1.5 POPULATION BIAS IN NOAA STORM EVENTS DATABASE REPORTS OF SEVERE WIND AND HAIL: IMPLICATIONS FOR RISK AND IMPACTS ASSESSMENT IN RURAL AREAS

Megan D. Walker
Blue Skies Meteorological Services, Gainesville, FL

1. INTRODUCTION

The National Oceanic and Atmospheric Administration (NOAA) Storm Events Database (SED) serves as an official record of severe and significant weather impacts in the United States. These data provide the foundation for our understanding of both the impact of individual storm events as well as the climatology of severe weather hazards across the country, and they are regularly relied upon by professionals across diverse disciplines, including meteorologists, insurers and reinsurers, engineers, planners, emergency managers, and the legal system in their quests to understand risks and impacts.

The Storm Events Database is arguably our most accurate and complete record of severe and significant weather in the United States. And yet, for a record to be included in the database, it must first be reported to the National Weather Service (NWS). This introduces the potential for a population bias in the dataset: in areas with lower population densities, there are fewer available observers to report severe weather impacts.

The presence of a population bias in the Storm Events Database could therefore lead to substantial under-representation of significant and severe weather in rural areas, which would suggest the need for a cautious and nuanced approach to estimating weather-related risks and impacts using SED data in regions with low population density.

While low population density regions may have relatively few weather observers, they are not necessarily characterized by low economic exposure. In recent decades, as the U.S. has expanded solar and wind power generation capacity, that development has been focused in rural areas, especially in the Great Plains, Midwest, and Southeast (Figure 1). In addition, domestic migration since 2010 has focused population growth primarily in suburban and adjacent rural counties in the Sunbelt and American West (Figure 2).

Many rural counties located away from population centers in the Great Plains and Midwest have experienced a net decrease in population since 2010; yet, these are precisely the areas where new solar and wind farms are being built. This pattern is especially evident in western Texas, Oklahoma, and Kansas as well as in southern Illinois (comparison of Figures 1 and 2). For this reason, net exposure may still increase in rural areas with decreasing population, as development of valuable energy infrastructure offsets the loss of residential and commercial infrastructure.

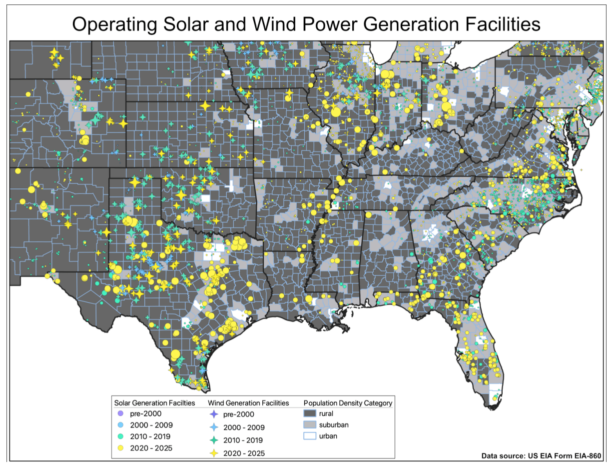

Figure 1: Locations of operating solar and wind generation facilities in the U.S., as of November 2025. The color of each marker indicates the decade in which the facility was installed, and the size of the marker is scaled by the nameplate capacity of the facility. Background shading is by county population density category, as described in the methodology. Data sources: U.S. EIA Form EIA-860; U.S. Census Bureau.

