P1.5

McGILL ALGORITHM FOR PRECIPITATION NOWCASTING BY LAGRANGIAN EXTRAPOLATION (MAPLE) APPLIED TO THE SOUTH KOREAN RADAR NETWORK. PART 2: REAL-TIME VERIFICATION FOR THE SUMMER SEASON

Hee Choon Lee *†, Yong Hee Lee†, Jong-Chul Ha†,Dong-Eon Chang†,
Aldo Bellon§, Isztar Zawadzki§, and Gyuwon Lee‡
†National Institute of Meteorological Research, KMA, Seoul, Rep. of Korea
§Department of Atmospheric and Oceanic Sciences, McGill University, Montreal, Quebec, Canada
‡Department of Astronomy and Atmospheric Sciences, Kyungpook National University,
Daegu, Rep. of Korea

1 INTRODUCTION

Regarding weather prediction, the general public is concerned essentially with whether it will rain or not. The onset time and the amount of rainfall are other inevitable matters of concern when rainfall is predicted. People in the fields of agriculture, transportation, general services and sports and leisure are equally concerned about the quality of meteorological information. However, in spite of this overall interest, the prediction of precipitation is one of the most difficult aspects of weather forecasting because of the large spatial and temporal variability of rain and of its imperfect parameterization. The main reason for this difficulty is that precipitation processes are not controlled by one specific factor, but by a multitude of factors. Although many other types of meteorological information are forecasted, quantitative precipitation forecast (QPF) is a primary issue in the assessment of numerical weather prediction (NWP) models. Ebert *et al.* (2003) assessed the QPFs from several operational NWP models and evaluated their predictabilities and characteristics with various verification techniques. Similar comparisons were performed through the Sydney 2000 Forecast Demonstration Project (Ebert *et al.* 2004). The predicted intensity, occurrence and location of rainfall were evaluated with the traditional verification methods. A prediction system must be objectively verified in order to assess further development aimed at improving the original algorithm. In addition to the traditional verification techniques that use a simple contingency table, a variety of verification techniques that consider the spatial characteristics of the prediction as well as the ensemble prediction and the socio-economic perspective have been developed and applied by Casati *et al.* (2008).

Since 2007, the National Institute of Meteorological Research (NIMR) of Korea Meteorological Administration (KMA) has developed a nowcasting system for the Korean Peninsula in collaboration with McGill University in order to enhance the short-term predictability of precipitation. The system called MAPLE (McGill Algorithm for Precipitation nowcasting by Lagrangian Extrapolation: Germann and Zawadzki 2002) uses radar composite maps to predict the location of precipitation echoes several hours in advance (up to 6 hours) using the variational echo tracking (VET) method and a semi-Lagrangian backward advection technique (Turner *et al.* 2004). This system has been operating in real-time since June 2008, the output being used in operations by KMA′s weather forecasters and hydrologists.

There are two main considerations in the evaluation of the predictability of the MAPLE system, namely, which technique and which ground truth ought to be used. Regarding the former, one of the options is the traditional verification technique that does not consider the spatial attributes of the predicted maps. The relation between forecast and observation is simply evaluated at each grid point by means of a simple contingency table. Small location errors can mislead the forecast performance in this approach. We seek a more objective verification technique in order to evaluate both the spatial distribution and the location of the pattern of predicted rainfall in relation

* *Corresponding author address:* Hee Choon Lee, National Institute of Meteorological Research, KMA, Seoul, Rep. of Korea 156-720; e-mail: hclee@kma.go.kr

to that observed. Therefore, in this study, we implement the Contiguous Rain Area (CRA) method proposed by Ebert and McBride (2000), one of the objective-oriented verification methods that is best suited for our intended purpose. This method quantitatively separates the location, volume and pattern error that are included in the total error. Grams *et al*. (2006) used a modified CRA technique to identify the systematic sources of error in the forecasts of convective systems during the 2002 International H_2O Project (IHOP) made with the Eta, MM5 and WRF models.

Regarding the other consideration, rainfall forecasts of NWP models are generally evaluated with rain gauge data using the traditional statistical methods (Savvidou *et al*. 2007). However,Tartaglione *et al*. (2008) demonstrated that some problems induced by the uncertainty in their representation of a mean area rainfall can occur when gauges are used. Unlike other observations such as temperature, the space-time characteristics of the ground truth should be similar to those of the predicted rainfall fields. Therefore, we have simply taken the observed radar rain rate pattern corresponding to the forecast time as ground truth.

