U.S. DEPARTMENT OF COMMERCE
Office of Inspector General

PUBLIC RELEASE

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

NWS's Verification System for Severe and Hazardous Weather Forecasting Needs Modernization

Inspection Report No. IPE-9255 / January 1998

Office of Inspections and Program Evaluations

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The National Weather Service (NWS) helps protect the nation's people and property from the dangerous effects of severe and hazardous weather by issuing severe weather and flood warnings, public forecasts, and advisories. Because severe storms can injure so many people and cause enormous economic damage and emotional distress, NWS's ability to issue reliable, accurate, and timely warnings and forecasts is crucial.

NWS determines how well it handles its forecasting and severe storm warnings through its verification process - a quality control process that essentially matches warnings to actual weather observations and compiles statistical results of forecasting performance. Specifically, severe weather warnings are issued by NWS forecasters at the 115 Weather Forecast Offices (WFOs) spread across the United States, Guam, and Puerto Rico. After issuing severe storm warnings, the WFOs collect information about the actual weather to "verify" the accuracy of the warnings. This verification creates a baseline of skill or accuracy against which later changes in forecast procedures and products can be measured, and also helps NWS officials and staff (1) measure NWS performance, (2) answer congressional, media, and other requests for information, (3) publish a historical climatological record (Storm Data), (4) monitor trends, and (5) improve forecaster performance.

The National Oceanic and Atmospheric Administration (NOAA), through NWS, is working to significantly improve short-term warning and forecast products for thunderstorms, tornadoes, hurricanes, floods, tsunamis, and geomagnetic storms. In pursuit of this goal, NWS has made major advancements in the 1990s with new technology and modernized operations. Improved radars, new computer systems, and other technological advances associated with NWS's modernization efforts have substantially improved its access to critical data and its ability to forecast weather events. But modernization has also increased the workload for a decreasing NWS workforce because more detailed data is used in forecasting and twice as many severe thunderstorm and tornado warnings are being issued now as were in 1990. As a result, NWS's severe storm verification process - which has remained largely unchanged since its inception in 1979 - now needs updating to better reflect NWS's modernized technology and field office structure, staffing, and workload requirements.

In conducting this inspection, we focused on (1) assessing whether NWS's severe storm verification process and statistics are valid and reliable measures of NWS severe storm forecasting performance and (2) determining whether NWS modernization efforts have improved the accuracy of such forecasting. NWS routinely claims that its modernization efforts, and its improved recruitment and training of a cadre of weather volunteers, have resulted in more accurate forecasting of severe weather and improved verification scores. Based on our evaluation, we agree that (1) NWS's modernization efforts have resulted in more accurate forecasting; (2) its verification statistics are vital indicators and, for the most part, valid measures

of performance; and (3) non-technological reasons, such as the more effective recruiting of spotters, greater reliance on HAM radio networks, and improved coordination with emergency, state, and local managers, have also resulted in better forecasting and verification. However, we believe that the reliability of NWS's verification statistics and process can be improved.

As a key part of our evaluation, we conducted a comprehensive survey that included (1) sending a questionnaire to the warning coordination meteorologists (WCMs) at all forecast offices[1] and (2) subsequently interviewing a representative number of WCMs on their questionnaire comments and related matters. WCMs are the primary liaisons between the WFOs and the external user community. Our survey questionnaire - which we developed in consultation with senior and cognizant NWS personnel - provided considerable insight into the NWS process and procedures for verifying severe and hazardous weather. With a 91 percent response rate to our questionnaire, we were able to benefit from an exceptionally broad range of experience and opinions for this part of the review.

Most NWS forecasters we surveyed believe the verification process is important but think that too much time is spent on verification, especially of small storms. With NWS offices continuing to be downsized, some question the thousands of staff hours devoted to verification at a time when it is increasingly difficult to perform the basic forecasting mission.

- The WCMs overwhelmingly see verification as important and beneficial, but they are concerned about the lack of time and staff for verification. The meteorologists stated that time spent obtaining weather event reports during and after storms for verification purposes detracts from performing basic forecasting, improving forecast skills, and receiving training.

- Although forecasters spend thousands of staff hours on verification, it is important to note that the total time reportedly spent on verification equals only about 20 full-time equivalent positions, a seemingly small number compared to the hundreds of forecasters.

- The majority of verification time is spent during the four- to six-month severe weather season when forecasters are most busy alerting the public to hazardous weather.

- Many hours are spent on verification during and immediately after storms, when the WFOs are busiest with observing and forecasting severe weather.

[1] *OIG Survey Questionnaire for Warning Coordination Meteorologists.* See Appendix A for survey questions and responses.

- Historically the verification program has lacked a timely feedback mechanism, and has been hampered by poor data input procedures, inadequate quality control, and antiquated data processing methods.

Based on the survey and other evaluation efforts, we have concluded that the verification of severe and hazardous weather forecasting can be done more efficiently. We believe that with certain key changes and improved internal controls, NWS can significantly enhance its verification data credibility and dependability.

- **NWS should increase quality control to reduce the subjectivity of verification information**. We found that (1) WFO event information is susceptible to inaccuracies and uneven quality; (2) WFOs lack a standardized review process; (3) local storm reports are consistently not entered in the correct format (which reduces the quality and timely dissemination of data to NWS and outside users); (4) verification training for NWS employees and its SKYWARN volunteers is inadequate; and (5) NWS automated data checks of verification data are incomplete. (See page 12.)

- **NWS should implement a real-time verification system and reduce the *Storm Data* backlog.** NWS is not (1) providing rapid feedback to WFOs, (2) eliminating redundant keying-in of local storm reports and event information, or (3) reducing the large backlog of *Storm Data* publications. NWS needs to implement a modern, real-time verification and data collection system and database, and reduce the *Storm Data* backlog. (See page 25.)

- **NWS should reassess wind and hail warning thresholds**. The number of warnings that just meet the minimum threshold level for verification has increased significantly over the last five years. They represent a disproportionately high percentage of the NWS/WFO severe and hazardous weather verification workload. This increase is caused in large part by NWS's new improved radars that detect precipitation and measure wind speed and direction with greater accuracy. As a result, NWS needs to thoroughly review its current warning thresholds and process to ascertain whether the number of marginally severe storm warnings issued can be reduced without any increased danger to public safety. (See page 30.)

- **NWS needs to strengthen its headquarters and regional office oversight roles and responsibilities**. NWS's Office of Meteorology has not actively coordinated or addressed national, regional, and local verification issues. The Office of Meteorology should reestablish its National Verification Committee to evaluate all of its programs, revise its verification manual to include new verification techniques, and revise its National Verification Plan to update the requirements and goals of each verification program area.

Moreover, the regional office role in the verification process should be strengthened to improve quality control and provide additional verification information. (See page 35.)

- **The National Hurricane Center (NHC) should expand its verification efforts and test its emergency backup preparedness**. NHC does not verify the accuracy of its forecasts to the actual hurricane or the wind data of each hurricane. NHC needs to (1) systematically verify and document the portion of each warning area that did or did not receive hurricane force winds and (2) systematically verify its hurricane model wind radii forecasts. In addition, it has been two years since NHC last tested its backup plan, which details how NHC would operate in an emergency. NHC should test it during the next off-season. (See page 39.)

On page 44, we offer a series of recommendations to address our concerns.

In response to our draft report, NOAA's Acting Chief Financial Officer/Chief Administrative Officer agreed with the report's findings and recommendations. NOAA has several actions underway to address our recommendations. The report includes comments from NOAA's December 1, 1997, response to our draft report. A full copy of that response is attached as Appendix B.

INTRODUCTION

The Office of Inspector General conducted an inspection of the National Weather Service's (NWS's) severe storm verification program to assess whether NWS's severe storm verification process and statistics are valid and reliable measures of NWS severe storm forecasting performance and whether NWS modernization efforts have improved the accuracy of such forecasting. Because severe storms can injure so many people and cause economic and emotional damage, NWS's ability to issue reliable, accurate, and timely warnings and forecasts is important. NWS determines how well it does its forecasting through its verification process, which matches warnings to observations and compiles statistical results of forecasting performance. NWS's severe storm verification information is gathered by the 115 Weather Forecast Offices (WFOs) across the country and then compiled by the Office of Meteorology (OM) in Silver Spring, Maryland. We conducted the inspection from November 15, 1996, to May 30, 1997, to determine whether verification operations - such as quality control, automation, and managerial oversight - were adequate.

Inspections are special reviews the OIG undertakes to give agency managers timely information about operations, including current and foreseeable problems. By highlighting problems, the OIG hopes to help managers move quickly to address them and to avoid similar problems in the future. Inspections are also conducted to detect and prevent fraud, waste, and abuse and to encourage effective, efficient, and economical operations. This inspection was conducted in accordance with the Inspector General Act of 1978, as amended, and the Quality Standards for Inspections issued by the President's Council on Integrity and Efficiency. At the conclusion of the inspection, we discussed our observations and recommendations with NWS's Assistant Administrator for Weather Services, Deputy Assistant Administrator for Operations, and Deputy Assistant Administrator for Modernization.

PURPOSE AND SCOPE

The purpose of our inspection was to assess whether NWS's severe storm verification process and statistics are valid, reliable measures of NWS severe storm forecasting performance. The scope of our inspection included (1) evaluating NWS's verification program, including reports, databases, and quality control procedures; (2) evaluating verification data for statistical accuracy; (3) surveying all Warning Coordination Meteorologists (WCMs) about the verification process and accuracy of forecasting; (4) contacting 20 meteorologists-in-charge (MICs) and numerous WCMs and forecasters; and (5) contacting Department of Defense personnel, private weather service operators and professional meteorologists, and international weather organizations to compare and contrast NWS policies, practices, and procedures.

We examined background documentation relating to NWS's mission, budget requirements, operating procedures, and management plans. We also examined specific documentation relating to NWS's OM, regional offices, WFOs, and the National Hurricane Center (NHC). We examined OM's Strategic Operating Plan and National Verification Plan; the regional offices' analyses of storm data event reports, which ultimately become the final verification results published by NWS; WFO warning and event logs, WCM Handbook, and storm data event reports; *Storm Data* information for four WFOs for July 1996; and NHC track and intensity verification statistics, NHC's backup plan, and NHC's Draft 1997 Strategic Plan.

We conducted most of our inspection work at six NWS sites: (1) NWS headquarters in Silver Spring, Maryland; (2) the Aviation Weather Center in Kansas City, Missouri; (3) the Storm Prediction Center in Norman, Oklahoma; (4) the National Hurricane Center in Miami, Florida; (5) the Eastern Region Headquarters in Bohemia, New York; and (6) the Central Region Headquarters in Kansas City, Missouri. We also visited seven WFOs in the eastern, central, and southern regions. WFO forecasters provided information on local operations and verification procedures, and WCMs and focal points provided answers to our OIG Verification Questionnaire. Our survey questionnaire - which we developed in consultation with senior and cognizant NWS personnel - provided a great deal of insight into the NWS process and procedures for verifying severe and hazardous weather. In addition, Air Force and Navy liaisons to the Department of Commerce provided information and statistics on Department of Defense verification programs, and Finnish and Australian experts provided background information on their verification programs.

BACKGROUND

NWS is responsible for protecting the nation's lives and property from severe storms by issuing severe weather and flood warnings, public forecasts, and advisories for the United States, its territories, adjacent waters, and ocean areas. With about 5,300 employees, NWS's ability to forecast the weather affects not only the lives and property of every American but also our nation's commercial interests. Catastrophic disasters resulting from tornadoes, hurricanes, and floods demonstrate the importance of providing timely weather forecasts and warnings. The United States reportedly experiences more severe local storms and flooding than any other nation. Eighty-five percent of all presidentially declared disasters result from severe weather events. In a typical year, the United States has about 10,000 violent thunderstorms, 5,000 floods, 1,000 tornadoes, and several hurricanes.

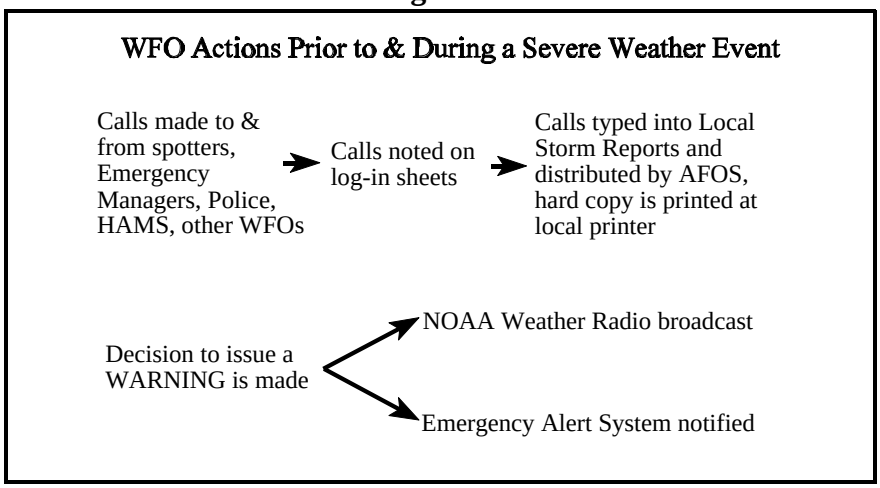
To improve the accuracy, timeliness, and efficiency of its weather forecasts and warnings, the National Oceanic and Atmospheric Administration, of which NWS is a part, has launched a major program to modernize NWS. NOAA's Modernization and Associated Restructuring (MAR) program, estimated to cost over $4 billion, is intended to modernize NWS weather observing, information processing, and communication systems. Such new technologies as the

Advanced Weather Interactive Processing System (AWIPS), the Next Generation Weather Radar (NEXRAD), the Automated Surface Observing System (ASOS), and the next generation Geostationary Operational Environmental Satellites (GOES) are key MAR elements. MAR will require significant changes in the current weather service infrastructure and operations, including verification.

Severe Local Storm Verification: What It Is & How It Is Done

Verification is the process of assessing the quality of forecasts by matching warnings to actual event observations and compiling measures of performance. Figure 1 illustrates the warning and initial event gathering that takes place *before* and *during* a severe local storm. During bad weather, the WFOs are in contact with people trained in identifying severe weather (referred to as SKYWARN spotters), emergency managers, law enforcement agencies, utility companies, and fire department personnel. Cooperative observers are another important source of information. In under-populated rural areas, WFOs often identify people who are willing to be telephoned by the WFOs about the weather. The importance of obtaining this information is to match what is happening on the radar to what is happening on the ground. At a minimum, the time, location, and weather observation are written onto a log-in sheet.

Figure 1



Time permitting, the event information is typed into a local storm report (LSR) and electronically sent to the Storm Prediction Center. WFOs print out a copy of LSRs for later use. Figure 2 is an LSR for Russell County, Alabama, for June 15, 1996.

