UrbanFootprint

TECHNICAL DOCUMENTATION

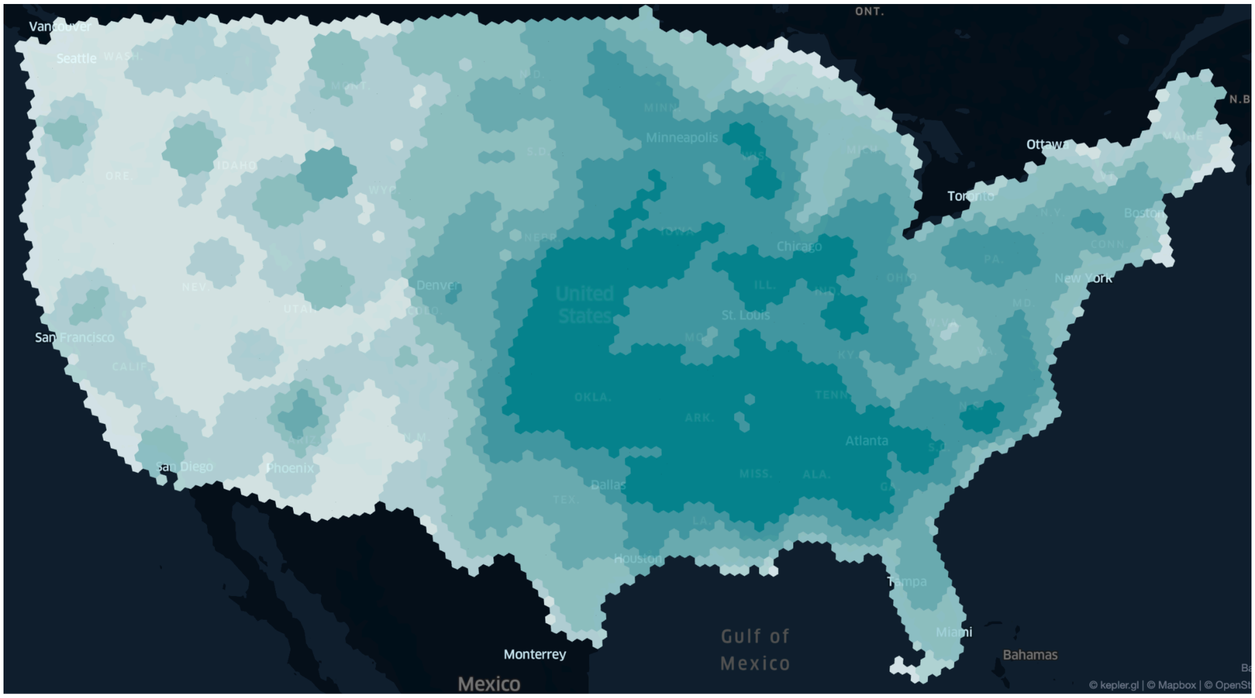
UrbanFootprint Tornado Winds Methodology

Last Updated: Dec 11, 2023

Overview

Tornadoes, characterized by their violent and unpredictable nature, are one of the most damaging natural phenomena affecting the United States. These intense vortexes are capable of causing widespread destruction within minutes, often with little warning. The unpredictability of tornado paths and the localized nature of their impact make them particularly difficult to model and anticipate, leading to significant challenges in emergency response and disaster preparedness.

The UrbanFootprint Tornado Winds Methodology aims to quantify the risk of tornado exposure. Our model is used to generate probabilities of exceedance of a given tornado magnitude for CONUS in an H3 zoom level 4 grid. At a high level, we use an improved version of the minimum assumption method using Bayes theorem and geospatial intersections to arrive at the probability that a tornado will impact a location. The figure below shows a sample wind risk map for EF0 tornadoes:



This dataset can be used as a general indicator of exposure to various tornado magnitudes.

