CORRIGENDUM

ELIZABETH E. EBERT

Bureau of Meteorology Research Centre, Melbourne, Australia

1. Introduction

A recent study by Ebert (2001) described a ''poor man's'' ensemble of quantitative precipitation forecasts (QPFs) from seven operational numerical weather prediction (NWP) models. This ensemble was shown to provide useful probabilistic and deterministic forecasts of rainfall in the short range, 1–2 days. Unfortunately, an inadvertent coding error led to incorrect probability of precipitation (POP) forecasts for the second day of the forecast period. As a result, one of the conclusions of the Ebert (2001) study was that ''probabilistic skill was limited to lower rain rates during the second 24 h.'' At first glance this appeared to be a plausible result because independent model QPFs diverge with time much more quickly than do fields of 500-hPa geopotential height, which has been examined in earlier experiments with poor man's ensembles (e.g., Atger 1999; Ziehmann 2000).

Upon discovery of the coding error, the 48-h POP forecasts were corrected and verified against the Australian daily rain gauge analysis as was done earlier. The new results are far more encouraging. Instead of deteriorating quickly with time, the 48-h POP forecasts were only slightly less skillful than the 24-h POP forecasts. They had Brier skill scores greater than zero (i.e., had useful predictive skill) for all rain rates out to and exceeding 50 mm day^{-1}, and outperformed the ECMWF Ensemble Prediction System (EPS) at day 2 as well as at day 1.

The purpose of this corrigendum is to correct the erroneous results published by Ebert (2001) concerning the 48-h probabilistic QPFs. The following text, figures, and table should be substituted for section 4 in her paper.

2. Probabilistic rain forecasts

Recent studies on using ensembles to predict the probability of precipitation used fairly large ensembles ($\geq$25 members) in order to try to represent the full range of likely scenarios and derive better estimates of rainfall probability (Du et al. 1997; Buizza et al. 1999a). In this section we examine whether our poor man's ensemble of seven members contains useful information for estimating the probability of rainfall.

The POP was estimated at each grid point as the proportion of model QPFs predicting rain at or above a given threshold. No attempt was made to give greater weight to models with greater skill or to calibrate the POP forecasts (e.g., Hamill and Colucci 1997). The rain threshold was varied from 1 to 50 mm day^{-1} to determine the ensemble's predictive ability for progressively higher rainfall.

Figure 3 shows a reliability diagram for the 24- and 48-h poor man's ensembles, for a rain threshold of 1 mm day^{-1}. The reliability of the two ensembles is nearly identical and shows a tendency to overforecast rain frequency. This is perhaps not surprising since the NWP models have a tendency to overestimate the area of rainfall greater than 1 mm day^{-1} (McBride and Ebert 2000).

The Brier skill score is plotted as a function of increasing rain threshold in Fig. 4. The BSS was computed using all forecasts pooled in space and time to ensure greater stability, particularly at high thresholds. In practice the daily averaged values of BSS are a few percent lower than the pooled values. For both 24- and 48-h forecasts the poor man's ensemble has predictive skill for daily rainfall up to and exceeding 50 mm day^{-1}. The Brier skill score is greatest when the rain threshold is lowest (all rain included) and decreases as progressively higher rain rates are isolated.

It is interesting to compare the probabilistic skill of the poor man's ensemble with that of an ensemble prediction system. Following the example of Mullen and Buizza (2001) probabilistic forecasts from the 51-member ECMWF EPS initialized at 1200 UTC were verified for the summer and winter seasons over Australia. Table 3 shows the Brier skill scores for both systems. During the warm season the poor man's 24-h ensemble achieved BSS values of 0.22 for a 1 mm day^{-1} rain threshold and 0.10 for rain exceeding 10 mm day^{-1}. The ECMWF 36-h ensemble did not perform as well, giving BSS values of 0.12 and -0.09 for 1 and 10 mm day^{-1} thresholds, respectively. These values are lower than those reported by Mullen and Buizza (2001) for the United States warm season. By the second day of the forecast

