U.S. DEPARTMENT OF COMMERCE NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION NATIONAL WEATHER SERVICE OFFICE OF SYSTEMS DEVELOPMENT TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 92-15

AN EXTRAPOLATIVE-STATISTICAL METHOD FOR FORECASTING RADAR REFLECTIVITY

David H. Kitzmiller and Jeffrey B. Ator

December 1992

AN EXTRAPOLATIVE-STATISTICAL METHOD FOR FORECASTING RADAR REFLECTIVITY

David H. Kitzmiller and Jeffrey B. Ator

1. INTRODUCTION

Convective rainstorms significantly affect many human activities. At the same time, such phenomena represent a particularly difficult forecasting problem because they develop and decay rapidly, and generally affect a small area for only a short period of time.

A number of numerical models of convective systems have been developed and demonstrated. These models simulate the development of individual thunderstorms or entire mesososcale convective systems. In an operational weather forecasting environment, however, neither the necessary observations nor computing power is usually available to run these models.

Yet it is possible to make use of radar observations to forecast future rainfall by extrapolation of the digitized reflectivity field. The movement of mature convective systems (the ones most likely to cause flash flooding and widespread severe weather outbreaks) is sufficiently conservative that linear extrapolation forecasts of their position are useful to at least 30 minutes, and possibly somewhat longer. The velocity of the system can be estimated from a sequence of earlier radar observations and its future position forecasted by assuming that the velocity will remain constant during the forecast projection period. The extrapolative forecast procedure is illustrated schematically in Fig. 1. The storm motion vector (SMV) between times t-30 and t₀ can be estimated objectively by calculating the displacement that yields the best pattern match between the two images. The forecasted precipitation area at time t_{30} (the hatched region in the figure) is made by displacing the t₀ image at this velocity.

This type of extrapolative forecast system has been implemented operationally by forecasting groups in a number of countries (Austin and Bellon, 1974; Conway, 1987; Takemura et al., 1987). At present, the Weather Surveillance Radar 1988 Doppler (WSR-88D) information processing system includes an echo centroid extrapolation product, though the product represents only forecasted positions for individual convective cells.

Saffle and Elvander (1981), hereafter referred to as SE81, demonstrated that it should be possible to make substantive improvements in extrapolative forecasts by treating future zero-tilt reflectivity (ZTR) within the storm system as a statistically-derived function of current reflectivity, echo top height (TOPS), and vertically-integrated liquid (VIL) estimates. In their approach, the basic extrapolative forecast steps (as shown in Fig. 1) were repeated on many sequences of archived radar data. The TOPS and VIL fields were extrapolated at the same velocity as the ZTR field. For each test forecast, grid-point values from the initial-time radar field, the field forecasted for valid time by extrapolation, and the valid-time radar field, were stored. An equation relating initial-time and extrapolated ZTR, TOPS, and VIL values to the valid-time observed ZTR was then determined by linear screening regression. In operations, this equation would be used to forecast the future ZTR values within the area to which the reflectivity field is extrapolated. SE81 found that the optimum predictor combination for future ZTR featured both ZTR and echo tops; it appeared that the deepest echoes were the ones most likely to feature high reflectivity after an interval of 30 minutes or more.

This study involves a repeat of SE81's original experiments in forecasting instantaneous ZTR values with a much larger sample of volumetric reflectivity observations than was then available. We have also incorporated a number of refinements to the methodology for estimating the storm motion vector, and expressed the reflectivity forecast in probabilistic terms. We will show that the extrapolative method yields useful skill to projections as great as 60 minutes.

2. DATA USED IN THIS STUDY

The radar data used in this study were collected at the WSR-57 site in Oklahoma City, Oklahoma (OKC), which has been equipped with Radar Data Processor II (RADAP-II) minicomputer equipment since the early 1980's. The RADAP II controls the radar during volumetric scanning observations, calculates and displays a variety of reflectivity-based radar products, and automatically archives data. In these experiments, we have used observations from the period 1985-1989. The cases are almost exclusively convective events, which are the ones most likely to cause flash flooding. All data were manually edited to remove anomalous propagation echoes.

