Department of Commerce • National Oceanic & Atmospheric Administration • National Weather Service

NATIONAL WEATHER SERVICE INSTRUCTION 10-1601
JANUARY 6, 2003

Operations and Services
Performance, NWSPD 10-16

VERIFICATION PROCEDURES

NOTICE: This publication is available at: http://www.nws.noaa.gov/directives/.

OPR: OS52 (W. Lerner) **Certified by:** OS5 (D. Wernly)
Type of Issuance: Initial.

SUMMARY OF REVISIONS: This directive supersedes Weather Service Operations Manual Chapter C-75, "National Verification Program," Issuance 00-06, dated April 29, 2000; and Operations Manual Letters 07-00, dated October 16, 2000; 08-00, dated October 24, 2000; and 09-01 (also filed with F-42), dated September 28, 2001.

signed by December 23, 2002
Gregory A. Mandt Date
Director, Office of Climate,
Water, and Weather Services

Verification Procedures

Table of Contents: Page

1. Public Forecast Verification Procedures.

1.1 Public (Zone) Forecast Elements.

1.1.1 Introduction. The Advanced Weather Interactive Processing System (AWIPS) Verification Program (AVP) automatically collects forecasts and observations used in the verification of individual elements in the public (zone) forecasts. The AWIPS Verification Program has continuity back to the Automation of Field Operations and Services (AFOS) era. In support of the AVP, the Meteorological Development Laboratory (MDL):

a. provides and maintains AVP data/information collection and collation software operating at Weather Forecast Offices (WFO);

b. collects and archives basic AVP data transmitted from the WFOs to the National Oceanic and Atmospheric Administration (NOAA) Central Computer Facility (NCCF) in Bowie, Maryland;

c. checks AVP data for inconsistencies and removes questionable data from the database;

d. performs all centralized data processing of national verification statistics from the AVP data archived at MDL; and

e. provides documentation of the structure and application of AVP and the centralized verification software.

The following elements are verified at specific sites: max/min temperature, probability of precipitation, precipitation type, cloud amount, snowfall amount, wind speed.

1.1.2 Verification Sites. All national verification statistics are computed for specific sites, collectively called the national network. A list of all sites in the national network appears on the National Weather Service (NWS) verification Web page.

1.1.3 Data Input. Public forecasts, forecast guidance, and verifying observations are collected twice a day at times corresponding to the 0000 and 1200 coordinated universal time (UTC) model cycles. If the Interactive Forecast Preparation System (IFPS) is used to prepare the public forecast, the forecaster enters no data manually, except to correct the database. The public forecast elements, given in section 1.1.6, are automatically decoded from the station digital forecast matrices (DFM) twice a day when the forecaster runs the Coded City Forecast formatter. Conversely, if the IFPS is not used for public forecast issuance, the forecaster uses the verification editor to enter all public forecast elements into the database for each verification site (including the snowfall site, if applicable) no later than 2 hours after forecast issuance. The verification editor contains a detailed help file.

Aviation (AVN) Output Statistics (MOS) guidance elements are automatically entered into the verification database in the local AWIPS. MDL collects, archives, and computes verification statistics for the AVN and Nested Grid Model (NGM) MOS for as long as each product is produced.

All verifying observations are automatically taken from the aviation routine weather reports (METAR), aviation selected special weather reports (SPECI), and supplementary climate data (SCD) and entered into the verification database. The forecaster may view data in the verification database through the verification editor. If necessary, the forecaster may use this editor to modify any erroneous forecast/observation data in the database. For further details concerning the setup and running of the AWIPS verification software at the WFO, see the AWIPS User Manual (electronically accessible via AWIPS).

1.1.4 Data Transmission to the NCCF. Approximately 5 days after the start of a model cycle, all national network verification data for that cycle are automatically transmitted to the NCCF. Local site data are not transmitted to the NCCF but are retained in the WFO database. Following this transmission, the database may not be edited.

1.1.5 Public Forecast Verification Reports. MDL computes monthly verification scores for each verification site in the national network. See appendix A for a summary of verification scores. See Dagostro (1985) for a discussion of the application of verification scores to individual forecast elements. NWS employees may access verification statistics from the *Stats on Demand* feature of the NWS Verification Web Page. This Web page is operated and maintained by the Office of Climate Water and Weather Services (OCWWS) Performance Branch. *Stats on Demand* accesses an interactive database and generates verification statistics customized to the user's request. The user requests data for any public weather element and the desired MOS guidance product for one or more:

a. months;

b. model cycles;

c. projections; and

d. verification sites (single site, multiple sites, regional, or national).

The MDL Evaluation Branch also operates and maintains a Web page with public forecast verification statistics displayed in the form of scatter plots and reliability diagrams.

Periodically, the OCWWS computes and disseminates results from the national verification database focusing on a particular scientific, management, or training issue.

1.1.6 Elements. Projections for public elements verified at specific points in time (e.g., precipitation type) are defined as the number of hours elapsed since the appropriate model cycle initialization. Projections for public elements verified across forecast periods approximately

12 hours long (e.g., max/min temperature) are called "forecast periods," i.e., first forecast period, second forecast period, etc.

a. Max/Min Temperatures.

(1) Projections: The first four forecast periods are verified.

(2) WFO Forecasts and MOS guidance: Daytime maximum (max) and nighttime minimum (min) temperatures are forecast in whole degrees Fahrenheit for the first four 12- or 13-hour forecast periods. Daytime is defined as 7 a.m. to 7 p.m. Local Standard Time (LST). Nighttime is defined as 7 p.m. to 8 a.m. LST.

(3) Observations: Daytime max and nighttime min temperatures, in degrees Fahrenheit, are inferred from the METARs. An algorithm, described in Beasley (1995), uses the 6-hour max/min temperatures (1xxxx and 2xxxx groups) and hourly temperature readings to derive a daytime max and nighttime min.

(4) Verification Statistics: The following statistics are available: mean absolute error (MAE), mean algebraic error, root mean squared error, and a frequency distribution of absolute error classes. The MAE is also computed under certain circumstances, e.g., whenever the forecaster changed MOS guidance by 4° F or more, whenever the observed temperature changed by 10° F or more within 24 hours. With these MAE data, WFO forecasts are compared to MOS guidance and climatology.

b. Probability of Precipitation (PoP). Probability of 0.01 inch or greater liquid equivalent precipitation within a 12-hour period: 0000 to 1200 UTC and 1200 to 0000 UTC in the Contiguous United States and Alaska; 0600 to 1800 UTC and 1800 to 0600 UTC in Hawaii.

(1) Projections:

The first four forecast periods are verified.

(2) WFO Forecasts and MOS guidance: The following forecast percentages are allowed in the digital forecast matrices and verification database: {0, 5, 10, 20, 30, ..., 80, 90, 100}. MOS guidance PoPs, given to the nearest percent, are automatically rounded to the nearest allowable value.

(3) Observations: From METAR, 12-hour precipitation amounts to the nearest hundredth of an inch are automatically recorded. Sometimes measurable snowfall is not detected by the Automated Surface Observing System (ASOS) heated tipping bucket precipitation gauge, resulting in a

precipitation report of none or trace in the METARs and the AWIPS verification database. When this occurs, the AWIPS verification editor may be used to change the precipitation amount in the verification database from zero or trace to 0.03 inch. Under these circumstances, the MDL quality control software will overlook the discrepancy between the verification database (0.03 inch) and the METAR reports (no measurable precipitation), resulting in an acknowledgment by the MDL software that measurable precipitation occurred during the edited 12-hour period.

(4) Verification Statistics: The following statistics are available: the Brier score, the Brier score whenever measurable precipitation occurred and did not occur, and forecast reliability statistics for each allowable PoP value. With these measures, WFO forecasts are compared to MOS guidance and climatology.

c. Precipitation Type. Precipitation type is only verified September through May.

(1) Projections: 18, 30, and 42 hours after model initialization. The 18-hour projection occurs during the first forecast period, the 30-hour projection occurs during the second forecast period, and the 42-hour projection occurs during the third forecast period.

(2) WFO Forecasts and MOS guidance: Data are entered by category, where:

1: Freezing precipitation (freezing drizzle and freezing rain);
2: Frozen precipitation (snow, snow grains, hail, ice pellets, ice crystals);
3: Liquid precipitation (drizzle and rain).

If mixed precipitation is forecast, the most critical element is entered. The order of priority for "most critical" is freezing, frozen and liquid. Thus, if freezing rain and snow are expected, a "1" is entered. MOS precipitation type is a single category.

(3) Observations: From METARs and SPECIs, all precipitation types on the verification hour and within the period ± 1 hour of the verifying hour are automatically recorded as three digits, XYZ, where:

X=1 if freezing precipitation occurred, otherwise X=0;
Y=2 if frozen precipitation occurred, otherwise Y=0;
Z=3 if liquid precipitation occurred, otherwise Z=0.

(4) Verification Statistics: Verification statistics for this element are not available on the Web; however, MDL issues semiannual office notes with verification summaries of this element.

d. Cloud Amount.

(1) Projections: 12, 18, and 24 hours after model initialization.

(2) WFO Forecasts, MOS guidance, and Observations: Data are entered as a category:

1: Clear (CLR), no clouds;
2: FEW or Scattered (SCT), greater than 0 to 4/8 sky cover;
3: Broken (BKN), greater than 4/8 to less than 8/8 sky cover;
4: Overcast (OVC), 8/8 sky cover, Vertical Visibility (VV).

