Trends in Severe Local Storm Watch Verification at the National Severe Storms Forecast Center

RICHARD W. ANTHONY AND PRESTON W. LEFTWICH, JR.

National Weather Service, National Severe Storms Forecast Center, Kansas City, Missouri

(Manuscript received 6 November 1991, in final form 1 May 1992)

ABSTRACT

Trends of tornado and severe thunderstorm watch verification for the period 1967–1990 are presented. Over the past 10 years the annual number of reported severe thunderstorm events has increased substantially. In comparison, the number of tornado events reported has remained relatively constant from year to year. During the period 1967 to 1990, the percentage of watches verifying has increased from 45% to 85%, while the probability of detection (POD) of severe local storm events has increased from about 0.35 to 0.50. Recent results show that the number of severe thunderstorm watches has steadily increased since 1985, the size of watches has slowly decreased, and lead time has slightly decreased.

Yearly fluctuations in tornado watch verification statistics appear to be best related to the number of outbreak tornadoes. Data suggest the Severe Local Storm Unit (SELS) performance in forecasting outbreak episodes and strong/violent tornadoes is improving at a faster rate than in forecasting isolated tornadoes. Improving trends in severe thunderstorm watch verification for the period 1967 to 1990 are also documented. In general, the results suggest a gradual improvement in forecast performance since the late 1960s. Several factors that impact the verification scores are also identified and discussed.

1. Introduction

This study examines the verification of severe local storm watches issued by the National Severe Storms Forecast Center (NSSFC). In addition to documenting trends in the accuracy of these forecasts, it updates an earlier work by Pearson and Weiss (1979).

In recent years, several studies have examined the quality of weather forecasts. For example, Murphy and Sabin (1986) and Glahn (1985) have shown that, on a national basis, there has been statistically significant improvement in the quality of precipitation and temperature forecasts over the past two decades. Murphy and Sabin (1986) also showed that improvements have been better for objective forecasts produced by numerical–statistical models than for subjective forecasts formulated by National Weather Service (NWS) forecasters.

By definition, a severe thunderstorm produces wind gusts greater than or equal to 50 kt (93 km h^{-1}) and/or hail greater than or equal to 3/4 inch (1.9 cm). Galway (1989) provides a thorough review of the evolution of severe thunderstorm criteria. At a specific location, a severe local storm is a rare event, and, to a large extent, documentation depends upon public reporting of a storm's occurrence. As discussed by Hales and Kelly (1985), there have been dramatic increases in the annual numbers of reported severe local storm events. Schwartz and Flueck (1988) identify numerous factors that may influence the number of severe events now being reported. Among these important factors are an increased effort by the NWS to both improve severe local storm forecast verification procedures and establish trained local spotter networks.

Figure 1, which depicts the annual number of severe local storm reports for 1967 through 1990, shows that the overall increases in the total number of events are primarily due to increases in damaging wind and hail reports since the late 1970s. In comparison, the annual numbers of tornado reports have shown much less variation. Although it is beyond the intent of this paper to thoroughly analyze the effects of these changes on verification of severe local storm watches, it is important to recognize that changes in the character of the event database impact the verification statistics.

Galway (1967) presented verification statistics for the period 1952–1966 and Pearson and Weiss (1979) presented verification results for the period 1967–1977. As an extension to these previous works, this paper presents verification results for tornado and severe thunderstorm watches for the period 1967–1990. Since the verification procedures have been modified over the past ten years a brief review of current procedures is in order.

Corresponding author address: Richard W. Anthony, NWS, National Severe Storms Forecast Center, Federal Building, Room 1728, 601 East 12th Street, Kansas City, MO 64106.

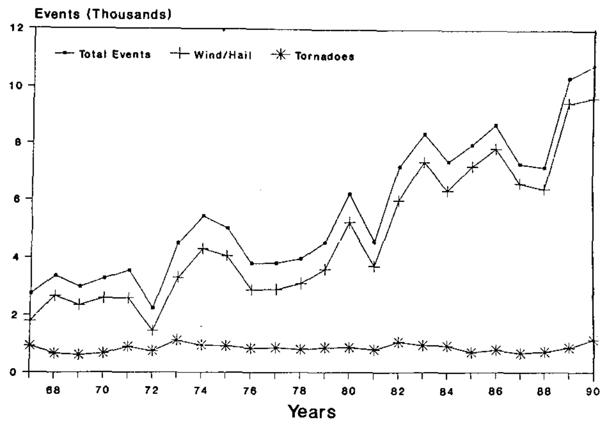

FIG. 1. Annual tornado and wind/hail severe local storm reports in the contiguous United States for 1967–1990.

2. Severe local storm watch verification procedures

This paper focuses on verification of severe local storm watches, an unscheduled public product of the Severe Local Storms (SELS) Unit. SELS issues two types of watches, tornado and severe thunderstorm. Typically, a tornado watch is issued when meteorological parameters favor tornado development. A severe thunderstorm watch is issued when meteorological parameters favor large hail and damaging thunderstorm wind gusts. The average watch is valid for approximately six hours and is a parallelogram that covers an area of approximately 59 800 km^2 (23 000 n mi^2). Currently, routine verification of severe local storm watches includes four basic statistics: the probability of detection (POD), which is the proportion of severe events correctly forecast; the percent verified (PV), which is the proportion of watches containing at least one report; and modified versions of both the critical success index (CSI) (Schaefer 1990; Weiss et al. 1980) and false-alarm ratio (FAR). The CSI is the ratio of successful predictions to the number of severe events and false alarms. The modified FAR for watch verification considers both the spatial and temporal distribution of severe reports. For spatial purposes, historically, each severe local storm report occurring within a watch area is considered to affect an area of approximately 42 000 km^2 (10 000 n mi^2), which is equivalent to a 5 × 5 array of manually digitized radar (MDR) grid blocks[1] [each MDR block is approximately 41 km (22 n mi) on a side and the size of the 5 × 5 array is about 42 000 km^2] (Fig. 2). Each block can be counted only once. The good area percentage (A) is defined as the area affected by severe local storms divided by the total area in the watch.

[1] As Galway (1967) explains, this criterion was based on initial Air Weather Service requirements for forecast precision.