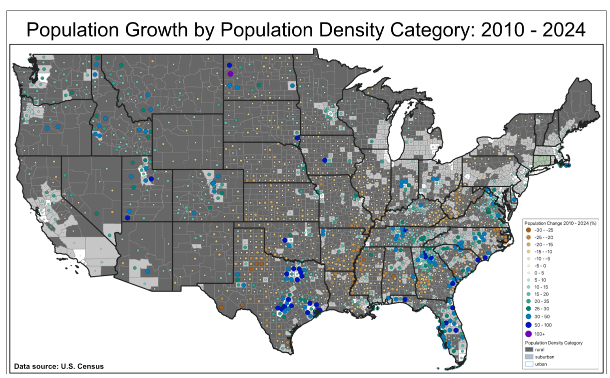

Figure 2: Percent change in county population from 2010 to 2024. Marker size and color indicate the direction and magnitude of the observed population change. Background shading is by county population density category, as described in the methodology. Data source: U.S. Census Bureau.

2. DATA AND METHODS

To test whether a population bias is present in the Storm Events Database, reports of severe wind and severe hail were examined from 1955 – 2024 across five metropolitan areas in the central United States that experience a relatively high frequency of both hazards: Dallas/Fort Worth, Oklahoma City, Tulsa, Kansas City, and St. Louis. The 91 counties that comprise these five

Corresponding author address: Megan D. Walker,
Blue Skies Meteorological Services, Gainesville, FL;
email: meganwr@blueskiesmeteorology.com

metropolitan areas are shown in Figure 3, shaded by the county population density in 2020.

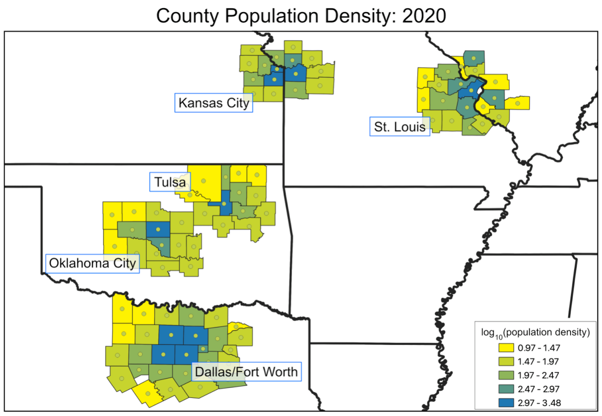

Figure 3: The study area consists of five metropolitan areas comprised of 91 counties. Counties are shaded by the $\log_{10}$ of the county population density in 2020. Data source: U.S. Census Bureau.

These metropolitan areas were chosen because each consists of a densely populated urban core surrounded by suburban areas of moderate population density and outlying rural areas (Figure 3), across which the severe weather climatology can be considered constant. In other words, across each metropolitan area, the population density varies while the frequency of severe wind and severe hail remains relatively steady.

The spatial densities of severe wind and hail reports (number of reports per square mile) were compared to the population density (number of people per square mile) in each county, thereby normalizing for the geographic area of each county.

This initial work used county population data from the 2020 U.S. Census. An underlying assumption of this analysis is therefore that the population density distribution across the five metropolitan areas has not shifted significantly over the study period. Put more simply, this analysis assumes that the counties that were most and least populous in 2020 were also most and least populous, respectively, in previous decades. To the author's knowledge, there have been no major shifts in the geographic locations of the urban center in the five metropolitan areas included in this study, which would support the validity of this assumption. However, the assumption will be tested in a planned update to this work that incorporates decadal Census data.

Throughout this analysis, counties are referred to as rural, suburban, and urban, with rural counties having a population density of less than 100 people per square mile, suburban counties having a population density between 100 and 1000 people per square mile, and urban counties having a population density of greater than 1000 people per square mile. The thresholds for these categories are based on evaluation of the population density distribution across the five metropolitan areas.

3. RESULTS AND DISCUSSION

3.1 Geographic Overview

When per-county severe hail reports density (Figure 4) and severe wind reports density (Figure 5) are mapped and compared to population density (Figure 3), a clear spatial pattern emerges. In all five metropolitan areas, the dense urban core counties have the highest density of both severe hail and severe wind reports, with storm reports density decreasing with distance from the urban core. This is the same spatial pattern observed with population density in these metropolitan areas, providing an early indication of the relationship between the three variables.