Figure 2

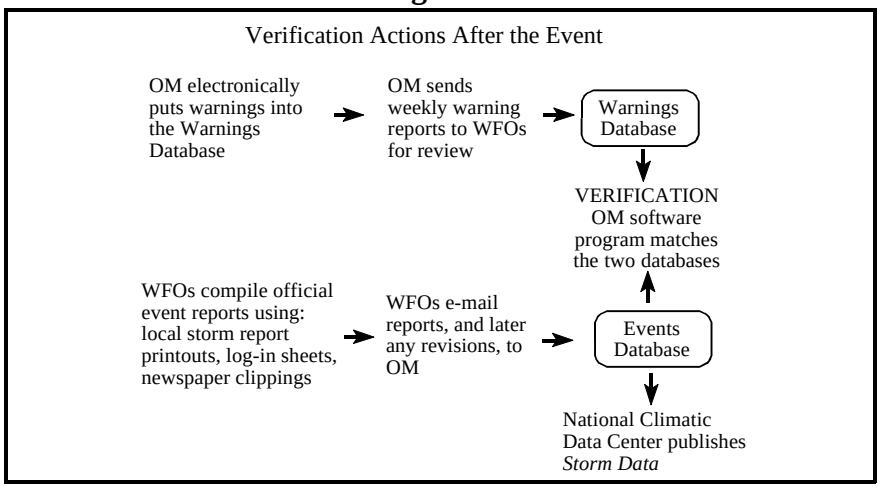
Local Storm Report

TIME (CDT)	.....CITY..... ....LOCATION....	STATE	...EVENT/REMARKS...
0600 PM 6/15/96	PHENIX CITY RUSSELL	AL	.75 INCH HAIL HAIL SIZE MARBLE TO DIME
0625 PM 6/15/96	SEALE RUSSELL	AL	WIND DAMAGE. SEVERAL TREES DOWNED NEAR FIRE DEPARTMENT ...POWER POLE DOWNED...TREE DOWNED ON HOUSE NEAR POST OFFICE. GENERALLY OCCURRED 6:25 TO 6:30 PM

After a decision is made to issue a warning (which can be at any time in the aforementioned process), the information is distributed to NOAA Weather Radio, AFOS, and the Emergency Alert System (EAS). Each WFO has the capability to record and broadcast NOAA Weather Radio information. EAS, formerly called the Emergency Broadcast System, is used by federal and local authorities to warn the public of an immediate threat to life and property. The system consists of all AM, FM, and network and cable television stations in a given area. The new system requires new equipment and offers information in both audio and digital format. The responsibility of the WFO is to alert, via EAS, the media and emergency managers, whose responsibility, in turn, is to notify the public of the warning.

The majority of verification activities take place after the severe weather system has passed through the WFO's county warning area of responsibility. Figure 3 illustrates how the warning and the event databases are compiled. For warnings, the Office of Meteorology electronically collects

Figure 3



warnings from AFOS and electronically transfers them into a central database. OM sends to the WFOs a list of the warnings issued each week to review and correct. Compilation of the events database is much more time consuming. Within 24 hours after the event, WFOs make additional calls to spotters and cooperative observers in counties where warnings were issued and no reports were gathered. That information, in conjunction with LSR printouts, the log-in sheet of event reports received or made during the storm, radar images, and newspaper clippings are used to compile the official storm data event reports. The WCM, or a meteorologist who is designated as the focal point, enters the event information using the storm data software. The WFOs have 60 days to electronically submit the storm data file, containing severe weather events for one month, to OM. After receiving all WFO event reports, OM runs a computer program matching events to warnings and computes national and WFO verification statistics.

Once all of the event information is compiled for a given month, it is electronically forwarded to Asheville, North Carolina, where the National Climatic Data Center, responsible for archiving weather data, publishes *Storm Data*, the official record of severe weather and unusual phenomena. Monthly issues of this report contain a chronological listing, by state, of occurrences of storms and unusual weather phenomena. Reports contain information on storm paths, deaths, injuries, and property damage. An "Outstanding Storms of the Month" section highlights severe weather events with photographs, illustrations, and narratives. The December issue includes annual tornado, lightning, flash flood, and tropical cyclone summaries.

Recent Changes in NWS's Severe Local Storm Verification

In 1979, NWS formally began verifying severe thunderstorm and tornado warnings at 52 Weather Service Forecast Offices (WSFOs) and about 198 Weather Service Offices (WSOs). The National Severe Storms Forecast Center (NSSFC) in Kansas City, Missouri, already conducting verification of severe weather watches, began conducting verification of warnings. Each WSFO sent warning and event information to NSSFC, which retyped the information into its verification program. In January 1996, OM replaced NSSFC as the verification focal point and 115 WFOs replaced the 52 WSFOs and 198 WSOs. The only change made by OM was to automate the process more - using a relational database software program and utilizing electronic mail.

Severe Local Storm Verification Statistics

The verification process measures NWS's forecasting performance, and it plays a crucial role in helping the agency improve its warning performance. The verification methods in use today build on the very early conceptual and methodological work first formulated more than a century ago. Verification then, as today, is defined by a "2 X 2" contingency table that refers to two possible forecasts and two possible observations. Figure 4 shows the two possible forecasts as a warning issued or no warning issued, and the two possible observations as the event occurs or does not occur. Most weather falls in the "W" box, where there is no severe weather event and no warning is issued. The optimum outcome, if a warning is issued or if severe weather is observed, is the "X" box.

Figure 4

2 X 2 Contingency Table

		Warning	
		YES	**NO**
Event	**YES**	X	Y
	NO	Z	W

From this information, NWS computes the following performance measures and statistics.

- The False Alarm Ratio (FAR) is the fraction of all warnings (X + Z) that are unverified (Z). A high FAR indicates that warnings are being issued but no event is taking place.

  $$FAR = \frac{Z}{(X+Z)}$$

- Probability of Detection (POD) is the fraction of all events (X +Y) that are warned events (X). POD measures how well events are covered by warnings. To achieve a high POD, warnings could be issued to cover the possibility that an event will occur, but that would be at the expense of the FAR.

  $$POD = \frac{X}{(X+Y)}$$

- Critical Success Index (CSI) is the fraction of all events plus the number of unverified warnings (X + Y + Z) that are warned events (X). Thus, the Critical Success Index is a function of POD and FAR.

  $$CSI = \frac{X}{(X + Y + Z)}$$

- Lead time is calculated by subtracting the time a warning was issued from the time an event reportedly occurs.

What Is a Severe Local Storm?

NWS has established thresholds for wind, wind damage, hail, and tornadoes to define severe local storms, and hence, what to issue warnings for. The thresholds are also used in the event collection process. Severe storm criteria are as follows:

- Thunderstorm wind gust of 50 knots (58 miles per hour) or greater.
- Thunderstorm wind damage that implies the occurrence of a severe thunderstorm.
- Hail equal to or greater than 3/4 inch in diameter.
- Tornado touching the ground and associated with a convective cloud.

To obtain verification, at least one of the above severe storm criteria must be observed along with the county location of the storm.

OIG Questionnaire Sent to Meteorologists Involved in Severe Storm Verification

To better understand past and current verification issues and problems, we sent a questionnaire to the WCM at each of the 115 operating WFOs. We received 105 responses, for a 91 percent response rate. Our 44-question survey asked about WFO verification methodology and practices and the strengths and weaknesses of the verification process. The questionnaires were answered by one individual or a mix of: WCMs, MICs, focal points for the storm data event information software, and/or junior meteorologists. If more than one person at a WFO completed a questionnaire, the senior person was included in the summary statistics. See Appendix A for a copy of the questionnaire and a summary of the responses.

National Hurricane Center Verification

NHC, one of seven national centers operated by NWS, issues hurricane watches, warnings, and forecasts for hazardous tropical weather conditions, primarily tropical cyclones,[2] for the Atlantic Ocean, Caribbean Sea, Gulf of Mexico, and the Eastern Pacific Ocean. Although NHC also prepares and distributes aviation, marine, and military advisories, providing hurricane information to the public, businesses, and federal, state, and local governments is its major responsibility. NHC coordinates with other NWS forecast offices; other federal, state, and local preparedness officials; and international weather service and government officials when tropical cyclones threaten the United States and foreign countries within its area of responsibility. NHC has an annual budget of about $2 million per year and 41 full-time employees.

NHC verifies its two major forecasts (track and intensity) by comparing the actual post-analysis of all track and intensity data. NHC issues a 72-hour track and intensity forecast every six hours for all tropical cyclones in the north Atlantic and Eastern Pacific oceans. NHC measures forecast accuracy by identifying forecast errors in nautical miles for track forecasts and knots for intensity forecasts. Track forecast errors are determined as the distance between a forecast position and the actual observation for the same time. Intensity forecast errors are the difference between the forecast and actual wind speed for the same time.

[2]A tropical cyclone is the general term for all circulating weather systems over tropical waters including tropical depressions (38 mph winds or less), tropical storms (39 to 73 mph winds), and hurricanes (74 mph winds or more).

OBSERVATIONS AND CONCLUSIONS

I. NWS Advances Associated With its Modernization Have Improved Weather Forecasting

NWS routinely claims that its modernization efforts, and its recruitment and training of a cadre of weather volunteers, have resulted in improved verification scores[3] and more accurate forecasting of severe and hazardous weather.[4] Based on our evaluation, we agree that: (A) NWS's verification statistics are vital indicators and, for the most part, valid measures of performance and (B) NWS's modernization, including technological and non-technological efforts, has resulted in more accurate forecasting.

A. NWS's verification statistics appear to be valid measures of performance

NWS primarily uses its calculated POD, FAR, and the CSI for its statistical and performance measurement reporting. (See page 6.) Based on our evaluation, including our discussions with a wide range of interested and involved parties and the survey results, we have concluded that the statistics derived from NWS's verification process appear to be valid measures of NWS's performance. Hence, we concur with a June 1996 General Accounting Office (GAO) report[5] that cites NWS's short-term warning and forecast and the hurricane warning processes as good examples of results-oriented performance measures. GAO has found that successful agencies have performance measures that are tied to program goals and can demonstrate the degree to which the desired results were achieved. NWS's mission depends on the quality and timeliness of observations, assessments, and information delivered for its warnings and forecasts. As a result, rather than simply count the number of forecasts it made, NWS determined that the most important business of its short-term warning and forecast weather services was to predict the time and location of weather events and to do so accurately.

---

[3] *Severe Local Storm Warning Verification for 1995,* NOAA Technical Memorandum SPC-1, June 1996.

[4] *National Weather Service Warning Performance Based on the WSR-88D,* Paul D. Polger, Barry S. Goldsmith, Richard C. Przywarty, and Joseph R. Bocchieri, February 1994.

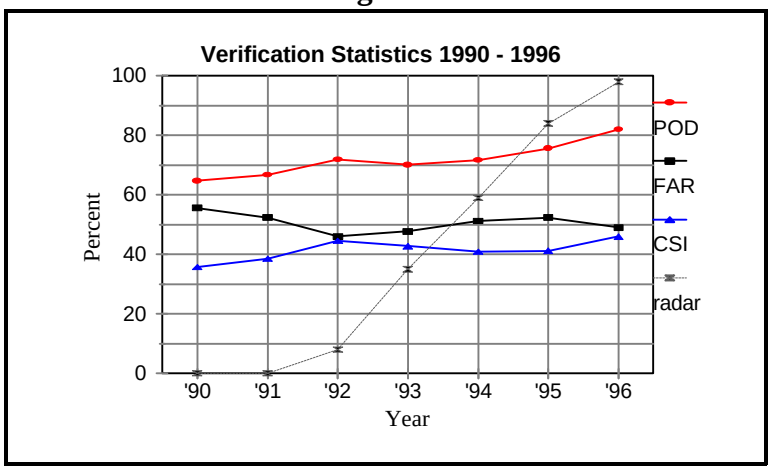
[5] *Effectively Implementing the Government Performance and Results Act,* June 1996, GAO/GGD-96-118.

NWS - wisely in our opinion - began to measure the extent to which it could increase the lead time it gave the public before severe events and the accuracy of its warnings. Ideally, for every severe weather event that occurs, a warning is in place, far enough in advance, for the public to take proper actions. A 100 percent probability of detection, the best possible outcome, means that for all severe events reported, a warning was in place. Lead time, the number of minutes between a warning and the event, is an appropriate measure of timeliness. A low false alarm rate (FAR) signifies that no "over-warning" is taking place. And, as previously mentioned, the critical success index (CSI) is a function of the POD and FAR, and the higher the number, the better the outcome.

B. NWS's modernization has resulted in more accurate forecasting

NWS has made major advancements in the 1990s with new technology and modernized operations. Improved radars, new computer systems, and other technological advances associated with NWS's modernization efforts have substantially improved its access to critical weather data and ability to forecast. Notwithstanding our belief that the statistics on POD, FAR, and CSI are valid measures, we initially questioned whether modernization efforts have resulted in more accurate forecasting. We subsequently determined that there has been a positive impact. Figure 5, for example, charts the POD, FAR, and CSI, for severe local storms from 1990 through 1996.

Figure 5



The dashed line, labeled radar, shows the implementation of the 120 radars that are a key element of the modernized NWS field structure. From 1992 through 1995, when the bulk of the radars were being delivered and accepted, the POD improved, increasing from 72 to 76 percent, while the FAR, for which an increase is undesirable, worsened, rising from 46 to 52

percent. We believe that the increase in the FAR or over warning occurred because forecasters were experiencing a learning curve with the new radars. Because the new radars detect precipitation and measure wind speed and direction in greater detail, forecasters were issuing warnings to cover the possibility that a severe event would occur. From 1992 to 1995, while the radars were being implemented, the number of warnings increased substantially. However, after four years, we believe that the forecasters' learning curve is about over. To support this, the POD and FAR both improved in 1996. Overall, the POD from 1990 to 1996 has increased from 65 to 82 percent. More importantly, the 1996 FAR, while still high, is lower than the pre-modernization scores, indicating that technology has improved forecasting.

While the emphasis of this discussion has been on NWS's new radars, it should be noted that several non-technological pieces of the modernization also have influenced the verification scores. For example, the new WCM position, that includes verification as a primary responsibility, focuses the WFO on the importance of collecting verification information. In addition, a number of new initiatives have been implemented in the past three to five years, according to our survey. Over half of the WFOs stated that they have taken steps to recruit more spotters. Other actions mentioned include expanding or starting HAM radio spotter networks, aggressively calling spotters to obtain ground truth reports, and coordinating more with emergency managers.

Conclusions

Our review of the verification of severe and hazardous weather forecasting found that NWS's verification statistics are valid measures of performance and that NWS's modernization, including technological and non-technological efforts, has resulted in more accurate forecasting. However, while the performance measures demonstrate that NWS appears to be making significant progress in attaining its mission, a key challenge for NWS and NOAA, is to show the impact of future funding requests on performance goals. The Government Performance and Results Act of 1993 (GPRA) states that "congressional policy making, spending decisions and program oversight are seriously handicapped by insufficient attention to program performance and results." As a GPRA pilot project, NOAA has attempted to demonstrate the impacts of different funding levels on its level of performance in its Fiscal Year 1998 budget request. However, we did not analyze this performance budget or determine if increased funding would result in higher levels of performance.

II. NWS Should Increase Quality Control to Reduce the Subjectivity of Verification Information

NWS's severe storm verification system has historically been hampered by poor data input procedures, inadequate quality control, antiquated data processing methods, and the lack of timely feedback to forecast offices. Our review identified four areas that have contributed to NWS event information being susceptible to uneven quality and inaccuracies: (A) lack of a standardized review process and data quality problems, (B) inadequate verification training of NWS employees and SKYWARN volunteers, (C) incomplete automated data checks, and (D) inadequate use of standardized local storm report software.

A. Lack of a standardized review process and data quality problems

Because NWS collects actual weather event information from a wide variety of sources, including the emergency managers in cities, NWS's "spotters," newspapers, and the general public, the quality of the information received can be as varied as the sources. Hence, information such as wind speed, time of event, and damage estimates vary widely, and experience has shown that it is highly subjective. In addition, the same people who issue warnings are also responsible for documenting that the severe weather actually occurred. Given the pressure to verify, this practice seems to have a built-in conflict of interest or at least gives the appearance of impropriety. However, most of the WCMs we interviewed stated that the same meteorologist does not "usually" issue and verify the same warning. Hence, they do not believe that this issue is a problem. But, here again, this situation highlights NWS's vital need for a standardized review process.

Our survey showed that while 50 percent of the WCMs review, to some degree, storm data event reports, warnings, log sheets, and Local Storm Reports, there is no standard process in place. Twenty-five percent of the offices make follow-up phone calls to spotters and other observers, and the other 25 percent take other steps or no steps at all. Although the WCM manual[6] requires WCMs to analyze and synthesize weather event information and prepare the monthly storm data event report, it does not describe how monthly data should be reviewed. As a result, the WFOs have different collection methods. More importantly, some NWS personnel stated that WCMs are poorly trained in verification and have minimal resources for quality control and post-event surveys.

[6] *Warning Coordination Meteorologist: Job Aid,* Revised August 1996.