The ASOS only reports clouds up to 12,000 feet above ground level (AGL). Human observers augment the METARs with cloud layers above 12,000 feet at only a limited number of sites; therefore, MDL estimates total cloud amount at every verification site using METARs and information from the Geostationary Operational Environmental Satellite (GOES) cloud product (Kluepfel et al., 1994).

(3) Verification Statistics: Verification statistics for this element are not available on the Web; however, MDL issues semiannual office notes with verification summaries of this element.

e. Snowfall Amount. Snowfall is only verified September through May. The snowfall verification sites are the WFOs which report snowfall in the SCD. When a WFO that reports snowfall is not co-located with a national network verification site, that WFO is used for just snowfall verification and no other elements.

(1) Projection: 1st forecast period, 0000 to 1200 UTC for the late afternoon forecast issuance, and 1200 to 0000 UTC for the early morning forecast issuance.

(2) WFO Forecasts: The forecast value, in inches, is entered into the database.

(3) MOS guidance (where applicable): The MOS guidance categorical forecast is entered into the database. The categories are:

0: No snow or a trace of snow;
1: 0.1 inch to less than 2 inches;
2: 2 to less than 4 inches;
4: 4 to less than 6 inches;
6: 6 or more inches.

(4) Observations: Snowfall amounts are taken from the SCD and recorded in whole inches. For amounts above 1 inch, the decimal value is truncated (e.g., 2.7 inches is recorded as “2”). For values ranging from 0.1 to 0.9 inch, the snowfall is recorded as “1 inch.” A trace or less is recorded as zero. This unusual “rounding” system is a workaround for a software bug that only allows observations to be entered into the verification database as integers, even though snowfall in the SCDs is reported to the nearest 0.1 inch. Since all verification statistics for snowfall are computed from contingency tables of forecast or guidance data versus observed data with the aforementioned categories, the actual observed snowfall is important only if it is recorded as the correct category. These rules ensure that the observation category is always correct even if the integer value is not properly rounded. For example, when the observed snowfall is 0.1 to 0.9 inch, the recorded integer value is 1 inch, which is category 1. When the observed snowfall is 1.1 to 1.9 inches, the decimal is truncated. This makes the integer value 1 inch, which is also category 1. A report of 2.6 inches is also truncated, this time to 2 inches or category 2.

(5) Verification Statistics: Verification statistics for this element are not available on the Web; however, MDL issues semiannual office notes with verification summaries of this element.

f. Wind Speed.

(1) Projection: 42 hours after model initialization (occurs during the 3rd forecast period).

(2) WFO Forecasts and MOS: The wind speed, in knots, is recorded at the verifying hour. Unlike the 12-, 18-, and 24-hour wind speeds, which are taken from the terminal aerodrome forecasts (TAF), the 42-hour wind speed is a public forecast element.

(3) Observations: From METARs and SPECIs, IFPS automatically records the wind speed, in knots, at the verifying hour. The highest sustained speed within the period ±3 hours of the verifying hour is also recorded.

(4) Verification Statistics: Verification statistics for this element are not available on the Web; however, MDL issues semiannual office notes with verification summaries of this element.

1.2 Winter Storm and High Wind Warnings. Perform winter storm and high wind warning verification manually at the WFO. For verification purposes, treat all winter storm, heavy snow, blizzard, heavy sleet, and ice storm warnings and events generically as **winter storm** warnings and events. Keep one set of statistics for winter storm verification and another for high wind verification.

1.2.1 Matching Warnings and Events. Treat each public forecast zone as a separate verification area. Therefore, count a warning covering three zones as three warned areas or three warnings. For verification purposes, define an event as a situation when weather conditions meet or exceed the local warning criteria set by the NWS region (e.g., 4 inches or more of snow in 12 hours or less, winds of 58 mph or greater for any duration).

Record warnings and events in separate databases. All listings in the event database must meet warning criteria. Do not record multiple events for a single zone. Count one verified warning and one warned event whenever the event occurs in a warned zone. Count one unwarned event if the event occurs in a zone with no warning. Record one unverified warning for each warned zone that does not experience an event. The following rules apply during special circumstances:

a. Count one verified warning and one warned event whenever a warning is issued up to 2 hours after the associated event reached warning criteria but prior to the end of the event.

b. Count one verified warning and one unwarned event whenever a warning is issued more than 2 hours after the associated event reached warning criteria but prior to the end of the event.

1.2.2 Extensions. Warnings are often extended in area and/or time. Count extensions of warnings to new areas (zones) as new warnings, i.e., one warning per zone. Do not verify any extensions in time of a zone already warned.

1.2.3 Lead Time. Compute a lead time for each zone that experiences an event. Subtract the time of warning issuance from the time when the event first met warning criteria in the zone. Set negative values to zero. If a zone experiences an event meeting warning criteria with no warning in effect, assign that event a lead time of zero. Compute average lead time from all the lead times listed in the event database, including zeroes.

1.2.4 Non-automated Warning Reports. The regional headquarters report the fiscal year's verification statistics to the OCWWS Performance Branch. Record winter storm and high wind warning data separately, using the format in Table 1. All reports for the current fiscal year are due by the fifth working day of February (October through December data), May (October through March data), August (October through June data), and November (entire previous fiscal year). The OCWWS Performance Branch subsequently collates regional data into national summaries.

Table 1. Example of format for reporting winter storm and high wind verification statistics.

Verification Statistics **Fiscal Year ____________** **Date of Report ___________** **Region ___________**	**Winter Storms**	**High Winds**
Number of Warnings Issued		
Number of Verified Warnings		
Number of Unverified Warnings		
Number of Events		
Number of Events with Warnings		
Number of Events without Warnings		
Average Lead Time		
Probability of Detection (POD)		
False Alarm Ration (FAR)		
Critical Success Index (CSI)		

2. Severe Weather Verification Procedures. This section describes the verification of all severe thunderstorm and tornado watches and warnings.

2.1 Warning Verification. The OCWWS Performance Branch is responsible for the operation and maintenance of the automated severe weather warning verification program.

2.1.1 Matching Warnings and Events. All warning data are automatically extracted from the warning products issued to the public. The basic area for a tornado or severe thunderstorm warning is the county. Therefore, for verification purposes, each county included in a warning statement is counted as a separate warning.

All event data are automatically taken from the final *Storm Data* reports prepared by the WFOs. Each severe weather warning may only be verified by a confirmed event meeting NWS warning criteria and occurring within the valid period and county represented by the warning. For verification purposes, multiple severe thunderstorm wind and hail events in the same county separated by less than 10 miles and 15 minutes are considered duplicates; therefore, only the first entry is recorded into the event database. This rule has the following exceptions:

a. Any event that causes death or injury is included in the event database.

b. Any event that causes crop or property damage in excess of $500,000 is included in the event database.

c. Any report of winds 65 knots or greater is included in the event database.

d. Any hail size report of 2 inches or greater is included in the event database.

e. An event is not considered a duplicate if it is the only event verifying a warning.

Any event not recorded in the verification database due to the aforementioned duplicate rule still appears in the publication *Storm Data*.

Warnings and events are recorded in separate databases. An event must meet warning criteria to be recorded in the event database. Whenever one or more recordable events occur in a warned county, the following are recorded: one verified warning; one warned event for each recordable event. One unwarned event is recorded for each event that occurs in a county with no warning. One unverified warning is counted for each warned county that does not experience an event.

Verification statistics are computed for tornado and severe thunderstorm warnings and events using one of three methods:

a. <u>Tornado only (TOR)</u>. A confirmed tornado is required to verify a tornado warning (using the TOR product). Likewise, to count as a warned event, a tornado must be covered by a tornado warning.

b. <u>Severe Thunderstorm only (SVR)</u>. A severe thunderstorm warning (using the SVR product) is only verified by a non-tornadic severe thunderstorm, and a non-tornadic severe thunderstorm is only counted as a warned event if it is covered by a severe thunderstorm warning. Therefore, a tornado event does <u>not</u> verify a severe thunderstorm warning, and a tornado warning does not cover a non-tornadic severe thunderstorm event.

c. <u>Generic</u>. All tornado and severe thunderstorm data are treated as a generic severe local storm. This means any tornado or severe thunderstorm warning may be verified by either a tornado or severe thunderstorm event. Likewise, to count as a warned event, any tornado or severe thunderstorm may be covered by either a tornado or severe thunderstorm warning.

2.1.2 <u>Lead Time</u>. The methodologies for computing the lead time in each county for tornado, severe thunderstorm, and generic severe thunderstorm/tornado events are identical. The time of warning issuance is subtracted from the time when an event meeting warning criteria was first reported in the county. Negative values are converted to zero. An event moving into a second county creates an additional event for the database. The lead time for the second event is based on the time the event first entered the second county. If one or more events occur in a county not covered by a warning, each unwarned event is assigned a lead time of zero. Average lead time is computed from all lead times listed in the event database, including zeroes.

2.1.3 Display of Verification Statistics. NWS employees access verification statistics through the *Stats on Demand* feature of the NWS Verification Web Page. *Stats on Demand* accesses an interactive database that provides verification statistics customized to the user's request. The user may request data by:

a. type of warning;

b. one or more dates (select beginning and ending date);

c. one or more counties, WFOs, states, NWS regions, or the contiguous United States;

d. severity of event, based on total cost of damage, number of fatalities, and/or tornado F-scale (optional).

2.1.4 Backup Mode for Warnings. When a WFO goes into backup mode, warnings are still sorted by county, so all warnings issued by the backup office are attributed to the primary WFO.

2.1.5 Preliminary Tornado Warning Reports. The regional headquarters report preliminary tornado warning verification statistics to the OCWWS Performance Branch. The report for the previous month is due by the 14th of each month. When the 14th of the month falls on a weekend or holiday, the statistics are due the last business day prior to the 14th. Present the statistics in the format given in Table 2. The OCWWS Performance Branch subsequently collates the regional data into a national summary.