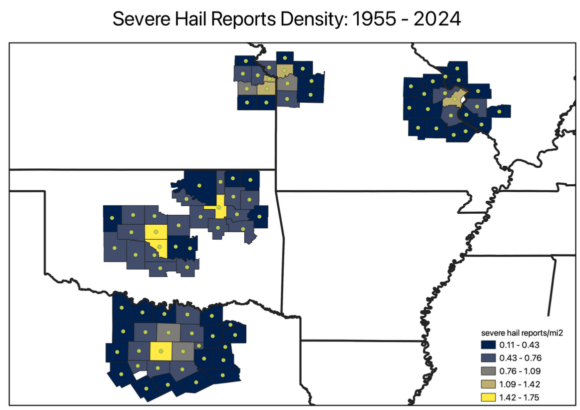

Figure 4: Spatial density of severe hail reports by county, 1955 - 2024. Data source: NOAA Storm Events Database.

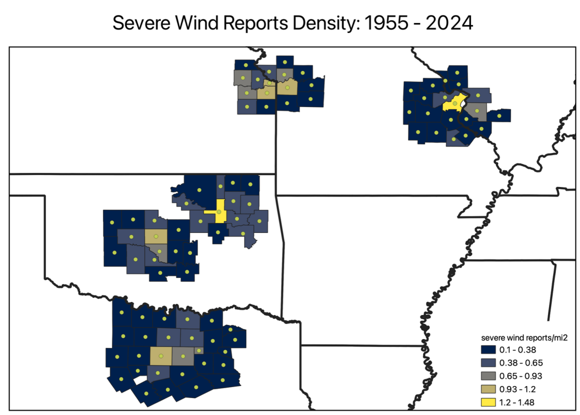

Figure 5: Spatial density of severe wind reports by county, 1955 - 2024. Data source: NOAA Storm Events Database.

Although all five metropolitan areas are relatively proximate and are climatologically prone to both severe wind and severe hail, regional differences in the relative frequencies of severe wind and severe hail are evident in the storm reports data (Figure 6). For example, the Oklahoma City metropolitan area (indicated by green markers in Figure 6) has experienced more severe hail than severe wind over the study period, whereas the St. Louis metropolitan area (indicated by red markers in Figure 6) has experienced more severe wind than severe hail. These observed regional differences over the 70-year study period are consistent with NOAA Storm Prediction Center 30-year severe wind and hail climatologies.

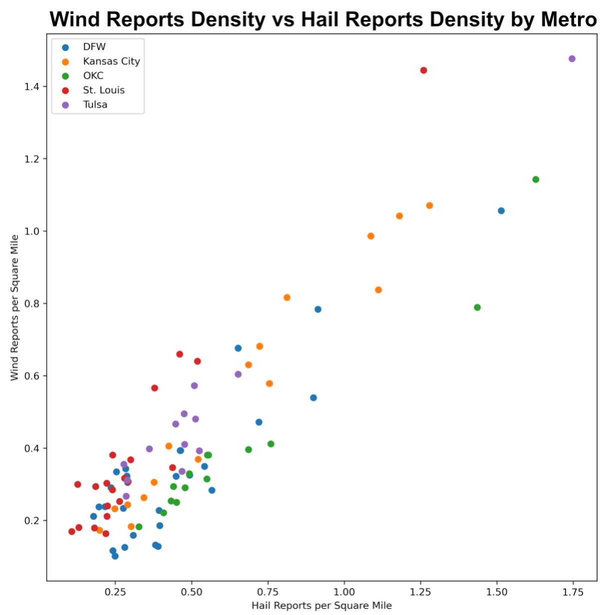

Figure 6: Spatial density of wind reports vs spatial density of hail reports for each county, grouped by metropolitan area, revealing slight regional differences in the relative frequency of severe wind and severe hail. Data source: NOAA SED.