WFO Data Problems

Although NWS stated that quality control procedures are used to ensure reliability,[7] we determined that the WFOs have five serious data quality problems: (1) they do not record the source of all events, (2) their wind and hail information is imprecise, (3) their event information is misleading, (4) the event information is sometimes recorded incorrectly, and (5) meteorologists sometimes verify their "own" severe and hazardous weather warnings thereby creating a potential conflict of interest.

1. WFOs do not record the source of all events

Based on our survey, 94 percent of the WFOs keep a log-in sheet of all calls received from outside parties and made by forecasters. We discovered that the source of the calls is not necessarily recorded, thus preventing a reassessment of the event information at a later time. Representatives of some WFOs estimate that 10 to 50 percent of the incoming calls are missing source information, such as a name or spotter number.

In reviewing copies of WFO log-in sheets, we also found that some user identifications and locations were missing. WFO MICs and WCMs stated that this occurs because forecasters do not have time during events to record all the necessary information. With 21,000 severe events in 1996, this clearly represents a serious quality control problem. To be effective, a system's input controls must ensure that data origination, authorization, and compilation are adequately controlled. But, by not recording the source of calls, NWS cannot systematically confirm or otherwise validate the accuracy of the data that is the basis for its national verification scores.

Although NWS's 115 WFOs need flexibility in inputting and reviewing event data, NWS can strengthen data accuracy by establishing a more standardized verification process. After discussing this problem with numerous MICs, WCMs, and forecasters, we have concluded that NWS needs to increase data integrity by requiring each WFO to record the source of event information. Some WFOs use recorders attached to telephone lines or pink-colored log-in sheets placed at every telephone that record the original source and time of event information to facilitate this process.

2. WFO wind and hail information is imprecise

Storm Data, which is the official record of severe weather published by NWS, describes the type of severe event; the time; the event characteristics, such as wind speed and hail size;

[7] *National Weather Service, Draft FY 1997 Annual Operating Plan,* October 1996.

and the damage estimates. *Storm Data* is used by researchers, insurance companies, and OM for research studies, processing insurance claims, and accumulating a climatological database. Forecasters compile storm data event information by analyzing reports received from personnel who have received spotter training, but just as often, from other non-trained individuals. Most individuals provide their judgment of severe events without instruments to measure wind speed and hail size. Because uncertainty exists about actual values assigned to events, such as maximum wind speed, the value of property damage, maximum hail size, and the time of an event, information is imprecise. NWS has acknowledged inconsistencies and uncertainties in the estimates received. As a result, the national database of wind speed and damage and hail lacks sufficient precision.

For wind speed and damage, more than 75 percent[8] of the events lack estimates. Over the last 10 years, 49,000 out of 69,000 wind events in *Storm Data* were labeled as "thunderstorm wind," which can mean anything from a large tree limb down to major structural damage. This shows how imprecise and subjective the reporting of events is, especially when it comes from such a diverse reporting group, including trained SKYWARN spotters, NWS meteorologists, or other public officials on one end to members of the general public on the other. The remainder of wind events that are assigned values consist of measured and estimated gusts, but the database does not distinguish between them. More importantly, NWS personnel stated that some WFOs appear to have arbitrarily assigned wind gusts values (50 knots for minimum verification) to reports of wind damage, making it more difficult to distinguish between actual wind gusts and assumed gust speeds.

After researching historical wind-related verification problems and speaking with WFO personnel, we believe that the severe event database should distinguish between measured and estimated wind gusts to increase objectivity and permit more rigorous analysis. Because the database does not distinguish between measured and estimated gusts, outside users perceive the estimated wind amounts as actual measures. While distinguishing between measured and estimated wind gusts, the severe event database should not force the WFOs to assign estimated gusts to all damage events. If a damage estimate is known, but wind speed is unknown, 50 knot winds should not be arbitrarily assigned. WFO personnel stated that the database lacks real wind values. They stated that although the database contains many estimates, once the estimates are in the database they become accepted as actual values. While OM now requires WFOs to provide a wind estimate and/or damage estimate, OM has not addressed the need to distinguish between actual and estimated in the database.

[8]*Severe Local Storm Climatology 1955-1995: Analysis of Reporting Trends and Implications for NWS Operations,* Steven J. Weiss and Michael D. Vescio, National Weather Association, 21st Annual Meeting, December 1996.

For hail, determining the actual size can be highly subjective because most observations are not measured. For example, with the pressure to verify warnings, WFOs stated that they call various sources and ask whether hail is dime size (approximate size for verification). However, such questions by forecasters are leading and very subjective. As a result, this creates a level of imprecision similar to wind speed estimates. From our survey, the WFOs confirmed the level of imprecision of hail estimates, but more importantly, stated that the increased number of low hail events requires considerable verification time. Specifically, the number of low and intermediate hail events (1.75 inches and less) has increased, while large hail events have remained relatively even. With NWS's new radars that detect smaller storms, warnings, and consequently events, have greatly increased. To reduce the subjectivity in hail size, we believe that (1) precise hail measurements should be taken and (2) the hail threshold needs to be evaluated (see page 30 for a discussion of the hail threshold issue).

NWS's SKYWARN program trains hundreds of weather spotters across the country to identify, record, and report the effects of severe thunderstorms and tornadoes based on the severe storm criteria. The WFOs train and retrain these spotters all year long. It is important that the WFOs continue to emphasize the importance of accurate data reports in spotter training. For example, WFOs should recommend that all trained spotters use measured - not estimated - hail reports to numerically verify warnings.

The NWS Director's Advisory Committee on Forecast Operations stated that spotters could be supplied with standardized calipers to provide accurate hail reports.[9] Although NWS personnel and spotters questioned whether NWS had funding to provide inexpensive calipers, some spotters already use inexpensive calipers or rulers to obtain more accurate measurements. The Advisory Committee even suggested that rulers or calipers could be supplied in spotter sessions by corporate sponsors. The Committee also suggested that spotters could provide hail estimates to WFOs during hail storms and then provide more accurate measurements afterwards. We agree with both of the Committee's recommendations. NWS should look into the merits and possibility of purchasing or obtaining inexpensive calipers or rulers to accurately measure hail size.

3. WFO event information is misleading

Because the WFOs issue warnings at precise times, determining the actual event time, or the best estimate, is very important. Verification of a warning and amount of lead time depend on when an event occurs. Information is misleading because the time of an event is often the most difficult aspect to verify.

[9] *Director's Advisory Committee on Forecast Operations Draft 1997 Report*, March 1997.

Warnings are verified by spotter and cooperative observer reports meeting the NWS warning criteria (see page 7) and occurring within the valid time and location of a warning. However, on a single event, a WFO may receive multiple reports with times varying by 30 minutes or more. This occurs since most damage reports are relayed after storms, and those making the reports are not always sure when the damage occurred. To be as accurate as possible, external sources such as emergency managers could review verification data; WCMs could spot check storm data event reports and "sign off" as completed before sending them to OM; WCMs and forecasters could compare the timing of storm reports with radar imagery, randomly reviewing reports from nearby counties for time continuity (cross-checking can reveal potential errors in report times); and WCMs and the *Storm Data* focal points could randomly review changes made to monthly storm data event reports.

4. WFO event information is recorded incorrectly

The most damaging if not the most common problem is event information that is recorded incorrectly. Before a WFO records an event time on a log-in sheet, changes can be made. NWS has no internal control or software that prevents the changing of event times. Many WFO personnel stated that they knew "other" WFOs who changed reported event times to fall within the warning time frame. When an event time falls on or after the warning time, the warning is verified. A "miss" occurs when the event time is before the warning time or if there's no event. As a result, minutes make the difference between a verified and an unverified warning.

WFO personnel, in discussions and questionnaires, stated that since the raw data is collected locally, the "fudging" of information to improve verification scores exists. For example, if a spotter calls in and reports hail at 4:10 P.M., and a warning is issued, perhaps based on that report, at 4:14 P.M., the event time could conceivably be altered to meet the warning time frame. Forecasters often write down event times on small pieces of paper during an event and record the information at a later time. Once information is formally recorded, an audit trail exists, and any change would be documented and have to be justified. It would be difficult to impossible to find examples of forecasters "fudging" data because, as we were told, "the pieces of paper are thrown away." However, based on our interviews and discussions, we believe that some changing of data does exist.

To begin correcting this problem, some WFOs, as stated on page 13, have started using recorders attached to telephone lines or pink-colored log-in sheets that record the original source and time of event information. This is a "best practice" that could be replicated at all WFOs. The Baltimore/Washington D.C. WFO has developed Local Storm Report software (see page 18) that can automatically record the source and time of event information. Being tested during our inspection, the software has the potential to be used at all WFOs. NWS needs to determine the status and potential of this software as soon as possible. To

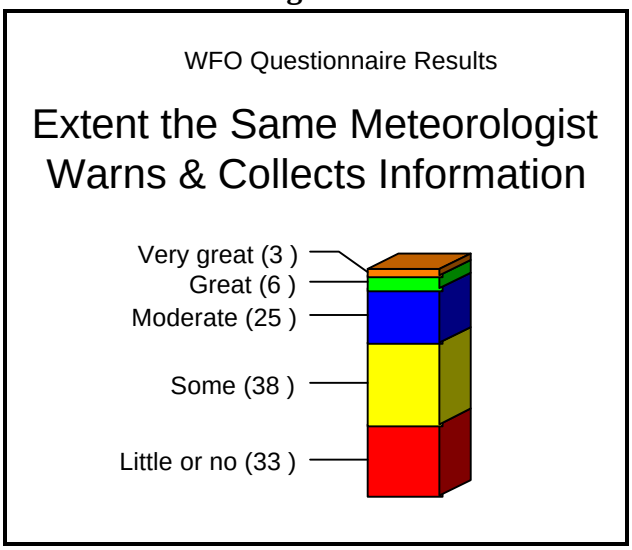
determine whether the WFOs changed data once it is logged in, we compared the June 1996 warnings and local storm reports to the June 1996 *Storm Data* reports for four WFOs. We were pleased to find that most warning and LSR times and events matched the *Storm Data* times and events. The remaining warning and LSR times and events differed as updated information became available.

5. Meteorologists sometimes verify their own severe storm warnings, thereby creating a potential conflict of interest.

Due to the relatively small number of staff at most WFOs and the methods they use to collect and verify information, many meteorologists issue and verify their own warnings. Obviously, this practice raises questions about a potential conflict of interest.

In discussing this practice with many of the warning coordination meteorologists who oversee the process, we found that they were not concerned about a potential conflict of interest because they believe that the same forecaster does not "usually" issue and verify the same warning. Based on our observations, we disagree with the WCMs' assessment. In our questionnaire, we asked each WFO the extent to which the same meteorologist issues, and then verifies, a given warning. Figure 6 shows the response of the 105 WFOs that answered our survey. Thirty-three WFOs estimated that zero to 10 percent of the warnings are verified by the meteorologist that issued them. Thirty-eight WFOs estimated that 11 to 20 percent are verified by the same meteorologist. Twenty-five WFOs stated that there was a moderate amount of overlap, with 31 to 40 percent of the warnings being verified by the same meteorologist. And finally, nine WFOs reported that as many as 41 percent or more of their warnings are verified by the meteorologist that issued them.

Figure 6



Clearly this data is at odds with the WCMs' impression of the situation. When we asked the WCMs in the survey whether this represents a conflict of interest, they overwhelmingly said no. As previously mentioned, we did not find any evidence that information was intentionally being altered to verify warnings and improve scores. Nonetheless, there is, at a minimum, an

appearance of a conflict of interest that should be addressed. The solution to this problem, however, is neither simple nor straightforward.

Ideally, an outside reviewer with no vested interest in the verification results would perform verification. But, realistically the added expense of implementing such an independent review system would be prohibitive. In addition, WFO personnel stated that the information from spotters and cooperative observers would continue to be gathered by forecasters during serious weather events to determine whether warnings need to be issued or extended. Thus, there was concern that it would be overly burdensome and confusing to the public if they had to call in or respond to both forecasters and independent verifiers at two telephone numbers regarding "ground truth." Here we agree with the WCMs that we spoke to who expressed concerns that this could disrupt WFO operations. Given staffing realities and the forecaster's potential conflict of interest of gathering information for dual purposes - issuing warnings and verifying events - we suggest that NWS institute less labor intensive and less operationally disruptive controls to reduce the likelihood of an actual conflict of interest.

A better solution would be to match and flag instances where the same forecaster issues and verifies a warning. One WFO, for example, is developing software that could identify cases where the same forecaster issued and verified a warning, thus flagging warnings that would be subject to a random third party verification check. The advantage of this software is that the match is done at the local level and flags would be raised early in the process. An intermediate solution, until such software is in place, is to add the name of the forecaster who collects event information to the storm data event database. Given that the warning forecaster is currently in the warning database file, when the two databases are matched by OM for verification purposes, the degree to which the same meteorologist is issuing and verifying the same warnings could be monitored and third party verification checks could be initiated.

Conclusions

NWS's 1997 Annual Plan states that "Quality control procedures are followed to ensure the highest possible reliability of the gathered data." Based on our observations, this is misleading. The Government Performance and Results Act of 1993, Public Law 103-62, will require NWS to document how its verification measures are derived and describe the means used to "verify and validate" its verification measures. But NWS lacks a standardized review process, and currently cannot ensure that it achieves the highest quality control. As a result, we believe that NWS needs to (1) move as quickly as possible to put in place quality control procedures to ensure the highest possible reliability of the gathered data and (2) describe in any future plan, testimony, or budget document how its verification measures are derived and their quality control limitations.

Verification quality control should be standardized by developing NWS guidelines at the national level and implementing them throughout the network. Specifically, we recommend that NWS reestablish its National Verification Committee (NVC), which oversaw verification roles and responsibilities in the 1980s (see page 35). NVC and the regional offices should then implement a standardized review process to ensure data integrity by:

- recording the source of event information,
- distinguishing between measured and estimated wind gusts,
- periodically sending verification data to external sources for review,
- spot-checking monthly data and "signing off" as completed before sending it to OM,
- comparing the time estimates of storm reports with radar imagery,
- randomly reviewing reports from nearby counties for time continuity,
- randomly reviewing changes made to monthly storm data event reports,
- using only measured hail diameters - not estimates - to numerically verify warnings,
- adding the name of the forecaster who collects event information to the storm data event database,
- developing software to flag warnings that are issued and verified by the same forecaster, and
- instituting third party reviews or verification checks when the same person has issued and verified a warning.

NWS agreed with most of our recommendations aimed at implementing a standardized review process. However, NWS opposes the addition of the name of the forecaster who collects event information to the event database and the development of software to prevent the same forecaster from issuing and verifying a warning. NWS believes that WFO staffing is too strained during severe events to prevent the same forecaster from issuing and verifying a warning. Although we agree that WFO staffing can become strained during severe events, we believe that NWS should still add the name of the forecaster who collects event information to the event database and develop software to flag when the same forecaster issues and verifies the same warning. By having this information and software available, the WFOs can, when feasible, better monitor who issues and verifies warnings and institute third party reviews or verification checks.

B. Inadequate verification training

Internal verification training for WCMs and forecasters and external training for spotters should improve data quality. Both groups need to receive current training on verification techniques and severe weather identification and reporting. Consequently, we included training questions in our survey and held numerous conversations with NWS personnel to assess the adequacy of training. Our survey determined that alternative training mechanisms were needed and specific data accuracy measures should be emphasized in spotter training.

WCMs receive some forecasting training soon after being selected, but training on verification event-gathering and measurement techniques, verification software, the type and severity of events, and on-site surveys has been mostly on-the-job. In response to our questionnaire, 33 percent of the WCMs listed weaknesses in the verification training they received, including: (1) too much emphasis on Eastern and Mid-Plains states weather rather than the more common forms of storms found throughout the country, (2) not enough instruction on how to write storm data event reports, and (3) inadequate guidance on how spotters and others should gather and report event information. Of the remaining 67 percent, 19 percent could not remember the verification part of the WCM course, 15 percent liked the training, and 33 percent either did not receive WCM training or did not answer the question.

