NOAA Technical Memorandum
NWS TDL 82

Detection of Severe Local Storm Phenonema by Automated Interpretation of Radar and Storm Environment

David H. Kitzmiller and Jay P. Breidenbach

Techniques Development Laboratory
Silver Spring, MD
August 1995

UNITED STATES DEPARTMENT OFCOMMERCE / National Oceanic and Atmospheric Administration / National Weather Service

NOAA TECHNICAL MEMORANDUMS

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 methology used in the National Meteorological Service and for the more effective utilization of weather forecasts by the ultimate user.

NOAA Technical Memorandums 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 elewhere at a later date. Publications 1 through 5 are in the former series Weather Bureau Technical Notes (TN), Technical Memorandums, Weather Bureau Technical Memorandum, (WBTM). Beginning with TDL 37, publications are now part of the series NOAA Technical Memorandums, National Weather Service (NWS).

Publications listed below are available from the National Technical Information Service, U.S. Department of Commerce, sills Bldg., 5285 Port Royal Road, Springfield, VA 22161. Prices on request. Order by accession number (given in parentheses).

ESSA Technical Memorandums

WBTM	TDL 21	Automatic Decoding of hourly Weather Reports. George W. Hollenbaugh, Harry R. Glahn, and Dale A. Lowry, July 1969, 27 pp. (PB-185-806)
WBTM	TDL 22	An Operationally Oriented Objective Analysis Program. Harry R. Glahn, George W. Hollenbaugh, and Dale A. Lowry, July 1969, 20 pp. (FB-186-129)
WBTM	TDL 23	An Operational Subsynoptic Advection Model. Harry R. Glahn, Dale A. Lowry, and George W. Hollenbaugh, July 1969, 26 pp. (PB-186-389)
WBTM	TDL 24	A Lake Erie Storm Surge Forecasting Technique. William S. Richardson and N. Arthur Pore, August 1969, 23 pp. (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, 9 pp. plus appendixes A and B. (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, 27 pp. plus appendix. (PB-189-105)
WBTM	TDL 27	An Operational Method for Objectively Forecasting Probability of Precipitation. Harry R. Glahn and Dale A. Lowry, October 1969, 24 pp. (PB-188-660)
WBTM	TDL 28	Techniques for Forecasting Low Water Occurrence at Baltimore and Norfolk. James M. McClelland, March 1970, 34 pp. (PB-191-744)
WBTM	TDL 29	A Method fcr Predicting Surface Winds. Harry R. Glahn, March 1970, 18 pp. (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, 103 pp. (PB-193-449)
WBTM	TDL 31	Persistence of Precipitation at 108 Cities in the Conterminous United States. Donald L. Jorgensen and William H. Klein, May 1970, 84 pp. (PB-193-599)
WBTM	TDL 32	Computer-Produced Worded Forecasts. Harry R. Glahn, June 1970, 8 pp. (PB-194-262)
WBTM	TDL 33	Calculation of Precipitable Water. L. P. Harrison, June 1970, 61 pp. (PB-193-600)
WBTM	TDL 34	An Objective Method for Forecasting Winds Over Lake Erie and Lake Ontario. Celso S. Barrientos. August 1970, 20 pp. (PB-194-586)
WBTM	TDL 35	Probabilistic Prediction in Meteorology; a Bibliography. Allan H. Murphy and Roger A. Allen, June 1970, 60 pp. (PB-194-415)
WBTM	TDL 36	Current High Altitude Observations--Investigation and Possible Improvement. M. A. Alaka and R. C. Elvander, July 1970, 24 pp. (COM-71-00003)
NWS	TDL 37	Prediction of Surface Dew Point Temperatures. R. C. Elvander, February 1971, 40 pp. (COM-71-00253)
NWS	TDL 38	Objectively Computed Surface Diagnostic Fields. Robert J. Bermowitz, February 1971, 23 pp. (COM-71-0301)
NWS	TDL 39	Computer Prediction of Precipitation Probability for 108 Cities in the United States. William H. Klein, February 1971, 32 pp. (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, 61 pp. (COM-71-00368)
NWS	TDL 41	Twice-Daily Mean Heights in the Troposphere Over North America and Vicinity. August F. Korte, June 1971, 31 pp. (COM-71-0286)
NWS	TDL 42	Some Experiments With a Fine-Mesh 500-millibar Barotropic Model. Robert J. Bermowitz, August 1971, 20 pp. (COM-71-00958)
NWS	TDL 43	Air-Sea Energy Exchange in Lagrangian Temperature and Dew Point Forecasts. Ronald M. Reap, October 1971, 23 pp. (COM-71-01112)
NWS	TDL 44	Use of Surface Observations in Boundary-Layer Analysis. H. Michael Mogil and William D. Bonner, March 1972, 16 pp. (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, 14 pp. (COM-72-10753)

(Continued on inside back cover)

NOAA Technical Memorandum
NWS TDL 82

Detection of Severe Local Storm Phenonema by Automated Interpretation of Radar and Storm Environment

David H. Kitzmiller and Jay P. Breidenbach

Techniques Development Laboratory
Silver Spring, MD
August 1995

UNITED STATES
DEPARTMENT OFCOMMERCE
Ronald H. Brown
Secretary

National Oceanic and
Atmospheric Administration
D. James Baker
Under Secretary

National Weather Service
Elbert W. Friday, Jr.
Assistant Administrator

Table of Contents

		Page
	Abstract	1
1.	Introduction	1
2.	Observational and Model Data Used in Algorithm Development	3
3.	Creation of a Statistical Development Sample for SWP3 and HAIL3	4
4.	Individual Radar and Environmental Indices as Predictors of Severe Weather	4
5.	Derivation of the SWP3 and HAIL3 Algorithms	7
6.	Graphical Presentation of the Algorithm Output	9
7.	Reliability of the SWP3 and HAIL3 Algorithms	9
8.	Skill of the Algorithms in Terms of Categorical Forecasts	10
9.	Differences between the SWP3 Algorithm and Operational SWP	12
10.	Conclusions and Implications for Operational Use of SWP3 and HAIL3	13
	References	14
	Table	17
	Figures	18

DETECTION OF SEVERE LOCAL STORM PHENOMENA BY INTERPRETATION OF RADAR AND STORM ENVIRONMENT INFORMATION

David H. Kitzmiller and Jay P. Breidenbach

Techniques Development Laboratory
Office of Systems Development
National Weather Service
Silver Spring, Maryland

ABSTRACT

Many operational features of the WSR-88D were incorporated specifically to aid forecasters in the detection of severe local storms (damaging winds, large hail, and tornadoes). One interpretive product, the Severe Weather Potential (SWP) algorithm, yields an index proportional to the probability that an individual thunderstorm cell will soon produce any severe weather phenomena. The SWP is based solely on radar information, namely vertically-integrated liquid (VIL) and storm horizontal extent.

Forecasters have long known that critical values of many radar indices for severe weather change with the storm environment. In particular, shallow storms with moderately high VIL values are much more likely to produce severe weather in the spring than in the summer. This work describes an automated solution to the problem of adapting severe/nonsevere VIL thresholds to environment conditions. New algorithms have been developed that incorporate radar data and estimates of upper-air temperature and wind vectors. These techniques produce the probabilities of general severe weather and of large hail within the area near a thunderstorm. The probabilities are valid within a square region 44 km on a side for 30 minutes after the radar observation.

Estimates of the algorithms' skill in terms of categorical (severe/nonsevere) forecast scores will be presented. Comparative verification tests indicate that the new radar/environmental severe weather algorithms can produce 10-15% fewer false alarms than does the operational SWP algorithm.

1. INTRODUCTION

An important element of the National Weather Service (NWS) mission is providing to the public warnings of severe local storms (thunderstorms featuring damaging winds, large hail, or tornadoes). Because thunderstorms develop, decay, and often move rapidly, forecasters rely on radar and satellite observations, as well as reports from dedicated spotters and the general public, in monitoring these weather systems.

Many operational features of the Weather Surveillance Radar 1988 (Doppler) (WSR-88D) were included specifically to aid in severe storm detection. The WSR-88D data processing system includes a number of automated interpretive aids to alert forecasters to storms that are likely to bear hail, or rotating storms which may produce tornadoes (Klazura and Imy 1993). One such interpretive product, the Severe Weather Potential (SWP) algorithm, (Kitzmiller et al. 1995) designed to assess the probability that individual storms are producing, or will shortly produce, damaging winds, large hail, or tornadoes.

The SWP is based on the storm cell's maximum vertically-integrated liquid (VIL) and its horizontal extent, both of which are strongly correlated to the probability of both large hail and damaging winds.

The SWP algorithm, like most of the other WSR-88D detection algorithms, was introduced primarily as a "first cut" discriminator of severe and nonsevere storms, not as a categorical assessment of storm severity. It indicates the majority of storms as very unlikely to produce severe weather, and a few storms as so dangerous that they probably warrant official warnings even if there is no time to examine other radar information or obtain independent confirmation from human observers. The remaining storms might be considered sufficiently threatening to warrant closer examination by radar and/or dedicated spotters. Our goal in the present work is to improve the SWP algorithm's ability to discriminate between innocuous and dangerous storms.

Forecasters and researchers have long recognized that, while VIL and severe weather probability are highly correlated over most of the continental U. S., the exact nature of the VIL/severe weather relationship is highly dependent on the season and on geographical location. In general, for a sample of storms in any limited range of VIL values, the probability of severe storm phenomena is positively correlated to the mid-tropospheric wind speed and temperature lapse rate, and negatively correlated to the mean temperature or humidity of the lower troposphere (Beasley 1986; Jendrowski 1988; Breidenbach et al. 1993; Kitzmiller et al. 1995). For example, over the Central Plains, relatively shallow storms with low VIL are much more likely to be severe in early spring than in summer; severe storms over the northeastern U. S. usually feature lower VIL values than severe storms over the Plains.

The currently-operational SWP algorithm incorporates only a limited amount of radar information and no environmental information. Recent efforts have focused on incorporating both additional radar predictors and upper-air data such as temperature, wind velocity, and stability. Breidenbach et al. (1992; 1995) have documented efforts at refining the SWP algorithm by incorporating radar data other than VIL; this work resulted in experimental "second-generation" algorithms. Efforts at refining the SWP by incorporating environmental information yielded promising results (Breidenbach et al. 1993). We have now developed two "third-generation" algorithms (SWP3 and HAIL3) that incorporate both radar and environmental predictors, and yield the probability that either any severe weather phenomenon, or large hail specifically, will occur within a 44-km square region centered on the thunderstorm for the next 30 minutes. They can be implemented on any platform in which radar reflectivity data and mandatory-level upper-air data are available.

To implicitly account for regional differences in severe weather climatology over the United States east of the Rocky Mountains, two sets of severe weather and hail algorithms were developed, one for the Central Plains and one for the Northeast. Results indicate that the new severe weather algorithms can significantly improve on the performance of the operational SWP product. We will document the expected performance of the algorithms in terms of categorical severe/nonsevere nowcasts based on the probability values. Algorithm output for two examples is included.

2. OBSERVATIONAL AND MODEL DATA USED IN ALGORITHM DEVELOPMENT

The methods used to develop SWP3 and HAIL3 are similar to those employed in the development of the currently-operational SWP algorithm, as documented in Kitzmiller et al. (1995), hereafter referred to as KMS95. A sample of radar and storm environment observations was collected and then collated with nearby severe local storm reports. Equations relating severe storm occurrence (the predictand) to radar and environment data (the predictors) were then developed.

Radar data

The radar data for the Central Plains development sample was taken from Radar Data Processor Version II (RADAP II) archives collected at Amarillo, Texas (AMA), Wichita, Kansas (ICT), and Oklahoma City, Oklahoma (OKC), between 1985 and 1991. Typically, new volumetric scans were available every 10 or 12 minutes. The RADAP II archive has been described in detail by McDonald and Saffle (1994) and by KMS95. The Plains sample contains data on over 6000 individual thunderstorms.

Radar data for the Northeast development sample was from the RADAP II unit at Binghamton, New York, (BGM) between 1988 and 1992, and from the WSR-88D unit at Sterling, Virginia, (KLWX) during 1992 and 1993. We obtained only VIL graphic images from WSR-88D; these were manually interpreted for cell locations, peak VIL values, and VIL horizontal coverage. This data sample contains data on nearly 700 storms.

Environmental data

Upper-air conditions were derived from analyses and 6-h and 12-h forecasts of the Nested Grid Model (NGM) (Hoke et al. 1989) archived by the Techniques Development Laboratory. These data were chosen in preference to radiosonde observations because they are readily available in gridded form and represent a robust estimate of atmospheric conditions at asynoptic times.

To objectively assign environmental conditions to individual storm cells, the environment was assumed to be constant over each 230-km radius radar umbrella, with values corresponding to those at the center. The NGM data were available from Techniques Development Laboratory archives at 6-h intervals: initial-time analyses at 0000 and 1200 UTC, and 6-h forecasts at 0600 and 1800 UTC. For radar data within one hour of an analysis or forecast valid time, the values were taken at that time. For data outside these 2-h windows, the conditions most favorable for strong convection (higher instability, humidity, wind speeds) at the bracketing valid times were used. Temperature, humidity, stability, and wind data were derived from 0000 or 1200 UTC analyses or 6-h forecasts; vertical velocity and divergence predictors were derived from 6- and 12-h forecasts (the initial fields are quasi-nondivergent).

