

PROBABILISTIC QUANTITATIVE PRECIPITATION FORECASTS BASED ON THE NCEP GLOBAL ENSEMBLE

Zoltan Toth and Yuejian Zhu

GSC at Environmental Modeling Center
National Centers for Environmental Prediction
NOAA/National Weather Service
Washington DC, USA

Summary: Through a number of examples we subjectively compare the performance of precipitation amount forecasts from a high resolution (T126) control prediction to that of probabilistic quantitative precipitation forecasts based on a lower (T62) resolution ensemble. It is important to note that the generation of the lower resolution ensemble requires approximately the same computer resources as running the higher resolution control. The results, supported by the evaluation of many other cases, indicate that the ensemble can extend the predictability of precipitation events by a day or two as compared to using only a high resolution control forecast. The improvement the ensemble offers is manifested in forecasts that (1) have useful information for longer lead times, (2) behave much more consistently in time, (3) have more accurate information about the spatial distribution of heavy rainfall, and (4) are associated with a flow dependent estimate of reliability.

1. INTRODUCTION

Due to the chaotic nature of the atmosphere, the quality of weather forecasts generally degrades with increasing lead time. The rate at which predictability is lost at different locations and at different times, however, varies and depends on the instability characteristics of the prevailing atmospheric flow. Knowledge about the day to day changes in the reliability of weather forecasts may be crucial for a wide range of users. A practical approach to quantifying predictability in a nonlinear forecast situation is to run an ensemble of forecasts from slightly perturbed initial conditions. At NCEP, the breeding method (Toth and Kalnay, 1993) is used to generate initial perturbations that are likely fast growing analysis errors. Including the control forecasts, a 17-member global ensemble is run operationally every day that provides forecast guidance out to 15 days ahead (Toth and Kalnay, 1997). Note that during the first 2–3 days, the forecast information can be enhanced by running a higher resolution experimental regional ensemble (Tracton and Du, 1998), embedded into the global ensemble forecasts, while beyond 15 days, the Climate Prediction Center of NCEP issues probabilistic guidance for temperature and precipitation.

2. ENSEMBLE GENERATED PQPF

Subjective (Toth et al., 1997) and objective evaluation of the global ensemble (Zhu et al., 1996) indicates that it can provide valuable forecast information otherwise not available through the traditional use of a single control forecast only. In particular, the frequency of different weather events within the ensemble of forecasts can be interpreted in probabilistic terms. We found that the ensemble generated probabilistic circulation forecasts can be easily calibrated, resulting in practically perfectly calibrated forecasts (i. e., forecast probabilities matching observed frequencies over the long run).

In this study, through a series of examples, we present a subjective evaluation of probabilistic quantitative precipitation forecasts (PQPFs) generated from the global ensemble over the US. The probabilistic forecasts are created by counting how many of the ensemble members exceed any given 24-hour accumulated precipitation amount, and then dividing that number by the total number (17) of ensemble forecasts. For example, when precipitation amount exceeds 1 inch in 11 out of a total of 17 members, the forecast probability for the 1 inch limit is 0.65 (11/17). This procedure is similar to that used by Petroliaigis et al. (1996) in their study of PQPF products generated by the ECMWF ensemble prediction system.

3. POTENTIAL PREDICTABILITY

Our first example is concerned with potential predictability. We assume that our numerical weather prediction (NWP) model is perfect and that errors are due only to the fact that the initial condition of the atmosphere is not known exactly. We use the operational ensemble to represent the uncertainty in the initial condition: for each member, except the controls, a small perturbation has been added to the analysis of the state of the atmosphere at initial time. Fig. 1 shows a series of control forecasts of 24-hour accumulated precipitation amount, with increasing lead time, for a precipitation event associated with the tail of a well predictable large scale cold front that passed over the Florida Peninsula. Comparing the 1-day lead time forecast, which is generally quite accurate and can be considered as a proxy for verification, to the longer lead time forecasts, we can see a rather familiar pattern: none of the forecasts