More importantly, over half of the WCMs stated that additional verification training was needed, preferring on-site materials, such as CD-ROMs and videos, and discussions at the annual WCM conferences. Because of time constraints and limited travel funds, they believed that formal classroom training is not practical. The regional WCMs, who oversee and coordinate the verification activities in each region, suggested that region-specific training should be provided through CD-ROMs, videos, and discussions at the annual WCM meetings. NWS should provide WCMs with the alternative training mechanisms discussed above to promote data quality.

C. Incomplete automated data checks and data recording

NWS national verification of severe local storm warnings was done at the National Severe Storms Forecast Center in Kansas City, Missouri, before OM took over this responsibility in January 1996. NSSFC received all warnings and event information from the field offices, manually retyping paper information and performing no data edit checks. However, to improve quality control, OM changed the severe storm verification system in 1996 from a mainframe to a personal computer, implementing electronic transfer of warning and event information, thereby reducing the number of errors made from retyping. OM also reduced the number of warnings and event information lost through paper transmission.

OM now performs some automated data edit checks, which continually enhance data quality by detecting errors caused by omissions, invalid entries, and other inaccuracies. For example, data checks ensure that data fields such as hail size and wind speed fall within assigned parameters. Even with OM's current data checks, we determined that there are still two data quality problems with the current process: (1) complete automated data checks have not been established and (2) changes to monthly storm data event reports are not recorded.

1. Complete automated data checks have not been established

Although OM has established some PC-based data checks, additional data checks are needed. OM's PC-based programs use current database management software that allows on-line, immediate data checks. OM established data checks in early 1997 for hail size, wind speed, date of event, and location of event. However, the severe weather event database has many other data fields that need quality control. Figure 7 outlines some potential data fields and data checks that can be performed.

OM should implement additional automated tests of national verification information. OM plans to establish these and other data checks when contractor funds are available. However, even additional automated data checks cannot prevent all of the data input problems cited earlier. When completed, current and historical data fields can be edited for improved quality control. NWS has years of severe weather data that has not been subjected to automated data checks. Although it is unlikely that edit changes from historical data would change the overall statistics, additional data checks will correct gross errors and improve quality control. For example, in *Storm Data* recently, a damage level of $50 million, instead of $500,000, was recorded.

Figure 7

Office of Meteorology Potential Data Checks

Field	Test
Year	Check for years outside the boundaries (e.g., 1968 instead of 1986).
Month	Check for values <1 or >12.
Day	Check for values <1 or >31.
Time	Check for values <0 or >2359.
WFO	Check for names that don't match a fixed set.
Event	Check for nonexistent event types and ensure that type and magnitude are compatible (e.g., hail must be associated with realistic hail size and not wind speed by mistake).
Latitude Longitude	All events must have a latitude and longitude.
Deaths	Check for unusual numbers. Compare with severity of the event.
Injuries	Check for unusual numbers.
Damage	Check for unusual numbers.
Warning Issue Time	Check that warning issue time is not after or the same as the warning expiration time.

2. Changes to monthly storm data event reports are not recorded

After the WFOs submit their monthly event data to OM, OM compiles the information and then resubmits it back to each WFO for review. The WFOs may have received revised or additional information or initially sent incorrect information. For example, a WFO may have received information to verify an event, change an event time, or change a damage or hail

size estimate. Documenting changes to any database maintains quality control. However, OM does not keep a historical record of monthly changes to storm data event reports, which would help ensure that changes to data fields are properly authorized and documented.

After receiving monthly event data in a temporary file, OM transfers the data to the appropriate WFO file, performs some data checks, and compiles national statistics. OM returns monthly data to the WFOs for review. If WFOs have changes to their monthly storm data event reports, they are sent back to the temporary file. However, OM takes the temporary file and replaces existing data fields with the new data, without recording what changes have been made in the original data. Not recording data changes is a serious internal control weakness.

Although OM personnel believe that changes in monthly data are minimal, they have no data to support their claim. Even if monthly changes are minimal, OM should document WFO changes to event times, event size (for hail), event speed (for wind), and damage estimates to record possible trends, determine if changes are increasing, and document whether changes are authorized by the WCMs. In addition, OM should create a new field in the database that allows WFOs to identify the source of information (e.g., trained spotter, emergency personnel, media, law enforcement, general public). This will document the source and reliability of information.

D. Inadequate use of local storm report software

Although Local Storm Reports (LSRs) are draft unofficial weather reports, they contain important and timely meteorological information used by other WFOs, the media, and private meteorologists. LSRs contain preliminary event information, specifically the time, county, state, and type and size of event. They are prepared by NWS forecasters from information received from spotters, power companies, county managers, and other sources. Because LSRs are a major source of verification information, LSR quality is essential.

In 1994, NWS established a standardized LSR format[10] to replace different LSR software packages used by WFOs. This format, PC-LOGIT, developed by the Storm Prediction Center (SPC), electronically decodes LSRs sent in by the WFOs. When PC-LOGIT is used, vital storm information is rapidly disseminated to other WFOs, SPC, regional offices, media, and others. However, because the WFOs do not consistently use PC-LOGIT to input LSRs, quality is reduced and timely storm information to NWS and outside users is lost. NWS personnel estimated that about one-third to one-half of the offices use non-standard software to enter LSR information.

[10] *National Weather Service Operations Manual,* Chapter 40, March 1994.

Upon contacting NWS personnel to determine why PC-LOGIT was not being used by all offices, we learned that the SPC-developed initial version of PC-LOGIT contained several problems. For example, it was not WINDOWS-based, and its database was not compatible with the storm data event database. Specifically, there was a line capacity problem, certain towns were not included in the database, and the program crashed if too many reports were entered. After additional work in 1995, SPC personnel completed an updated version of PC-LOGIT and sent it to select WFOs for testing. In February 1996, the updated version was released to all WFOs, after SPC made other WFO recommended changes. However, about 30 to 50 percent of the WFOs still use different formatting software in place of PC-LOGIT to transmit LSRs. In late 1996, SPC was moved to Norman, Oklahoma, where PC-LOGIT changes have been overseen. SPC personnel stated that because some WFOs have different software, they have fought proper formatting techniques that would increase electronic decoding at SPC. SPC and some WFOs both stated that PC-LOGIT needs better marketing so that NWS uses one uniform formatting software package and database.

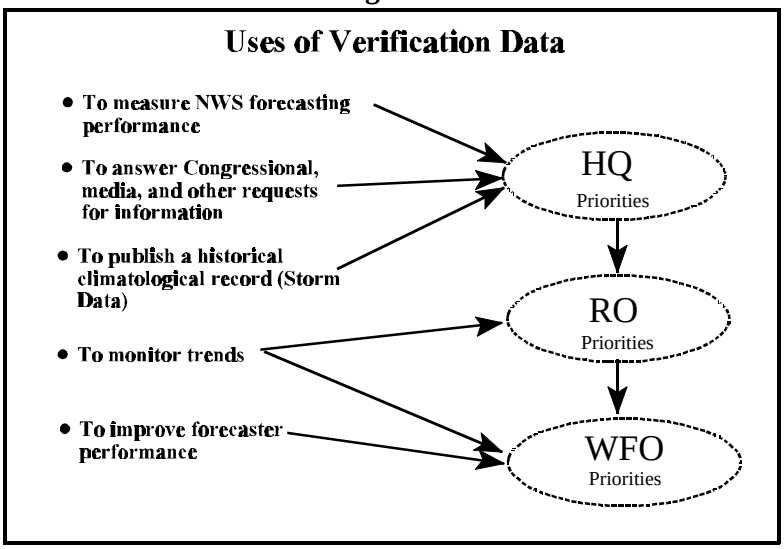
Because some WFOs do not use PC-LOGIT, SPC must manually reformat data. SPC's initial policy was to manually enter significant reports that were not properly formatted. However, because this formatting is labor intensive, SPC now limits its time spent reformatting reports. As a result, data errors occur and the SPC database is incomplete. More importantly, SPC prepares its nationwide LSR report without these office reports. The hourly, nationwide report gets sent to the WFOs, emergency managers, and private meteorologists. By not having all office reports, some WFOs are not fully informed about approaching weather, compromising their warning decision-making. SPC plans to continually upgrade the software and improve the compatibility with other programs, such as the storm data event software. While software changes are made during the next year, the regional offices should ensure that PC-LOGIT is received and used by all field offices, enabling all WFOs to receive SPC weather reports that affect their office.

NWS agreed with our recommendations to (1) determine the status of local storm report software being developed at the Baltimore/Washington D.C. WFO, (2) describe the means used to verify and validate its verification measures, (3) provide WCMs with alternative training mechanisms, (4) emphasize the importance of data accuracy in spotter training, (5) implement specific automated tests of national verification information, (6) document changes made to monthly storm data reports, (7) create a new field in the event database, and (8) require that PC-LOGIT be used by all WFOs.

III. NWS Should Implement a Real-time Verification System and Reduce the *Storm Data* Backlog

Within NOAA, verification information has a number of applications and uses (see Figure 8). At NWS headquarters, verification statistics are used to evaluate the agency's overall performance. The information is also used to respond to geographically related questions about weather raised by the Congress, the media, other Federal agencies, and insurance representatives, to name a few. The event information ultimately becomes the official record of severe and hazardous weather phenomena that researchers use for analyzing trends. Regional NWS offices primarily use the statistics to monitor WFO trends and performance. Finally, at the WFOs, verification statistics are used to appraise office performance and assess forecaster biases to under- or over-forecast in certain situations. Verification data, however, in accordance with a NOAA agreement with the employees union, is not to be used to prepare individual forecasters' personal performance ratings.

Figure 8



Our inspection identified two problems with NWS's verification process that have hindered the agency's ability to use the verification information most effectively or to improve or revamp the verification process: (A) verification requires redundant keying-in of LSRs and storm data event information and does not provide rapid feedback to WFOs and (B) publication of *Storm Data* is not timely.

A. Verification requires redundant keying-in of storm data event information and does not provide rapid feedback to WFOs

NWS's National Verification Plan identified six specific goals for the verification program.[11] We believe that the two key goals, "minimize workload, particularly at the local level" and "provide rapid feedback to forecasters and local management," have not been realized. To achieve the greatest benefit, the feedback should be received when the event is still fresh in the individual's mind.

Currently, during severe weather, WFOs enter local storm reports into the AFOS system. The purpose of LSRs is to alert SPC and neighboring WFOs of the characteristics of local weather. The weather reports are collected from people calling in and the WFO calling spotters and contacts for reports. LSRs are considered draft, unofficial reports. For verification purposes, 24 hours to two months after the severe weather event, paper printouts of the LSRs are used in conjunction with follow-up phone call log-in sheets and newspaper clippings, to compile the official storm report. Thus, to prepare the official storm data event reports, LSRs are entered into the storm data software program. We did not determine the full magnitude of the double entry of the same data, but it is significant. For example, we obtained LSRs and the storm data information for one office. Forty-four percent of the storm data event reports were identical to the LSRs, which means that changes have not been made.

OM has the responsibility to match the warning database to the event database. The WFOs have 60 days from the end of the month to send, by electronic mail, the storm data information to OM. Many WFOs send their reports in late. OM waits until all storm data reports are received because the matching process between the two databases is time consuming. As a result, several months elapse before the WFOs obtain verification feedback about their forecasting performance.

OM has made software improvements by automating the collection and compilation of storm data information. The current system takes the WFO storm data event file and electronically inserts it into the central event database. Previously, WFO event information was retyped into the central database. The automated process reduces the workload where the information is centrally compiled, but does not affect the WFO workload. WFO's must continue to type identical information into LSRs and the storm data event reports.

[11] *Report of the National Verification Task Team: National Verification Plan,* National Weather Service, June 1982.

Forecasters need to receive timely feedback on the accuracy of their severe storm warnings in order to help improve skills. They are primarily interested in evaluating their individual forecasting performance, identifying the strengths and weaknesses of their forecasts, and identifying any consistent bias to under- or over-forecast in certain weather situations.[12] An operational verification system should provide near-real-time access to forecasts and observations, so forecasters can quickly see where their warnings are on target and where they have been less accurate. This more instant feedback helps the forecaster improve his performance. Under the current process, forecasters cannot automatically compare areas warned with LSRs for timely visual feedback and preliminary verification. Entering LSRs, obtaining preliminary verification, submitting storm data event reports, and receiving official verification information are all separate functions. Forecasters obtain post-storm verification if they manually compare verifying observations to warnings issued. However, this is a time-consuming process for busy forecasters.

NOAA's Space Environment Laboratory (SEL) implemented a verification system with near-real-time access to its space forecasts and observations.[13] Another example of a real-time verification system is the Finnish Meteorological Institute's computer program that is used for its daily forecasts. The goal of the Finnish system is to give all duty forecasters access to a verification database via a user-friendly, graphical interface.[14] Neither SEL nor the Finnish system are directly transferable to the severe storm verification process, which relies on human reporting, not instruments, for corroboration.

However, NOAA has done considerable work, since 1993, preparing the requirements for a new verification process in AWIPS.[15] The AWIPS requirements specified WFO local verification capability rather than the current centralized verification. Forecasters would create, maintain, and archive separate warning and event logs to reduce the keying-in of information and provide rapid feedback. We found several cases where WFOs, with their limited resources, are developing software to eliminate the redundant entry of LSRs into the

[12] *An Operational Forecast Verification System at the Space Environment Laboratory,* Kent A. Doggett, NOAA Space Environment Laboratory, March 1997.

[13] *An Operational Forecast Verification System at the Space Environment Laboratory,* March 1997.

[14] *Real-Time Verification of Operational Weather Forecasts at the Finnish Meteorological Institute,* Perti Nurmi, The Evaluation of Space Weather Forecasts: Proceedings of a Workshop at Boulder, CO., June 1996.

[15] *Technique Specification Package 89-08-R1,* National Weather Service, Office of Systems Development, April 1, 1993.

storm data software and will provide preliminary verification scores. Although WFOs stated that they had waited many years for AWIPS, they have begun to develop their own in-house applications because of the AWIPS development delays.

We believe NWS needs to implement an integrated, real-time verification and data collection system and database as soon as possible. The system should accept LSRs, provide the foundation for storm data reports, and provide quick initial verification feedback to forecasters. In order to accomplish this, NWS will need to develop software for inputting LSRs that is compatible with its storm data software (precluding the need for double entry of event information). NWS should use the WFO development efforts mentioned above to determine which system, or combination of prototype systems, best meets current requirements.

B. Publication of the official *Storm Data* product is not timely

Event information collected during the verification process becomes the official record of severe weather events and unusual phenomena. By the time NWS publishes *Storm Data*, the official record, the information is nine months old. For example, *Storm Data* for October 1996 was printed in mid-July 1997.

The *Storm Data* backlog has been steadily increasing since January 1996 for three reasons. First, when the national verification program operations moved from NSSFC to OM in January 1996, there was a publication delay. Second, the transition to new verification software caused a publication backlog. Third, some WFOs are late with their data submissions, and OM waits until all, or a majority of, reports are in before running the software program that matches the warnings and events databases.

We believe that with the transition from NSSFC to OM completed, the backlog should be reduced. In fact, between the new software capability and the OM staff's resourcefulness, OM has expedited the process by pre-formatting event information and sending a print-ready file to NCDC. However, late data submissions by the WFOs continue to be a problem. We do not believe the WFOs need more time, as they already have 60 days from the last day of the month to submit monthly event data. In our survey, 79 percent of the WFOs stated that this was enough time. Of the 17 percent (four percent stated that it depends on their current workload) that said it was not enough time, most stated that up to 90 days are needed. We believe 60 days is enough time, although exceptions could be made during the busiest of the severe weather months.