Because a specific valid time is assigned to each watch, the temporal distribution of reports is also incorporated into the FAR calculations. A watch is divided into half-hour time intervals, starting at the beginning of the valid time. Severe reports occurring within a valid watch are considered to affect three half-hour periods: the half-hour in which it occurs, the half-hour before, and the half-hour after the time of occurrence. If any of these periods is not in the valid period of the watch, it is not counted. Each half-hour interval can be verified only once. The cumulative period of time affected by severe local storms is divided by the valid period of the watch to determine the good time percentage (T). The total FAR for watches is then defined as

$$\mathrm{FAR} = 1 - (A \times T), \quad (1)$$

where A is the proportion of the watch area that is affected by severe local storm reports and T is the good time percentage. Additional details of watch verification can be found in Weiss et al. (1980).

3. Verification results

a. Combined severe local storm watch statistics

Although Galway (1967) and Pearson and Weiss (1979) published only tornado watch verification statistics, recent NSSFC summaries (e.g., Leftwich 1989) considered both tornado and severe thunderstorm watches. Figure 3 shows the number of tornado and severe thunderstorm watches issued by year, and the combined total number of watches for the period 1967–1990. In recent years, there has been a general increase in the total number of watches. This has resulted primarily from an increase in severe thunderstorm watches. This tendency roughly mirrors the event trend noted in Fig. 1. Although there is no apparent long-term trend in the number of tornado watches, during

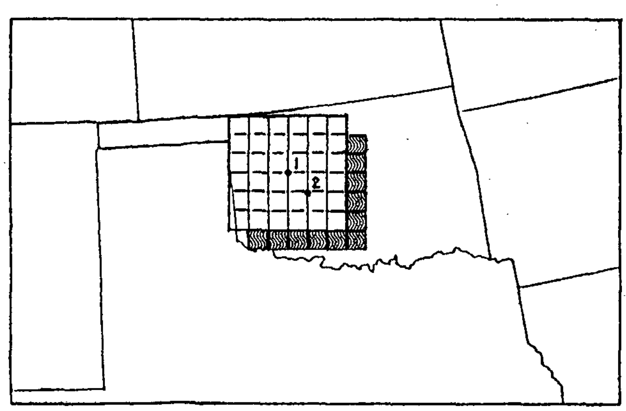

FIG. 2. MDR elements in Oklahoma. A report in block 1 will verify the plain MDR boxes. A second report in block 2 will verify the additional 11 stippled boxes.

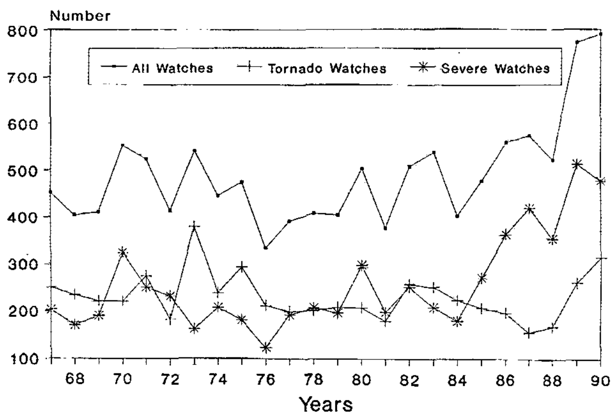

FIG. 3. Annual number of tornado and severe thunderstorm watches issued for 1967–1990.

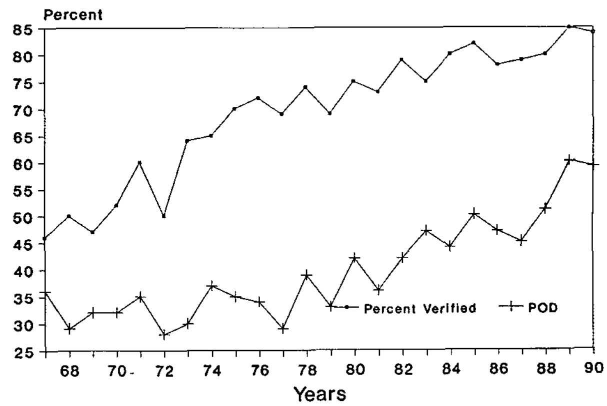

FIG. 5. Percent of all watches verified and probability of detection (POD) for all events 1973–1990.

the period 1982 to 1987 there was a slight decrease, followed by an increase during 1989 and 1990. Again, this mirrors the trend in tornadoes. Pearson and Weiss (1979) found that the mean watch size was slowly increasing. However, recent data (Fig. 4) show a reversal of the earlier trend. Note the approximately 15% reduction in size since 1985.

The issuance of either type of watch, by definition, denotes the threat of severe local storms. An overall analysis of severe local storm forecasting is addressed when all watches are considered together. For the combined watch verification statistics (Fig. 5) any severe event will verify any type of watch. The percentage of watches that contain at least one severe event has increased from 45% in 1967 to nearly 85% in 1990. This represents about a 40% increase over the 24-year period. Likewise, the probability of detection (POD), which is the percentage of severe events occurring within valid watch areas, has increased from 0.35 in 1967 to 0.60 in 1990. These results suggest a general improvement in performance during the period 1967–1990. Some of the improvement may be due to factors such as an increased number of severe event reports and reduction of lead times over recent years to about 30 min. Other factors, such as improved numerical models, use of interactive computers, and more proficient recognition of severe local storm-producing synoptic patterns, have also produced improvement. Additional evidence of improvement in watch verification is illustrated in Fig. 6. In the 24-year period from 1967 to 1990, the FAR for all watches has improved from around 93% to about 60%. Also, during the same period, the CSI has increased from approximately 0.06 to about 0.32.

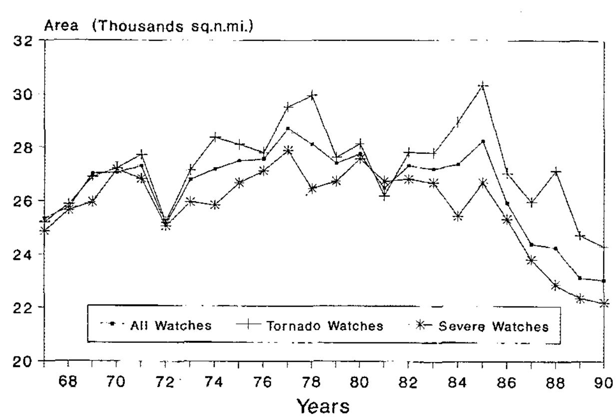

FIG. 4. Annual average watch area for 1967–1990.

