast, the frequency bias of the MSM shows a low value for high thresholds. Figure 1 (b) shows that, for a spatial scale of 80 km, the over-target rate of the LFM becomes similar to that of the MSM. Moreover, these LFM scores are larger than those of the MSM for the 15 mm/3 h, 20 mm/3 h and 30 mm/3 h thresholds.

In this study, the frequency of high-resolution forecasts was small for low thresholds and large for high thresholds. Moreover, for large thresholds, the FSS of low-resolution forecasts was lower than that of

[1] E-mail: ikuta@met.kishou.go.jp
[2] E-mail: honda.yuuki@met.kishou.go.jp

high-resolution forecasts. This FSS behavior could represent the characteristics of high-resolution precipitation forecasts which were usually underestimated as a result of displacement error. Figure 1 (b) shows that the increase in spatial scale gave LFM scores that were larger than those of the MSM at the 15 mm/3 h, 20 mm/3 h and 30 mm/3 h thresholds. These results indicate that introducing an adequate spatial scale in precipitation verification successfully relaxes the requirement for perfect matching. Accordingly, it can be seen that the FSS represents be an appropriate approach for verification of the high-resolution model.

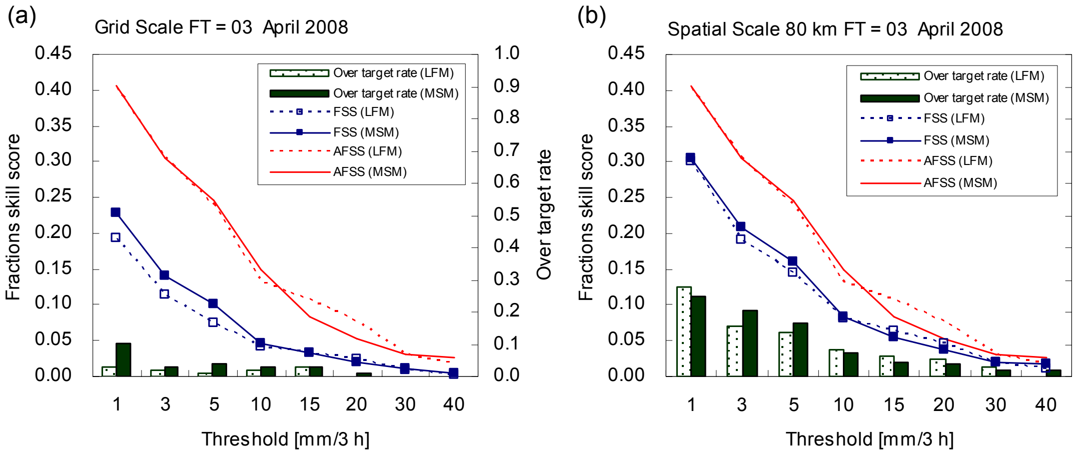

Figure 1. Fractions skill score (line plots with symbols), asymptotic fractions skill score (line plots) and rate of over-target skill (bar charts) are shown as a function of threshold and three hour accumulation. Each score was calculated in (a) the grid scale, and (b) the spatial scale of 80 km.

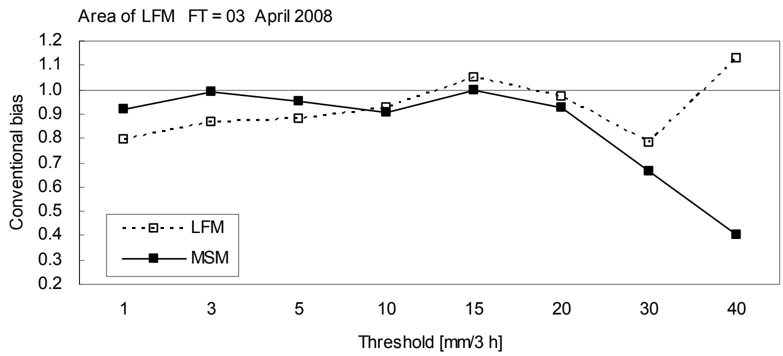

Figure 2. Conventional bias $\overline{f_M}/\overline{f_O}$ (line plots with symbols) is shown as a function of threshold for April 2008, where over-bars $(\overline{\ })$ indicate the monthly average.

References

Yates, E., S. Anquetin, V. Ducrocq, J.-D. Creutin, D. Ricard and K. Chancibault, 2006: Point and areal validation of forecast precipitation fields. *Meteorol. Appl.*, **13**, 1-20.

Damrath U. 2004: Verification against precipitation observations of a high density network – what did we learn? In International Verification Methods Workshop, Montreal, 15–17 September 2004.

Ebert EE. 2002: Fuzzy verification: giving partial credit to erroneous forecasts. In NCAR/FAA Verification Workshop: Making Verification More Meaningful, NCAR, Boulder, 30 July – 1 August 2002.

Casati, B., Ross, D.B. Stephenson, 2004: A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorol. Appl.*, **11**, 141-154.

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

Ebert EE, 2008: Fuzzy verification of high resolution gridded forecasts: A review and proposed framework. *Meteorol. Appl.*, **15**, 51–64.