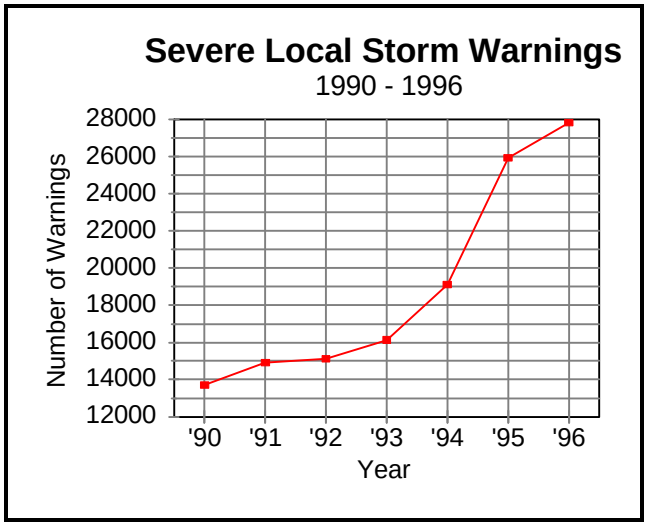
NWS needs to reduce its backlog of *Storm Data* and ensure that the publications are issued within 90 days. NWS guidelines (Chapter C-42) should be revised to provide for a "cutoff date" for WFO submission of data. After the cutoff date, OM should send the reports to the NCDC, with outstanding reports labeled "not received." Proper action, such as a memorandum to the applicable regional WCM, should be taken against offices that do not submit their data in a timely fashion. An addendum with the missing data from WFOs who missed the prior publication deadline could be added to the next month's publication.

NWS agreed that an integrated, real-time verification and data collection system and database is needed. However, NWS believes that it should be designed to operate within the AWIPS environment. Until AWIPS is implemented, NWS believes that PC-LOGIT and the storm data software may provide an interim solution. In addition, NWS reported that it has already taken steps to reduce the backlog of monthly *Storm Data*, and hopes to eliminate it by the end of 1997.

IV. NWS Should Reassess Wind and Hail Warning Thresholds

The number of severe weather warnings has more than doubled from 1990 to 1996 (see Figure 9).[16] Most of the increase in warnings involves the less severe storms. While advances in technology and ground sitings have lead to the increase in the number of warnings, some NWS personnel and the media feel that the thresholds are too low, and that the events associated with less intense storms pose little risk to the public.

Figure 9



NWS's multi billion-dollar modernization effort has equipped the WFOs with more powerful satellites, more sophisticated radars, and better computer models. The Doppler radars detect precipitation and measure wind speed and direction in greater detail. Meteorologists attribute the increased accuracy and speed in prediction to the modernization.[17] However, another result of the new technologies and increased geographical radar coverage has been the large increase in the number of warnings, which have more than doubled, from 13,696 in 1990 to 27,838 in 1996.

Impact of Increased Warnings

The NWS Director's Advisory Committee on Forecast Operations (DACFO), originally formed to uncover technical field problems and issues related to field operations, has assumed more modernization responsibilities. DACFO is composed of field forecasters and issues an annual report with recommendations. While the purpose of issuing warnings is to both save lives and protect property, DACFO has long believed that the thresholds for severe local storms are too low. The combination of low thresholds and better technology has resulted in the forecasters having more information and issuing more warnings than ever before.

[16] 1990-95 figures from NOAA Technical Memorandum NWS SPC-1; 1996 figures provided by OM.

[17] "Ahead of the Weather," William J. Cook, *US News & World Report*, April 29, 1996.

While the public may benefit some from more storm warnings, the increasing volume of severe thunderstorm and tornado warnings has had three major negative impacts. First, issuing too many warnings taxes the media's ability to disseminate timely and effective warnings through the Emergency Alert System. One WCM noted in an internal NWS E-mail that NWS "...will alienate the broadcast media at a time when we have the opportunity to communicate our warnings better than we ever have before. Too many warnings and we get turned off or ignored." One NWS employee told us that some stations are turning off EAS because too many warnings are being issued. While another employee disagreed that the EAS was being turned off, the two employees agreed that too many warnings is a problem in the northern and southern plains states, where most warnings are issued.

Second, the increase in warnings affects WFO resources needed for verification, requiring additional staff time to collect and submit event information. Although our survey confirmed that forecasters spend thousands of hours on verification, the total number of hours is equal to only about 20 full-time positions. While this is a relatively small number, the impact of verification on WFO resources can be significant for several reasons: time spent on verification varies from office to office, from a few to hundreds of hours; verification time is compressed into a four- to six-month severe weather season; and verification demands occur during and after storms, when offices are at their busiest.

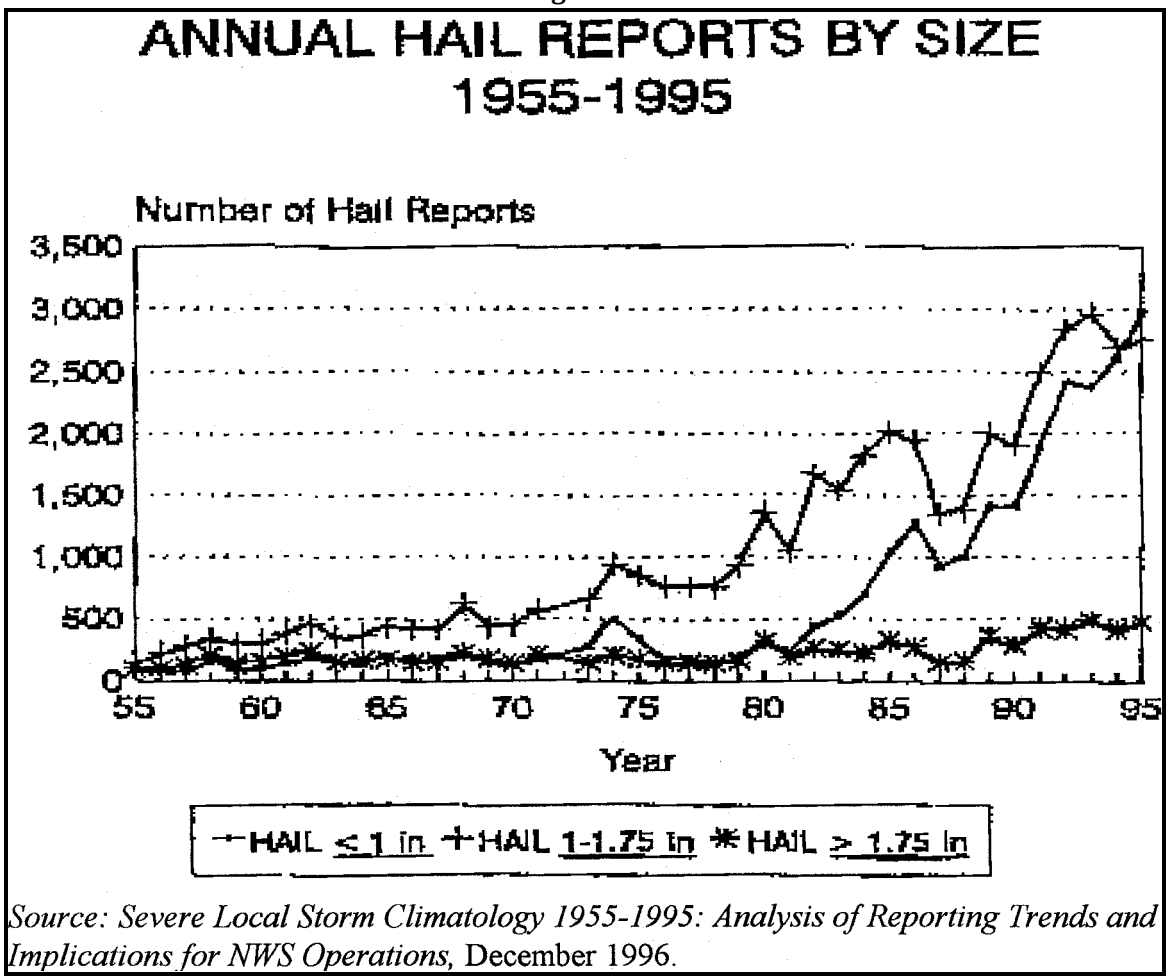
Third, although there is no solid evidence of this, NWS personnel contend that in some parts of the country, the public has become complacent to warnings because they are so often issued for non-severe storms. They also stated that these marginal hail events (.75-inch hailstones) are not a threat to most of the population. Figure 10, on the following page, shows how the number of marginal hail events began increasing rapidly in the 1980s to the point that this group is now reported more than any other event.[18] The data suggests that a change in the marginal hail climatology has occurred; however, the increase in 1980 coincided with NWS beginning its severe storm verification program.

Recommendations of DACFO

The Advisory Committee's concerns that the thresholds are too low for severe thunderstorms first appeared in its 1988 annual report. In 1994, 1995, and 1996, DACFO recommended changing the criteria for severe thunderstorms, reporting that: "Criteria for a severe thunderstorm are too low for some parts of the country. In Kansas, 3/4 inch hail and 50 knot

[18] *Severe Local Storm Climatology 1955-1995: Analysis of Reporting Trends and Implications for NWS Operations*, December 1996.

Figure 10



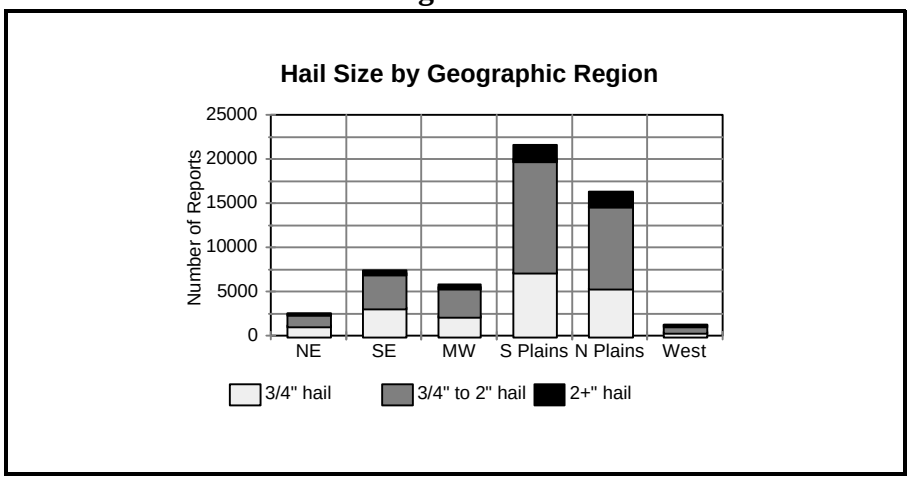
Source: Severe Local Storm Climatology 1955-1995: Analysis of Reporting Trends and Implications for NWS Operations, December 1996.

winds are so common that they are not perceived as significant events by the public. As a result, many warnings that technically verify may effectively be 'false alarms,' and credibility of the entire warning program suffers."[19]

Another DACFO recommendation is to allow regional discretion in defining severe thunderstorm criteria for local climatology. This recommendation has some credibility for the southern and northern plains states, where there are a large number of hail reports, over half of them larger than the 3/4-inch threshold (see Figure 11).

[19]*Twenty Fifth DACFO Final Report,* page 13, July 1997.

Figure 11



One reason for not implementing a regional program involves the problems with visitors being unfamiliar with regional thresholds. For example, a person from the Northeast, driving through the Southern plains states, might not be aware that the hail threshold in Texas is 1.5 inches. While this argument has some merit, NWS has experience with the application of different thresholds across the country. Snow thresholds in the winter storm warning program vary by state. While we are not advocating WFO or even state-by-state thresholds, we believe further exploration of a macro-level breakdown of warning thresholds, based on geographical weather patterns, is justified.

Another DACFO recommendation, and one discussed by NWS for more than a decade, is to have a multi-tiered Severe Weather Warning system. This would consist of having a separate category of enhanced public warnings for extremely severe events. The Air Force Global Warning Center, responsible for issuing weather warnings to pilots at Air Force bases around the world, uses such a system. The Air Force has broken its convective weather warning system into four categories: tornado, severe thunderstorm, moderate thunderstorm, and general thunderstorm. The Air Force threshold for the severe thunderstorms category is three-quarter inch hail. The benefit of a NWS multi-tiered system is that it gives the public more refined information for better decision-making. A proponent of the multi-tiered system gave the following example. A car dealership, when alerted that a thunderstorm is approaching, has to decide whether to move all cars inside, move some inside and others under the tarmac, or to leave them outside. Since hail size is directly related to the potential amount of car damage, the more specific the information, the better the decision-making.

NWS has not wanted to confront these issues until its modernization and associated restructuring near completion. However, potential media and EAS problems and our interest in this issue have compelled NWS to establish a working group in June 1997 to recommend

changes to severe thunderstorm warnings. The working group consists of NWS field and regional personnel, and scientific, academic, electronic media, insurance company, and engineering personnel. It is expected to issue a report on its findings in October 1997.

Conclusions

Although we support NWS's current working group, we recommend that the agency expand it to include more external experts. Making any change to the severe thunderstorm thresholds is likely to be controversial. We believe a panel consisting of university and other scientists, experts in the field of aviation, emergency managers, and other users should be convened to address the following options:

- The appropriateness of and problems associated with the current national warning thresholds.
- The pros and cons of a geographically based regional warning system with differing thresholds.
- The pros and cons of a multi-tiered severe weather warning system to facilitate alerting the public.

We are not recommending any of the above changes. Rather, we are recommending that NWS thoroughly review and, as appropriate, operationally test changes to the current warning criteria. The various options - changing the nationwide threshold, moving to a regional threshold system, or implementing a multi-tiered warning system - should be operationally tested in a number of states. NWS should document the impacts of the proposed changes on the various groups that would be most affected by different threshold levels. NWS's working group should also evaluate the effects of using larger wind events for verification purposes and developing alternative verification methods or measures.

NWS believes that its review team consisted of a broad cross section of the "hazards community" suggesting that this was sufficient. We believe more should be done since 17 of the 24 team members were NOAA employees. NOAA acknowledged that the composition of the four subteams was narrowly defined such that not every team reviewed all possible options. In fact, it is our understanding that only the science and operations team (seven NWS employees) reviewed and subsequently rejected the geographically based regional warning system option. At a minimum, this creates an appearance of bias. We therefore encourage NOAA to see what it can do to seek input from external sources before it rejects or proposes changes in the warning thresholds. In addition, we continue to believe that NWS should operationally test changes to the current warning criteria.

V. NWS Needs to Strengthen its Headquarters and Regional Office Oversight Roles and Responsibilities

A. OM has not provided the necessary leadership and oversight for national, regional, and local verification issues and efforts

The Office of Meteorology, which took over responsibility for the National Verification Plan from the National Severe Storms Forecast Center in January 1996, is responsible for overseeing and leading the program through guidance, training, and program evaluation. The Verification Program began in 1980, when NWS established a National Verification Task Team to prepare an NVP to meet NWS national, regional, and local verification needs. With eight major forecast programs,[20] NWS lacked a comprehensive national program to address verification weaknesses, such as poor data input procedures, inadequate quality control, antiquated data processing methods, and the lack of timely feedback at the local level.

In June 1982, the team completed its work by issuing the National Verification Plan. As a result, in 1983 the National Verification Committee was established to implement the plan, advise management on the uses of verification statistics, and monitor field responses to the verification programs. The committee met semiannually in 1983 and 1984 and then annually through 1989. The committee chairman tried to hold meetings in 1990 and 1991, but funding was not available. As a result, the committee never fully implemented the national plan, leaving a considerable gap in NWS's ability to provide national oversight for verification efforts and to measure the program's effectiveness.

Some NWS personnel told us that a national group, like the old National Verification Committee, should be reestablished, emphasizing that user services are more important than obtaining verification statistics. Specifically, the group could evaluate all OM programs, including customer service, warnings, communications, verification, and regional practices. Currently, NWS uses ad-hoc groups to evaluate services. OM's National Verification Manager has focused on compiling and disseminating national verification statistics at the expense of national, regional, and local verification policies and oversight. With only one full-time contract employee, the manager has had little time for key verification issues, or to provide overall leadership and oversight. However, he plans to revise the verification chapters in NWS's Operations Manual to include new verification techniques and prepare a new verification plan updating the requirements and goals of each verification program area.