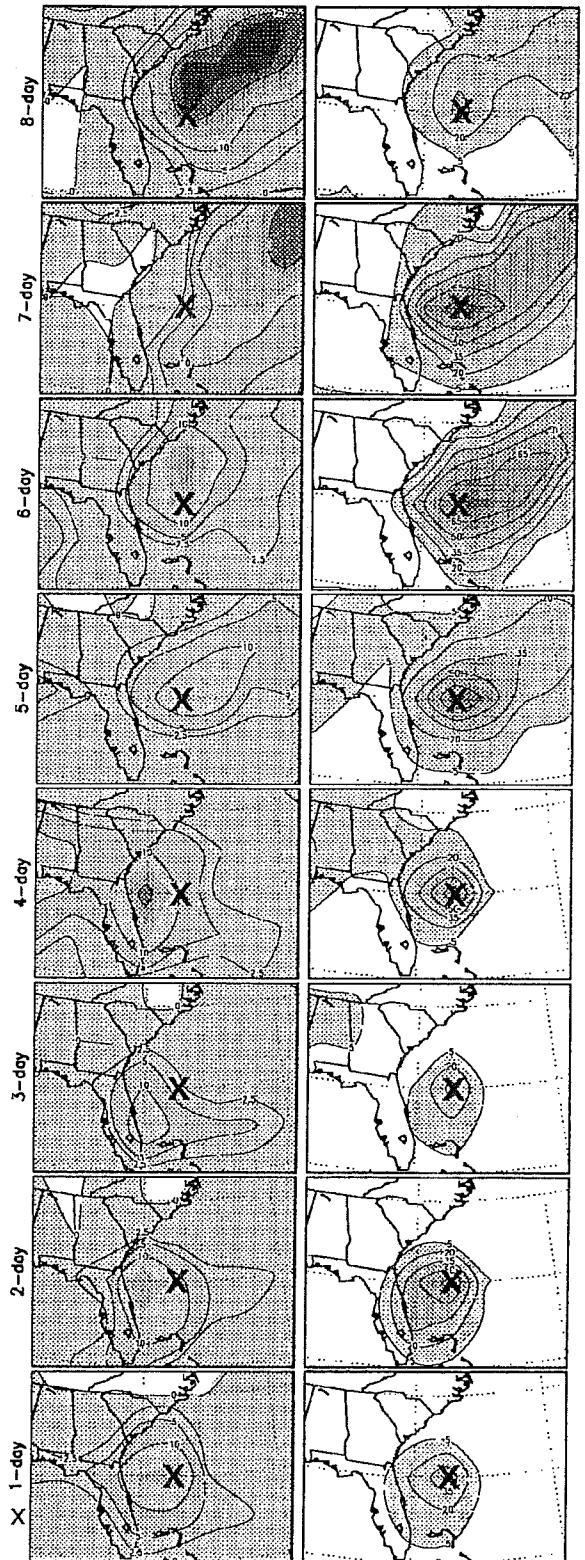


Fig. 1. 24-hr accumulated precipitation forecasts ending 1997 February 25 1200 UTC with 1–8 days lead time by the MRF model (from bottom to top, left column), and associated PQPF exceedance probabilities for the half inch limit from the global ensemble (right column). The location of the maximum in the MRF forecast at 1-day lead time is marked by a cross on all charts.

TOTH AND ZHU: PQPF BASED ON THE NCEP GLOBAL ENSEMBLE

(except at 5-day lead), gave the right location for the maximum amount of precipitation. The successive forecasts are different not because of model deficiencies (since all forecasts are made with the *same* model) but rather due to minor differences in their initial conditions.

On the other hand, if we look at the PQPFs that are based on 17 ensemble forecasts, we see a very clear signal: all 8 forecasts with different lead times give the maximum value of probability at the location where the precipitation had its maximum in the 1-day lead control forecast. This remarkable agreement, in contrast with the scatter of the location of maximum values in the control forecasts, highlights the potential advantage of running an ensemble of forecasts, instead of a single higher resolution control. Apparently, the ensemble is able to filter out much of the initial value related uncertainty. The ensemble, in terms of PQPFs, can better identify the most likely location and amount of precipitation events, whereas the control just picks one of the possible scenarios (which can often be quite different from the most likely one). We emphasize that the ensemble consists of mainly low resolution (T62) forecasts and therefore requires approximately the same computer resources as running the high resolution (T126) control.