2.2 Watch Verification. The Storm Prediction Center (SPC) is responsible for verifying the tornado and severe thunderstorm watches it issues. The area defined by a tornado or severe thunderstorm watch is defined as the verification area without regard to the number of counties affected. Weiss et al. (1980) describe how SPC accounts for variations in the size of convective watch areas. All event data are taken from the OCWWS database. Statistics are stratified for tornado and severe thunderstorm watches combined and for tornado watches only.

Table 2. Example of format for reporting preliminary tornado and flash flood warning verification statistics. Flash flood warning verification is explained in section 4.1

Preliminary Verification Statistics Month and Year ______________ Region ____________	**Tornadoes**	**Flash Floods**
Number of Warnings Issued		
Number of Verified Warnings		
Number of Unverified Warnings		
Number of Events		
Number of Events with Warnings		
Number of Events without Warnings		
Average Lead Time		
POD		
FAR		
CSI		

3. Marine Forecast Verification Procedures.

3.1 Coded Marine Forecasts.

3.1.1 Introduction. Marine forecasts are verified at fixed point locations at fixed points in time. The Marine Prediction Center (MPC), Tropical Prediction Center (TPC), and WFOs with marine forecast responsibility will issue coded marine verification forecasts (MVF) twice a day for each verification site in their individual coastal waters (CWF), offshore (OFF), Great Lake near shore (NSH), and Great Lake open lake (GLF) forecast areas. Discontinue the issuance of the MVF in the absence of operational verification sites within your area of responsibility. WFOs with marine responsibility are listed in Tables 3 and 4. Background material on marine forecast verification methodology are provided in Burroughs (1993), and software documentation appears in Burroughs and Nichols (1993).

Table 3. Coastal WFOs with marine responsibility.

<u>Eastern Region WFOs</u> Caribou, ME (CAR) Portland, ME (GYX) Boston, MA (BOX) New York City (OKX) Philadelphia (PHI) Baltimore, MD/Washington DC (LWX) Wakefield, VA (AKQ) Morehead City, NC (MHX) Wilmington, NC (ILM) Charleston, SC (CHS) <u>Southern Region WFOs</u> Jacksonville, FL (JAX) Melbourne, FL (MLB) Miami, FL (MFL) Key West, FL (EYW) Tampa Bay Area, FL (TBW) Tallahassee, FL (TAE) Mobile, AL (MOB) New Orleans, LA (LIX) Lake Charles, LA (LCH) Houston/Galveston, TX (HGX) Corpus Christi, TX (CRP) Brownsville, TX (BRO) San Juan, PR (TJSJ)	<u>Western Region WFOs</u> Seattle, WA (SEW) Portland, OR (PQR) Medford, OR (MFR) Eureka, CA (EKA) San Francisco, CA (MTR) Los Angeles, CA (LOX) San Diego, CA (SGX) <u>Alaska Region WFOs</u> Juneau, AK (PAJK) Anchorage, AK (PAFC) Fairbanks, AK (PAFG) <u>Pacific Region WFOs</u> Honolulu, HI (PHFO) Guam (PGUM) Pago Pago (NSTU)

Table 4. Great Lakes WFOs with marine responsibility.

<u>Eastern Region WFOs</u> Cleveland, OH (CLE) Buffalo, NY (BUF)	<u>Central Region WFOs</u> Duluth, MN (DLH) Marquette, MI (MQT) Gaylord, MI (APX) Detroit, MI (DTX) Green Bay, WI (GRB) Milwaukee, WI (MKX) Chicago, IL (LOT) Grand Rapids, MI (GRR)

3.1.2 Verification Sites. The WFOs with marine responsibility, MPC, and TPC will use any functioning buoy or Coastal Marine Automated Network (C-MAN) station residing within their respective forecast areas as a verification site. Remove any buoy or C-MAN that becomes inactive, i.e., no data available for verification. WFOs with Great Lakes marine responsibility will discontinue the MVF after the buoys are removed from the lakes for the winter. A list of national marine verification sites appears on the national verification Web page.

3.1.3 Coded Forecast Format. Code the MVF in accordance with the format in Table 5. Issue the MVF no later than 2 hours after issuing the CWF, OFF, NSH, or GLF, using forecast values meteorologically consistent with the worded forecasts, remembering the winds and waves in the MVF are intended only for the sensors of the buoys and C-MAN stations. See Table 6 for a sample CWF with the corresponding MVF.

Table 5. Definitions of code used in the MVF. See text for detailed explanation.

CODE FORMAT *%%F nn(space)xxxxx(space)*t_1t_1*/WW/ddff/*hh/t_2t_2/WW/ddff/*hh [LF][LF]$$*	
%%F	Code for computer and delimiter for operational forecast
nn	Forecaster number
xxxxx	Buoy/C-MAN identifier
t_1t_1	Time, in hours (UTC), of the midpoint of the valid period for the 16- to 20-hour forecast, i.e., 06 or 18 UTC
WW	Warning status: NO: No advisory or warning SC: Small craft advisory GL: Gale warning ST: Storm warning TS: Tropical storm warning HR: Hurricane warning HF: Warning for hurricane force winds in the absence of a hurricane
dd	Wind direction
ff	Wind speed
hh	Significant wave height
t_2t_2	Valid time for 30-hour forecast (UTC) Time, in hours (UTC), of the midpoint of the valid period for the 28- to 32-hour forecast, i.e., 06 or 18 UTC
[LF][LF]$$	End bulletin code (2 line feeds followed by turn off code)

Table 6. Examples of marine products.

Example of a segment of a Coastal Waters Forecast:

PZZ550-011600-
POINT ARENA TO PIGEON POINT OUT TO 20 NM-

300 AM PDT WED JUN XX XXXX

...GALE WARNING...

.TODAY...NW WIND 35 KT. WIND WAVES 4 TO 6 FT. NW SWELL 10 TO 12 FT. PATCHY FOG.
.TONIGHT...NW WIND 25 KT...DECREASING TO 10 TO 20 KT BY MORNING. WIND WAVES 3 TO 5 FT. NW SWELL 8 TO 10 FT. PATCHY FOG.
.THURSDAY...NW WIND 10 TO 20 KT. WIND WAVES 2 TO 4 FT. NW SWELL 6 TO 9 FT. PATCHY FOG.

Example of Corresponding Coded MVF:

FXUS56 KMTR 011030
MVF001

%%F56 46013 18/GL/3235/12/06/SC/3623/10
%%F56 46042 18/GL/3235/12/06/SC/3623/10

Note: SC (Small Craft Advisory) is indicated in the coded forecast for the second verification period based on the wind speed exceeding the small craft threshold, even though the corresponding coastal waters forecast has no "small craft advisory" headline. It was superceded by the "gale warning headline" due to the first period conditions.

A detailed explanation for each MVF entry is given below:

a. Forecaster Number (*nn*). Do NOT use comparative verification as an individual performance measure. However, once verification statistics become available for individuals, forecasters will review them for feedback, self-improvement, and knowledge.

b. Buoy/C-MAN Identifier (*xxxxx*). See section 3.1.2, Verification Sites.

c. Valid Periods (t_1t_1 and t_2t_2). The first valid period (t_1t_1) (UTC) in the MVF is 18 hours ± 2 hours following the 0000 or 1200 UTC model cycle, i.e., 1600 to 2000 UTC today for today's 0000 UTC cycle and 0400 to 0800 UTC tomorrow for today's 1200 UTC cycle. Therefore, the *WW, dd, ff,* and *hh* values immediately after t_1t_1 are 16- to 20-hour forecasts. The second valid period (t_2t_2) in the MVF is 30 hours ± 2 hours following the 0000 or 1200 UTC model cycle, i.e., 0400 to 0800 UTC tomorrow for today's 0000 UTC cycle and 1600 to 2000 UTC tomorrow for today's 1200 UTC cycle. Therefore, the *WW, dd, ff,* and *hh* values immediately after t_2t_2 are 28- to 32-hour forecasts.

d. Warning/Advisory Status for Winds and Waves (*WW*). Enter the warning/advisory status with one of the 2-character abbreviations, explained below. This entry should represent the worst conditions expected during the appropriate valid period. For example, if a gale warning is issued due to forecast winds increasing to gale force during the second forecast period of the offshore forecast, enter "NO" advisory/warning for the 16- to 20-hour forecast (first forecast period) and "GL" advisory/warning for the 28- to 32-hour forecast (second forecast period) even though the headline near the beginning of the offshore forecast reads "Gale Warning."

A forecaster may enter a certain advisory/warning category in the MVF (e.g., gales), but forecast a mean speed (section 3.1.3.f, wind speed, *ff*) less than the minimum threshold for that warning category. Both entries are legitimate because the gale warnings are issued for the maximum forecast speed during the valid period, and the forecast wind speed is for the mean sustained wind speed expected during the valid period. The following 2-character entries are allowed:

(1) NO: No warning or small craft advisory. Enter "NO" as the placeholder when wind speed is not forecast. Enter "NO" when a small craft advisory in the near shore forecast is issued solely for waves since none of the Great Lake C-MAN stations measure wave height.

(2) SC: Small craft advisory. Small craft advisories are only issued for CWFs and NSHs.

(3) GL: Gale warning.

(4) ST: Storm warning.

(5) TS: Tropical storm warning.

(6) HR: Hurricane warning.