[20] Agriculture, air pollution, aviation, fire weather, hydrology, marine, public, and severe weather.

We agree that NWS should reestablish a national committee, revise its Operations Manual, and prepare a revised verification plan. More importantly, NWS needs to strengthen the role of its National Verification Manager and OM authority and responsibility to oversee field verification efforts. Specifically, NWS needs to ensure that its National Verification Manager proactively reviews, with the national committee, national verification issues, such as warning criteria and verification techniques. The manager should ensure that verification "best practices" are shared with and used by regional and local offices. For example, the manager should evaluate how well verification data is obtained and reviewed for quality control.

NWS agrees that a more formal process should be established for general oversight on the development of new programs and specific verification issues. However, NWS believes that it is not cost effective to reestablish meetings of a National Verification Committee with the same charter as the original charter. NWS plans for its existing working groups to handle issues similar to those addressed by the NVC, and will assign one person from each region to represent that region's strategic vision on verification. These individuals will meet or hold conferences with the national verification program manager at regular intervals to discuss general verification issues and follow-up on specific action items. By assigning regional personnel to meet with the national manager, NWS will establish a de facto national committee that meets the intent of our recommendation.

B. The regional office role in verification needs to be expanded

The NWS regional offices' role in verification is generally very limited, except for major events. We determined that the regional Warning Coordination Meteorologists review verification statistics for office trends and certain LSRs for event information. The regional WCMs stated that overseeing WFO operations was their role, but that due to staffing limitations, they are unable to authenticate thousands of individual events reported by the WFOs. However, some NWS personnel felt that regional personnel were out of touch with WFO operations and real-time weather situations. After our discussions with the regional WCMs, they suggested that they could improve quality control of verification statistics, and they could increase local, state, and private sector alliances to provide additional verification information.

Regional Office Quality Control Opportunities

NWS's Operations Manual, Chapter C-72, states that the regional offices (ROs) should review verification output and assist their field offices in using it to improve field

performance.[21] The ROs also are supposed to help WFOs interpret policy and procedures for data collection. WFO personnel stated that because OM provides monthly statistics to their regional headquarters, regional personnel just want to "graph scores and not fix the system's flaws." More importantly, they felt that because the regional WCMs were usually trained and worked in the pre-MAR environment, they lack insight into current WFO operations, particularly in severe weather situations.

To enhance regional WCM knowledge of WFO operations, WFO personnel suggested that regional WCMs work in a WFO for one week every two to three months. In prior years, the regional WCMs visited WFOs to gain insight into WFO operations. However, limited funding has seriously reduced these visits. Rather than work two or three WFO shifts each year, which the regional WCMs believe is counterproductive, the regional WCMs believe that visiting their WFOs and observing new initiatives is more efficient. Without these visits, they agree that they could lose touch with WFO operations. We believe that the regional WCMs and WFOs can significantly improve their communication through direct office visits. Although limited funding for office visits can be a problem, NWS should provide funding for some regional WCM visits to WFOs in Fiscal Year 1998.

After discussions with regional WCMs, local WCMs, and forecasters, we believe that the ROs should conduct random audits, or "spot checks," of event and warning information to help improve data accuracy and reliability. The ROs should also establish partnerships to randomly send event information to an outside entity (e.g., emergency managers or local university meteorology departments) to evaluate the adequacy and validity of that data.

Although NWS states that a uniform process, for office reviews, can be established across the country, NWS does not mention whether or when such a process will be established. NWS needs to clarify its answer in its Action Plan to our office.

Regional Office Partnerships

The WCM Handbook requires WCMs and other WFO personnel to contact utility companies, 911 centers, state police, and insurance companies for assistance with event collection information.[22] The ROs could help their WFOs by establishing alternative sources

[21] "National Watch/Warning Verification Program," *Weather Service Operations Manual,"* Chapter C-72, April 21, 1987.

[22] *Warning Coordination Meteorologist: Job Aid*, Revised August 1996.

of event information, especially for damage values which are required for the storm data reports and which take a substantial amount of time to obtain, are generally of poor quality, and have little to do with the verification of severe weather. More reliable data could be obtained by ROs coordinating information from common sources, such as insurance companies.

NWS's Operations Manual Chapter C-49 outlines recommended contacts for fostering partnerships.[23] NWS regional personnel stated that C-49 is flexible enough to cover many partnership possibilities, without forcing any one of them. For example, coastal WFOs work with the U.S. Coast Guard and use Coast Guard damage reports, whereas many inland WFOs may be able to develop partnerships with the Army Corps of Engineers along major waterways to obtain damage estimates. There is a wide range of possibilities. Regional personnel stated that they can try to foster working relationships with telephone, power, or insurance companies, and most will work with them because it is mutually beneficial. However, they cannot demand such cooperation. In addition, some police agencies and nuclear power companies probably would want to ensure that sensitive personal and customer information would not be released with nonsensitive information. We believe that the ROs should develop procedures to gather information from utility companies, 911 emergency centers, and insurance companies so that hail and wind damage information is routinely shared with NWS.

NWS stated that its regional and national headquarters have provided lists of potential sources or data for WCM use. However, NWS does not state whether and when procedures to gather such information will be developed. NWS needs to clarify its answer in its Action Plan to our office.

[23] "Warning Coordination and Hazard Awareness Program," *Weather Service Operations Manual*, Chapter C-49, August 7, 1991.

VI. NHC Should Expand its Verification Efforts and Test its Emergency Backup Preparedness

Hurricanes, which strike the United States about two to three times every summer, are greater in size and duration than severe thunderstorms and tornadoes. Hurricanes can persist for many days and affect thousands of lives and billions of dollars of property. The National Hurricane Center issues advisories, watches, and warnings for hurricanes over the Atlantic Ocean, the Eastern Pacific Ocean, the Caribbean Sea, and the Gulf of Mexico. After hurricane season, which usually runs from May to November, NHC conducts an annual post-seasonal review and prepares an annual hurricane report for issuance to interested parties. Among the information that NHC compiles and distributes is the verification of each hurricane's track and intensity. NHC's comparison of hurricane track and intensity forecasts and the actual track analysis has been done for 25 years. However, we believe that some improvements can be made in the hurricane verification process.

A. Hurricane verification should include hurricane warning area ("margins of error") and wind radii

Currently, the NHC documents the "end points" of a hurricane warning area, the actual hurricane track, and wind intensity. NHC does not verify the accuracy of the forecast or compare it to the actual hurricane or the wind radii of each hurricane. NHC would like to verify more forecasts, but ground observations are limited. For example, forecasting intensity remains the forecaster's greatest problem because of limited observations from ships and other equipment, and because understanding the physical process of changing winds is very difficult. To address these limitations, two of NHC's key performance goals for 2000 are to (1) reduce the hurricane warning areas and (2) begin verifying wind radii analyses and other marine and aviation products.[24] When combined with better dissemination, substantial savings to the public will be gained as less damage will be sustained and unnecessary preparation and evacuation costs will be avoided.

[24] *Draft 1997 Strategic Plan for the Tropical Prediction Center,* National Hurricane Center, February 1997.

NHC Warning Areas

NOAA's Strategic Plan calls for a 10-percent reduction in hurricane warning-area size and a comparable 10-percent savings in damages avoided and unnecessary preparation costs.[25] In order to reduce hurricane warning areas, NHC needs to know the portion of each warning area that did not receive hurricane force winds. However, NHC does not systematically document and verify the hurricane warning areas.

Although NHC usually issues hurricane warnings 24 hours in advance for about 300 nautical miles, the average hurricane covers only about 100 nautical miles. Since NHC's average forecast error for 24 hours is about 100 nautical miles, NHC adds an additional 100 miles to the north and south of the hurricane track for a 300 nautical mile average warning area. NHC would like to reduce its warning areas but does not want to risk hurricanes striking unwarned people.

NHC personnel stated that there has never been a hurricane where NHC has not issued a warning. NHC argues that a false alarm area is inherent in issuing a warning because the hurricane is going to strike within the warned area, but not the entire area. However, such over-warning not only results in a loss of credibility, but more importantly, is very costly. NOAA estimates that the cost to the public of posting hurricane warnings on the eastern seaboard for each hurricane is about $200 million. This estimate includes the cost of boarding up homes, closing down businesses and manufacturing plants, and evacuating oil rigs. It does not include economic losses due to disruption of commercial activities, such as lost sales and canceled reservations. Conceptually, by reducing the warned area by 10 percent - the NOAA goal - the public could save at least $20 million per storm.

We believe that NHC should systematically verify and document the portion of each warning area that did and did not receive hurricane force winds. NHC's Director agrees, and has directed NHC's verification specialist to analyze the data from prior hurricanes to begin understanding the time involved and the data needed. More importantly, this will provide an initial baseline, or average size, of over-warned areas. Although NHC's verification specialist believed that performing this analysis could be time consuming, he thought that it could be done with "a small allocation of funds to support a student appointment to help with the creation of the required data bases."

[25] *The 1995-2000 NOAA Strategic Plan,* National Oceanic and Atmospheric Administration, July 15, 1993.

NHC Wind Radii

NHC stated that the average size of a hurricane warning area is consistent with the 100-mile average track error at 24 hours and the average hurricane size. The average size of a hurricane should comprise the wind radii, which is the horizontal distribution of winds extending out from a hurricane's center. However, wind radii determination has been difficult because surface wind observations from ships, land stations, and buoys in the vicinity of tropical cyclones are very limited. Therefore, the initial values of the wind radii are usually estimated from satellite imagery, NWS radars, and U.S. Air Force (USAF) and NOAA aircraft observations, although these platforms do not measure surface winds directly. In addition to the lack of surface observations, NHC currently has no objective guidance models to predict the evolution of the tropical cyclone wind field. As a result, NHC usually determines the 12-, 24- and 36-hour radii forecasts by extrapolation and simple rules of thumb. Because of these uncertainties, the wind radii initial estimates and forecasts have never been verified.

Because NHC has not quantitatively assessed and statistically verified wind radii, it cannot begin to evaluate how hurricanes intensify or weaken or how the warning margins of error can be reduced. NHC is required to estimate the radii of 34, 50, and 64 knot winds in four quadrants (northeast, southeast, southwest, and northwest) at the starting time of each forecast. These radii are important to people with boating and marine interests and ship operators who need to avoid areas with hurricane and gale force winds. NHC is also tasked with forecasting the wind radii at 12, 24, and 36 hour intervals. These radii are important to emergency managers who need to know when gale force winds will begin, and they also have an impact on the extent of the size of the watch and warning areas.

To begin learning how to analyze wind radii, the Hurricane Research Division (HRD) of NOAA's Atlantic Oceanographic and Meteorological Laboratory has recently collected a database of surface observations from the past 20-25 Atlantic land-falling storms. NHC is working with HRD to use this data to evaluate wind radii estimates, and results should be available for presentation at the 1998 Interdepartmental Hurricane Conference. NHC should continue to validate its wind radii for cases where surface data is available. HRD is planning a Hurricanes at Landfall Project, which should provide extensive verification data for a few cases. New instrumentation to measure surface winds is being added to the USAF Reserve aircraft. NHC should develop methods for using this new data source for wind radii forecasting and verification. In addition, the Geophysical Fluid Dynamics Laboratory (GFDL) hurricane model provides a forecast of the hurricane wind field, and wind radii could be determined from these predictions. NHC should begin to systematically verify the GFDL model wind radii forecasts.

B. NHC's backup plan should be tested annually

It has been two years since NHC has tested its backup plan, which outlines the actions that will be taken if NHC processing and operations are interrupted by fire, power outages, water damage, or other events. NHC provides not only tropical cyclone forecasting, but aviation, marine, and satellite forecasting as well. In an emergency, the plan calls for NHC to shift its hurricane forecast responsibilities to the Hydrometeorological Prediction Center (HPC) in Camp Springs, Maryland; marine forecast responsibilities to the Marine Prediction Center in Camp Springs, Maryland; aviation forecast responsibilities to the Aviation Weather Center in Kansas City, Missouri; and satellite product responsibilities to NESDIS's Synoptic Analysis Branch in Suitland, Maryland. Because of funding cuts, HPC has lost about 30 percent of its staff. As a result, NHC management questions whether HPC has the resources to act as a backup center. We also found that the backup plan for NHC's marine, aviation, and satellite operations had not been recently tested, although the three designated backup centers have NHC's backup product formats and software available for immediate use.

NHC is concerned about operations interruptions from a tropical storm or hurricane, fire and water damage, a power outage, and even a computer virus. For example, in April 1997, the Storm Prediction Center experienced an electrical outage when construction equipment severed power lines. SPC stated that this outage did not affect its services because SPC exercised its back-up arrangement with the USAF Global Weather Center. However, although the outage was short, a longer outage could severely hamper a center's operations.

Federal Information Processing Standards Publication 87 states that unless backup plans are continually reviewed and tested, they may fail when needed. Personnel should test backup plans at the designated backup site by making copies of all needed data and other information. The test should demonstrate that the backup site remains unharmed in a simulated catastrophe or disruption of service. A backup plan for any data processing activity, regardless of its operation, should address three areas:

- *Emergency response procedures*, to cover the appropriate response to a disastrous event.

- *Backup operations procedures*, to ensure that essential data processing operational tasks can be conducted after disruption to the primary data facility.

- *Recovery action procedures*, to facilitate the rapid restoration of a data processing facility following physical destruction, major damage, or loss of data.

In addition to Publication 87, both the NWS Operations Manual Chapter 41[26] and the National Hurricane Operations Plan[27] dictate the transfer of hurricane warning responsibilities. Chapter 41 states that maintaining a successful backup program requires, as a minimum, annual testing of the plan. During our visit, NHC provided a draft copy of its backup plan, which stated that backup procedures will be tested before each hurricane season.

NHC management stated that drafting and testing a backup plan had been a low priority. In June 1997, an HPC staff member received training on NHC's Automated Tropical Cyclone Forecasting System (ATCFS), which runs models, displays the "working best track," and graphically outlines NHC products. He will prepare a manual so that HPC staff can run the ATCFS at HPC. HPC has a dedicated computer and is getting the latest version of the system software, and will have a dedicated terminal or workstation for displaying satellite imagery/gridded data. HPC personnel stated that during the next off-season, messages, models, and data could be tested. They did not want to disrupt normal hurricane operations now. We agree. NHC should, however, test its backup plan at HPC during the next off-season to determine whether HPC is a viable backup center. If not, another backup center should be selected immediately or appropriate action should be taken to make HPC a viable backup center.

NHC agreed to systematically verify the portion of each warning area that did and did not receive hurricane force winds. NHC also stated that it is testing various methods for verifying wind radii forecasts. NHC also will test its backup plan during the 1998 preseason.

[26]"Tropical Cyclone Program," Chapter, C-41, *Weather Service Operations Manual*, May 1, 1996.

[27]*National Hurricane Operations Plan*, Office of the Federal Coordinator for Meteorological Services and Supporting Research, NOAA, May 1996.

RECOMMENDATIONS

We recommend that the Acting Assistant Administrator for Weather Services direct that:

1. NWS improve verification quality control by developing NWS guidelines at the national level and implementing them throughout the network. Specifically, NWS should reestablish its National Verification Committee and, in conjunction with OM, WFOs, and the regional offices, implement a standardized review process to ensure data integrity that will include:

   - recording the source of event information.
   - distinguishing between measured and estimated wind gusts.
   - sending verification data to external sources for review.
   - spot checking monthly data and "signing off" as completed before sending it to OM.
   - comparing the timing of storm reports with radar imagery.
   - randomly reviewing reports from nearby counties for time continuity.
   - randomly reviewing changes made to monthly storm data reports.
   - using only measured hail diameters - not estimates - to numerically verify warnings.
   - adding the name of the forecaster who collects event information to the storm data event database.
   - developing software to prevent the same forecaster from issuing and verifying a warning.
   - instituting third party reviews or verification checks when the same person has issued and verified a warning.

