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

TDL OFFICE NOTE 96-1

ONE-HOUR FORECASTS OF RADAR-ESTIMATED RAINFALL BY AN EXTRAPOLATIVE-STATISTICAL METHOD

David H. Kitzmiller

January 1996

ONE-HOUR FORECASTS OF RADAR-ESTIMATED RAINFALL BY AN EXTRAPOLATIVE-STATISTICAL METHOD

David H. Kitzmiller

ABSTRACT

A common approach to short-range precipitation forecasting involves the extrapolation of gridded radar reflectivity fields. In general, the future rainrate at any given point is forecasted by assuming that the current rainrate pattern inferred by radar will move at some known velocity, and that the rainrates within the echo region remain constant in time or decrease at an assumed rate. The echo velocity may be estimated by pattern correlation between the most recent radar image and earlier ones, or it is assumed to be equal to the environmental wind at some level between 850 and 500 mb.

As a refinement of this basic technique, a large number of extrapolative rainfall amount forecasts were prepared and then statistically correlated with the observed rainfall, as estimated by radar, during the valid period. The relationships between the purely extrapolative forecasts and the observed amounts can then be used in interpreting other extrapolative forecasts prepared operationally. The resulting extrapolative-statistical approach to rainfall forecasting implicitly accounts for echo decay and uncertainty in the extrapolation process, and is a form of the Model Output Statistics technique often used to produce forecasts of sensible weather from numerical weather prediction model output. Both low-level reflectivity and vertically-integrated liquid (VIL) are used as predictors of rainfall amount. In practice, the extrapolative-statistical algorithm produces probabilities that the radar-estimated rainfall will exceed 0.1, 0.25, 0.5, and 1 inch, within each box of a 4-km grid, during the next hour. A categorical rainfall forecast can then be derived from the probabilities. The operational WSR-88D Z-R relationship was used to convert reflectivity to rainrates.

Because it incorporates input from a variety of statistical predictors, the extrapolative-statistical approach yields improvements over a purely extrapolative one, at the expense of a modest increase in computing time. This note shows that the use of multiple statistical predictors measureably improves forecasts of rainfall amounts of 0.5 inch and greater.

1. INTRODUCTION

Quantitative forecasts of rainfall, particularly for situations involving convective storms, represent a difficult forecast problem. It is now possible, in theory, to use time-dependent physical models to forecast the development and intensity of convective storms in real time. However, the interpretation of such model forecasts of precipitation can be difficult, because users often have little documentation on bias in the forecasts and what range of observed rainfall values can be expected for any forecasted value.

To produce short-term forecasts (less than 3 h) of many weather phenomena, stochastic or statistical models can be used. Though they may not incorporate explicit models of atmospheric physical processes, these models may be based on prognostic relationships derived empirically from a substantial number of historical cases. Statistical models are also generally inexpensive to run operationally. The rainfall forecasting model described below requires only radar reflectivity data, and can produce forecasts in less than 1 minute on platforms as small as a personal computer.

A particular class of statistical models, namely extrapolative models, have been used for operational forecasts of rainfall at a number of forecasting centers (Austin and Bellon 1974; Conway 1987; Takemura et al. 1987). These generally use radar reflectivity, sometimes augmented by satellite measurements or automated raingage networks, to specify a rainrate field. A forecast of rain accumulation at any one point is then derived by forecasting the movement of the rainrate pattern, under the assumption that the shape of the pattern is constant in time, or that it changes according to some known principle. An extrapolative rain forecast algorithm was developed for use in the WSR-88D processing system (Walton and Johnson 1986) and tested in real time on a similar system by Kelsch (1990), but it has not been implemented operationally at this time.

The rainfall forecast algorithm described here uses a statistical method to refine the basic extrapolative approach. First, extrapolative forecasts of reflectivity and rainfall fields were prepared from archived data for a large number rainfall events. These forecasts were then statistically correlated with observed rainfall fields. The correlation procedures provide guidance on interpreting the purely extrapolative forecasts. They show, for any given forecasted rain amount, the range of observed rainfall amounts that can reasonably be expected, and the probabilities that some fixed rainfall thresholds will be exceeded.

It is logical to expect that rainfall in the 0-1 h time frame can be forecasted by one or more predictors such as the current rainrate and VIL, and by forecasts of the rainrate and VIL. The future rainrate and VIL can be estimated by extrapolating the current fields, and assuming that the initial echo pattern's shape and rainrate remain fixed during the forecast period. We refer to the initial-time rainrate and VIL as persistence predictors, and to the rainfall and VIL forecasts as extrapolation predictors. Additional candidate predictors included spatially-smoothed rainfall and VIL forecasts, from which local noise has been removed. Also, forecasts valid during the 0-30 and 30-60 minute portions of the valid period were included. The 0-30 minute forecast is probably more accurate in an absolute sense than the full 60-minute forecast, and might contribute independent information.

We do not know a priori what combination of predictors has the most information with respect to rainfall. Both extrapolation and persistence predictors have certain advantages. The persistence quantities have high absolute accuracy, while the extrapolation predictors contribute information about the short-term movement of the precipitation system. The performance of the candidate predictors may be evaluated by correlating them with observed rainfall over many historical cases for which verification data are available. An objectively-derived optimum combination of the candidates can be selected by some process such as screening regression.

In this development effort, the forecast fields were defined on a 4-km local map grid of 230 km radius, centered on the radar site. The echo velocity during the forecast period was estimated by pattern matching between reflectivity maps at 10, 20, and 30 minutes prior to forecast initial time. For any one box within the grid, the forecasted rainfall amount was derived from the time-average rainrate over it, under the assumption that the echo pattern shape and rainrate remain constant, while the pattern itself moves at a fixed velocity. The rainfall forecasts were then correlated with radar-estimated rainfall amounts in the 0-1 h period after initial time. By forward-selection linear screening regression, a combination of these potential predictors was selected which explained a large percentage of the variance in observed rainfall.

From the forecast/observation dataset, we derived expressions for the probability that the observed rainfall will exceed 0.1, 0.25, 0.5, and 1 inch during the 0-1 h period. The probability values are also used to derive a categorical forecast rainfall field (essentially a forecast of the isohyets) by applying thresholds to the probabilities. We will demonstrate the degree of improvement that the extrapolative-statistical method offers over purely extrapolative forecasts.

This method is a form of Model Output Statistics (MOS), as described by Glahn and Lowry (1972). An application of this extrapolative-statistical (EXSTAT) approach for forecasting radar reflectivity at 30 and 60 minutes was documented by Saffle and Elvander (1981) and by Kitzmiller and Ator (1993). Another probabilistic forecast model for precipitation amount, based on extrapolation, was developed by Andersson and Ivarsson (1991). The Andersson-Ivarsson model's probabilities are based on the advection of a rainrate field inferred from initial-time reflectivity, rather than on a sample of historical radar observations. The probabilities are based on the assumption that the initial rainrates are always correct; our model accounts for errors in the extrapolation rainfall forecasts based on historical observations.

2. DATA USED IN DEVELOPMENT AND VALIDATION

The radar data used in this study were collected at the WSR-57 site in Oklahoma City, Oklahoma (OKC), which was first equipped with Radar Data Processor II (RADAP II) minicomputer equipment during the early 1980's. The RADAP II controlled the radar during volumetric scanning operations, digitized the reflectivity measurements, and automatically archived the observations. The most common scanning strategy used yielded a new volumetric scan every 10 minutes. For the development work described here, we used observations from the period 1985-1991. The cases were largely spring and summer convective events, which are those most likely to cause flash flooding in that part

of the Plains. Cases were manually edited to eliminate those featuring anomalous propagation echoes.

We wished to demonstrate the application of this method to WSR-88D as well as RADAP II data. Therefore a number of sample forecasts were prepared and verified with WSR-88D Archive Level II data from the sites at Twin Lakes and Norman, Oklahoma, and at St. Louis, Missouri.

3. PREPARATION OF REFLECTIVITY FORECASTS

To collect data for the development, we obtained sequences of scans that were continuous for 90 minutes. The nominal initial time was 30 minutes into the sequence; data prior to 30 minutes were used in deriving the echo motion, and those afterward for estimating the verifying rainfall amounts. A total of 392 sequences were used.

For each sequence, a storm motion vector (SMV) was derived by using the pattern-matching method described in the next section. Forecasts of the reflectivity and VIL fields were made at 10-minute intervals through 60 minutes. For the sake of simplicity, reflectivity was converted to nominal rainrates and integrated to create rainfall amounts. The VIL forecasts were temporally averaged.

Reflectivity was converted to rainrate through the WSR-88D Z-R relationship:

$$R = (Z/300)^{.714} \tag{1}$$

where R is rainrate in mm h^{-1} and Z is reflectivity in $mm^6\ m^{-3}$. The VIL was derived by integrating the Marshall-Palmer Z-M relationship with respect to height; this relationship is given by:

$$M = (Z/100)^{4/7} \tag{2}$$

where M is mixing ratio in g m^{-3} (Greene and Clark 1972; Marshall and Palmer 1948).

The statistical predictor/predictand sample was created by drawing the predictor values and the verifying rainfall value from every fifth box in each of the analysis grids. Only boxes between 20 and 80 nm from the radar were used. We thus obtained a sample of 16,463 individual cases.

4. DERIVATION OF STORM MOTION VECTORS

A binary-correlation pattern-matching technique was used to estimate the SMV for each test case. This method, described by Saffle and Elvander (1981) and by Ciccione and Pircher (1984), is computationally efficient and works well when the entire echo region does not change size or shape appreciably between images. The method reduces two radar grid maps to all "0" or "1" points according to some criterion. In the present work, this criterion was a ZTR value of 30 dBZ. The later map grid was shifted relative to the earlier one

until the binary correlation coefficient between the two was maximized. This coefficient (BC) is given by:

$$BC = N_{MATCH} / \sqrt{N_1 N_2} \tag{3}$$

where N_1 and N_2 are the number of "1" boxes (boxes with 30 dBZ echoes) in the first and second images, and N_{MATCH} is the number of boxes in which the earlier map and the shifted later map are "1" simultaneously. Note that BC approaches 1 as N_1 approaches N_2, and as N_{MATCH} approaches N_1 or N_2. Thus the binary correlation approaches unity if the reflectivity pattern undergoes only small changes in size and shape between the two images.

The process of finding the optimum shift between two binary patterns is shown schematically in Fig. 1, in which a dot indicates a "1" point and a blank square a "0" point. Shifting the pattern in t_0 to the left two boxes and up one box maximizes NMATCH. Since the patterns in both t_0 and t_1 have 17 "1" points, and the maximum value of NMATCH is 15, the maximum binary correlation between the images is is 0.88.

The SMV was estimated using two possible pairings of the images prior to initial time, namely t_0 and t_{-30}, and t_0 and t_{-20}. The 30-minute image pairing was used whenever it was possible to realize a binary correlation of at least 0.4 between the two images. If the maximum correlation was less than 0.4, the 20-minute pair was matched. If the value of BC was still less than 0.4, a suitable SMV from any pairing of images between 30 and 60 minutes before t_0 was used. In practice, we were able to obtain a reasonable SMV for all the cases included here.

5. RELATIONSHIPS BETWEEN FORECASTS AND RAINFALL OBSERVATIONS

The relationship between purely extrapolative forecasts of rainfall and observed (radar-estimated) rainfall appears in Fig. 2a. The solid curve shows the 50th percentile (median) observed rainfall, the dashed curve shows the amount that is exceeded in 75 per cent of the cases, and the dotted curve the amount exceeded in 25 per cent of the cases. For forecasted amounts above 0.25 inch, the forecasts are biased somewhat high (the observed amount is generally lower than that forecasted). Forecasts of time-averaged (0-60 minutes) VIL are also strongly correlated to the estimated rainfall (Fig. 2b), which is logical since our verifying observation is essentially a temporal average of reflectivity. The percentile values for average VIL above 15 kg m^{-2} had to be estimated from less than 100 cases, and thus the curves are not smooth. Both of these predictors explain about 40% of the variance in observed rainfall, and the spread between the 25th and 75th percentile values indicate substantial absolute errors in forecasting larger amounts.