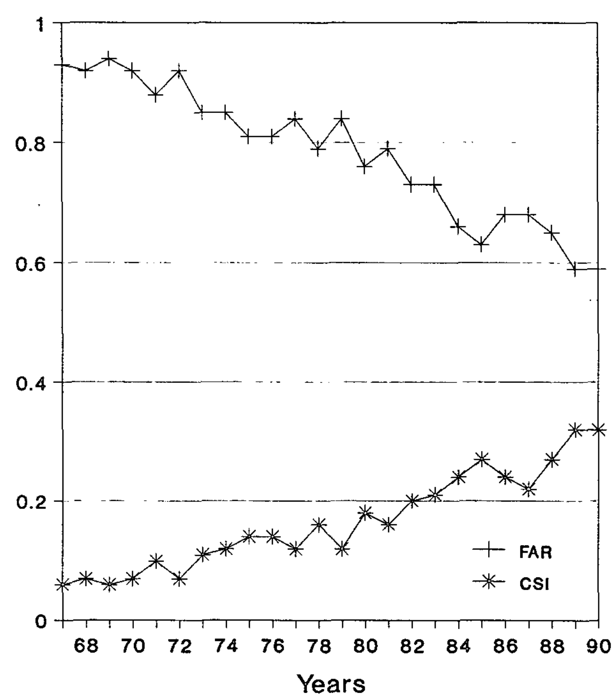

FIG. 6. False alarm ratio (FAR) and critical success index (CSI) for all watches 1976–1990.

b. Tornado watch statistics

Issuance of tornado watches emphasizes the threat of tornadoes. To assess tornado forecasting skill, tornado watches are verified separately. That is, a tornado watch is considered to be verified only when a tornado occurs within the valid time period and within the watch area. The average yearly number of tornadoes reported for the period 1967–1990 is 838. Figure 7a depicts the number of tornadoes that were reported each year during this period. Yearly values range from 1133 in 1990 to 608 in 1969. Figure 7b depicts the annual percent departure from the normal number of tornadoes and the percent departure from the normal number of outbreak tornadoes. In this paper, an outbreak consists of 10 or more tornadoes occurring in an organized temporal and spatial manner using Galway's (1977) criteria. Visual inspection of Figs. 7a and 7b suggests a slight long-term decrease in number of tornadoes, except for the years 1989 and 1990. That is, the slope of a fitted straight line is slightly negative. However, the slope of such a trend line does not test significantly different from zero via a *t* test (Walpole and Myers 1978) even at the 10% level. The recent negative tendency may be a result of a better understanding of microburst-type events that cause damage that once was attributed to tornadoes, but is now classified as straight-line wind damage, thus reducing the number of confirmed tornadoes (Fujita 1981). The record number of total severe reports, which occurred in 1990, is likely related to a number of factors, such as the return to above-normal number of tornadoes, an increased number of severe weather spotters, and National Weather Service emphasis on forecast verification.

Individual tornadoes are generally short-lived and affect a small area. Thus, verification of tornado watches via tornadoes focuses only on the percentage of watches verified and the associated POD. These statistics are presented for the period 1967–1990 in Fig. 8. Note that from 1970 to 1984 the percentage of watches that were verified increased from 30% to 53%. Although the upward trend reversed in 1985, substantial recovery was noted the last three years. The reduction in the percentage of tornado watches verified in the mid-1980s appears to correlate with the dramatic decrease in tornadoes reported from 1984 to 1987 (Fig. 7a).

SELS began issuing both tornado and severe thunderstorm watches in 1964. The question of whether SELS has demonstrated the ability to distinguish between primarily a tornado threat or severe thunderstorm threat is an important one. Data for the period 1967–1990 show that only about 10% of the reported tornadoes occur in severe thunderstorm watches, while nearly 32% of the tornadoes reported are in tornado watches. These data suggest that once the severe local storm threat is correctly diagnosed, SELS does demonstrate significant ability to distinguish between tornado and severe thunderstorm situations.

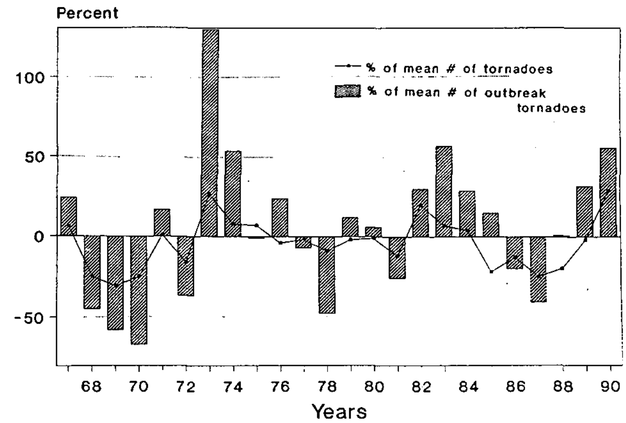

FIG. 7b. Percent departure from mean number of tornadoes (838) and mean number of outbreak tornadoes (189) for 1967–1990.

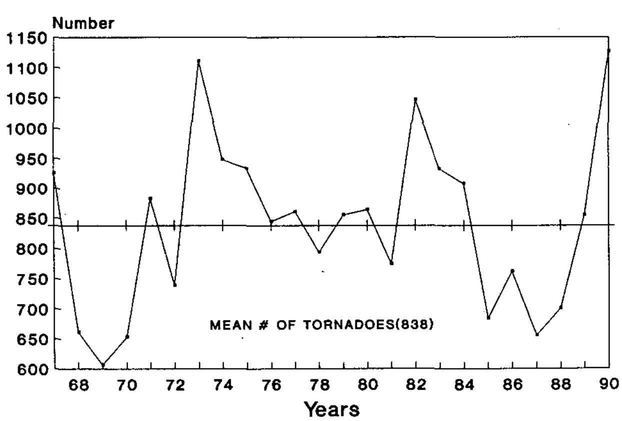

FIG. 7a. Annual number of reported tornadoes for 1967–1990. Mean number of tornadoes during this period is 838 (solid horizontal line).