Severe local storm reports

Reports of high convective wind gusts (those causing damage or measured in excess of 50 kt), large hail (2 cm or greater in diameter), and tornadoes are logged by the National Severe Storms Forecast Center (NSSFC). We collated these reports with storm cell data by mapping the reports to the VIL analysis

grid. For each cell, the number of all severe reports, the number of large hail reports, and the largest reported hail diameter were noted.

3. CREATION OF A STATISTICAL DEVELOPMENT SAMPLE FOR SWP3 AND HAIL3

The statistical data sample was derived by objectively interpolating radar reflectivity and VIL data to a 4-km cartesian grid centered on the radar (the 4-km analysis VIL is available as a graphic product in the WSR-88D suite). Thunderstorm centroids appear as local maxima in the VIL field. A storm "cell" is considered to be a region 28 km square (7 grid boxes).

The collection of radar images described above contained numerous individual storms, and many that were close to each other in both space and time. In order to limit the statistical sample to fairly active storms and to maintain a degree of statistical independence, a number of restrictions were placed on storm size and proximity. Only storm cells with at least two grid boxes with a VIL of 10 kg m^{-2} or more were considered for inclusion. Any two cells in the final dataset were separated from each other by at least 28 km, or by 45 minutes in time. Where spatial or temporal overlaps were found, only the larger cell (in terms of maximum VIL) was included.

The final development dataset contains storm cell chacteristics including maximum VIL, the number of map grid boxes with VIL in excess of 10 and 20 kg m^{-2}, and the operational SWP value. For cases derived from RADAP II data, additional radar predictors including the partial VIL above the freezing level and above 15 kft MSL, and the heights of the 40- and 50-dBZ echo tops, were also noted. It was not possible to derive these additional predictors from the WSR-88D VIL graphic images from the LWX site.

The severe storm report log was then examined to locate any severe weather or hail events reported near the storms, and the number of associated severe storm and hail reports were recorded. A storm was considered to be severe (or a large hail producer) if at least one report (or one large hail report) occurred from 10 minutes before to 30 minutes after the nominal radar observation time. This convention should account for both events in progress and those about to develop.

These collection procedures yielded a Plains data sample of 6068 cells, of which 8% were severe and 5.5% featured large hail. The Northeast sample consisted of 668 cells, of which 20% were severe and 6% featured large hail. Thus most severe events in the Plains involved large hail, while over the Northeast, damaging wind events were predominant. These features of severe weather climatology might be due partly to physics and partly to land development patterns. Many reported wind events in the mid-Atlantic region are associated with falling trees and damage to structures. The scarcity of wind events over the Plains could be due in part to a relatively low density of construction even near population centers and to sparse forestation.

4. INDIVIDUAL RADAR AND ENVIRONMENTAL INDICES AS PREDICTORS OF SEVERE WEATHER

Table 1 contains a complete list of the candidate predictors included in statistical tests and regression procedures used to develop SWP3 and HAIL3. Most are commonly known; explanations for the others follow here.

Partial VIL indices differ from "total" VIL in that the integration of the radar-estimated volumetric rainwater mixing ratio is carried out only above some reference level above the nominal terrain height. As shown by Breidenbach et al. (1995), the partial VIL above 15 kft MSL (PVIL15) is less affected by range from the radar than is VIL, and PVIL15 has almost as much information with respect to severe weather. The partial VIL above the freezing level (PVILFR) should logically be correlated with large hail, since supercooled water is crucial to the hail formation process. "VIL size" predictors (SVG10, SVG20) represent storm areal extent, in terms of the number of analysis grid boxes covered by VIL in excess of 10 and 20 kg m^{-2}.

Most of the environmental predictors tested here have already been used in assessing the severe weather threat on the meso- and synoptic scales. That is, the predictors are correlated to the chance that at least one severe storm will develop within a region as large as several thousand square miles. We found that some of these indices are also correlated to the probability that an individual storm will produce severe weather, independent of any radar information. The Vertical Totals and Total Totals indices (Miller 1972) are calculated from mandatory-level radiosonde data and have long been used in assessing the environmental potential for severe weather outbreaks. The "surface" variations of these indices and of the K index (George 1954) are calculated by replacing the 850-mb temperature and/or dew point by the surface values; we approximated the surface values by data from the model's nominal 1000-mb level. While this level is sometimes below the true terrain surface, the derived indices still realistically reflect the model forecast of low-level stability conditions. The 1000-mb height is an important predictor of severe storm potential in operational Model Output Statistics guidance (Reap and Foster 1979).

A number of indices based on wind velocity and vertical wind profiles are known to be useful predictors of storm intensity. Wind speeds at and above 500 mb are correlated with both severe storm and large hail potential (Reap and Foster 1979; Doswell 1980; Dessens 1986). Veering of the wind with height is also a characteristic of severe storm conditions; the thermal advection index (Kitzmiller and McGovern 1989) is proportional to clockwise vertical wind shear and overall wind speeds in the 850-700 mb layer.

Certain indices (freezing level height, 500-mb temperature, and 1000-500 mb thickness) were included primarily as large hail indicators; experience and statistical analyses have shown that thunderstorms developing in particularly cold environments are especially likely to produce large hail (Wagenmaker 1992; Kitzmiller and Breidenbach 1993).

Fig. 1 shows the correlation ratio (Panofsky and Brier 1968) between some of the candidate radar and environmental predictors and severe weather occurrence. This correlation statistic is essentially the fraction of the predictand variance explained by the predictor without any a priori assumptions about the nature of the predictor/predictand relationship. If this relationship happens to be nearly linear, then the correlation ratio approaches the square of the linear correlation coefficient between the variables. Here, the predictand is binary (0 for nonsevere cells, 1 for cells with one or more severe reports).

As might be expected, the radar-based predictors possess more information than do the environmental predictors. The radar observes the morphology of

individual storms, and thus yields highly specific information about a small region near those storms. The environmental data indicates primarily the likelihood that any one of a number of storms may produce severe phenomena. However, this information may be substantially independent of the radar data, and thus provides some knowledge about the situation that the radar does not.

For the Plains data sample (Fig. 1a), the best individual predictors all involved VIL, including partial VIL and the operational SWP algorithm. Each of these predictors explained more than 20% of the predictand variance. Other commonly-used radar indices, including the 40- and 50-dBZ echo top height, explained between 15 and 20% of the predictand variance. The best environment indices, such as surface Total Totals and the 500-mb u-wind component, explained about 6% of the variance. This is still physically significant, considering that no data on individual storms is incorporated in these indices.

Results similar to those for the Plains were obtained by testing the Northeast sample (Fig. 1b). Only VIL-based radar predictors were used here, since other volumetric radar indices could not be derived from the LWX VIL graphic products. Among the environmental predictors, the 700- and 500-mb wind velocity were more informative than stability. This might be due to the fact that most severe events in the East involve damaging winds, and that convective transfer of momentum from the mid-troposphere to the surface might be an important mechanism in driving these events.

The relative value of the radar predictors with respect to hail occurrence (Fig. 2) is similar to their value with regard to general severe weather, though the correlation ratios are smaller. The best environmental predictors, in terms of the correlation ratio, now involve stability or temperature. This might reflect the increased tendency of storms in especially cold, unstable environments to create large hail, and for hailstones to experience less melting after leaving the storm updraft.

These and earlier results (KMS95) show that VIL indices have information with respect to both hail and wind events. It is probable that VIL is correlated to hail because hailstorms have especially large volumes of supercooled liquid water, and because wet hailstones themselves are highly effective radar backscatterers. At the same time, intense convective downdrafts are often initiated by fallout of suspended liquid water, thus storms with relatively high VIL are those most likely to cause damaging surface winds (Srivastava 1985; Wolfson 1990).

As noted in KMS95, VIL is not an effective indicator for most tornadic storms. Even storms in larger tornado outbreaks often feature rather low VIL values; it appears that the organization of the rotating updraft in tornadic storms does not necessarily cause a deep core of high reflectivity. Thus we believe that forecasters should rely primarily on indications of environmental vertical wind shear and mesocyclones in attempting to detect tornadoes.

The amount of information contained in the better radar-based predictors is well represented by severe weather and hail relative frequency charts, as shown in Fig. 3. In both data samples, storms with VIL values less than 20 kg m^{-2} feature less than half the climatic relative frequency of either large hail or general severe weather; for storms with VIL in excess of

50 kg m^{-2}, severe weather is likely, and a substantial number of the storms feature hail.

Further examination of the VIL/severe weather relationship for the Plains showed that it is nonlinear (Fig. 4); the severe weather relative frequency can be more closely approximated by the square of the VIL than by VIL itself. Accordingly, we submitted the square of VIL as a candidate predictor in the screening regression procedures later used to derive SWP3 and HAIL3 in their final forms.

Similar plots of severe storm and hail relative frequency as functions of environmental parameters over the Plains appear in Fig. 5. Note that the only radar information explicitly incorporated in these nowcasting relationships is the minimum VIL and size criteria. Yet, the environmental information alone does contain significant information on the likelihood that individual storms will produce severe weather. The 500-mb temperature (Fig. 5a) can be correlated with hail probabilities as low as 1% and as high as 25%, even in the absence of detailed radar information. In a similar manner, hail potential is inversely correlated to freezing level height (Fig. 5b). Finally, the "surface" total totals index is fairly highly correlated to both severe weather and hail probability over the Plains (Fig. 5c).

The single greatest environmental influence on individual storm severity in the Northest appeared to be mid-tropospheric wind speed, in particular that at the 700-mb level (Fig. 6). In conditions with light winds (< 9 m s^{-1}), just under 10% of the storms produced severe phenomena, while for speeds in excess of 20 m s^{-1}, over 30% did. This could be an indication of convective transfer of horizontal momentum from the middle troposphere to the ground. Reap and Foster (1979) noted the correlation between high upper-air winds and conditional severe local storm probability. We again note that the optimum choice of predictors for the Northeast might change as more cases are added to the sample, but the present findings are consistent with Reap and Foster's earlier results.

Two-predictor histograms clearly illustrate that independent information is available from both radar and environmental data, as noted earlier by Breidenbach et al. (1993). We determined correlation ratios for all possible combinations of radar and environmental predictors. The histograms shown in Fig. 7 were those that explained the greatest percentage of predictand variance. General severe storm probability is shown as a function of VIL and freezing level height and as a function of SVG20 and 700-mb wind speed (Figs. 7a, 7b respectively). The influence of environmental factors is most apparent for moderately strong storms (VIL between 40 and 50, and SVG20 of 4 or more). Likewise, hail probability over the Plains is clearly dependent on both VIL and freezing level height (Fig. 7c); a similar relationship was evident in the Northeast data sample.

5. DERIVATION OF THE SWP3 AND HAIL3 ALGORITHMS

The correlation results documented in Section 4 guided us in our final derivation of the new algorithms, by indicating the predictors and predictor combinations that explained the most predictand variance. The final versions of the probability equations themselves were derived by statistical linear regression on the data.

While we had found that the partial VIL predictor PVILFR had more information than VIL itself, was also found that an interactive combination of VIL and freezing level height had as much information as did PVILFR. We decided to use VIL in preference to PVILFR because VIL is already available operationally. We also found that certain combinations of VIL and environmental predictors, specified a priori, explained more predictand variance than did a combination chosen by standard forward selection screening, in which the first predictor chosen is that with the highest linear correlation.

The relationships shown in Fig. 7 were utilized in constructing regression formulas relating the predictor variables to event probabilities. For example, in the Plains data sample, both severe storm and hail relative frequency are well-approximated by a biquadratic function of MAXVIL and freezing level (Kitzmiller and Breidenbach 1993). This function is a linear combination of terms in $MAXVIL^2$, MAXVIL, and the product of MAXVIL and freezing level.

The regression procedures yielded the following equations for SWP3 (general severe weather probability):

$$SWP3(\%) = -16.37 + 2.33\ SVG20 + 1.02\ WSPD_{700} + .646\ MAXVIL \quad (1)$$

for the Northeastern U. S., and:

$$SWP3(\%) = -16.49 + .025\ MAXVIL^2 - .00206\ (MAXVIL \times FRZLVL) + .365\ U\text{-}WIND_{500} + .341\ (SFC\ TOTAL\ TOTALS) \quad (2)$$

for the Plains. Here, SVG20 is the number of 4-km grid boxes with VIL > 20 kg m^{-2}, WSPD700 is the 700-mb wind speed in m s^{-1}, FRZLVL is the freezing level height in m MSL, and $U\text{-}WIND_{500}$ is the west-east 500-mb wind component in m s^{-1}. The Total Totals index predictor is in °C. The expression in (1) explains 18.6% of the predictand variance, and that in (2) explains 25.7%. In (2), most of the predictand variance is explained by the $MAXVIL^2$ and MAXVIL x FRZLVL terms. The forward-selection procedure itself selected the 700-mb windspeed predictor in (1), indicating the importance of strong mid-tropospheric winds in causing strong convective wind events over the eastern United States.

The following equations were derived for HAIL3 (probability of 2 cm hail):

$$HAIL3(\%) = 14.22 + .03\ MAXVIL^2 - .0031\ (MAXVIL \times FRZLVL) \quad (3)$$

for the Northeast, and:

$$HAIL3(\%) = -375.43 + .019\ MAXVIL^2 - .00619\ (MAXVIL \times FRZLVL) + 2.057\ MAXVIL + .066\ THICK_{1000\text{-}500} \quad (4)$$

for the Plains. Here, $THICK_{1000\text{-}500}$ is the 1000-500 mb thickness in m. The expression in (3) explains 22.4% of the predictand variance, and that in (4) explains 24.9%.