Given the errors associated with direct forecasts of rainfall amount, it may be preferable to treat the forecasting problem as one of answering one or more yes/no questions. Will it rain more than 0.1 inch? Will it rain more than 0.5 inch? This information may be more useful to a forecaster who is most concerned with whether some critical amount will be reached, rather than with the absolute amount. Furthermore, it is logical to expect that an objective forecasting system designed to yield information on whether or not some specific threshold will be reached can be tuned more precisely than a system that attempts to forecast the rainfall directly.

Accordingly, we developed a set of four equations relating the candidate predictors of rainfall to the probability that the amounts 0.1, 0.25, 0.5, and 1 inch will be exceeded. When specific, or categorical, forecasts of rainfall amount are needed, they can be derived by comparing the probabilities to predetermined threshold values.

The relationships between extrapolation forecasts of rainfall and the relative frequency (or probability) of observed rainfall exceeding the four thresholds mentioned above appear in Fig. 3. Here, the forecasted rainfall field has been averaged over a 3x3 box region. This horizontal smoothing reduces random "noise" in the forecast field and improves the correlation with the predictand. When the extrapolative forecast is for 0.1 inch, there is a 50% probability that 0.1 inch will be reached or exceeded; when the forecast is from 0.25 to 0.5 inch, there is about a 40% chance that that amount will be observed. Forecasts of 1 inch have only about a 30% chance of being realized. This shows that the extrapolative model has less skill in forecasting the higher, rarer amounts.

We have treated the extrapolative rainfall forecasts for the 0-30 and 30-60 minute periods as predictors in their own right; because of errors in the extrapolation process, it is likely that there is more reliable information in the first 30 minutes of the period than in the final 30. We found that both of these predictors were strongly correlated to the 60-minute observed rainfall, and that the 0-30 minute forecast was indeed the more strongly correlated. As shown in Fig. 4, a 30-minute rainfall forecast of 0.5 inch or above was likely to be followed by 0.5 inch or more. This predictor was selected by the screening regression procedure used to derive expressions for the probability of amounts greater than 0.5 and 1 inch.

The observed relative frequency of rainfall in excess of 1 inch as a function of time-average VIL is shown in Fig. 5. Time-averaged VIL values above 10 kg m^{-2} are associated with a rather high probability of rainfall exceeding 1 inch, since VIL's in this range are usually observed only in and near convective cells.

The relative magnitudes of the statistical correlation between some of the candidate predictors and two of the binary predictands is shown in Fig. 6. The correlation statistic is the nonlinear correlation ratio (Panofsky and Brier 1968). This statistic shows approximately what fraction of the predictand variance is explained by the predictor without any assumptions as to the nature of the relationship between them (e.g., linear, exponential, or even monotonic). In the event that the predictor and predictand are linear with respect to each other, the correlation ratio approaches the square of the linear correlation coefficient. Note that even the initial-time reflectivity and VIL fields possess information with regard to the next hour's rainfall, an indication that reflectivity persistence alone indicates something about rainfall potential. However, for both the 0.1- and 1-inch predictands, extrapolation forecasts improve on the information in persistence. For the 0.1-inch predictand, the optimum predictor is the spatially-averaged rainfall extrapolation forecast. For the 1-inch predictand, forecasts of the spatially-averaged VIL field and the 0-30 minute rainfall forecast are the predictors most highly correlated with the event. These findings imply that loading of convective updrafts with rainwater, and the existence of high reflectivity immediately over and near the forecast grid box at initial time,

are important in controlling whether large rain amounts, rather than lesser amounts, occur there.

6. EXPRESSIONS FOR RAINFALL AMOUNT PROBABILITIES

We obtained expressions relating the various predictors to the probability of reaching rainfall thresholds by a combination of methods. Certain linearized predictors were created by defining curves or sets of straight lines that best describe the relationship between the 0-60 minute rainfall forecast and the event relative frequency. For example, the relationships between forecasted rainfall amount and the relative frequency of ≥ 0.1 and ≥ 0.25 inch rainfall, shown in Fig. 3, suggested a set of lines with varying slopes. These linearized predictors were then submitted along with those in the basic set to a forward-selection linear screening regression procedure.

In the expressions below, rainfall amounts are in .01 inch and VIL is in kg m^{-2}. For the 0.1-inch predictand, the best fit to the dependent data was obtained by defining two straight lines relating rainfall amount to event relative frequency, and truncating the output probability value at 90%:

$$
\begin{aligned}
P(\geq 0.1) &= 1 + 4.2[\mathrm{RAIN60}]_S, \quad \text{if } [\mathrm{RAIN60}]_S \leq 15; \\
&= 65 + 0.706\{[\mathrm{RAIN60}]_S - 16\}, \text{ if } 15 < [\mathrm{RAIN60}]_S \leq 50; \\
&= 90, \text{ if } [\mathrm{RAIN60}]_S > 50.
\end{aligned} \tag{4}
$$

where P is in per cent and $[\mathrm{RAIN60}]_S$ is the extrapolation forecast of 60-minute rainfall, averaged over the adjacent 3x3 grid box region. This set of linear functions explains 49% of the predictand variance.

A similar set of expressions was derived for the probability of 0.25 inch:

$$
\begin{aligned}
P(\geq 0.25) &= 0.5 + 1.39[\mathrm{RAIN60}]_S, \quad \text{if } [\mathrm{RAIN60}]_S \leq 50; \\
&= 71 + 0.31\{[\mathrm{RAIN60}]_S - 51\}, \text{ if } 50 < [\mathrm{RAIN60}]_S \leq 80; \\
&= 85, \text{ if } [\mathrm{RAIN60}]_S > 80.
\end{aligned} \tag{5}
$$

This function explains 42% of the predictand variance.

For the 0.5-inch predictand, the following expression involving both the 30- and 60-minute rainfall forecasts was derived:

$$P(\geq 0.5) = 0.27 + 0.41(\mathrm{RAIN30}) + 0.22(\mathrm{RAIN60}) \tag{6}$$

where RAIN30 is the 0-30 minute rainfall amount forecast and RAIN60 is the 0-60 minute rainfall forecast. Both rainfall forecasts are for the grid box at which the probability is defined; these predictors are not spatially averaged. The expression in (6) explains 30% of the predictand variance.

The expression for the probability of rainfall in excess of 1 inch is:

$$P(\geq 1) = -.135 + 1.87[\mathrm{VIL60}]_{TS} + 0.14(\mathrm{RAIN30}) \tag{7}$$

where $[VIL60]_{TS}$ is the 0-60 minute VIL forecast, averaged over that time period and over the 3x3 grid box region adjacent to the box in question. The expression in (7) explains 21% of the predictand variance.

7. SKILL LEVEL OF PROBABILITY FORECASTS

The skill of the probability values obtained from (4)-(7) above can be assessed by converting them to categorical (yes/no) forecasts and scoring the forecasts. The most common method of deriving categorical forecasts from probabilities is by applying a threshold value: all probabilities below the threshold are interpreted as "no," those at and above the threshold as "yes." Such forecasts can be scored in terms of probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI). The calculation and characteristics of these scores have been discussed by Donaldson et al. (1975) and by Schaefer (1990). Another measure of forecast utility is bias, which is the ratio of the number of "yes" forecasts to "yes" observations.

The skill of algorithms of this type appears higher in the dependent (development) dataset than when evaluated on independent cases. This inflation of skill occurs because the probability equations incorporate not only physical relationships common to all precipitation systems, but statistical properties peculiar to the development data sample.

To improve our assessment of the skill level, we employed a form of cross validation (Elsner and Schmertmann 1994). In our approach, three new development data samples were created by withholding, in turn, all cases from 1985-87, from 1989, and from 1991. Each new development subsample had from 12,000 to 14,000 cases. Prototype probability equations were derived from each of the subsamples, by forward-selection linear screening regression. The equations were then evaluated on the data withheld from development and the forecasts verified and scored. Thus, each forecast was prepared from an algorithm that had been developed from cases observed during other calendar years. There were 5460 cases in the verification sample.

Scores for each of the four probability equations are shown in Figs. 7-10. The POD, FAR, and CSI scores for all possible probability thresholds from 1% to 50% are shown. For example, for the $P(\geq 0.1)$ equation (Fig. 7), a yes/no threshold of 20% yields a POD of 0.84 and an FAR of 0.42. That is, 84% of the 0.1-inch precipitation events were covered by "yes" forecasts, and 42% of all "yes" forecasts were false alarms in which the observed rainfall was less than 0.1 inch.

For the higher rainfall thresholds, skill in terms of the critical success index decreases. When forecasting the occurrence of higher, rarer rainfall amounts, it becomes necessary to issue more false alarms in order to correctly forecast the same percentage of "yes" events. Thus the peak CSI for the 0.1-inch forecasts was 0.54, those for the 0.25- and 0.50-inch forecasts about 0.4, and that for the 1-inch forecasts was 0.25 (Figs. 7-10).

8. COMPARISION OF EXTRAPOLATIVE-STATISTICAL WITH PURELY EXTRAPOLATIVE FORECASTS

As is apparent from our results, purely extrapolative forecasts themselves possess considerable skill even without the addition of a statistical compo-

nent. However, we can demonstrate that extrapolative-statistical forecasts are indeed more skillful than ones that could be derived by simply applying thresholds to extrapolative forecasts.

In these experiments, a set of purely extrapolative forecasts was evaluated in the same manner as probability forecasts. That is, the POD, FAR, and bias were determined over a range of forecasted rainfall amount thresholds, just as those scores had been determined for each in a range of probability thresholds. We then compared the FAR and bias values yielded by the probabilistic forecasts to those from the extrapolative forecasts, when the POD values were identical. When both algorithms yield the same POD, the more skillful one will yield a lower FAR value, or a bias closer to unity. Because the probability forecasts were evaluated on an essentially independent data sample, any improvement offered by the EXSTAT method must be due to higher skill, and not simply fitting of the model to the development data sample.

In Fig. 11a, we show the biases associated with various POD values for both the EXSTAT algorithm and the extrapolation model for forecasts of ≥ 0.1 inch of rain. The data sample was the same one used in the cross-validation experiment described in Section 7. For POD values above 0.8, the EXSTAT algorithm does produce lower biases (and fewer false alarms). At lower POD values, the EXSTAT algorithm actually produces a higher bias. This is probably because the probability model was specifically tuned to yield the most precise specification of rain/no rain, rather than being tuned to yield the lowest possible bias for some given POD. Moreover, the probability forecasts were generally a function of the purely extrapolative forecasts alone; the regression procedure added little further information.

At higher amount thresholds, the EXSTAT algorithms generally yield superior skill for most POD values. As shown in Fig. 11b,c, the EXSTAT algorithm can yield biases about three fourths the magnitude of those from the extrapolation algorithm. The improvement is most dramatic for the 1-inch amount. Thus the slightly higher computational load imposed by the need to extrapolate both the reflectivity and VIL fields appears to be rewarded with noticeably reduced biases in categorical forecasts.

10. CATEGORICAL RAINFALL FORECASTS

For some purposes, it is desirable to express the rainfall forecasts in terms of amount rather than yes/no or probability. One of the simplest ways of deriving categorical amounts from probabilities is to compare the probabilities of to predetermined threshold values. Thus we may say that if P(≥ 0.1) exceeds a threshold of 30%, then the rainfall will be at least 0.1 inch. If the 0.1-inch threshold is exceeded, and if P(≥ 0.25) exceeds another threshold, say 27%, then the rainfall is forecasted to be at least 0.25 inch, and so on. This procedure essentially creates an isohyetal field from maps of the four probability fields.