4 . A TYPICAL EXAMPLE

Following the highly predictable case above, here we present a more typical example of ensemble-based PQPF. The precipitation event is again related to a well predictable cold front, now stretching across the eastern half of the continental US. The observed precipitation (estimated based on raingage and radar measurements, Baldwin and Mitchell, 1996) is shown in Fig. 2. At short (1-day) lead time the high resolution global control forecast could capture well the large scale precipitation event (Fig. 3). However, the regional scale details (for example, heavy precipitation off the Louisiana coast and southwest of Florida) are not captured at the T126 resolution.

The lower resolution ensemble (Fig. 3) provides very similar information to the control forecast at day 1. However, at longer (5-, and especially at 7-day) lead times the advantage of the ensemble approach becomes evident. The control forecast at 5-day lead time is very similar to the 1-day forecast and consequently verifies very well on the large scales. Such a good performance for the MRF forecast is not typical and a forecaster would not necessarily have strong confidence in a 5-day precipitation forecast. In this case, however, the ensemble PQPFs did not change much from 1-day to 5-day lead time either. All ensemble members still agree fairly well with the control, suggesting an *unusually high* confidence in the 5-day forecast.

At day 7 the control forecast for the southern part of the precipitation pattern is still rather skilful. However, the northern half of the heavy precipitation area is substantially displaced to the northeast, as compared to the verification. In contrast, the axis of highest probabilities from the ensemble is virtually unchanged, providing an excellent warning for heavy precipitation. Note also that the area of probabilities at day 7 (say, the 20% contour) is much enlarged from day 5, suggesting more uncertainty in the forecasts, in line with the poorer performance of the control forecast.

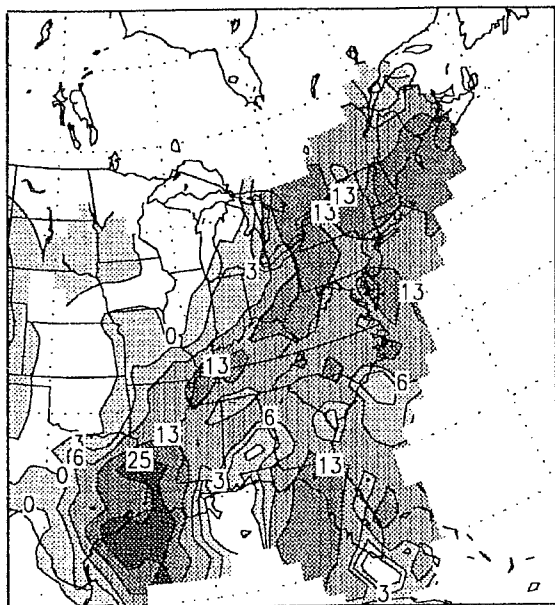


Fig. 2. Observed 24-hr accumulated precipitation amount for the period ending 1997 March 26 1200 UTC.

5. THE WINTER OF 1996/97 FLOOD IN CALIFORNIA

Series of heavy precipitation events along the west coast caused serious floods in different parts of California during the winter of 1996/97. Most of these precipitation events were predicted well by the NCEP global forecast system. Here we will focus on one event in late December – early January to highlight the advantages of using an ensemble as compared to relying on a single forecast only.

