NWS Tornado Warnings: A Regional Analysis

Ronald Kennedy Jr.
California State University – Long Beach, Long Beach, California

Harold Brooks
NOAA National Severe Storms Laboratory, Norman, Oklahoma

James Correia Jr.
OU/CIMMS & NOAA/SPC, Norman, Oklahoma

[1]National Weather Center Research Experiences for Undergraduates Program
Norman, Oklahoma

ABSTRACT

Sergeant JP Finley used a 2 x 2 contingency table to forecast tornado occurrences. This table allowed a user to obtain probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), Success ratio (SR), and the Bias. These metrics help evaluate help provide the value and quality of a forecast. Recent studies of national weather service tornado warning performance for the Central, Eastern, and Sothern regions from 1986 - 2011. This study found a major decrease in false alarm ratio (FAR), probability of detection (POD) starting 2012 for the Central region, while the Southern and Eastern region displayed no change in false alarm ratio (FAR) or probability of detection (POD) until the following year. A warning forecaster has lots of information to consider before issuing a warning. They have radar, spotter reports, environmental conditions, historical data sets, etc. These elements all play a role in the weight of evidence required to issue a warning or not. If there is enough weight of evidence to issue a warning that goes above the threshold set in place most times the forecaster will issue a warning. This of the paper will be the regional analysis of regional performance focusing specifically on the time-period around the change in 2012.

1. Introduction

In 1884 JP Finley, a forecaster used a 2 x 2 contingency table using yes or no forecasting to forecast tornadoes. Finely used this method covering 18 districts in the central and eastern united states during the March, April, and May. Forecasts were produced twice a day and valid for 8 hours periods from 0700 and 1500 LT. A on the table translates to a hit, the tornado was forecasted and occurred. B, on the table, is the false alarm ratio, a tornado did not happen but was in the forecast to occur. C, on the table, is a missed event, a tornado occurred but not forecasted. D represents the correct negatives. No tornado occurred or forecasted. This 2 x 2 Contingency Table allows a forecaster with a method to evaluate the quality of a forecast. For example, the formula to find the probability of detection (POD) can show the percentage of the events forecasted happened. Using Finley's numbers, we can compute A/A+B (28/51) equals a POD of 0.549. Slightly more than half of the tornadoes that occurred were correctly predicted to occur. The formula for false alarm ratio (FAR) B/A+B+C shows the percentage of events that were forecasted and not observed. From the table we can calculate the FAR, 72/100 equals 0.72. Showing 72% of the forecasts for tornadoes turned out to be false alarms, no tornado occurred. Critical Success Index (CSI), the intersection of warnings and events divided by the union, A/A+B+C (28/123) equals 0.227. Of the tornado events that were either forecasted or observed, 23% of those were correctly forecast. Lastly, we have Bias, the number of forecasts divided by the number of events. (A+B)/(A+C) (100/51) equals 1.96 Tornadoes were predicted roughly twice as

[1] *Corresponding author address:* California State University – Long Beach, Pacific303, 4835 Pacific Coast Hwy, Long Beach, CA 90804, Ronald.Kennedy@student.csulb.edu.

often as they occurred. These numbers help forecasters also evaluate the value of a forecast. The cost or benefit a population may or may not incur from a decision based on the forecast. The quality of a forecast deals with how close the forecast matches the observation. This information is important in decision making, especially for large metropolitan areas. The National Weather Service tornado warning system has changed over time. Jimmy Correia Jr. and Harold Brooks studied tornado warning performance from 1986-2016 period for lead time, the probability of detection, false alarm ratio, and warning duration. Using metrics for mean lead time for tornadoes warned in advance, fraction of tornadoes warned in advance that work in a consistent way across the official changes in policy for warning issuance, including points in time when unofficial changes took place. The mean lead time for tornadoes warned in advance was comparatively constant from 1986-2011, while the fraction of tornadoes warned in advance increased through about 2006, and the false alarm ratio slowly decreased. The largest changes in performance took place in 2012 when the default warning duration decreased, and there is an increased emphasis on reducing false alarms. As a result, the lead time, probability of detection, and false alarm ratio all decrease in 2012. The change in the probability of detection, false alarm ratio, bias, success ratio, and critical success index that occurred in 2012 is the focus of this project. A regional analysis of what Brooks and Correia (2018) did nationally and the time period around the change that occurred in 2012.