For both trial time projections, individual sequences of radar images from 85 or more separate calendar days were used. Each image sequence consisted of five volumetric scans. The first four images were taken at 30, 20, 10, and 0 minutes before the initial time; these were used to estimate the motion vector. The fifth scan was taken at valid time, either 30 or 60 minutes after initial time. Our data sample features 816 sequences for the 30-minute forecasting experiment and 658 sequences for the 60-minute experiment. The digitized ZTR, VIL, and TOPS fields were objectively interpolated to a 4 x 4 km cartesian grid centered on the radar site. Data within 20 km of the radar were excluded to avoid the effects of ground clutter; data beyond 180 km were excluded because earth curvature effects can adversely affect the VIL and TOPS calculations at such range.

3. DERIVATION OF STORM MOTION VECTORS

A binary-correlation pattern-matching procedure was used to estimate storm motion vectors in this study. This method, described by SE81 and by Ciccione and Pircher (1984), is economical, and works well when the entire echo region does not change size or shape appreciably. Here, the "0-1" binary criterion was taken to be the 40-dBZ level in the ZTR field.

Initially, the SMV was estimated as the mean of the vectors calculated between three pairs of images, that is, from t_{-30} to t_{-20} , from t_{-20} to t_{-10} , and from t_{-10} to t_0 . The subscripts refer to time, in minutes, relative to the initial time t_0 . By employing image pairs only 10 minutes apart, we hoped to minimize the effects of changes in the echo patterns. However, it was found that this approach tended to underestimate the echo speed, due to discretization error when the echo movement was small relative to the grid spacing.

2

Better results were obtained when the SMV was estimated from the t_0 and t_{-30} images, or the two possible pairings at 20-minute separation. After further tests, it was decided to select the extrapolation SMV according to which of these three pairs of images had the highest binary correlation coefficient (that is, the pair having the closest match in terms of shape and size of the reflectivity region). Since the 30-minute movement vector should logically be least affected by discretization error, it was used in preference to the 20-minute vectors, unless one of the 20-minute image pairs had a binary correlation coefficient exceeding the 30-minute coefficient by at least 0.05.

In cases where none of the pairings had a binary correlation as high as 0.50, a motion vector from an earlier pair of images, as far back as 120 minutes from t_0 , was used. Finally, if no earlier pattern-match estimate of the SMV was available, the 700-hPa wind vector from the Nested Grid Model (NGM) forecast or analysis was used. Other tests, carried out in conjunction with this study, had shown that this wind vector was generally the best fully independent estimate of the echo motion that could be obtained from the NGM upper-air winds.

4. VALIDATION OF THE REFLECTIVITY EXTRAPOLATION TECHNIQUE

To verify that the extrapolation method outlined above yields significantly skillful forecasts of reflectivity, we prepared analyses in which both extrapolative and persistence forecasts were verified with observed data. The extrapolative forecasts were made by advancing an initial-time (t_0) field for 30- and 60-minute projections. The persistence forecast was simply the t_0 reflectivity field itself. We expected that the set of extrapolative forecasts would be the one more highly correlated to verifying observations.

The experimental forecasts utilized data from over 280 sequences of radar images. All points of the 4 x 4 km analysis grid within approximately 130 km of the radar were incorporated. The forecasts and verifying observations were reduced to categorical values according to whether the reflectivity reached, or did not reach, a threshold of 18 dBZ. For the 30-minute forecast experiment, over 3,818,000 individual grid-point cases were tested, while over 2,882,000 cases were available for the 60-minute experiment.

The terms for possible outcomes are shown in Fig. 2. In the figure, X indicates the number of grid points at which both forecasted and observed reflectivity exceeded 18 dBZ ("hits"), Y denotes the number of grid points at which the observed, but not the forecasted, reflectivity exceeded 18 dBZ ("misses"), and Z indicates cases in which forecasted but not observed reflectivity exceeded the threshold ("false alarms"). The final category, W, denotes cases in which neither forecasted nor observed reflectivity reached 18 dBZ.

The outcome values are often used in calculating forecast scores as outlined by Donaldson et al. (1975). The probability of detection (POD), is defined by X/(X+Y), the false alarm ratio by Z/(X+Z), and the critical success index (CSI) by X/(X+Y+Z).