The purpose of this study is thus to evaluate the predictability of the MAPLE system over the Korean Peninsula and to provide the guidelines for its practical interpretation. The evaluation of the prediction is performed for the summer period of 2008 using the CRA technique. Section 2 of this paper provides the characteristics of the MAPLE system, section 3 describes the CRA technique and section 4 summarizes the results of the CRA analysis.

2 NOWCASTING MODEL

The MAPLE system provides the prediction of rain echoes over several hours. The Variational echo tracking (VET) method, originally suggested by Laroche and Zawadzki (1995), is used to estimate the optimum field of motion vectors as described by Germann and Zawadzki (2002). The sensitivity of the many VET input parameters on the resultant field of motion is examined in a companion paper by Lee *et al*. (2009). Three radar composite maps of maximum reflectivity (along the vertical) are used to estimate the field of motion. The model domain is (1024 $\times$ 1024) pixels at 1 km horizontal resolution which covers the Korean radar network. The echo motion vectors are calculated over (25 $\times$ 25) sub-areas that cover a sub-domain of (800 $\times$ 800) pixels, each sub-area thus representing an area of (32 km by 32 km). In order to apply the semi-Lagrangian advection scheme, a velocity vector at each grid point is then derived by bilinear interpolation using the field of (25 $\times$ 25) vectors as described by Germann and Zawadzki (2002).

MAPLE does not incorporate a dynamic process for precipitation, which is a fundamental limitation of its prediction. Germann *et al*. (2006) discusses the sources of uncertainty in the MAPLE prediction associated with the growth or dissipation of precipitation echoes as well as with changes of the echo motion field over the forecast time. Turner *et al*. (2004) uses a wavelet transform method for filtering the non-predictable scales of precipitation in order to reduce forecast errors.

In this study, we evaluate the MAPLE prediction without applying filtering. Even though the total forecast time is up to 6 hours, the detailed analysis is performed for only up to 3 hours. The 6-h forecasts were generated every 10 minutes with a 10-min time step. The verification period covers the summer season of 2008, (June to August inclusively). A total of 852 forecasts was used for verification.

3 METHODOLOGY

Ebert and McBride (2000) propose a CRA technique that is based on the concept of contiguous rain areas defined as a region bounded by a specified isopleth in the union of the observed and forecast rain field. Predicted and observed radar composites for CRA analysis are smoothed in order to reduce the negative impact of noisy radar reflectivity measurements at the small scale. The amount of smoothing is simply achieved by averaging the 1 km resolution data over a (5 $\times$ 5) pixel neighborhood.

The main advantage of applying this technique is to decompose the total error into its components due to location (or displacement), volume (or intensity), and pattern structure. The total mean squared error $\mathrm{MSE}_{\mathrm{total}}$ [$\mathrm{mm}^2\mathrm{h}^{-2}$] can thus be written as the sum of each error component as follows.

$$\mathrm{MSE}_{\mathrm{total}} = \mathrm{MSE}_{\mathrm{displacement}} + \mathrm{MSE}_{\mathrm{volume}} + \mathrm{MSE}_{\mathrm{pattern}}$$

The total MSE is defined as

$$\text{MSE}_{\text{total}} = \frac{1}{N}\sum_{i=1}^{N}(f_i - o_i)^2$$

where f_i and o_i are the forecast and observed rainfall at grinpoint i, repectively. N is the number of gridpoints in the verification domain. The verification domain in the CRA technique is extended in order to include the shifted forecast field in the initial area bounded by the specified isopleth. The displacement error is determined by shifting the forecast field with respect to the observed field in the north-south and west-east directions until the cross-correlation coefficient is maximized or until the total squared error is minimized.

The $\text{MSE}_{\text{displacement}}$ is then obtained by subtracting the mean squared error between the forecast and observed field after this optimum shift (and denoted by $\text{MSE}_{\text{shift}}$) from the total mean squared error computed prior the shift at its original forecast position, that is,

$$\text{MSE}_{\text{displacement}} = \text{MSE}_{\text{total}} - \text{MSE}_{\text{shift}}$$

The mean squared error after the optimum shift is recalculated as

$$\text{MSE}_{\text{shift}} = \frac{1}{N}\sum_{i=1}^{N}(f_i^{'} - o_i)^2$$

where $f_i^{'}$ is the shifted forecast at gridpoint i.