2. Data & Methods

A warning forecaster has lots of information to consider before issuing a warning. They have radar, spotter reports, environmental conditions, historical data sets, etc. These elements all play a role in the weight of evidence required to issue a warning or not. If there is enough weight of evidence to issue a warning that goes above the threshold set in place most times the forecaster will issue a warning. The warning decision threshold shift shows to affiliate with a consistent overall skill, shown on the performance diagram. The performance diagram helps display the geometric relationship between the four measures of the yes, no forecast performance. The probability of detection (POD), false alarm ratio (FAR) or it's opposite, the success ratio (SR), bias and critical success index (CSI). For a good forecast probability of detection (POD), success ratio (SR), bias, and critical success index (CSI) move closer to the upper right corner on the diagram. A perfect forecast is in the top right-hand corner of the diagram. Subsequently, the closer to the bottom left corner, the more times the forecast is never right. Changes in the direction depending on to the left or right determine differences in probability of detection (POD) and success ratio (SR) and therefore bias and critical success index (CSI). The plots on a performance diagram provide a picture of the differences in performance. Optimal increases in accuracy are attained by moving at 45 degrees; this maintains an unbiased forecast by simultaneously increasing in detection and reducing false positives. How far up by the 45-degree angle represents how often the forecast is correct. When comparing the horizontal line versus the vertical line provides over-forecast and under-forecast statistics. This is measured by plotting the forecast quality measure relative to a reference forecast. Bias and CSI show at every point on the performance diagram as POD and FAR pairs and vice versa. A high basis will cause the probability of detection (POD) to rise. For example, if a forecaster forecast all the time the POD will reach one the max showing an even is never missed. If a forecaster does not want a high false alarm ratio (FAR) and never forecasts an event the probability of detection (POD) will go down to 0, representing no events have been predicted. The Bias measures the ratio of the frequency of forecast events to the frequency of observed events. Indicates whether the forecast system tends to under-forecast (BIAS<1) or over-forecast (BIAS>1) events. Does not measure how well the forecast relates to the observations, only measures comparative occurrences. Lastly, the success ratio Gives information about the possibility of an observed event, given that it was forecast. It is sensitive to false alarms but ignores misses.

3. Results

The probability of detection POD for the eastern region has the lowest consistent POD with a random peak in 2004. For the central region, the POD drops in 2012, but eastern and southern region stay the same. False alarm ratio (FAR) shows overall a consistent FAR from 1986 – 2017 with a slow drop beginning in the more recent years. Regionally the false alarm ratio (FAR) in the eastern region shows a sudden decrease for the 2004 year. While the central region drops in 2012, southern and eastern region increase in false alarm ratio (FAR). This change only affecting central region occurred without any policy changes nationally that impact how warnings were issued. The 2011 Joplin, MO service assessment played a role in this sudden decrease by empathizing the reduction of false alarms. The changes in the regions with the national performance diagram show, the county-based warnings are the grey scale, older warnings are the light colors, and recent warnings are the black colors. For the storm-based warnings, the red is the early period of the warning era, and the blue is the later more recent period of the warning era. The cluster of dots noticeably move down the critical success index space representing the quality of the forecast is the same, but the bias is reducing. The performance diagram by region overall appear similar, seems central and southern region overall show a higher probability of detection (POD) than eastern, but they are about the same. In the southern and eastern region, the year of 2012 shows the later storm-based warning era behaving like the early storm-based warning era. The central region storm-based warning era point for 2012 shows the opposite. With the year of 2012 moving away from the previous warning era, displaying the threshold change within only the central region. In the performance diagram, for the years 2011 – 2013 by region, is consistent with the previous diagrams. For the year 2012 in the central region, the bias was reduced, consequently, lowering the false alarm ratio (FAR) for the region. Whereas, southern and eastern did not follow reducing their false alarms until the following year.

4. Discussion

This study is focused on the years 2011 – 2013, with attention on the drop-in probability of detection and false alarm ratio for Central, Eastern, and Southern regions. This change occurred in the Central region first, then a year later the same change occurred in the Southern and Eastern region. For further study it would be interesting to look at specific local weather forecast offices starting in the Central region and moving to Southern and Eastern. The concentration by region and local weather forecast office would allow insight on why the local forecast offices raised their threshold for warning although no official change in warning was put in place by the national weather service. With data from this research and the local forecast office insight, this would allow further understanding on why the probability of detection and false alarm ratio dropped.