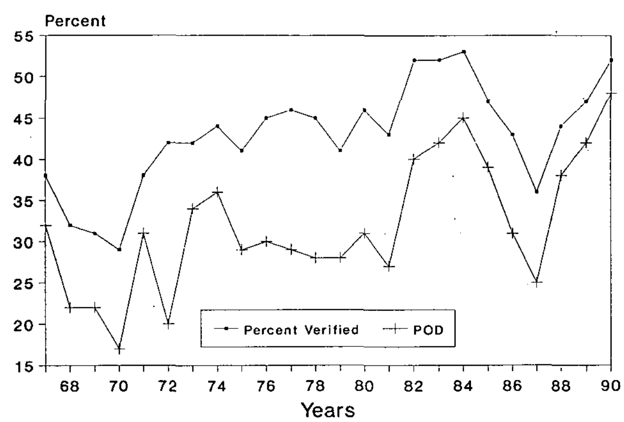

FIG. 8. Percent of tornado watches verified and probability of detection (POD) of tornadoes for 1967–1990.

c. *Tornado watch verification during tornado outbreaks*

As an extension to the data compiled by Galway (1977) and the statistics presented by Pearson and Weiss (1979), additional statistics are presented for tornado outbreak days. As previously mentioned, an outbreak consists of 10 or more tornadoes occurring in an organized temporal and spatial manner. The temporal and spatial criteria are those defined for three outbreak types by Galway (1977). Thus, the occurrence of 10 or more tornadoes on a given day does not by itself constitute an outbreak. In this section, verification was performed on tornado watches that were issued prior to or during the time of the outbreak and in the area of the outbreak.

Intuitively, from Fig. 9, which depicts the number of outbreak and nonoutbreak tornadoes for the period 1967–90, and Fig. 7b, one sees that the range in total number of outbreak tornadoes is slightly greater than that of the nonoutbreak tornadoes. Figure 10 suggests a correlation between the POD for tornadoes and the yearly number of outbreak tornadoes. To verify this observation, a linear correlation coefficient between normal variates of the POD of tornadoes and normal variates of the number of outbreak tornadoes was computed for the period 1967 to 1990. A linear correlation coefficient of 0.62 was found to be statistically significant at the 1% level, based on the analysis of variance procedure (Panofsky and Brier 1965). For comparison, a similar procedure was followed for the percentage of tornado watches verified and the number of outbreak tornadoes. In this case, a linear correlation coefficient of 0.51 was not statistically significant at the 1% level. It is reasonable to expect that the POD is often higher when there is a greater number of outbreak tornadoes. One watch covering an outbreak may contain many tornadoes and result in a higher POD. The watch itself, however, can be "verified" only once, and thus, an outbreak will not influence the overall percentage of watches verified as much as it does the POD. Furthermore, in the midst of tornado outbreaks, storms are ongoing at the times watches are issued. Uncertainty concerning the timing and location of further development is reduced. Thus, it is likely that watches issued during outbreaks will attain better verification scores than many of the isolated watches.

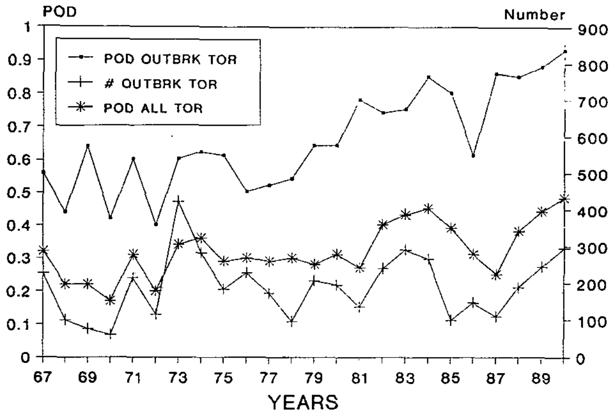

FIG. 10. Probability of detection (POD) for outbreak tornadoes and all tornadoes and number of outbreak tornadoes for 1967–1990. Right scale refers to number of tornadoes and left scale to POD.

However, the rather dramatic increase in the POD for outbreak tornadoes (Fig. 10) since 1976 is not reflected in the number of outbreak tornadoes. In fact, like nonoutbreak tornadoes, the number of outbreak tornadoes has generally shown no long-term increase. Figure 10 also shows the POD for all tornadoes. The average POD during the 24-year period for the outbreak tornadoes is 0.64, but during the last eight years the average has increased to 0.79. Also note the similarities in the yearly trends of the two PODs and the apparent gradual improvement in forecast skill.

On average, more than 70% of the watches issued prior to or during outbreak situations have verified since 1967 (Fig. 11). It is noteworthy that the results

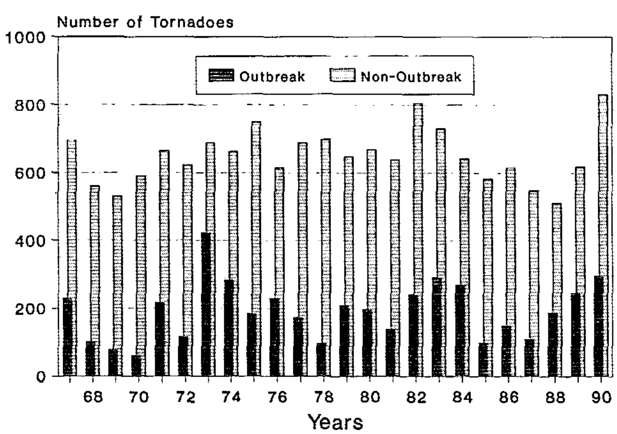

FIG. 9. Annual number of nonoutbreak tornadoes and outbreak tornadoes for 1967–1990.

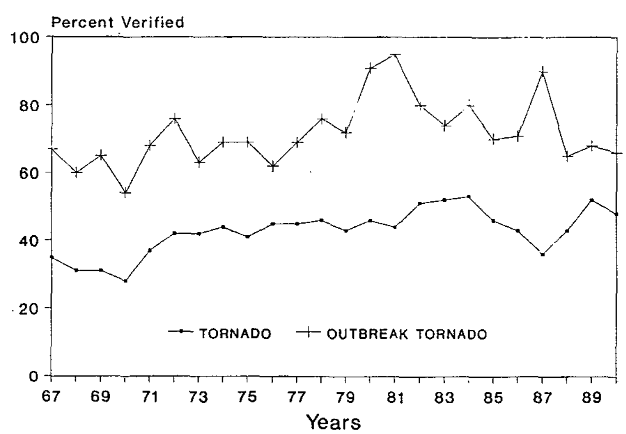

FIG. 11. Percent verified for all tornado watches and watches during outbreaks for 1967–1990.

for watches issued on outbreak days show a higher degree of skill than the verification for all tornado watches.