The similar selection of predictors in (3) and (4) indicate that hail-bearing storms over both the Plains and the Northeast have similar VIL signatures and that environmental temperature has an important influence in the hail

process. The similarity between (2) and (4) reflects the dominance of hail events within the Plains sample of severe weather reports.

6. GRAPHICAL PRESENTATION OF THE ALGORITHM OUTPUT

Figs. 8 and 9 illustrate a potential real-time display product featuring VIL and the probabilities obtained from Eqns. 1-4. The VIL analysis appears in the background; the event probabilities are plotted immediately above and below the peak VIL values within the larger storm cells. Severe storm reports from 10 minutes before to 30 minutes after the radar valid time are also plotted (W for wind, H for large hail) .

Fig. 8 shows the analysis for storm cells within an instability line extending from southeastern Pennsylvania through Maryland to central Virginia. Radar data were from the Sterling, Virginia, WSR-88D. The NGM 700-mb wind speed forecast over the umbrella at this time was approximately 17 m s^{-1} (~34 kt), and the freezing level approximately 360 dm MSL (11,800 ft). The most intense storms were in the northern half of the line; peak VIL values in these storms ranged from 35 to 50 kg m^{-2}. Two wind events and one hail event were associated with the Pennsylvania storms between 2030 and 2100 UTC. Wind and hail events were also reported in association with the storms over the District of Columbia and northern Virginia.

A summertime situation with much warmer conditions and lighter wind speeds is illustrated in Fig. 9. These late-evening storms were observed by the Operational Support Facility WSR-88D at Norman, Oklahoma. The NGM forecast indicated a freezing level height of 479 dm (15,700 ft), a 500-mb u-wind of 10.7 m s^{-1} (21 kt), and a surface Total Totals index of 68°C. However, VIL values in the more intense cells shown here were above 50 kg m^{-2}; the cell with a severe probability of 99 featured a VIL of 85. Five hail events were produced by the larger storms between 0415 and 0430 UTC; more large hail was reported to the southwest of these storms between 0420 and 0430, as the storms propagated in that direction. These storms continued to produce large hail for the next two hours.

7. RELIABILITY OF THE SWP3 AND HAIL3 ALGORITHMS

The probability equations (1) - (4) generally represent a good fit to the dependent data samples. To test their reliability within these samples, the equations were evaluated for each case (storm cell), and the verifying observations were averaged over each 10% probability range to see how closely the average forecasted probability approximated the actual event relative frequency. In Figs. 10 and 11, the observed relative frequency is plotted versus the mean forecasted probability in each 10% range. The dashed line represents perfect reliability, or zero probability bias. The number of cases in each probability category is also shown.

For both the Plains and Northeast SWP3 algorithms, the true severe weather probability is slightly lower than the forecasted value when the forecasts are less than 35% and above 60%; between these values SWP3 underforecasts the probability (Fig. 10). The Plains HAIL3 equation has little systematic bias overall and no serious misrepresentation of the probabilities for forecasts below 70% (Fig. 11a). The Northeast hail algorithm yields forecasts of greater than 30% in less than 10% of the cases, and at these higher values there are some marked departures from zero bias (Fig. 11b). It is likely that

the coefficients in the Northeast HAIL3 equation would change if a substantial number of new cases were added to the development sample.

8. SKILL OF THE ALGORITHMS IN TERMS OF CATEGORICAL FORECASTS

Though SWP3 and HAIL3 provide probabilistic guidance, their performance is most easily evaluated by examining categorical (severe/nonsevere) forecasts based on the probabilities. Categorical forecasts are generally derived by setting some fixed threshold probability value, and forecasting all storm cells with probabilities at or above the threshold to be severe. All other cells are assumed to be nonsevere. This verification exercise is most useful when a range of possible thresholds, from low to fairly high, are examined.

The performance of these forecasts may be described by three commonly-used measures, the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) (Donaldson et al. 1975; Schaefer 1990). Let x be the number of severe events correctly forecasted to be severe, w be the number of nonsevere events correctly forecasted, z the number of nonsevere events incorrectly forecasted to be severe, and y the number of severe events incorrectly forecasted to be nonsevere. We may refer to x as the number of "hits," z as the number of "false alarms," and y as the number of "missed events." The POD is the percentage of all severe events correctly forecasted to be severe:

$$POD = x / (x + y). \tag{5}$$

The FAR is the percentage of severe forecasts that are false alarms:

$$FAR = z / (x + z). \tag{6}$$

The CSI is the percentage of all severe cases and severe forecasts that are "hits:"

$$CSI = x / (x + y + z). \tag{7}$$

Both the POD and FAR decrease if the severe/nonsevere probability threshold is lowered. For rare events such as severe local storms, the CSI reaches a peak value near thresholds that yield neither too low a POD nor too high an FAR.

The performance of SWP3 in terms of POD, FAR, and CSI for a range of severe/nonsevere thresholds between 1% and 50% is shown in Fig. 12. As might be expected from the distribution of forecast probability values, SWP3 does not have high skill in the absolute sense. That is, for thresholds yielding a fairly high POD (> 0.7), the majority of severe forecasts are false alarms (FAR > 0.5). The skill of the Northeast algorithm appears higher than that of the Plains algorithm, largely because a substantially higher fraction of the Northeast are severe. Thus for any given threshold up to about 40%, the Northeast SWP3 algorithm (Fig. 12b) yields a higher POD, and about same FAR, as does the Plains algorithm (Fig. 12a).

Because these forecast scores were derived from the dependent data sample, and the predictor selection partly reflects statistical relationships peculiar to that sample, the scores probably reflect a higher degree of skill than would be realized in an indepdent set of cases. To improve our skill esti-

mates, we derived another set of trial forecasts and scored them through a cross-validation procedure (Elsner and Schmertmann 1994). In classic cross validation, prototype forecast algorithms are derived from all except one case in the data sample, and the resulting prototype is then verified against the case withheld. This process is repeated until all cases have been verified against a prototype algorithm derived from independent cases.

However, this classic approach implicitly assumes that all cases are statistically independent. While our dataset features only distinct storms, it does contain many storms observed within hours of each other, under similar environmental conditions. We therefore divided the Northeast data into 35 subsamples, and the Plains data into 121 subsamples, with the latest observation in one subsample separated by at least 3 days from the earliest observation in the next. The sizes of the subsamples ranged from one storm cell to as many as 150. We carried out the procedure by withholding entire subsamples, one at a time.

This practice does not insure complete statistical independence of the subsamples, since particular convective regimes may persist for over a month. It does insure that identical environmental conditions, and storms within the same convective outbreak, do not appear in both the development and verification samples. The use of fewer and larger subsamples, for example full calendar years, did not seem justified because it would significantly reduce the size of the development samples, leading to underestimation of skill.

As shown in Fig. 13, the skill of the algorithms when applied to nominally independent data decreases measureably. For any given yes/no probability threshold, the cross-validation method indicates POD values 1-5 percentage points lower, and FAR values slightly higher, than were indicated by verifying within the dependent data.

The FAR curves in Fig. 13 show that FAR decreases below 0.5 near an SWP3 value of 30. Thus the distribution of probability values is such that the majority of storms are severe when SWP3 is greater than 30 over the Plains, and greater than 35 for the Northeast. In both cases, approximately 40% of the severe storms have SWP3 values above these thresholds.

The skill levels of the Northeast and Plains HAIL3 algorithms are very similar to each other, with the Plains algorithm generally giving a slightly lower FAR than the Northeast algorithm for any one POD (Fig. 14). These scores were derived by the cross-validation method described above. At higher thresholds, it sometimes happens that the FAR actually increases as the threshold probability increases (as shown for thresholds of 37% and 40% in Fig. 14b). This occurs when the increase in the threshold value causes the number of hits to decrease more rapidly than number of false alarms.

The distribution of HAIL3 values is such that for all Plains storms with HAIL3 above 35, over half will produce large hail. Similarly, about half of all Northeast storms with HAIL3 above 37 will produce large hail. Approximately 30% of storms that produce large hail have HAIL3 values above these thresholds.

9. DIFFERENCES BETWEEN THE SWP3 ALGORITHM AND OPERATIONAL SWP

There are several important differences between the development methodologies and performance characteristics of the operational SWP and the new SWP3 algorithm. As noted above, SWP3 utilizes environmental data (specifically, freezing level height, 1000-500 mb layer thickness, surface Total Totals index, and the 700- and 500-mb wind vectors) as well as radar information to determine severe weather potential. As will be shown below, the incorporation of environmental data yields an operationally significant increase in the skill over that available from radar data alone.

Also, the definition of the "valid area" for the output probability has been changed. For SWP, this area was within roughly 20 km of the track of the storm centroid during the 20-minute period after the start of the volume scan. Much of the radar and severe storm data were collated manually. In the development of SWP3, the verification zone is a region 44-km square, centered on the storm cell's current position. This alternative definition was chosen to simplify the process of identifying threat regions relative to county boundaries or geographic features. The storm is considered to be severe if an event was reported in this region from 10 minutes before to 30 minutes after the nominal time of the radar volume scan. All radar and storm observations were matched objectively.

Because the new verification zone does not move with the storms, it is effectively smaller than the one defined for SWP. Despite the somewhat longer lead time accounted for, the percentage of cells featuring severe weather is lower than that within the statistical sample used to develop SWP.

Finally, we note that SWP3 and HAIL3 provide event probabilities, while SWP is an index proportional to severe weather probability. The new algorithms were specifically tailored to the climatology of specific regions, and thus should yield reliable event probabilities as long as storm reporting procedures do not change significantly.

Comparative verification shows that the skill of SWP3 is indeed higher than that of SWP. We noted earlier that our goal was to reduce the total amount of forecaster time that must be spent in examining storms in detail, by either radar or other means, in order to determine if severe weather warnings are warranted. To demonstrate the improvement, we determined the number of false alarms that would be issued by both SWP and SWP3 in the process of achieving the same POD value. As shown in Fig. 15a, where the comparison was made within the Plains developmental data sample, the SWP3 algorithm consistently yielded 15-20% fewer false alarms than did a radar-only algorithm. The improvement was less for the Northeast sample (Fig. 15b), where the SWP3 algorithm offered clearly better skill only for POD values above 0.75.

Because this comparison was made within the development sample, where the forward-selection procedure for additional predictors insures improvement in hindcasting, the degree of improvement offered by the additional environmental predictors is probably overestimated. To better assess the amount of improvement that would likely be realized within independent data, we applied the cross-validation technique outlined above to prototype operational SWP algorithms (using only radar predictors). The "operational" prototypes were derived by using only the VIL predictors currently available. These five predictors are: size of VIL greater than 10, 15, 20, and 25 kg m^{-2}, and VIL-

weight, which is the product of maximum VIL and the size of VIL greater than or equal to 10 kg m^{-2}. As before, the probabilities were converted to categorical forecasts, by using a range of yes/no thresholds.

We again present the results by showing the FAR associated with fixed POD values for both radar-only and radar/environment equations. We should logically expect that the more skillful algorithm will still produce fewer false alarms for values of the POD above 0.5. We found that SWP3 does generally yield fewer false alarms, though its improvement over SWP is considerably smaller than within the dependent data (Fig. 16).

For the Plains sample (Fig. 16a), both algorithms yield about the same number of false alarms until the POD criterion is relaxed to about 0.65. This implies that the incorporation of environmental data most consistently adds information mainly where the operational SWP algorithm would indicate probabilities between 25 and 35%. Improvement to lower probabilities is not consistent, while improvement to higher probabilities becomes important only at much higher yes/no thresholds (and lower POD values).

For the Northeast sample (Fig. 16b), the results suggest that the incorporation of environmental data most consistently improves forecasts at POD values above 0.75, or yes/no threshold probabilities below 25%. Most of the prototype equations derived from the Northeast data featured 500- or 700-mb wind speed as a predictor, and the upper-air winds appear to be the most important consideration in determining whether moderately-strong storms over this region produce strong surface wind gusts.

Though the operational WSR-88D product suite contains a hail detection algorithm, it was intended as an indicator of hail of <u>any</u> size reaching the surface. Also, this product could not be reliably duplicated with RADAP II data. Therefore, we did not attempt to compare it to HAIL3.

10. CONCLUSIONS AND IMPLICATIONS FOR OPERATIONAL USE OF SWP3 AND HAIL3

We have found that the incorporation of environmental data should result in an operationally significant improvement to the SWP algorithm currently available as a WSR-88D interpretive product. Since VIL itself is already widely used as a general severe weather indicator, we hope that our findings on VIL/storm environment/severe weather relationships may also be useful to forecasters who do not have access to data acquisition/display systems in which both radar and environment data are available in real time.

However, the improvement appear significantly larger in the dependent data than in a set of independent cases. This implies that more data should be collected and analyzed, and that other radar data, particularly Doppler velocity, should be considered in further refining both hail and high wind detection algorithms. Because of operational limitations on the acquisition of Doppler data in storm regions, such as range folding and beam overshooting of the subcloud layer, it is likely that examination of reflectivity patterns will continue to be an important aspect of severe weather detection for some time.