(7) HF: Warning for hurricane force winds in the absence of a hurricane.

e. Wind Direction (*dd*). Enter mean wind direction for the MVF valid period in tens of degrees, e.g., enter "12" for a wind from 120 degrees. If a wind shift or variable winds are expected during the period, enter the forecast direction at the midpoint hour of the valid period (i.e., 0600 or 1800 UTC). If the wind direction is less than 100 degrees, place a zero in the tens digit, e.g., enter "07" for a wind from 70 degrees. When the wind speed equals 100 knots or more, add 50 to wind direction, e.g., enter "57" for a wind from 70 degrees when the speed is 100 knots or more. Enter "99" if wind is forecast to be variable based on regional guidelines or the wind direction is not forecast due to missing observation data.

f. Wind Speed (*ff*). Enter mean wind speed for the MVF valid period to the nearest knot, not to the nearest 5 knots, as expressed in the worded forecasts. Do not forecast 99 knots. If the wind speed is less than 10 knots, enter a zero in the tens digit place, e.g., enter "06" for 6 knots. For speeds of 100 knots or more, subtract 100 from the forecast speed and add 50 to the wind direction. For example, given a 110 knot wind from 270 degrees, enter "77" for wind direction and "10" for wind speed. If the wind speed is not forecast due to missing observation data, enter "99" for wind speed, and also enter "NO" as the placeholder in the warning/advisory position.

g. Significant Wave Height (*hh*). Enter mean significant wave height for the MVF valid period in feet. If less than 10 feet, place a zero in the tens digit, e.g., enter "08" for 8 feet. If the significant wave height is not forecast due to missing observation data, enter "99" as the placeholder.

3.1.4 Data Archiving and Processing. The Environmental Modeling Center archives marine forecast/observation data at the NCCF in Bowie, Maryland, and computes quarterly verification scores, which are linked to the NWS Verification Web Page.

3.1.5 Verification Statistics. Verification statistics are computed for warning/advisory category, wind direction, wind speed and significant wave height. These statistics are based on a series of 5 hourly buoy or C-MAN observations within the MVF valid periods. A summary of each element follows.

a. Warning/Advisory Status for Winds and Waves. The warning/advisory status is verified against the highest of the five hourly wind speed observations during the MVF valid period.

(1) The lower threshold that defines small craft advisories (SCA) is set locally or regionally, and these values are programmed into the marine verification software. Either the observed lower wave height threshold for an SCA or the observed lower wind speed threshold for an SCA verifies the advisory. A 33-knot observed wind is the upper threshold for verifying an SCA. SCAs are only issued for CWFs and NSHs.

(2) A 34- to 47-knot wind verifies a gale warning.

(3) A 48- to 63-knot wind verifies a storm warning.

(4) A 34- to 63-knot wind verifies a tropical storm warning.

(5) A wind exceeding 63 knots verifies a hurricane warning or a warning for hurricane force winds in the absence of a hurricane.

The advisory/warning categories in the CWFs and NSHs are verified in 5x5 contingency tables of forecast categories versus observed categories. The warning categories in the OFFs and GLFs are verified in 4x4 contingency tables of forecast categories versus observed categories.

b. Wind Speed. The coded forecast to the nearest knot is verified against the mean of the five hourly wind speed observations during the MVF valid period. Verification statistics are computed from the information contained in 7x7 contingency tables of forecasts versus observations. The wind speed categories are:

1: Less than 8 knots,
2: 8 to 12 knots,
3: 13 to 17 knots,
4: 18 to 22 knots,
5: 23 to 27 knots,
6: 28 to 32 knots,
7: Greater than 32 knots.

c. Wind Direction. Variable forecasts (coded ‘99’) are not verified. Each forecast is verified with a time-averaged observation from the valid period of the MVF, omitting any observation with a reported wind speed less than 8 knots. Under most circumstances, this is the unit vector resultant of the five hourly reported directions during the forecast valid period. If any of the remaining 8-knot or greater winds varied in direction from any of the others in the valid period by more than 90 degrees, then the forecast is verified with the wind direction at the midpoint hour of the valid period, i.e., 0600 or 1800 UTC. If that midpoint hour wind speed was less than 8 knots and the reported directions varied by more than 90 degrees, then wind direction for that valid period is not verified.

Verification statistics are computed from the information contained in 8x8 contingency tables of forecasts vs. observations. The categories are defined as the eight points of the compass:

North: 338 to 22 degrees,
Northeast: 23 to 67 degrees,
East: 68 to 112 degrees,
Southeast: 113 to 157 degrees,
South: 158 to 202 degrees,
Southwest: 203 to 247 degrees,
West: 248 to 292 degrees,
Northwest: 293 to 337 degrees.

d. Wave Height. The coded forecast to the nearest foot is verified against the mean of the five hourly significant wave height observations during the MVF valid period. Verification statistics are computed from the information contained in 7x7 contingency tables of forecasts versus observations. The categories are:

1: Less than 3 feet,
2: 3 to 5 feet,
3: 6 to 8 feet,
4: 9 to 12 feet,
5: 13 to 16 feet,
6: 17 to 20 feet,
7: Greater than 20 feet.

3.2. Coastal Flood and Lakeshore Flood Warnings. WFOs with marine forecast responsibility perform coastal flood warning/lakeshore flood warning (CFW) verification manually.

3.2.1 Matching Warnings and Events. Treat each public forecast zone as a separate verification area. Therefore, count a warning covering three zones as three warned areas or three warnings.
Only the following reportable events in a WFO's monthly *Storm Data* report verify a CFW.

a. Storm surge. Only storm surges from extratropical storms verify a CFW.

b. Seiche.

c. High astronomical tide.

Treat minor coastal or lakeshore flooding, such as nuisance flooding, as a non-event for verification purposes.

Record warnings and events in separate databases. All listings in the event database must meet warning criteria. Do not record multiple events for a single zone. Count one verified warning and one warned event whenever the event occurs in a warned zone. Count one unwarned event if the event occurs in a zone with no warning. Record one unverified warning for each warned zone that does not experience an event. The following rules apply during special circumstances:

a. Count one verified warning and one warned event whenever a warning is issued less than 2 hours after the associated event reached warning criteria but prior to the end of the event.

b. Count one verified warning and one unwarned event whenever a warning is issued more than 2 hours after the associated event reached warning criteria but prior to the end of the event.

3.2.2 Extensions. Warnings are often extended in area and/or time. Count extensions of warnings to new areas (zones) as new warnings, i.e., one warning per zone. Do not verify any extensions in time of a zone already warned.

3.2.3 Lead Time. Compute a lead time for each public forecast zone that experiences a coastal or lakeshore flood event. Subtract the time of warning issuance from the time when the event first met warning criteria in the zone. Set negative values to zero. If a zone experiences an event meeting warning criteria with no warning in effect, assign that event a lead time of zero. Compute average lead time from all the lead times listed in the event database, including zeroes.

3.2.4 Non-Automated Warning Reports. The regional headquarters report verification statistics to the OCWWS Marine and Coastal Weather Services Branch. Use the format in Table 7. All reports for the previous extra-tropical storm season, i.e., July through June, are due by the fifth working day of August.

3.3 Special Marine Warnings (SMW). The OCWWS Performance Branch operates and maintains the automated SMW verification program. Any SMW issued for a coastal or Great Lake marine zone, Lake Okeechobee, or Lake Pontchartrain is verified.

3.3.1 Matching Warnings and Events. All warning data are automatically taken from the warning products issued to the public. Verifying events for SMWs are taken from the monthly *Storm Data* reports. Each marine zone represents a separate verification area. Therefore, a warning covering two zones counts as two warned areas or two separate warnings. Only the following reportable events in a WFO's monthly *Storm Data* report verify the SMW:

a. 3/4 inch or greater marine hail.

b. Marine thunderstorm wind, 34 knots or greater.

c. Waterspouts.

Table 7. Example of format for reporting CFW verification statistics.

Verification Statistics **Fiscal Year ______________** **Date of Report ____________** **Region ____________**	**CFW**
Number of Warnings Issued	
Number of Verified Warnings	
Number of Unverified Warnings	
Number of Events	
Number of Events with Warnings	
Number of Events without Warnings	
Average Lead Time	
POD	
FAR	
CSI	

Warnings and events are recorded in separate databases. An event must meet warning criteria to be recorded in the event database. Whenever one or more recordable events occur in a warned marine zone, the following are recorded: one verified warning; one warned event for each recordable event. One unwarned event is recorded for each event that occurs in a zone with no warning. One unverified warning is counted for each warned zone that does not experience an event.

3.3.2 Lead Time. The lead time for each event is computed separately for each marine zone by subtracting the time of warning issuance from the time when warning criteria were first met in the zone. If one or more events occur in a zone with no warning in effect, each unwarned event is assigned a lead time of zero. Average lead time is computed from all lead times listed in the event database, including the zeroes.

3.3.3 Display of Verification Statistics. NWS employees access verification statistics through the *Stats on Demand* feature of the NWS Verification Web Page. *Stats on Demand* accesses an interactive database that provides verification statistics customized to the user's request. The user may request data by:

a. one or more dates (select beginning and ending date);

b. one or more marine zones, WFOs, states, NWS regions, or the contiguous United States;

c. severity of event, based on total cost of damage and/or number of fatalities (optional).

3.3.4 <u>Backup Mode for Warnings</u>. When a WFO goes into backup mode, warnings are still sorted by county, so all warnings issued by the backup office are attributed to the primary WFO.

4. <u>Hydrologic Verification Procedures</u>. Hydrologic verification consists of the verification of flash flood warnings (FFW) and River Forecast Center (RFC) river stage forecasts.

4.1 <u>Flash Flood Warnings</u>. The OCWWS Performance Branch is responsible for the operation and maintenance of the automated FFW verification program.

