P2.7 **Nonmeteorological Factors in Warning Verification**

Sarah M. Davis and James G. LaDue

Cooperative Institute for Mesoscale Meteorological Studies and Warning Decision Training Branch

1. Introduction

The purpose of this study is to determine if non-meteorological factors, specifically county population density, distance to the nearest WSR-88D and time of day, that influence short-term warning performance for tornado, hail and convectively driven wind events. The current National Weather Service (NWS) Warning Verification Program focuses on warnings issued and whether or not those warnings verify (NWS 2003). These statistics do not address potential nonmeteorological factors. Other studies (Doswell, et.al. 1990, Murphy, et.al. 1987, Schaefer, J.T. 1990) have investigated the relationship between severe weather reports and population density, and even proximity to major highways. This paper applies a similar concept over a four-year period for several County Warning Areas (CWA) within the .NWS. The results and methodology, as products of this study, may be useful to help other CWAs determine the spatial characteristics of their warning performance.

2. Methodology

The basic unit of a warning is the county (NWSI 2003). Whenever a warning is issued, the specific counties the warning covers are mentioned in the text of the warning. If a warning covers five counties, then five warnings, one for each county, are recorded in the verification database. Also recorded in the database is whether or not an event occurred during the duration of the warning. If an event occurs in one warned county but not a neighboring warned county then the warning in the one county verifies and the warning in the neighboring county goes unverified. If two events in the same county are separated by less than 10 miles and 15 minutes then they are counted as one event in the database. Exceptions to this rule are: any event that causes death or injury, any event that causes crop or property damage over $500,000, any report of winds in excess of 65 knots, and any report of hail size greater than 2 inches in diameter. Also, an event is counted if it is the only one verifying a warning for a given county. These exceptions are included in the database regardless of proximity in time or space to each other (NWSI 2003).

This study employs the generic approach to warning verification as defined in Table 5 of NWSI 10-1601, where any severe event verifies a severe warning. For example, a tornado event would verify a severe thunderstorm warning and a severe thunderstorm would verify a tornado warning (NWSI, 2004 Section 2.1.1). Verification data was collected for a four-year period (1998-2001) for five separate County Warning Areas (CWAs). The CWAs used in this research are: AMA (Amarillo), LUB (Lubbock), OUN (Oklahoma City), TSA (Tulsa), and OHX (Memphis). Each CWA was broken down into its individual counties and five parameters were calculated for each county: Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI), warning density, and event density. The population density of each county was obtained from the Census 2000 website, and the parameters were compared with the population density.

* Corresponding Author's Address: Sarah M. Davis, Warning Decision Training Branch, 3200 Marshall Ave. Suite 202, Norman, OK 73152.

To calculate the POD, FAR and CSI the National Weather Service uses a specialized 2X2 contingency table:

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

Table 1: Specialized contingency table from NWSI 10-1601.

This contingency table implies that one of four things can happen: 1. An event is observed and forecasted (A), 2. An event is observed and not forecasted (B), 3. An event is forecasted and not observed (C), and 4. No events are observed or forecasted (X).

The Probability of Detection is the ratio of the number of events correctly forecasted over the number of actual events:

$$POD = \frac{A}{A + B} \quad 1]$$

For POD the best score is 1 and the worst score is 0.

The False Alarm Ratio is the number of incorrect forecasts over the total number of forecasts:

$$FAR = \frac{C}{A + C} \quad 2]$$

For FAR the best score is 0 and the worst score is 1.

The Critical Success Index is calculated by dividing the number of correct forecasts by the number of events plus the number of incorrect forecasts:

$$CSI = \frac{A}{A + B + C} \quad 3]$$

For CSI the best score is 1 and the worst score is 0.

The National Weather Service provided the verification database for the time period for this study. From this database the number of events, number of warnings, number of unwarned events, and the number of unverified warnings was calculated for each county and entered into an Excel spreadsheet. Using this data, the POD, FAR, and CSI was calculated for each county. The total area (in square miles) and total population of each county was obtained from the Census 2000 website (United States Census Bureau, 2003) and entered into the spreadsheet. Within the spreadsheet the area and population density were calculated. The warning density (#warnings/county area) and the event density (#events/county area) were also calculated. From these calculations scatter plots were generated to relate these parameters as a function of population density. Using data provided by the Geography Network website (ESRI, 2003) the parameters calculated were entered into a shape file and mapped using ArcView GIS (Geographic Information Systems) software. In addition, water bodies, town locations, county boundaries, interstate highways, and WSR-88D locations were plotted for geographic reference. The mean distance of the counties were broken into 50 km bins by determining the county's position relative to 50 km range rings around the nearest radar.