We have developed a set of thresholds that produce reasonable agreement between forecasted and observed rainfall fields for a variety of cases. Our aim was to determine thresholds that produce a reasonably high probability of detection (at least 0.6) and not too large a false alarm ratio (no higher than 0.7). Note that these POD and FAR values are valid only for the subset of cases where the forecasts have reached the next lower category. For example,

we can detect 1-inch events only when the thresholds for 0.1, 0.25, and 0.5 inch have been reached. Therefore, for the purposes of calculating the POD, the total number of 1-inch events considered is a fraction of the total number.

By experimenting within the development sample of 16463 cases, we arrived at the following thresholds:

For 0.1 inch: $P(\geq 0.1) \geq 27\%$;

For 0.25 inch: $P(\geq 0.25) \geq 25\%$;

For 0.5 inch: $P(\geq 0.5) \geq 21\%$;

For 1 inch: $P(\geq 1.0) \geq 18\%$.

A verification of the categorical forecasts produced by these thresholds appears in Fig. 12a. The outcome of all forecasts for the lowest rainfall category (< 0.1 inch) appears in the leftmost column, outcomes for all forecasts of 0.1-0.24 inch appear in the second column from the left, and so on. As can be seen from the marginal totals, the forecasts are biased toward larger values than observed; this is a result of our desire to achieve a reasonably high POD, even when the FAR exceeded 0.5. For cases when the forecast is greater than 0.1 inch, the observed amount is most likely to fall into the next lower category.

For all forecasts greater than or equal to 0.1 inch, 31% fell into the correct category, and 82% within one category of the correct one. Of all forecasts for amounts greater than or equal to 0.25 inch, 68% were within one category of the verifying one.

To validate the EXSTAT algorithm's performance for WSR-88D base data, we prepared and verified forecasts for four events observed by the radars at Twin Lakes, Oklahoma; Norman, Oklahoma; and St. Louis, Missouri. These cases all included fairly intense convection. The verification results (Fig. 12b) included all grid boxes between 20 and 75 nm from the radar, rather than every fifth one as in the development sample. These events included substantially greater coverage by 0.5 inch and larger amounts than appeared in the development sample. Still, we again found that 33% of the forecasts for ≥ 0.1 inch verified, and 82% were within one category of the correct one. For forecasts ≥ 0.25 inch, 75% were within one category of the correct one.

It should be noted that forecasting for rainfall categories rather than for specific amounts sometimes masks the existence of very large errors, particularly for the few extreme amounts (> 2 inches) sometimes observed. However, the thresholding procedure generally produces a useful isohyetal field, as will be shown.

11. RAINFALL PROBABILITY FIELDS FOR A DEMONSTRATION CASE

The relationship between initial-time reflectivity and VIL, and the resulting rainfall probabilities, appear in Figs. 13-15. The initial-time maps (Fig. 13) are from the WSR-88D Operational Support Facility radar at Norman, Oklahoma, at 1800 UTC, 11 May 1992. The radar is near the center of the radar

image. An area of mature thunderstorms was creating light to moderate rainfall over the region to the northeast of the site, with reflectivities mostly less than 40 dBZ (Fig. 13a). A few new thunderstorms, with VIL values above 30 kg m^{-2}, were developing immediately to the south and north of the radar (Fig. 13b). All activity appeared to be moving to the northeast at about 16 m s^{-1}, and this velocity was used in preparing the forecast fields.

The probability fields derived from these data and the forecast echo velocity appear in Figs. 14-15. The probability of rainfall exceeding 0.1 inch (Fig. 14a) exceeds 90 per cent over much of the region initially covered by reflectivity in excess of 30 dBZ. For the higher amounts, the probabilities decrease; the highest probability for 1+ inch is only 30 per cent, which is associated with a high-VIL storm southeast of the radar (Fig. 15b). A forecaster concerned with the chance that rainfall might exceed 0.5 in during the period 1800-1900 UTC would probably focus his/her attention on the areas where the corresponding probability exceeds about 20% (Fig. 15a), namely in two areas roughly north and south of the radar site.

A categorical rainfall forecast, and the verifying rainfall field, appear in Fig. 16. The forecast equations apply only within about 90 nm (180 km) of the radar, and the forecast field is truncated on its eastern edge (Fig. 16a). The forecast clearly indicates the heavy rainfall associated with thunderstorms immediately north and south of the radar. The lighter rainfall to the northeast of the radar is overforecasted in terms of extent and intensity, while a newly-developing storm that produced a heavy rain signature far to the south of the radar site is outside the forecast region (Fig. 16b).

12. CONCLUSIONS AND FUTURE WORK

Our results demonstrate the manner in which the EXSTAT method improves on purely extrapolative forecasts of reflectivity, at the expense of a modest increase in computing time. The use of a variety of statistical predictors of rainfall particularly improves forecasts of rainfall amounts in excess of 0.5 inch.

The skill of these forecasts is somewhat exaggerated by our use of a purely radar-based predictand. We have essentially forecasted time-averaged reflectivity rather than rainfall as might be observed by a gage network. Thus forecast errors due to temporal and spatial variations in the Z-R relationship are not accounted for. Therefore, we have begun work on refining the algorithm by utilizing rainfall estimates that incorporate both WSR-88D and gage data. These 4-km, 1-h estimates are produced by the WSR-88D Stage III precipitation analysis operation (Hudlow 1988) conducted at the Arkansas/Red River Basin River Forecast Center in Tulsa, Oklahoma. These estimates should provide a predictand with higher absolute accuracy than can be obtained from radar alone.

ACKNOWLEDGEMENTS

Robert Saffle provided the initial impetus for this work, and gave helpful advice about getting it underway. Jeffrey Ator and Bryon Lawrence created the computer code for binary pattern matching and extrapolation forecasting. National Weather Service personnel at the Oklahoma City forecast office

created the digital radar archive in raw form, and Melvina McDonald and other Techniques Development Laboratory personnel rearchived it in final form.

REFERENCES

Andersson, T., and K.-I. Ivarsson, 1991: A model for probability nowcasts of accumulated precipitation using radar. J. Appl. Meteor., 30, 135-141.

Austin G. L., and A. Bellon, 1974: The use of digital weather radar records for short-term precipitation forecasting. Quart. J. Royal Meteor. Soc., 100, 658-664.

Charba, J. P., 1977: Operational system for predicting severe local storms two to six hours in advance. NOAA Technical Memorandum 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. Proceedings Second International Symposium on Nowcasting, Norrkoping, European Space Agency, 241-246.

Conway, B. J., 1987: FRONTIERS: An operational system for nowcasting precipitation. Proceedings Symposium on Mesoscale Analysis and Forecasting, 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. Preprints Ninth Conference on Severe Local Storms, Norman, Amer. Meteor. Soc., 321-326.

Elsner, J. B., and C. P. Schmertmann, 1994: Assessing forecast skill through cross validation. Wea. Forecasting, 9, 619-624.

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11, 1203-1211.

Greene, D. R., and R. A. Clark, 1972: Vertically-integrated liquid--A new analysis tool. Mon. Wea. Rev., 100, 548-552.

Hudlow, M. D., 1988: Technological developments in real-time operational hydrologic forecasting in the United States. Journal of Hydrology, 102, 69-92.

Kelsch, M., 1990: Estimating flash flood potential with radar. Preprints 16th Conference on Severe Local Storms, Kananaskis Park, Amer. Meteor. Soc., 621-626.

Kitzmiller, D. H., and J. B. Ator, 1993: Extrapolative-statistical forecasts of radar reflectivity. National Weather Digest, 18, No. 2, 12-19.

Marshall, J. S., and W. M. Palmer, 1948: The distribution of raindrops with size. J. Meteor., 5, 165-166.

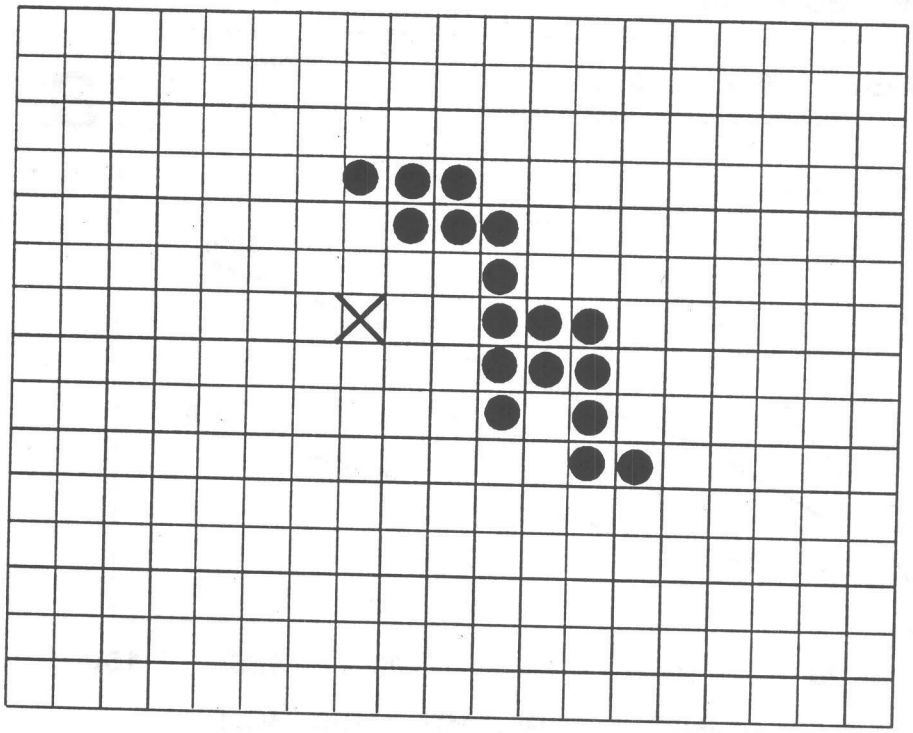
Panofsky, H. A., and G. W. Brier, 1968: Some Applications of Statistics to Meteorology. Pennsylvania State University Press, University Park, 224 pp.

Saffle, R. E., and R. C. Elvander, 1981: Use of digital radar in automated short range estimates of severe weather probability and radar reflectivity. Preprints Seventh Conference on Probability and Statistics in Atmospheric Sciences, Monterey, Amer. Meteor. Soc., 192-199.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. Wea. Forecasting, 5, 570-575.

Takemura, Y., Y. Makihara, K. Takase, and K. Aonashi, 1987: Operational experiment in very short range forecasting of precipitation. Proceedings Symposium on Mesoscale Analysis and Forecasting, Vancouver, European Space Agency, 239-244.

Walton, M. L., and E. R. Johnson, 1986: An improved precipitation projection procedure for the NEXRAD flash-flood potential system. Preprints 23rd Conference on Radar Meteorology and Conference on Cloud Physics, Snowmass, Amer. Meteor. Soc., JP62-JP65.



