A Statistical Approach to Short-Term Thunderstorm Outlooks

RICHARD P. MCNULTY[1]

Techniques Development Unit, National Severe Storms Forecast Center, Kansas City, MO 64106

(Manuscript received 19 May 1980, in final form 13 April 1981)

ABSTRACT

A multiple discriminant analysis program is used to obtain prediction functions for general and severe thunderstorm activity during April and July. Four predictors, lifted index, mean low-level mixing ratio, K index and mean 200–300 mb divergence, are statistically combined for an area east of the Rockies. Forecasts are valid for a 12 h period. Tests on dependent and independent data using a probability of detection (POD), a false alarm rate (FAR), and a critical success index show stability for the prediction functions. Operational use of the prediction functions is examined from two approaches, one involving POD and FAR and the other involving a more conventional probabilty approach. Semi-operational results from April and July 1979 evaluations show statistical successes comparable to the dependent data results.

1. Introduction

Statistical prediction has been used in thunderstorm forecasting for many years. Tillotson (1951) used a statistical approach to develop a thunderstorm forecasting scheme for Denver for the month of September. In more recent years David (1973) applied screening regression to numerical model output and 0600 GMT surface data to develop a 12–36 h prediction. This product is routinely produced at the National Severe Storms Forecast Center (NSSFC) in Kansas City, Missouri.

Many years of work on Model Output Statistics (MOS) thunderstorm outlook-guidance have led to the current operational National Weather Service product described by Reap and Foster (1979). This outlook guidance is produced at the National Meteorological Center once a day from 0000 GMT model output for a period 12–36 h after model initialization time. It is disseminated via both teletype and facsimile.

Both the David and Reap/Foster outlook products provide guidance to the NSSFC Severe Local Storms Unit (SELS) for preparation of the 0830 GMT Convective Outlook (AC) which delineates areas of thunderstorm and severe thunderstorm potential until 1200 GMT the following day. There is no additional AC guidance provided to NSSFC for updating the outlook at 1500 and 1930 GMT. A 2–6 h forecast of severe local storm probability, whose "primary intent . . . is to aid forecasters at the NSSFC in the issuance of tornado and severe thunderstorm watches" (Charba, 1979), combines LFM data (6, 9, 12 and 18 h forecasts from the 1200 GMT run) with surface and radar observations. This product is transmitted every 3 h starting around 1530 GMT. Thus, although there is short-term watch guidance at hand, updated outlook guidance based on the new 1200 GMT upper air data is not available. The scheme described here was devised to fill this gap.

With the most recently observed radiosonde data as input, thunderstorm and severe thunderstorm outlook guidance for 0–12 h after observation time is produced. It is realized that the value of observed data for prediction purposes decreases with time away from observation time, particularly beyond 8–12 h. Nevertheless, the results demonstrate skill in the 0–12 h period and thus show the feasibility of using this approach on an operational level.

The intent of the forecast is to differentiate between areas of severe convection, non-severe convection and no convection. Multiple discriminant analysis (MDA) is a statistical tool apropos to this problem. The procedure generates one or more discriminant functions whose values distinguish between two or more groups (or types) of phenomena. In this investigation, two and three group discriminations are examined in an attempt to distinguish thunderstorm occurrence from non-thunderstorm occurrence, severe thunderstorm situations from non-severe thunderstorm situations, and severe from non-severe from no thunderstorm occurrence. MDA combines a specified number of prediction parameters in equation form (the discriminant function). It should be realized that two group MDA reduces to multiple linear regression

[1] Present affiliation: National Weather Service Forecast Office, Topeka, KS 66616.

(Glahn, 1978; Randerson, 1977). For details of the MDA process, the reader is referred to Miller (1962).

2. The data

In the work described here, four parameters physically important to thunderstorm occurrence were chosen. These parameters are considered by SELS forecasters as the most basic and most reliable parameters needed to construct a convective outlook. The parameters are the mean 200–300 mb divergence (DIV) (McNulty, 1978); the SELS lifted index (LI) (Galway, 1956); the mean mixing ratio in the lowest 100 mb (R); and the K index (George, 1960). These variables represent three of the four major ingredients needed for synoptic-scale thunderstorm guidance preparation [see description by McNulty (1978)].