4.1.1 <u>Matching Warnings and Events</u>. All warning data are automatically extracted from the warning products issued to the public. FFWs are issued by county. Since each county specified in a warning represents a separate verification area, a warning covering three counties is counted as three warned areas or three warnings. Events are automatically taken from the final *Storm Data* reports prepared by the WFOs. *Storm Data* reports entered as the event type "flash flood" verify an FFW.

For verification purposes, multiple flash flood events in the same county separated by less than 30 minutes are considered duplicates; therefore, only the first entry is made to the event database. This rule has the following exceptions:

a. any event that causes death or injury is included in the event database;

b. any event that causes crop or property damage in excess of $500,000 is included in the event database;

c. an event is not considered a duplicate if it is the only event verifying a warning.

Any event not recorded in the verification database due to the aforementioned duplicate rule still appears in the publication *Storm Data*.

Warnings and events are recorded in separate databases. An event must meet warning criteria to be recorded in the event database. Whenever one or more recordable events occur in a warned county, the following are recorded: one verified warning; one warned event for each recordable event. One unwarned event is recorded for each event that occurs in a county with no warning. One unverified warning is counted for each warned county that does not experience an event.

4.1.2 <u>Lead Time</u>. The lead time for each flash flood event is computed separately for each county by subtracting the time of warning issuance from the time when the event first occurred in the county. Negative values are converted to zero. If one or more events occur in a county with no warning in effect, each unwarned event is assigned a lead time of zero. Average lead time is computed from all lead times listed in the event database, including zeroes.

4.1.3 Display of Verification Statistics. NWS employees access FFW verification statistics through the *Stats on Demand* feature of the NWS Verification Web Page. *Stats on Demand* accesses an interactive database that provides verification statistics customized to the user's request. The user may request data by:

a. one or more dates (select beginning and ending date);

b. one or more counties, WFOs, states, NWS regions, or the contiguous United States.

4.1.4 Backup Mode for Warnings. When a WFO goes into backup mode, FFWs are still sorted by county, so all FFWs issued by the backup office are attributed to the primary WFO.

4.1.5 Preliminary Flash Flood Warning Reports. The regional headquarters report to the OCWWS Performance Branch no later than the 14th of each month preliminary flash flood warning verification statistics for the previous month. When the 14th of the month falls on a weekend or holiday, the statistics are due the last business day prior to the 14th. Present the statistics in the format given in Table 8. The OCWWS Performance Branch subsequently collates the regional data into a national summary.

4.2 RFC River Stage Forecasts. The RFCs operate the river stage forecast verification software, and the OCWWS Hydrological Services Division maintains policy. For a selected set of locations, both stream level observations (stage) and stage forecasts issued by RFCs are posted to a verification database at each RFC. Forecast values are matched with concurrent observations. From these pairs, verification statistics measuring the performance of the forecast system are calculated. The initial phase of river forecast verification is based on calculations of mean, mean absolute, and root mean square differences between observed and forecast values for each verification site on the river. Monthly verification statistics are automatically sent from the RFCs to the OCWWS Performance Branch.

NWS employees access verification statistics through the *Stats on Demand* feature of the NWS Verification Web Page. *Stats on Demand* accesses an interactive database and generates verification statistics customized to the user's request. The system allows verification statistics for locations to be grouped together by forecast lead time as well as hydrologic characteristics, i.e., (1) locations responding rapidly to rainfall, (2) locations with intermediate responses, and (3) locations with slow responses.

Table 8. Example of format for reporting preliminary tornado and flash flood warning verification statistics. Tornado warning verification is explained in section 2.1

Preliminary Verification Statistics Month and Year ______________ Region ____________	Tornadoes	Flash Floods
Number of Warnings Issued		
Number of Verified Warnings		
Number of Unverified Warnings		
Number of Events		
Number of Events with Warnings		
Number of Events without Warnings		
Average Lead Time		
POD		
FAR		
CSI		

5. Quantitative Precipitation Forecasts (QPF). Quantitative precipitation forecast verification statistics are found on the National Precipitation Verification Unit (NPVU) web page, which is operated and maintained by the OCWWS Performance Branch. The QPFs used in verification are issued on a 10 kilometer (km) grid by each RFC. The HPC issues gridded QPF guidance. QPFs are verified with quantitative precipitation estimates (QPE) issued by the RFCs. The QPEs are 4 km multi-sensor quality controlled estimates of observed precipitation. Monthly, all QPFs and QPEs are re-mapped to a 32 km AWIPS grid and used to compute verification statistics for each RFC area. Statistics are computed by the NPVU for model QPF guidance, HPC guidance, and all RFC-produced QPFs.

HPC also computes verification statistics for its QPFs and corresponding model QPFs. These data are posted to the HPC web page.

6. Aviation Verification Procedures.

6.1 Terminal Aerodrome Forecast Verification. The WFOs will use two software packages for TAF verification: the AWIPS Verification Program (AVP) and Aviation Verify.

6.1.1 AWIPS Verification Program.

a. Introduction. The AVP runs at the WFO on AWIPS and automatically collects forecasts and observations used in the verification of TAFs. The AWIPS Verification Program has continuity back to the AFOS era. In support of the AVP, MDL:

(1) provides and maintains AVP data/information collection and collation software operating at WFOs;

(2) collects and archives basic AVP data transmitted from the WFOs to the NOAA NCCF in Bowie, Maryland;

(3) checks AVP data for inconsistencies and removes questionable data from the database;

(4) performs all centralized data processing of national verification statistics from the AVP data archived at MDL; and

(5) provides documentation of the structure and application of AVP and the centralized verification software.

b. Verification Sites. National verification statistics are computed at specific sites, collectively called the national network. A list of all TAF sites in the national network appears on the NWS Verification Web Page. Each WFO has at least one TAF national network verification site. For continuity, additional sites from the AFOS era are also included, giving some WFOs more than one verification site.

c. Data Input. Each day, the scheduled 0600 and 1800 UTC TAFs are automatically decoded and stored in the AWIPS Verification Program database. AVN MOS guidance elements are automatically entered into the verification database. MDL collects, archives, and computes verification statistics from the AVN and NGM MOS for as long as each product is produced. The MOS guidance elements from the 0000 UTC cycle are paired with the 0600 UTC TAF elements, and the MOS guidance elements from the 1200 UTC cycle are paired with the 1800 UTC TAF elements. All verifying data taken from the METARs are automatically entered into the verification database.

The forecaster may view data recently entered into the verification database through the verification editor. If necessary, the forecaster may use this editor to modify any erroneous forecast/observation data in the database. For further details concerning the setup and running of the AWIPS verification software at the WFO, see the AWIPS User Manual (electronically accessible via AWIPS).

d. Data Transmission to the NCCF. Approximately 5 days after the start of a model cycle, all national network verification data for that cycle are automatically transmitted to the NCCF. Local site data are not transmitted to the NCCF but are retained in the WFO database. Following this transmission, the database may not be edited.

e. TAF Verification Reports. MDL computes monthly verification scores for each verification site in the national network. See appendix A for a summary of verification scores. Dagostro (1985) applies verification scores to individual forecast elements.

NWS employees may access verification statistics through the *Stats on Demand* feature of the NWS Verification Web Page. *Stats on Demand* accesses an interactive database and generates verification statistics customized to the user's request. The user may request data for any TAF element and the desired MOS guidance product for one or more:

(1) months,

(2) TAF beginning times,

(3) projections,

(4) verification sites (single site, multiple sites, regional, or national).

f. Elements. Only regularly scheduled TAFs beginning at 0600 and 1800 UTC are verified. Projections for TAF elements are defined as the number of hours elapsed since the beginning time of the TAF. The verification software evaluates the prevailing portion of the TAF and does not recognize temporary change (TEMPO) groups and probability (PROB) groups. Amended forecasts are not verified.

(1) Ceiling Height.

(a) Projections: 3, 6, 9, and 15 hours.

(b) TAF: The TAF ceiling height at each verifying hour is recorded in hundreds of feet AGL. Also, the following are recorded:

96: Ceiling above 9000 feet,
97: Unlimited ceiling.

(c) MOS Guidance: The MOS guidance (AVN and NGM) at each verifying hour is recorded as a category:

1: Less than 200 feet AGL,
2: 200 to 400 feet AGL,
3: 500 to 900 feet AGL,

4: 1000 to 3000 feet AGL,
5: Greater than 3000 to 6500 feet AGL,
6: Greater than 6500 to 12,000 feet AGL,
7: Greater than 12,000 feet AGL.

(d) Observations: From METAR, ceiling height at each verifying hour is recorded in hundreds of feet AGL. For ASOS observations with no augmented clouds above 12,000 feet:

96: Unlimited ceiling (no ceiling below 12,000 feet),
97: Ceiling above 9000 feet.

(e) Verification Statistics: NWS employees may run *Stats on Demand* to generate and display contingency tables and verification statistics for the TAF and one of the MOS guidance products. The contingency tables consist of TAF/MOS guidance data versus observations using the following categories:

1: Less than 500 feet,
2: 500 to 900 feet,
3: 1000 to 3000 feet,
4: Greater than 3000 feet.

A second breakdown of categories is also available:

A: Less than 200 feet,
B: 200 feet or above.

(2) Visibility.

(a) Projections: 3, 6, 9, and 15 hours.

(b) TAFs: The TAF visibility at each verifying hour is recorded in statute miles and fractions thereof. TAF visibilities above 6 statute miles are recorded as "7."