Corresponding author address: Elizabeth E. Ebert, Bureau of Meteorology Research Centre, GPO Box 1289K, Melbourne 3001, Australia.
E-mail: e.ebert@bom.gov.au

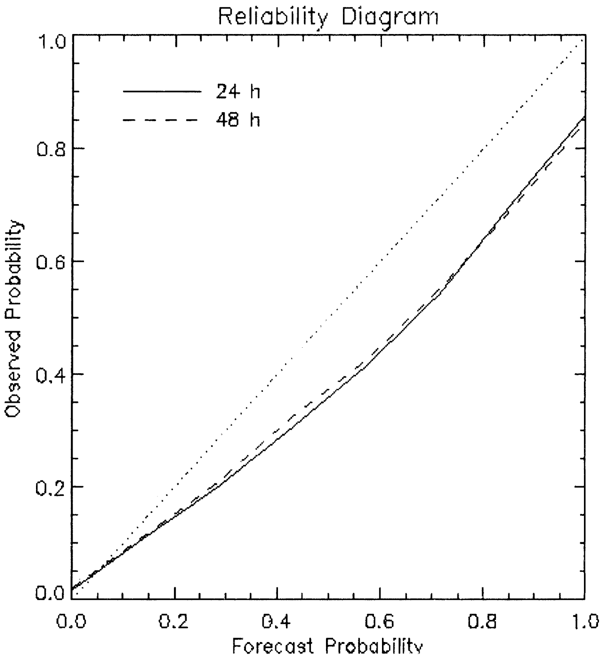

FIG. 3. Reliability diagram for POP forecasts from ensembles of 24-h QPFs (solid line) and 48-h QPFs (dashed line).

TABLE 3. Seasonal Brier skill scores for the poor man's ensemble and the ECMWF ensemble prediction system for (a) the warm season (Dec–Feb) and (b) the cool season (Jun–Aug).

	Poor man's ensemble		ECMWF EPS	
	24-h forecasts	48-h forecasts	36-h forecasts	60-h forecasts
(a) Warm season (Dec–Feb)				
Rain ≥ 1 mm day^{-1}	0.22	0.18	0.12	0.13
Rain ≥ 10 mm day^{-1}	0.10	0.08	−0.09	−0.06
(b) Cool season (Jun–Aug)				
Rain ≥ 1 mm day^{-1}	0.50	0.43	0.36	0.35
Rain ≥ 10 mm day^{-1}	0.44	0.33	0.27	0.22

the skill of the poor man's ensemble decreased slightly, while the scores for the ECMWF EPS were similar to its day 1 values. Cool season Brier skill scores over Australia were higher for both ensembles. For 1 and 10 mm day^{-1} thresholds the BSSs were 0.50 and 0.44 for the 24-h poor man's ensemble and 0.36 and 0.27 for the 36-h ECMWF EPS. Although we cannot verify 24-h forecasts from the ECMWF EPS, our results and also those of Mullen and Buizza (2001) suggest that they are usually slightly poorer than the 36- and 60-h forecasts due to insufficient spread of the ensemble. The poor man's ensemble of independent model QPFs appears to have greater probabilistic skill *for 1- and 2-day forecasts* than the larger ensemble of QPFs from the ECMWF EPS.

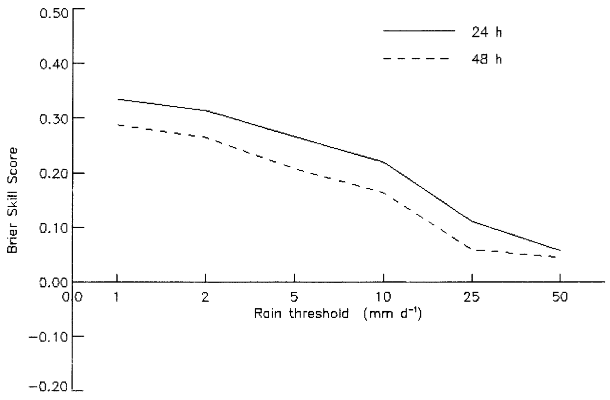

FIG. 4. Brier skill score as a function of rain threshold, for ensembles of 24-h QPFs (solid line) and 48-h QPFs (dashed line).