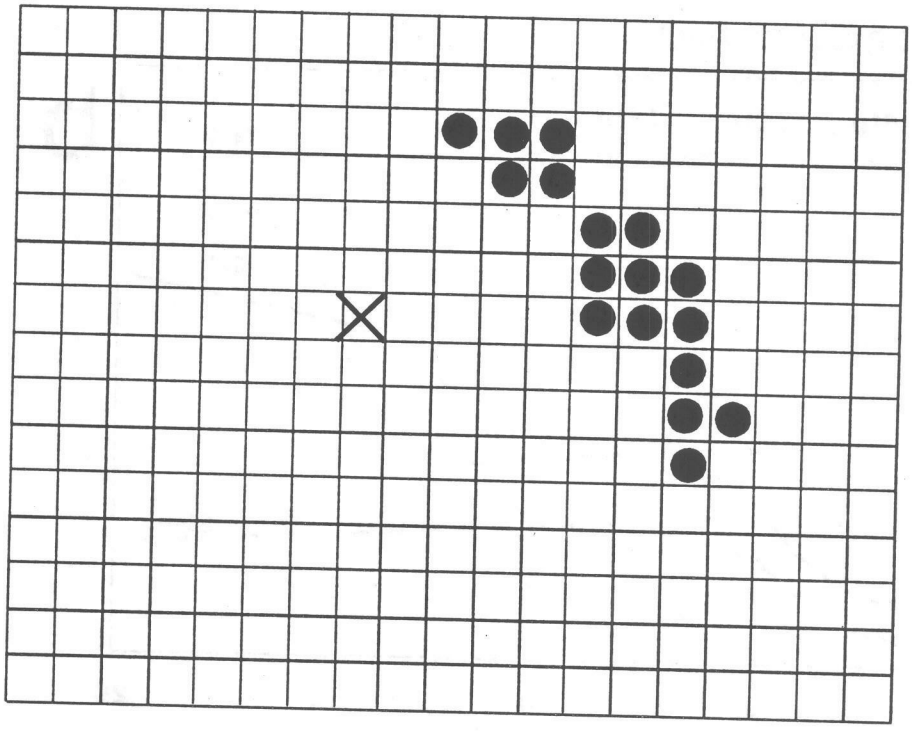
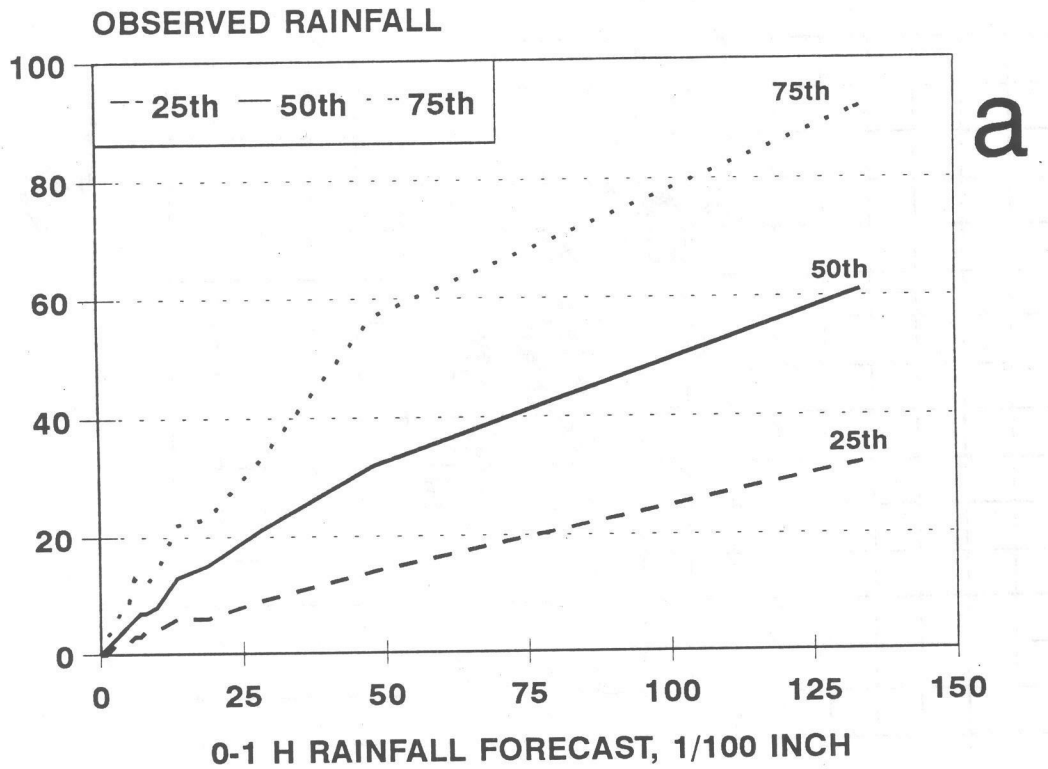


Figure 1. Two binary patterns, with "on" or "1" points shown as black dots, "off" or "0" points as empty boxes. Shifting pattern t_0 two boxes right and one box up maximizes the binary correlation with pattern t_1.



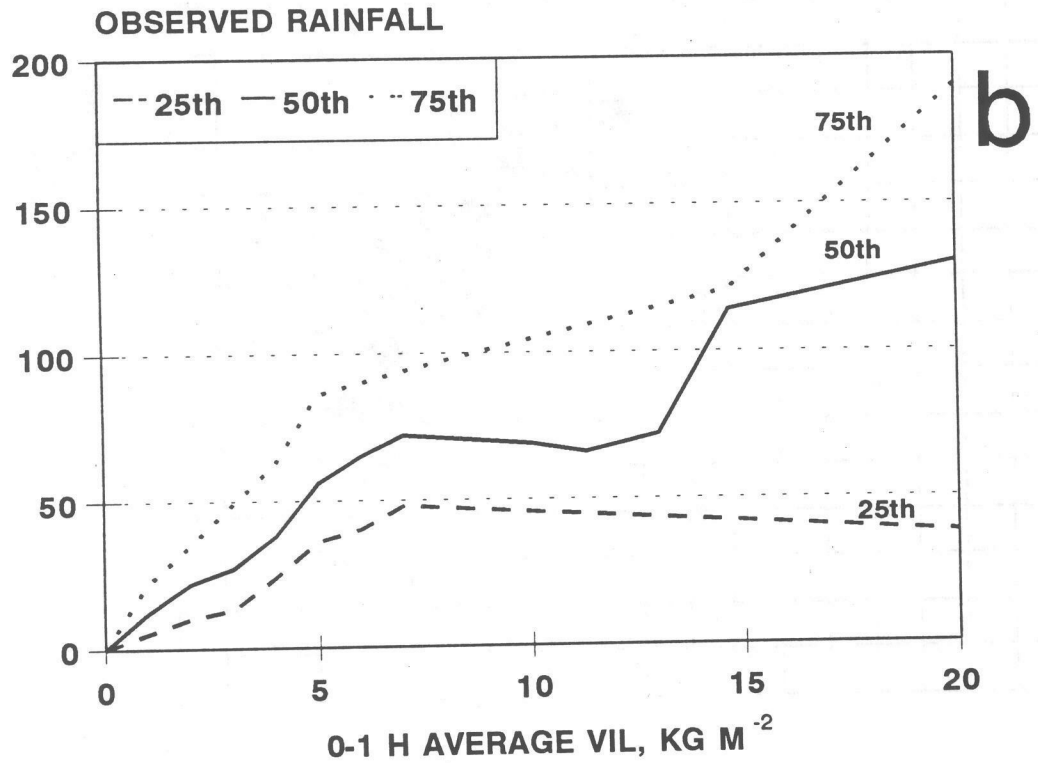


Figure 2. Distribution of observed 1-h rainfall amounts as a function of (a) a purely extrapolative rainfall forecast of reflectivity and (b) time-averaged VIL.

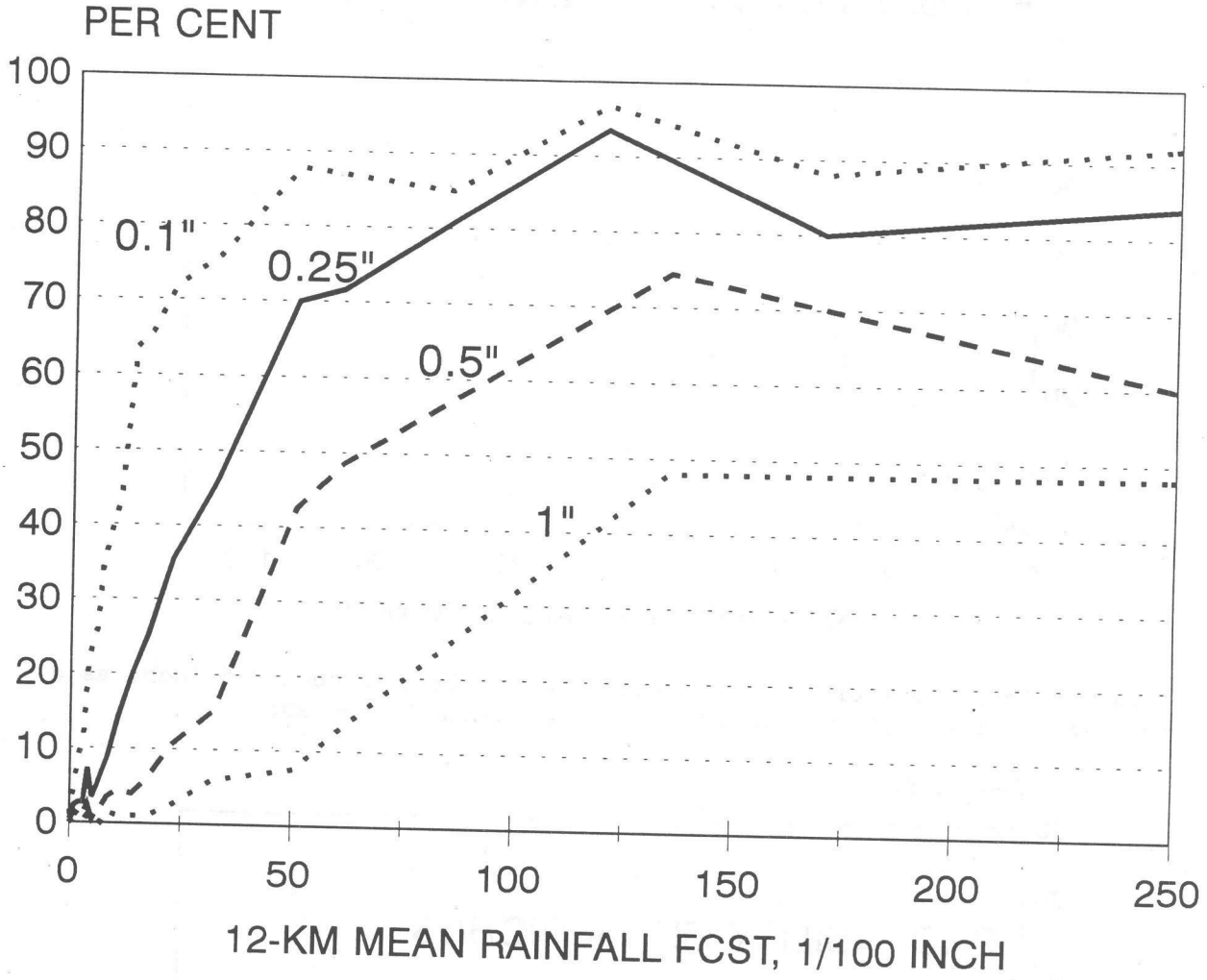


Figure 3. Percentage of cases in which various 1-h rainfall thresholds are exceeded, as a function of spatially-averaged, extrapolative rainfall forecasts.