The volume or bias error component is obtained as the square of the difference between the mean forecast $\boldsymbol{F}$ and observed $\boldsymbol{O}$ rain rate of the CRA entity after the shift.

$$\text{MSE}_{\text{volume}} = (\boldsymbol{F} - \boldsymbol{O})^2$$

The pattern error component is the subtraction of the volume error component from the mean square error obtained after shifting.

$$\text{MSE}_{\text{pattern}} = \text{MSE}_{\text{shift}} - \text{MSE}_{\text{volume}}$$

The threshold of 1.0 mmh^{-1} is used in identifying entity and error analysis.

4 RESULTS

The traditional statistical method is applied for the verification of rainfall rates $>$ 1.0 mmh^{-1} (Fig.1). The threat score (TS) for 30 min forecasts is about 0.55 with an e-folding time that is greater than 90 minutes. The threat score is computed as (Wilks 2006)

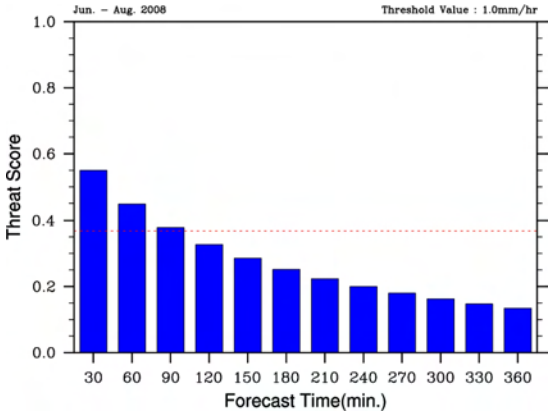

FIG. 1. Threat score for the MAPLE system verified with a 1 mmh^{-1} threshold over the 2008 summer season. The red dotted line denotes the 1/*e* level.

$$\text{TS} = \frac{a}{a+b+c}$$

where *a*, *b* and *c* are respectively the number of *hits*, *misses* and *false alarms* defined by a selected threshold. The threat score exceeds 0.3 for up to a 2-h forecast.

4.1 *Displacement Analysis*

The location errors of the MAPLE forecasts are calculated through the CRA analysis. Fig. 2 shows the distributions and histograms of the CRA displacement errors for forecasts of 30 minutes up to 3 hours. As expected, their distribution becomes wider with forecast time, indicating the larger discrepancy between the forecast and actual positions of the rain echoes. Nonetheless, most errors are restricted within a 20-km range up to a 3-h forecast. These results show that the MAPLE system adequately predicts the location of rainfall echoes. However, as the forecast time increases, it becomes evident that the distribution of the displacement error is skewed towards the southwest.

This feature is well illustrated in Fig. 3 that shows the 95% elliptic confidence region of the displacements for the each forecast time. The major axes and the centers of the ellipses are tilted and displaced toward the southwest, but the latter remain within a 10 km distance for up to a 3-h forecast. This bias indicates that the ob-

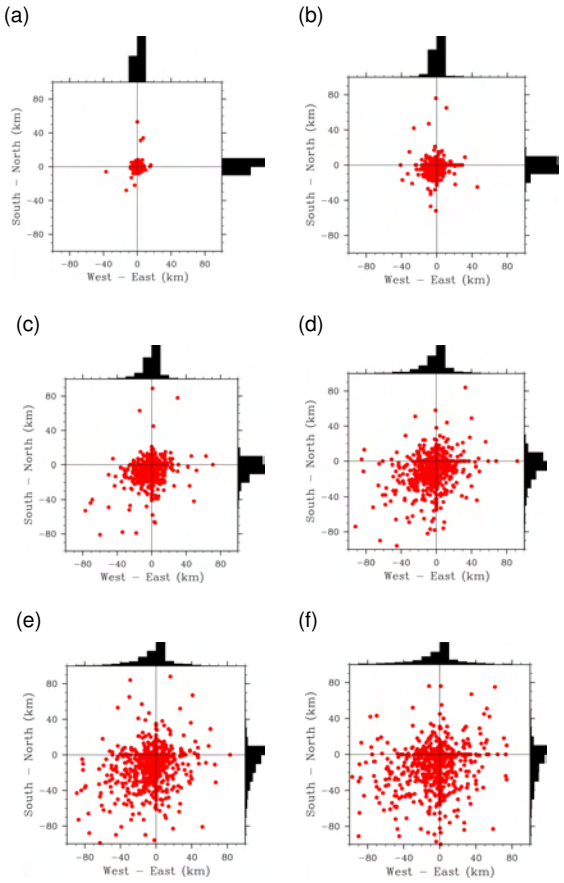

FIG. 2. Distribution and histogram of CRA displacements for each forecast time. (a) 30 min., (b) 60 min., (c) 90 min., (d) 120 min., (e) 150 min., and (f) 180 min.