A different approach was used to evaluate the influence of time of day on warning verification. Using an online almanac, the average sunrise and sunset for each month was determined for the major city in each CWA. This information was used to stratify the POD, FAR, and CSI into daytime and nighttime values.

3. Results

When the number of events versus the population density are plotted (Figure 1) there appears to be a correlation between the two, however this correlation is stronger in some CWAs than others. LUB has the strongest correlation with an R^2 of 0.71 while AMA has the weakest with an R^2 value of 0.26. The maps of event density by county (Figure 2) indicate that four of the five major cities studied had more events than their surrounding counties. The only exception was Tulsa which had far fewer events.

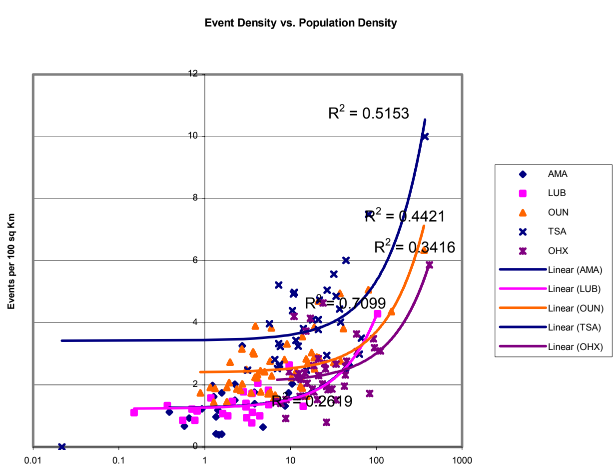

Figure 1: Event Density vs. Population Density for each CWA. Population Density is shown on a logarithmic scale. Also shown is the linear best fit with the R squared value for each CWA.

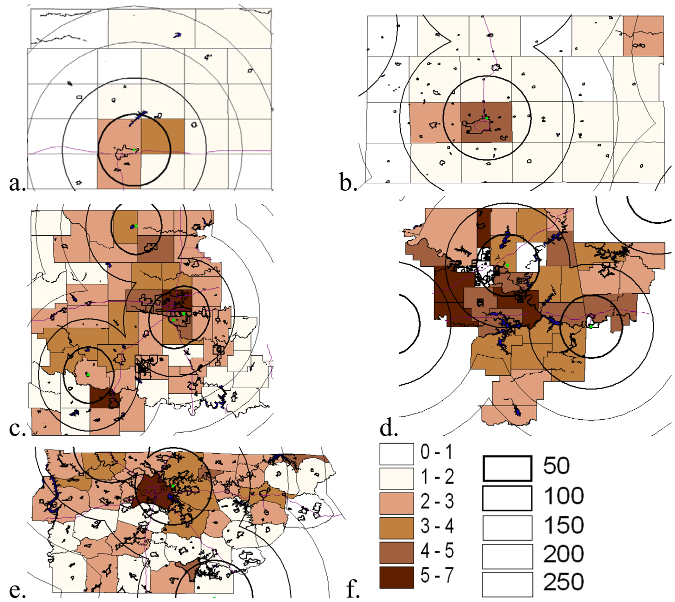

Figure 2: Map of Number of events per 100 square km by county for each CWA, a) AMA, b) LUB, c) OUN, d) TSA, e) OHX, and f) legend. The circles around the radar sites show a county's approximate distance from the radar in 50 kilometer intervals. Major towns/cities are outlined, and Interstate highways are shown for each CWA. CWAs are not shown on the same map scale.

The plot of warning density versus population density (Figure 3) shows that the correlation between population density and warning density is less than event density. OUN has the maximum correlation between population and warnings with an R^2 value of 0.44 while LUB has the lowest with an R^2 value of 0.14.