These four parameters are combined into all possible sets of three or four variables. Data for each set were then used as the input to the MDA procedure. The resulting discriminant functions were tested to see what criteria maximized the critical success index (CSI) for that function [CSI, as described by Donaldson *et al.* (1975), also has been termed "threat score" by Palmer and Allen (1949), and "ratio of verification" by Gilbert (1884)].

Two dependent data sets were based on April and July 1977 data. A predictand sample was constructed from values of the four parameters noted above which occurred every 12 h at the nine starred rawinsonde stations in Fig. 1 and at the locations of tornadoes listed in the SELS log. An indication of the occurrence of thunderstorms (severe or nonsevere) was also included. In keeping with the synoptic scale nature of the prediction technique, thunderstorms were considered to have occurred if they were within a radius of one latitude degree of a rawinsonde station during a period extending from one hour before to 12 h after observation time. These circles were chosen such that the number of surface observation stations within the circle was maximized while keeping the rawinsonde station near the center of the area. Verification data were obtained from surface hourly observations and from the SELS log. Two independent data sets were constructed in a similar fashion from April and July 1978 data.

For thunderstorm versus no thunderstorm occurrence, the following discriminant functions maximized the CSI:

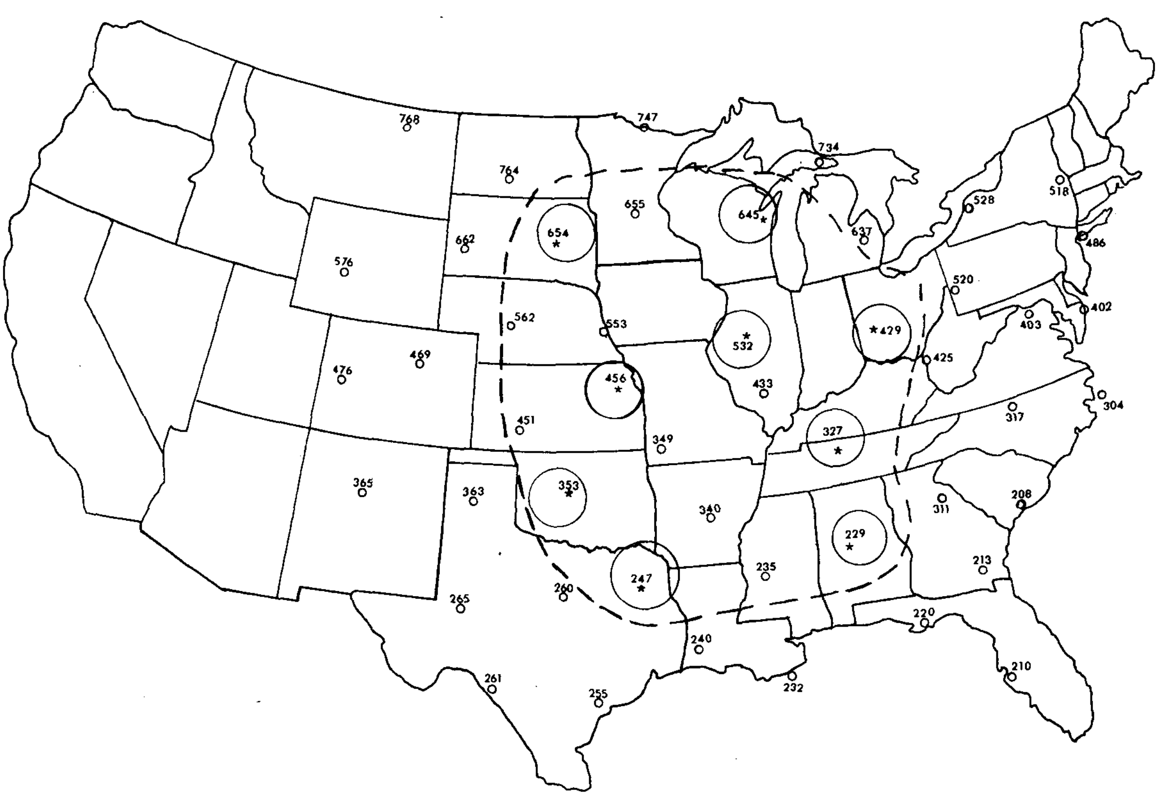

FIG. 1. Stations used in development sample (stars) and in April/July 1979 evaluation. Dashed line represents restricted evaluation area (see text).

TABLE 1. Values of POD, FAR and CSI for the dependent and independent data sets for the various prediction functions defined in the text.

	T1 ≥ 6	Y1 ≥ 1	J1 ≥ 2	Z1 ≤ 260
Dependent sample				
POD	0.876	0.935	0.875	0.787
FAR	0.349	0.323	0.309	0.492
CSI	0.596	0.647	0.629	0.447
Independent sample				
POD	0.720	0.767	0.802	0.808
FAR	0.346	0.403	0.280	0.450
CSI	0.522	0.506	0.611	0.486

April: (570 data point sample)