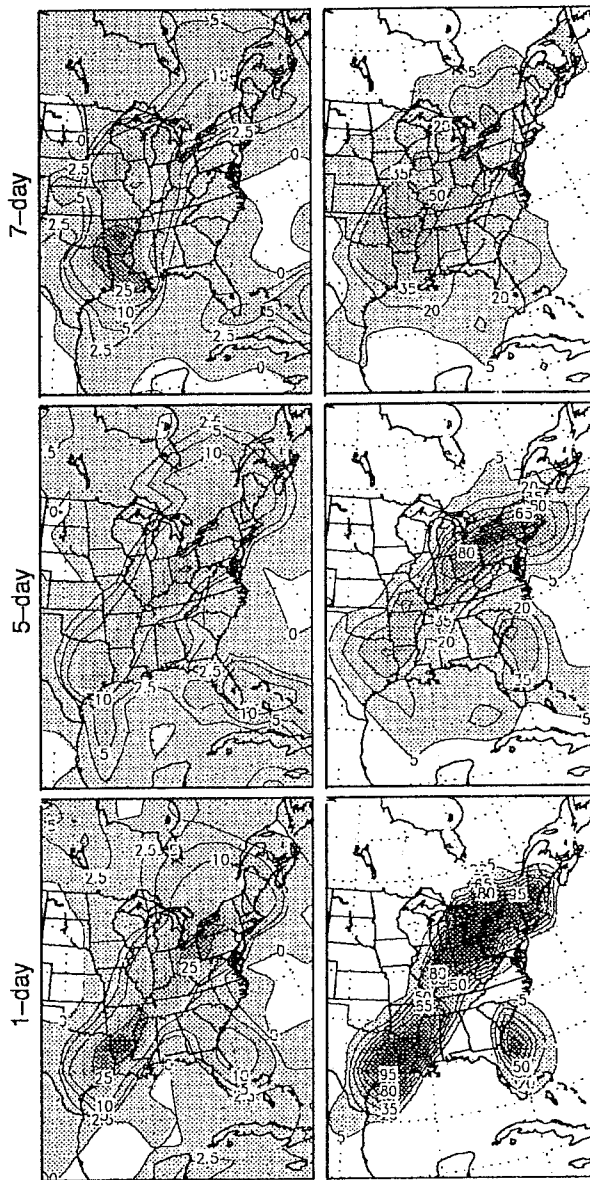


Fig. 3. 24-hr accumulated precipitation forecasts ending 1997 March 26 1200 UTC with 1-, 5-, and 7-day lead time by the MRF model (from bottom to top, left column), and associated PQPF exceedance probabilities for the half inch limit from the global ensemble (right column).

Fig. 4 shows the temporal evolution of observed precipitation in a northern California mesoscale size area between 24 December and 3 January, along with the 1-, 3-, and 5-day MRF control forecasts. The 1- and 3-day control forecasts verify very well whereas the 5-day forecasts missed much of the heavy precipitation that was observed between 30 December and 2 January. In contrast, the ensemble-based probabilities for precipitation above 1 inch (Fig. 5) are high throughout the whole event even at 5-day lead time. Equally important, the ensemble predicted very well around which date the heavy precipitation would start and end: notice that the probabilities are high only for 4–5 days over the period of largest accumulated observed precipitation, sharply rising before, and dropping after the event.

TOTH AND ZHU: PQPF BASED ON THE NCEP GLOBAL ENSEMBLE

Fig. 6 offers another way of comparing the behavior of the control and ensemble forecasts during this precipitation event. Shown are the control accumulated precipitation amount forecasts, along with the ensemble-based PQPFs with different lead times, all forecasts valid on 2 January for a small region of maximum observed precipitation. Again we can see that the ensemble gives consistently high probabilities for precipitation amounts exceeding 1 inch up to 6 days in advance, whereas the control forecast predicts large amounts consistently only from 3 days lead time.

The ensemble is not only able to extend the predictability of precipitation events in time but often also offers a more meaningful indication for the possible spatial distribution of precipitation. Fig. 7 shows the observed precipitation on 3 January. The control and ensemble forecasts at 1-day lead time (Fig. 8) capture well the main features of the precipitation pattern that day: heaviest precipitation along the (1) central California–Nevada border, with more than 10 mm rain extending to the area of (2) the Great Salt Lake, and another area of heavy precipitation in (3) western Washington state.

It is remarkable how similar the 8-day lead time ensemble-based PQPF is to the 1-day chart (and consequently to observations): The three main centers of precipitation activity still have 65%, 47%, and 40% probability of half inch or more rain associated with them. Note that the observed pattern is really only one realization among many others possible so it is not expected to match exactly the highest forecast probabilities. While the ensemble PQPF highlights the most likely distribution of heavy rain given the initial state and its uncertainty, the high resolution control forecast exemplifies only one of the possible scenarios. Not surprisingly, it has much less correspondence with the observed precipitation distribution than does the ensemble-based PQPF chart (cf. Fig. 7, top panels.)