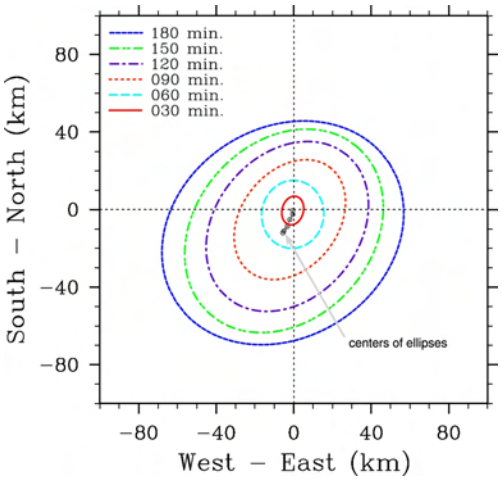

FIG. 3. 95% elliptic confidence region of the CRA displacements for the each forecast time.

served echoes are usually southwest of their location predicted by MAPLE. However, this feature is not considered to be a systematic bias caused by the MAPLE system and is regarded as negligible in relative magnitude.

4.2 *Rainfall Area and Volume*

The analysis of rainfall area and volume is performed not only in order to evaluate the MAPLE system but also in order to understand the behavior of the precipitation patterns in the 2008 summer season. Figure 4 and 5 show the difference between the forecast and observed rain area and volume respectively. The length of the bars indicates the number of cases (frequency) that the given area or volume fall in the category of % difference shown along the x-axis. The percentages are calculated from the difference between the forecast ***F*** and observed ***O*** average values expressed as

$$\%\ \text{difference} = (\boldsymbol{F} - \boldsymbol{O})/\boldsymbol{O} \times 100$$

Both rain areas and rain volumes are stratified into five size categories of the observed quantity. The size of categories are in terms of the observed quantity.

The percentage difference in rain area and volume are generally unbiased in all rainfall cases

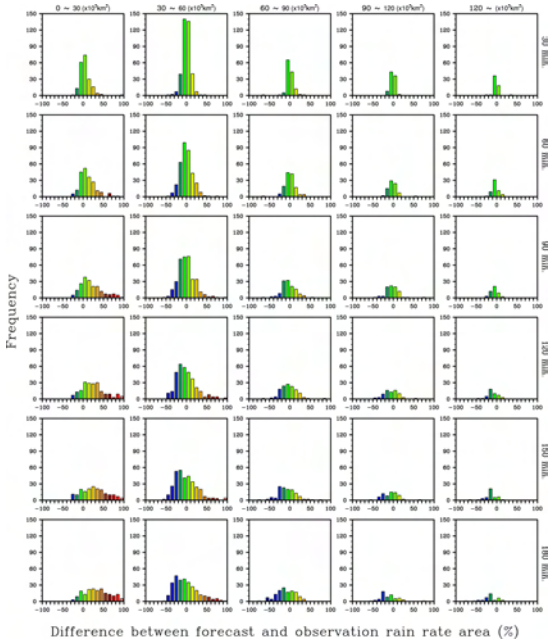

FIG. 4. Frequency of the percentage difference between the forecast and observed area of CRAs.

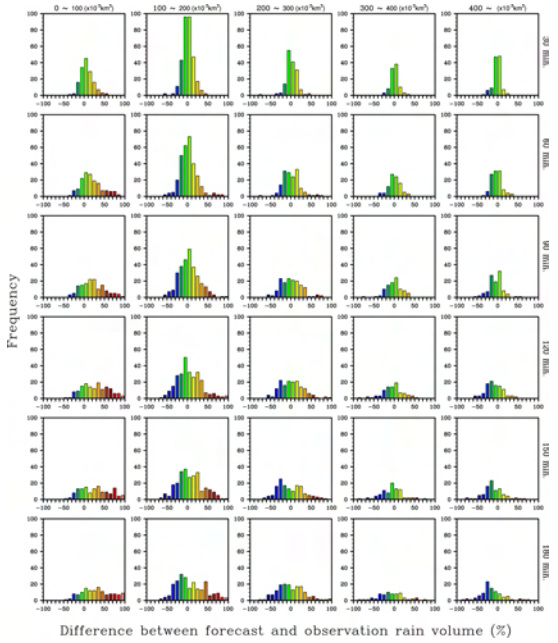

FIG. 5. Frequency of the percentage difference between the forecast and observed volume of CRAs.