$$T1 = DIV - 0.0604\ LI + 0.741\ R + 0.0352\ K.$$

July: (567 data point sample)

$$J1 = DIV - 0.296\ LI + 0.0772\ K.$$

For severe thunderstorm situations, two group MDA (thunder versus severe thunder) gave higher CSI values than three group MDA (no thunder, thunder, severe thunder), resulting in the following discriminant functions:

April: (177 data point sample)

$$Y1 = DIV - 0.443\ LI + 0.134\ R - 0.0218\ K.$$

July: (312 data point sample)

$$Z1 = DIV + 18.1\ LI + 17.4\ R + 2.11\ K.$$

Although it is known that CSI has a strong dependence on the relative frequency of the event within the sample, comparative evaluation of several functions via CSI is valid if the same sample is used with each function (as is done here).

3. General thunderstorm results

For the first set of equations in Section 2, a spectrum of probability of detection (POD), false alarm rate (FAR) and CSI values was calculated for categorical forecasts based on various threshold values of T1 and J1. For T1 ≥ 6 the CSI was maximized at 0.596 while for J1 ≥ 2, the CSI attained a value of 0.629. Values of POD, FAR and CSI for both the dependent and independent data sets are given in Table 1. Some decrease in POD and CSI is expected for the independent sample (using dependent sample criterion at best CSI), but the small changes show some stability in the forecast functions.

Values of T1 or J1 can be used in two ways. First, T1 or J1 values can be thought of in terms of POD, FAR and CSI values. Figs. 2a and 2b show curves of POD, FAR and CSI based on the combined dependent/independent sample. The value of T1 or J1 used to make a categorical forecast can be chosen on the basis of office forecast philosophy. For example, if most of the activity is to be covered, a high POD may be desired. T1 ≥ 5 gives a POD of ~90% with a FAR ~44%. If the forecast philosophy calls for a lower FAR, say, 30%, T1 ≥ 7.3 should be used with a corresponding decrease in the POD. These curves provide a degree of flexibility in the use of the forecast functions.

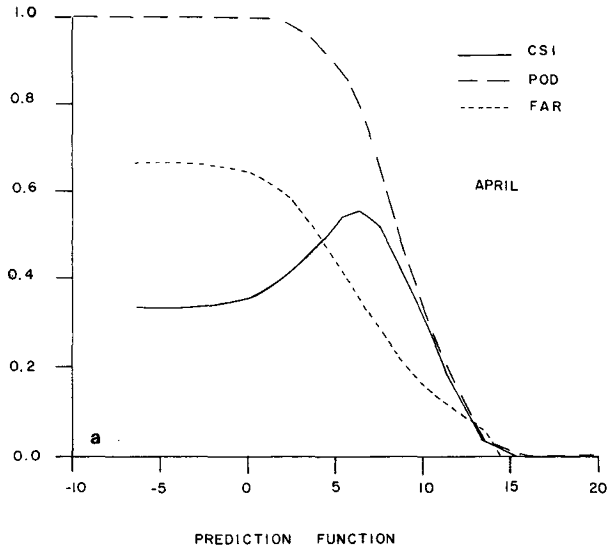

FIG. 2a. Values of POD, FAR, and CSI for the combined dependent/independent sample for April plotted as a function of T1.