6. ACKNOWLEDGMENTS

This work would not have been possible without the funding from the National Science Foundation, Grant AGS 1560419. The corresponding author would like to thank Dr. Daphne LaDue and the National Weather Center Research Experience for Undergraduates for the opportunity to conduct research. I would also like to thank Jimmy Correia Jr, Harold Brooks, Melanie Schroers, for all their help and guidance during this summer program.

.

7. REFERENCES

Murphy, A.H., 1996: The Finley Affair: A Signal Event in the History of Forecast Verification. *Wea. Forecasting,* **11**, 3–20, https://doi.org/10.1175/1520-0434(1996)011<0003:TFAASE>2.0.CO;2

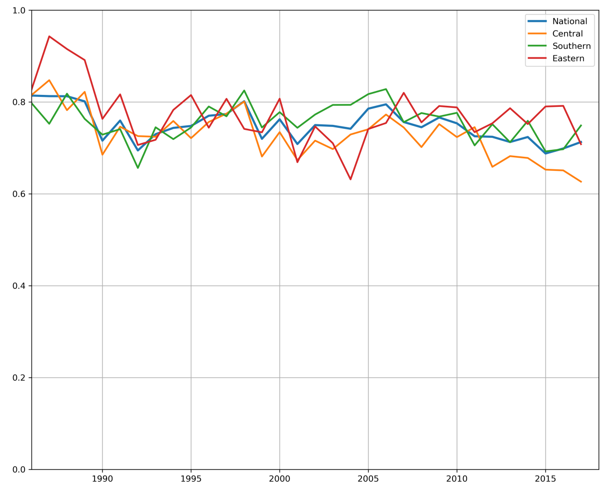
Brooks, H.E., 2004: TORNADO-WARNING PERFORMANCE IN THE PAST AND FUTURE: A Perspective from Signal Detection Theory. *Bull. Amer. Meteor. Soc.,* **85**, 837–844, https://doi.org/10.1175/BAMS-85-6-837

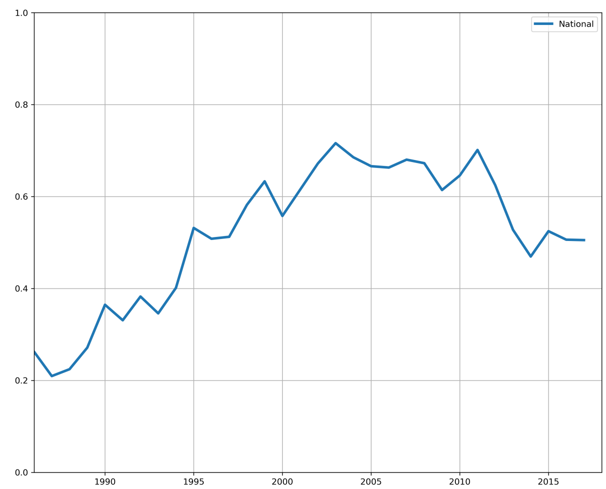
Simmons, K.M. and D. Sutter, 2009: False Alarms, Tornado Warnings, and Tornado Casualties. *Wea. Climate Soc.,* **1**, 38–53, https://doi.org/10.1175/2009WCAS1005.1

.

A perfect false alarm ratio would be 0. This has stayed consistent, but towards the later years it begins to decrease slowly.

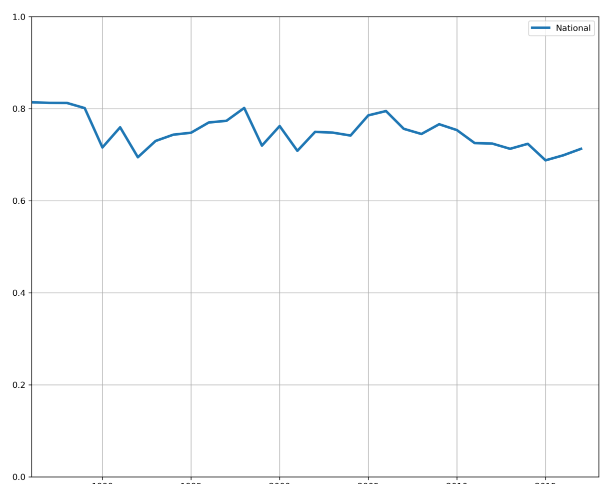
APPENDIX A.

