NOAA Technical Memorandum NWS TDL-65

OPERATIONAL SYSTEM FOR PREDICTING SEVERE LOCAL STORMS TWO TO SIX HOURS IN ADVANCE

Silver Spring, Md.
May 1977

noaa NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / National Weather Service

NOAA TECHNICAL MEMORANDA

National Weather Service, Techniques Development Laboratory Series

The primary purpose of the Techniques Development Laboratory of the Office of Systems Development is to translate increases of basic knowledge in meteorology and allied disciplines into improved operating techniques and procedures. To achieve this goal, the laboratory conducts applied research and development aimed at the improvement of diagnostic and prognostic methods for producing weather information. The laboratory performs studies both for the general improvement of prediction methodology used in the National Meteorological Service and for the more effective utilization of weather forecasts by the ultimate user.

NOAA Technical Memoranda in the National Weather Service Techniques Development Laboratory series facilitate rapid distribution of material that may be preliminary in nature and which may be published formally elsewhere at a later date. Publications 1 through 5 are in the former series, Weather Bureau Technical Notes (TN), Techniques Development Laboratory (TDL) Reports; publications 6 through 36 are in the former series, ESSA Technical Memoranda, Weather Bureau Technical Memoranda (WBTM). Beginning with TDL 37, publications are now part of the series NOAA Technical Memoranda, National Weather Service (NWS).

Publications listed below are available from the National Technical Information Service (NTIS), U.S. Department of Commerce, Sills Bldg., 5285 Port Royal Road, Springfield, Va. 22151. Prices vary for paper copy; $1.45 microfiche. Order by accession number, when given, in parentheses at end of each entry.

ESSA Technical Memoranda

WBTM TDL 9 A Survey of Studies of Aerological Network Requirements. M. A. Alaka, June 1967. (PB-174-984)

WBTM TDL 10 Objective Determination of Sea Level Pressure From Upper Level Heights. William Klein, Frank Lewis, and John Stackpole, May 1967. (PB-179-949)

WBTM TDL 11 Short Range, Subsynoptic Surface Weather Prediction. H. R. Glahn and D. A. Lowry, July 1967. (PB-175-772)

WBTM TDL 12 Charts Giving Station Precipitation in the Plateau States From 700-Mb. Lows During Winter. Donald L. Jorgensen, August F. Korte, and James A. Bunce, Jr., October 1967. (PB-176-742)

WBTM TDL 13 Interim Report on Sea and Swell Forecasting. N. A. Pore and W. S. Richardson, December 1967. (PB-177-038)

WBTM TDL 14 Meteorological Analysis of 1964-65 ICAO Turbulence Data. DeVer Colson, October 1968. (PB-180-268)

WBTM TDL 15 Prediction of Temperature and Dew Point by Three-Dimensional Trajectories. Ronald M. Reap, October 1968. (PB-180-727)

WBTM TDL 16 Objective Visibility Forecasting Techniques Based on Surface and Tower Observations. Donald M. Gales, October 1968. (PB-180-479)

WBTM TDL 17 Second Interim Report on Sea and Swell Forecasting. N. A. Pore and W. S. Richardson, January 1969. (PB-182-273)

WBTM TDL 18 Conditional Probabilities of Precipitation Amounts in the Conterminous United States. Donald L. Jorgensen, William H. Klein, and Charles F. Roberts, March 1969. (PB-183-144)

WBTM TDL 19 An Operationally Oriented Small-Scale 500-Millibar Height Analysis Program. Harry R. Glahn and George W. Hollenbaugh, March 1969. (PB-184-111)

WBTM TDL 20 A Comparison of Two Methods of Reducing Truncation Error. Robert J. Bermowitz, May 1969. (PB-184-741)

WBTM TDL 21 Automatic Decoding of Hourly Weather Reports. George W. Hollenbaugh, Harry R. Glahn, and Dale A. Lowry, July 1969. (PB-185-806)

WBTM TDL 22 An Operationally Oriented Objective Analysis Program. Harry R. Glahn, George W. Hollenbaugh, and Dale A. Lowry, July 1969. (PB-186-129)

WBTM TDL 23 An Operational Subsynoptic Advection Model. Harry R. Glahn, Dale A. Lowry, and George W. Hollenbaugh, July 1969. (PB-186-389)

WBTM TDL 24 A Lake Erie Storm Surge Forecasting Technique. William S. Richardson and N. Arthur Pore, August 1969. (PB-185-778)

WBTM TDL 25 Charts Giving Station Precipitation in the Plateau States From 850- and 500-Millibar Lows During Winter. August F. Korte, Donald L. Jorgensen, and William H. Klein, September 1969. (PB-187-476)

WBTM TDL 26 Computer Forecasts of Maximum and Minimum Surface Temperatures. William H. Klein, Frank Lewis, and George P. Casely, October 1969. (PB-189-105)

WBTM TDL 27 An Operational Method for Objectively Forecasting Probability of Precipitation. Harry R. Glahn and Dale A. Lowry, October 1969. (PB-188-660)

WBTM TDL 28 Techniques for Forecasting Low Water Occurrences at Baltimore and Norfolk. James M. McClelland, March 1970. (PB-191-744)

WBTM TDL 29 A Method for Predicting Surface Winds. Harry R. Glahn, March 1970. (PB-191-745)

WBTM TDL 30 Summary of Selected Reference Material on the Oceanographic Phenomena of Tides, Storm Surges, Waves, and Breakers. N. Arthur Pore, May 1970. (PB-192-449)

WBTM TDL 31 Persistence of Precipitation at 108 Cities in the Conterminous United States. Donald L. Jorgensen and William H. Klein, May 1970. (PB-193-599)

WBTM TDL 32 Computer-Produced Worded Forecasts. Harry R. Glahn, June 1970. (PB-194-262)

(Continued on inside back cover)

NOAA Technical Memorandum NWS TDL-65

OPERATIONAL SYSTEM FOR PREDICTING SEVERE LOCAL STORMS TWO TO SIX HOURS IN ADVANCE

Jerome P. Charba

Techniques Development Laboratory
Silver Spring, Md.
May 1977

UNITED STATES
DEPARTMENT OF COMMERCE
Juanita M. Kreps, Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION
Robert M. White, Administrator

National Weather
Service
George P. Cressman, Director

CONTENTS

OPERATIONAL SYSTEM FOR PREDICTING SEVERE LOCAL STORMS TWO TO SIX HOURS IN ADVANCE

Jerome P. Charba
Techniques Development Laboratory
National Weather Service, NOAA
Silver Spring, MD 20910

ABSTRACT. We have developed and implemented in the National Weather Service (NWS) an objective method yielding probabilities of tornadoes, large hail, and damaging surface winds 2-6 hr in advance. The probabilities are for square areas about 90 n.mi. on a side, covering most of the United States east of the Rocky Mountains.

The probabilities are produced by multiple linear regression equations. The independent variables or predictors in these equations are derived from routinely observed surface atmospheric variables, manually-digitized radar data, localized climatic frequencies of the predictand, and large-scale numerical model output. Predictor quantities computed from these data are optimally positioned relative to the predictand areas and "linearized" in order to enhance their correlations with the predictand.

Verification of 74 days of operational spring season forecasts showed the probabilities have skill relative to climatic frequency and persistence of severe storms. Subjective verification of five severe storm cases, quasi-objectively selected, showed that envelopes of the 15% and 25% probability isoplets matched up best with reported severe weather. These threshold probabilities, arrived at subjectively, agreed with the threshold probability found to produce the best Threat score. When the probability forecasts were subjectively compared against official NWS tornado and severe thunderstorm watches for individual cases, their accuracies appeared similar to those of the latter.

1. INTRODUCTION

Since 1952 the National Weather Service's (NWS's) National Severe Storm Forecast Center (NSSFC) has been issuing tornado and severe thunderstorm "watch" forecasts. Watches are issued for rectangular areas typically covering about 75,000 km^2; their valid periods generally fall within the range from 0 to 8 hr into the future. Issuance of a watch for tornadoes or severe thunderstorms accompanied by large hail and/or damaging surface winds means that one or more of these phenomena are likely to occur within the area delineated. Galway (1975) has presented statistics on the performance of the watches in capturing tornado-related deaths.

Over the years, NSSFC forecasters have issued watches largely on the basis of a subjective assessment of atmospheric conditions related to severe storms. A set of predictive parameters or indices has gradually emerged from postmortem case studies and operational forecasting experience. Most predictive parameters currently in use are contained in a report by Miller (1972). One problem with the forecasting parameters being used is that their order of importance has not been well established. Another problem is that, in operational practice, some of the parameters are not evaluated quantitatively; rather, they are recognized subjectively by careful examination of conventionally analyzed charts. This means that issuance of watches is still based largely on the skill and experience of individual line forecasters. However, perhaps this situation may change gradually due to the recent introduction of the objective forecasting technique discussed in this report.*

In early 1972, the Techniques Development Laboratory (TDL) undertook a limited effort to develop an objective method of producing short range probabilities of tornadoes and severe local storms. The goal was to provide severe weather forecasters with objectively produced guidance forecasts that would aid them in the preparation of watches issued to the public. Our initial approach was to exploit the predictive value of surface data observed hourly, since these data are timely and abundant (Charba and Livingston, 1973). We further recognized that forecasters had been making extensive use of hand-analyzed charts based upon these data and that objective methods of analysis and extraction of useful predictive information could possibly relieve them of much tedious map analysis. Using as input basic observed surface meteorological variables plus forecasts of 500-mb temperature from the Limited-Area Fine-Mesh model (LFM) (Howcroft and Desmarais, 1971), we developed a multiple regression equation yielding 2-6 hr probabilities of tornadoes and severe local storms. Probability forecasts for 2300-0300 GMT were initially transmitted to NSSFC in real time during the spring and summer of 1974. As a result of an encouraging response to this product, new equations were derived, and the program was expanded in 1975 to provide two additional forecasts, one for 1700-2100 GMT and another for 2000-0000 GMT (Charba, 1975). The predictor input remained restricted, however, to observed surface data and 500 mb temperature forecasts.

For use during the 1976 season, we developed an improved and expanded prediction model which we refer to as the second generation version. New predictor variables based upon much expanded data sources were incorporated

*NSSFC also issues a 12-36 hr outlook of thunderstorms and severe local storms. Objectively derived, probability forecasts have been available to aid the forecasters in preparing this product for several years (see David, 1973; Reap and Foster, 1975).

along with several refinements designed to improve the correlations between the predictors and predictand. During 1976, probabilities were produced for the same three periods as those of 1975 and transmitted by teletype bulletin. This article focuses on the development of the prediction equations used in this new model and the performance of the operational forecasts run during the 1976 spring season. We shall discuss the performance of the forecasts by means of skill scores as well as through a subjective examination of individual severe storm cases. We shall use the official watches issued to the public as a standard upon which to evaluate the performance of the predicted probabilities in the individual cases.

2. DEVELOPMENT OF PREDICTION EQUATIONS

The forecasting method we've used is based upon a hybrid combination of the classical statistical (Klein 1970) and model output statistics (MOS) (Glahn and Lowry 1972) approaches to objective weather forecasting. The predictand was defined as the occurrence or nonoccurrence of one or more tornadoes, hail $\geq$ 3/4 in. dia., or surface wind gusts in excess of 50 kt in square boxes approximately 90 n.mi. on a side--an area which is about one-third that of an average sized watch. Reports of severe weather were taken from NSSFC's edited severe weather data tapes. The predictand took on a value of 1 for an occurrence and 0 otherwise. The configuration of a single predictand box in relation to the computational grid is illustrated in figure 1. It should be pointed out that a predictand box was defined for each of the + marks within the irregular area outlined in figure 1. Since the distance between these grid points is exactly one-half the length of a side of the predictand box, there is considerable overlapping of predictand areas. This overlapping is the price paid for the choice of retaining the high horizontal resolution in many of the predictor fields (and, therefore, in the forecast probability fields) and, at the same time, insuring that the predictand frequencies don't become very small and discontinuous spatially. Small frequencies, in general, result in low peak forecast probabilities which, as a consequence, could undermine credibility in the product. In the case of severe weather, some of the spatial variability in frequency of occurrence is due to nonuniform reporting of severe storm events. Therefore, the overlapping is a form of smoothing which has the effect of improving the correlation between predictors and predictand.

Separate predictor/predictand data samples and separate prediction equations were derived for the Gulf Coast and non-Gulf Coast regions shown in figure 1. We delineated the Gulf region to encompass that part of the grid where the nearby presence of the Gulf of Mexico has a persistent and dominant control over airmass properties, particularly near the surface. Further geographical data stratification was not considered feasible since only two spring seasons of data (March 15 to June 15 of 1974 and 1975) were available. The samples were formed by combining 132 points in the Gulf region and 426 points in the non-Gulf region. Within a region, a single prediction equation was derived from the pooled sample; real-time forecasts at all points within a region are based upon this equation.