Intensities of tornadoes are indicated by F scale (Fujita 1981) with values ranging from F0 (weakest) to F5 (most violent). Strong and violent (F2–F5) tornadoes cause most fatalities (Kelly et al. 1978). Therefore, POD values based on tornado intensity were computed. The POD as a function of tornado intensity relative to all valid severe local storm watches is given in Fig. 12. The POD for strong/violent (F2–F5) tornadoes, which has remained at or above 0.7 for the last six years, is consistently higher than values for weak (F0–F1) tornadoes (Leftwich 1989). Results for the past several years are in agreement with the earlier findings of Ostby and Higginbotham (1982).

d. Severe thunderstorm watch statistics

Any severe local storm event occurring during the valid period and within the watch area will verify a severe thunderstorm watch. As shown in Fig. 1, the annual numbers of hail and thunderstorm wind events have increased dramatically over the past 24 years. Further, over the past ten years these events have nearly doubled the values of the previous decade. This increase is also reflected in the number of severe thunderstorm watches issued (Fig. 3). The average number of severe thunderstorm watches prior to 1984 was about 200 per year, but since 1984 the number has increased substantially to approximately 350 per year with a record number (505) issued in 1989.

The combination of smaller watches, shorter lead times, shorter watch valid times (or reduced "effective watch duration"), and more events has likely contributed to an increase in number of watches. However, further study is required to quantify such a relationship.

Since 1985, the percentage of severe thunderstorm watches verified has shown little change (values near 80%), but since 1984 the POD for all events has improved to nearly 40% (Fig. 13). Figure 13 also depicts a slow improvement in the FAR and CSI over the last several years.

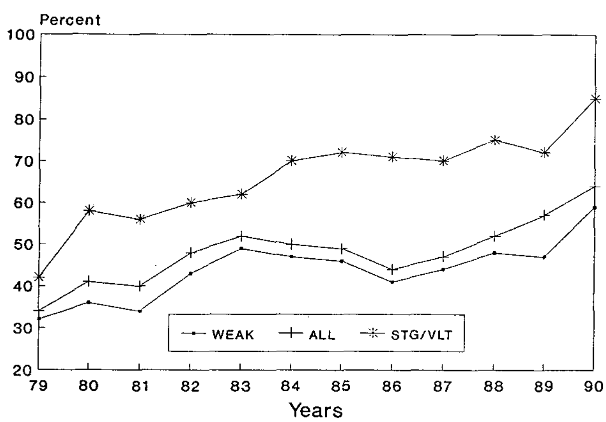

FIG. 12. Probability of detection (POD) for tornadoes based on F scale. Weak (F0–F1), Stg/Vlt (F2–F5), all (F0–F5) for 1979–1990.

e. Lead-time statistics

Technological advancements in communication capabilities over the past decade at both the NSSFC and at NWS field offices have significantly improved the ability to disseminate public weather information in a timely manner. However, adequate time is still needed for storm spotters to arrive at observing sights, and for local disaster-preparedness agencies to activate contingency plans prior to the development of severe local storms.

In the early years of severe storm forecasting, watches may have been issued several hours prior to the expected onset of severe convection. In recent years, the watch lead time, which is defined as the interval from watch issuance time to the time the watch becomes valid, has remained relatively constant at about 30–35 min (Fig. 14). Schwartz and Flueck (1988) documented lead-time trends for the period 1955–1984. They noted that since the late 1970s uniform lead times have become the rule.

In addition to the standard lead time (LT), Weiss and Kelly (1982) introduced statistical measures of projection time (PT) and projection ratio (PR). PT is defined as the interval between watch issue time and the time of the first severe storm report within the watch area. Projection ratio, which is defined as the ratio PT/LT, allows comparison of lead time and projection time. If $PR = 1$, the first severe report within the watch area occurs at the start of the valid period. If $PR \gg 1$, the first report occurs well into the valid period, whereas if $PR < 1$, the first report occurs prior to the start of the watch. The PT can be considered the "effective lead time," since it measures the interval between the issuance of the watch and the first report of severe weather within the watch. The frequency distribution of watch lead times that Weiss and Kelly (1982) presented for 1980 is compared with data presented for 1981–1990 in Table 1. The results are quite similar. Data show that 46% of all watches have lead times of 31–45 min. Another one-third (35%) have lead times of 16 to 30 min. Very few watches had lead times less than 15 min or greater than 1 h.

The projection time for all watches and tornado watches is also depicted in Fig. 14. The time interval between watch issue time and the time of the first severe local storm report is about 2 h for tornado watches and about 80 min for all watches. Note that the projection time for tornado watches only and for combined tornado and severe thunderstorm watches is significantly greater than their respective lead times. Also of interest is the trend toward a shorter PT for all watches while the PT for tornado watches has generally increased.

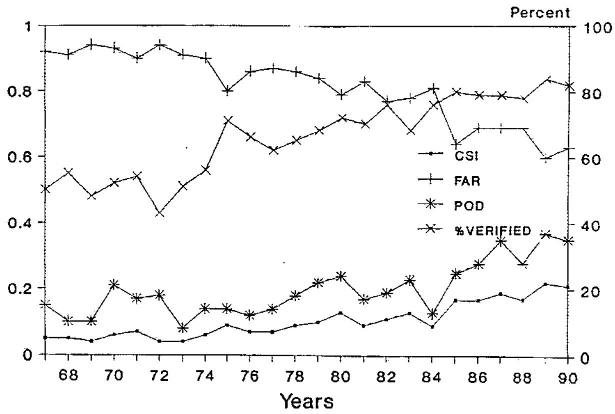

FIG. 13. Severe thunderstorm watch verification for 1967–1990. Left scale refers to CSI–FAR. Right scale refers to percent verified.