Source Datasets

Storm Events Database

- **Source:** National Oceanic and Atmospheric Administration (NOAA), National Centers for Environmental Information. Obtained through BigQuery
- **Link:** Storm Events Database
- **Description:** The Storm Events Database is an integrated database of severe weather events across the United States from 1950 to this year, with information about a storm event's location, azimuth, distance, impact, and severity, including the cost of damages to property and crops. It contains data documenting:
    - The occurrence of storms and other significant weather phenomena having sufficient intensity to cause loss of life, injuries, significant property damage, and/or disruption to commerce
    - Rare, unusual, weather phenomena that generate media attention, such as snow flurries in South Florida or the San Diego coastal area
    - Other significant meteorological events, such as record maximum or minimum temperatures or precipitation that occur in connection with another event.

Specifications

Tornado Magnitudes

We present the probabilities that a point parcel will experience a tornado of a given Enhanced Fujita (EF) scale magnitude or higher.

Spatial Resolution

Our results are presented on a zoom level 4 H3 grid.

Spatial Extent

We present the risk related to non-cyclonic winds for the entirety of CONUS.

Methodology

Input dataset

The Storm Events Database provides, among other things, a comprehensive list of tornadoes that have affected the United States, including:

- Date and time
- Peak magnitude
- Width
- Path (or a single pair of latitude and longitude coordinates)

It is likely that tornadoes that occurred in the earliest days of the National Weather Service were underreported. Given this, we only consider tornadoes that occurred on or after January 1st, 1973. After this date, we have a good degree of confidence that all tornadoes were reported.

Minimum Assumption Method

Our approach builds upon the original work by Schaefer et al[1] in providing a “minimum assumption method” for estimating the probability that a tornado will affect a point parcel on a given grid cell. At its most basic form, the minimum assumption method estimates this probability as:

$$P_{i,j} = \frac{A_{i,j}}{A_i Y}$$

Here, $P_{i,j}$ represents the probability that a point parcel at a grid cell i will be hit by a tornado of magnitude j or above in a year, $A_{i,j}$ is the total area of that grid cell affected by that magnitude of tornadoes or above, A_i is the total area of the cell and Y is the number of years in the historical record.

The original work by Schaefer et al had three main gaps that we address in our methodology:

1. For any cell that had not experienced tornadoes of a given magnitude, the probability was estimated at zero. Given that tornadoes generally are rare events, and even more so for high magnitude tornadoes, it is more likely that the true probability is very, very low, but not zero.
2. Schaefer et al did not use geospatial operations to determine what portion of the tornado affected each grid cell. Rather, they (and the follow-up work by

---

[1] Schaefer, Joseph T., Donald L. Kelly, and Robert F. Abbey. "A minimum assumption tornado-hazard probability model." *Journal of Applied Meteorology and Climatology* 25.12 (1986): 1934-1945.

Standohar-Alfano et al[2]) simply counted each tornado if it originated at a grid cell. Given that tornadoes sometimes have very long paths, this approach leads to undercounting for tornadoes that pass through a cell and overcounting for tornadoes that originate at a cell.

3. Tornado winds do not have uniform magnitudes, but rather have more intense wind speeds toward the center that decrease as you move away from the center.

We tackle each of these three gaps in the following sections.

Bayesian Approach for Estimating the Probability of Exceedance

We define $P(T_i|L)$ as the probability that a point parcel is hit by a tornado of magnitude i given that you are at a location L, and $P(T|L)$ as the probability that a point parcel is hit by a tornado *of any magnitude* given that you are at a location L. From Bayes theorem, the following relationship holds:

$$P(T_i|L) = \frac{P(T_i|L,T)P(T|L)}{P(T|L,T_i)}$$

Clearly, if you're hit by a tornado of magnitude i, then you're hit by *a* tornado, and so $P(T|L,T_i) = 1$, which leaves us with $P(T_i|L) = P(T_i|L,T)P(T|L)$. We can estimate $P(T|L)$ in a straightforward manner by taking the area of all tornadoes that have hit the cell and dividing that by the cell area and the time period (essentially applying the minimum assumption method without taking into account the magnitude). For $P(T_i|L,T)$, we assume that, given that a point parcel is hit by a tornado, the probability that it is hit by a tornado of magnitude i is conditionally independent of your location. This would mean that $P(T_i|L,T) = P(T_i|T)$. Using this assumption, one way of estimating $P(T_i|T)$ is to take the total area affected by tornadoes in CONUS (contiguous United States), and divide that by the total area affected by tornadoes of magnitude i or higher.