2. NWS determine the status of LSR software that is being developed at the Baltimore/Washington D.C. WFO and determine its potential use by all WFOs.

3. NWS describe the means used to "verify and validate" its verification measures and eliminate the phrase "Quality control procedures are followed to ensure the highest possible reliability of the gathered data" from its 1998 Annual Plan.

4. NWS provide WCMs with alternative training mechanisms that promote data integrity, including CD-ROMs, videos, and discussions at annual WCM meetings.

5. WFOs continue to emphasize the importance of data accuracy in spotter training. This should include helping WFOs purchase or obtain inexpensive calipers or rulers to accurately measure hail size.

6. OM implement specific automated tests of national verification information, including year, month, day, time, WFO, event, latitude and longitude, deaths, injuries, damage, and warning issue time.

7. OM document changes made to monthly storm data reports to ensure that changes are authorized.

8. OM create a new field in the event database and direct that WFOs use it to identify the source of information for their storm data.

9. The regional offices require that PC-LOGIT is to be used by all WFOs until an integrated system is developed.

10. NWS implement an integrated, real-time verification and data collection system and database to accept LSRs, provide the foundation for storm data reports, and provide timely initial or first-cut feedback to forecasters. To eliminate the redundant input of data, NWS should develop software for inputting LSRs that is compatible with its storm data software.

11. NWS reduce its backlog of the monthly *Storm Data* publication and ensure that the publication is issued within 90 days after each month. NWS also needs to revise its guidelines (Chapter C-42), documenting a "cutoff date" for WFO submission of data. After the cutoff date, OM should send the reports to NCDC, with outstanding reports labeled "not received," with sanctions for unreasonably late submissions.

12. NWS expand its review of the severe thunderstorm warning criteria to include more external input. The various options - such as changing the nationwide threshold, moving to a regional threshold system, or implementing a multi-tier warning system - should be operationally tested in a number of states. NWS should document the impacts of the proposed changes on the various groups that would be most affected by different threshold levels. NWS should consider and evaluate the effects of using larger wind events for verification purposes and developing alternative verification methods or measures.

13. NWS reestablish its National Verification Committee, revise Chapters 72 and 73 of its Operations Manual, and prepare a revised verification plan. More importantly, NWS needs to strengthen the role of its National Verification Director and OM authority and responsibility to oversee field verification efforts. Specifically, OM's director should review national verification issues, such as warning criteria and verification techniques, and ensure that verification "best practices," such as WFOs using recorders attached to telephone lines or pink-colored log-in sheets to record the original source and time of event information, are disseminated and used by regional and local offices.

14. The regional offices conduct random reviews, or "spot checks," of event data and warning information. The offices also should establish partnerships to randomly send event information to outside entities to evaluate its adequacy and validity. In addition, the regional offices and NWS headquarters should develop procedures to gather information from utility companies, 911 emergency centers, and insurance companies so that hail and wind damage information is routinely shared among offices. Although limited funding for office visits can be a problem, NWS should provide funding for some regional WCM visits to WFOs in Fiscal Year 1998.

15. NHC systematically verify (a) the portion of each warning area that did and did not receive hurricane force winds and (b) its Geophysical Fluid Dynamics Laboratory model (wind radii) forecasts.

16. NHC test its backup plan at HPC during the next off season to determine whether HPC is a viable backup center. If not, another backup center should be selected immediately or activities taken to make HPC viable.

Appendix A: OIG Survey Questionnaire for Warning Coordination Meteorologists

Questions & Summary Responses

(1) What region are you in?

19	Eastern Region
27	Southern Region
33	Central Region
22	Western Region
3	Alaska Region
1	Pacific Region

(2) If you are someone other than the WCM, what is your position?

No summary compiled.

(3) How many years have you been doing severe weather verification at your current office?

4.2 years average

(4) How many county-based warnings did your office issue in 1996 for thunderstorms, tornadoes, and flash floods? (numbers from OM)

Region	WFO Average	Region Total
Eastern	224	4,262
Central	278	8,917
Southern	384	9,985
Western	74	1,564

(5) How many warnings did your office issue in 1996 for winter storms and high winds?

No summary compiled.

(6) What is the population estimate for your county-warning area?

Region	WFO Average	Region Total
Eastern	4,483,151	85,179,880
Central	1,260,354	41,591,698
Southern	2,095,227	48,190,233
Western	1,882,495	33,884,916

(7) How many counties are in your county-warning area?

26.8	Average
16.2	Standard Deviation

(8) How many counties used to be in your county-warning area?

17.7	Average
15.3	Standard Deviation
10	Not applicable (new office)
1	Non-response

(9) What percentage of your county-warning area is rural?

79.1	Average
17.7	Standard Deviation
6	Non-response

(10a) Does your office use Local Storm Reports for verification purposes?

Number of Responses	Answer
101	Yes
3	No
1	Non-Response

(10b) Does your office make calls to locate damage sites and verify time of occurrence?

Number of Responses	Answer
105	Yes
0	No

(10c) Does your office use a log-in sheet for incoming calls and calls initiated by your office?

Number of Responses	Answer
99	Yes
6	No

(10d) Does your office use newspaper clippings for verification purposes?

Number of Responses	Answer
105	Yes
0	No

(10e) Does your office use radar signatures for verification purposes?

Number of Responses	Answer
24	Yes
81	No

(11) Does your office compute preliminary verification statistics for severe local storms, winter storms, river floods, etc.?

Number of Responses	Answer
66	Yes
39	No

(12) How many new spotters were trained in 1996? How many retrained spotters? How many spotters are in your spotter network?

	New	Retrained	Total
Average	312	353	1192
Std Dev	456	488	1366
Unknown	11	20	10

(13) What percentage of your warnings are verified by people calling in, WFO calling out, other?

	Call in	Call out	Other
Average	37	48	15
Std Dev	19	22	18
Min	5	0	0
Max	90	95	90
Mode	20	60	10
Non-response	6	6	6

(14) What are the major constraints to verifying more warnings?

Number of Times Constraint was listed	Constraint
85	Low Population, No one to report events
30	Lack of Staff/Time
5	Events Occurring at Night
4	"Cultural" Differences of Spotters
4	Emergency Manager Issues
3	Quality of Reports from Spotters, EMs, etc.
2	Odd Shaped County Boundaries
2	Lack of Automated Weather Stations/Equipment
2	False Alarm Rate
10	Other
147	Total Number of Constraints

(14a) What percent of warnings are not verified by your office (due to the above constraints estimate)?

Number of Responses	Answer
7	0% to 10%
23	11% to 20%
31	21% to 30%
13	31% to 40%
4	41% to 50%
4	51% to 60%
4	61% to 70%
8	71% to 80%
4	81% to 100%
7	Don't know/blank

(15) What type of severe event(s) are most difficult to verify? (multiple answers listed)

Number of Responses	Answer
41	Thunderstorms
21	Hail
17	Tornados
16	Flash Floods
12	Winds
10	All marginal storms
10	All convective storms
6	Down burst winds
4	Damage reports
1	Winter storms
105	Total people responding

(16) Does the same meteorologist issue and verify the same warning?

Number of Responses	Answer
14	Yes
37	No
54	Sometimes

(16a) If so, do you think this is a problem and why?

Number of Responses	Answer	
5	Yes	(no narrative summary was compiled)
68	No	
32	Non-response	

(16b) To what extent does the same meteorologist issue and verify the same warning in your office?

Percent	Number of Responses	Answer
31.4%	33	A. Little or no extent 0-10%
36.2%	38	B. Some extent 11-20%
23.8%	25	C. Moderate extent 21-40%
5.7%	6	D. Great extent 41-60%
2.9%	3	E. Very great extent 61-100%

(17) What new initiatives has your office made in collecting events over the past 3 to 5 years? (144 new initiatives)

Number of Responses	Answer
52	Spotter recruitment efforts
17	Expand or start HAM radio spotter network
13	Aggressive calling to obtain ground truth information
12	Increased coordination with Emergency Managers
11	Purchase and/or increase use of news-clipping services
11	Purchase and/or increase use of rural phone directories and lists
10	More and better spotter training
6	Internet solicitation of reports
3	Assignment of dedicated staff to gather ground truth information
3	More on-site surveys
3	Use of radar to establish location
3	Established a local verification database
7	Not applicable, new office
2	Non-response

(18) Is 60 days enough time to submit event reports to headquarters?

Percent	Number of Responses	Answer
79%	83	Yes
17%	18	No
4%	4	Depends

(18a) If no, how many days are needed?

Number of Responses	Answer
12	60 to 90 (including those that said 90)
7	90 to 120 (including those that said 120)
1	90 to 180

(19) How do you identify and verify events where warnings were not issued?

Number of Responses	Answer
70	Spotter reports
59	Newspaper clippings
28	Emergency manager reports
26	Radio and TV
11	Same as warning
10	Call out to field for reports
9	Reports from sheriff and law personnel
8	Site surveys
3	Considered a missed event
2	Radar images
2	From damage reports
2	Other
230	Total reports

(20) Who in your office reviews the accuracy of data submitted for verification?

Number of Responses	Answer
50	WCM
21	WCM & Focal Point
16	WCM & Others
6	Others
6	Focal Point
3	WCM & MIC
3	No one

(21) What specific steps as a WCM, do you perform to ensure that warning and event information is accurate? (Top 4 responses)

Number of Responses	Answer
52	Conduct a review of the following: storm data reports, warnings, log-in sheets
28	Make follow up phone calls to Emergency Managers, police and fire departments, and spotters
13	Review radar data
12	Conduct on-site evaluations and surveys

(22) How do you review additions, deletions and revisions to monthly or yearly summary statistics?

No summary was compiled.

(23) Does the false alarm rate or concerns about the FAR impact warning decision making?

Number of Responses	Answer
39	Yes
65	No
1	Non-response

(24) What suggestions do you have to improve event collection for severe local storms? Winter storms?

Number of Responses	Answer
71	Expand the spotter network
21	More staff, more time
17	Expand the HAM radio network
18	Continue fine tuning radar procedures
20	Increase the number of real time reports
16	More training of spotters
17	Automate reporting
16	Aggressive calling out
8	Misc suggestions
4	More automated observation equipment
2	1-800 phone line
2	Raise the storm warning criteria
8	None
9	Non-response

(25) What strengths or weaknesses of severe local storm verification training did you note in the WCM course?

Number of Responses	Answer
16	Respondents listed a strength of the course, of those:
10	Tim Marshall's talk of damage assessment
35	Respondents listed a weaknesses of the course, of those:
8	Too much emphasis on East/Mid Plains states and not enough on other regions
5	Not enough time spent on verification
5	Not enough time spent on filling out storm reports
3	Not enough information on spotters/event information gathering
54	Other Responses:
20	Respondent's could not remember
14	Replied "none" (no verification training or no comment)
20	Blank or N/A

(25a) Is follow-up training needed? If so, what should it emphasize?

Number of Responses	Answer
60	Total number of respondents that said Yes
28	Respondents that said No
17	Did not respond or said question was not applicable

TYPES OF TRAINING DESIRED:

10	Event information gathering techniques
5	Software training
5	Type/severity of event
3	Regional specific training
3	Survey training

METHODS OF TRAINING PREFERRED:

7	On-site training materials (CD-ROM, video, newsletter)
4	At the WCM conference (round table discussion, panel)

(26) How many hours per month are spent by the WCM verifying warnings and preparing storm data (do not include time spent preparing and sending local storm reports to the Storm Prediction Center)? Please give two numbers representing the busy and slow months.

Answer combined with number 27.

(27) How many hours per month are spent by the remaining WFO staff on verification (do not include staff time spent preparing and sending local storm reports to the Storm Prediction Center)? Please give the total number for office, not including your own, for busy and slow months.

Answer combined with number 26.

WFO TOTAL (Questions 26 + 27)

BUSY MONTH

55.5 Average hours per month

SLOW MONTH

10.2 Average hours per month

(28) How many hours should you and the person(s) listed in #27 spend on verification?

BUSY MONTH

26.3 Average hours per month
34.6 Standard Deviation

14 As much as needed
3 Too much
2 Don't know
8 Non-response

SLOW MONTH

4.9 Average hours per month
6.9 Sample Standard Deviation
0 Minimum
40 Maximum

OTHER VERBAL RESPONSES

13 As much as needed
1 Less
2 Don't know
9 Non-response

(29) Does your office benefit from the time spent on the verification process?

Percent	Number of Responses	Answer
77%	81	Yes
16%	17	Some
7%	7	No

(30) What other suggestions do you have for training and resources pertaining to the verification program?

No summary compiled.

(31) How does the new Office of Meteorology verification process compare to the old NSSFC process?

Percent	Number of Responses	Answer
53%	55	Better
12%	13	Worse
16%	17	Same
3%	3	Both
16%	17	N/A = unknown, too soon to know, not applicable

(32) For November 1996, how many warnings were missing from the weekly warning reports?

0.03	Average
0	Minimum
2	Maximum

(33) Could sampling "ground truth" effectively replace the current 100% verification methodology? Why or why not?

Number of Responses	Answer
9	Yes
51	No
5	Maybe, in some cases
7	Don't Know
23	Don't Understand the question
10	Non-response

(34) Could "radar signatures" replace the current 100% verification methodology? Why or why not?

Number of Responses	Answer
6	Yes
86	No
6	Maybe, in some cases
0	Don't Know
3	Don't Understand the question
4	Non-response

(35) How do you use the information (statistics) derived from the severe local storms verification program? (multiple uses reported)

Number of Responses	Answer
62	To critique performance and identify areas of improvement/training needs
9	To monitor year-to-year trends
6	To provide information to media, emergency managers
11	Other
12	Do not use them
9	Blank or ambiguous response

(36) How should verification information or scores be used in evaluating forecaster or office performance?

Forecaster Performance

84%	No
12%	Maybe
4%	Yes

Office Performance

82%	No
10%	Maybe
8%	Yes

(37) Should national verification be a higher or lower priority for the NWS?

Number of Responses	Answer
17	Higher
32	Lower
53	Stay the Same
3	Non-response

(38) Please list at least 3 major strengths of the severe local storm verification process.

A total of 240 strengths were listed. The following are the top five strengths:

Number of Responses	Answer
42	Provides an evaluation of office effectiveness
41	Provides an organized summary of severe events/climatological database
26	Improves working relationships with other agencies, county officials, and spotters
24	Enhances local studies and algorithm refinement and assess office training needs
21	Provides feedback to forecasters

(39) Please list at least 3 major weaknesses of the severe local storm verification process.

A total of 227 weaknesses were listed. The following are the top four weaknesses:

Number of Responses	Answer
34	Verification is a very time consuming process
33	Sparsely populated areas negatively affect verification scores
28	Verification statistics can be improperly used to evaluate and compare offices
14	100% verification is impossible

(40) What would you change in the national verification process?

Too many unique responses to summarize**.**

(41) What impact, if any, will the watch decentralization program have on verification efforts?

Number of Responses	Answer
40	No impact
31	Workload & time constraint concerns
13	Unknown
6	Will lead to improved products
5	Other possible negative impacts
10	Did not respond
105	Total responses

(42) Please provide any other comments or suggestions that you have on improving the verification processes?

No summary compiled.

WFO'S THAT REPORT ON HURRICANES:

(43) As it relates to verification, what suggestions do you have to improve event collection for hurricanes?

Eleven full responses, thus no summary was compiled.

(44) What suggestions do you have for improving the hurricane verification process?

Five full responses, thus no summary was compiled.

Appendix B: NWS Response to Report

UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
CHIEF FINANCIAL OFFICER/CHIEF ADMINISTRATIVE OFFICER

DEC 1 1997

MEMORANDUM FOR: Frank DeGeorge
Inspector General

FROM: Andrew Moxam
Acting Chief Financial Officer/
Chief Administrative Officer

SUBJECT: OIG Draft Inspection Report: NWS's Verification System For Severe and Hazardous Weather Forecasting Needs Modernization, Report No.IPE

Thank you for the opportunity to review and comment on the draft audit report concerning the National Weather Service's (NWS) severe weather verification process. We found the report to be thorough and comprehensive. It reflects a good understanding of this very complex process. We are pleased that you have concluded that the use of modernization technology, better recruitment and training of weather volunteers, and improved coordination with emergency management officials have resulted in more accurate forecasts and improved verification of severe weather events.