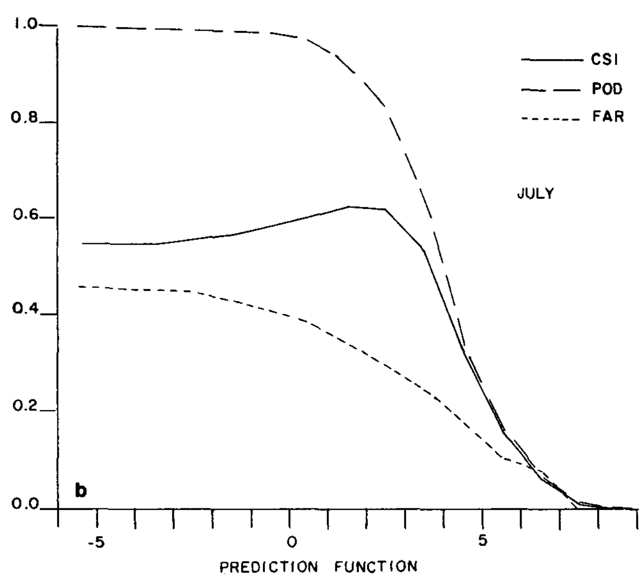

FIG. 2b. As in Fig. 2a except during July as a function of J1.

TABLE 2. Probability values derived from normal distributions fit to the thunderstorm data; T1, April; J1, July.

Function Value	Prob (T1)	Function Value	Prob (J1)
15.5	0.979	5.5	0.779
14.5	0.969	4.5	0.786
13.5	0.955	3.5	0.745
12.5	0.933	2.5	0.640
11.5	0.898	1.5	0.453
10.5	0.843	0.5	0.227
9.5	0.762	−0.5	0.074
8.5	0.650	−1.5	0.016
7.5	0.510		
6.5	0.362		
5.5	0.231		
4.5	0.134		
3.5	0.071		
2.5	0.036		

A second approach to T1 or J1 is to convert the value to the probability of a thunderstorm event occurring (not to be confused with a percent of areal coverage). To obtain this value, normal distribution curves were fit to the distributions of T1 and J1 for thunderstorm and no thunderstorm occurrence (dependent sample). From this normal fit for thunderstorm occurrence, an expected number of thunderstorms can be calculated for each value of T1 or J1. Similarly, from a normal fit to the no thunderstorm occurrence data, an expected number of nonthunderstorms can be found. The probability of thunder is then the ratio of the expected number of thunderstorms to the total number of expected events. Table 2 gives normal probability values associated with the prediction function. The probabilities of thunderstorm occurrence provide a more conventional forecast parameter than that of T1 or J1.

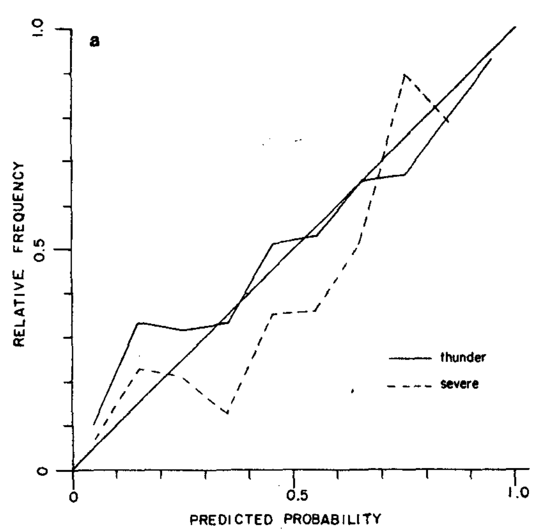

FIG. 3a. Relative frequency of occurrence of general (solid) and severe (dashed) thunderstorms for predicted probability categories for April independent data.