As noted above, these techniques do not possess high absolute accuracy in identifying severe storms; that is, a high probability of detection is associated with a high false alarm rate. Thus the algorithms are intended

primarily to alert forecasters to sudden or unexpected severe storm development. Other considerations, such as three-dimensional storm structure, storm motion, and real-time spotter reports must be used to decide which storms actually warrant warnings, and where the warnings should be valid. At the same time, forecasters can be confident that storms with very high SWP3 values (70% or more) probably warrant issuance of a warning without further examination. Meanwhile, the vast majority of storms are assigned very low probabilities (< 5%), and these are very unlikely to be severe within the next 30 minutes.

For specification of large hail in severe weather warnings, absolute skill is again rather low. Forecasters can be confident that storms with probabilities in excess of 50% will generally produce hail shortly, and may wish to specifically mention hail as a threat in statements to the public.

We have used NGM data as a robust source of upper-air information within these equations. While any operational implementation of SWP3 and HAIL3 should include a facility for forecaster updates of the environmental data, it should be noted that the algorithms include only predictors that have a fairly broad spatial structure function, and thus do not change quickly with time. It is necessary to correct only gross errors in the environmental values.

ACKNOWLEDGEMENTS

We are indebted to Robert Saffle and Wayne McGovern, both formerly employed in the Mesoscale Techniques Branch of the Techniques Development Laboratory, for their expertise and their support of this work. Melvina McDonald developed the RADAP II data archive used here. Manual reduction of the WSR-88D VIL graphic images was carried out expertly by Bryon Lawrence and Mary Scarzello.

REFERENCES

Beasley, R. A., 1986: An analysis of operational RADAP II parameters, corresponding synoptic variables, and concurrent severe weather events in Oklahoma. M. S. Thesis, Univ. of Oklahoma, 223 pp.

Breidenbach, J. P., D. H. Kitzmiller, R. E. Saffle, and Wayne E. McGovern, 1992: Evaluation of new radar predictors and their use in severe weather potential algorithms. Preprints 26th International Conference on Radar Meteorology, Norman, Amer. Meteor Soc., 163-166.

_____, D. H. Kitzmiller, and R. E. Saffle 1993: Joint relationships between severe local storm occurrence and radar-derived and environmental variables. Preprints 13th Conference on Weather Analysis and Forecasting, Vienna, Virginia, Amer. Meteor Soc., 588-591.

_____, _____, and _____ 1995: The use of second-generation radar reflectivity predictors in the development of a second-generation severe weather potential algorithm. Wea. Forecasting, 10, 369-379.

Dessens, J., 1986: Hail in southwestern France I: Hailfall characteristics and hailstorm environment. J. Appl. Meteor., 25, 35-47.

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

Doswell, C. A. III, 1980: Synoptic-scale environments associated with High Plains severe thunderstorms. Bull. Amer. Meteor. Soc., 61, 1388-1400.

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

George, J. J., 1954: Weather Forecasting for Aeronautics. Academic Press, 673 pp.

Hoke, J. E., N. A. Phillips, G. J. DiMego, J. J. Tucillo, and J. G. Sela, 1989: The regional analysis and forecast system of the National Meteorological Center. Wea. Forecasting, 4, 323-334.

Jendrowski, P. A., 1988: Regionalization of the NEXRAD Severe Weather Probability algorithm. Preprints 15th Conference on Severe Local Storms, Baltimore, Amer. Meteor. Soc., 205-208.

Kitzmiller, D. H., and W. E. McGovern, 1989: Wind profiler observations preceding outbreaks of large hail over northeastern Colorado. Wea. Forecasting, 5, 78-88.

_____, _____, and R. E. Saffle, 1995: The WSR-88D Severe Weather Potential Algorithm. Wea. Forecasting, 10, 141-159.

_____, and J. P. Breidenbach, 1993: Probabilistic nowcasts of large hail based on volumetric reflectivity and storm environment characteristics. Preprints 26th International Conference on Radar Meteorology, Norman, Amer. Meteor. Soc., 157-159.

Klazura, G. E., and D. A. Imy, 1993: A description of the initial set of analysis products available from the NEXRAD WSR-88D system. Bull. Amer. Meteor. Soc., 74, 1293-1311.

McDonald, M., and R. E. Saffle, 1994: Revised RADAP II archive data user's guide. TDL Office Note 94-2, National Weather Service, NOAA, U.S. Department of Commerce, 18 pp. [Available from Techniques Development Laboratory, W/OSD2, National Weather Service, 1325 East West Highway, Silver Spring, Md.]

Miller, R. C., 1972: Notes on analysis and severe storm forecasting procedures of the Air Force Global Weather Center. Air Weather Service Tech. Rep. 200 (Rev.), 102 pp. [NTIS AD 744 042].

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

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

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

Srivastava, R. C., 1985: A simple model of evaporatively driven downdraft: Application to a microburst downdraft. J. Atmos. Sci., 42, 1004-1023.

Wagenmaker, R. B., 1992: Operational detection of hail by radar using heights of VIP-5 reflectivity echoes. National Weather Digest, 17, Number 2, 2-15.

Wolfson, M. M., 1990: Understanding and predicting microbursts. Preprints 16th Conference on Severe Local Storms, Kananaskis Park, Amer. Meteor. Soc., 340-351.

Table 1. List of candidate severe local storm predictors utilized in this study.

Radar Predictors:

Storm maximum VIL (MAXVIL)
Number grid boxes with VIL > 10 kg m^{-2} (SVG10)
Number grid boxes with VIL > 20 kg m^{-2} (SVG20)
Operational SWP output
Partial VIL above 15 kft MSL (PVIL15)
Partial VIL above freezing level (PVILFR)
Height of 40-dBZ level
Height of 50-dBZ level
Maximum reflectivity above 24 kft MSL

Environmental predictors:

Temperature at 1000, 850 mb
Dew point at 1000, 850 mb
K index, surface K index
Vertical Totals index
Total Totals index, surface Total Totals index
1000-500 mb lifted index
700-500 mb temperature lapse rate
Surface-500 mb mean RH
Precipitable water
Freezing level height (FRZLVL)
Freezing level pressure
1000-mb height
1000-500 mb thickness ($THICK_{1000-500}$)
Wind speed at 850, 700, 500, 300 mb
U- and V-winds at 850, 700, 500 mb
Divergence at 850, 300 mb
Moisture divergence at 850 mb
850-500 mb thermal advection index
Vertical velocity at 950, 850, 700, 500 mb

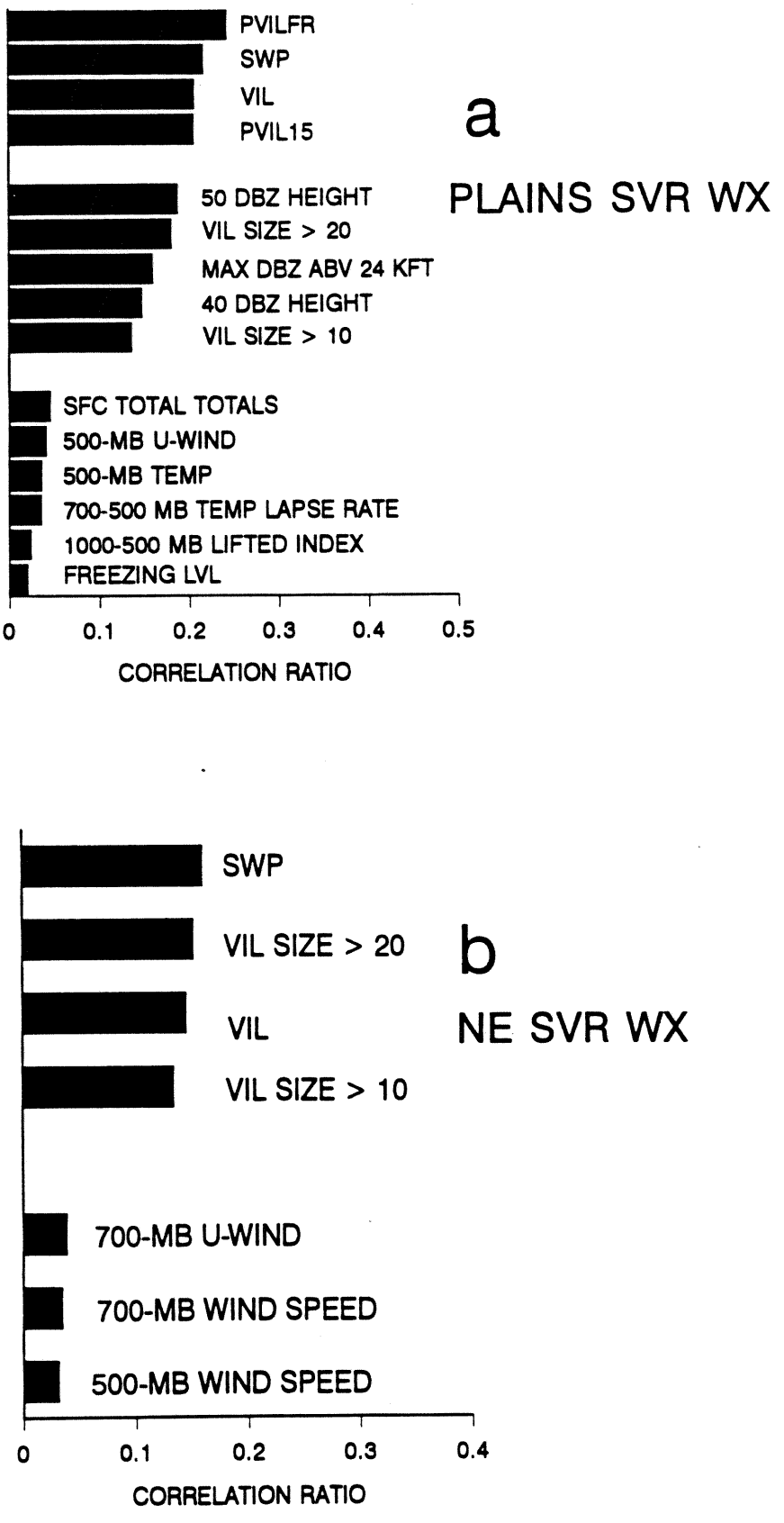
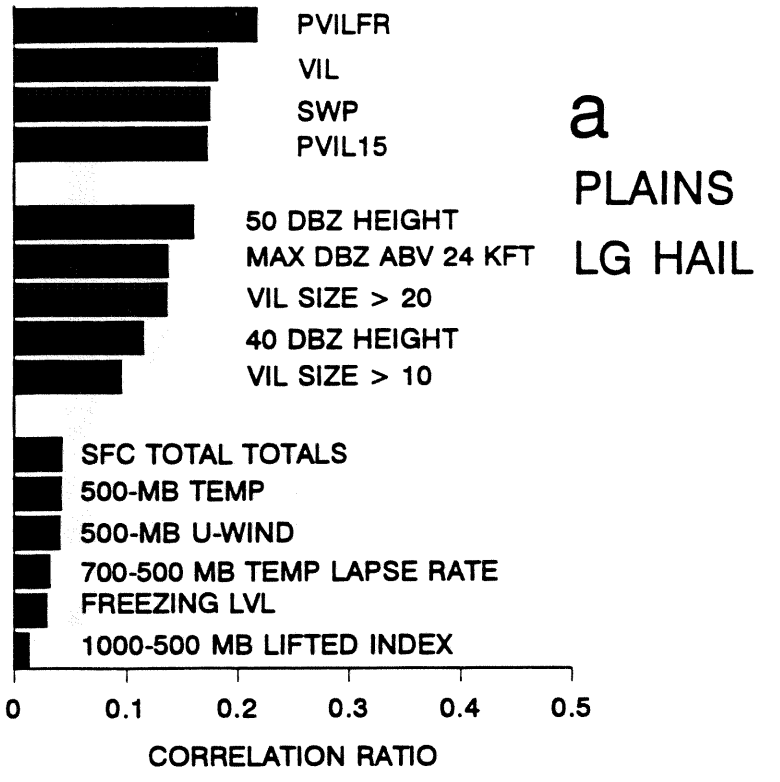


Figure 1. Correlation ratios with respect to general severe weather for various radar and environmental predictors. Statistics were derived from (a) the Central Plains sample and (b) from the Northeast sample.