3.2 Correlation between Population Density and Storm Reports Density

As suggested by Figures 3, 4, and 5, there are strong correlations between hail reports density, wind reports density, and population density. The Pearson's correlation coefficient is 0.72 between population density and hail reports density and is 0.74 between population density and wind reports density, indicating that population density alone explains approximately 50% of the variation in storm reports density. The correlation between wind reports density and hail reports density across all metropolitan areas is 0.90.

Figure 7 explicitly shows the relationship between these three variables, with the spatial density of wind reports plotted vs the spatial density of hail reports for all 91 counties in the study area, while the population density of each county is indicated by the color of the marker. The counties with high densities of wind and hail reports also tend to have high population densities.

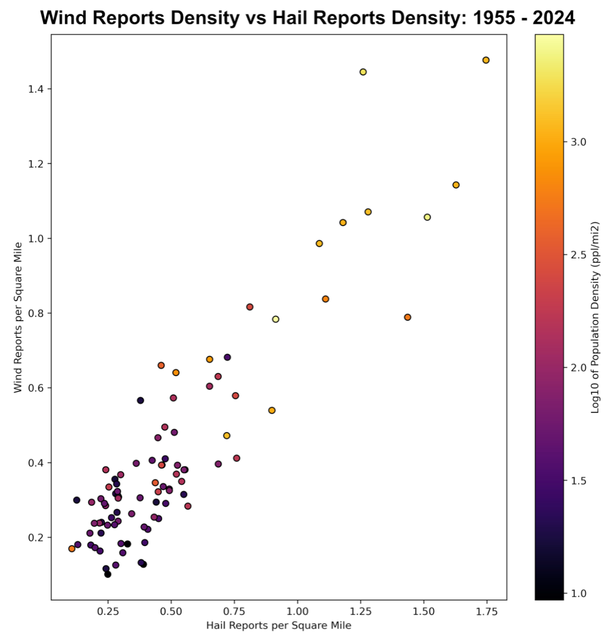

Figure 7: Spatial density of wind reports vs spatial density of hail reports for all 91 counties in the study area. The population density of each county is indicated by the color of the marker. Data sources: NOAA SED; U.S. Census Bureau.

3.3. Magnitude of the Population Bias

For this analysis, population bias is defined as a ratio of storm reports densities. For a given county, the population bias is calculated as the ratio of the observed storm reports density (SRD) of that county to the observed SRD of the most densely populated county within that same metropolitan area ("urban core county"). For county X located in metropolitan area A:

$$Population\ Bias_X = \frac{SRD\ of\ county\ X}{SRD\ of\ urban\ core\ county\ in\ A} \quad (1)$$

The population bias will therefore be 1 for the most densely populated county in each metropolitan area and will be less than 1 for any county with a SRD less than that of the urban core county within the same metropolitan area. The population bias was calculated separately for wind and hail.

Figure 8 shows the density of wind and hail reports in the urban core county of each metropolitan area (i.e. the denominator in the calculation of population bias). Figure 9 shows the population bias of each county plotted versus $\log_{10}$ of that county's population density. The average population biases for the three different population density categories (rural, suburban, and urban) were also calculated and are labeled in Figure 9.

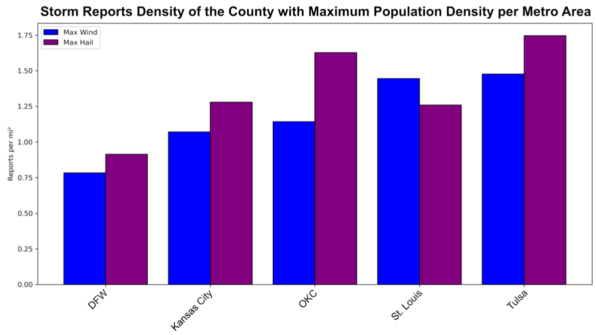

Figure 8: Spatial density of wind reports (blue) and hail reports (purple) of the county with the maximum population density in each metropolitan area. Data sources: NOAA SED; U.S. Census Bureau.