The mean daily ROC area is plotted in Fig. 5. To give an idea of the variability in skill among daily forecasts the vertical bars accompanying the 24-h ensemble values indicate ±1 standard deviation about the mean (daily standard deviations were similar for the 48-h ensemble). Using a value of 0.8 as a standard for a good probabilistic forecast, Fig. 5 shows that on a daily basis both the 24- and 48-h poor man's ensembles usually meet this standard for all but the heaviest rainfall occurrences.

3. Concluding remarks

Based on the earlier erroneous results, Ebert (2001) stated that, "Although we do not have ensembles for forecast periods beyond 48 h, we suspect that the usefulness of POP forecasts from the poor man's ensemble may be limited to the short term, 1–2 days." The verification of the corrected 48-h POP forecasts shown here suggests a different conclusion, namely that *POP forecasts from a poor man's ensemble may, in fact, continue to be useful beyond the 1–2-day range investigated here.*

The author sincerely apologizes for the misleading

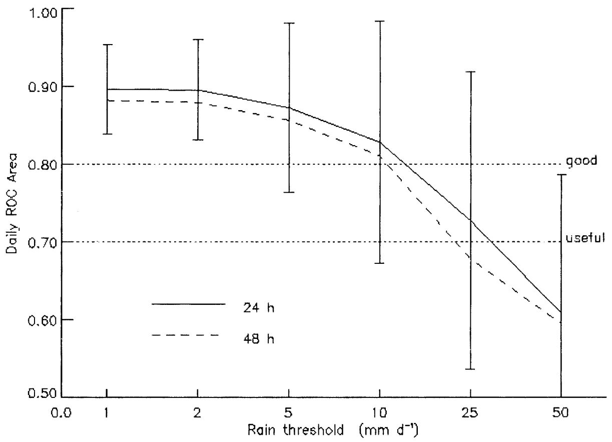

FIG. 5. Daily mean area under the ROC curve as a function of rain threshold, for ensembles of 24-h QPFs (solid line) and 48-h QPFs (dashed line). The vertical bars accompanying the 24-h ensemble values indicate ±1 standard deviation about the mean.

result published earlier and hopes that the correct (and more positive) performance of the poor man's ensemble at day 2 will encourage further research into and use of this approach.

REFERENCES

Atger, F., 1999: The skill of ensemble prediction systems. *Mon. Wea. Rev.,* **127,** 1941–1953.

Buizza, R., A. Hollingsworth, F. Lalaurette, and A. Ghelli, 1999a: Probabilistic predictions of precipitation using the ECMWF Ensemble Prediction System. *Wea. Forecasting,* **14,** 168–189.

Du, J., S. L. Mullen, and F. Sanders, 1997: Short-range ensemble forecasting of quantitative precipitation. *Mon. Wea. Rev.,* **125,** 2427–2459.

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.,* **129,** 2461–2480.

Hamill, T. M., and S. J. Colucci, 1997: Verification of eta–RSM short-range ensemble forecasts. *Mon. Wea. Rev.,* **125,** 1312–1327.

McBride, J. L., and E. E. Ebert, 2000: Verification of quantitative precipitation forecasts from operational numerical weather prediction models over Australia. *Wea. Forecasting,* **15,** 103–121.

Mullen, S. L., and R. Buizza, 2001: Quantitative precipitation forecasts over the United States by the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.,* **129,** 638–663.

Ziehmann, C., 2000: Comparison of a single-model EPS with a multi-model ensemble consisting of a few operational models. *Tellus,* **52A,** 280–299.