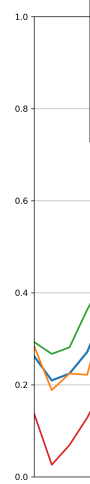




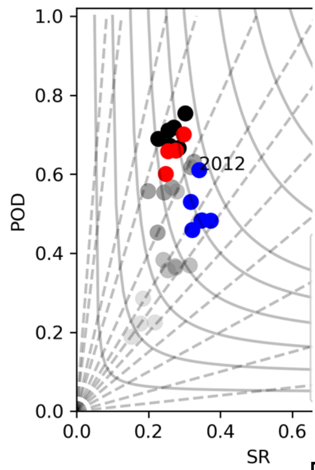
Then in 2011 – 2013 the probability of detection drops significantly. Eastern often has the lowest probability of detection consistently, with a random peak in 2004. For the years 2011 – 2013 there is a drop in POD that starts in the Central region first and then follows in Southern and Eastern the following years.

This graphic shows the probability of detection nationally from 1986 – 2017. A perfect POD would be 1. As you can see since 1986 Probability of detection has improved, peaked and in 2008 is the transition from county-based warnings to storm-based warnings.

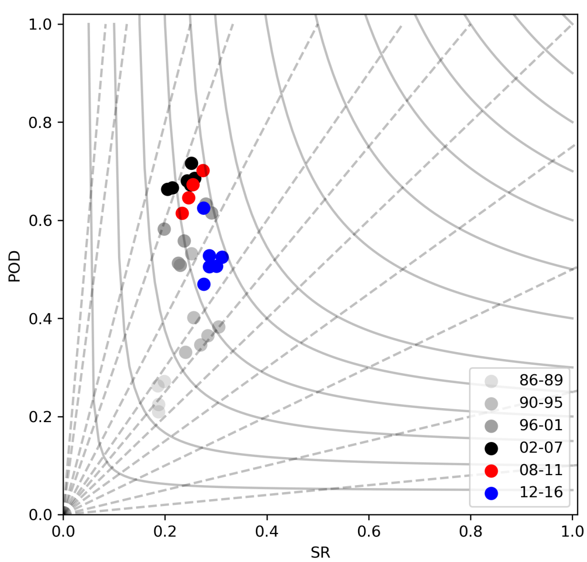




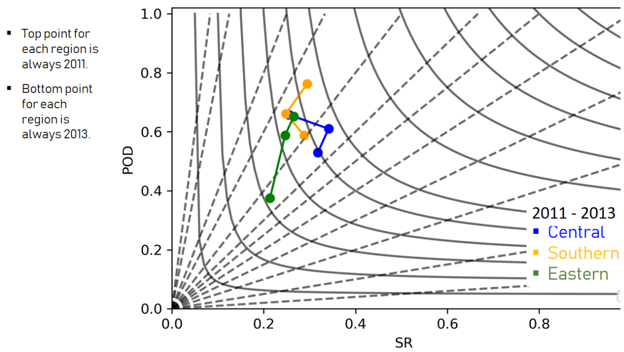
Eastern region stands out for the 2004 year. The false alarm ratio dropped significantly. In Central region false alarm ratio drops starting in 2012, while the other regions reduce later.



Central and Southern region overall show a higher POD. The location of the 2012 blue late era dot is different for the Central region. It's away from the reds (early storm-based warning era) and in the Southern and Eastern regions it is with the (early storm-based warning era) red dots. Causing it to look like 2012 belongs with the early.



The county-based warnings are all grey scale. Old warnings are light and recent warnings are black. The red is the early period and the blue is the late period. We see the cluster move down in the CSI space showing the quality of the forecast is the same, but the bias is reducing.



Keep in mind the top point for each region is always 2011 and the most bottom point for each region is always 2013. Reducing the bias reduces the false alarm ratio. Central region reduced their bias first then eastern and southern followed.