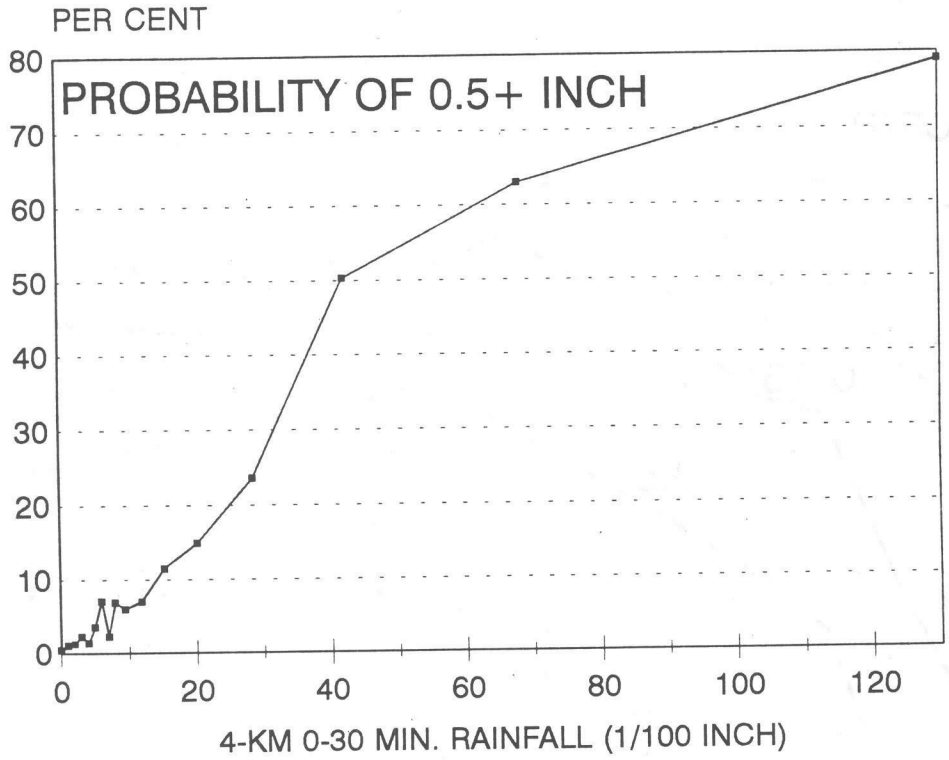


Figure 4. Percentage of cases in which 1-h rainfall exceeds 0.5 inch, as a function of the 0-30 minute extrapolative rainfall forecast.

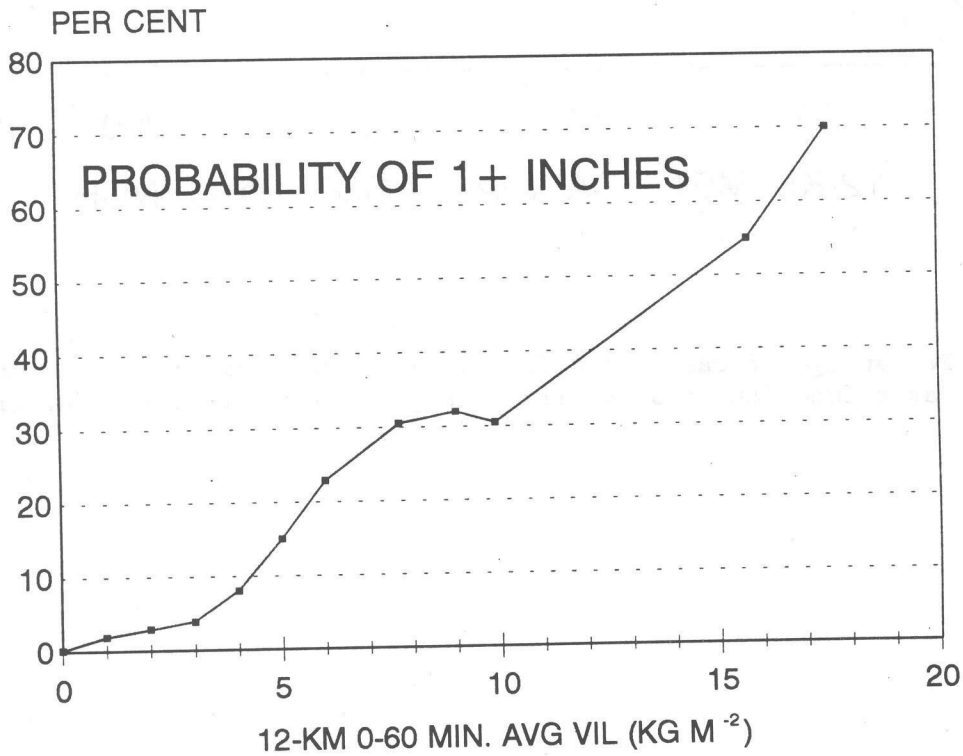
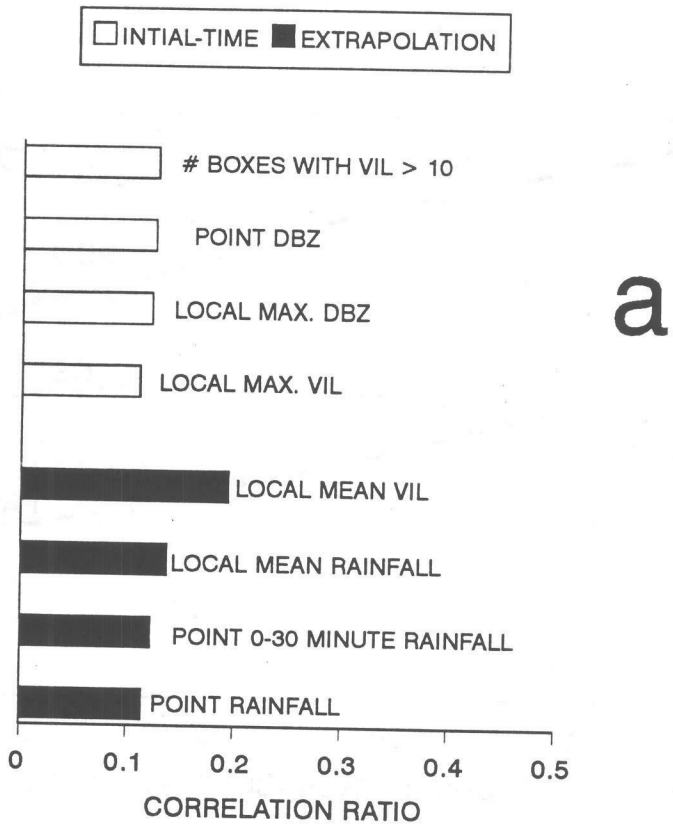


Figure 5. As in Fig. 4, except the threshold is 1 inch; and the predictor is the spatially-averaged VIL forecast.



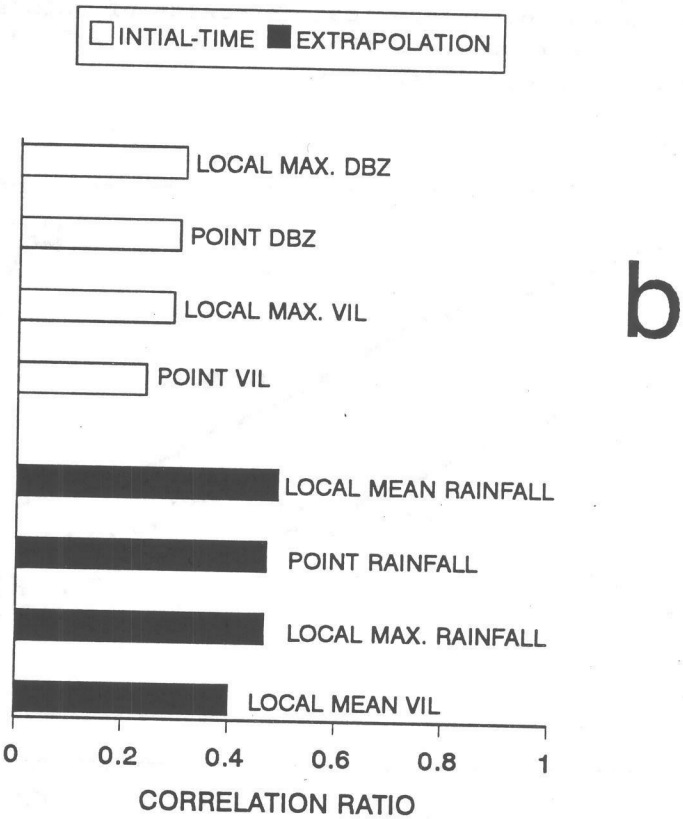


Figure 6. Nonlinear correlation ratios for various rainfall probability predictors, with respect to (a) a 1-inch threshold and (b) a 0.1-inch threshold.

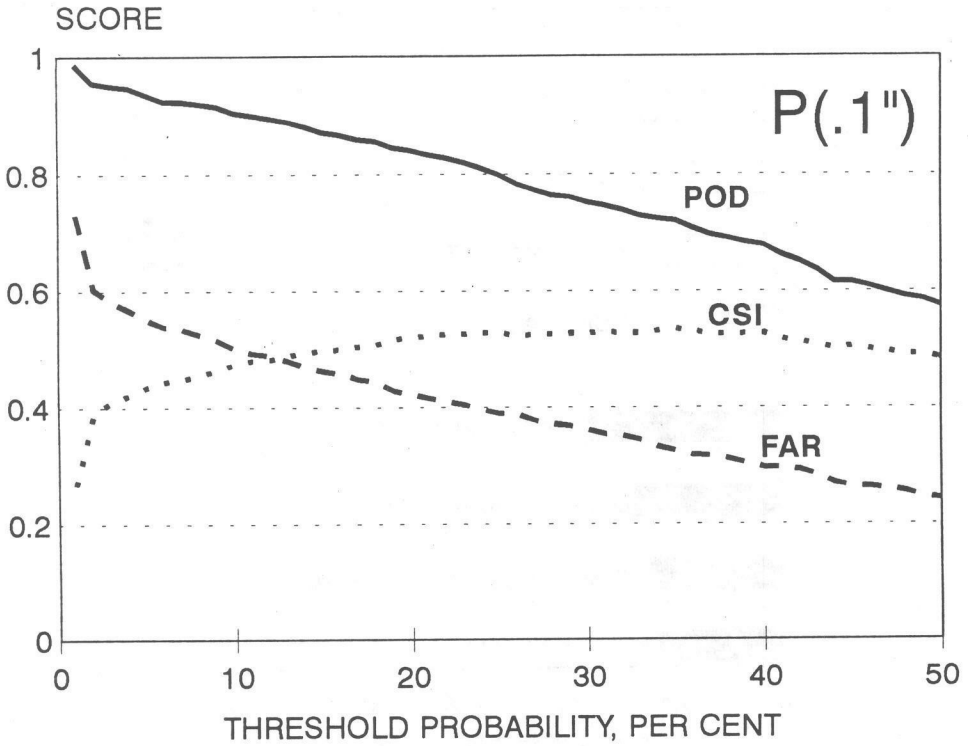


Figure 7. Skill scores for categorical forecasts of rainfall exceeding 0.1 inch. Solid line represents POD, dashed line FAR, dotted line CSI. Results are based on cross validation.