Potential predictors were developed from four data sources, namely, basic surface variables observed hourly, forecasts of basic upper air variables, manually digitized radar (MDR) data, and the predictand relative frequency of occurrence. The MDR data available operationally in the NWS are integers ranging from 1 to 9 that describe the intensity and coverage of radar echoes within squares 40-45 n.mi. on a side (see Moore et al. 1974). The valid times of the various data relative to a predictand period are illustrated in figure 2 for the 2000-0000 GMT predictand. Hereafter, we shall refer to the forecasts produced by the data inputs and valid times shown in figure 2 as the 1800-GMT forecasts since the probabilities for 2000-0000 GMT are heavily dependent on observed surface data for 1800 GMT. Correspondingly, the 1500-GMT forecasts are for 1700-2100 GMT and the 2100-GMT forecasts are for 2300-0300 GMT. Therefore, for each of the three forecast cycles, the valid times of the input data relative to the <u>beginning time</u> of the predictand are: MDR data approximately 2½ hr earlier, observed surface data 5 and 2 hr earlier, and LFM forecasts 1 hr later; the time of predictand frequency naturally corresponds to that of the predictand.

All potential predictors were evaluated at the grid points shown in figure 1. Therefore, we first had to interpolate the observed surface and LFM forecast variables to these points. In the case of the observed surface data (specifically, MSL pressure, and surface temperature, dewpoint, and wind), this interpolation was performed with the Cressman (1959) objective analysis method complete with data error checking. We used a biquadratic method to interpolate LFM grid point values to our grid.

The potential predictor variables derived from the grid point data are listed in table 1 along with the data source(s) from which each variable is obtained. Many of these variables are commonly used in thunderstorm forecasting, (e.g., see George 1960 and Miller 1972) and no discussion of them is necessary. However, some of the variables, the stability indices in particular, require elaboration. For instance, the K-index (variable No. 29), is normally defined as:

$$K = T_{850} + T_{d,850} - (T_{700} - T_{d,700}) - T_{500}, \quad (1)$$

where T is the temperature and T_d the dewpoint temperature at the pressure surfaces indicated by the subscripts. The modified K-index, K_{mod} (No. 30), is defined as:

$$K_{mod} = \overline{T} + \overline{T}_d - (T_{700} - T_{d,700}) - T_{500}, \quad (2)$$

where $\overline{T} = \frac{1}{2}(T_{sfc} + T_{850})$ and $\overline{T}_d = \frac{1}{2}(T_{d,sfc} + T_{d,850})$. The Total Totals index, TT (No. 31), is conventionally defined as:

$$TT = T_{850} + T_{d,850} - 2T_{500}; \quad (3)$$

we replaced T_{850} by $\overline{T}$ and $T_{d,850}$ by $\overline{T}_d$ and thus defined the modified Total Totals index, TT_{mod} (No. 32). The Showalter index (Showalter 1953), SI (No. 33), is conventionally defined as:

$$SI = T_{500} - T^*, \tag{4}$$

where T* is the temperature of an air parcel lifted dry adiabatically from 850 mb to saturation, then moist adiabatically to 500 mb. In the definition of the modified Showalter index (No. 34), the parcel is lifted from the surface rather than from 850 mb. In the case of each of these stability indices, and in other derived predictor quantities where surface variables are involved in the computation, the surface variables are based upon observed surface data. On the other hand, whenever a variable above the surface is involved it is based upon an LFM forecast which, incidentally, is valid 3 hr later than the observed data (see fig. 2). Note that the modification to these indices lies in the incorporation of the latest observed surface temperature and dewpoint; so, in a sense, we are adjusting these indices for errors in LFM forecasts of temperature and moisture at 850 mb. As will be seen later, the modified indices are generally selected as predictors ahead of those conventionally defined.

Continuing the elaboration of perhaps unfamiliar variables in table 1, the sign associated with variable No. 37 indicates whether the wind veers (positive sign) or backs (negative sign) with height. Variable No. 42, the 3-hr MSL pressure change, is the only variable incorporating surface data observed 3 hr earlier. The MDR variables (No. 43) are defined later. The predictand relative frequency variable (No. 44) was computed for individual predictand boxes and for each day from 6 spring seasons (1970-1975) of predictand data. Two of these 6 seasons, 1974 and 1975, formed the sample used in the development of the equations. In the procedure used to obtain the daily varying frequencies, monthly frequencies were computed initially, then spatially smoothed, and finally interpolated timewise to the day. This procedure gave spatially smooth, daily varying frequencies as desired. The main purpose for incorporating the predictand frequency as a predictor is that it could account for variability in severe storm occurrence on scales smaller than can be determined from predictors based upon currently available observed or forecast data.

Each of the computed variables in table 1 was smoothed to remove wavelengths of 4 grid lengths (170 n.mi.) and less; a 5-point hanning filter (Shuman 1957) was used.

For various causes, some well understood and others not, a predictor variable may correlate best with a weather phenomenon when it is positioned at some point other than where the phenomenon is occurring. For example, a squall line is often found to the east of the thermal gradient and convergence zone characteristic of a cold front. The need for optimal positioning of predictors is especially important when these predictors are based upon observed data since their valid time is earlier than that of the predictand.

We devised an objective procedure for determining the best position for each potential predictor. The best position was defined as that point for which the predictor had the highest linear correlation with the predictand from among 30 points surrounding the predictand box. The positions of 10 of the more prominent variables are shown in figure 3. These positions or offsets, as we call them, were determined only from the 1800 GMT data sample. We assumed that corresponding offsets determined from data ± 3 hr from 1800 GMT would differ little, and therefore, we applied the 1800 GMT offsets to the 1500 and 2100 GMT predictors as well.

Because fields of radar echoes are spatially very discontinuous, we decided to define several potential predictors from the MDR data. An individual potential predictor was defined as the MDR code for a particular MDR box in the vicinity of the predictand box. To determine which boxes (predictors) would be most useful, we made a special screening regression run in which 30 candidates were offered. The first four MDR boxes selected (see fig. 4) were chosen as the potential predictors to be later screened in combination with the other variables in table 1.

The screening regression technique used to derive the prediction equations relates the predictand to a weighted linear combination of predictor variables. We know, however, that in general the relationships are actually nonlinear. In practice, some of the nonlinearity can be accounted for by converting the continuous variables into sets of binary variables and screening the latter in combination with the former as discussed by Glahn and Bocchieri (1976). However, this too has drawbacks; for instance, (1) one does not know, _a priori_, how to choose the binary limits, (2) the total number of potential predictors can become intractably large when the new binary variables are created, and (3) some of the information contained in a continuous variable is lost upon conversion into a discrete set of binary variables.

Our attempt to account for the nonlinearity of predictor/predictand relationships in the dependent sample involved transforming the predictor variables in such a way that the new variables are, in fact, linearly related to the predictand. We went about this by first plotting, on rectangular coordinate axes, the predictand relative frequencies for small intervals along the predictor axis. The data used were from the dependent sample. Sample plots which show considerable nonlinearity are shown in figure 5.* A functional curve was then fitted to these points. Actually, we fit the plotted data points exactly by connecting adjacent points with straight lines as shown in figure 5. Each predictor variable was then transformed or "linearized" by interpolating from these piecewise linear functions the predictand relative frequency (ordinate value) corresponding to each predictor value (abscissa value). The new variable to be screened in place of the original one is made up of the interpolated values. Among the variables in table 1, only the ones derived from observed surface data and LFM output were linearized. MDR variables were not linearized because of the small range and discreteness of the

---

*Considerable care was taken to insure that the plotted points well represent the relationship over the entire range of the predictor variable but yet that each point was based upon a sufficient number of cases so that it could be considered as approximately representative of the population.

MDR code values; the predictand climatic frequency used as a predictor was not linearized because it was found to be approximately linearly related to the predictand frequency in the dependent sample.

One possible pitfall in applying this linearizing technique lies in "over-fitting" the dependent sample. This is especially dangerous when the dependent sample is small, because predictor/predictand curves based upon such a sample would generally not match up well with curves based upon independent samples. Consequently, this could result in poor forecast performance on independent samples. In our work we have found at least two benefits in using linearized variables (as opposed to their "raw" continuous and binary forms) as predictors in regression equations. First, they extended the range of forecast probabilities and, second, they improved the forecast skill (if only slightly) on independent data.

In the development of the severe weather regression equations, all the variables in table 1 were screened together. However, we also derived "backup" equations for which MDR variables in table 1 were withheld as possible predictors. During the operational usage, the backup equations are needed to produce probabilities at all grid points where the primary equations cannot be used, i.e., points where MDR data are nonexistent or happen to be missing.

Table 2 gives the sample size, reduction of variance (RV), and the predictand relative frequency for each of the regression equations used operationally. Note that the samples used in the derivation of the primary equations were much smaller than those of the backup equations. This is because missing MDR predictors depleted the primary equation samples. For this reason it is not really meaningful to compare the RVs between the primary and backup equations.

The table shows the RVs were lower for the Gulf than the non-Gulf region for each of the three times. It also shows the RVs to be lowest at 1500 GMT for both regions and highest at 1800 GMT for the non-Gulf region. In general, the RVs are seen to vary in the same manner as the predictand frequency.

As an interesting sidenote, notice in table 2 that the predictand frequencies for the primary equation samples are generally slightly higher than those for the backup equations. This tends to verify a suspected bias in the primary equation sample brought about by the loss of cases that had missing MDR predictor data. For instance, it is known that, in most cases, MDR data are missing because NWS network radars were not operating. Radars are usually shut down when there are no significant echoes. Thus, the cases for which MDR data are available are biased towards stormy weather and this is reflected in our predictand samples.

Typical of all equations derived are those for 1800 GMT. Tables 3a and 3b list the predictors in their order of selection and the cumulative reduction of variance for the 1800 GMT primary equations for the non-Gulf and Gulf regions, respectively. Note that, although the increments are small after inclusion of the ninth predictor, the cumulative reduction of variance increases appreciably out to 15 terms in both equations. In fact, all the equations summarized in table 2 contain 14 to 16 terms. Of course, the only reliable method of determining the optimal number of terms is to test candidate equations with different numbers of terms on an independent sample. We performed such a test on an independent sample described in detail in the next section. A plot of a verification score, P, (discussed in detail in the next section) as a function of the number of predictors per equation for the 1800 GMT non-Gulf primary and backup equations is shown in figure 6. For the purpose of the present discussion, it will suffice to state that the lower the P-score the better the match between the probabilities and the observed predictands. The plot shows that the best scores were obtained with equations that had more than 14 terms. In applications of this regression method to other predictands, the maximum number of predictors has generally been around 10 or 12 (e.g., see Bocchieri and Glahn 1972). The greater number of contributing terms found here is likely a reflection of the broad assortment of data sources upon which the predictors were based. These different data sources likely contain independent information about occurrences of severe weather which would be reflected in the derived variables. Therefore, even similar variables, when based upon different data sources, contain independent information and thus add to the overall reduction of variance of the predictand.

It is interesting to note the relative number of times forecast and observed data appear in the predictors. Recall that the forecast data are from the LFM (e.g., a 9-hr forecast from 1200 GMT in the case of the 1800 GMT equations), while the data falling in the observed category includes surface observations, MDR data, and predictand frequency. We find that for the non-Gulf and Gulf equations in tables 3a and 3b the forecast and observed data are incorporated into about an equal number of predictors--about 10 each. However, when one considers only the first three predictors of both equations, surface observations appear to play a leading role. This finding confirms the expected importance of observed data in this short range prediction problem.

Several characteristic features can be discerned in the predictor variables contained in the equations in tables 3a and 3b. Note the importance of the modified stability indices, such as the Modified Total Totals and Showalter. Moisture divergence at the surface and equivalent potential temperature (Θ_E) advection at the surface and 850 mb also rank high. It should be pointed out that the inclusion of both the surface and 850-mb Θ_E advection as predictors implies differential advection. For instance, note the signs of the linear correlation coefficients of these variables from figure 3. We should also add that tests have shown these Θ_E advection variables, as well as the 500-mb vorticity advection, would not have been selected as predictors had they not been offset as discussed earlier.

It has recently been documented that operational forecasting of tornadoes and severe local storms is more difficult in the Gulf Coastal states of Arkansas, Mississippi, Alabama, etc. than in states to their north and west (Galway, 1975). Since our Gulf region (fig. 1) encompasses much of the problem area, it's useful to examine the equations to see whether there are important differences in the predictors between the Gulf and non-Gulf regions. Tables 3a and 3b show that, while the order of selection of a few variables is different, 12 of 15 predictors are common to the equations of both regions. Furthermore, of the first five predictors, three are common to both equations. The main differences seem to be: (1) the observed surface wind is more important in the non-Gulf equation (note the second and third predictors of both equations); and (2) the 500-mb vorticity advection appears in the non-Gulf equation but not in the Gulf equation. Such differences are expected, however, since severe storm outbreaks are usually associated with well organized extratropical cyclones and their associated jet streams, and these systems often lie to the north of the Gulf Coastal states, particularly in late spring.