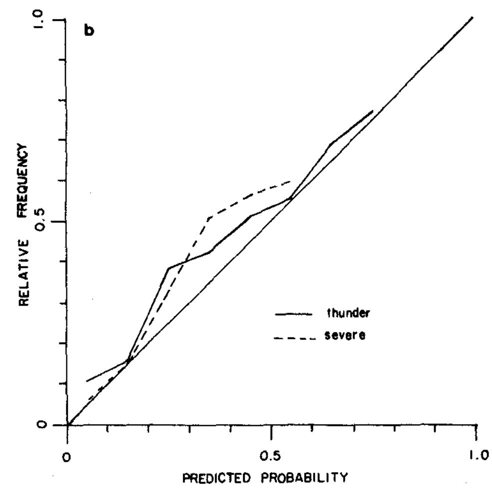

FIG. 3b. As in Fig. 3a except for July independent data.

To evaluate the probabilities so produced, reliability diagrams, henceforth called Murphygrams (Figs. 3a and 3b), were produced from the independent data sample. The increasing relative frequency of thunder (solid line) with increasing predicted probability shows that the probability functions used are fairly reliable despite the tendency to underforecast at lower probability values during April and for the entire sample during July.

4. Severe thunderstorm results

From the second set of equations in Section 2, CSI values were maximum for categorical forecasts of severe thunderstorms, given thunder occurs, for Y1 ≥ 1 and Z1 ≤ 260. Dependent and independent sample values shown in Table 1 are encouraging

TABLE 3. Probability values derived from normal distributions fit to the severe thunderstorm (given thunder occurs) data; Y1, April; Z1, July.

Function Value	Prob (Y1)	Function Value	Prob (Z1)
9.5	0.898	100	0.976
8.5	0.897	120	0.948
7.5	0.887	140	0.898
6.5	0.871	160	0.819
5.5	0.839	180	0.710
4.5	0.789	200	0.585
3.5	0.711	220	0.463
2.5	0.598	240	0.359
1.5	0.454	260	0.279
0.5	0.300	280	0.221
−0.5	0.168	300	0.180
−1.5	0.081		

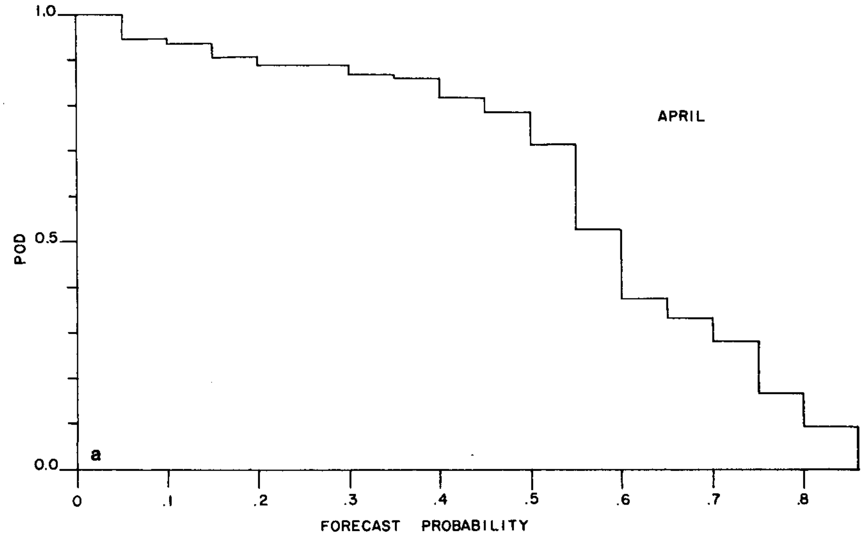

FIG. 4a. Severe thunderstorm POD estimates based on April 1979 probability forecasts.

but overall, not quite as good as those for general thunderstorm prediction.