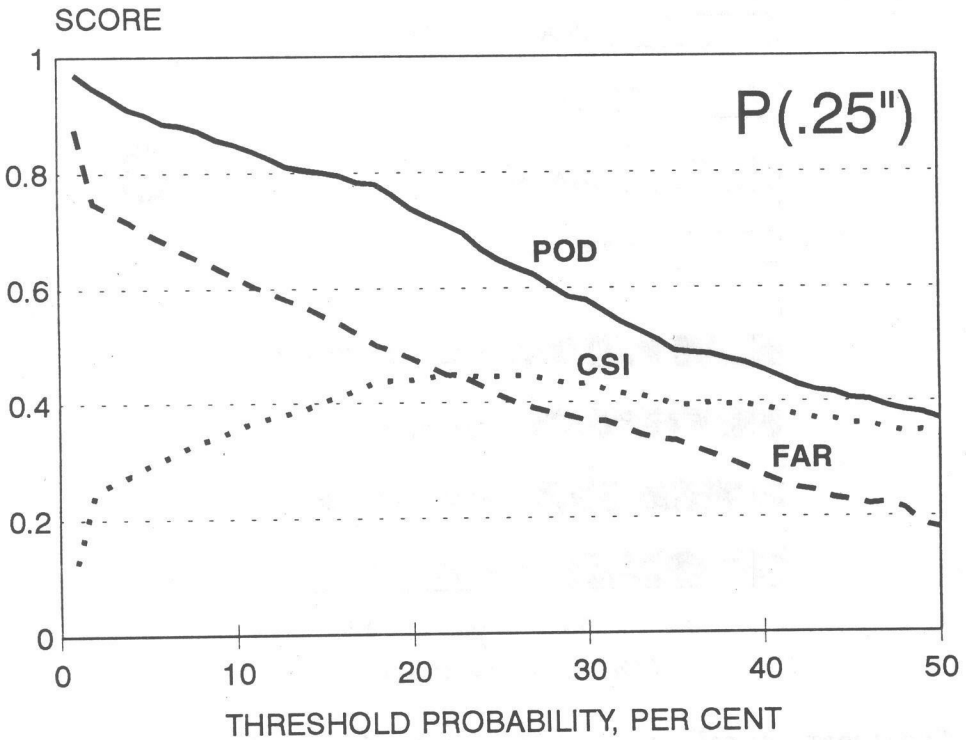


Figure 8. As in Fig. 7, except for a 0.25-inch threshold.

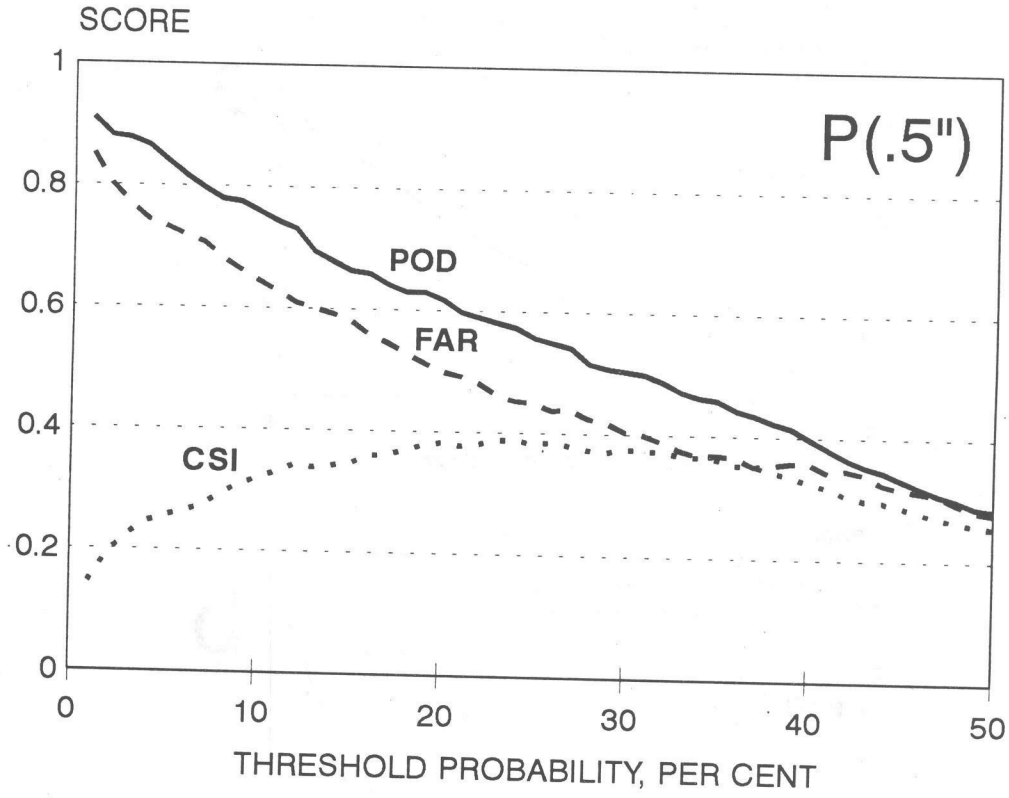


Figure 9. As in Fig. 7, except for a 0.5-inch threshold.

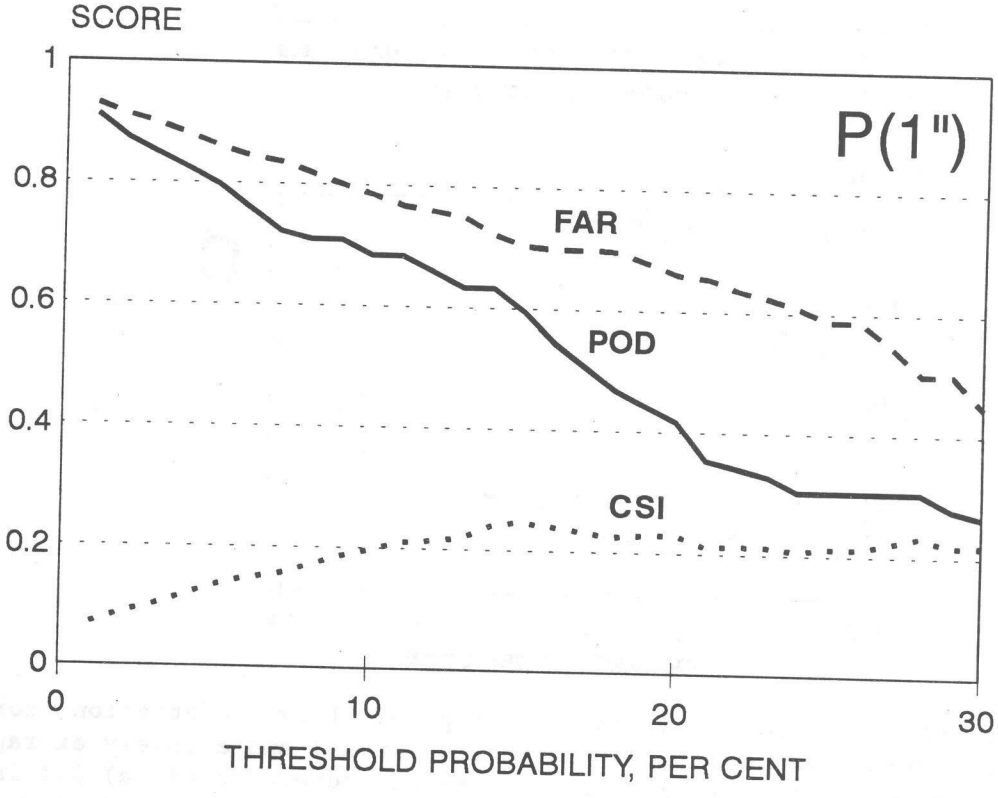


Figure 10. As in Fig. 7, except for a 1-inch threshold.

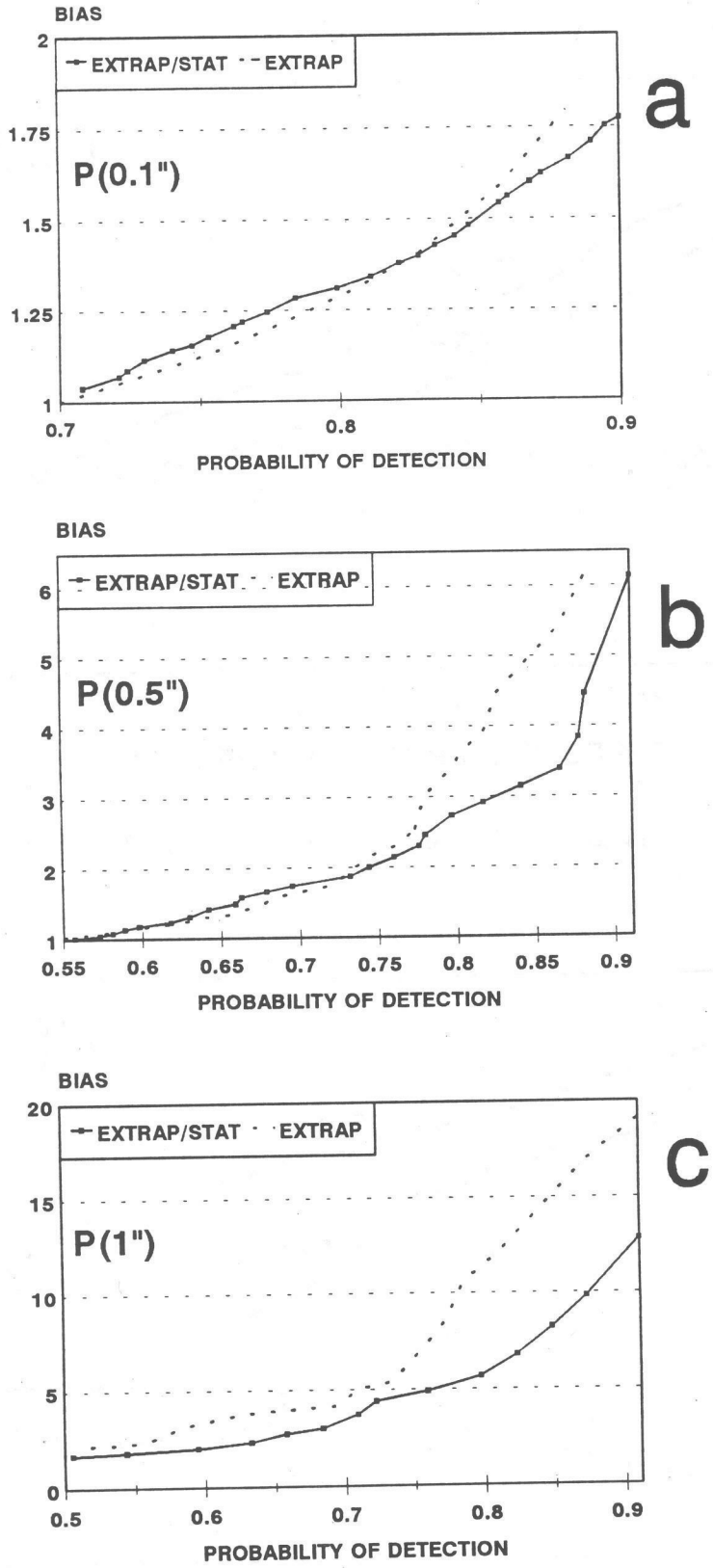


Figure 11. Forecast bias as a function of probability of detection, for extrapolative-statistical forecasts (solid curves) and a purely extrapolative forecasts (dotted lines). Scores are for forecasts of (a) 0.1 inch, (b) 0.5 inch, and (c) 1 inch. Results are based on cross validation.