The results of the 30-minute forecasting experiment, for both extrapolation and persistence forecasts, are shown in Fig. 3. Note that the number of cases in each outcome category is shown in thousands. The categorical forecast scores (Fig. 4) indicate that the extrapolative forecasts do possess more skill than the persistence forecasts. The POD is higher (0.70 compared to 0.66), while the FAR is lower (0.25 compared to 0.29). The CSI, which represents overall skill, was higher for the extrapolation forecasts (0.57 compared to 0.52). Within this sample, 32% of the verifying observations exceeded the 18 dBZ reflectivity threshold.

The differences in forecast scores were slightly larger for the 60-minute forecasts, as might be expected given the greater time projection. As shown in Fig. 5, the POD values for extrapolative and persistence forecasts were 0.55 and 0.50, respectively, the FAR values were 0.40 and 0.43, and the CSI values 0.42 and 0.36. Within this sample, only 21% of the verifying observations exceeded the reflectivity threshold. This percentage was smaller than in the 30-minute experiment; while all sequences were initiated at times when there was some convective activity within the verification region of the umbrella, a greater portion of the echoes exited the region during the longer projection interval.

Finally, the 30-minute forecasting experiment was repeated, but the reflectivity threshold was increased to 40 dBZ, a considerably higher value which typically constitutes a small fraction of the echo region, even in convective events. The extrapolation forecasts were clearly superior to persistence (Fig. 6), though the CSI values for the extrapolation and persistence forecasts were only 0.21 and 0.13, considerably lower than for forecasts for the 18 dBZ level. This is probably a consequence of the fact that the 40 dBZ region generally occupies a small fraction of the radar umbrella. Only 3.1% of the verifying values reached this reflectivity threshold. Also, the 40 dBZ region often represents the core region of convective updrafts, which develop and decay rapidly. The 18 dBZ area usually envelopes the strong updrafts, and changes size and shape more slowly.

5. CREATION OF THE STATISTICAL PREDICTOR-PREDICTAND DATASET

For each sequence of radar images, the extrapolation process outlined above was carried out, and the extrapolation forecast image was compared to the observed image at valid time. Grid-point values of ZTR, VIL, and TOPS, from both the extrapolated and initial time images, were stored in a separate dataset as candidate predictors. It is logical to expect that the initialtime data would contribute substantial information to forecasts of 20 minutes or less, which will be considered later in this study. The valid time ZTR values at the same grid points were stored as predictand data. Local averages and maxima of all values were also stored as new candidate predictors and alternate predictands.

This extrapolative-statistical method can be expected to yield information on reflectivity changes only within existing echo areas. Since the method cannot forecast changes in echo region shape or the formation of new echoes, data at other points in the valid time image do not contribute truly useful information to the regression procedure. Therefore, only ZTR, VIL, and TOPS values from grid points where the <u>extrapolated</u> ZTR was nonzero (the hatched area in Fig. 1) were stored and entered in the regression procedure. The values were drawn from every fifth grid point (20-km nominal spacing) in the north-south and east-west directions.

6. VALID-TIME REFLECTIVITY AS A FUNCTION OF EXTRAPOLATED RADAR PARAMETERS

The expected 30-minute projection valid-time ZTR (VALZTR) as a function of extrapolated ZTR (EXZTR) and TOPS (EXTOPS) is shown in Fig. 7. In this analysis, all extrapolated fields and VALZTR were averaged over 12 x 12 km square regions. Data from over 26,000 individual cases were included.

It is apparent that the extrapolated echo characteristics are significantly correlated to the VALZTR; the nonlinear correlation ratio (Panofsky and Brier, 1968) indicated that EXZTR (Fig. 7a) explained 22% of the variance in VALZTR. Within this sample, the mean of EXZTR was 35 dBZ, while the mean of VALZTR was 30 dBZ, with a standard deviation of 14 dBZ. It is understandable that VALZTR is less than EXZTR because of echo decay and errors in forecasting the system movement. The EXTOPS predictor (Fig. 7b) explained 23% of the variance. Though these are rather small reductions of variance, it should be noted that the verifying region is also small, only 144 $\rm km^2$.

The value of the extrapolation procedure is apparent when reductions in variance with respect to VALZTR are compared for various predictors (Fig. 8). The extrapolation predictors all explained about twice the percentage of the variance that the initial-time predictors did. Note that this comparison is somewhat different from the one described in the previous section; here, only grid points at which the extrapolation forecast indicated a nonzero reflectivity were considered.