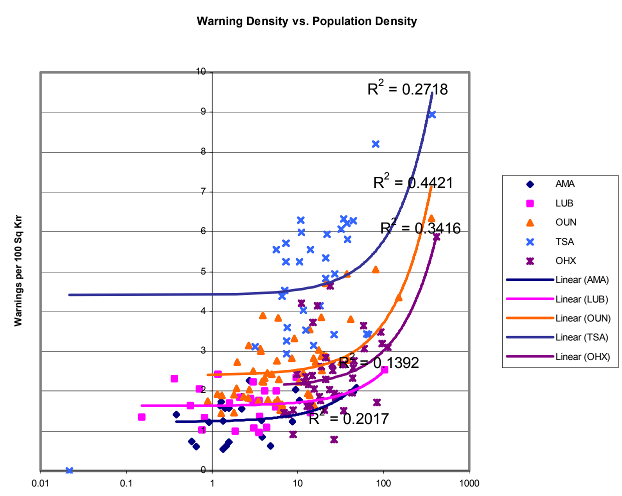

Figure 3: Warning Density vs. Population Density for each CWA. Population Density is shown on a logarithmic scale. Also shown is the linear best fit with the R squared value for each CWA.

The maps of warning density (Figure 4) hint that there are more warnings for the major city in each CWA, but the smaller cities do not seem to have more warnings than the more sparsely populated counties.

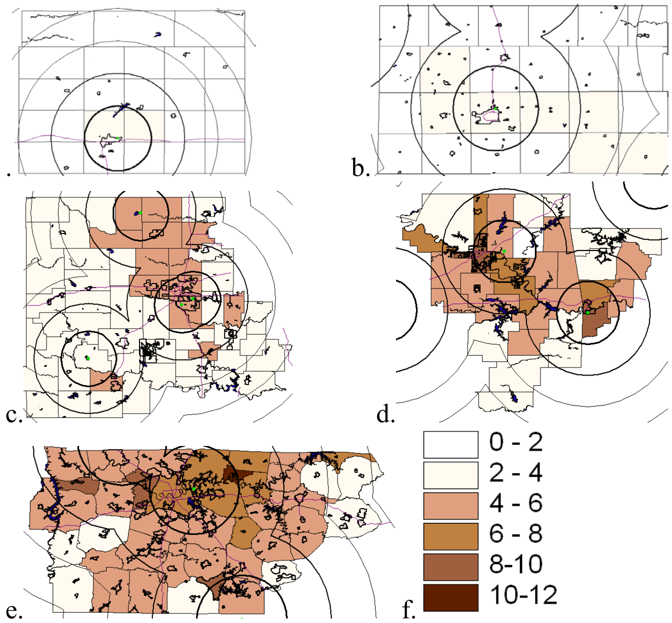

Figure 4: Map of Number of warnings per 100 square km by county for each CWA, a) AMA, b) LUB, c) OUN, d) TSA, e) OHX, and f) legend.

In order to avoid remove CWA-specific biases from the database, a normalization system was used for representation of each county's POD, FAR, and CSI. The overall POD, FAR, and CSI was calculated for each CWA over the entire four year period. The performance measure for each individual county was subtracted from the overall performance measure and multiplied by 100. For POD and CSI, a positive difference means that a particular county performed better than its overall CWA. For FAR, a negative difference means that a particular county performed better than the overall CWA. There is no consistent relationship between a particular performance measure for a county and that county's distance from a radar or population density.

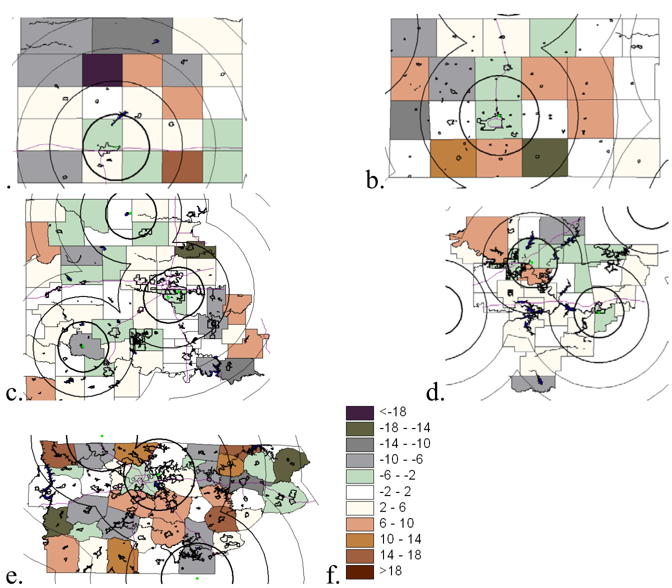

Figure 5: Map of POD difference by county for each CWA for a) AMA, b) LUB, c) OUN, d) TSA, e) OHX, and f) legend. Shown on each map is each county's difference from the CWA's overall POD. Better performing counties are shown in browns and lesser performing counties are shown in grays.