3. VERIFICATION

These spring season prediction equations were implemented operationally at the National Meteorological Center (NMC) on the National Oceanic and Atmospheric Administration's (NOAA's) IBM 360/195 computer in mid-April 1976. Normally, most of the grid point probabilities produced would be based upon the primary equations with the remainder based upon the backup equations. Unfortunately, because of transmission difficulties, the MDR data were not available to the computer program in real time until the following June. Therefore, the backup equations were used exclusively for producing the operational forecasts for most of the spring season.

To test the regression equations on independent data, we recomputed the spring 1976 operational forecasts from archived data. These archived data were identical to the real-time data except that they included MDR data. In the recomputed sample of "operational" forecasts, about 73% of the individual probability values were computed from the primary equations and the rest from the backup equations. The sample of forecasts available for verification contained 74 days from the period March 31 to June 14, 1976. When one considers that the relative frequency of the predictand event was only 2-3%, it is unlikely that different independent samples of such size would produce closely similar verification statistics. Therefore, the statistics to be presented for the independent sample must be regarded as somewhat preliminary. Corresponding verification statistics from the dependent sample, which contained 127 days, will also be presented for comparison. Because of the preliminary nature of the independent sample verification statistics, we shall also devote considerable discussion to subjective evaluation of the performance of the forecasts for a selected set of individual severe storm cases.

Since the prediction equations for 1500, 1800, and 2100 GMT were similar, we shall discuss the verification of only the 1800 GMT forecasts.

Two quantities were computed as measures of the forecast performance, namely, the bias and skill score. The bias was defined as the sum of all forecast probabilities divided by the sum of the predictand observations. Unbiased forecasts would therefore have a value of one. The score used to measure the skill of these probability forecasts is based upon a quantity which is one-half the score defined by Brier (1950). This quantity, P, which has become commonly used in the NWS to verify probability forecasts, is defined as:

$$P = \frac{1}{N} \sum_{i=1}^{N} (F_i - O_i)^2, \tag{5}$$

where F_i is the forecast probability, O_i is the observed event (1 when the event occurs, 0 otherwise) for case i, and N is the total number of forecasts. If P_F and P_C are the P-scores for the operational probability and climatic frequency forecasts, respectively, then the skill score, SS, is:

$$SS = \frac{P_C - P_F}{P_C} \times 100(\%). \tag{6}$$

Thus, SS is the percentage improvement in P-score of the probabilities over that of the predictand climatic frequency forecasts. A negative score would mean the operational forecasts were inferior to climatology and a positive score would mean the forecasts were better than climatology. In terms of fit to the observations, the best climatic frequency available for this verification was the smoothed predictand relative frequency for individual predictand boxes discussed in the previous section. Since this quantity was also selected as a predictor in all operational equations, the skill score essentially measures the forecasting skill of other predictors.

Table 4 gives separate values of the bias and skill score for probabilities from the dependent and independent samples. When one compares skill scores for the two samples, one finds a result which appears rather paradoxical. That is, for the non-Gulf region, the skill dropped from the dependent sample to the independent sample as would be expected but, for the Gulf region, it improved! The same kind of pattern of changes is seen for the biases for the two regions. A clue to the explanation of these results can be found in the predictand frequencies. First, note that for the non-Gulf region, the frequency for the independent sample was about half of what it was in the dependent sample but, for the Gulf region, the frequencies for the two samples were nearly the same. These findings bear out a concern held at the outset of the verification experiments, namely, that the samples (particularly the independent sample) are too small to give consistent and representative statistical results. Nevertheless, the verification statistics do demonstrate that the operational forecasts improved appreciably upon forecasts of the climatic frequency of the predictand and that, on the average, this improvement for the independent sample was not greatly different from that obtained for the dependent sample.

Since these forecasts project out only to 6 hr into the future, one might expect forecasts of persistence to have skill. We computed the bias and skill score for persistence for the independent sample in the non-Gulf region. Persistence was defined identically to a predictand except, of course, it was based upon earlier data. An 1800 GMT persistence forecast for 2000-0000 GMT was defined from severe weather data for the period 1400-1800 GMT. (Note that this persistence quantity could not have been used as a predictor since the authenticated reports of severe weather used here are not available in real time.) The bias and skill score values obtained were 0.14 and -8.16%, respectively. The negative skill score means persistence forecasts were inferior to climatic frequency forecasts. The inferior skill and great under-forecasting bias of persistence is brought about by the short lifetimes of severe storms and, more so, by the large diurnal increase in storm frequency over the period from 1400 GMT to 0000 GMT (e.g., see Wolford 1960; Williams 1976).

One goal to strive for in probability forecasting is that the probabilities be reliable. That is, when a certain probability of an event is forecast on many occasions, it should agree closely with the frequency of occurrence of the event. Sanders (1967) has shown that minimizing the difference between the probability and the observed frequency will lower the P-score. A plot of the reliability for small probability intervals for the 1800 GMT non-Gulf operational forecasts is shown in figure 7. Most of the differences between the probabilities and observed frequencies are seen to be less than 10%. Interestingly, the plot exhibits a systematic bias; below 40% the probabilities are higher than the observed frequencies while above 50% the opposite is true. A similar pattern of differences was also found in the case of the dependent sample (not shown) although the differences in the range 0-40% were all less than 5%. Since this systematic error appears in the probabilities for the dependent sample, its cause must lie in a shortcoming in the method of developing the equations. So far, this shortcoming has not been precisely identified.

Since severe storm occurrences are rare events, and the independent sample of forecasts was small, probably a more meaningful method of evaluating this product is through subjective examination of a series of severe storm cases. How meaningful this approach is depends upon how the cases are selected. We had 2½ months worth of forecasts (March 31 to June 14, 1976) to choose from. In our selection procedure, we restricted our search to April days only and to cases that had 10 or more severe weather events reported within the forecasting grid for the period 2000-0000 GMT. (The first restriction was arbitrary; May and June cases had not been examined before the decision was made to consider only April cases.) Seven days in April met our storm occurrence criterion. From those seven cases we chose five which are shown in figures 8a-8e. Both the probability isopleths and the severe weather occurrences were plotted by computer.

After these cases were selected, we obtained the tornado and severe thunderstorm watches issued for the public by NSSFC and superimposed on the 1800 GMT probability maps those watches whose issue times (IT) and valid times (VT) corresponded best with the valid period of our predictand. We emphasize that our product was not transmitted to NSSFC until after April 14 and that, while the 1800 GMT probabilities are normally available by 1845 GMT, the transmission times were unreliable during the first few weeks following implementation of the computer program. Because of the unreliable transmission times and forecasters' nonfamiliarity in the use of these new probabilities as guidance, it is unlikely the probabilities influenced the watches issued for the cases selected. All the watches plotted in figures 8a-8e were for tornadoes except the one indicated in figure 8d, which was for severe thunderstorms.

The data cutoff time for an 1800 GMT probability forecast is 1820 GMT. Thus, the comparison of the probabilities and the watches is most meaningful when the ITs of the latter fall within and immediately around the period 1820-1845 GMT.

The case for April 14 (fig. 8a) exhibited highest probabilities in Kansas and southeastern Nebraska. An area delineated by the 15% isopleth in these states verified well, while a smaller area similarly defined in northwest Texas did not. NSSFC issued two watches, one at 2115 GMT and another at 2255 GMT. The positions of the probability maxima and these watches roughly agree (fig. 8a). Interestingly, the 2255 GMT watch became valid at the same time it was issued and, since there were numerous reports of severe weather within the box during the period 2000-0000 GMT, it may have been issued partly in response to real-time reports of those storms.

The April 15 case, shown in figure 8b, exhibited a probability maximum of just over 90% which equalled the highest value attained during the entire spring season of 1976. Note that numerous severe storm occurrences were clustered in and around this maximum. Two other much weaker probability maxima are present but only the one around southern Lake Michigan had severe weather associated with it. Again there is remarkable similarity in the positions of watches and areas of high probability. The only place where there is clear disagreement is in Minnesota where the nonissuance of a watch proved to be a wise choice. Note that of the four watches plotted, three were issued 1 to 3 hr after 1830 GMT, the time when the probabilities became available. Thus, later data was available upon which to base them.

For April 19 (fig. 8c), the 15% isopleth captured all reports while the 25% isopleth, which enclosed a much smaller area, captured most of them. Tornado watches issued 30 minutes to 1 hr after the probabilities had become available were roughly coincident with these areas. Considering the ITs and VTs of the two watches, the 1945 GMT watch may have been issued partly in response to real time reports of the tornadoes seen to the east of the earlier watch.

On April 23 (fig. 8d), while peak probabilities were in excess of 55%, most of the reports were of hail and wind storms. The watches are seen to coincide well with the 25% probability envelope, except for the one issued at 2056 GMT for western Oklahoma, which apparently did not verify. On April 24 (fig. 8e), most of the reported occurrences are enclosed by the 25% isopleth. Overall, the probabilities seem to verify better than the watches for this case. Note that the watch for southeastern Arkansas and northwestern Mississippi was issued about 4 hr after the probabilities had become available; it may have been issued in response to real time reports of tornadoes to the east of the earlier watch.

From this subjective examination, it is difficult to assess the relative skill of the probabilities versus the watches issued to the public. One problem seriously hampering the comparison is that the issue and valid times of the watches are highly variable. Furthermore, the majority of watches had a valid period of about 6 hr, which is 2 hr longer than that of the probabilities. From figures 8a-8e, note also that many of the watches were issued after 1845 GMT and, therefore, their valid periods extended well beyond 0000 GMT. These incompatibilities between the probability forecasts and the watches prevent definitive conclusions regarding their relative skill for these cases. However, all things considered, it does seem fair to say that the probabilities appeared roughly comparable in forecast accuracy with the watches.

As has been seen, tornado and severe thunderstorm watches are issued for small areas where these storms are expected to be a serious threat. Therefore, in order for a forecaster to make effective use of the probabilities as guidance, he would need a threshold probability to delineate the threat areas.

One objective method of determining such a threshold probability involves maximizing some verification score appropriate for categorical forecasts. This means determining a threshold probability which, when used to convert the probabilities into categorical statements, yields the highest score. A score that gives a measure of the likelihood or "threat" of a predictand occurrence when it is forecast to occur is the Threat score (Palmer and Allen 1949). The Threat score is identical with the critical success index defined by Donaldson et al. (1975) except for an empirical factor. The Threat score, TS, of the occurrence of event A is defined as:

$$TS = \frac{A}{A+B+C}, \qquad (7)$$

where A is the number of categorical forecasts of an occurrence for which the event is observed, B is the number of forecasts of an occurrence for which the event is not observed, and C the number of observed occurrences which were not forecast. For a single forecast map, the relationship of A, B, C is illustrated in figure 9. It shows that the TS for the event occurrence is the area over which the event is forecast and subsequently observed, divided by the envelope of the forecast and observed areas. The TS has a range from 0, when all forecasts are incorrect, to 1 when all forecasts are correct.

We evaluated the TS for 2.5% threshold probability intervals as shown in figure 10. The dependent and independent samples were combined for these computations to increase the sample size. The peak TS for the Gulf and non-Gulf regions is seen for probability thresholds of 16% and 22%, respectively. The forecast bias (now defined as the number of events categorically forecast divided by the number of events observed) for these optimum thresholds was 1.7 for the Gulf region and 1.2 for the non-Gulf region. The finding that the threshold probability producing the peak TS results in a tendency to overforecast the event, is consistent with previous experience at TDL with this procedure.

Recall from the earlier subjective evaluations of the individual cases that the areas delineated by the 15% and 25% probability isopleths matched up best with areas over which severe storms were observed. This finding is consistent with the optimum threshold probabilities determined objectively. This consistency is gratifying, particularly since the subjective evaluations of the cases were made before the objectively-derived thresholds were available.

4. CONCLUSIONS AND FUTURE PLANS

The evidence seen in the subjective comparative verification of the probabilities and tornado and severe thunderstorm watches issued by NSSFC indicates the probabilities have a level of accuracy roughly comparable with the watches. This leads us to conclude that the probabilities, along with the threshold values known to maximize the Threat score, should be a valuable aid in the preparation of watches issued to the public.

Admittedly there are incompatibilities between this product and watches, and these hamper use of the probabilities as guidance in preparing watches. One incompatibility is that, while the issue times of the probabilities are fixed, issue times of watches are variable. It may be, however, that if, through its use, forecasters find this product to be reliable as guidance, their issuance times of watches will eventually correspond more nearly with those of the probabilities. Another area of incompatibility lies in the valid periods. Most watches currently being issued have valid periods of about 6 hr (Galway, personal communication). This means the areal extent of these watches must be large, particularly for rapidly traveling convective storm systems. Obviously, if watch areas are to be reduced in size their valid periods must also be reduced. On the other hand, if, during the next few years, operational constraints will not permit changes in issue times and valid periods of watches, then the problems just discussed could be relieved by making these objective guidance forecasts available to forecasters more frequently; they could be produced as often as hourly.