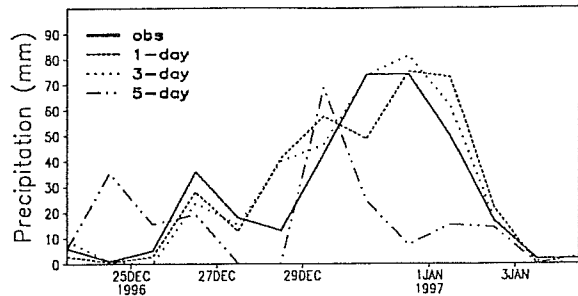


Fig. 4. 24-hr accumulated precipitation forecasts for 2.5 by 2.5 degree box centered at 40N, 122.5W by the MRF model with 1-, 3-, and 5-day lead times, along with observed amounts for periods ending at 12Z each day.

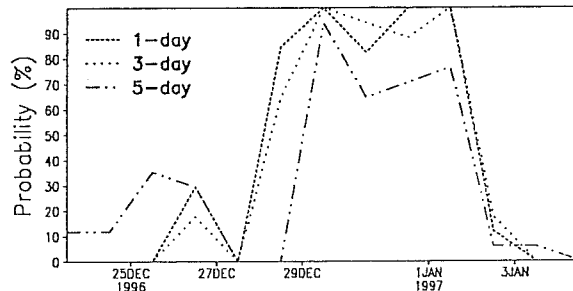


Fig. 5. Same as Fig. 4 except for ensemble-based PQPF exceedance probabilities for the one inch amount.

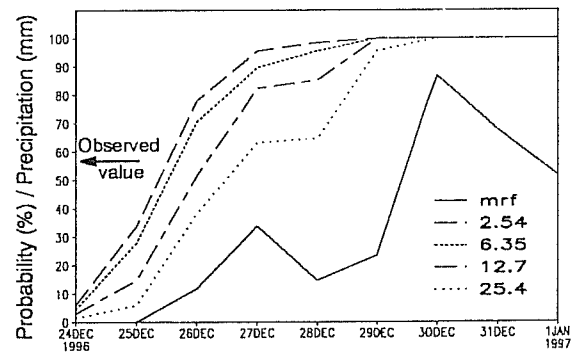


Fig. 6. 24-hr accumulated precipitation forecasts for 5 by 5 degree box centered at 38.75N, 121.75W by the MRF model with different initial dates, for the period ending 1997 January 2 1200 UTC, along with associated PQPF exceedance probabilities from the global ensemble for 0.1, 0.25, 0.5, and 1.0 inch thresholds.

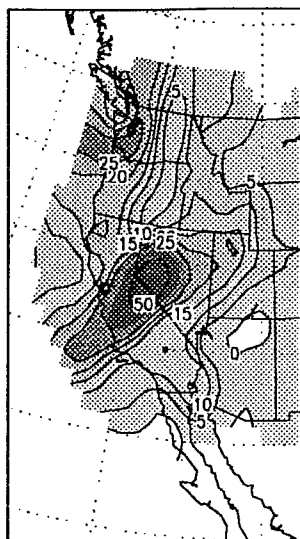


Fig. 7. Observed 24-hr accumulated precipitation amount for the period ending 1997 January 3 1200 UTC.

6. DISCUSSION

A subjective evaluation of probabilistic quantitative precipitation forecasts based on the NCEP global ensemble revealed that these products have useful forecast

information, often surpassing that contained in guidance from a single high resolution forecast (created at the same cost as the lower resolution ensemble). 1–15 days lead time PQPFs, based on the NCEP global ensemble, are regularly used at the Hydrometeorological Prediction Center and the Climate Prediction Center of NCEP, and at other places. The PQPF products have also been used in the daily operations of the FASTEX, NORPEX and CLAJET field experiments. A number of PQPF products can be found on the web at: <http://sgj62.wwb.noaa.gov:8080/ens/enhome.html>.

We should point out that the present study has several limitations. Most importantly, the PQPF guidance based on the raw ensemble forecast frequency information (section 2) is biased due to deficiencies both in model and ensemble formulations. In the future, we plan to fit a 3-parameter Gamma distribution (Lehman, 1997) to derive more accurate probability distributions based on the ensemble forecasts. The forecast probability distributions will then be compared to observed precipitation distributions from the past season. This will facilitate the calibration of the PQPFs, thus making them more reliable (i. e., make the forecast probabilities match the observed frequencies of different precipitation events over the long run).