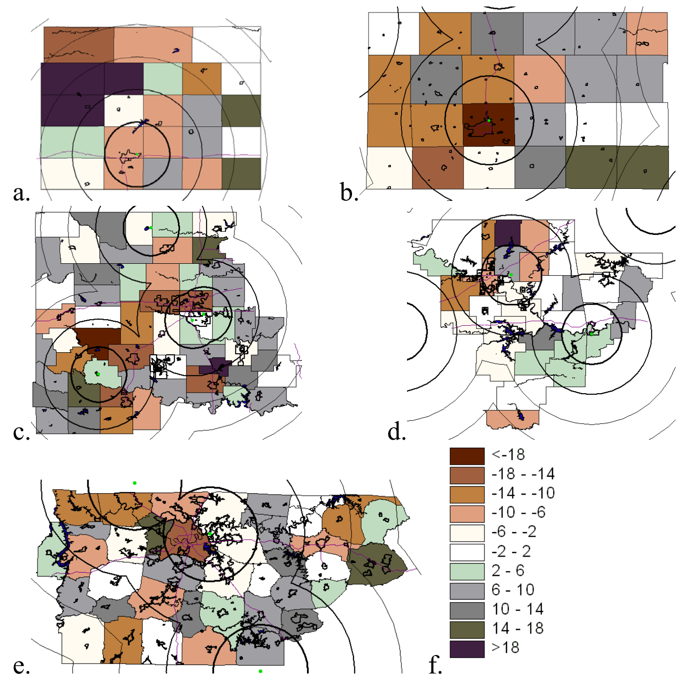

Figure 6: Map of FAR difference by county for each CWA for a) AMA, b) LUB, c) OUN, d) TSA, e) OHX, and f) legend. Shown on each map is each county's difference from the CWA's overall FAR. Better performing counties are shown in browns and lesser performing counties are shown in grays.

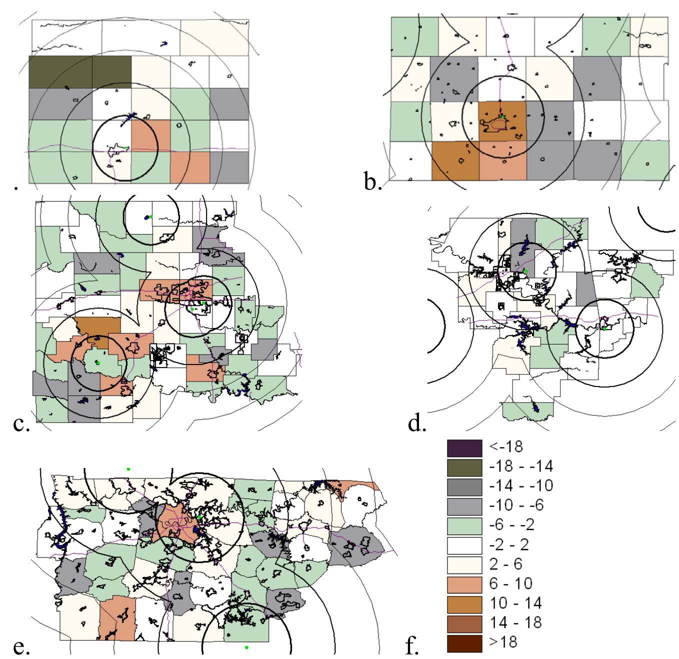

Figure7: Map of CSI difference by county for each CWA for a) AMA, b) LUB, c) OUN, d) TSA, e) OHX, and f) legend. Shown on each map is each county's difference from the CWA's overall CSI. Better performing counties are shown in browns and lesser performing counties are shown in grays.

The graph of population density versus CSI (Figure 8) also confirms that there is no relation between CSI and population density. Similar charts of POD and FAR versus population density are not included in this report, however they show the same results.

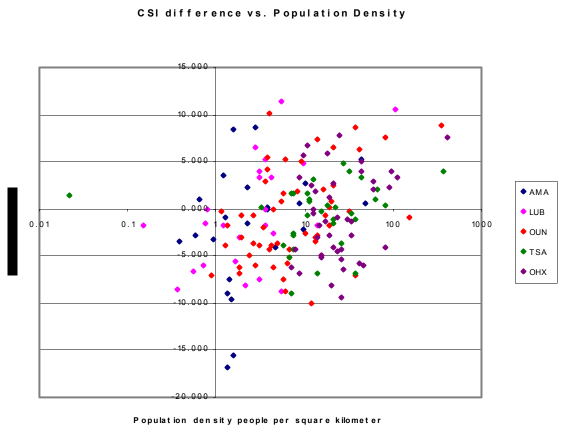

Figure 8: Graph showing the CSI difference versus population density for each CWA. Population density is shown on a logarithmic scale.