for forecasts up to one hour, but as the latter increases, a bias develops that depends on the size category. There is a larger probability of positive differences, that is, overestimation, with lead time when the area or volume is small. This simply means that the forecast areas or volumes were more likely to be larger at the time of forecast. Conversely, for the observed larger size categories, an underestimation is more likely with the longer forecast times, implying that the echo pattern was more likely to be weaker at the moment of forecast. This does not imply that large areas or volumes become larger with time because the stratification has been done in terms of the observed, not of the forecast quantity. In general terms, we can state that observed CRA feature at both ends of the spectrum at verification time are more likely to have originated from the middle part of the spectrum.

4.3 *Error Analysis*

The most important characteristic of a CRA analysis is its ability to decompose the total forecast error into three types of components. The total mean square error (MSE) for rain rate forecasts of the MAPLE system ranges between about 19 mm^2h^{-2} and 35 mm^2h^{-2} for up to 3

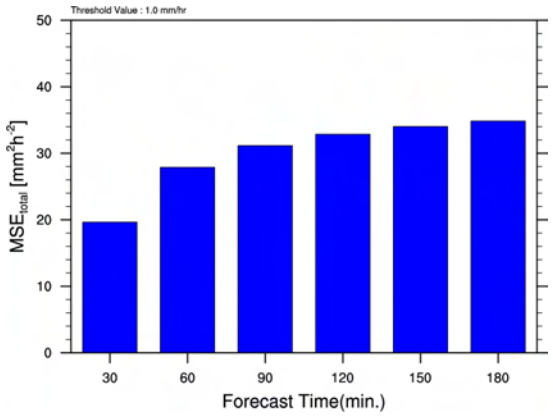

FIG. 6. Mean squared error of the MAPLE system for the 2008 summer season (June to August inclusively). The threshold used for the CRA definition is 1.0 mmh^{-1}.

hours as shown in Fig. 6. The MSE_{total} rapidly increase for the first half hour. The rapid increase in the first 30 minutes can be attributed to the inability to accurately forecast the intensity and location of echo features at the smaller scale. A noticeable increase is maintained until 90 minutes and then the rate of increase diminishes for the longer forecast times.

Figures 7 and 8 show the distribution and ratios of the error components of the total error. In general, each component has narrow distributions with the 25^{th} and 75^{th} percentiles of less than 10% from the median and a few extremes indicated by the maximum and minimum percentage values of the error (Fig. 7). The relative contribution of the pattern error to the total error shows a decrease with forecast time while that of the displacement error shows a slow increase. The volume error has a small magnitude but shows an increase similar to that of the displacement error.

The mean contribution of each component to the total error is shown in Fig. 8. The pattern error contributes nearly 90% of the total error while the maximum value of the displacement error barely reaches 7% for a 3-h forecast. The remaining 3% is attributed to the volume error. The relatively small contribution of the displacement error indicates that the MAPLE prediction of the location of precipitation echoes is reliable. Most of the total error is caused by the pattern error implying that the main source of forecast errors is caused by the change in the rainfall pattern during the forecast period not in the position or volume of precipitation echoes.