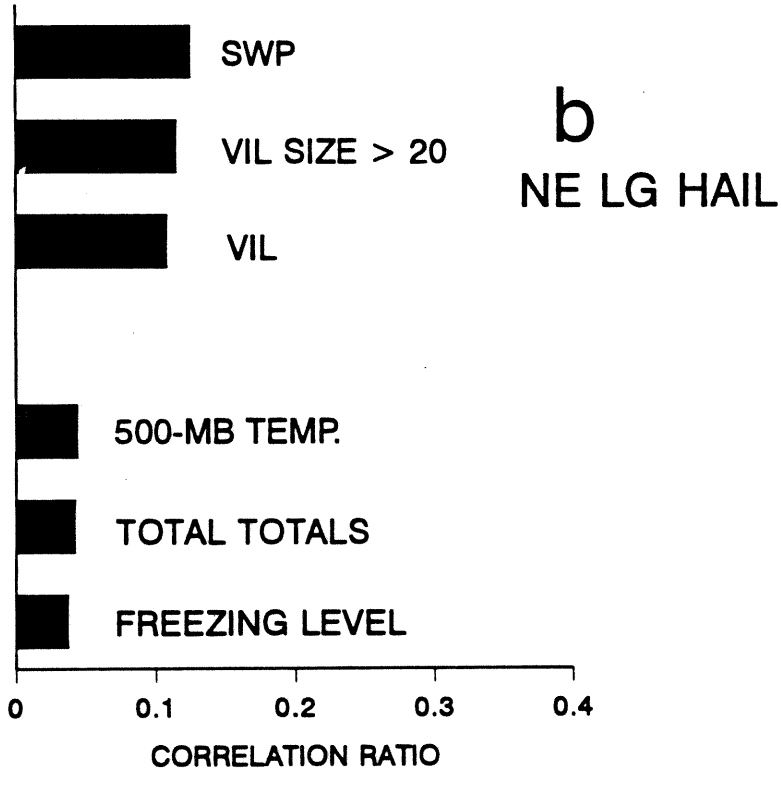
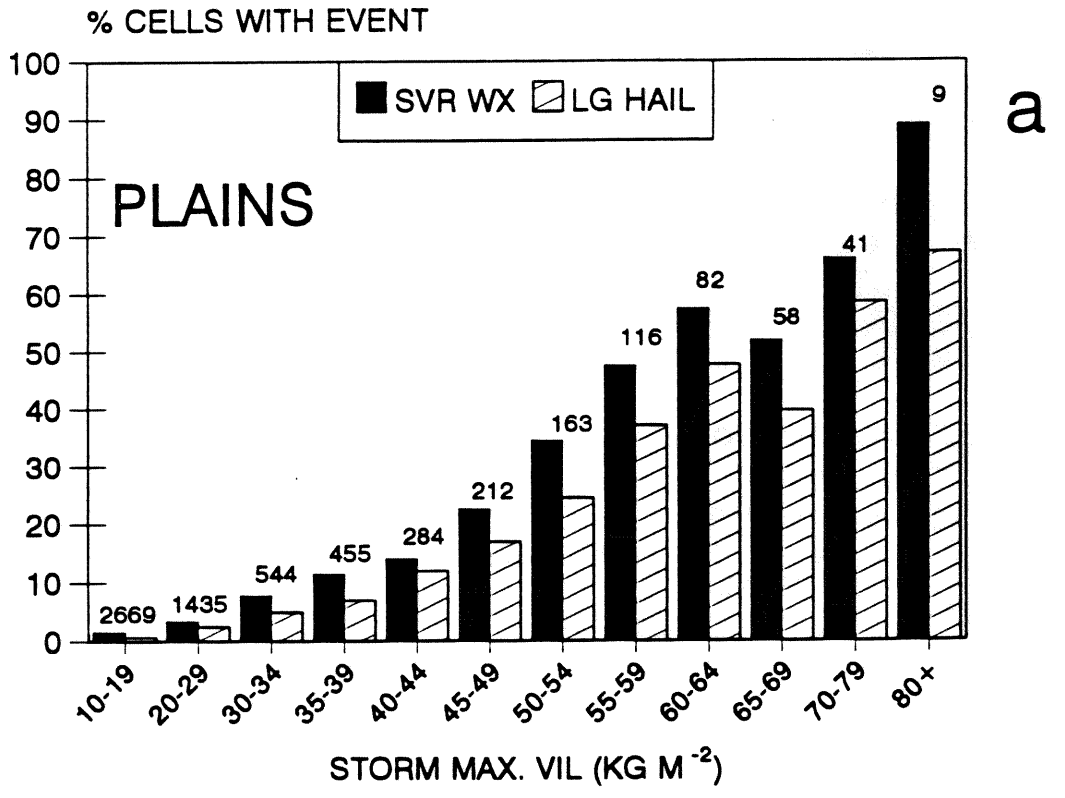


Figure 2. As in Fig. 1, except that the predictand is large (> 2 cm diameter) hail reports.



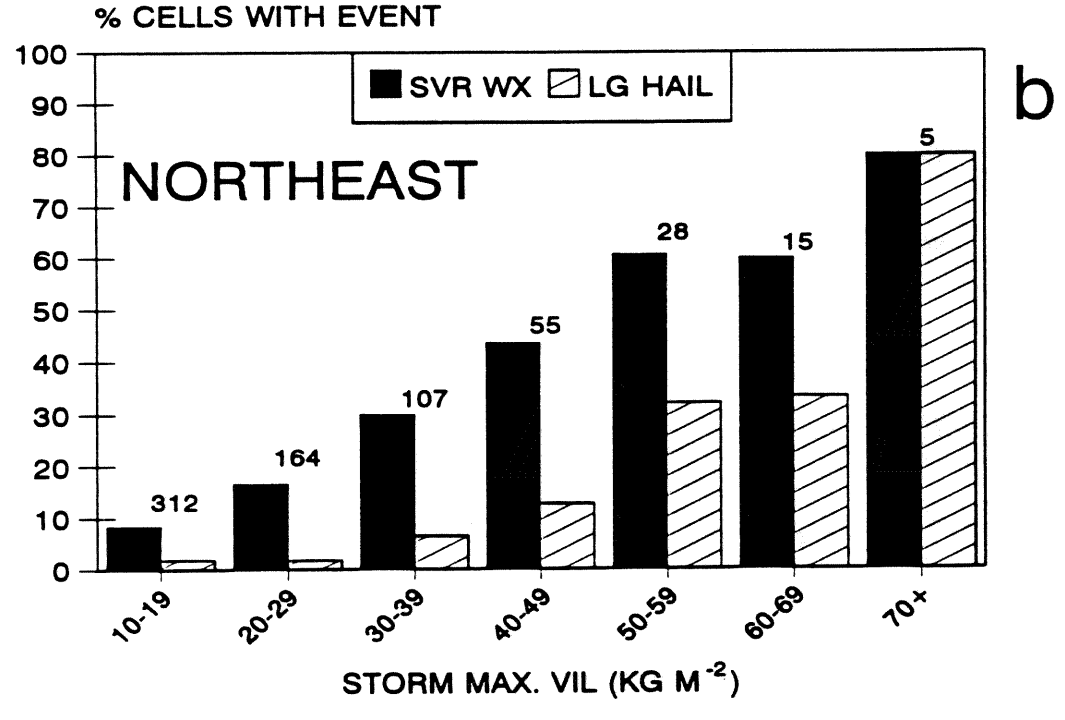


Figure 3. Percentage of storm cells with reports of general severe weather (solid bars) and large hail (hatched bars), as a function of storm maximum VIL. Data in (a) are from the Central Plains, data in (b) from the Northeast. Numbers above the bars indicate the number of storm cells in that VIL category.

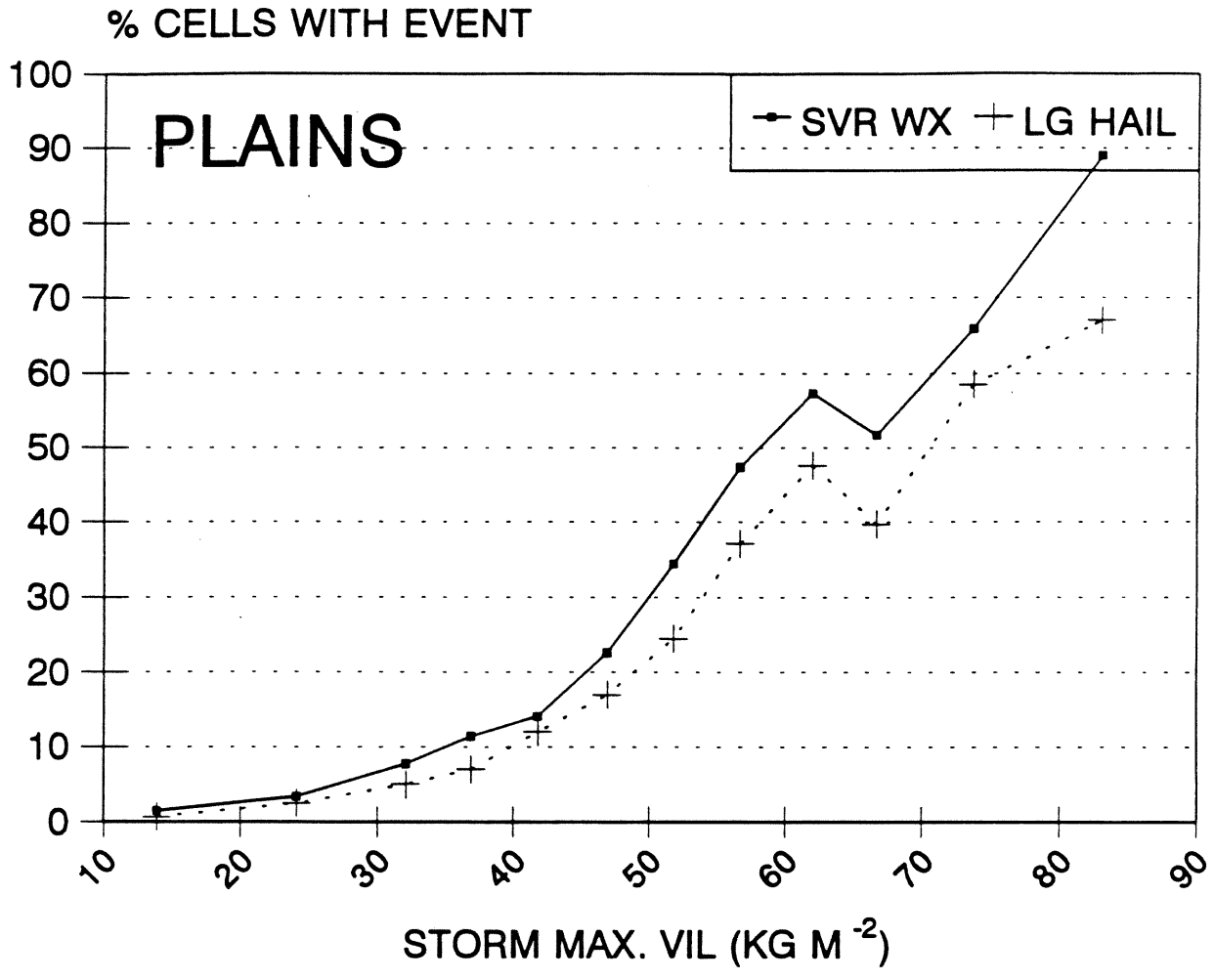


Figure 4. Percentage of storm cells with reports of general severe weather (solid line) and large hail (dashed line), as a function of storm maximum VIL over the Plains.

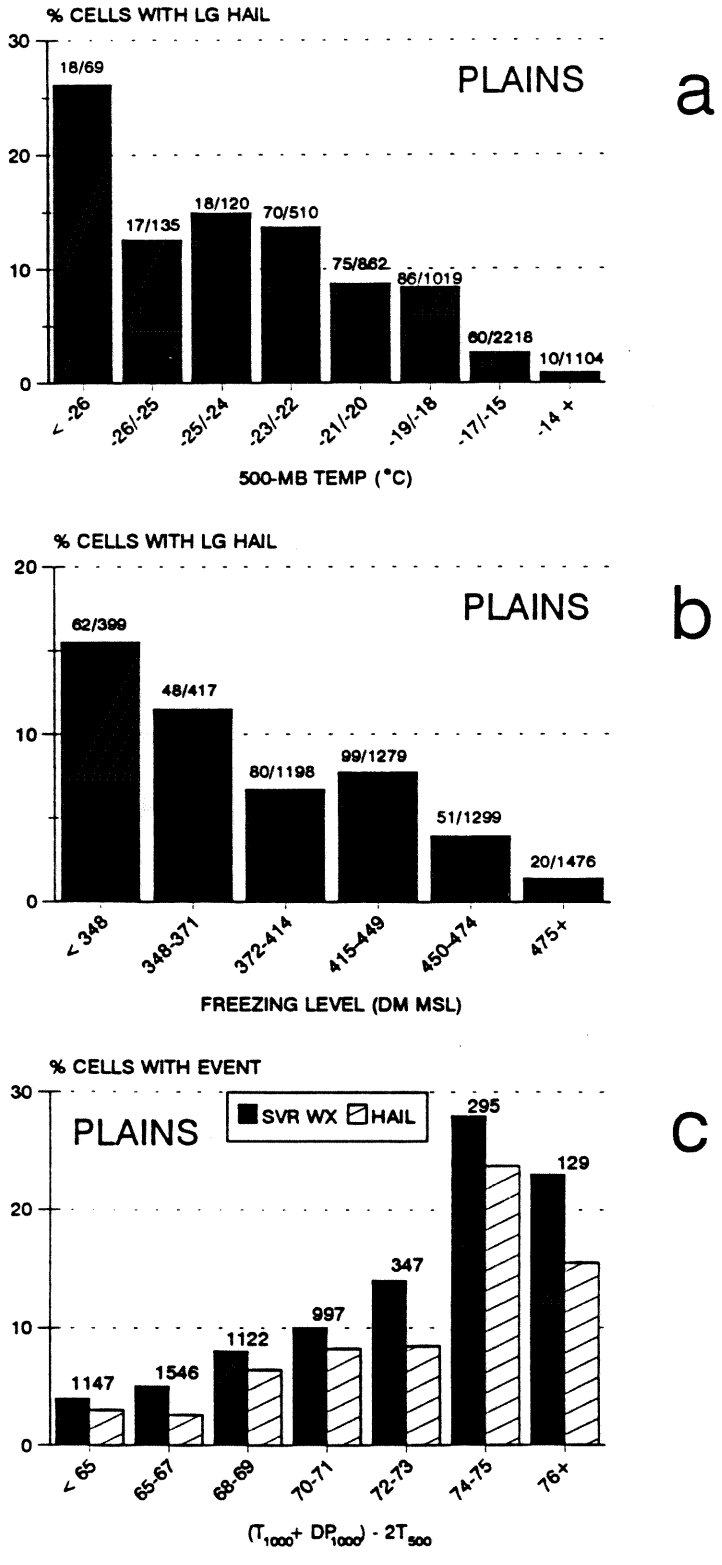


Figure 5. Percentage of storm cells with reports of large hail or severe weather as functions of environmental variables over the Plains. Environmental predictors are (a) 500-mb temperature, (b) freezing level height, and (c) surface Total Totals index. In (a) and (b), numbers over the bars indicate the number of cells with large hail and the total number of cells in each category; in (c), numbers indicate the number of cells in the category.