The percentage of watches verifying for the different lead-time categories for the period 1981–1990 averaged nearly 70%, which is similar to 1980 results. Weiss and Kelly (1982) noted that 61% of watches verifying in 1980 had the initial severe local storm event occur after the start of the valid period. During the period 1981–1990, however, the percentage where PR > 1 has averaged about 75% for all watches issued. Watches that are issued when there is ongoing severe thunderstorm activity could be categorized as "continuation" watches rather than initial threat, or independent, watches. Stratification of data in this manner, after the fact, would be a monumental task and far beyond the scope of this paper. One could argue that continuation watches require less skill than initial watches. Although this undoubtedly plays a role in verification results, follow-on watches still require the forecaster to determine 1) if the severe storms will diminish or persist, and 2) if they do persist, where they will propagate. These are not trivial questions. Statistical analysis still

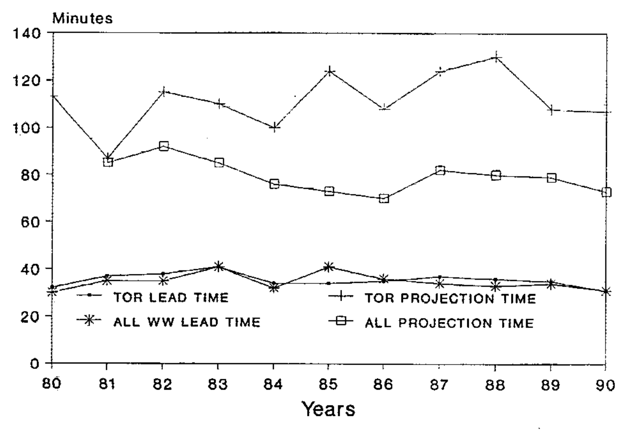

FIG. 14. Mean annual lead time and projection time for all watches and tornado watches for 1980–1990.

TABLE 1. Frequency distribution (%) of lead time in minutes for severe local storm watches issued in 1980 and 1981–1990. Data for 1980 from Weiss and Kelly (1982).

Min	1980 (%)	1981–1990	(%)
0–15	5	305	6
16–30	36	1942	35
31–45	45	2540	46
46–60	9	543	10
>60	5	190	3

supports success in SELS's effort to have watches in effect prior to the initial severe local storm event.

4. Impact of numerical model guidance and technology on severe local storm verification

Some of the improved forecast skill can be attributed to the improved synoptic-scale Numerical Weather Prediction (NWP) models (Sanders 1988; Junker et al. 1989). Automated short-range forecasts of the probability of severe local storm occurrence are produced by statistical methods incorporating diagnostic analysis of observed meteorological fields and LFM model output (Charba 1979). This product has provided additional guidance to forecasters.

Some of the improved skill can also be attributed to the use of interactive computer systems at NSSFC (Mosher 1985, 1991; Browning 1991). Over the past decade, SELS has made significant use of interactive computer technology. In 1982 the Centralized Storm Information System (CSIS) (Anthony et al. 1982) was installed at the NSSFC. Some of the more basic capabilities utilized by the forecaster include surface and upper-air plots, temporal change field plots, skew *T*–log*p* sounding displays, and isentropic and cross-section analyses. Of the parameters computed by CSIS, perhaps the most widely used are surface-based lifted index (Edman 1989; Hales and Doswell 1982) and moisture convergence (Hirt 1982; Xiang and Beckman 1985). The CSIS is also capable of ingesting radar data. In 1990, approximately 67 radar stations were available. CSIS is capable of brightness normalizing and remapping radar scope presentations to satellite projection. Radar images can be color enhanced and superimposed on other data. Presentations from several radars can be composited into a single image.

Of all the display and analysis capabilities of CSIS, by far the most significant operational impact has been due to its flexibility in displaying and manipulating satellite data. Conventional satellite data are available as frequently as every 15 min and within 6 min after data collection time. Hourly satellite water vapor data are also available through CSIS (Mosher 1988; Beckman 1987). The forecaster can use the equilibrium temperature (Schaefer et al. 1982), tropopause temperature, or any other desired temperature to develop specific color enhancements. In addition, image enhancement can also be interactively altered, through the full range of colors, by use of "joystick" controls.

In spring 1990, the NSSFC began using the VAS Data Utilization Center (VDUC) (Suomi et al. 1983; Browning 1991). In addition to improved display and analysis capabilities, the VDUC is directly linked with the National Meteorological Center (NMC) computers. This allows rapid access to numerical model output that can be displayed and animated.

Although forecasters are certain that interactive computer systems have had a very positive impact on SELS forecast procedures and operations (Ostby 1984; Mosher 1985, 1991), it is nearly impossible to objectively quantify their impact on either forecasts or verification scores. The same is true for factors such as increased severe local storm reports and improved NWP. Therefore, the goal of this study has been to summarize trends in verification statistics, while acknowledging that many influences exist.

It is apparent, however, that future improvements in severe local storm forecasting must be largely based on a more rigorous understanding of subsynoptic features and scale interactions. Improvement will come as a result of utilizing better models and combining data from new sources, such as ASOS (McNulty et al. 1985); lightning data (e.g., Mosher and Lewis 1990); wind profilers (Augustine and Zipser 1987); WSR-88D Doppler radar (e.g., Wilson et al. 1980; Baer 1991); improved satellite systems (e.g., Shenk et al. 1986); and the eventual incorporation of these data into mesoscale models (e.g., Koch 1988).

5. Summary and conclusions

Verification statistics for severe local storm watches have been presented. Data include performance trends for severe thunderstorm watches, tornado watches, and for all watches. In addition, verification results for outbreak tornadoes are also documented. Results suggest a gradual improvement in most aspects of severe local storm forecasting.

Forecasting severe local storms involves the prediction of relatively rare events. While the verification data may not suggest as high a degree of skill as in forecasting other weather phenomena, such as precipitation, it should be remembered that the severe local storm forecaster is dealing with situations that are climatologically very rare and a chance for positive verification is much lower than is true for other types of forecasting (House 1963; Galway 1967; Pearson and Weiss 1979; Murphy 1991).