Figure 9 shows the CSI difference for daytime and nighttime CSI for each CWA. AMA, LUB and OUN tended to have higher scores during the day while TSA and OHX tended to have better scores during the night. The raw number of events and warnings during the day and night for each CWA are plotted in Figure 10. There is a maximum of severe events around OUN and TSA, while they taper off to the east and west. An interesting note here is that as one travels from east to west the majority of severe events shift from daytime to nighttime.

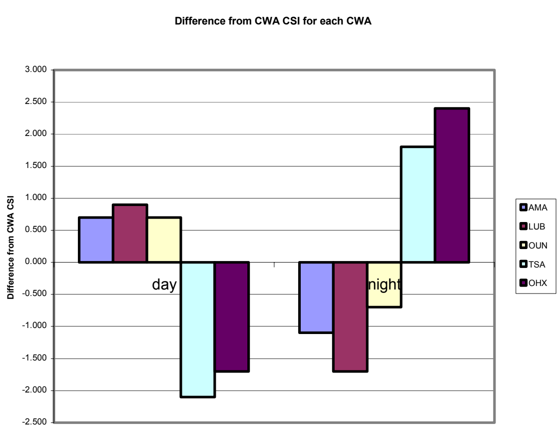

Figure 9: Graph showing CSI deviation for both daytime and nighttime.

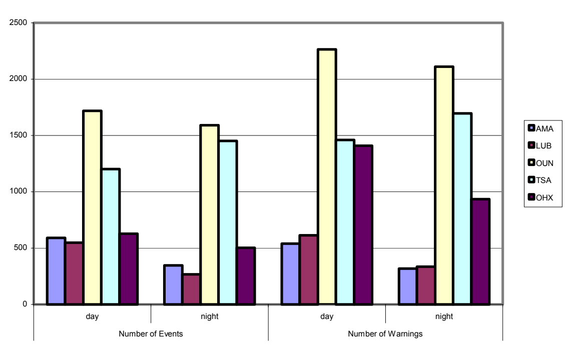

Figure 10: Chart showing the number of events and the number of warnings for each CWA in both the daytime and nighttime.

4. Conclusion

The correlations between population density and warning density (Figure 3) are significantly weaker than the correlations between population and event density (Figure 1). This result suggests that while population density may have an effect on the number of reported events, warning density is more a function of individual office's guidance on when to warn or not warn. In addition, POD, FAR, and CSI are not affected by population density or distance from the nearest radar. AMA, LUB, and OUN all had more warnings and events during the day than the night, and as the number of warnings increases the CSI increases. For this reason, the daytime CSI was better than the nighttime for these CWAs. TSA had more warnings and events during the night than the day and therefore the CSI was better in the night than during the day. OHX had more events and warnings during the day, however the CSI for OHX tended to be better during the night (Figure 9 and Figure 10). These results suggest that time of day has a varying effect between CWAs. Regardless of these results, this study introduces a methodology to assist offices with identifying and potentially improving warning performance in specific areas.

5. Works Cited:

Doswell, C.A., Davies-Jones, R., Keller, D.L. 1990: On Summary of Skill in Rare Event Forecasting Based on Contingency Tables. *Weather and Forecasting* Volume 5, December 1990, p 576-585.

ESRI, cited 2003: Geography Network. [Available online at http://www.geographynetwork.com]

Murphy, A.H., Ehrendorfer, M. 1987: On the Relationship between the Accuracy and Value of Forecasts in the Cost-Loss Ratio Situation. *Weather and Forecasting* Volume 2, September 1987, p 243-251.

Schaefer, J.T. 1990: The Critical Success Index as an Indicator of Warning Skill. *Weather and Forecasting* Volume 5, December 1990, p 570-575.

United States Census Bureau, cited 2003: U.S. Census 2000 [Available online at http://www.census.gov]

Verification Procedures: NWS Instruction 10-1601. Updated July 20, 2004. Accessible online at: http://www.nws.noaa.gov/directives/ .