Changes to our current forecasting procedures planned for the future include (1) redefining the present predictand and (2) developing an improved method of positioning the predictors. The present predictand combines occurrences of tornadoes, hail, and damaging wind such that a report of one or more of these phenomena constitutes a predictand occurrence. In the future we intend to form separate predictands for tornadoes and for hail and damaging winds combined. Separate probabilities for these new predictands should provide a better guidance product since watches are categorized in this manner. Beyond this measure of predictand stratification, we also plan to develop separate probabilities for "family" tornado outbreaks. The basis for this decision is the known capability of forecasting tornadoes more accurately when they occur in large numbers clustered in relatively small areas (Galway 1975).

During daily subjective verification of the operational spring season probabilities, we've noted that on occasions when convective systems moved either very fast or were nearly stationary the positions of peak probability were more likely to be in error than when the systems traveled near their typical speeds. The error occurs because the positions of the predictors relative to the predictand areas are determined by the average movement of the storm systems. In the future, we'll investigate the use of techniques to vary predictor positions according to synoptic conditions.

These changes plus the accumulation of longer data samples and additions of new or improved sources of predictor input data should result in an improved product.* New data sources expected to become available within the next few years include quantitative satellite data and output from a numerical boundary layer prediction model currently being developed at TDL. As for the data currently being used, improvements are expected in the MDR data and the LFM forecasts. Improved MDR data should result from a new code and a finer grid now being implemented in the NWS, while more accurate LFM forecasts are to be expected from current efforts at NMC to develop improved model initialization and physics along with a finer grid.

ACKNOWLEDGMENTS

I give my sincere appreciation to Dr. Mikhail A. Alaka for the valuable personal and scientific guidance he rendered throughout the course of this study. Mr. Mark Jacobs and Ms. Maryrose Loftus contributed excellent computer programming assistance, Messrs. Denis Sakelaris and William Herrmann aided in the preparation of the figures, and Ms. Barbara Howerton typed the manuscript.

REFERENCES

Bocchieri, J. R., and H. R. Glahn, 1972: Use of model output statistics for predicting ceiling height. Mon. Wea. Rev., 100, 869-879.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. Mon. Wea. Rev., 78, 1-3.

Charba, J. P., 1975: Operational scheme for short range forecasts of severe local weather. Preprints Ninth Conf. on Severe Local Storms, Am. Meteor. Soc., Boston, Mass., 226-231.

Charba, J. P., and M. J. Livingston, 1973: Preliminary results on short range forecasting of severe storms from surface predictors. Preprints Eighth Conf. on Severe Local Storms, Am. Meteor. Soc., Boston, Mass., 226-231.

Cressman, G. P., 1959: An operational objective analysis system. Mon. Wea. Rev., 87, 367-374.

*However, one must also carefully consider that any change to an existing statistical forecasting scheme, whether it involves the developmental methodology or the introduction of new input data, may carry along with it the undesirable consequence that the samples for development and/or evaluation of the product are reduced in size.

David, C. L., 1973: An objective method for estimating the probability of severe thunderstorms using predictors from the NMC (PE) numerical prediction model and from observed surface data. Preprints Eighth Conf. on Severe Local Storms, Am. Meteor. Soc., Boston, Mass., 223-225.

Donaldson, R. J., Jr., R. M. Dyer, and M. J. Kraus, 1975: An objective evaluator of techniques for predicting severe weather events. Preprints Ninth Conf. on Severe Local Storms, Am. Meteor. Soc., Boston, Mass., 321-326.

Galway, J. G., 1975: Relationship of tornado deaths to severe weather watch areas. Mon. Wea. Rev., 103, 737-741.

George, J. J., 1960: Weather Forecasting for Aeronautics, Academic Press, New York, 673 pp.

Glahn, H. R., and J. R. Bocchieri, 1976: Testing the limited-area fine mesh model for probability of precipitation forecasting. Mon. Wea. Rev., 104, 127-132.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics in objective weather forecasting. J. Appl. Meteor., 11, 1203-1211.

Howcroft, J., and A. Desmarais, 1971: The limited area fine mesh (LFM) model. NWS Tech. Proc. Bull., No. 67, 11 pp. (Available from NWS Hdqrs., Silver Spring, Md.).

Klein, W. H., 1970: The forecast research program of the Techniques Development Laboratory. Bull. Am. Meteor. Soc., 51, 133-142.

Miller, R. C., 1972: Notes on analysis and severe-storm forecasting procedures of the Military Warning Center. Air. Wea. Service, Tech. Rep. 200 (Rev.), Air Force Global Weather Central, Offutt AFB, Neb.

Moore, P. L., A. D. Cummings, and D. L. Smith, 1974: The National Weather Service manually digitized radar program and its application to precipitation probability forecasting. Preprints Fifth Conf. on Forecasting and Analysis, Am. Meteor. Soc., Boston, Mass., 69-74.

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

Reap, R. M., and D. S. Foster, 1975: New operational thunderstorm and severe storm probability forecasts based on model output statistics (MOS). Preprints Ninth Conf. on Severe Local Storms. Am. Meteor. Soc., Boston, Mass., 58-63.

Sanders, F., 1967: The verification of probability forecasts. J. Appl. Meteor., 6, 756-761.

Showalter, A. K., 1953: A stability index for thunderstorm forecasting. Bull. Am. Meteor. Soc., 34, 250-252.

Shuman, F. G., 1957: Numerical methods in weather prediction. II. Smoothing and filtering. Mon. Wea. Rev., 85, 357-361.

Williams, R. J., 1976: Surface parameters associated with tornadoes. Mon. Wea. Rev., 104, 540-545.

Wolford, L. A., 1960: Tornado occurrences in the United States. Tech. Paper No. 20, U.S. Weather Bureau, Washington, D.C,, 71 pp.

Table 1.--Potential severe weather predictor variables. The data source(s) for each variable is given in the right hand column.

	Variable	Data Source(s)
1.	Sfc u-component	Sfc obs
2.	Sfc v-component	Sfc obs
3.	BL u-component	LFM
4.	BL v-component	LFM
5.	850-mb u-component	LFM
6.	850-mb v-component	LFM
7.	500-mb u-component	LFM
8.	500-mb v-component	LFM
9.	MSL pressure	Sfc obs
10.	700-mb vertical velocity	LFM
11.	Sfc mixing ratio	Sfc obs
12.	850-mb mixing ratio	LFM
13.	700-mb mixing ratio	LFM
14.	850- to 500-mb mean mixing ratio	LFM
15.	850- to 500-mb mean temp.-dew point spread	LFM
16.	Sfc equiv. pot. temp.	Sfc obs
17.	Sfc equiv. pot. temp. x horizontal gradient of sfc equiv. pot. temp.	Sfc obs
18.	850-mb equiv. pot. temp.	LFM
19.	700-mb equiv. pot. temp.	LFM
20.	850- to 700-mb mean equiv. pot. temp.	LFM
21.	(Sfc minus 700-mb) equiv. pot. temp.	Sfc obs + LFM
22.	(850-mb minus 700-mb) equiv. pot. temp.	LFM
23.	Sfc equiv. pot. temp. advection	Sfc obs
24.	850-mb equiv. pot. temp. advection	LFM
25.	700-mb equiv. pot. temp. adevction	LFM
26.	Sfc moisture divergence	Sfc obs
27.	BL moisture divergence	Sfc obs + LFM
28.	850-mb moisture divergence	LFM
29.	K index	LFM
30.	Modified K index	Sfc obs + LFM
31.	Total Totals index	LFM
32.	Modified Total Totals index	Sfc obs + LFM
33.	Showalter index	LFM
34.	Modified Showalter index	Sfc obs + LFM
35.	500-mb wind speed	LFM
36.	Mag. of 850- to 500-mb wind shear	LFM
37.	Signed mag. of 850- to 500-mb wind shear	LFM
38.	(500-mb minus 850-mb) wind direction.	LFM
39.	BL vorticity	LFM
40.	500-mb vorticity	LFM
41.	500-mb vorticity advection	LFM
42.	Three-hr MSL pressure change	Sfc obs
43.	MDR variables (see Fig. 4)	MDR data
44.	Predictand relative frequency	Predictand data

Table 2.--Statistics from the development of the severe weather regression equations for both the Gulf and non-Gulf Coast regions (see fig. 1). Primary and backup equation statistics are listed except for 1500 GMT; for this time the primary equations were not used since MDR variables were not selected as predictors.

Time (GMT)	Equation	Region	Sample Size	Reduction of Variance (%)	Predictand Frequency (%)
1500	Backup	Gulf	15,972	6.51	1.51
	Backup	Non-Gulf	51,546	9.88	1.70
1800	Primary	Gulf	11,143	13.04	3.36
	Backup	Gulf	16,764	11.82	3.18
	Primary	Non-Gulf	30,403	21.16	3.67
	Backup	Non-Gulf	54,102	18.79	3.35
2100	Primary	Gulf	10,814	14.25	2.93
	Backup	Gulf	17,028	12.33	2.93
	Primary	Non-Gulf	34,482	17.98	3.32
	Backup	Non-Gulf	54,954	17.06	3.28

Table 3a.--Predictor variables selected for inclusion in the 1800 GMT primary severe weather regression equation for the non-Gulf region. The predictors are listed in the order selected with the data from which each was derived in parentheses. The cumulative reduction of variance with each added predictor is given in the right hand column. The subscript to the MDR variable corresponds to the convention adopted in figure 4.

	Variable	Cumulative Reduction of Variance (%)
1.	Modified Total Totals index (Sfc obs + LFM)	11.9
2.	Sfc moisture divergence (Sfc obs)	15.5
3.	Sfc equiv. pot. temp. advection (Sfc obs)	16.7
4.	850-mb equiv. pot. temp. advection (LFM)	17.7
5.	Showalter index (LFM)	18.5
6.	500-mb vorticity advection (LFM)	19.0
7.	850-mb equiv. pot. temp. (LFM)	19.3
8.	Modified Showalter index (binary) (Sfc obs + LFM)	19.6
9.	(Sfc minus 700-mb) equiv. pot. temp. (Sfc obs + LFM)	19.8
10.	MDR_1 (MDR data)	20.0
11.	Modified Showalter index (Sfc obs + LFM)	20.3
12.	Sfc equiv. pot. temp. (Sfc obs)	20.6
13.	Predictand relative frequency (Predictand data)	20.8
14.	500-mb wind speed (LFM)	21.0
15.	B.L. moisture divergence (Sfc obs + LFM)	21.2

Table 3b.--Same as table 3a except for the Gulf region.

	Variable	Cumulative Reduction of Variance (%)
1.	Modified Total Totals index (Sfc obs + LFM)	5.3
2.	850-mb equiv. pot. temp. advection (LFM)	7.8
3.	Sfc moisture divergence (Sfc obs)	8.8
4.	500-mb wind speed (LFM)	9.5
5.	Predictand relative frequency (Predictand data)	10.1
6.	850-mb equiv. pot. temp. (LFM)	10.6
7.	Modified Showalter index (binary) (Sfc obs + LFM)	11.1
8.	MDR_2 (MDR data)	11.5
9.	Boundary layer moisture divergence (Sfc obs + LFM)	11.8
10.	(Sfc minus 700-mb) equiv. pot. temp. (Sfc obs + LFM)	12.1
11.	Showalter index (binary) (LFM)	12.3
12.	MSL pressure (binary) (Sfc obs)	12.5
13.	850-mb moisture divergence (binary) (LFM)	12.7
14.	Sfc equiv. pot. temp. (Sfc obs)	12.9
15.	Modified K index (Sfc obs + LFM)	13.0

Table 4.--Statistics from verification of the probability forecasts for 2000-0000 GMT as produced by the 1800 GMT operational regression equations. The dependent sample contained 127 days from two spring seasons (mid-March to mid-June of 1974-75) while the independent sample contained 74 days from the period March 31 to June 14, 1976. See text for definition of the bias and skill score.

