

Historically, tornado warnings were issued and verified on a county-wide basis. Starting in fiscal year 2008 (October 2007), a new methodology was implemented. Storm-based tornado warnings are now issued and verified solely for the areas impacted by the warning and event, rather than for the entire county or counties containing the threat. The new methodology is probably presenting a greater challenge to forecasters, as hit rates have dropped to near 0.7 during this period. This isn't surprising since the warnings are issued for smaller, more geographically-focused areas than before. The customer service is better, but the more specific warnings are harder to verify.

Large swings in performance continue to occur with geography. Scores tend to be better in areas with widespread, organized tornado activity. For example, from the Texas and Oklahoma panhandles to central Oklahoma, Kansas, and eastern Nebraska (an area often known as tornado alley), hit rates for the past three years have ranged from 0.7 to 0.9, whereas outside of the central and southern United States, hit rates rarely exceed 0.5. False alarm ratios are lower in tornado alley (0.5 to 0.7), whereas the national false alarm ratios have been running between 0.7 and 0.8.

FYs 2012 and 2013 had an unusually low number of tornadoes reported. The total number reported during 2013 was less than half the total number reported during 2011. A low number of events tends to result in much lower forecast scores.



Historically, tornado warnings have been issued and verified on a county-wide basis. Starting in fiscal year 2008 (October 2007), a new methodology was implemented. Storm-based tornado warnings are now issued and verified solely for the areas impacted by the warning, rather than for the entire county or counties containing the threat.

Geographically, lead times tend to be higher in areas with more organized and frequent tornado activity. For example, in central Oklahoma, the heart of "tornado alley," lead times averaged just over 19 minutes in fiscal years 2010 and 2011. However, in areas outside the central and southern United States, lead times are often well below 10 minutes.

FYs 2012 and 2013 had an unusually low number of tornadoes reported. The total number reported during 2013 was less than half the total number reported during 2011. A low number of events tends to result in much lower forecast scores, which was true during 2012-13.



Historically, severe thunderstorm warnings (which include tornado warnings) have been issued and verified on a county-wide basis. Starting in fiscal year 2008 (October 2007), a new methodology was implemented. Storm-based severe thunderstorm warnings are now issued and verified solely for the areas impacted by the warning, rather than for the entire county or counties containing the threat.



Historically, severe thunderstorm warnings (which include tornado warnings) have been issued and verified on a county-wide basis. Starting in fiscal year 2008 (October 2007), a new methodology was implemented. Storm-based severe thunderstorm warnings are now issued and verified solely for the areas impacted by the warning, rather than for the entire county or counties containing the threat.

FYs 2012 and 2013 had an unusually low number of severe thunderstorms reported. A low number of events tends to result in much lower forecast scores, which was true during this period.

Historically, flash flood warnings have been issued and verified on a county-wide basis. Starting in fiscal year 2010 (October 2009), a new methodology was implemented. Storm-based flash flood warnings are now verified solely for the areas impacted by the warning, rather than for the entire county or counties containing the threat. The new methodology may be presenting a greater challenge to forecasters, as the hit rate dropped sharply to between 0.76 and 0.80 for fiscal years 2010 thru 2012. This is likely due to the warnings being issued for smaller, more geographically-focused areas than before. The customer service is better, but the more specific warnings are harder to verify. Lead times have continued their gradual rise, but 2012 saw a set-back to 53 minutes. Persistent, widespread droughts across the southern United States during 2012 likely helped cause the lower scores. Flash floods are harder to predict during widespread drought conditions because the forecasters have fewer opportunities to predict them.

The NWS has steadily cut the track error for which it forecasts the centers of hurricanes over the past 40 years. Today's track errors have been cut to less than one third the track errors typically measured during the 1970s and to less than one half the track errors typically measured during the 1990s.

The hit rate for winter storm warnings has held steady since we started measuring it in 1999, but the lead time for warning these events has nearly doubled.

Accuracy for 3- and 6-hour aviation forecasts for ceiling and visibility (combined) is depicted. The hit rate gives the percentage of time that observed ceilings less than 1000 feet or visibilities less than 3 miles were correctly forecast. The hit rate for these events has held relatively constant since 2006. The widespread droughts of the past three years have made ceilings and visibilities more difficult to forecast; however, the hit rates have remained within one standard deviation of the goal. The false alarm ratios, while a little higher than the 2006-2010 trend, have remained just below (better than) the goal.

Marine Forecasts

12- and 24-hour forecasts for marine winds and waves have steadily increased in accuracy over the past 15 years; however, those trends have begun to plateau during the past five to ten years. The use of the WaveWatch III model for forecasting wave heights has added an additional "boost" to these scores during the past two years. The 81% correct value in 2013 was quite extraordinary, but the new model is likely to make these extraordinary numbers more commonplace in the future.

New verification software is currently being developed that will enable the NWS to verify not just the first 24 hours of the marine forecast, but the entire forecast out to seven days. Including Days 2 thru 7 in these statistics will lower these scores considerably; however, the new system will enable NWS forecasters and managers to track the performance of the *entire* marine forecast. The new verification system will also include a lot of new data points that were not used in the past.

NOTE: IMPORTANT! Due to staffing issues within the Performance Branch, the update of PFM verification data has been discontinued indefinitely. (*Data only available from 01/2004 to 05/2012*).

In 2011, 2-day temperature forecasts had the same level of accuracy that 1-day forecasts had 25 years ago. In 2011, 3-day temperature forecasts had the same level of accuracy that 2-day forecasts had 15 years ago. In 2011, 3-day temperature forecasts had the same level of accuracy that 1-day forecasts had 40 years ago. In 2011, 5-day temperature forecasts had the same level of accuracy that 4-day forecasts had 10 years ago. In 2011, 5-day temperature forecasts had the same level of accuracy that 3-day forecasts had 20 years ago.

In 2011, 5-day temperature forecasts had the same level of accuracy that 2-day forecasts had 40 years ago.

In 2011, 7-day temperature forecasts had the same level of accuracy that 5-day forecasts had 20 years ago.