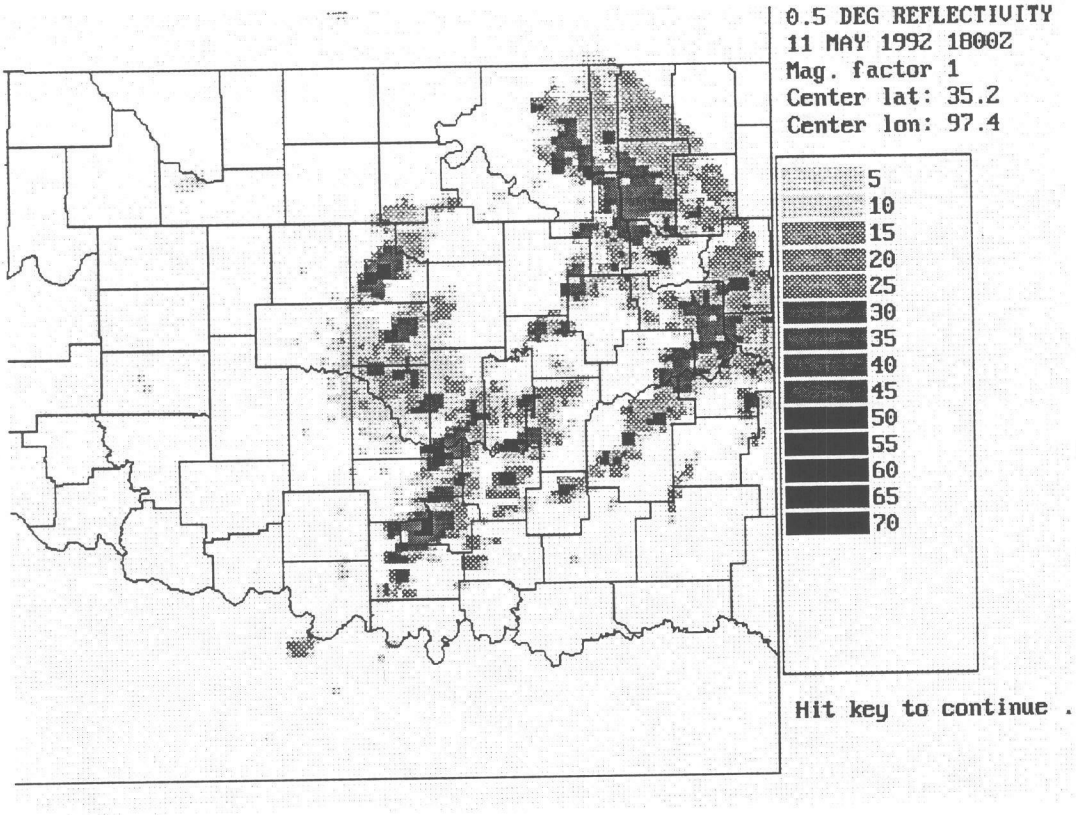
a

OBSERVED \ FORECASTED	<.1	.1-.24	.25-.49	.5-.99	1+	TOTALS:
<.1	13962	554	127	71	4	14718
.1-.24	303	349	163	75	6	896
.25-.49	87	102	125	114	26	454
.5-.99	41	29	33	113	59	275
1+	16	10	13	36	45	120
TOTALS:	14409	1044	461	409	140	

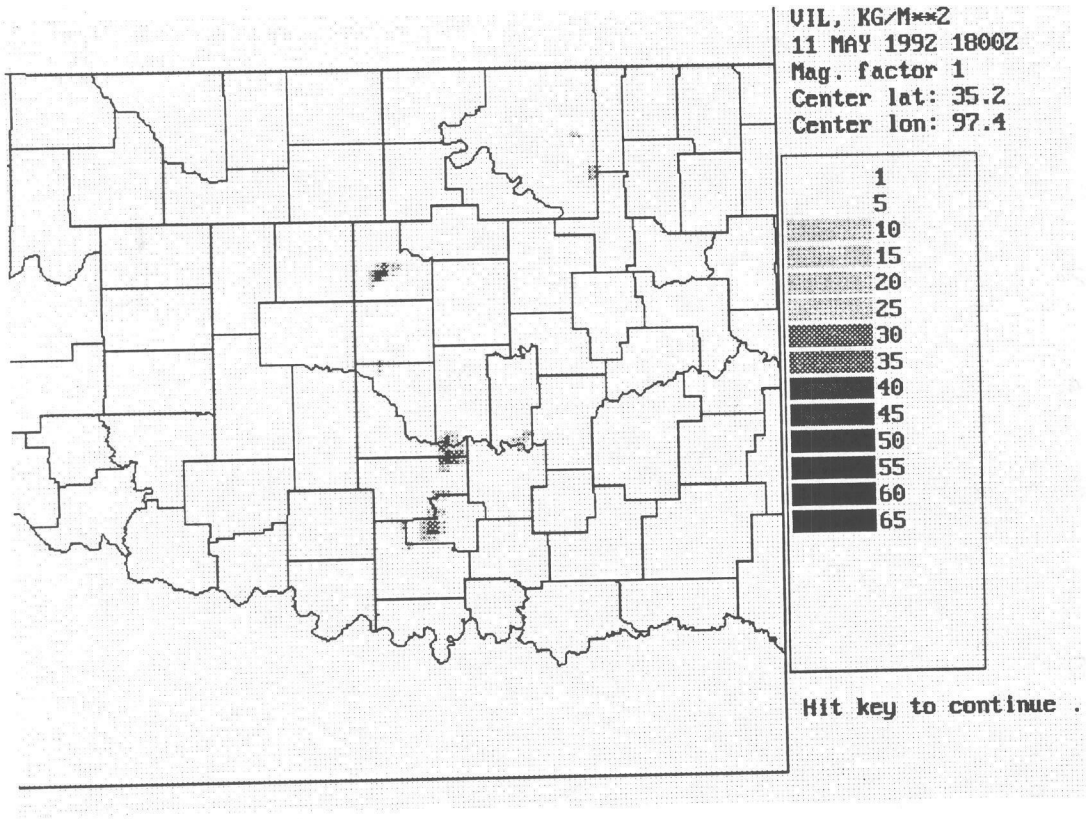
b

OBSERVED \ FORECASTED	<.1	.1-.24	.25-.49	.5-.99	1+	TOTALS:
<.1	27147	1553	375	251	33	29359
.1-.24	524	871	516	371	27	2309
.25-.49	166	231	262	723	62	1444
.5-.99	109	101	200	748	166	1324
1+	27	41	67	340	565	1040
TOTALS:	27973	2797	1420	2433	853	

Figure 12. Verification contigency tables for categorical rainfall amount forecasts. Results in (a) are from the dependent data sample based on RADAP II radar observations; those in (b) are from four independent cases based on WSR-88D observations.

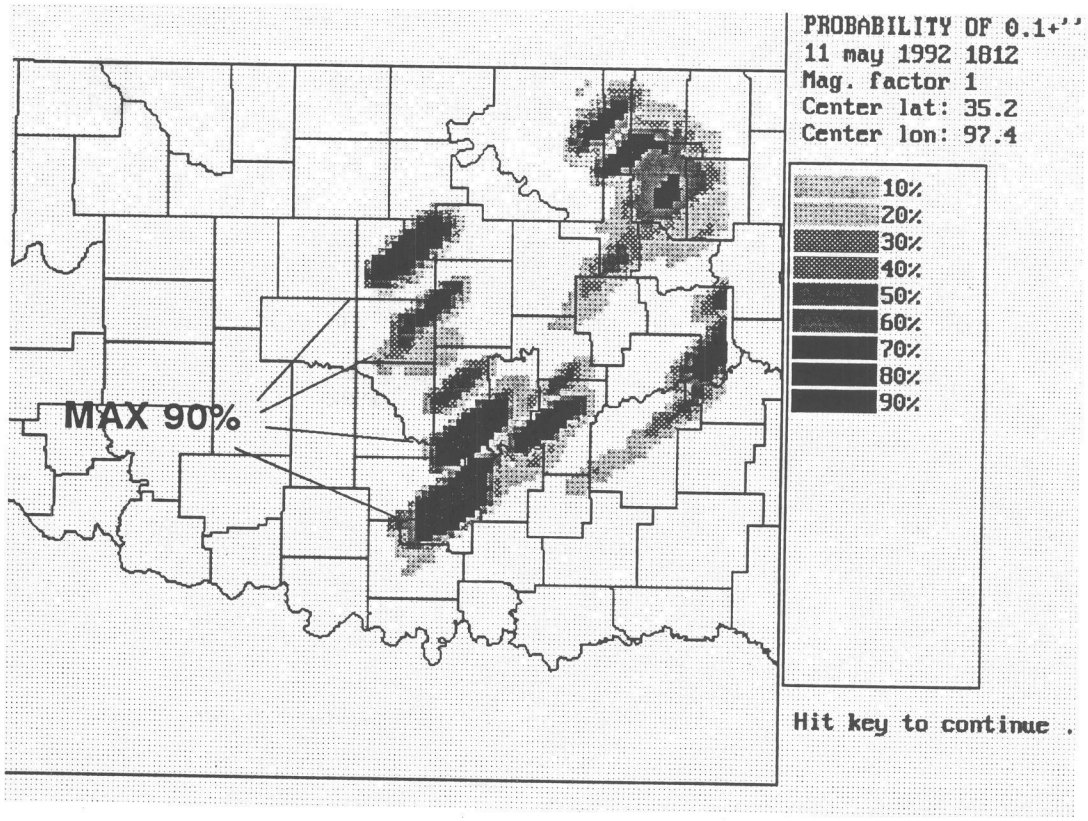


a

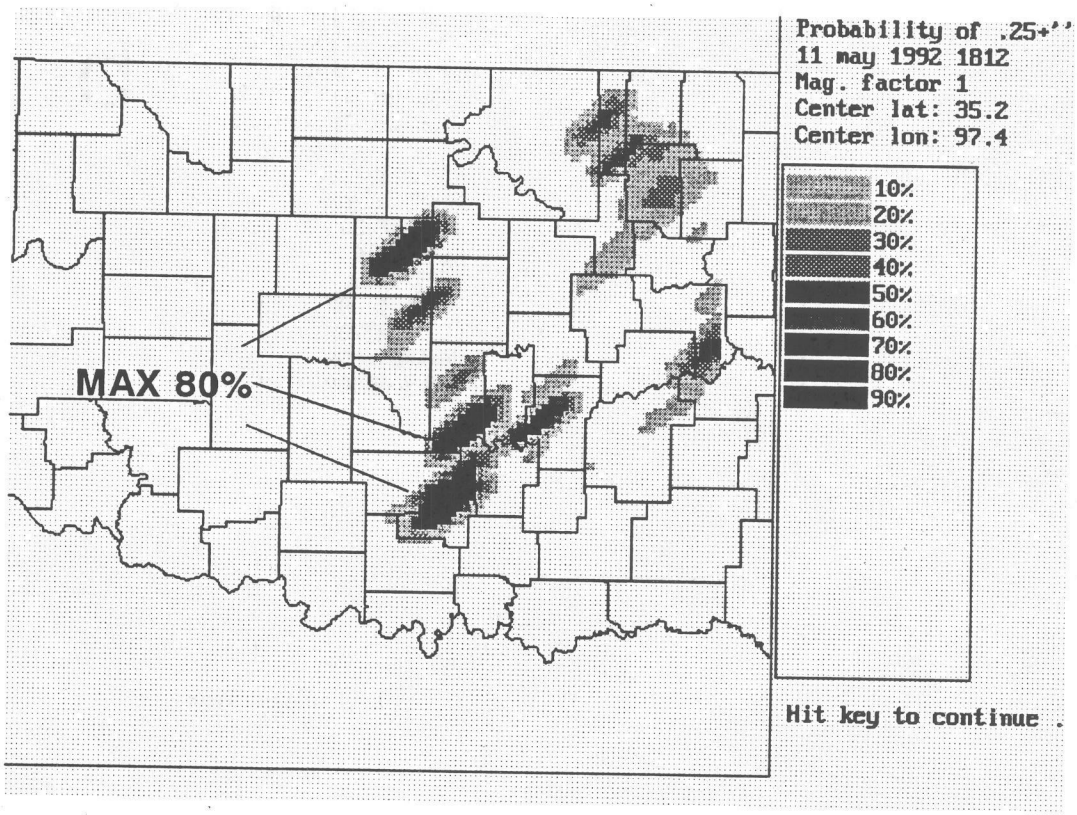


b

Figure 13. Precipitation distribution at 1800 UTC, 11 May 1992, as observed by the Norman, Oklahoma, WSR-88D. The field in (a) is 0.5° reflectivity, that in (b) is VIL.