	Region	Sample Size (cases)	Bias	Skill Score (%)	Predictand Frequency (%)
Dependent Sample	Gulf	16,764	1.25	12.88	3.18
	Non-Gulf	54,102	1.18	18.64	3.35
Independent Sample	Gulf	9,768	0.98	14.18	2.98
	Non-Gulf	31,524	1.75	12.30	1.65

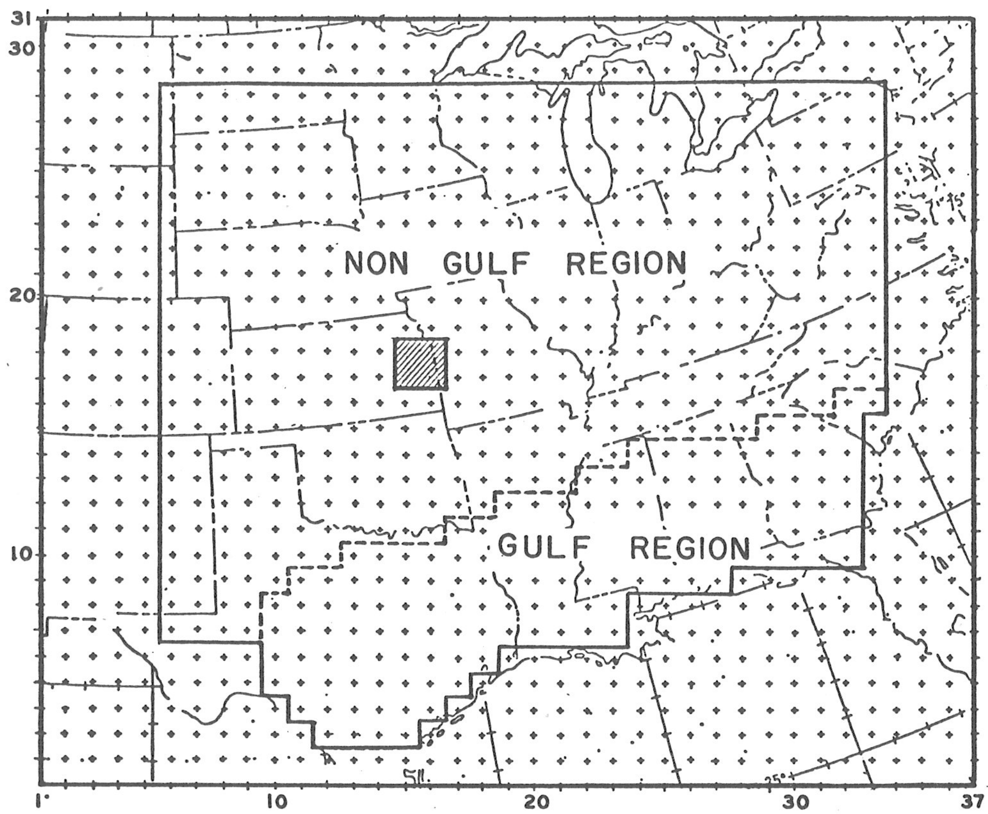


Figure 1.--Computation grid and severe weather predictand domain. The 31x37 grid is used both in the objective analysis of observed surface data and in the computation of the derived predictors. Severe weather predictands are restricted to the area enclosed by the heavy line; the hatched box illustrates the area of an individual predictand. Separate developmental samples were generated for the Gulf and non-Gulf regions.

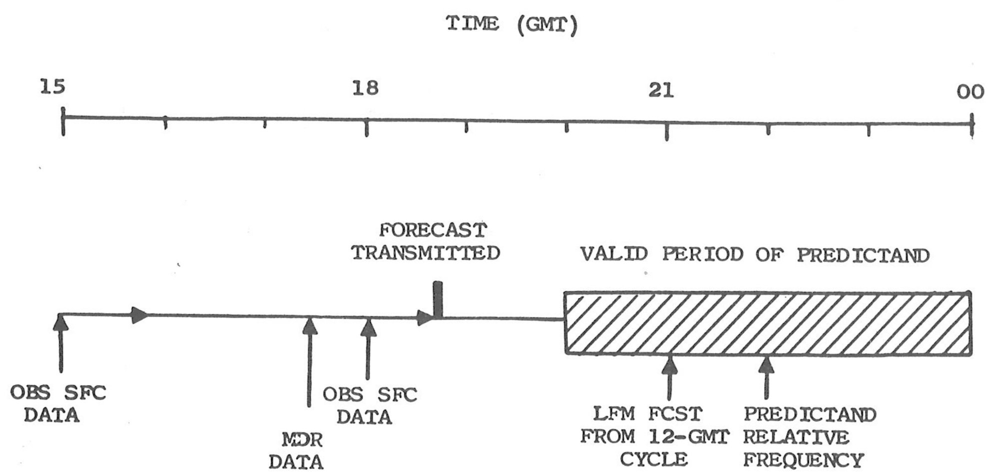


Figure 2.--Types of input data and their valid times relative to the valid period of the predictand for the 1800 GMT forecast cycle.

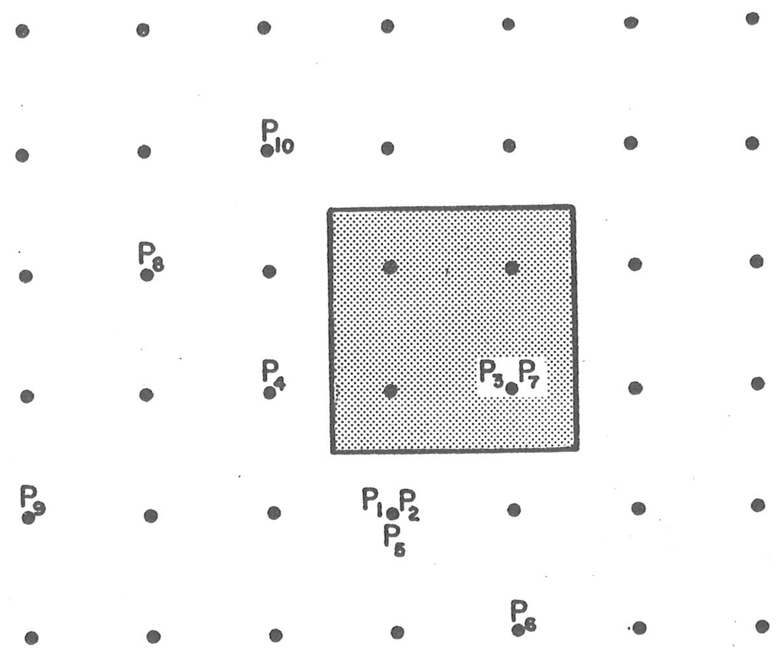


Figure 3.--The P's denote grid point positions of different predictor variables relative to the predictand box. The position of each variable was determined from the field of space-lagged linear correlation coefficients where the correlation is between that variable and the predictand; the highest correlation specified the position or offset of the variable relative to the predictand box (shaded). The subscripts to the P's are in the order of decreasing magnitude of the linear correlation coefficients. Each variable is identified below; the trailing information in parenthesis gives the sign of the correlation coefficient and the data from which the variable was computed:

P_1 - Modified Total Totals index (+; Sfc obs + LFM)
P_2 - Modified Showalter index (-; Sfc obs + LFM)
P_3 - Showalter index (-; LFM)
P_4 - Sfc moisture divergence (-; Sfc obs)
P_5 - (Sfc minus 700-mb) equiv. pot. temp. (+; Sfc obs + LFM)
P_6 - Equiv. pot. temp. (+; Sfc obs)
P_7 - Sfc. equiv. pot. temp. advection (+; Sfc obs)
P_8 - 500-mb vorticity advection (+; LFM)
P_9 - 850-mb equiv. pot. temp. advection (-; LFM)
P_{10} - 500-mb wind speed (+; LFM)

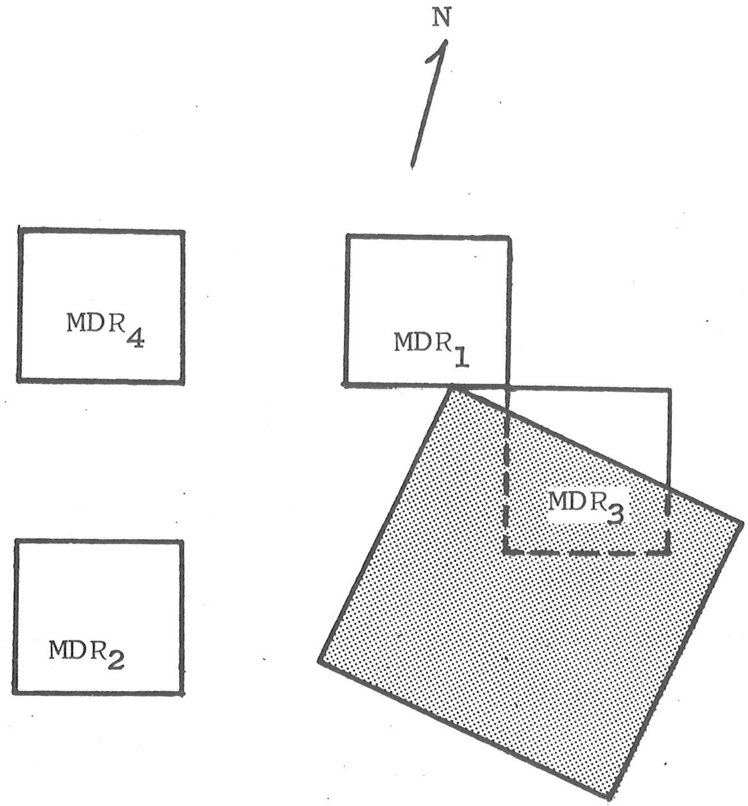
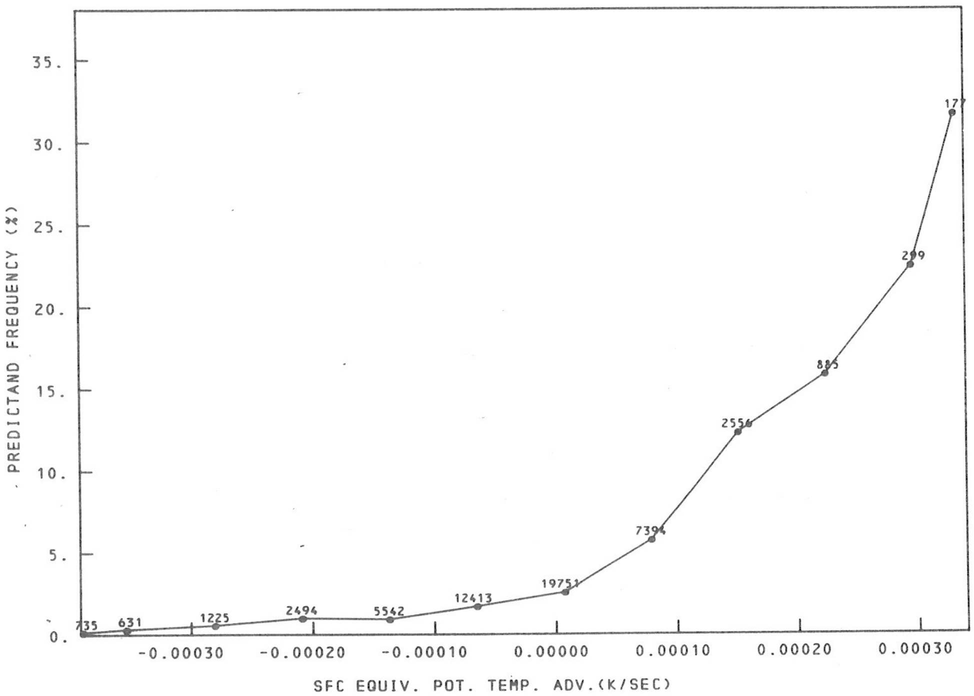


Figure 4.--Possible predictors defined from MDR data. The subscripts denote the order of selection of four MDR predictor boxes in a special screening regression run wherein 30 MDR boxes surrounding the predictand box (shaded) were offered.