Consistent results were obtained when this analysis was repeated for 60-minute extrapolation forecasts. In this data sample, the mean and standard deviation of VALZTR were 26.1 and 15.4, respectively. The corresponding reductions in variance (Fig. 9) were lower than for the 30-minute forecasts, as would be expected. Again, the extrapolation predictors explained substantially more of the variance in VALZTR than initial-time values did.

Forward screening linear regression was used to derive forecast equations relating VALZTR to the various predictors at 30- and 60-minute projections. An alternate predictand, ZTR averaged over a 20 x 20 km square region, was also tested. Generally, for both predictands, only two or three predictors contributed substantially to the reduction of variance, and these were usually EXTOPS and EXZTR.

Verification of these forecasts on independent radar data suggested that forecasts of the ZTR field itself might not be sufficiently accurate for operational use. The mean forecast error was nearly 10 dBZ, which corresponds to a rainrate error of factor four to five. This result suggested that the extrapolative forecast system might be more useful if it produced probabilistic, rather than continuous, forecasts.

7. PROBABILISTIC FORECASTS FOR REFLECTIVITY IN EXCESS OF A 40 DBZ THRESHOLD

The probabilistic approach involves defining the predictand as unity if VALZTR is greater than or equal to 40 dBZ, and zero otherwise. The screening regression process then yields a regression estimate of event probability (REEP) of reaching or exceeding the 40-dBZ threshold. Such forecasts should be useful in field operations, since they would provide the forecaster with the <u>potential</u> of rainfall in excess of a significant rate, rather than a simple expected value with no explicit statement of the level of certainty. The 40-dBZ level corresponds to a rainrate of approximately 0.5 inches per hour and is often used to delineate "convective" from "stratiform" rainfall.

The results of this approach are illustrated in Fig. 10, which shows the probability that the ZTR will exceed 40 dBZ within a 12 x 12 km square region, given various values of ZTR and VIL at initial time. The ZTR predictor is EXZTRMX, defined as the local maximum ZTR value within a 12 x 12 km region centered on the grid point of interest, and extrapolated forward by 30 minutes. The VIL predictor, EXVILMX, was similarly defined as the extrapolation of the local maximum in VIL.

As shown in Fig. 10a, if the initial local maximum ZTR was 22.5 dBZ, there was only a 5% probability that the same subregion within the moving echo area would have a reflectivity in excess of 40 dBZ 30 minutes later. The probability increased to over 50% if the initial echo was 46-48 dBZ, and to over 70% if the initial echo was 53 dBZ or greater. Similar results are shown in Fig. 10b; it is apparent that current VIL is also an effective predictor of future reflectivity. If the initial VIL exceeded 10 kg m⁻², there was at least a 60% probability that the echo region would feature ZTR in excess of 40 dBZ after 30 minutes. In this sample, 32% of the cases had valid-time ZTR above the threshold, and EXVILMX explained 25% of the predictand variance. The EXZTRMX predictor explained 24% of the variance.

The reduction of variance with respect to this binary predictand, for both extrapolated and initial-time predictors, appears in Fig. 11. As was the case for a continuous ZTR predictand, the extrapolated predictors were much more highly correlated with later high-reflectivity occurrence than were the initial-time fields. The extrapolated echo top predictor was no longer the best in terms of reduction of variance. It is possible that, while extrapolated TOPS was the predictor most highly correlated to future ZTR over the entire range of ZTR values, the extrapolated ZTR and VIL fields were more highly correlated to future ZTR within the range near the 40 dBZ value, which is of importance in defining the binary predictand.

The relationship between extrapolated VIL and VALZTR at the 60-minute projection is shown in Fig. 12. Here, the probability of 40 dBZ echo occurrence, within a 12 x 12 km region at the 60-minute projection, is shown as a function of EXVILMX. At this greater time projection, EXVILMX explains less variance in the predictand, as could be expected; the histogram indicates that this predictor cannot delineate probability values less than 12% or greater than 50%. The overall relative frequency of 40-dBZ echoes in this sample was only 25%; this is due to the tendency of echoes to move out of the verification region during the forecast period. The EXVILMX predictor explained 9% of the predictand variance.