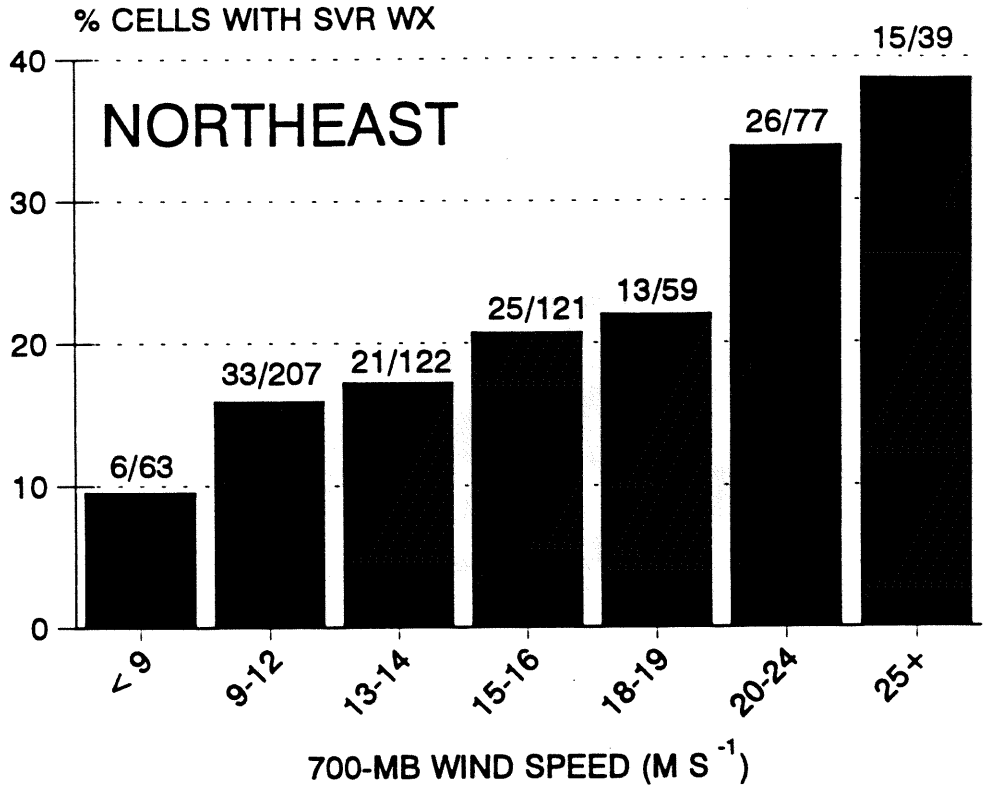


Figure 6. As in Fig. 5, except that the environmental predictor is 700-mb wind speed. Data are from the Northeast.

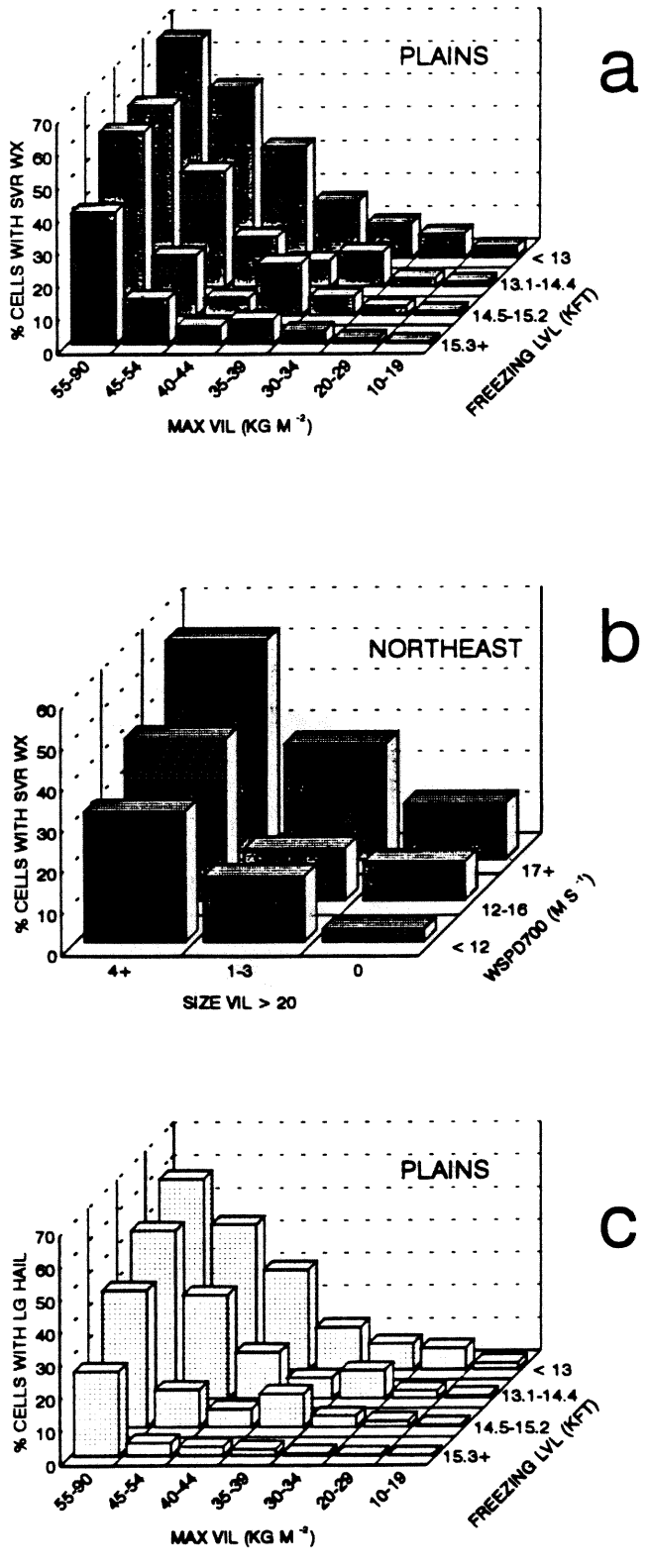


Figure 7. Percentage of storm cells with severe weather or large hail as a function of VIL and environmental data considered simultaneously. In (a) the predictors are maximum VIL and freezing level height, predictand is severe weather; in (b) predictors are size of VIL > 20 and 700-mb wind speed, predictand is severe weather; in (c) predictors are maximum VIL and freezing level height, predictand is large hail.

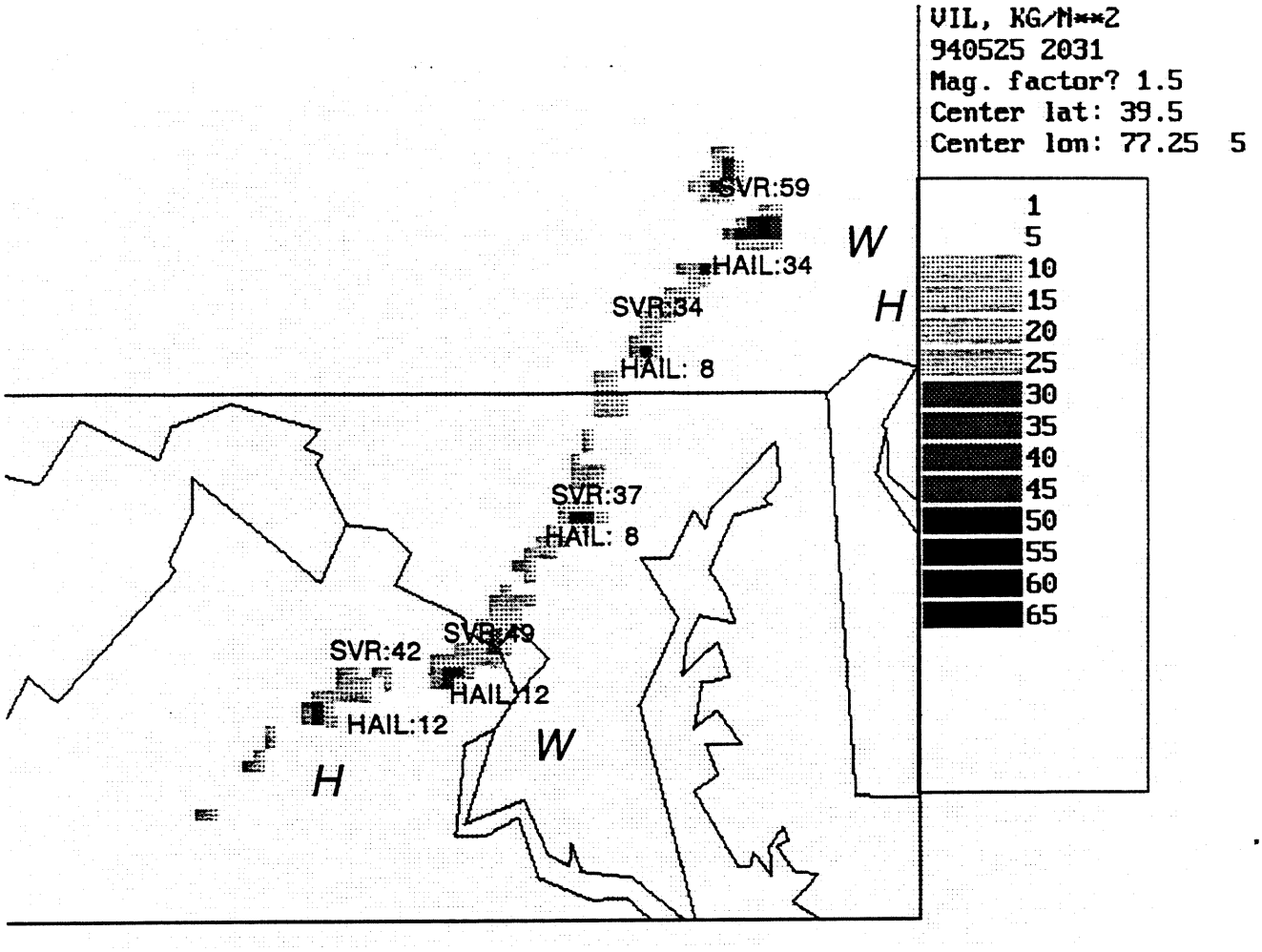


Figure 8. Values of severe weather and hail probability plotted over a VIL analysis. Data are from the Sterling, Virginia, WSR-88D, at 2031 UTC, 25 May 1994. Severe weather events (H for large hail, W for wind) during the period 2020-2100 UTC are also plotted.