The best way to use Y1 and Z1 is to convert these values to a probability of occurrence of severe thunderstorms. As before, normal curves are fit to the Y1 and Z1 distributions of severe and non-severe thunderstorm occurrence. The curves are then used to calculate the conditional probability of severe thunderstorms, given that thunder occurs. The probability of severe thunderstorms is then obtained from the product of the conditional probability of severe thunderstorms and the probability of thunderstorm occurrence. These probabilities as a function of Y1 and Z1 are given in Table 3.

The dashed lines on the Murphygrams (Figs. 3a and 3b) show the reliability of the severe thunderstorm probabilities. Results show a tendency to overforecast during April and underforecast during July.

5. Operational evaluation

Use of the discriminant functions given in Section 2 and conversion to probabilities via normal distribution curves are conducive to computer processing. Values of T1, J1, Y1 and Z1, as well as probability of general and severe thunderstorms, are calculated at NSSFC for rawinsonde stations and displayed in map-plot format for forecaster analysis. These values provide the forecaster with guidance for preparation of a convective outlook product for the next 12 h.

This scheme was tested on a semi-operational basis during April and July 1979. Each day 1200 GMT rawinsonde data were used to evaluate the discriminant functions. It was assumed that values calculated are representative of a continuous function over the map so that isolines of T1/J1 and probability could be drawn.

For the April data, an estimate of the POD was made by calculating the ratio of surface observation stations (within the T1 = 6 isoline) reporting thunderstorms or lightning to all stations reporting the same.

For the entire area shown in Fig. 1, the POD estimate was 0.751, a number close to the independent sample POD given in Table 1 (0.720). If the area of interest is restricted to a region about only the nine stations used to determine the discriminant functions (dashed line on Fig. 1), the POD estimate increases to 0.781. Thus, while it is realized

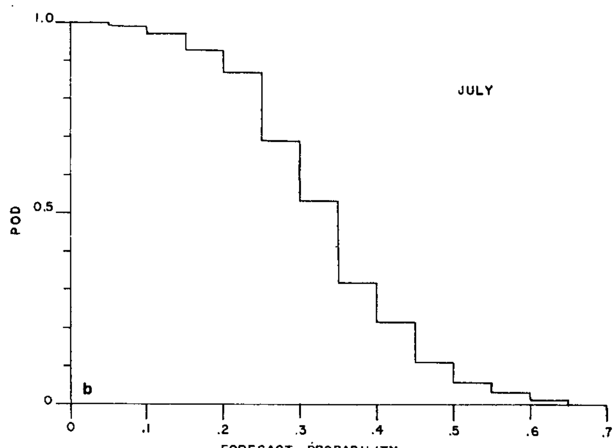

FIG. 4b. As in Fig. 4a except for July 1979.

that extension of the discrimination function to stations not in the original sample introduces potential sources of error, the results justify this extension. Further, if only cases of extensive thunderstorm occurrence are examined (30 or more stations reporting thunder), over 80% of the thunder reports are in the area of T1 ≥ 6. Based on these tests it is concluded that the map-analysis presenta-potential sources of error, the results justify this extension. Further, if only cases of extensive thunderstorm occurrence are examined (30 or more stations reporting thunder), over 80% of the thunder reports are in the area of T1 ≥ 6. Based on these tests it is concluded that the map-analysis presenta-tion is comparable to the statistics based on the dependent/independent sample results.

From the forecast probability of severe thunderstorms associated with each severe event which occurred, an estimate of the POD can be made by taking the ratio of the number of severe events with forecast probabilities greater than a specified value to the total number of severe events. Distributions of POD as a function of forecast probability are shown in Figs. 4a and 4b for April and July, respectively. POD's of 80% can be attained by using a categorical criterion of about 0.4 for April and 0.2 for July. A value nearer 90% requires a criterion of 0.15 for both months.