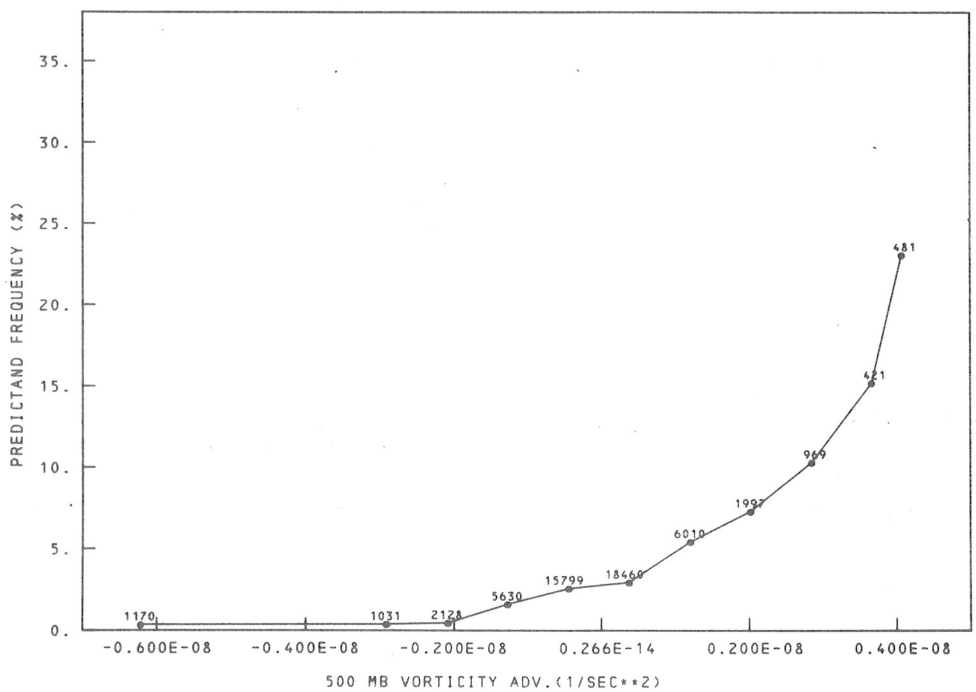
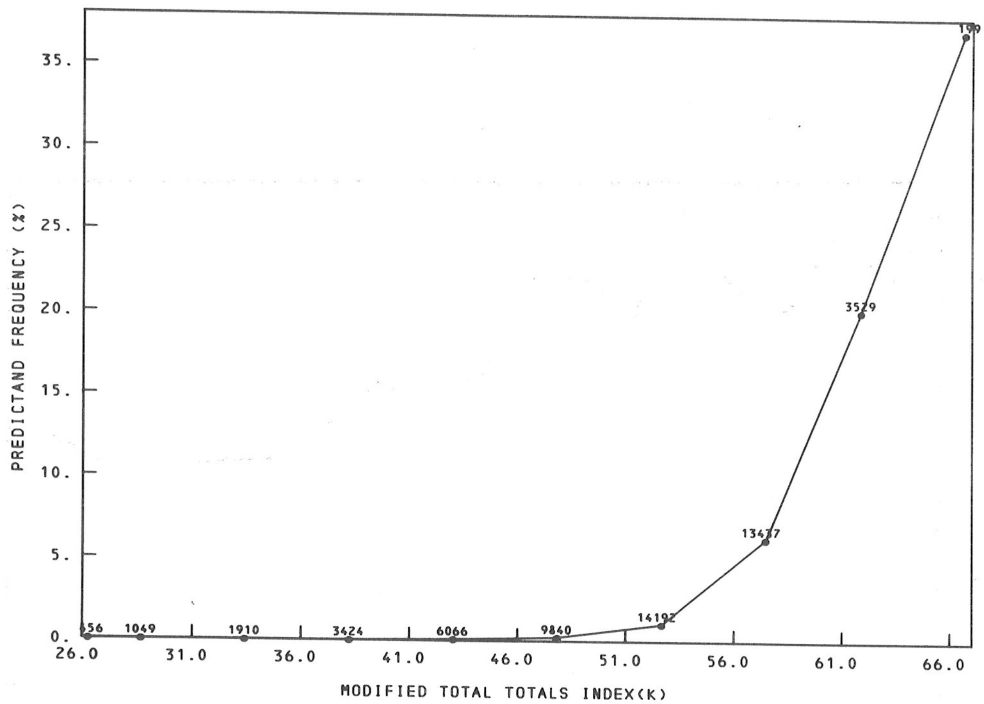


Figure 5.--Plots of predictand relative frequency versus selected potential predictor variables as determined from the dependent data sample. Among the variables in table 1, these exhibited the greatest nonlinearity. The number above each plotted point is the number of predictor/predictand data pairs used to define the point.



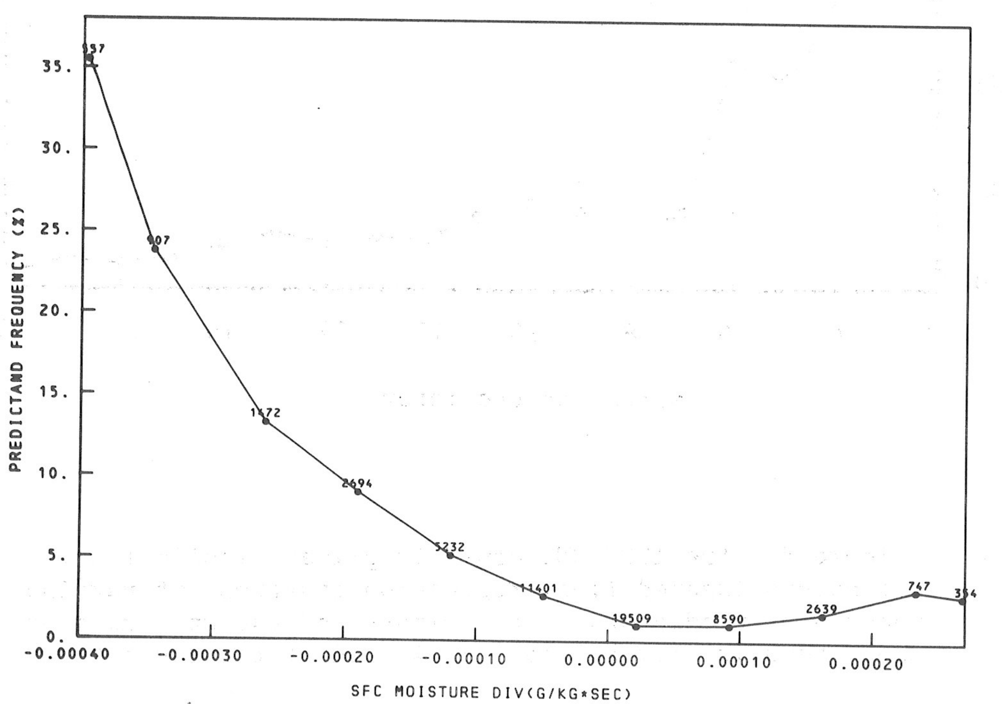


Figure 5.--Continued.

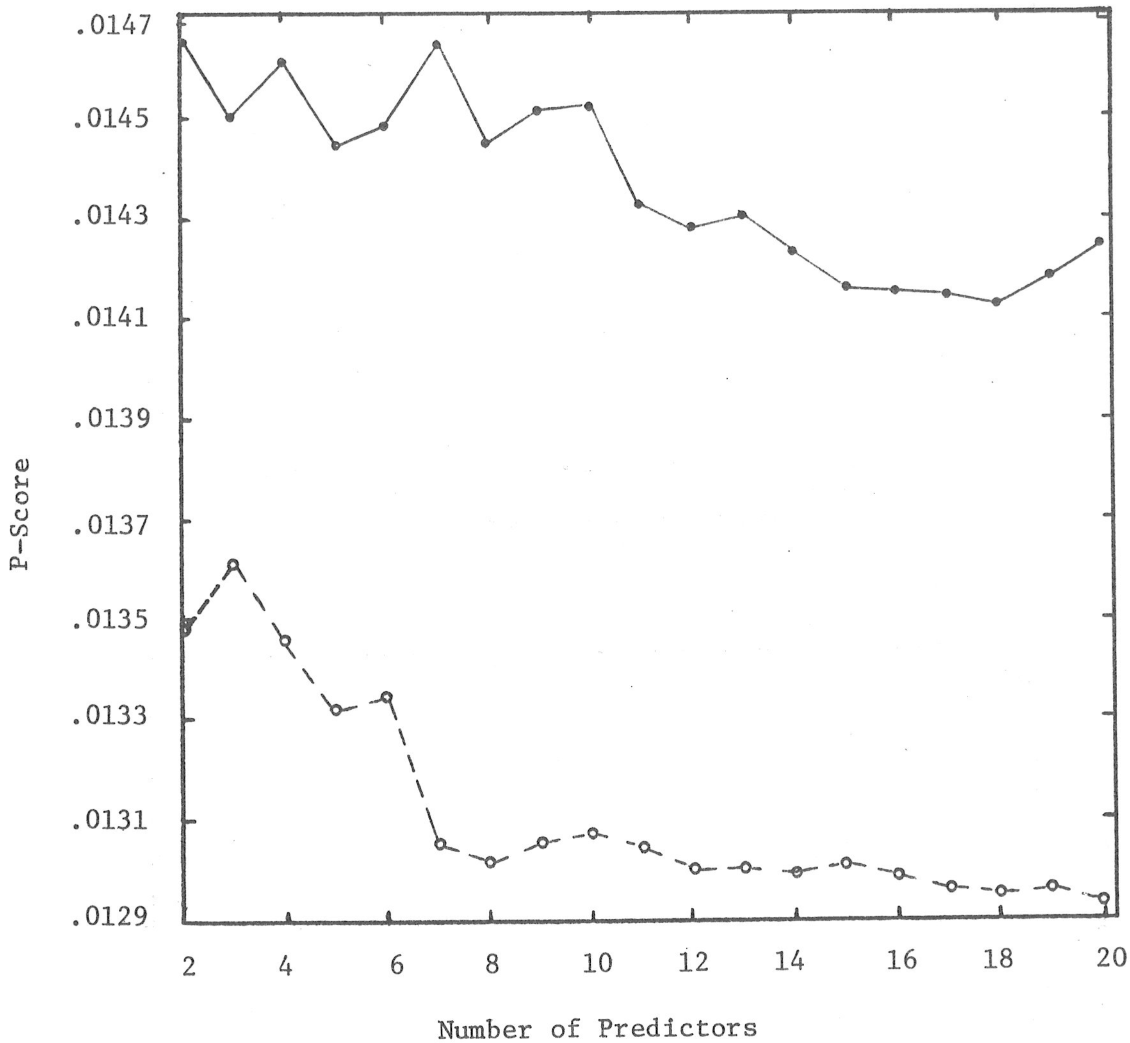


Figure 6.--P-Score for the 1800 GMT non-Gulf primary (solid line) and backup (dashed line) regression equations of varying numbers of predictors. The primary and backup equations used operationally had 15 and 14 predictors, respectively.

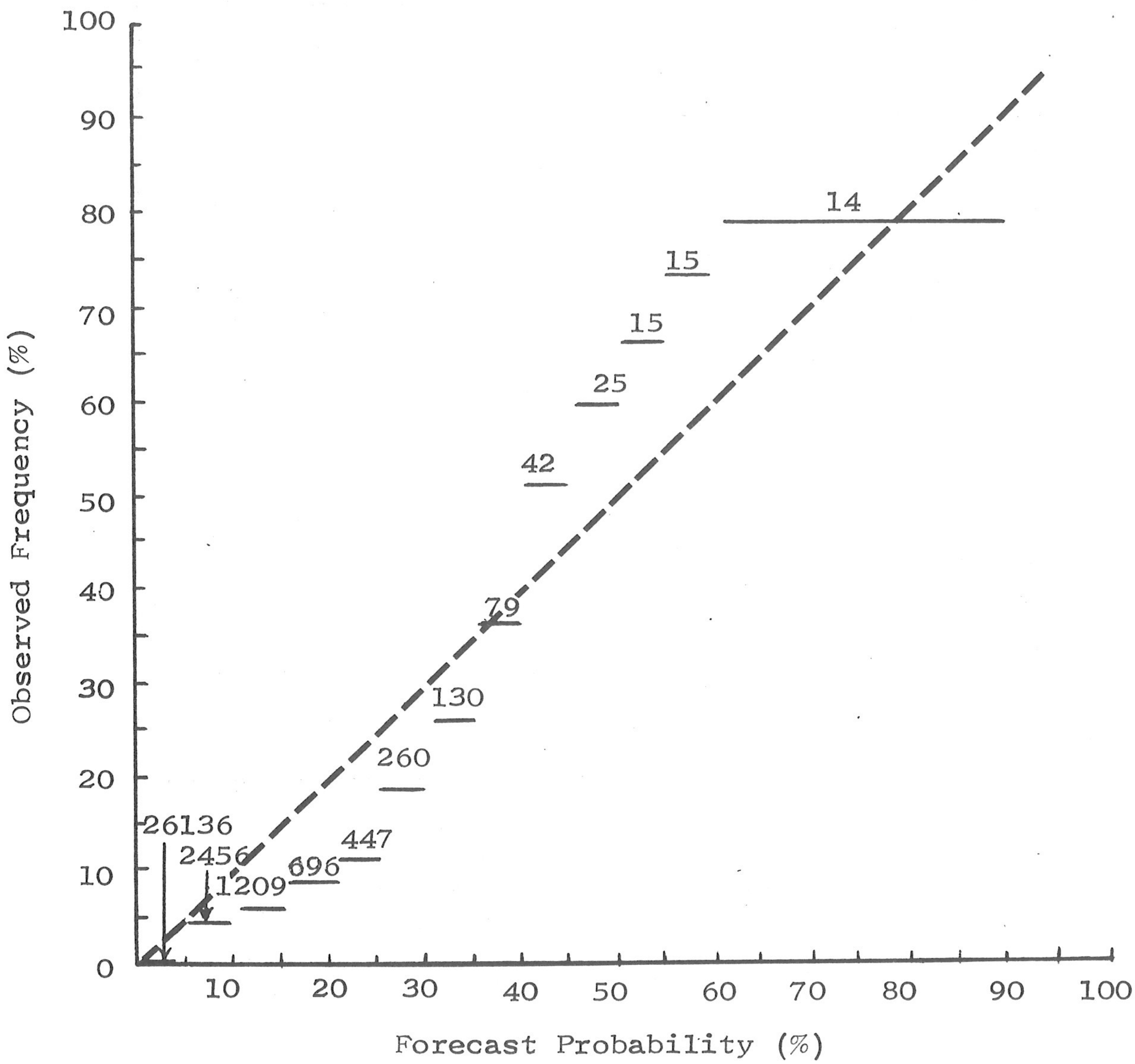


Figure 7.--Reliability of the 1800 GMT probability forecasts for the non-Gulf region (independent sample). Observed frequencies are plotted for 5% probability intervals up to 60%; the forecast probability range from 60 to 90%, where there were only 14 cases, as shown, was considered as a single interval. The dashed line depicts perfect reliability.