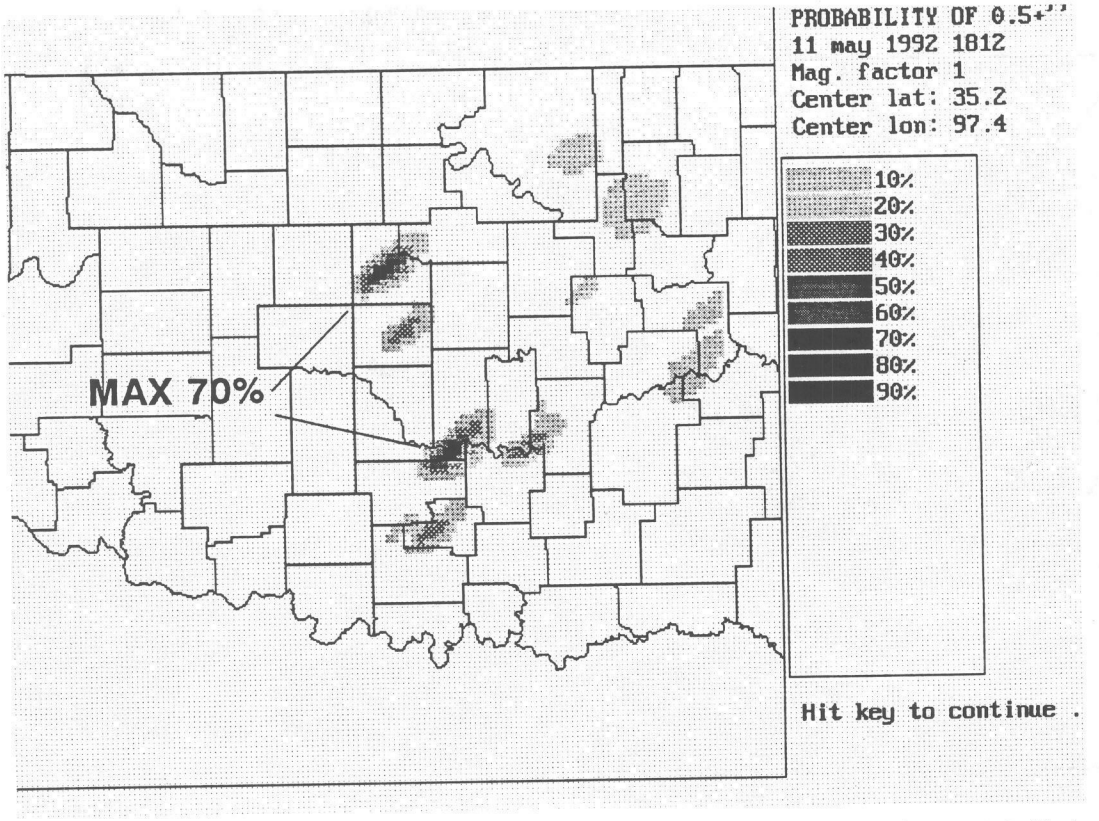


a

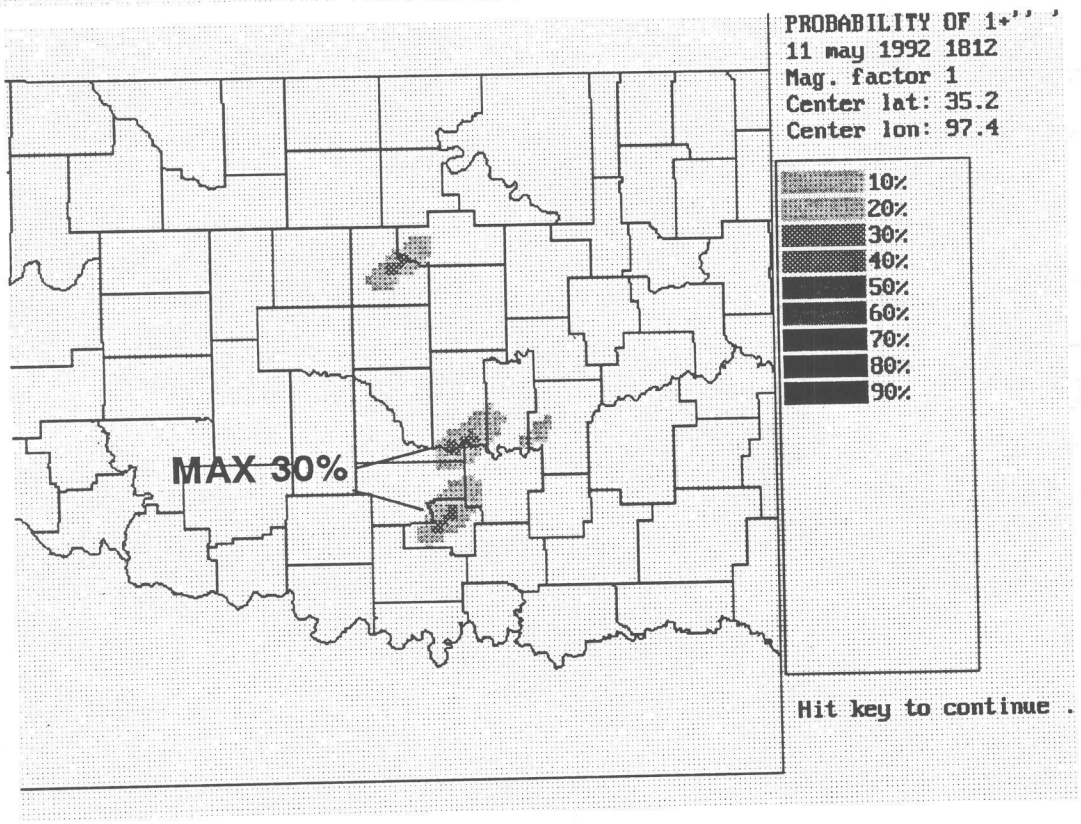


b

Figure 14. Rainfall probabilities based on reflectivity and VIL in Fig. 13. Probabilities are for (a) 0.1 inch and (b) 0.25 inch.

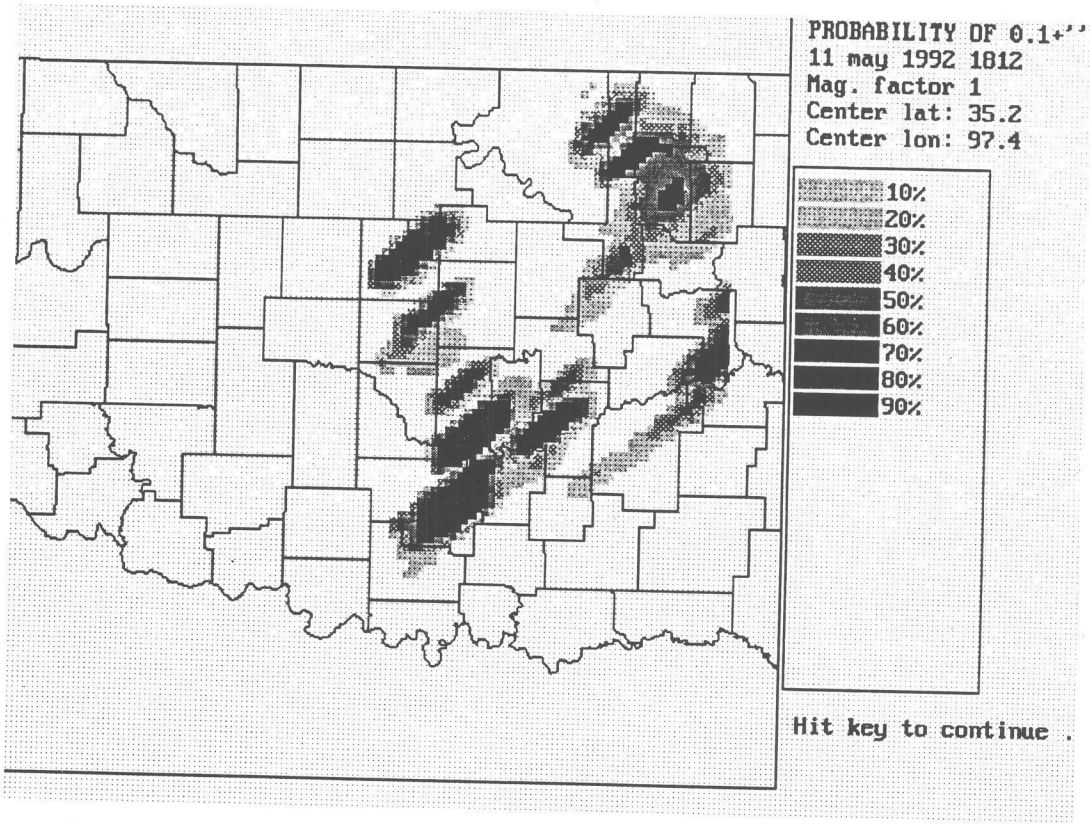


a

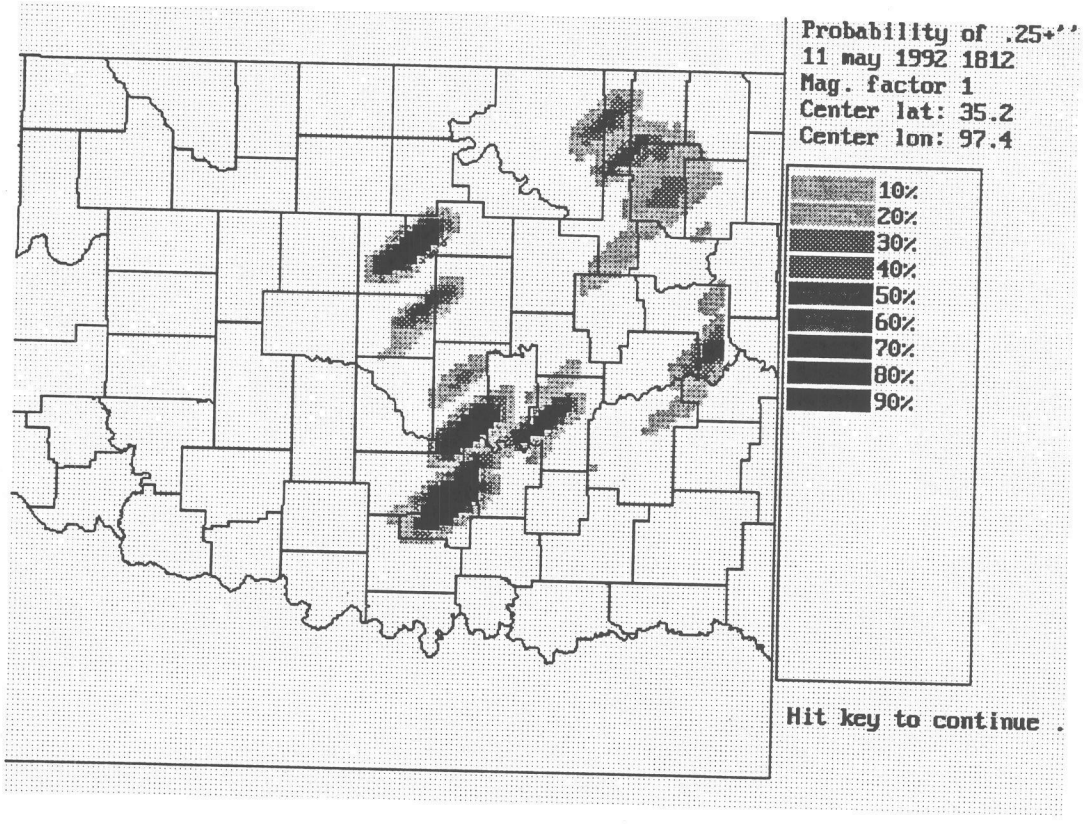


b

Figure 15. As in Fig. 14, except that rainfall thresholds are (a) 0.5 inch and (b) 1 inch.



a



b

Figure 16. Categorical rainfall forecast based on probabilities in Figs. 14-15 (a), and verifying amounts (b).