(c) MOS Guidance: The MOS guidance (AVN and NGM) at each verifying hour is recorded as a category. When the NGM MOS is used, the categories are:

1: Less than ½ statute mile,
2: ½ through less than 1 statute mile,
3: 1 through less than 3 statute miles,
4: 3 through 5 statute miles,
5: 6 or more statute miles.

When the AVN MOS is used, the categories are:

1: Less than or equal to 1/4 statute mile,
2: Greater than 1/4 to ½ statute mile,
3: Greater than ½ through less than 1 statute mile,
4: 1 through less than 3 statute miles,
5: 3 through 5 statute miles,
6: 6 statute miles,
7: Greater than 6 statute miles.

(d) Observations: From METAR, the visibility at each verifying hour is recorded in statute miles and fractions thereof. Visibilities above 7 miles are recorded as "8."

(e) Verification Statistics: NWS employees may run *Stats on Demand* to generate and display contingency tables and verification statistics for the TAF and one of the MOS guidance products. The contingency tables consist of TAF/MOS guidance data versus observations using the following categories:

1: Less than 1 statute mile,
2: 1 through less than 3 statute miles,
3: 3 through 5 statute miles,
4: Greater than 5 statute miles.

A second breakdown of categories is also available:

A: Less than or equal to 1/4 statute mile,
B: Greater than 1/4 statute mile.

(3) Wind Direction and Speed.

(a) Projections: 3, 9, and 15 hours.

(b) TAFs and MOS Guidance: From the TAFs and MOS at each verifying hour, the wind direction is recorded to the nearest ten degrees relative to true north and the sustained wind speed is recorded to the nearest knot.

(c) Observations: From METARs, the wind direction and sustained speed at each verifying hour are recorded (same units as forecasts).

(d) Verification Statistics: Verification statistics for this element are not yet available on the web; however, MDL issues semiannual office notes with verification summaries of this element.

6.1.2 Aviation Verify. WFOs are required to use Aviation Verify, a PC-based program, to evaluate TAF performance. Aviation Verify represents a philosophical shift in how TAFs are evaluated. The use of TEMPO and PROB groups in TAFs means two separate forecasts may be valid for a given terminal at the same time, the prevailing conditions and the TEMPO or PROB conditions. In the AWIPS verification program, just the prevailing conditions are evaluated even though the TEMPO or PROB forecast conditions often have a larger impact on operations and flight planning. Aviation Verify is an operational program that considers all forecasts in effect for every 5-minute segment of the TAF. Separate TEMPO and PROB group evaluation is also done. This philosophical shift will be reflected in a national TAF verification program under development. Until the new national program is complete, Aviation Verify is an interim operational TAF verification program with the following characteristics:

a. Verification Sites. All terminals for which a WFO issues TAFs are verified.

b. Data Collection. Since Aviation Verify is a PC-based program, all raw data are collected from AWIPS through Local Data Acquisition and Dissemination (LDAD) system. Ceilings, visibilities, wind direction, and wind speed data are collected from all scheduled TAFs, METARs, SPECIs, the NGM MOS, and the AVN MOS. Southern Region Headquarters ingests and temporarily stores for the current month all raw data for the entire Nation.

c. Data Transmission. At the beginning of each month, raw data from the previous month are automatically transmitted from Southern Region Headquarters to the OCWWS Aviation Services Branch.

d. Reports.

(1) WFO Results. Verification statistics for each forecaster or the entire WFO may be computed for a designated period.

(2) Regional Results. Using the raw data archived at each regional headquarters office, verification statistics for the NWS region or any subset of the region may be computed for a designated period using the program RAMVer.

(3) National Results. Using the data collected at OCWWS, verification statistics may be computed for the entire nation or any desired subset of the nation for a designated period using the program RAMVer.

Further details are available in the Aviation Verify user guide, located on the NWS Verification Web Page.

6.2 Aviation Weather Center (AWC) Verification Procedures.

6.2.1 Background. The AWC uses the automated Real-Time Verification System (RTVS), created specifically for verifying AWC's manually produced forecasts and various associated automated forecast algorithms. RTVS is a new software system which is continuously under review and revision as more and better sources of aviation verification observations are implemented. Verification techniques are under constant scrutiny in an effort to improve upon the subjectivity of pilot reports and other observations/observation products used in many aviation forecast verification procedures. Additionally, the RTVS' convective verification procedures are often revised and refined in an effort to provide the AWC with the best possible statistics for describing the accuracy of its convective forecasts. The National Convective Weather Diagnostic algorithm is currently used to verify AWC's convective products. While RTVS provides a baseline and a starting point for verification trend monitoring, the statistics are subject to change as RTVS evolves into a more mature system meeting the AWC's needs. Statistics are also prone to substantial monthly and seasonal variability based on the subjectivity and unreliable frequency of pilot reports. No standardized observing network exists for verifying aviation forecast variables, such as icing and turbulence. Despite these problems, statistics are presented as 12-month running averages.

6.2.2 Domestic Products Verified and Statistics Calculated.

a. Airman's Meteorological Information (AIRMET).

(1) Icing (AIRMET Zulu) and Turbulence (AIRMET Tango). The following verification statistics, defined in appendix A section 4.4 **(link)**, are calculated separately for AIRMET Zulu and AIRMET Tango: POD, POD of no observations (POD[N]), the percent area of AIRMET coverage across the domestic airspace (% Area), and the percent volume of AIRMET coverage across the domestic airspace.

(2) Instrument Flight Rules (IFR) Conditions (AIRMET Sierra). The following verification statistics are calculated: POD, FAR, and % Area.

b. Convective Forecasts.

(1) Convective Significant Meteorological Information (SIGMET). The following verification statistics are calculated: POD, FAR, % Area.

(2) Collaborative Convective Forecast Product: The following verification statistics are calculated: POD, FAR, and % Area.

7. Tropical Cyclone Verification Procedures. The National Hurricane Center (NHC) and the Central Pacific Hurricane Center (CPHC) verify tropical cyclone track and intensity forecasts.

7.1 Tropical Cyclone Forecasts/Advisories. NHC and CPHC issue Tropical Cyclone Forecast/Advisory products. The Tropical Cyclone Forecast/Advisory product will be referred to as the TCM product in this instruction. The first TCM product associated with each tropical system is normally issued when meteorological data indicate the formation of a tropical or subtropical cyclone. Subsequent advisories are issued at 0300, 0900, 1500, and 2100 UTC. Special forecasts/advisories are issued if significant changes to the forecast occur. Each advisory product contains 12-, 24-, 36-, 48-, and 72-hour forecast positions and maximum sustained wind speed.

7.1.1 Verification Elements. The following TCM elements are verified at 12, 24, 36, 48, and 72 hours:

a. Wind Speed**.** Wind speed is the primary product element used to verify a storm's *intensity*. The wind speed is a tropical cyclone's estimated maximum, one minute, sustained wind speed ten meters above the ground. This wind speed is generally referred to as the "surface wind speed". This maximum wind speed can occur anywhere within the cyclone's circulation. The forecast maximum wind speed is rounded to the closest five knots.

b. Location. The location is represented by the latitude and longitude of a storm's center, reported to the nearest tenth of a degree. Storm location is the product element used to verify a storm's *track*. The Track Forecast error is reported in nautical miles.

7.1.2 Verification Process. Each TCM product provides an estimated, current tropical cyclone location and wind speed. The location and wind speed are estimated by NHC and CPHC using some combination of: surface land observations, radiosonde observations (if available), Global Positioning System dropwindsondes (if available), other reconnaissance aircraft data (if available), ship reports, C-Man stations, moored buoys, oil rig platforms, radar data, satellite data, and synoptic analysis. Thus, the estimated storm location and wind speed serve as a *preliminary* means for verifying TCM products issued in the prior 12, 24, 36, 48, and 72 hours.

After each tropical cyclone season hurricane specialists review all available data and may refine the earlier-established estimates for tropical cyclone location and wind speed. The set of refined locations and wind speeds are referred to as the storm's final Best Track.

Verification is performed by comparing all TCM wind speed and location forecasts with the storm's final Best Track for tropical storm, hurricane, and subtropical stages.

7.2 Model Verification. A variety of models are run operationally and provide forecasted tropical cyclone tracks. Several models provide forecasted tropical cyclone intensities. The models range in complexity from simple statistical models to three-dimensional primitive equation models.

7.2.1 Verification Elements. The following model elements may be verified at 12, 24, 36, 48, and 72 hours:

a. Wind Speed. Wind speed is the primary product element used to verify a storm's *intensity*. The wind speed is a tropical cyclone's maximum, one minute, sustained wind speed ten meters above the ground. This wind speed is generally referred to as the "surface wind speed". This maximum wind speed can occur anywhere within the cyclone's circulation. The forecast maximum wind speed is rounded to the closest five knots.

b. Location. The (storm) location is represented by the latitude and longitude of a storm's center, reported to the nearest tenth of a degree. Storm location is the product element used to verify a storm's *track*. The Track Forecast error is reported in nautical miles.

7.2.2 Verification Process. Model forecasts for wind speed and/or location are verified with a storm's final Best Track.

7.3 Verification Reports. NHC and CPHC maintain storm-specific and seasonal verification statistics and post this information on the following Web sites, respectively:

http://www.nhc.noaa.gov

http://www.prh.noaa.gov/pr/hnl/cphc/pages/cphc.shtml

8. Climate Verification Procedures. The Climate Prediction Center (CPC) verifies its medium range and seasonal outlooks. Temperature and precipitation for the following forecasts are verified using first-order (CLIMAT international exchange) stations in the continental United States:

a. 6-10 day,

b. Week 2 (8-14 day),

c. monthly (issued with a 0.5 month lead time),

d. seasonal (0.5 to 12.5 month lead time).