8. VALIDATION OF 30- AND 60-MINUTE PROBABILISTIC REFLECTIVITY FORECASTS

To obtain a more comprehensive view of the possibilities of this potential forecasting system, we prepared and then verified probabilistic forecast equations for the two predictands illustrated above. For both the 30- and 60-minute projections, two separate equations were developed from subsets of the available data samples. Forecasts were then generated for independent samples and verified. The resulting equations for 30-minute ZTR are:

(1)P30 = -66.5 + (2.05 EXZTRMX) + (2.31 EXTOPMX)

based on data from 1985, 1987, and 1988, and

P30 = -50.3 + (1.98 EXZTRMX) + (1.97 EXVILAV)

based on data from 1985, 1988, and 1989. Here, P30 is the probability that a ZTR value of 40 dBZ or more will be observed within a 12 x 12 km region 30 minutes later. The predictors EXZTRMX and EXTOPMX are the maximum observed ZTR (dBZ) and TOPS (km AGL) values within a 12 x 12 km region extrapolated to valid time. EXVILAV is the mean VIL value within a 12 x 12 km region extrapolated to valid time.

Within their respective dependent datasets, (1) explained 21% of the predictand variance and (2) explained 25%. In both equations, most of the reduction of variance was contributed by the EXZTRMX term. Though our earlier analysis indicated that EXVILMX had the highest nonlinear correlation to future ZTR, EXZTRMX had a higher linear correlation. EXVILMX might have been selected by this screening procedure if it had been entered in a "linearized" form (see Reap and Foster, 1979, or Charba, 1977 for an explanation of the linearization process).

Forecasts for cases in calendar year 1989 were then prepared from (1) and for cases in 1987 from (2). The reliability of the probability forecasts within the independent data samples is illustrated in Fig. 13. Though there was some tendency to overforecast when forecasted values were greater than 70%, no strong bias appeared in the sample as a whole.

The probabilistic forecasts may be objectively reduced to categorical (yes/no) by applying a threshold probability. An analysis of the scores that might be achieved by such categorical forecasts appears in Fig. 14. As before, POD represents probability of detection, FAR the false alarm ratio, and CSI the critical success index or threat score. Each score is shown for a range of possible threshold probabilities, from 1 to 40%. The peak CSI is achieved at a threshold of 35%, at which the POD is 0.71 and the FAR is 0.42. The bias at this threshold (not shown) is 1.2. Thus, categorical forecasts from this system could detect approximately 70% of the 40 dBZ echoes, while the number of "yes" forecasts would exceed the number of "yes" observations by about 20%. To achieve a POD of 0.8, it would be necessary to apply a threshold probability of 29, and accept an FAR of 0.48 and bias of 1.5.

This process was repeated to obtain forecast equations for a 20 \pm 20 km region at a 60-minute projection. The resulting test equations are:

P60 = -32.1 + (1.40 EXZTRMX) + (2.04 EXTOPMX)

based on data from 1985, 1988, and 1989, and

P60 = -39.6 + (1.48 EXZTRMX) + (2.17 EXTOPMX)

based on data from 1985, 1987, and 1988. The definition of P60 is analogous to that of P30 in (1) and (2). The reduction of variance for these equations

(3)

(4)

(2)

was 0.10 for (3) and 0.13 for (4). Again, EXZTRMX contributed the most to the reduction of variance.

The results of reliability and verification tests appear in Figs. 15 and 16. Categorical forecasts scores (Fig. 16) were not radically different from those for the 30-minute forecasts, since a larger verification region was specified for these 60-minute forecasts (400 km² rather than 144 km²). The peak CSI is still achieved at a threshold of 35%, where the POD is 0.71, but the FAR and bias are larger (0.50 and 1.4, respectively).

9. DISCUSSION

This study confirms the results of SE81, who reported significant skill at forecasting future ZTR as a statistically-derived function of current ZTR and the volumetric reflectivity indices VIL and TOPS. Though deterioration of skill is evident for 60-minute forecasts, the forecasts at least as far as 30 minutes are clearly useful. It should be possible to improve the forecasts for projections beyond 30 minutes by the incorporation of new predictors involving environmental conditions and time rates of change of reflectivity characteristics. It might also be useful to state the probability that reflectivity will exceed some threshold during a future period of time, such 30 to 60 minutes.