Automated verification procedures at NSSFC have been in place since the late 1970s. Although the current verification scheme provides some measure of forecast accuracy, it is recognized that it does not address all characteristics of severe local storm watches. For example, more accurate account can be made for the "nonforecast scenarios." In the formulation of the CSI matrix (Donaldson et al. 1975), there is no attempt to incorporate the correct "no severe" forecasts. Although many of these are trivial, others involve difficult forecast situations that are close to severe thresholds. Thus, it would seem proper to take into account by some method these "none forecast/none observed" cases. Experiments in this area have been discussed by Doswell et al. (1990) and Schaefer (1990). Also, verification statistics stratified by season or geographic area will be addressed in future work.

Proper and meaningful interpretation of verification statistics is exceptionally difficult. It is complicated by a number of interrelated factors influencing the forecast itself. These factors include personnel changes and associated experience level of the individual forecasters; forecast dissemination procedures; and perhaps more important, varying philosophies pertaining to optimum watch characteristics. Since the decision-making process in severe local storm forecasting is highly subjective, it is impossible to explain or confirm every change in skill level suggested by the statistics.

Heuristically, one would expect that the larger the watch area, the better chance a watch has to verify. This trend was observed in the verification statistics presented by Pearson and Weiss (1979). Recent data suggest SELS's severe local storm forecast ability is not as dependent on watch size as it was in earlier years, however. In fact, recent watch-size trends in general (Fig. 4) show that mean areas have slowly decreased for both tornado and severe thunderstorm watches, reversing the trend noted by Pearson and Weiss (1979). Although the watch-size trend is now downward, the POD, FAR, and CSI (Figs. 5, 6) have continued to gradually improve. The trend of a decreasing number of tornadoes for 1982 to 1987 (Fig. 7), mirrored in Fig. 9, results in lowering tornado watch verification for 1982 to 1987. These data illustrate the connection between more active periods and better apparent forecast performance.

Pearson and Weiss (1979) suggested that most of the variation in the total number of tornadoes resides in the number of outbreak tornadoes. Indeed, Figs. 7 and 9 show that the annual number of nonoutbreak tornadoes is relatively constant, but the number of outbreak tornadoes varies substantially. From Fig. 10 one sees that since 1975 the POD has generally increased while the number of outbreak tornadoes during the mid-1980s has generally declined. This suggests that the outbreak POD is not simply a result of a larger number of tornadoes; however, the years with a high number of outbreaks have higher POD. Comparing Figs. 7 and 10, one can see that the years with a higher number of tornadoes (reflecting more outbreak events) tend to have the highest probability of detection (Schaefer 1990).

Since the late 1970s the number of documented severe thunderstorm reports has increased dramatically, while in the mid-1980s the number of reported tornadoes has remained relatively constant. The decrease in tornado watch verification results for 1985–87 appears to mirror the reduced number of outbreak tornadoes. Although outbreak tornadoes in recent years have decreased slightly in number, the POD for outbreak tornadoes has increased to more than 75%. These data suggest that SELS's skill level in forecasting outbreak episodes is improving at a faster rate than forecasts for nonoutbreak tornadoes. Further evidence to support this conclusion is found in the POD of tornadoes stratified by F scale. In general, the POD for strong/violent tornadoes, which are typically associated with outbreak situations, has shown a more pronounced improvement compared to the POD trend for weak tornadoes.

Finally, forecasters are certain that interactive computer systems have had a positive impact on forecast procedures and operations. It is nearly impossible to objectively quantify this impact, however. The same is true for factors such as increased severe local storm reports and improved guidance from numerical models. Therefore, this study has summarized trends in verification statistics, while acknowledging that many influences exist.

Acknowledgments. The authors wish to thank Steven Weiss for his encouragement and reviews of this manuscript. Appreciation is also extended to Fred Ostby, Barry Schwartz, and the reviewers for their comments and suggestions for this manuscript.

REFERENCES

Anthony, R. W., W. Carle, J. Schaefer, R. Livingston, A. Siebers, F. Mosher, J. Young, and T. Whittaker, 1982: The centralized storm information system at the NOAA Kansas City complex. Preprints, *Ninth Conf. on Weather Analysis and Forecasting,* Seattle, Amer. Meteor. Soc., 40–43.

Augustine, J. A., and E. J. Zipser, 1987: The use of wind profilers in a mesoscale experiment. *Bull. Amer. Meteor. Soc.,* **68,** 4–17.

Baer, V. E., 1991: The transition from the present radar dissemination system to the NEXRAD Information Dissemination Service (NIDS). *Bull. Amer. Meteor. Soc.,* **72,** 29–33.

Beckman, S. K., 1987: Operational use of water vapor imagery. *NOAA Tech. Memo. NWSCR-87,* National Weather Service, Kansas City, MO, 15 pp. [NTIS Accession No. PB88-140751].

Browning, P. A., 1991: The VDUC interactive computer system at the National Severe Storms Forecast Center. Preprints, *Seventh Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology,* New Orleans, Amer. Meteor. Soc., 204–207.

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

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

Doswell, C. A., D. L. Keller, and S. J. Weiss, 1990: An analysis of the temporal and spatial variation of tornado and severe thunderstorm watch verification. Preprints, *16th Conf. on Severe Local Storms,* Kananaskis Park, Alberta, Canada, Amer. Meteor. Soc., 294–299.

Edman, D. A., 1989: A comparison of the surface based hourly lifted indices produced by the NSSFC CSIS and AFOS ADAP algorithms. Preprints, *12th Conf. on Weather Analysis and Forecasting,* Monterey, CA, Amer. Meteor. Soc., 93–96.

Fujita, T. T., 1981: Tornadoes and downbursts in the context of generalized planetary scales. *J. Atmos. Sci.,* **38,** 1511–1534.

Galway, J. G., 1967: SELS forecast verification 1952–1966. Preprints, *Fifth Conf. on Severe Local Storms,* St. Louis, Amer. Meteor. Soc., 140–145.

——, 1977: Some climatological aspects of tornado outbreaks. *Mon. Wea. Rev.,* **105,** 477–484.

——, 1989: The evolution of severe thunderstorm criteria within the weather service. *Wea. Forecasting,* **4,** 585–592.

Glahn, H. R., 1985: Yes, precipitation forecasts have improved. *Bull. Amer. Meteor. Soc.,* **66,** 820–830.

Hales, J. E., and C. A. Doswell, 1982: High resolution diagnosis of instability using hourly surface lifted parcel temperatures. Preprints, *12th Conf. on Severe Local Storms,* San Antonio, TX, Amer. Meteor. Soc., 172–175.