The number of stations used in the verification varies between 60 and 100, depending on variable and time period. Temperature and precipitation for the extended lead seasonal forecasts (1.5 to 12.5 month lead time) are verified using climate division data from the National Climatic Data Center. Because these data only become available after 2 to 3 months, verification of these forecasts is delayed. A version of the Heidke Skill score (link to appendix A, section 2.c) is computed for climate forecast verification.

9. Model Verification Procedures. The Environmental Modeling Center verifies its numerical models. As part of its World Meteorological Organization responsibilities, the National Centers for Environmental Prediction Central Operations (NCO) sends monthly numerical model verification statistics to all World Forecast Centers. NCO also provides model verification statistics to the annual Numerical Weather Prediction report.

10. Use of Verification Information in Evaluating Forecaster Performance. Verification scores are not used to establish criteria for rating the forecasting and warning performance element of an individual's performance plan. Such use of the verification program is not appropriate because objectively derived verification scores by themselves seldom fully measure the quality of a set of forecasts. A forecaster demonstrates overall skill through his or her ability to analyze data, interpret guidance, and generate forecasts of maximum utility. Individual forecaster verification data is private matter between office management and employees and will be safeguarded.

To properly utilize forecast verification scores in the performance evaluation process, managers use scores as an indicator of excellence or of need for improvement. For example, a skill score which is "clearly above average" may be used, in part, to recognize excellence via the awards system. However, NWS managers at all echelons should be aware no two forecasters, offices, or management areas face the same series of weather events. Factors which must be taken into account include the number of forecasts produced, availability and quality of guidance, local climatology, and the increased level of difficulty associated with rare events. There is no substitute for sound supervisory judgment in accounting for these influences.

11 . References.

Beasley, R.A., 1995: AFOS-era forecast verification. NOAA Techniques Development Laboratory Computer Program NWS TDL CP 95-2, National Weather Service, NOAA, U.S. Department of Commerce, 50 pp.

Burroughs, L.D., 1993: National marine verification program - verification statistics. OPC Technical Note/NMC Office Note No. 400, National Weather Service, NOAA, U.S. Dept. of Commerce, 48 pp.

Burroughs, L.D., and R.E. Nichols, 1993: National marine verification program - concepts and data management. OPC Technical Note/NMC Office Note No. 393, National Weather Service, NOAA, U.S. Dept. of Commerce, 17 pp.

Dagostaro, V.J., 1985: The national AFOS-era verification data processing system. TDL Office Note 85-9, National Weather Service, NOAA, U.S. Department of Commerce, 47 pp.

Kluepfel, C.K., A.J. Schreiner, and D.A. Unger, 1994: The satellite-derived cloud cover product (sounder). NWS Technical Procedures Bulletin No. 410, NOAA, U.S. Department of Commerce, 15 pp.

Weiss, S.J., D.L. Kelly, and J.T. Schaefer, 1980: New objective verification techniques at the National Severe Storms Forecast Center. Preprints, 8th Conference on Weather Forecasting and Analysis, Denver, Colorado, American Meteorological Society, 412-419.

APPENDIX A - Verification Scores

Table of Contents: Page

1. Introduction. Verification scores are applied at the local, regional, and national levels. Different scores may be applied to the same data. The type of score selected for use depends upon the objective. Frequently used scores are given in this manual and presented within the context of specific elements and events subject to verification. An excellent reference for verification scores is Wilks (1995).

In general terms, the scores are measures of accuracy and skill. **Accuracy** is a measure of how much a forecast agrees with the event or element being forecast. The smaller the difference between the forecast and observation, the greater the accuracy. **Skill** is a measure of improvement of a forecast over an established standard. Examples of standards often used for comparison include the climatological frequency (or value), persistence, or forecasts made by another process (e.g., model output statistics). The greater the improvement, the greater the skill.

2. Generalized Contingency Table. A forecast/observation contingency table is often developed to summarize all variables by category. The following generalized contingency table has m mutually exclusive and exhaustive categories. The element X_{ij} gives the number of times the observation was in the *ith* category and the forecast was in the *jth* category. The row and column totals have the subscript p. Various scores can be computed from the elements in a contingency table such as:

Table A-1. Generalized Contingency Table.

	Forecast Category				
Observed Category	1	2	...	m	Total
1	X_{11}	X_{12}	...	X_{1m}	X_{1p}
2	X_{21}	X_{22}	...	X_{2m}	X_{2p}
...	...	...	...	...	...
m	X_{m1}	X_{m2}	...	X_{mm}	X_{mp}
Total	X_{p1}	X_{p2}	...	X_{pm}	X_{pp}

a. Percent Correct *(PC)* is the percentage of time a correct forecast was made *(j=i)* regardless of the category.

$$PC = \frac{\sum_{i=1}^{m} X_{ii}}{X_{pp}} \times 100$$

b. Bias by Category *(BIAS)* measures the tendency to overforecast (*BIAS* greater than 1) or underforecast (*BIAS* less than 1) a particular category, *i*. For each contingency table, *m* values of bias exist.

$$BIAS_i = \frac{X_{pi}}{X_{ip}}$$

c. Skill Score *(SS)*. A skill score measures the fraction of possible improvement of the forecasts over some standard or test set of forecasts.

$$SS = \frac{NC - E}{T - E}, \textit{ where:}$$

$$NC \ (number \ correct) = \sum_{i=1}^{m} X_{ii}; \quad T = X_{pp}$$

and E represents the standard or test set of forecasts.

c. Heidke Skill Score *(HSS)* When the test forecasts are the values expected by chance, computed from the marginal totals of the contingency table, the score is known as the Heidke skill score. A perfect Heidke skill score is one. Zero is indicative of no skill, and a negative score indicates skill worse than random forecasts.

$$HSS = \frac{NC - E}{T - E}, \; where:$$

$$NC \; (number \; correct) = \sum_{i=1}^{m} X_{ii}; \quad T = X_{pp}; \quad E = \sum_{i=1}^{m} \frac{X_{ip} X_{pi}}{T}$$

The CPC uses a version of the Heidke skill score for its main verification statistic. This is calculated by the formula:

$$HSS = \frac{NC - CH}{NT - CH} * 100$$

where, HSS is the Heidke skill score, NC is the total number of locations for which the forecast was correct, NT is the total number of locations for which a forecast was made, and CH is the number of locations which would be forecast correctly, on average, by chance. In a three class system (which is the how all CPC forecasts are characterized), one third of the locations are expected to be correct by chance. Thus if 99 locations are forecast, 33 are expected to be correctly forecast. This statistic results in scores of 100 if all locations are forecast correctly, zero if 33 are forecast correctly, and -50 if all locations are forecast incorrectly.

3. Specialized Contingency Table. The following contingency table applies when there are only two outcomes (yes or no) for a specific event or forecast. The number of correct forecasts for the specific event is given by *A*, and the number of events observed but not forecast is given by *B*. The number of forecasts which did not verify is represented by *C*. The number of times the specific event was neither forecast nor observed is represented by *X*.

Table A-2. Specialized Contingency Table

		Forecasts	
		Yes	No
Events	Yes	*A*	*B*
	No	*C*	*X*

The scores most frequently computed from this table are:

a. Probability of Detection *(POD)* is the fraction of actual events *(A+B)* correctly forecast *(A)*. In the case of warnings, the *POD* is the number of warned events divided by the total number of events. The more often an event is correctly forecast, the better the score. The best possible score is 1, the worst possible score is 0.

$$POD = \frac{A}{A + B}$$

b. False Alarm Ratio *(FAR)* is the fraction of all forecasts *(A+C)* which were incorrect *(C)*. In the case of warnings, the *FAR* is the number of false alarms (unverified warnings) divided by the total number of warnings. The more often an event is forecast and does not occur, the worse the score. The best possible score is 0, the worst possible score is 1.

$$FAR = \frac{C}{A + C}$$

The *POD* and *FAR* are most often used in the verification of watches and warnings. However, it is possible to apply the *POD* and *FAR* to many events and forecasts related to public and aviation elements. Two examples are the *POD* for ceilings below 1000 feet and the *FAR* for forecasts of freezing rain.

Overforecasting an event will achieve a high *POD* but at the expense of a high *FAR*. Overall success can be expressed by the critical success index *(CSI)*.

c. Critical Success Index is the ratio of correct forecasts *(A)* to the number of events *(A+B)* plus the number of incorrect forecasts *(C)*.

$$CSI = \frac{A}{A + B + C}$$

The best possible score is 1, the worst is 0. The relationship among *POD*, *FAR*, and *CSI* can be expressed as follows:

$$CSI = [(POD)^{-1} + (1 - FAR)^{-1} - 1]^{-1}$$

In the case of severe thunderstorm watches and warnings, the value of *A* varies depending upon whether it is taken from the warning or the event database. This is true because multiple events within a single county are sometimes counted as separate events in the event database, whereas only one warning can be in effect for a particular county at the same time. For this reason, the number of warned events in the event database, denoted below as A_e, may exceed the number of verified warnings in the warning database, denoted below as A_w. Using these conventions, the definitions of *POD* and *FAR* are

$$POD = \frac{A_e}{A_e + B}$$

$$FAR = \frac{C}{A_w + C}$$

Given these expressions for *POD* and *FAR* and the *CSI* formula, expressed in terms of *POD* and *FAR*, the *CSI* becomes:

$$CSI = \frac{A_w A_e}{A_w A_e + A_w B + A_e C}$$

4. Scores Computed for Specific Forecast Elements. Other scores may be computed, where N = number of cases; f_i = the *ith* forecast, and o_i = the *ith* observation (matching the forecast).