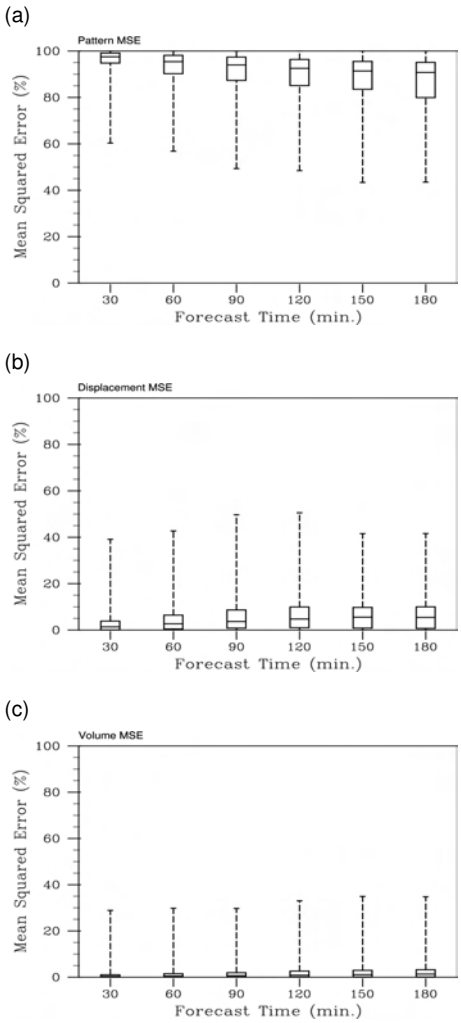

FIG. 7. The (a) pattern, (b) location and (c) volume error distribution over the forecast time as a % of the total error. The box represents the interquartile range, from the 25th to the 75th percentile, and the line through this box represents the median. The dotted line extends from the 25th and 75th percentiles to the outermost minimum and maximum values of the sample.

4.4 *Contingency Analysis*

The contingency table shown in Table 1 is used to evaluate the forecasts of the MAPLE system in terms of ′events′. The displacement error and the forecast mean rain rate are the two variables considered in the evaluation of the skill in predicting the location and quantity of precipitation. The threshold for the displacement error is chosen to be 20 km, a value that is also used by Ebert *et al.* (2004). The selection for the amount of rainfall

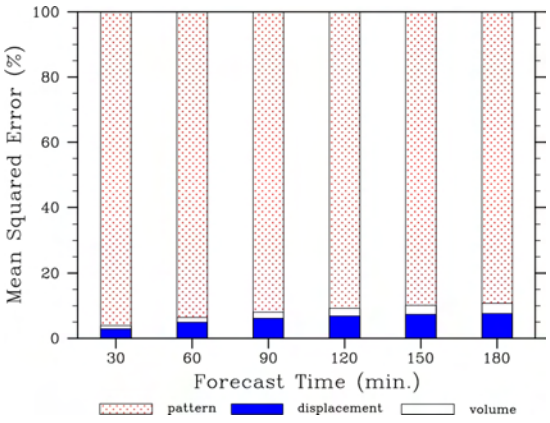

FIG. 8. The distribution of the pattern, location and volume errors as % of the total error.

depends on whether the differences between the forecast ***F*** and observed ***O*** mean rain rates are within acceptable limits so as to be qualified as a good quantitative forecast. We have required that the difference in mean rain rate (***F***-***O***) be within 10% of the observed value. An event is designated a ′hit′ (HT) if both errors are within the thresholds established above. An ′underestimate′ (UE) is categorized when the displacement error is within 20 km but the forecast mean rate difference $<$ -10%. Conversely an ′overestimate′ (OE) is recorded when the displacement is within 20 km but the difference exceeds +10%. If the forecast mean rain rate is within $\pm$10% but the displacement from observation exceeds 20 km, then the event is designated a ′missed location′ (ML). A ′missed event′ (ME) or ′false alarm′ (FA) is tabulated when, in addition to an erroneous location, the underestimate or overestimate is considerable, that is, each exceeds the selected threshold.

Figure 9 shows the result of this comparison over a 3-h period for CRAs with a mean rain rate threshold of 1 mmh^{-1}. The total number of the events for each forecast time is shown in the upper part of the figure. The frequency of the six categories just defined is plotted as a percentage of the total number of events for each forecast time. The frequencies of FA, ML and ME associated with a large displacement error show a gradual linear increase only after the first hour. This behavior reflects similar results obtained from the error decomposition analysis (Figs. 7 and 8). Both OE and UE do not change significantly with forecast time. Even though HT shows a sharp decrease from 80% at 30 minutes to just over 20%

TABLE 1. Contingency table used for event verification according to Ebert and McBride (2000). ***F*** and ***O*** indicate the forecast and observed mean rain rate, respectively.

		Difference between ***F*** and ***O***		
		Less than -10%	Within 10%	More than 10%
Displacement	$\leq$ 20km	Underestimate	Hit	Overestimate
	$>$ 20km	Missed Event	Missed Location	False Alarm

at 2.5 hours, it remains the most probable category throughout this period. The crossover of HT and ME or FA occurs around 2.5 hours. We can thus advise our forecasters that the MAPLE system provides reliable rainfall prediction up to 2.5 hours in terms of both the location and quantity of precipitation.