We agree with the report's findings and recommendations. As acknowledged in the report, several actions are already underway to improve the severe weather verification process. Your recommendations will help to enhance our efforts in that regard.

Attachment

Printed on Recycled Paper

GENERAL COMMENTS

We have some minor revisions to recommend. These changes do not modify the recommendations, but correct some minor factual details and clarify the work that needs to be performed.

- Page 5, lines 2-3 of "Recent Changes in NWS's Severe Local Storm Verification" Section: The verification that started in 1979 was for all National Weather Service (NWS) offices that had warning responsibility. This included over 200 offices, Weather Service Offices as well as Weather Service Forecast Offices.

- Page 8, line 1: There are four non-Washington area National Centers for Environmental Prediction (NCEP) Centers, or seven NCEP service centers.

- Page 23, lines 2, 10, 14, 16, and 18: The PC-LOGIT program was developed by the Storm Prediction Center (SPC). The National Severe Storms Forecast Center was disbanded on October 1, 1995, and replaced by the SPC and the Aviation Weather Center.

- Page 24, Figure 8: Other Federal agencies (Department of Energy, National Research Council, etc.) should be included as users of verification data.

- Page 26, paragraph 2, line 1: The Space Environment Laboratory is not the Space Environment Center.

- Page 39, lines 7-10: The statement that begins with "As a result, NHC's ..." is inaccurate and should be deleted.

- Page 39, last line before the NHC Wind Radii section: Replace the phrase "existing resources" with "a small allocation of funds to support a student appointment to help with the creation of the required data bases."

- Pages 39-40: Replace the section entitled NHC Wind Radii with the following:

2

The average size of the hurricane warning area is consistent with the 100-mile average track error at 24 hours and the average hurricane size. The size of a hurricane is quantified by wind radii, which measure how far winds of various thresholds extend outward from the storm center. National Hurricane Center (NHC) is required to estimate the radii of 34, 50 and 64 kt winds in four quadrants (northeast, southeast, southwest, northwest) at the starting time of each forecast. NHC is also tasked with forecasting the wind radii at 12, 24 and 36 h. These radii are important to emergency managers, who need to know when gale force winds will begin, and they also have an impact on the extent of the watch and warning areas. The wind radii are also important to ship operators who need to avoid areas with hurricane and gale force winds.

Surface wind observations from ships, land stations, and buoys in the vicinity of tropical cyclones are very limited. Therefore, the initial values of the wind radii are usually estimated from satellite imagery, NWS radars, and U.S. Air Force (USAF) and NOAA aircraft observations, even though these platforms do not measure surface winds directly. In addition, there are currently no objective guidance models that can reliably predict the evolution of the tropical cyclone wind field, so the 12-, 24- and 36-hour radii forecasts are usually determined by extrapolation and simple rules of thumb. Because of these uncertainties, the wind radii initial estimates and forecasts have never been verified.

NHC should begin a program to evaluate its wind radii. The Hurricane Research Division (HRD) of NOAA's Atlantic Oceanographic and Meteorological Laboratory has recently collected a data base of surface observations from the past 20-25 Atlantic land-falling storms. NHC is working with HRD to use this data set to evaluate its wind radii estimates, and results should be available for presentation at the 1998 Interdepartmental Hurricane Conference. NHC should continue to validate its wind radii for cases where surface data is available. HRD is planning a Hurricanes at Landfall Project, which should provide extensive verification data for a few cases. New instrumentation to measure surface winds is being added to the USAF Reserve aircraft. NHC should develop methods for using this new data source for wind radii forecasting and verification. In addition, the Geophysical Fluid Dynamics Laboratory (GFDL) hurricane model provides a forecast of the hurricane wind field, and wind

radii could be determined from these predictions. The GFDL model does not provide skillful predictions of the maximum winds of hurricanes, but its ability to forecast wind radii is unknown. NHC should begin to systematically verify the GFDL model wind radii forecasts.

Page 41, paragraph 2: It should be noted that this SPC outage did not hamper NWS services because the SPC exercised its back-up arrangement with the USAF Global Weather Center.

RECOMMENDATIONS

RECOMMENDATION 1 - Improve verification quality control by developing NWS guidelines at the national level and implementing them throughout the network.

Response - The NWS acknowledges the need for a more formal process for general oversight on the development of new programs and specific verification issues. See our response to Recommendation 13 regarding the reestablishment of the National Verification Committee and strengthening the Office of Meteorology's (OM) oversight role.

We agree with most of the specific recommendations to ensure data integrity but oppose several others. OM will take steps to implement recording the source-of-event information, distinguishing between measured and estimated wind gusts, randomly reviewing changes made to monthly storm data reports, and spot checking monthly data at the Weather Forecast Office (WFO) and signing it off as completed before sending it to OM.

We believe that comparing the timing of storm reports with radar imagery in most cases may be a good idea; but in many cases, mountainous terrain will contribute to beam blockage. Radar reports alone should not be used as a verifying report. Randomly reviewing severe event reports from nearby counties for time continuity may or may not be valid. Many severe events may occur over isolated areas, especially in warm season convection. Even in progressive line situations, time discontinuities may have physical reasons, such as discrete propagation.

It is a reasonable goal to use only measured hail diameters, but one should not necessarily rely on calipers or other measuring devices. Other common, and known objects, such as coins, golf balls, and baseballs are very good for comparison. See the response to recommendation 5 for additional information.

The NWS opposes the addition of the name of the forecaster who collects event information to the storm data event data base and the development of software to prevent the same forecaster from issuing and verifying a warning. In a severe weather situation at a WFO where staffing is strained to the limit and time is at a premium, it is difficult enough to follow up on severe weather events by anyone in the office, much less having to pick another forecaster. This recommendation places an unreasonable hardship on the WFO. In addition, the warning process involves an entire office and assigning an individual name to the warning is unrealistic.

RECOMMENDATION 2 - NWS determine the status of local storm reports (LSR) software that is being developed at the Baltimore/Washington D.C. WFO and determine its potential use by all WFOs.

Response - NWS will comply.

RECOMMENDATION 3 - NWS describe the means used to "verify and validate" its verification measures and eliminate the phrase "Quality control procedures are followed to ensure the highest possible reliability of the gathered data" from its 1998 Annual Plan.

Response - The NWS 1998 Annual Plan was published prior to release of the OIG draft report. However, the NWS will eliminate the phrase in question from all future plans and describe the means used to "verify and validate" its verification measures.

RECOMMENDATION 4 - NWS provide Warning Coordination Meteorologists (WCMs) with alternative training mechanisms that promote data integrity, including CD-ROMs, videos, and discussions at annual WCM meetings.

Response - The NWS attempts to develop new training and training tools for WCMs to utilize on an annual basis. Such training aids have been in the offing for the past several years, but the expense and budget pressures have precluded the

NWS from moving forward this past fiscal year. Budget pressures have also limited OM and the regions from conducting regional and National WCM conferences. Historically, the biannual WCM conference allocated a substantial block of time to data integrity and other verification issues. Regional WCM meetings that have been held during the in-between years have also dealt with verification issues. OM has been represented at these meetings.

RECOMMENDATION 5 - WFOs continue to emphasize the importance of data accuracy in spotter training. This should include helping WFOs purchase or obtain inexpensive calipers or rulers to accurately measure hail size.

Response - We will continue to work with the NWS Regions to reinforce the need for WFOs to emphasize the importance of data accuracy in spotter training. This includes guidelines that address the degree of accuracy desired, what measuring tools can best meet the defined level of accuracy desired and, personal safety during large hail events. It must be remembered that hail is melting constantly and that any significant handling of the hail in an attempt to get an acceptably accurate measurement may result in the further loss of hail size information.

RECOMMENDATION 6 - OM implement specific automated tests of national verification information, including year, month, day, time, WFO, event, latitude, deaths, injuries, damage, and warning issue time.

Response - NWS will comply with the recommendation. These tests will be applied to current, as well as historical, data.

RECOMMENDATION 7 - OM document changes made to monthly storm data reports to ensure that changes are authorized.

Response - NWS will create a separate data file that contains all modified reports and the corresponding original report. Examination of this file will provide information on the type and frequency of changes.

RECOMMENDATION 8 - OM create a new field in the event data base and direct that WFOs use it to identify the source of information for their storm data.

Response - NWS will comply.

RECOMMENDATION 9 - The regional offices require that PC-LOGIT be used by all WFOs until an integrated system is developed.

Response - NWS will work with SPC and the NWS Regions to ensure the uniform use of PC-LOGIT.

RECOMMENDATION 10 - NWS implement an integrated, real-time verification and data collection system and data base to accept LSRs, provide the foundation for storm data reports, and provide timely initial or first-cut feedback to forecasters. To eliminate the redundant input of data, NWS should develop software for inputting LSRs that is compatible with its storm data software.

Response - NWS is in agreement on the need for an integrated, real-time verification and data collection system and database to accept LSRs, to provide the foundation for storm data reports, and to provide timely initial or first-cut feedback to forecasters. For such a system to be effective, however, it should be designed to operate within the Advanced Weather Interactive Processing System (AWIPS) environment. A system based on the existing programs PC-LOGIT and the storm data software may provide an interim solution. This possibility will be investigated as indicated in the response to recommendation 2.

RECOMMENDATION 11 - NWS reduce its backlog of the monthly Storm Data publication and ensure that the publication is issued within 90 days after each month; revise its guidelines (Chapter C-42) documenting a "cutoff date" for WFO submission of data. After the cutoff date, OM should send the reports to the National Climatic Data Center (NCDC), with outstanding reports labeled "not received," and with sanctions for unreasonably late submissions.

Response - NWS has taken steps to reduce the backlog of monthly Storm Data. A new version of the storm data entry program generates reports in a format that can be printed with little or no modification. This has significantly reduced the time required to prepare the monthly Storm Data, and NWS expects to eliminate the backlog by the end of 1997. Once the backlog is eliminated, NWS will adhere to strict cutoff dates to ensure the publication remains timely. The latest procedures and cutoff dates will be documented in the next revision to WSOM Chapter F-42. OM will work with the NWS Regions to identify unreasonably late submissions of storm data reports and take appropriate management action to ensure compliance.

7

RECOMMENDATION 12 - NWS expand its Working Group on Warning Thresholds to include NWS local and regional personnel, university and other scientists, aviation experts, emergency managers, and other users. The group should thoroughly review and, as appropriate, operationally test changes to the current warning criteria and document the impacts of proposed changes on the various groups that would be most affected by different threshold levels. NWS's working group should also evaluate the effects of using larger wind events for verification purposes and developing alternative verification methods or measures.

Response - The Severe Thunderstorm Warning Criteria Review Team consists of people representing a broad cross section of the hazards community, including:

- NWS Warning Coordination Meteorologists and Science Operations Officers
- NWS Training Specialists from the NWS Training Center and Operational Support Facility/Operations Training Branch
- Center for Operational Meteorology, Education, and Training (COMET)
- The Storm Prediction Center
- Office of Atmospheric Research
- U.S. Air Force
- American Meteorological Society's Board of Broadcast Meteorology
- Social Scientists
- National Coordinating Council for Emergency Management
- American Association of Wind Engineers
- Institute for Home and Business Safety
- Small Aircraft Manufacturer's Association

This large group was divided into four subteams:

- Science Operations
- Science Research
- Social Science/Customer Service
- External Users

Each team was asked a series of questions to help identify current and anticipated needs and understand the purpose of severe thunderstorm warnings and what constitutes a threat to

life and property. Options of changing thresholds, using multi-tier thresholds, using regional criteria, and the impact of such changes were considered. Each team member weighed these questions with respect to their own discipline as well as the overall impact to society.

Recommendations from the team will be rendered later this fall. Those recommendations will include input from each subteam, an overall recommendation, and a concept for implementation.

RECOMMENDATION 13 - NWS reestablish its National Verification Committee, revise Chapters 72 and 73 of its Operations Manual, and prepare a revised verification plan; strengthen the role of its National Verification Director and OM authority and responsibility to oversee field verification efforts. Specifically, OM's director should review national verification issues, such as warning criteria and verification techniques, to ensure that verification "best practices," such as WFOs using recorders attached to telephone lines or pink-colored log-in sheets to record the original source and time of event information, are disseminated and used by regional and local offices.

Response - The NWS agrees that a more formal process should be established for general oversight on the development of new programs and specific verification issues. However, the NWS does not believe it is cost-effective to reestablish meetings of a National Verification Committee with the same charter as the original committee. There are already working groups in existence that are handling issues similar to those addressed by the National Verification Committee. For example, the Design, Development, and Testing Team for Verification has responsibility for the development and implementation of public and aviation verification programs in the AWIPS era. Also, the severe weather verification program continues to evolve in coordination with OM, SPC, NCDC, and the regional offices. Individual offices have been involved in testing the major software modifications before they are implemented on a national basis. Each of these efforts requires a level of oversight that is beyond the purview of a national committee.

OM will request that one person be identified from each region to represent that region's strategic vision on verification. These individuals will meet, or hold telephone/video conferences, with the national verification program manager at

regular intervals to discuss general verification issues and ensure follow-through on specific action items. Also, OM will issue an updated national verification plan and the complete the update of the verification chapters by July 1, 1998.

RECOMMENDATION 14 - The regional offices conduct random reviews, or "spot checks," of event data and warning information. The offices also should establish partnerships to randomly send event information to outside entities to evaluate its adequacy and validity and to develop procedures to gather information from utility companies, 911 emergency centers, and insurance companies so that hail and wind damage information is routinely shared among offices. Although limited funding for office visits can be a problem, NWS should provide funding for some regional WCM visits to WFOs in fiscal year 1998.

Response - Individual regions that have established procedures for office reviews will be encouraged to include "spot checks" of event data and warning information as part of that review. Ideally, a uniform review process can be established across the country.

Regional and national headquarters have provided lists of potential sources of data for WCM use. The concept of funding WCM visits is not a new concept. In FY 94, OM provided each WCM with a $1,000 discretionary fund for such local visits. In recent fiscal years, it was proposed that this discretionary fund be increased gradually to $3,000. Our budget shortfalls have forced the lowering of this item on the list of budget priorities. The regions suffered similar budget reductions and could not maintain the program. Within new funding, the NWS will examine ways to provide the resources necessary to conduct some of these activities.

RECOMMENDATION 15 - NHC systematically verify (a) the portion of each warning area that did and did not receive hurricane force winds and (b) its GFDL model (wind radii) and hurricane forecasts against surface observations.

Response - The Tropical Prediction Center (TPC)/NHC is in the process of testing various methods for verifying wind radii forecasts. It will report at the American Meteorological Society Annual Meeting in Phoenix (January 1998) on the results of verification of recent radii forecasts against "buoys of opportunity." Radii forecasts from recent years have been compiled for possible verification. The most elementary verification is against subsequent "0-hour"

forecasts, which are, of course, themselves imperfect. Verification at landfall brings its own set of problems because there are only one or two landfalls in a typical year.

As a statistical problem, the sample sizes are always going to be small. While NHC will schedule a continuing verification as recommended, statistical significance problems will detract from the utility of the results.

RECOMMENDATION 16 - NHC test its back-up plan at Hydro-meteorological Prediction Center (HPC) during the next off-season to determine whether HPC is a viable back-up center. If not, another back-up center should be selected immediately or activities taken to make HPC viable.

Response - TPC will schedule an exercise of the backup of TPC by HPC during the preseason 1998. The actual dates and other details will be worked out at the NOAA Hurricane Conference, November 18-21, 1997. It is anticipated that a small team from TPC will visit HPC to oversee the creation and dissemination of test versions of products issued by both NHC and the Tropical Analysis and Forecasting Branch.