If one accepts climate as essentially constant across each metropolitan area, this storm reports density ratio (population bias) describes how the number of observed storm reports in a given county compares to the number of storm reports that theoretically would have been observed had the population density of that county been as high as the urban core county. In other words, it describes how many storm reports may be "missing" as a result of a lower population density.

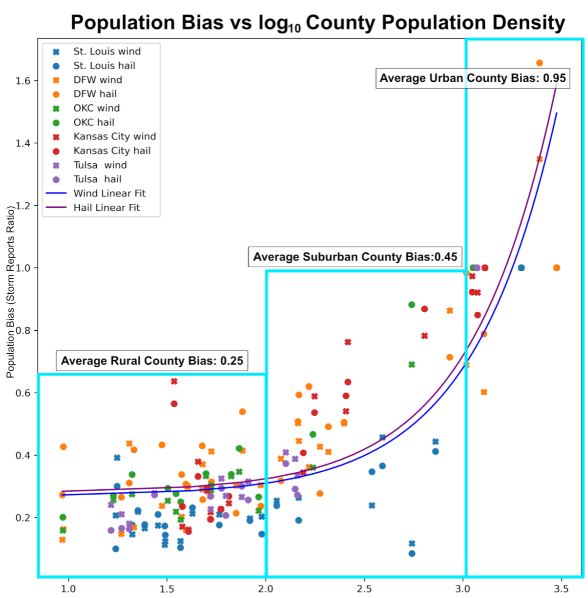

Figure 9: Population bias (storm reports density ratio) vs $\log_{10}$ of population density for wind ("x" marker) and hail ("o" marker). The color of the marker indicates the metropolitan area in which each county is located. Also shown are the linear regression lines for wind (blue) and hail (purple). Data sources: NOAA SED; U.S. Census Bureau.

On average, rural counties reported only 25% as many wind and hail reports per square mile as the urban core county within the same metropolitan area. Suburban counties reported on average 45% as many wind and hail reports per square mile as the urban core county within the same metropolitan area.

Figure 9 also shows how population bias scales with population density. With a logarithmic scale on the x-axis, the population bias data trace a broad exponential curve, indicating a roughly linear relationship between population bias and population density. Linear regressions fitted to the data had similar slopes, intercepts, RMSE, and R^2 values for wind and hail, further indicating that population biases behave similarly for both wind and hail reports with respect to population density (P_d).

$$Wind\ Population\ Bias = (4.11 * 10^{-4})P_d + 0.2685 \quad (2)$$
$$R^2 = 0.699; RMSE = 0.141$$

$$Hail\ Population\ Bias = (4.39 * 10^{-4})P_d + 0.2799 \quad (3)$$
$$R^2 = 0.694; RMSE = 0.152$$

5. CONCLUSIONS

A clear and consistent population bias was found in the NOAA Storm Events Database for the five central-U.S. metropolitan areas examined. Urban counties had, on average, four times more reports of severe wind and severe hail per square mile than neighboring rural counties, and the ratio of storm reports densities between counties was found to scale linearly with population density. This ratio was similar for both wind and hail reports.

These findings suggest that we may be underestimating the occurrence of severe wind and hail in rural areas, thereby underestimating storm-related risks. The accuracy of severe weather risk and impacts assessments in rural areas would therefore likely be improved by incorporating additional indicators of storm-related impacts that are less dependent on population density (e.g. NWS warnings and radar data).

6. REFERENCES

National Oceanic and Atmospheric Administration (NOAA): Storm Events Database

NOAA, National Weather Service, Storm Prediction Center: Severe Weather Maps, Graphics, and Data Page

United States Census Bureau: County Intercensal Population by Characteristics: 2010-2020

United States Census Bureau: County Population Totals and Components of Change: 2020-2024

United States Energy Information Administration: Form EIA-860 detailed data with previous form data (EIA-860A/860B)