4.1 Temperature, Wind Speed and Direction, and Wave Height. Scores frequently computed for forecasts of temperature, wind speed and direction, and wave height include:

a. Mean Error *(ME)* indicates whether collective forecast values were too high or too low. This is also called the mean algebraic error.

$$ME = \frac{1}{N}\sum_{i=1}^{N} (f_i - o_i)$$

b. Mean Absolute Error *(MAE)* measures error without regard to the sign (whether positive or negative).

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |f_i - o_i|$$

c. Root Mean Square Error *(RMSE)* weights large errors more than the MAE.

$$RMSE = \sqrt{\frac{1}{N}\left[\sum_{i=1}^{N} (f_i - o_i)^2\right]}$$

d. Measuring Errors Against Some Standard. The above measures of accuracy *(ME, MAE, RMSE)* may also be computed for some forecast standard, such as Model

Output Statistics *(MOS)* guidance, climatology *(CLI)*, or persistence *(PER)*. For example, the *MAE* for *MOS* guidance forecasts (m_i) is

$$MAE_{MOS} = \frac{1}{N} \sum_{i=1}^{N} |m_i - o_i|$$

Forecast skill is determined by measuring the improvement of forecasts over a forecast standard. For example, the *MAE* may be used to compute the percent improvement of forecasts over *MOS*, $I(MAE)_{MOS}$.

$$I(MAE)_{MOS} = \frac{MAE_{MOS} - MAE}{MAE_{MOS}} \; x \; 100$$

Other examples include $I(RMSE)_{MOS}$, $I(MAE)_{CLI}$, and $I(RMSE)_{PER}$.

4.2 Probability of Precipitation. Scores typically computed for probability of precipitation verification include:

a. Brier Score *(BS)* measures the mean square error of all PoP intervals forecast. The standard NWS Brier score, defined below, is one-half the original score defined by Brier (1950).

$$BS = \frac{1}{N} \sum_{i=1}^{N} (f_i - o_i)^2$$

where, f_i = forecast probability for the *ith* case, o_i = observed precipitation occurrence (0 or 1), and N = the number of cases.

NWS forecasts (f_i) are expressed as one of the following values: 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. Hughes (1980) explains the Brier score and probability forecasting in detail.

b. Climatological Brier Score (BS_{CLI}) is an application of the Brier score to forecasts, c_i, consisting of climatic relative frequencies, *RF* (see below).

$$BS_{CLI} = \frac{1}{N} \sum_{i=1}^{N} (c_i - o_i)^2$$

c. Improvement over Climate Based on Brier Score $(I(BS)_{CLI})$ measures the improvement gained from actual forecasts versus climatological values.

$$I(BS)_{CLI} = \frac{BS_{CLI} - BS}{BS_{CLI}} \times 100$$

d. MOS Brier Score *(BS_{MOS})* is analogous to BS_{CLI}, except the Brier score is computed for MOS forecasts.

$$BS_{MOS} = \frac{1}{N}\sum_{i=1}^{N}(m_i - o_i)^2$$

where, m_i = MOS guidance probability the for the *ith* case. MOS guidance probabilities *(m_i)* are forecast to the nearest 0.01; however for NWS PoP verification, the m_i values are rounded to one of the following values: 0, 0.02, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0.

e. Improvement over MOS Based on Brier Score *($I(BS)_{MOS}$)* is analogous to $I(BS)_{CLI}$, except this score measures the improvement of the forecast over MOS.

$$I(BS)_{MOS} = \frac{BS_{MOS} - BS}{BS_{MOS}} \times 100$$

f. Relative Frequency of the Event *(RF)* is the fraction of the time the event occurred.

$$RF = \frac{1}{N}\sum_{i=1}^{N} o_i$$

g. Reliability, a measure of bias, compares the average forecast of the event with the relative frequency of the event. The reliability may be determined overall or by forecast interval, e.g., 10 percent PoP intervals.

$$\frac{1}{N}\sum_{i=1}^{N} f_i \quad \textit{compared with} \quad \frac{1}{N}\sum_{i=1}^{N} o_i \ ,$$

where, N is the total number of events or the number of events in the interval. If the average forecast of the event is larger (smaller) than the relative frequency of the event, the event was overforecast (underforecast).

h. Bias, Threat Score, POD, and FAR, when applied to QPF verification, are computed from gridded data for specific precipitation amount thresholds, e.g. 0.01 inch, 0.25 inch, 0.50 inch, 1.00 inch, etc. Bias *(B)* and Threat Score *(TS)* (Gilbert 1884; Junker et al. 1989; Schaefer 1990) (also known as the *CSI*) are defined as follows:

$$B = \frac{F}{O}$$

$$TS = CSI = \frac{H}{F + O - H}$$

where, *F* is the number of points forecast to have at least a certain amount (threshold) of precipitation, *O* is the number of points observed to have at least the threshold amount, and *H* is the number of points with correct forecasts for that threshold of precipitation. When the bias is less [greater] than unity for a given threshold, the forecast is under [over] forecasting the areal coverage for that amount. Geometrically, the threat score for a given threshold amount represents the ratio of the correctly predicted area to the threat area. Threat area is defined as the envelope of forecast and observed areas for that threshold. A perfect forecast yields a threat score of one, and a forecast with no areas correctly predicted receives a zero. The threat score, therefore, provides a measure of how accurately the location of precipitation is forecast within the valid period of the forecast. To receive a high threat score, forecast precipitation must be accurate—both spatially and temporally. For example, if a 1.00-inch isohyet is forecast, and all the observed rainfall within that area ranges from 0.8 to 0.99 inch, the forecaster's 1.00-inch threat score would be zero. However, the 0.8 to 0.99 inch area would favorably affect the 0.5-inch threat score. Also, a forecast area that is adjacent to an observed area with no overlap produces a zero threat score, and forecasts that are incorrect by just a couple of hours may receive little or no credit. Closely related to the threat score are *POD* and *FAR* which are expressed as:

$$POD = \frac{H}{O}$$

$$FAR = \frac{F - H}{F}$$

i. Equitable threat score *(ETS)* (Gandin and Murphy 1992, Messinger 1996) is similar to the threat score except the expected number of hits in a random

$$ETS = \frac{H - E}{F + O - H - E}$$

forecast, *E,* is subtracted from the numerator and denominator: where $E=FO/N$ and N is the number of points being verified. E is substantial for low precipitation categories, i.e., 0.10 inch or less in 24 hours, small at intermediate categories, and negligible for high categories, i.e., 1 inch or more in 24 hours.

4.3 Ceiling Height and Visibility. The Log Score *(LS)* is used for verifying ceiling height and visibility forecasts. It emphasizes accuracy in the more critical lower ceiling height and visibility ranges.

$$LS = \frac{50}{N} \sum_{i=1}^{N} \left| Log_{10}\left(\frac{f_i}{o_i}\right) \right|$$

Where f_i is the category of the *ith* forecast and o_i is the category of the *ith* observation. Note, f_i and o_i may also be used to represent the actual respective forecast and observed values of the element (i.e., ceiling height in feet, visibility in statute miles). Persistence is often used as the reference standard for evaluating ceiling height and visibility forecasts. The last hourly observation available to the forecaster before dissemination of the terminal aerodrome forecast defines the persistence forecasts of ceiling height and visibility to which the TAFs are compared.

4.4 Aviation Weather Center (AWC) Verification Statistics. The following statistics are used for verifying AWC forecasts:

a Probability of Detection *(POD)*. Same as section 3a of this appendix.

b. False Alarm Ratio *(FAR)*. Same as section 3b of this appendix.

c. Probability of Detection of "No" Observations *(POD[N])* is an estimate of the proportion of "no" observations that were correctly forecast (i.e., PIREPs which include reports such as negative icing or negative turbulence). Based on the contingency table presented in section 3 of this manual,

$$POD(N) = \frac{X}{X + C}$$

where, C = the number of forecasts which did not verify; X = the number of times the specific event was neither forecast nor observed.

d. Percent Area *(% Area)* is the percentage of the forecast domain's area where the forecast variable is expected to occur. It is the percent of the total area with a YES forecast.

e. Percent Volume *(% Vol)* is the percentage of the forecast domain's volume where the forecast variable is expected to occur. It is the percent of the total volume with a YES forecast.

5. References.

Brier, G.W., 1950: Verification of forecasts expressed in terms of probability. Monthly Weather Review, 78, 1-3.

Gandin, L.S., and A.H. Murphy, 1992: Equitable skill scores for categorical forecasts. Monthly Weather Review, 120, 361-370.

Gilbert, G.F., 1884: Finley's tornado predictions. American Meteorological Journal, 1, 166-172.

Hughes, L.A., 1980: Probability forecasting - reasons, procedures, problems. NOAA Technical Memorandum NWS FCST 24, National Weather Service, NOAA, U.S. Department of Commerce, 84 pp.

Junker, N.W., J.E. Hoke, and R.H. Grumm, 1989: Performance of NMC's regional models. Weather and Forecasting, 4, 368-390.

Messinger, F., 1996: Improvements in precipitation forecasts with the Eta regional model at the National Centers for Environmental Prediction: The 48-km upgrade. Bulletin of the American Meteorological Society, 77, 2637-2649.

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

Wilks, D.S., 1995: Statistical Methods in the Atmospheric Sciences. Academic Press, San Diego, CA, 467 pp.

APPENDIX B - Glossary of Terms

Storm Data - NOAA's official publication which documents the occurrence of storms and other significant weather phenomena having sufficient intensity to cause loss of life, injuries, significant property damage, disruption to commerce, and other noteworthy meteorological events.