Fig. 5 shows an analysis of a sample thunderstorm prediction function for 1200 GMT 8 April 1979. Verifying thunderstorm observations are shown as carets. Fig. 6 is a sample severe thunderstorm probability forecast. Severe storms [i.e., tornadoes (triangles); wind storms (boxes), gusts ≥ 50 kt, (26 m s^{-1}); hailstorms (circles), hailstones ≥ ¾ inches, (~2 cm)] occurring during the forecast period are also shown.

6. Summary

The use of MDA to develop a short-term outlook guidance product has been demonstrated. Such a product fills a gap in objective guidance not currently available at the operational level. The approach gives results which show a fair degree of accuracy.

While the limitations of using observations static in time are realized, the results demonstrate their usefulness in short period forecasting. Several shortcomings of this approach should be noted, specifically, the tendency for activity to drift downstream from areas of maximum probability. This effect is due to the static nature of the predictors. Stations over south Texas, outside the area of the dependent data set, tend to have probabilities that are too high. This is associated with the more unstable values of LI and higher values of R that are more prevalent there than over the development area.

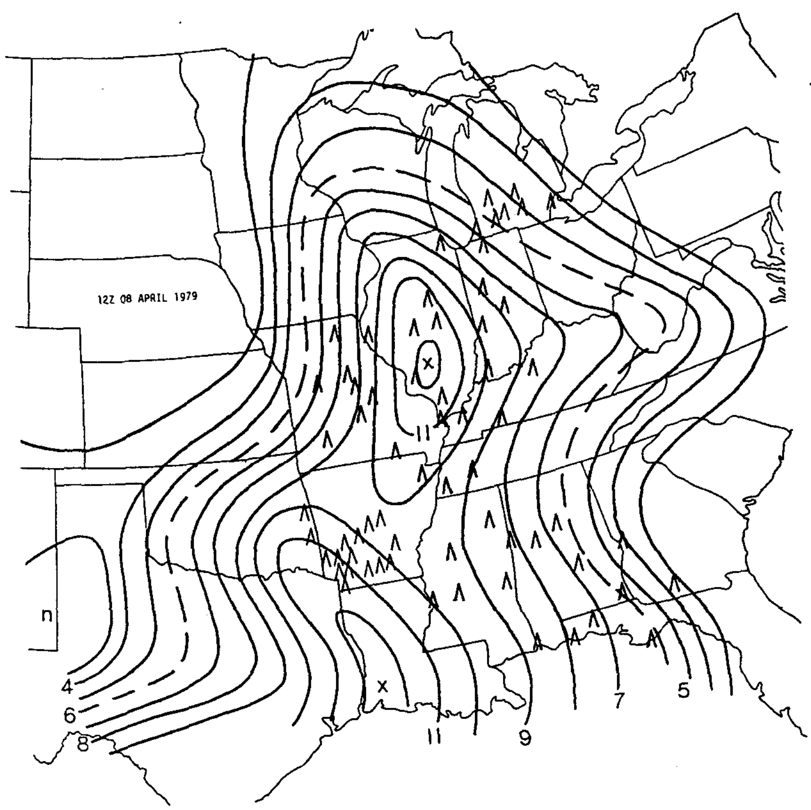

FIG. 5. Thunderstorm prediction function (T1) analysis for 1200 GMT 9 April 1979. Carets indicate verifying observation locations.