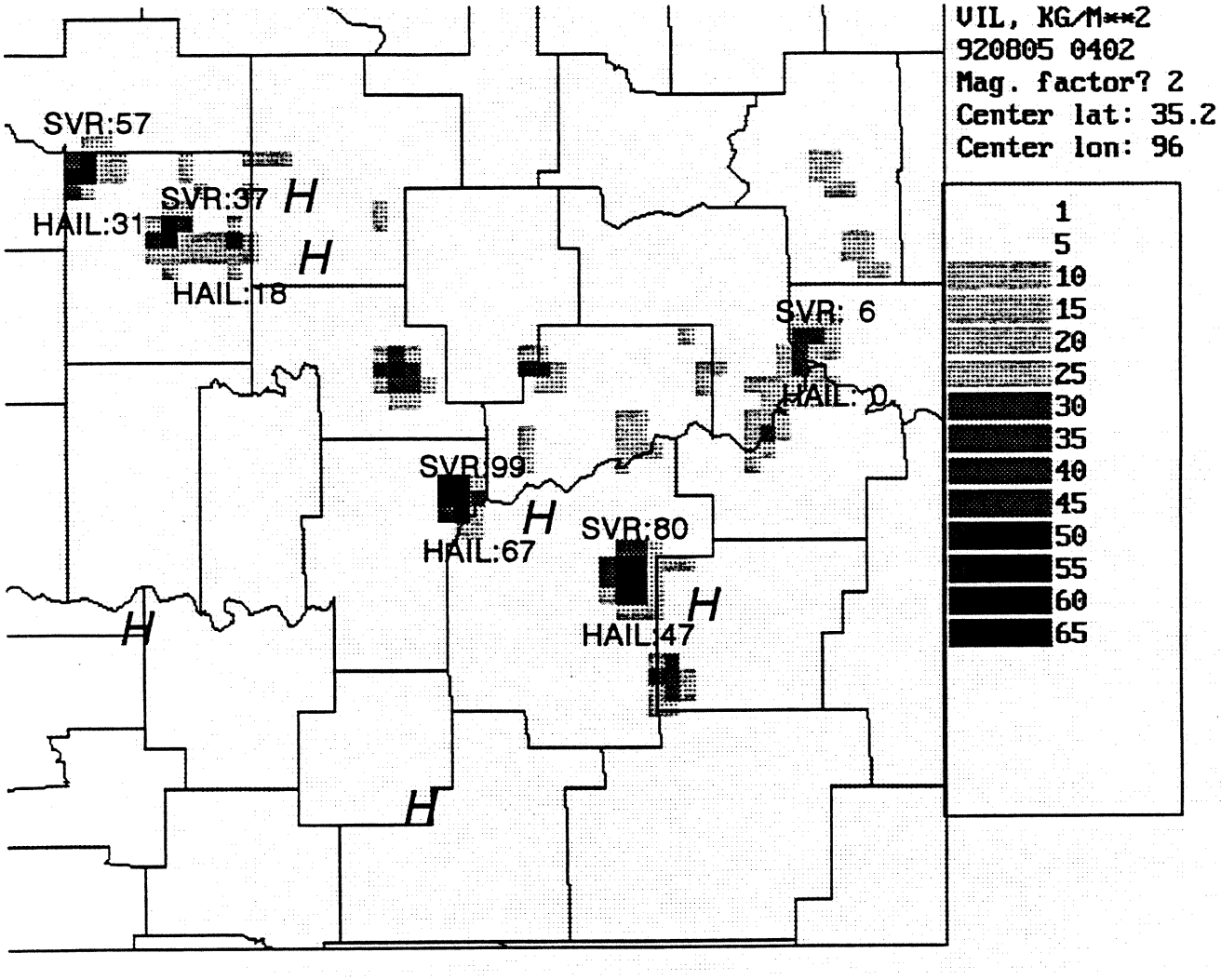
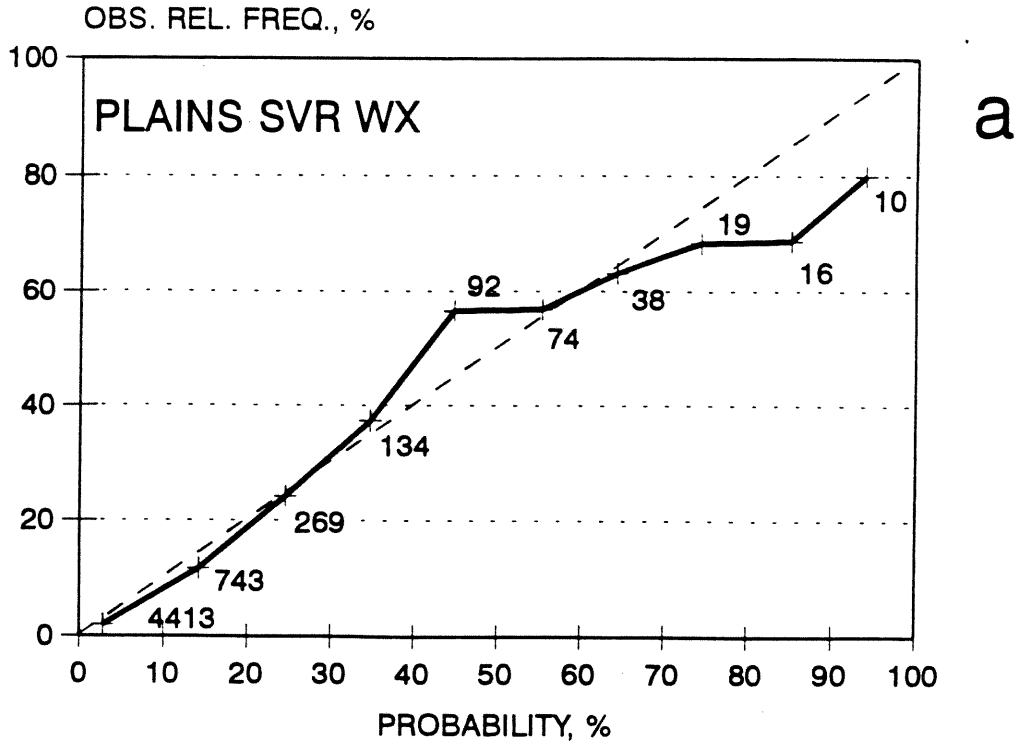


Figure 9. As in Fig. 8, except that data are from the Norman, Oklahoma, WSR-88D at 0402 UTC, 5 August 1992.



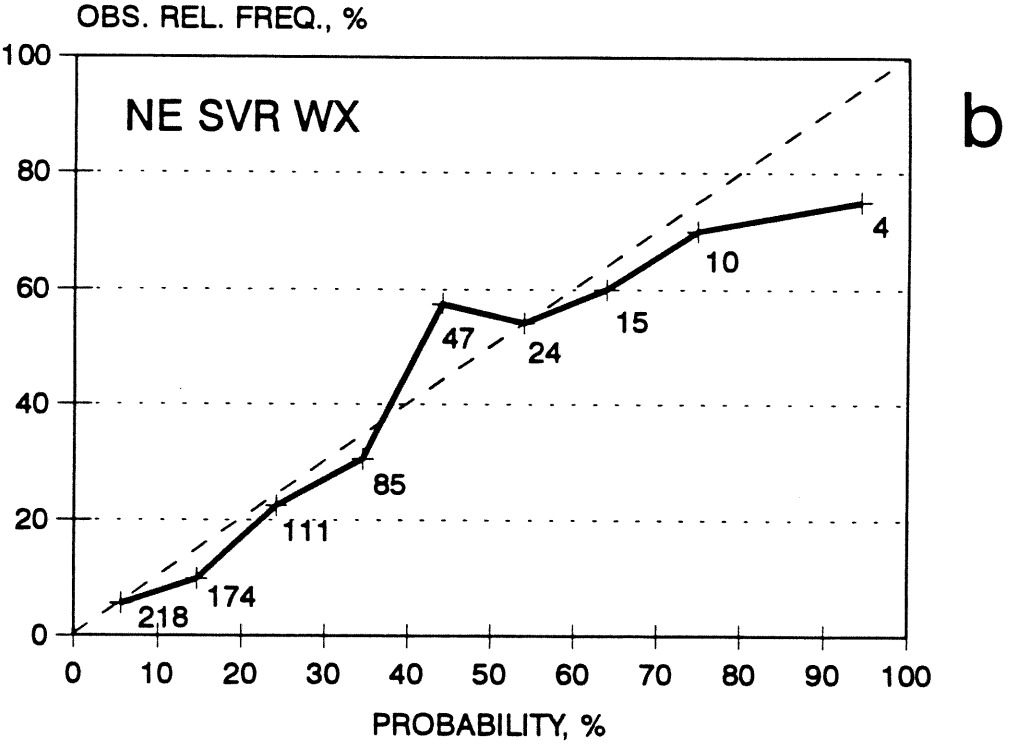
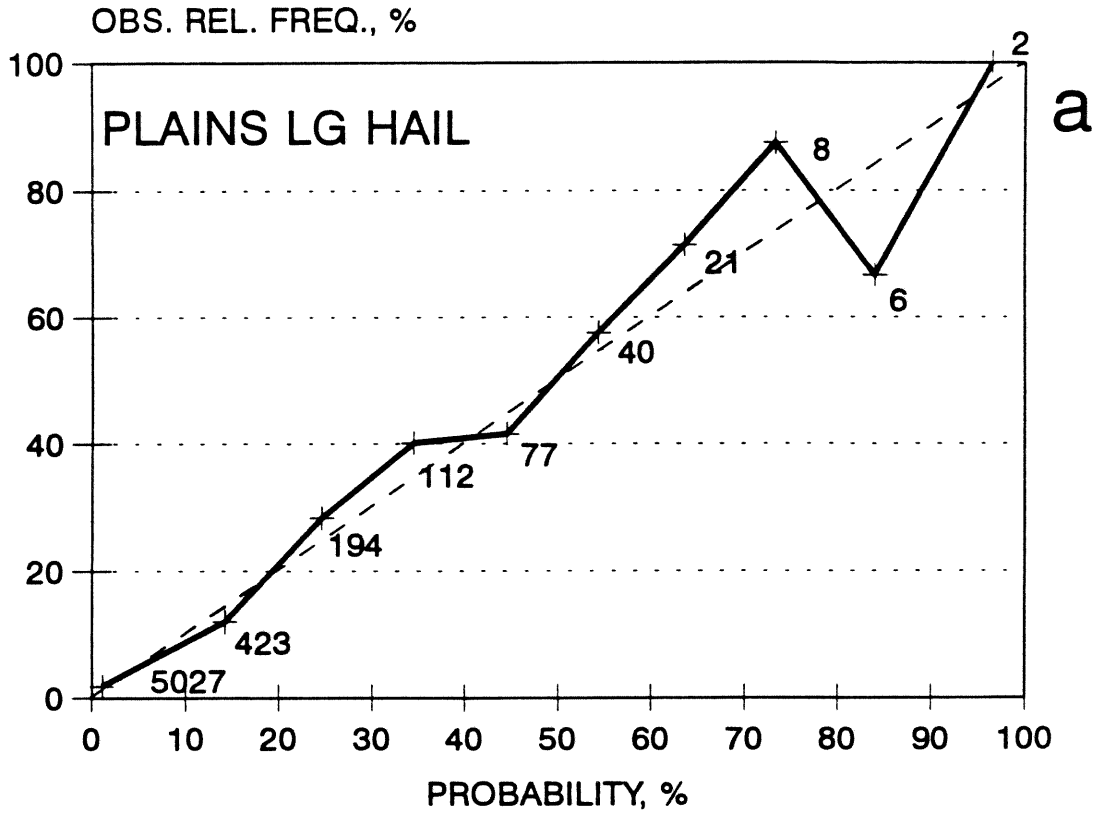


Figure 10. Reliability of SWP3 algorithm for (a) Plains and (b) Northeast cases. The number of cases in each 10 percentage-point category is plotted next to the observed relative frequency value.



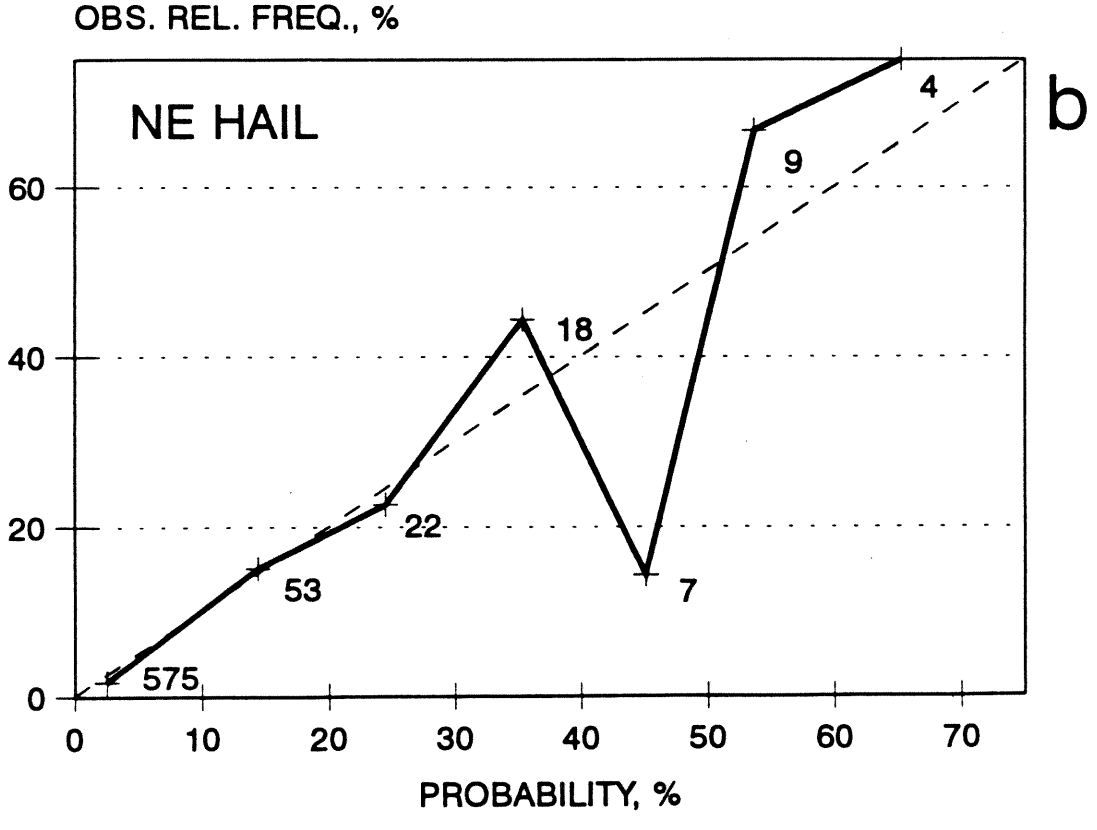
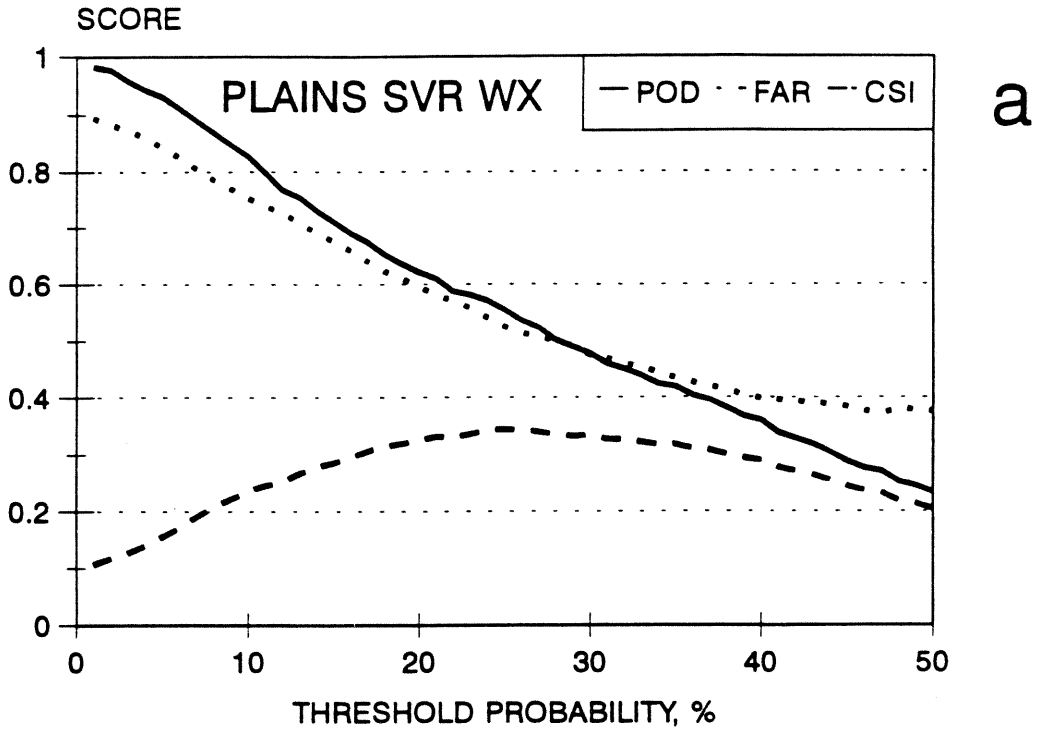


Figure 11. As in Fig. 10, except for the HAIL3 algorithm.