Figure 10 shows a similar analysis for several discrete rainfall thresholds from 1 up to 8 mmh^{-1}. The number of cases for each threshold is again provided on the top part of each plot for all the six forecast times. This figure reveals that HT remains the main category for most rates $<$ 6 mmh^{-1} for up to a 2-h forecast. The frequencies of FA and of OE decrease towards zero for higher rates simply because it is rare to have predicted rates that were still higher than the already high rates observed. The UE frequencies increase for rates $>$ 5 mmh^{-1} for all forecast times underline the common occurrence that weaker rates were instead predicted when such moderate rates were observed. The ME frequency for higher rain rates is particularly prominent in the 3-h forecast, emphasizing the tendency for any nowcasting system at longer forecasting intervals to both underestimate the occurrence of observed higher rainfalls and to misplace their location. This characteristic of MAPLE forecasts, that is, underestimation with strong precipitation, is consistent with the results of rain area and volume analysis shown in Figs. 4 and 5.

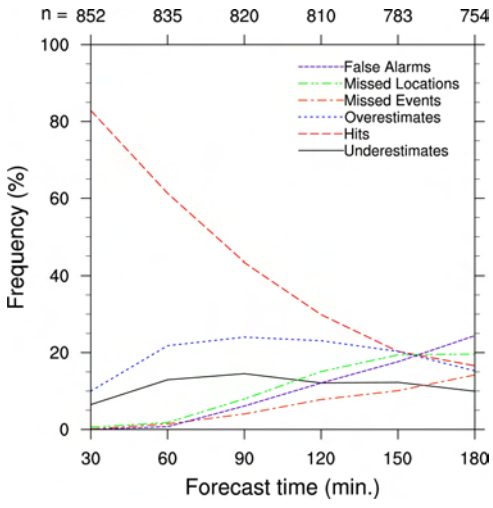

FIG. 9. The event verification for CRAs with rates $\geq$ 1 mmh^{-1}. The numbers on top of the plot indicate the number of events for each forecast time.

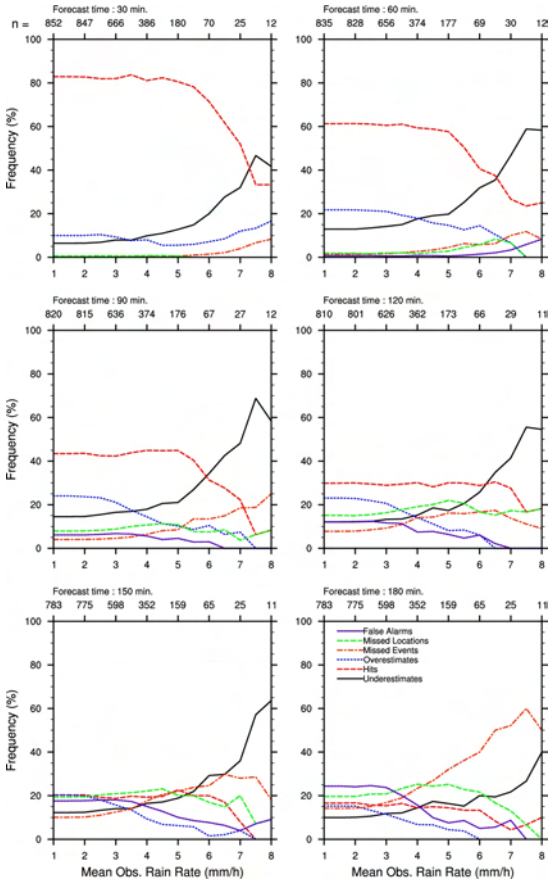

FIG. 10. The event verification as a function of the forecast time and of the observed rain rate threshold. The number of events for each interval of rain rate and of forecast time is provided on top of each plot.

5 DISCUSSION

The MAPLE nowcasting system, originally developed by McGill University for the continental United States, has been adapted for the smaller Korean radar network in 2007. Its evaluation has been performed over the 2008 summer season using the object-oriented Continuous Rain Area (CRA) technique after applying a (5×5) smoother to the original 1-km resolution data. Being an advection model according to the same optimally derived velocity field throughout the entire period of the forecast, MAPLE is not meant to predict the growth or dissipation of precipitation echoes. However, in spite of this limitation, the threat score, one of the traditional stochastic skill scores, remains above 0.3 for up to 2 hours (Fig. 1). The evaluation of the quantitative predictability of MAPLE has been illustrated by means of an analysis of rain area and volume (bias). The differences between forecast and observation become biased as the forecast time increases with MAPLE overestimating CRAs with small area or volume and underestimating those with larger area and volume (Figs. 4 and 5). However, this simply implies that small areas (volumes) observed at verification time were more likely to be larger at the moment of forecast, and observed larger areas (volumes) were more likely to be smaller at the moment of prediction.