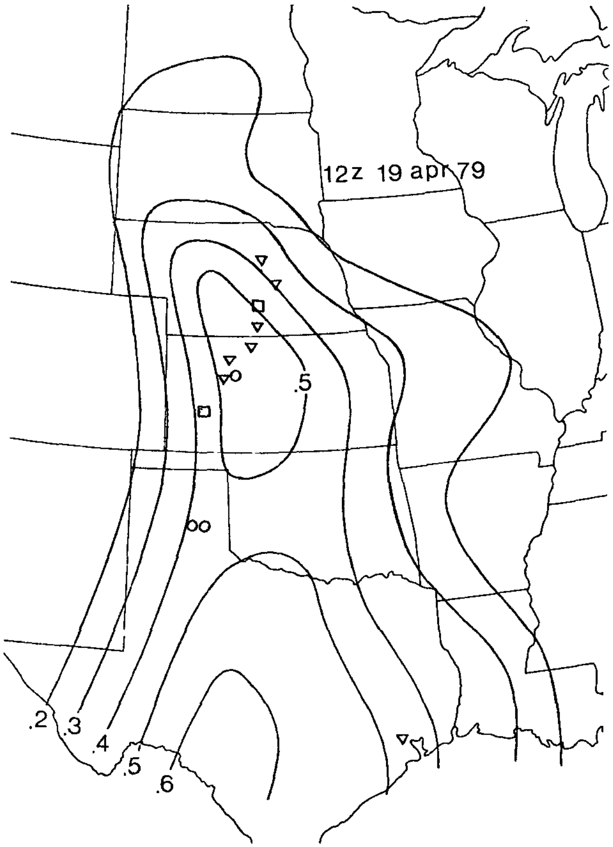

FIG. 6. Severe thunderstorm probability analysis for 1200 GMT 19 April 1979.

Further work along this line will redevelop the discriminant functions using a verification period of 3 to 12 h. This will allow a 3 h period for operational processing of the data before the valid period of the forecast.

By stratifying the data into monthly periods, month-to-month changes in forecast variables can be taken into account. A set of monthly equations covering the major portion of the severe weather season (March–August) should provide a degree of improvement over the seasonal approach to severe thunderstorm prediction that is currently used.

Acknowledgments. Thanks go to Dr. J. T. Schaefer for his review and recommendations on this paper. Special thanks go to Beverly D. Lambert for typing the manuscript.

REFERENCES

Charba, J. P., 1979: Two to six hour severe local storm probabilities: An operational forecasting system. *Mon. Wea. Rev.*, **107**, 268–282.

David, C. L., 1973: An objective method for estimating the probability of severe thunderstorms using predictions from the NMC (PE) numerical prediction model and from observed surface data. *Preprints 8th Conf. Severe Local Storms*, Denver, Amer. Meteor. Soc., 223–225.

Donaldson, R. J., R. M. Dyer and M. J. Kraus, 1975: An objective evaluator of techniques for predicting severe weather events. *Preprints 9th Conf. Severe Local Storms*, Norman, Amer. Meteor. Soc., 321–326.

Galway, J. G., 1956: The lifted index as a predictor of latent instability. *Bull. Amer. Meteor. Soc.*, **37**, 528–529.

George, J. J., 1960: *Weather Forecasting for Aeronautics*. Academic Press, 673 pp. (see pp. 411–415).

Gilbert, G. K., 1884: Finley's tornado predictions. *Amer. Meteor. J.*, **1**, 166–172.

Glahn, H. R., 1978: Comments on "Determining the relative frequency of occurrence of local cumulonimbus activity through discriminant analysis." *Mon. Wea. Rev.*, **106**, 265–267.

McNulty, R. P., 1978: On upper tropospheric kinematics and severe weather occurrence. *Mon. Wea. Rev.*, **106**, 662–672.

Miller, R. G., 1962: Statistical prediction by discriminant analysis. *Meteor. Monogr.*, No. 25, Amer. Meteor. Soc., 54 pp.

Palmer, W. C., and R. A. Allen, 1949: Note on accuracy of forecasts concerning the rain problem. U.S. Weather Bureau (unpublished).

Randerson, D., 1977: Determining the relative frequency of occurrence of local cumulonimbus activity through discriminant analysis. *Mon. Wea. Rev.*, **105**, 709–712.

Reap, R. M., and D. S. Foster, 1979: Automated 12–36 hour probability forecasts of thunderstorms and severe local storms. *J. Appl. Meteor.*, **18**, 1304–1315.

Tillotson, K. C., 1951: An objective aid for forecasting September thunderstorms at Denver, Colo. *Mon. Wea. Rev.*, **79**, 27–34.