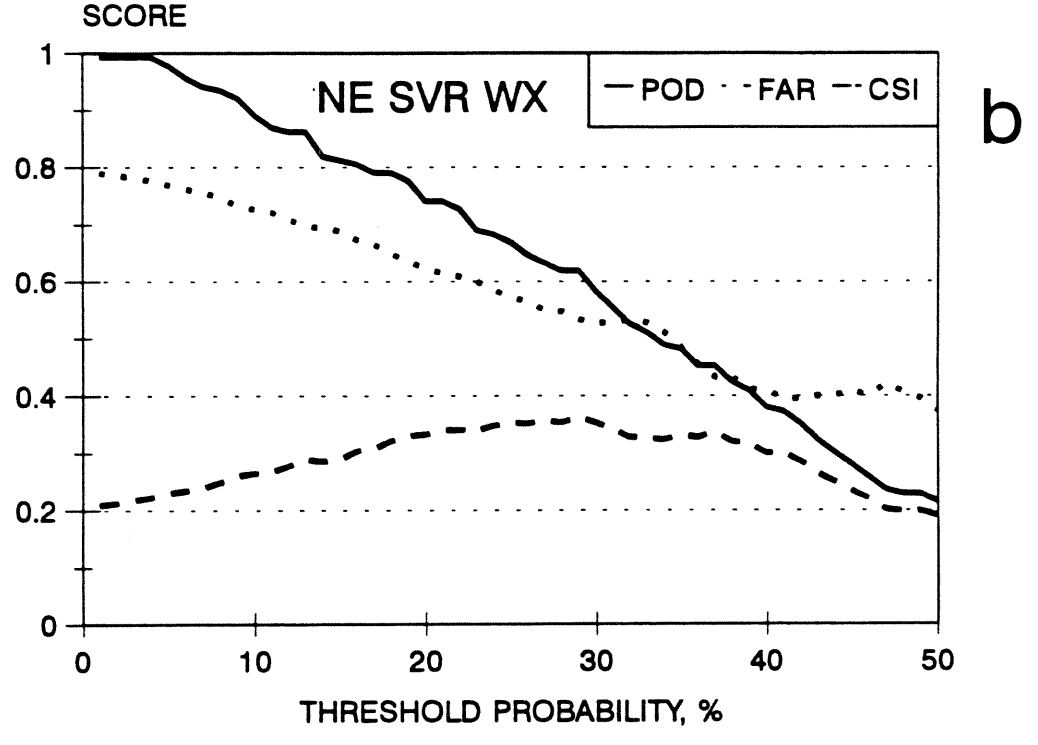


Figure 12. Categorical forecast scores for SWP3 algorithms for (a) the Plains and (b) the Northeast. Scores were determined from cases within the development data sample.

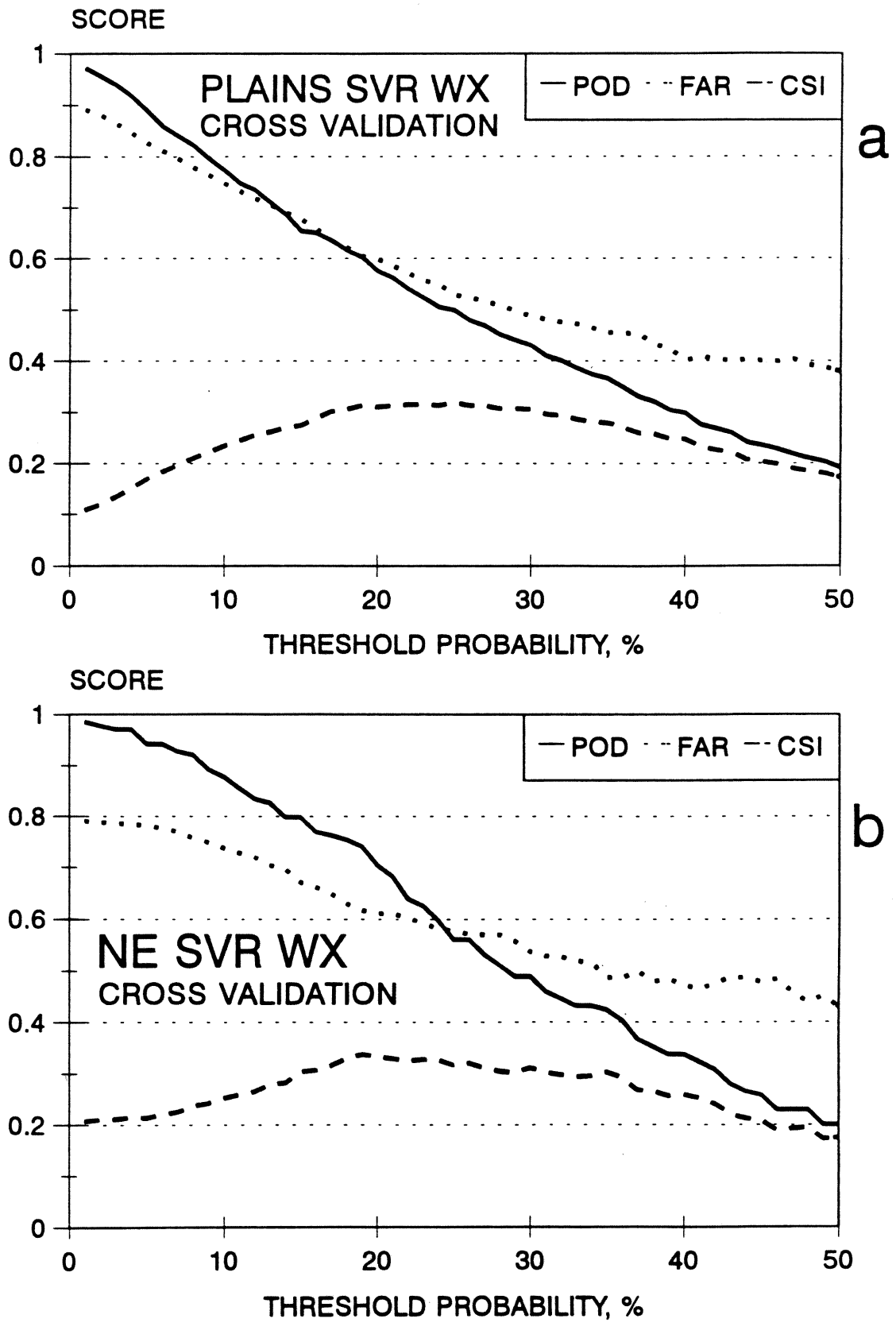
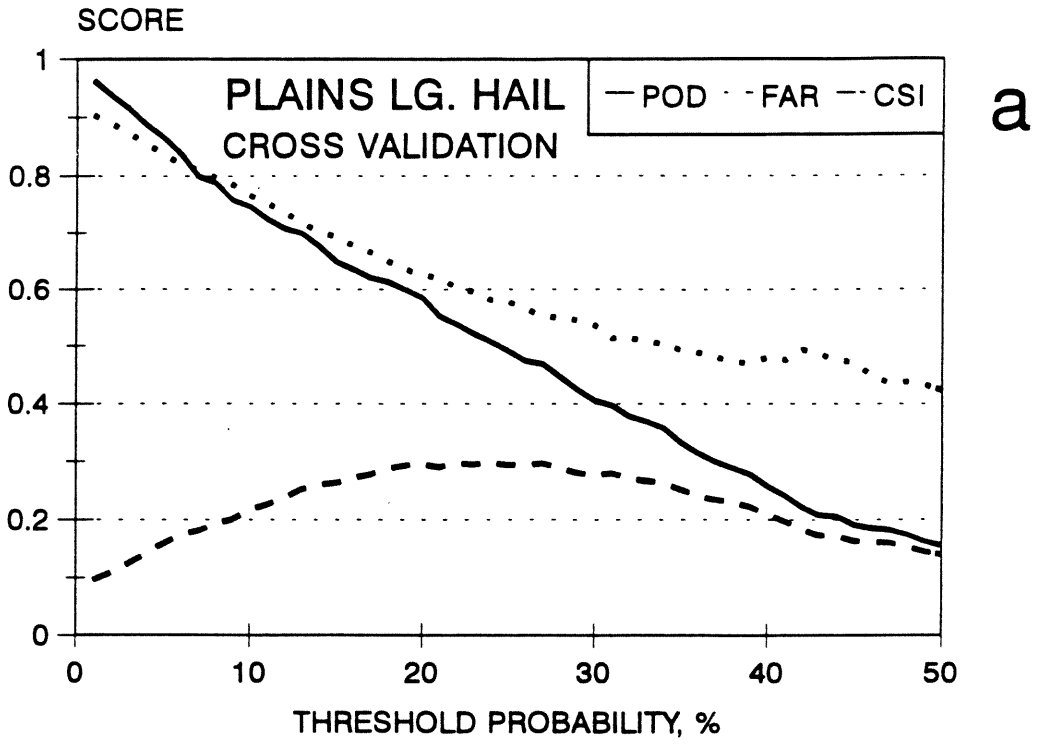


Figure 13. As in Fig. 12, except that a cross-validation approach was used; thus the scores are more representative of those that would be achieved within an independent data sample.



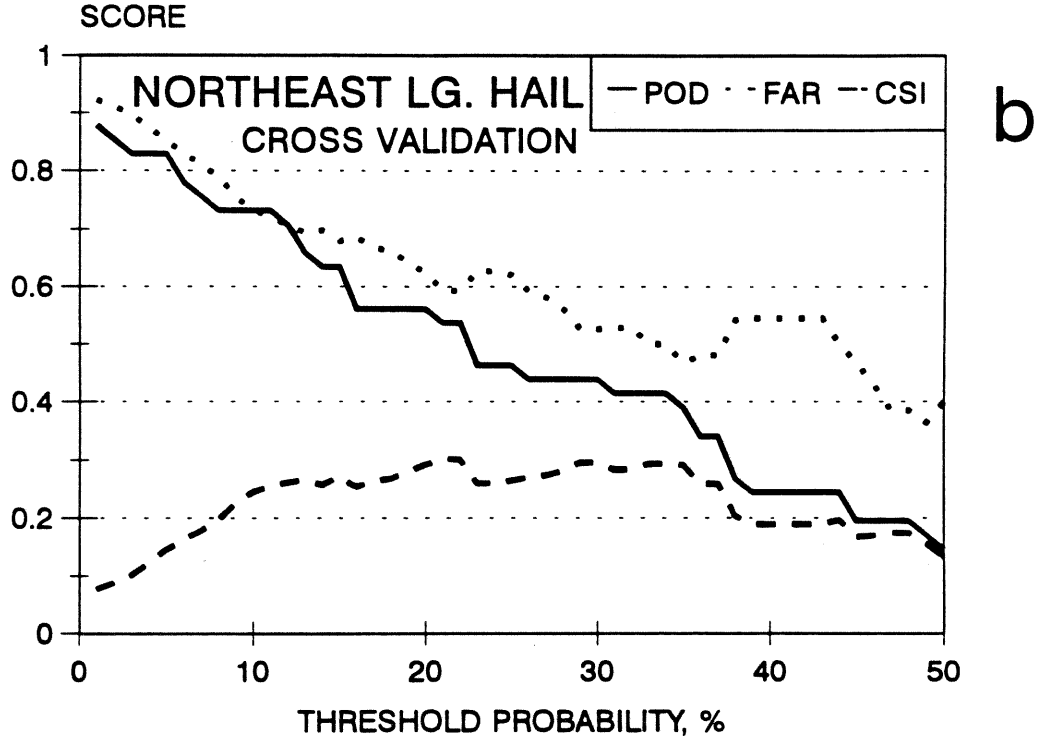


Figure 14. As in Fig. 13, except scores for the HAIL3 algorithm.