This should have the desirable effect of making the probabilities of exceedance of all magnitudes for any location that has experienced any tornado non-zero (but still very low). The only cells that will have 0 probability of exceedance are the ones that have had no tornadoes at all in the past seventy years.

Of course, this is no proof that tornadoes actually do follow this conditional independence assumption. It is likely that there are meteorological, topographical, or other factors that differentiate the probabilities of tornadoes of some magnitudes instead of others. However, given that this approach does not have a large effect on the lower magnitude

---

[2] Standohar-Alfano, Christine D., and John W. van de Lindt. "Empirically based probabilistic tornado hazard analysis of the United States using 1973–2011 data." *Natural Hazards Review* 16.1 (2015): 04014013.

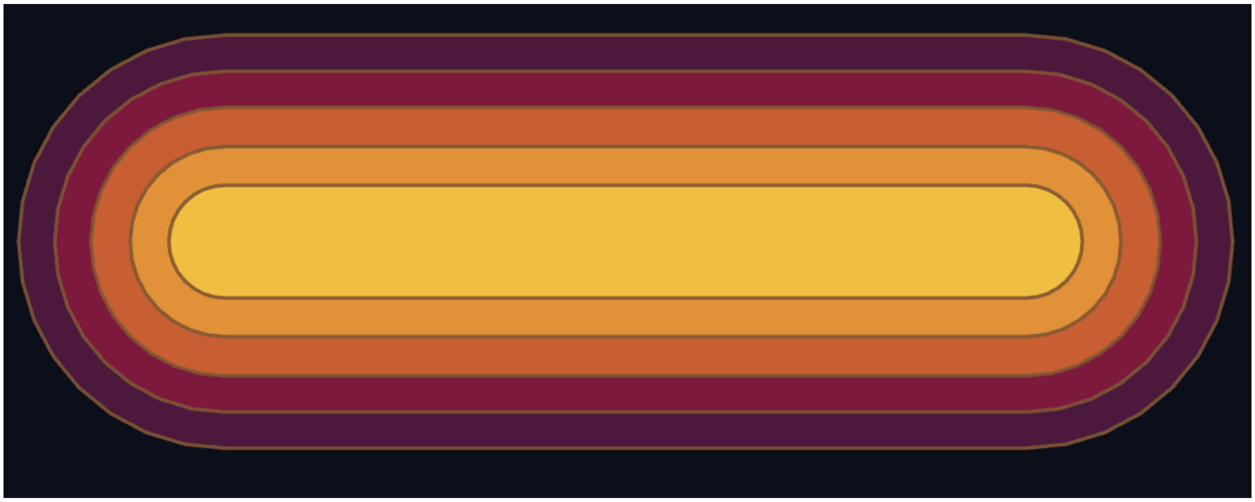
tornadoes (which are the ones that we have most certainty of), giving a non-zero estimate for the probability of higher magnitude tornadoes is solely an improvement.

Using Geospatial Operations for Estimating the Affected Area

We use the geography functionality built into BigQuery to better estimate the affected area within a grid cell. For every tornado, we buffer the path by half of the width, and intersect those geometries with our H3 zoom level 4 spatial index. Using this approach, only the portions of each tornado that are truly in each H3 cell are counted towards the probability of exceedance.

Area Corrections for Tornado Magnitudes

We follow the approach proposed by Standohar-Alfano et al to adjust the tornado geometries based on a wind profile that decreases as you move away from the center. An illustrative example of what this looks like for an EF4 tornado is shown below:



Additional Implementation Details

We applied the improved minimum assumption method on a zoom level 4 H3 spatial index. We follow a similar approach to Schaefer et al in using a sliding window for the purposes of smoothing. Specifically, we use a k-ring size of 2, where for each cell we consider not just the tornado geometries within the cell, but also within all of its neighbors that are a distance of at most 2 cells away. In terms of the zoom level and size of the k-ring, these values are consistent with those used by Schaefer et al.