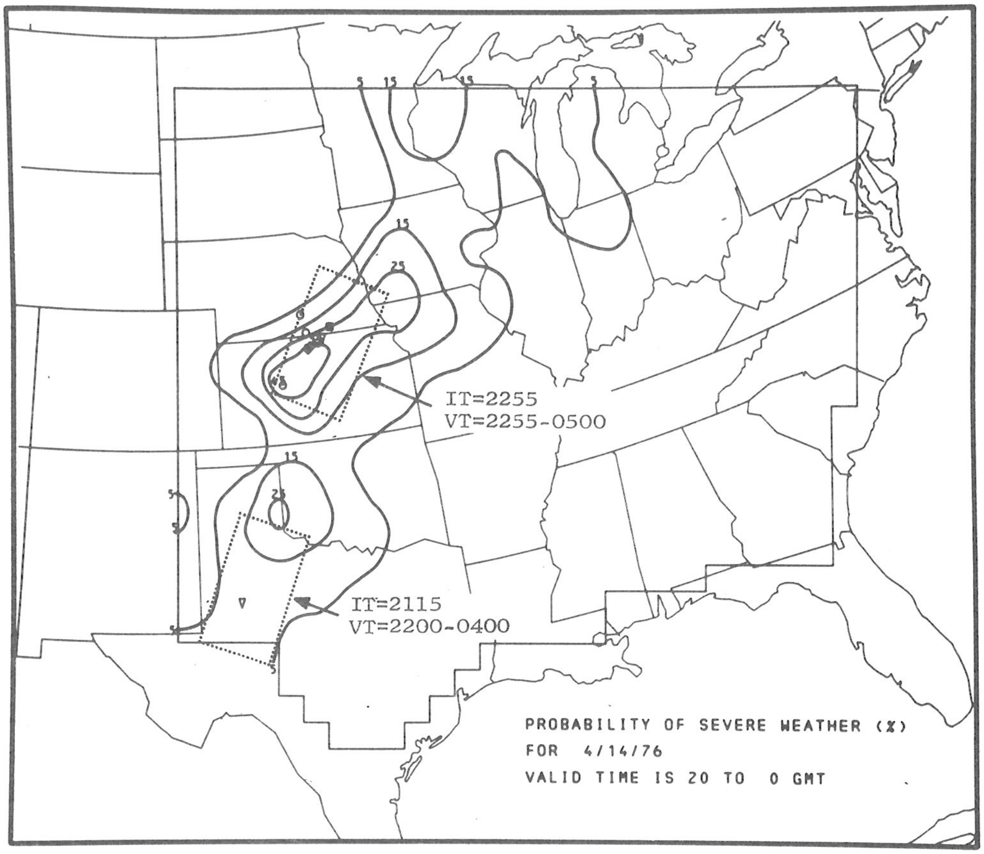


Figure 8a.--Sample operational probability forecast along with verifying reports of severe weather. This forecast is valid for the period 2000-000 GMT as are the plotted reports of severe weather. The symbols denote the following: ∇ - tornado; o - hail ≥ 3/4 in. dia.; □ - wind damage or surface winds > 50 kt. These probabilities, produced by the 1800 GMT equations, are available in real time by 1830 to 1845 GMT. Tornado watches issued by NSSFC whose issue times (IT) and valid times (VT) correspond closest to that of this forecast are indicated as the dotted rectangles.

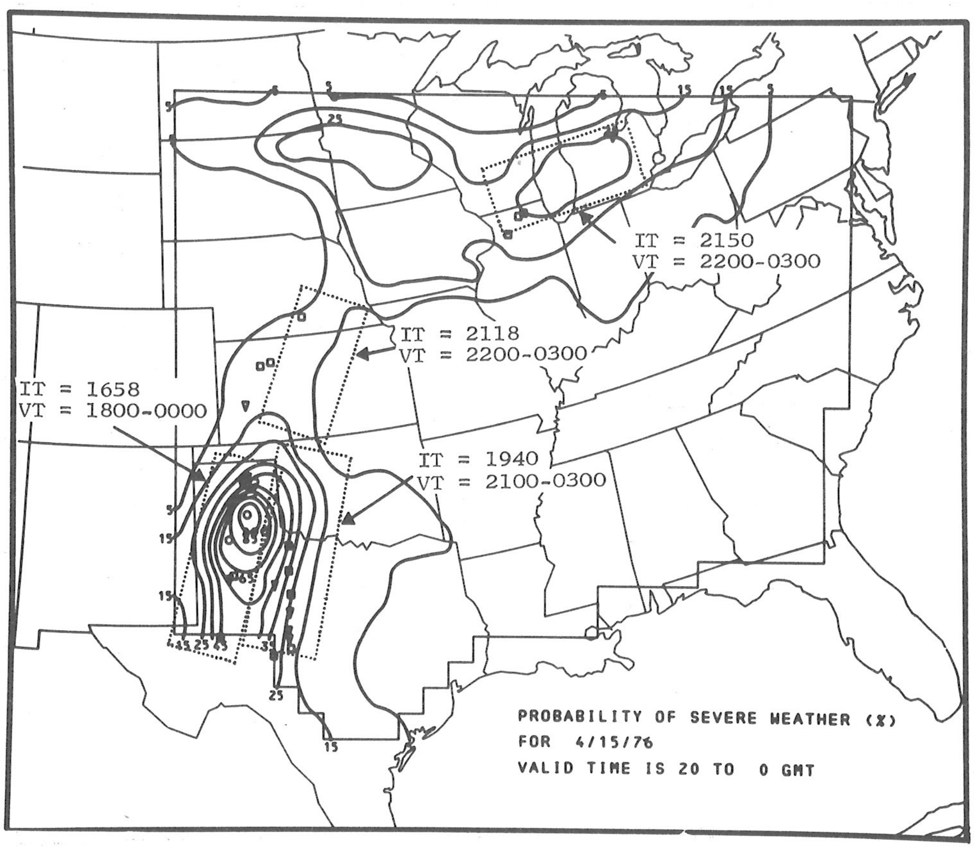


Figure 8b.--Same as figure 8a.

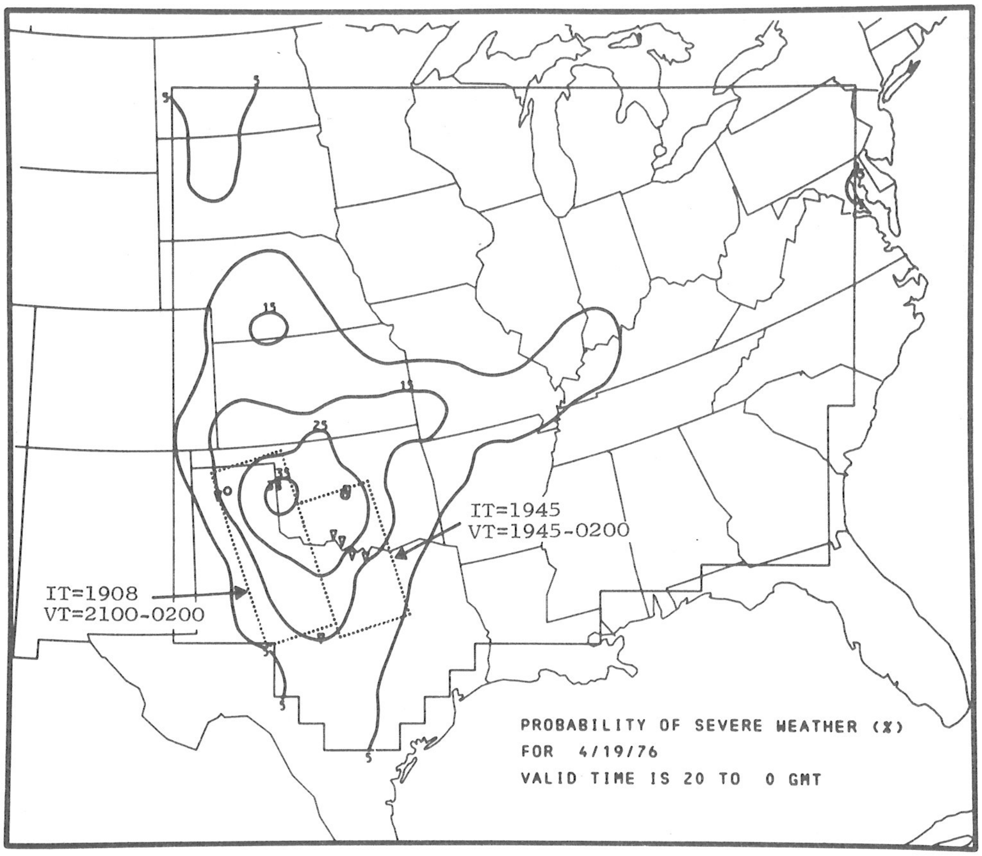


Figure 8c.--Same as figure 8a.

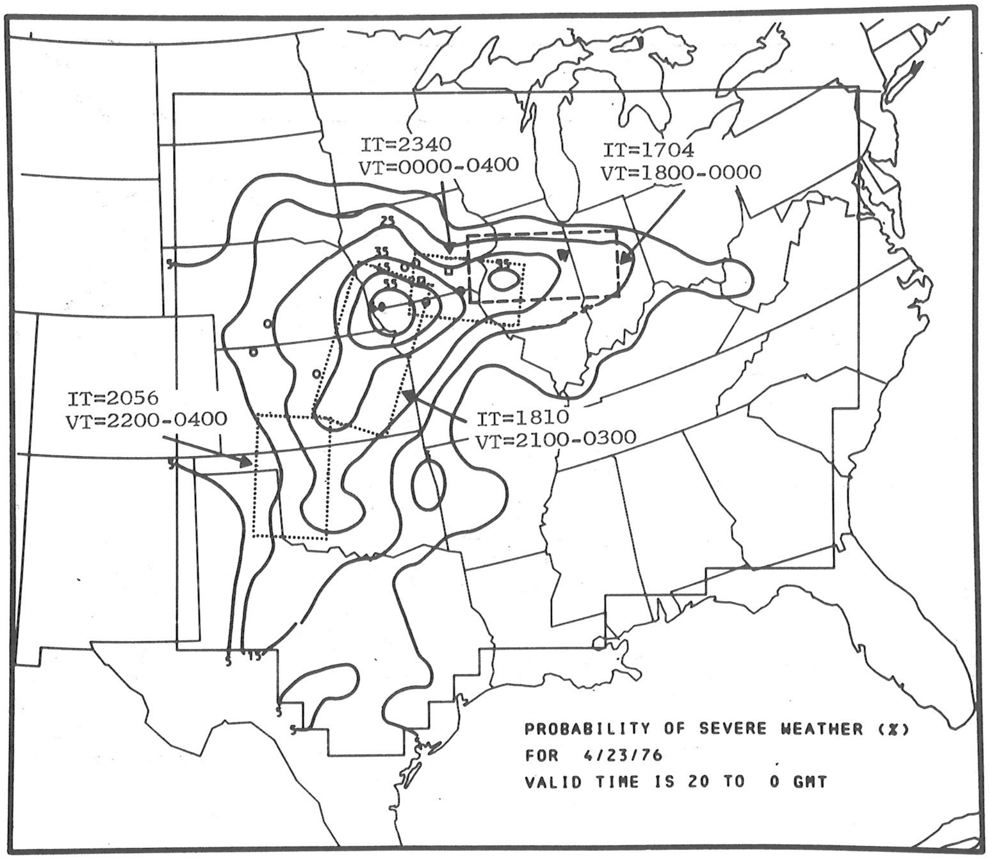


Figure 8d.--Same as figure 8a except that the dashed rectangle is a severe thunderstorm watch.