——, and D. L. Kelly, 1985: The relationship between the collection of severe thunderstorm reports and warning verification. Preprints, *14th Conf. on Severe Local Storms,* Indianapolis, IN, Amer. Meteor. Soc., 13–16.

Hirt, W. D., 1982: Short-term prediction of convective development using dewpoint convergence. Preprints, *Ninth Conf. on Weather Forecasting and Analysis,* Seattle, Amer. Meteor. Soc., 201–205.

House, D. C., 1963: Forecasting tornadoes and severe thunderstorms. *Meteorol. Mongr.,* **5,** 141–155.

Junker, N. W., J. E. Hoke, and R. H. Crumm, 1989: Performance of NMC regional models. *Wea. Forecasting,* **4,** 368–390.

Kelly, D. L., J. T. Schaefer, R. P. McNulty, C. A. Doswell III, and R. F. Abbey, Jr., 1978: An augmented tornado climatology. *Mon. Wea. Rev.,* **106,** 1172–1182.

Koch, S. E., 1988: Mesoscale modeling of the severe thunderstorm environment. Preprints, *Eighth Conf. on Numerical Weather Prediction,* Baltimore, Amer. Meteor. Soc., J42–J51.

Leftwich, P. W., 1989: Verification of severe local storm forecasts issued by the National Severe Storms Forecast Center: 1988. NOAA Tech. Memo. NWS NSSFC-24, National Weather Service, Kansas City, MO, 10 pp. [NTIS Accession No. PB90-140211/AS].

McNulty, R. P., J. T. Schaefer, M. D. Manker, T. T. Townsend, H. Bogin, and D. M. Entriken, 1985: ASOS and thunderstorms. Preprints, *14th Conf. on Severe Local Storms,* Indianapolis, IN, Amer. Meteor. Soc., 101–104.

Mosher, F. R., 1985: Impacts of interactive processing systems on the forecasting ability of the National Severe Storms Forecast Center. Preprints, *Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* Los Angeles, Amer. Meteor. Soc., 117–123.

——, 1988: Assessment of the utility of VAS data products for severe local storms forecasting. Preprints, *15th Conf. on Severe Local Storms,* Baltimore, Amer. Meteor. Soc., 182–185.

——, 1991: Improved forecaster productivity through the use of interactive technology. *Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* New Orleans, Amer. Meteor. Soc., 419–424.

——, and J. S. Lewis, 1990: Use of lightning location data in severe storm forecasting. Preprints, *16th Conf. on Severe Local Storms,* Kananaskis Park, Alberta, Canada, Amer. Meteor. Soc., 692–697.

Murphy, A. H., 1991: Probabilities, odds, and forecasts of rare events. *Wea. Forecasting,* **6,** 302–307.

——, and T. E. Sabin, 1986: Trends in quality of National Weather Service Forecasts. *Wea. Forecasting,* **1,** 42–55.

Ostby, F. P., and D. M. Higginbotham, 1982: Tornado predictability as a function of geography and intensity. Preprints, *12th Conf. on Severe Local Storms,* San Antonio, TX, Amer. Meteor. Soc., 176–179.

——, 1984: Use of CSIS in severe weather prediction. *Proc. Soc. of Photo-optical Instrumentation Engineers (SPIE)* Vol. 481, Recent Advances in Civil Space Remote Sensing, Arlington, VA, 78–83.

Panofsky, H. A., and G. Brier, 1965: *Some Applications of Statistics to Meteorology.* The Pennsylvania State University, 224 pp.

Pearson, A., and S. J. Weiss, 1979: Some trends in forecast skill at the National Severe Storms Forecast Center. *Bull. Amer. Meteor. Soc.,* **60,** 319–326.

Sanders, F., 1988: The impact of numerical weather prediction on weather forecasting. Preprints, *15th Conf. on Severe Local Storms,* Baltimore, Amer. Meteor. Soc., 52–53.

Schaefer, J. T., 1990: The Critical Success Index as an indicator of warning skill. *Wea. Forecasting,* **5,** 570–575.

——, C. A. Doswell, D. W. McCann, T. W. Schlatter, and H. B. Wobus, 1982: Thermodynamic analysis procedure at the National Severe Storms Forecast Center. Preprints, *Ninth Conf. on Weather Analysis and Forecasting,* Seattle, Amer. Meteor. Soc., 304–309.

Schwartz, B. E., and J. A. Flueck, 1988: On the nature and verification of severe weather watches. Preprints, *15th Conf. on Severe Local Storms,* Baltimore, Amer. Meteor. Soc., 161–164.

Shenk, W. E., V. Suomi, F. Von Bun, F. Hall, W. Esaias, and M. Maxwell, 1986: Ideas for a future earth observing system from geosynchronous orbit. Preprints, *Second Conf. on Satellite Meteorology/Remote Sensing and Applications,* Williamsburg, VA, Amer. Meteor. Soc., 508–513.

Suomi, V. E., R. Fox, S. S. Limaye, and W. L. Smith, 1983: McIDAS III: A modern interactive data access and analysis system. *J. Appl. Meteor.,* **22,** 766–778.

Walpole, R. E., and R. H. Myers, 1978: *Probability and Statistics for Engineers and Scientists.* Collier Macmillan Publishers, 580 pp.

Weiss, S. J., and D. L. Kelly, 1982: An examination of tornado and severe thunderstorm occurrence in relation to the lead time of NWS severe local storm watches. Preprints, *12th Conf. on Severe Local Storms,* San Antonio, TX, Amer. Meteor. Soc., 463–466.

——, ——, and J. T. Schaefer, 1980: New objective verification techniques at the National Severe Storms Forecast Center. Preprints, *Eighth Conf. on Weather Forecasting and Analysis,* Denver, Amer. Meteor. Soc., 412–419.

Wilson, J., R. Carbone, H. Baynton, and R. Serafin, 1980: Operational application of meteorological Doppler radar. *Bull. Amer. Meteor. Soc.,* **61,** 1154–1168.

Xiang, X., and S. K. Beckman, 1985: The analysis of surface moisture convergence/divergence fields and severe weather based on certain satellite cloud patterns. Preprints, *14th Conf. on Severe Local Storms,* Indianapolis, IN, Amer. Meteor. Soc., 93–96.