We now intend to extend this extrapolative-statistical approach to quantitative precipitation forecasting. It is possible that the 30- and 60-minute rainfall accumulation fields are less volatile and more spatially continuous than the instantaneous ZTR field. Rainfall amount will be treated as a function of extrapolated ZTR, VIL, and TOPS over the duration of the forecast period. The next phase of this research effort is now underway.

10. ACKNOWLEDGEMENTS

We wish to thank Wayne E. McGovern and Robert E. Saffle for their suggestions on both the conduct of these experiments and this note. Karen Yip assisted in the preparation of the manuscript.

11. REFERENCES

- Austin G. L., and A. Bellon, 1974: The use of digital weather radar records for short-term precipitation forecasting. <u>Quart. J. Royal Meteror. Soc.</u>, 100, 658-664.
- Charba, J. P., 1977: Operational system for predicting severe local storms two to six hours in advance. <u>NOAA Technical Memorandum</u> NWS TDL-65, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 24 pp. [NTIS PB 271 147/1]
- Ciccione, M., and V. Pircher, 1984: Preliminary assessment of very short term forecasting of rain from single radar data. <u>Proceedings of the Second</u> <u>International Symposium on Nowcasting</u>, Norrkoping, European Space Agency, 241-246.
- Conway, B. J., 1987: FRONTIERS: An operational system for nowcasting
 precipitation. <u>Proceedings Symposium on Mesoscale Analysis and Forecast ing</u>, Vancouver, European Space Agency, 233-238.

- Donaldson, R. J., R. M. Dyer and M. J. Kraus, 1975: An objective evaluator of techniques for predicting severe weather events. <u>Preprints Ninth</u> <u>Conference on Severe Local Storms</u>, Norman, Amer. Meteor. Soc., 321-326.
- Panofsky, H. A., and G. W. Brier, 1968: <u>Some Applications of Statistics to</u> <u>Meteorology</u>. Pennsylvania State University Press, University Park, 224 pp.
- Reap, R. M., and D. S. Foster, 1979: Automated 12-36 hour probability forecasts of thunderstorms and severe local storms. <u>J. Appl. Meteor.</u>, 18, 1304-1315.
- Saffle, R. E., and R. C. Elvander, 1981: Use of digital radar in automated short range estimates of severe weather probability and radar reflectivity. <u>Preprints Seventh Conference on Probability and Statistics in</u> <u>Atmospheric Sciences</u>, Monterey, Amer. Meteor. Soc., 192-199.
- Takemura, Y., Y. Makihara, K. Takase, and K. Aonashi, 1987: Operational experiment in very short range forecasting of precipitation. <u>Proceedings</u> <u>Symposium on Mesoscale Analysis and Forecasting</u>, Vancouver, European Space Agency, 239-244.



Figure 1. Schematic diagram of extrapolative forecast process for radar fields. Cross-hatched region at right ("t₀ extrap") represents area over which precipitation is forecasted at t_{+30} .



Figure 2. Possible outcomes of categorical (yes/no) forecasts for radar reflectivity.



Figure 3. Number of grid blocks within each forecast/verification category, for 30-minute radar reflectivity forecasts. Numbers in parantheses are for persistence forecasts (non-extrapolated fields); numbers represent thousands of cases.



Figure 4. Scores for 30-minute extrapolation and persistence forecasts of reflectivity of 18 dBZ or greater.



Figure 5. As in Fig. 4, except for 60-minute forecasts.



Figure 6. As in Fig. 4, except for forecasts of 40 dBZ reflectivity.









Figure 8. Reduction in variance with respect to 12 x 12 km average ZTR, by various initial-time and extrapolated volumetric radar predictors. Forecasts are for 30 minutes.



Figure 9. As in Fig. 8, except for 60-minute forecasts.









Figure 11. Reduction of variance with respect to binary reflectivity predictand (ZTR>40 dBZ), by various initial-time and extrapolated predictors. Predictors and predictand are all based on local maxima within a 12 x 12 km region.



Figure 12. As in Fig. 10b, except for 60-minute forecasts.

16







Figure 16. As in Fig. 14, except for 60-minute forecasts valid within 20 x 20 km regions.