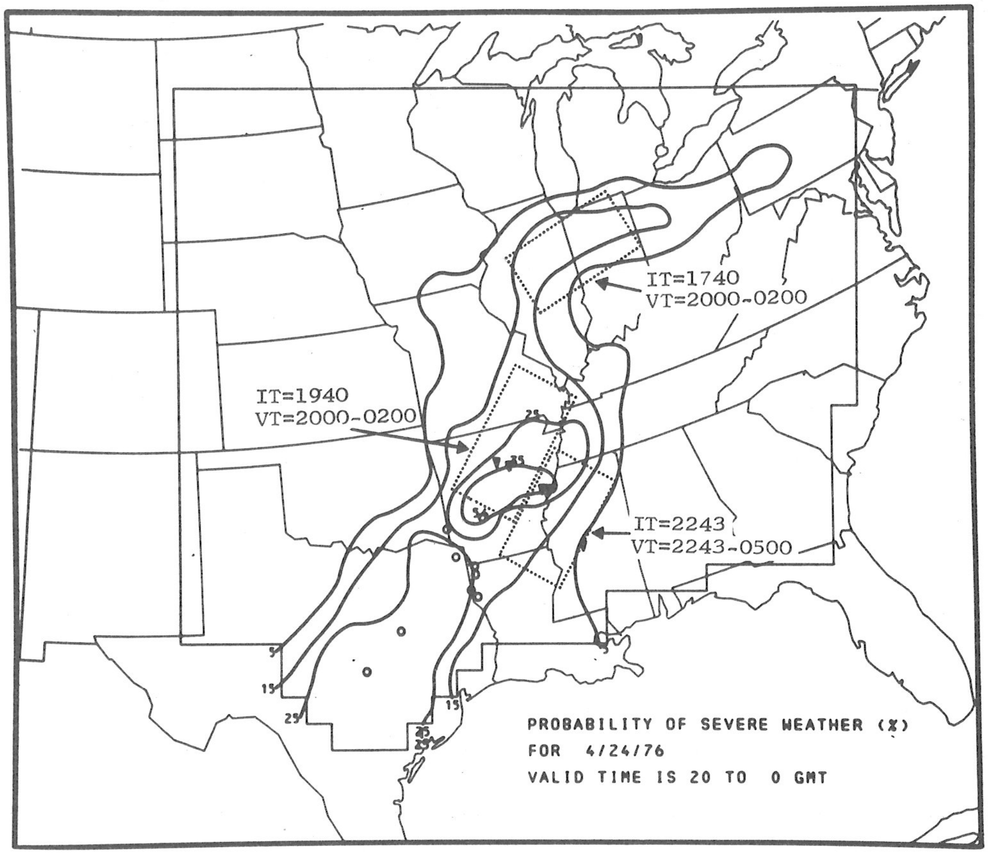


Figure 8e.--Same as figure 8a.

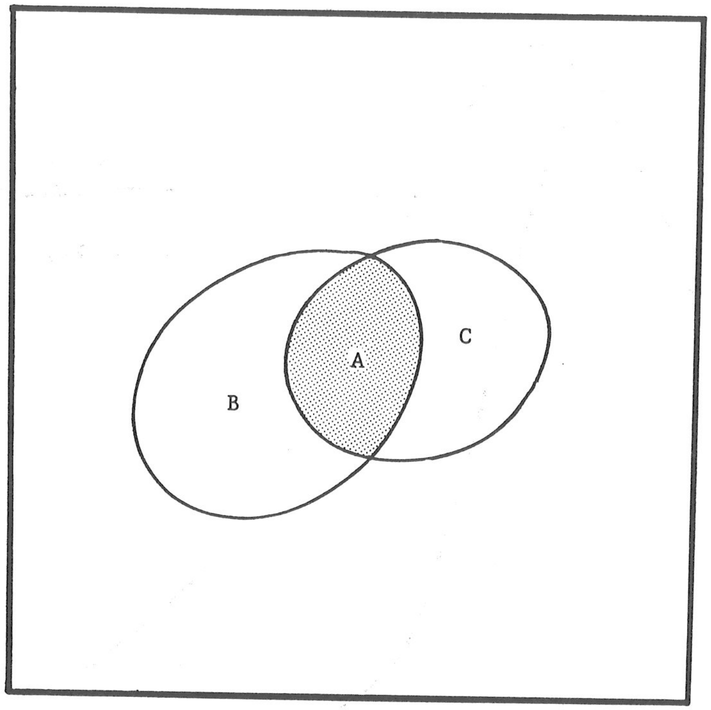


Figure 9.--Schematic illustrating the terms in the Threat score. The predictand event is forecast to occur within areas A + B while the event is observed to occur within areas A + C.

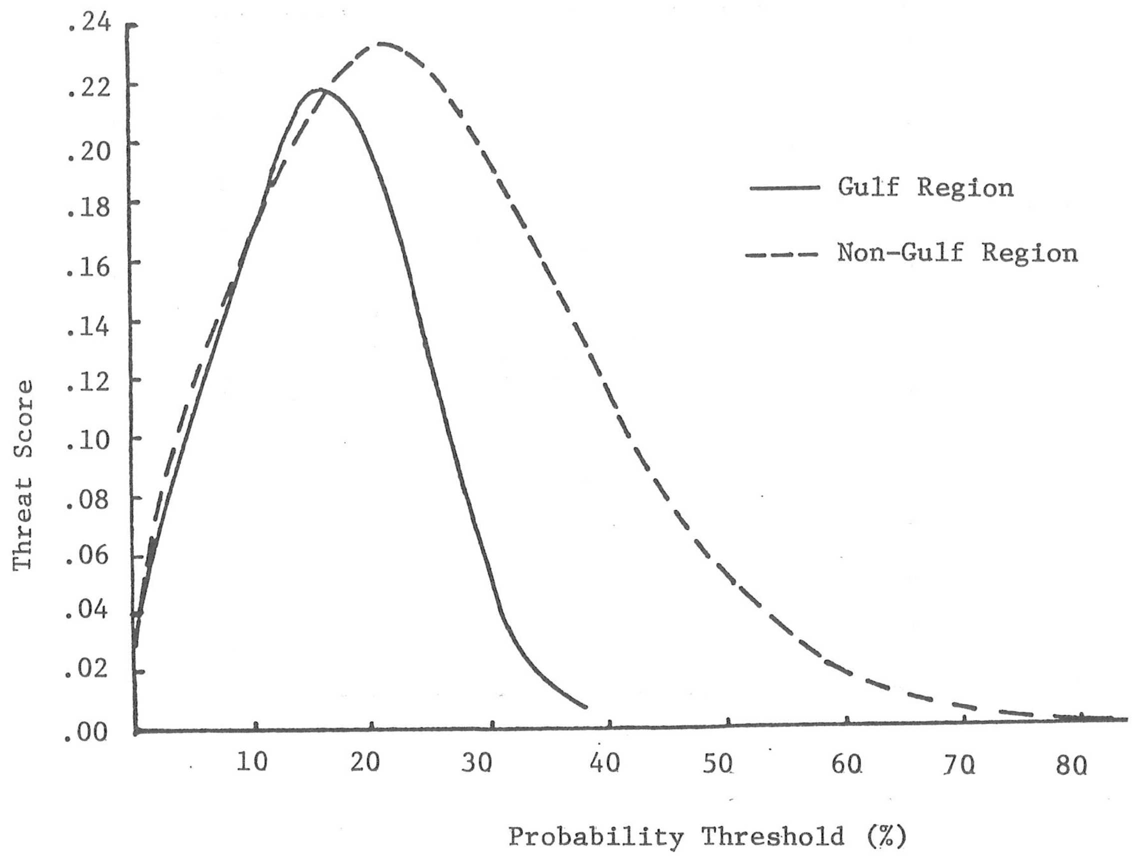


Figure 10.--Threat score versus probability thresholds for the 1800 GMT Gulf and non-Gulf region forecasts (dependent and independent samples combined).

(Continued from inside front cover)

WBTM TDL 33 Calculation of Precipitable Water. L. P. Harrison, June 1970. (PB-193-600)
WBTM TDL 34 An Objective Method for Forecasting Winds Over Lake Erie and Lake Ontario. Celso S. Barrientos, August 1970. (PB-194-586)
WBTM TDL 35 Probabilistic Prediction in Meteorology: a Bibliography. Allan H. Murphy and Roger A. Allen, June 1970. (PB-194-415)
WBTM TDL 36 Current High Altitude Observations--Investigation and Possible Improvement. M. A. Alaka and R. C. Elvander, July 1970. (COM-71-00003)

NOAA Technical Memoranda

NWS TDL-37 Prediction of Surface Dew Point Temperatures. R. C. Elvander, February 1971. (COM-71-00253)
NWS TDL-38 Objectively Computed Surface Diagnostic Fields. Robert J. Bermowitz, February 1971. (COM-71-00301)
NWS TDS-39 Computer Prediction of Precipitation Probability for 108 Cities in the United States. William H. Klein, February 1971. (COM-71-00249)
NWS TDL-40 Wave Climatology for the Great Lakes. N. A. Pore, J. M. McClelland, C. S. Barrientos, and W. E. Kennedy, February 1971. (COM-71-00368)
NWS TDL-41 Twice-Daily Mean Monthly Heights in the Troposphere Over North America and Vicinity. August F. Korte, June 1971. (COM-71-00826)
NWS TDL-42 Some Experiments With a Fine-Mesh 500-Millibar Barotropic Model. Robert J. Bermowitz, August 1971. (COM-71-00958)
NWS TDL-43 Air-Sea Energy Exchange in Lagrangian Temperature and Dew Point Forecasts. Ronald M. Reap, October 1971. (COM-71-01112)
NWS TDL-44 Use of Surface Observations in Boundary-Layer Analysis. H. Michael Mogil and William D. Bonner, March 1972. (COM-72-10641)
NWS TDL-45 The Use of Model Output Statistics (MOS) To Estimate Daily Maximum Temperatures. John R. Annett, Harry R. Glahn, and Dale A. Lowry, March 1972. (COM-72-10753)
NWS TDL-46 SPLASH (Special Program To List Amplitudes of Surges From Hurricanes) I. Landfall Storms. Chester P. Jelesnianski, April 1972. (COM-72-10807)
NWS TDL-47 Mean Diurnal and Monthly Height Changes in the Troposphere Over North America and Vicinity. August F. Korte and DeVer Colson, August 1972. (COM-72-11132)
NWS TDL-48 Synoptic Climatological Studies of Precipitation in the Plateau States From 850-, 700-, and 500-Millibar Lows During Spring. August F. Korte, Donald L. Jorgensen, and William H. Klein, August 1972. (COM-73-10069)
NWS TDL-49 Synoptic Climatological Studies of Precipitation in the Plateau States From 850-Millibar Lows During Fall. August F. Korte and DeVer Colson, August 1972. (COM-74-10464)
NWS TDL-50 Forecasting Extratropical Storm Surges For the Northeast Coast of the United States. N. Arthur Pore, William S. Richardson, and Herman P. Perrotti, January 1974. (COM-74-10719)
NWS TDL-51 Predicting the Conditional Probability of Frozen Precipitation. Harry R. Glahn and Joseph R. Bocchieri, March 1974. (COM-74-10909/AS)
NWS TDL-52 SPLASH (Special Program to List Amplitudes of Surges From Hurricanes) II. General Track and Variant Storm Conditions. Chester P. Jelesnianski, March 1974.
NWS TDL-53 A Comparison Between the Single Station and Generalized Operator Techniques for Automated Prediction of Precipitation Probability. Joseph R. Bocchieri, September 1974. (COM-74-11763/AS)
NWS TDL-54 Climatology of Lake Erie Storm Surges at Buffalo and Toledo. N. Arthur Pore, Herman P. Perrotti, and William S. Richardson, December 1974. (COM-75-10587/AS)
NWS TDL-55 Dissipation, Dispersion and Difference Schemes. Paul E. Long, Jr., May 1975. (COM-75-10972/AS)
NWS TDL-56 Some Physical and Numerical Aspects of Boundary Layer Modeling. Paul E. Long, Jr., May 1975. (COM-75-10980)
NWS TDL-57 A Predictive Boundary Layer Model. Wilson A. Shaffer and Paul E. Long, Jr., May 1975.
NWS TDL-58 A Preliminary View of Storm Surges Before and After Storm Modifications for Alongshore-Moving Storms. Chester P. Jelesnianski and Celso S. Barrientos, October 1975. (PB247362)
NWS TDL-59 Assimilation of Surface, Upper Air, and Grid-Point Data in the Objective Analysis Procedure for a Three-Dimensional Trajectory Model. Ronald M. Reap, February 1976. (PB-256082/AS)
NWS TDL-60 Verification of Severe Local Storm Warnings Based on Radar Echo Characteristics. Donald Foster, June 1976.
NWS TDL-61 A Sheared Coordinate System for Storm Surge Equations of Motion with a Mildly Curved Coast. C. P. Jelesnianski, July 1976. (PB261956)
NWS TDL-62 Automated Prediction of Thunderstorms and Severe Local Storms. Ronald M. Reap and Donald S. Foster, April 1977.
NWS TDL-63 Automated Great Lakes Wave Forecasts. N. Arthur Pore, February 1977.
NWS TDL-64 Operational System for Predicting Thunderstorms Two to Six Hours in Advance. Jerome P. Charba, March 1977.

NOAA--S/T 77-2670