Notwithstanding the advantages of the ensemble PQPFs demonstrated in this study, we note that the success of the ensemble approach hinges upon the quality of the NWP model forecasts. During the summer months, for example, when model performance is generally poorer, the ensemble has less to offer. This highlights the need for further model improvements. The improvements, however,

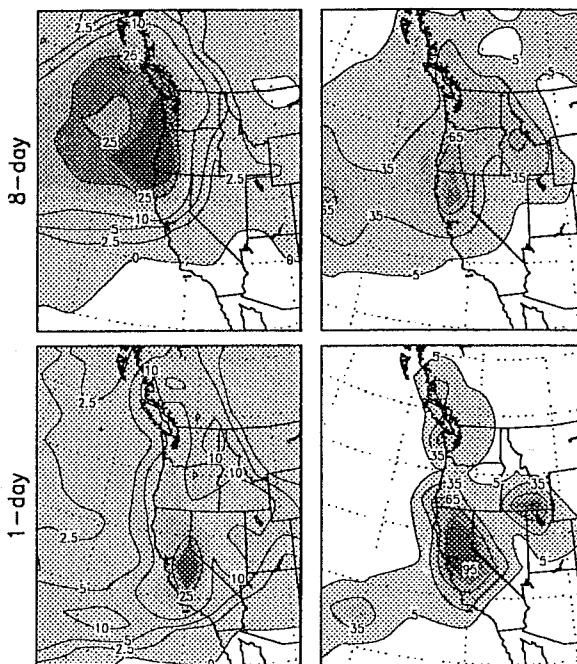


Fig. 8. 24-hr accumulated precipitation forecasts ending 1997 January 3 1200 UTC with 1- (bottom), and 8-day lead time by the MRF model (left column), and associated PQPF exceedance probabilities for the half inch limit from the global ensemble (right column).

TOOTH AND ZHU: PQPF BASED ON THE NCEP GLOBAL ENSEMBLE

should not be sought solely through increased model resolution, and especially not at the expense of running an ensemble: as we saw, using the same computer resources for running an ensemble, instead of a double resolution control, offers potentially much more benefit to the users. This conclusion is true not only for heavy precipitation guidance but also for circulation forecasts in general (Toth et al., 1998).

7. ACKNOWLEDGEMENTS

We are grateful to our colleagues, Eugenia Kalnay, Steven Tracton, Richard Wobus and Stephen Lord for their helpful comments and encouragement during the course of this work.

8. REFERENCES

Baldwin, M.E., and K. E. Mitchell: The NCEP hourly multi-sensor U.S. precipitation analysis. Preprints, 11th NWP Conf Norfolk, VA, 19–23 Aug. 1996, J95–96.

Lehman, R. L., 1997: Modeling precipitation data with Pearson Type III distributions: Overview of new interactive software "PFIT". Preprints of the 13th Conference on Hydrology, 2–7 February 1997, Long Beach, California, p. J150–J153.

Petroligis, T., R. Buizza, A. Lanzinger, and T. N. Palmer, 1996: Extreme rainfall prediction using the ECMWF ensemble prediction system. Research Department Technical Memorandum No. 227, p. 43.

Toth, Z., and Kalnay, E., 1993: Ensemble Forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteorol. Soc.*, **74**, 2317–2330.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319.

Toth, Z., E. Kalnay, S. Tracton, R. Wobus, and J. Irwin, 1997: A synoptic evaluation of the NCEP ensemble. *Weather and Forecasting*, **12**, 140–153.

Toth, Z., Y. Zhu, T. Marchok, S. Tracton, and E. Kalnay, 1998: Verification of the NCEP global ensemble forecasts. Preprints of the 12th Conference on Numerical Weather Prediction, 11–16 January 1998, Phoenix, Arizona, 286–289.

Tracton, M. Steven, and J. Du, 1998: Short-Range Ensemble Forecasting (SREF) at NCEP/EMC. Proceedings of the Fifth Workshop on Meteorological Operational Systems, 13–17 November 1997, in print.

Zhu, Y., G. Iyengar, Z. Toth, M. S. Tracton, and T. Marchok, 1996: Objective evaluation of the NCEP global ensemble forecasting system. Preprints of the 15th AMS Conference on Weather Analysis and Forecasting, 19–23 August 1996, Norfolk, Virginia, p. J79–J82.