While the general stochastic verification methods estimate only the total error, the CRA technique decomposes the total error into its three error component defined as the pattern, displacement and volume (or bias) error. The pattern error due to the fine scale structure of the precipitation patterns accounts for 90% or more of the total error throughout the 3-h forecast period, (Fig. 8). The error in the predicted location increases with forecast time but is mainly distributed within 20 km from the forecasted location up to 3 hours (Figs. 2 and 3), and remains within about 7% of the total error throughout the forecast period, (Fig. 8). The volume (or bias) error is the smallest of the three contributors at less than 3%. The event verification indicates that MAPLE maintains reliable forecasts up to 2.5 hours in terms of both location and amount of rainfall (Figs. 9 and 10). We suggest that this threshold may be considered by forecasters as a guideline for the upper limit of MAPLE forecasting skill in South Korea.

Acknowledgements

The authors would like to thank Dr. Ebert of Bureau of Meteorology Research Center for her help in programing and understanding the concept of CRA. This research was carried out as a part of research project 'NIMR-2009-C-1' supported by the National Institute of Meteorological Research (NIMR) in Korea.

References

Casati, B., L. J. Wilson, D. B. Stephenson, P. Nurmi, A. Ghelli, M. Pocernich, U. Damrath, E. E. Ebert, B. G. Brown, and S. Mason, 2008: Forecast verification: current status and future directions. *Meteorol. Appl.*, **15**, 3–18.

Ebert, E. E., and J. L. McBride, 2000: Verification of precipitation in weather systems: determination of systematic errors. *J. Hydrol.*, **239**, 179–202.

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

——, L. J. Wilson, B. G. Brown, P. Nurmi, H. E. Brooks, J. Bally, and M. Jaeneke, 2004: Verification of nowcasts from the WWRP Sydney 2000 Forecast Demonstration Project. *Wea. Forecasting*, **19**, 73–96.

Germann, U., and I. Zawadzki, 2002: Scale-dependence of the predictability of precipitation from continental radar images. Part I: Description of the methodology. *Mon. Wea. Rev.*, **130**, 2859–2873.

——, ——, and B. Turner, 2006: Predictability of precipitation from continental radar images. Part IV: Limits to prediction. *J. Atmos. Sci.*, **63**, 2092–2108.

Grams, J. S., W. A. Gallus Jr., S. E. Koch, L. S. Wharton, A. Loughe, and E. E. Ebert, 2006: The use of a modified Ebert-McBride technique to evaluate mesoscale model QPF as a function of convective system morphology during IHOP 2002. *Wea. Forecasting*, **21**, 288–306.

Laroche, S., and I. Zawadzki, 1995: Retrievals of horizontal winds from single-doppler clear-air

data by methods of cross correlation and variational analysis. *J. Atmos. Oceanic Technol.*, **12**, 721–738.

Lee, H. C., A. Bellon, I. Zawadzki and A. Kilambi, 2009: McGill Algorithm for precipitation nowcasting by Lagrangian extrapolation (MAPLE) applied to the South Korean radar network. Part 1: Sensitivity studies of the variational echo tracking (VET) technique. Paper P 1.2 of this 34th Radar Conference.

Savvidou, K., S. C. Michaelides, A. Orphanou, P. Constantinides, J.-P. Schulz, U. Voigt, and M. Savvides, 2007: Verification of precipitation forecasts by the DWD limited area model LME over Cyprus. *Adv. Geosci.*, **10**, 133–138.

Tartaglione, N., S. Mariani, C. Accadia, S. Michaelides, and M. Casaioli, 2008: Objective verification of spatial precipitation forecasts. *Precipitation: Advances in Measurement, Estimation and Prediction*, S. Michaelides, Eds., Springer, 453–472.

Turner, B. J., I. Zawadzki, and U. Germann, 2004: Predictability of precipitation from continental radar images. Part III: operational nowcasting implementation (MAPLE). *J. Appl. Meteor.*, **43**, 231–248.

Wilks, D. S., 2006: *Statistical methods in the atmospheric sciences*. Academic Press, 627pp.