

Tropical Storm Paper

Brooke Anderson & Matthew Hughes

May 25, 2020

Statement of key issue

Extensive physical exposure data is often available for tropical cyclones as they near and cross communities in the United States. This data can come both from established monitoring networks, like [NOAA network name?], but also may result from data collection efforts during or after the storm by atmospheric scientists and engineers seeking to characterize a storm. Researchers studying the human impacts of these storms, including epidemiologists, economists, and social scientists are interested in this data as well, but it is often unavailable in a way that is accessible for them to use. Resolving physical exposure and human impact datasets is challenging because the human impact measurements and physical exposure data do not have congruent resolutions.

In this paper, we will provide examples of how tropical cyclone exposure data can be misaligned, or require aggregation, both spatially and temporally to allow them to be integrated with data characterizing human impacts of the storm. Further, we will go through several common temporal and spatial resolutions in which outcome data from past studies of the human impacts of tropical cyclones in the United States have been available. We will discuss how differences in resolution between exposure and outcome datasets can create challenges in measuring and inferring the association between tropical cyclone exposure and human impacts, including potential biases and reduced precision in estimates of these associations. We will finish by discussing tools and challenges in integrating exposure and outcome datasets in the face of these differences in resolution in the original data and paths forward in resolving these challenges.

Implications of not improving this integration

Differences in resolution between exposure and outcome datasets create challenges in measuring and inferring the association between tropical cyclone exposure and human impacts. When researchers study the economic, social, and health impacts that tropical storms and hurricanes have in locales like the Gulf Coast of the United States, it is important to select an appropriate spatial and temporal scale to adequately classify exposure. Mis-matches in the spatial and temporal scale of exposure data versus outcome data creates challenges when measuring and inferring associations between tropical cyclone exposures and human impacts. This is a problem because it gives an inaccurate picture of how communities and individuals' health are impacted by these storms. Two potential implications are that this mis-match can introducing bias in estimated associations and that it can reduce precision in estimates of those associations.

Misclassification error / measurement error. One pathway for problems is through misclassification / measurement error bias. Misclassification error occurs when exposure and outcome variables are measured in categories and the wrong category is assigned to a particular case/observation - for example when a case that is exposed is incorrectly categorized as unexposed. For researchers who study tropical storms, it is common to categorize exposures by declaring a county or parish as exposed or unexposed based on whether the storm track went through the county. Failure to classify exposure accurately allows misclassification bias to move the results of the study further from the true parameter (for example, classifying certain observations as exposed to a storm when they really were not, or vice-versa). Measurement error occurs when the variables being measured are continuous, such as the amount of precipitation or the wind speed that was measured

during a tropical cyclone.

Non-differential misclassification error. Non-differential misclassification refers to misclassification of either the exposure or the outcome, that is unrelated to the other (Aschengrau and Seage 2013). Environmental epidemiology studies are often prone to non-differential misclassification error, because the methods of assessing exposure are not always congruent with the way that researchers conduct human impact studies. The effect of misclassifying exposures will often, though not always, bias the results of outcome towards the null (Armstrong 1998). In effect, this will weaken or obscure any associations that are present that the researcher may hope to observe in the data (Armstrong 1998). **[Add an example from tropical cyclone impacts studies of a case where you might get non-differential misclassification error.]**

Differential misclassification error. Differential misclassification error occurs when the misclassification of the outcome is related to the misclassification of the exposure or vice versa (Aschengrau and Seage 2013). While non-differential misclassification often (though not always) has the effect of moving the observed association or parameter towards the null, differential misclassification can move the observation in either direction. Differential misclassification in tropical cyclone impact studies occurs often in self reported data as in (Lieberman-Cribbin et al. 2017) where study subjects were asked to report their own flooding exposure and their mental health symptoms of depression, anxiety, and PTSD. It is reasonable to believe that self perceived exposure to hurricane related flooding would not be independent from perceived negative mental health symptoms and thus potentially contribute to differential misclassification error in this situation. **[Define differential misclassification error.] [Add an example from tropical cyclone impacts studies of a case where you might get non-differential misclassification error.]** Test for push

Dichotomizing continuous exposure measurements. Sometimes, researchers use an agreed upon threshold to split a continuous metric into a binary classification (exposed or unexposed). For example, a county may be classified as exposed or unexposed based on local winds exceeding a threshold (e.g. gale-force winds or higher). S.C. Grabich et al. 2016 classified hurricane exposure in a Florida county using maximum wind speed. Maximum wind speed is a continuous variable, but the study used binary categorizations to divide it into tropical wind speeds, classified as greater than 39 miles per hour, and hurricane wind speeds, classified as greater than 74 miles per hour. Florida counties experiencing maximum wind speeds below 39 miles per hour were considered unexposed.

Researchers typically categorize or dichotomize continuous variables in several situations for several reasons. They do this typically because it simplifies the data and allows for easier analysis and interpretation (Naggara et al. 2011). Additionally, it is very common in clinical settings to categorize continuous variables, for example hypertensive or not hypertensive, overweight or not overweight, dead or alive, etc. (Van Walraven and Hart 2008).

Despite several advantages to dichotomizing continuous variables that we just discussed, the general consensus in epidemiology is not to do it. Statistical power is lost because so much information is lost when categorization occurs (Van Walraven and Hart 2008). This makes sense when you consider that continuous variables allow you to observe nuance in the data and perceive a dose response relationship between the predictor and response variables, should one exist. This effect is masked when researchers categorize data, and even more so when a smaller number of categorical variables are used (for example dichotomization itself at 2). Generally, if you are going to categorize continuous data, it is better to use 3 or more categories rather than just two. An example of a paper that used three different bins was (Kinney et al. 2008), which explored the risk of autism after a pregnancy that included exposure to a tropical storm in the state of Louisiana. The study authors classified tropical storm exposure as severe, intermediate, and low exposure, and these exposure classifications were determined based on whether a mother lived in a Louisiana parish that had both of the exposure factors of interest: storm intensity and storm vulnerability. Storm vulnerability in this case was based on another dichotomy: whether or not the storm center passed through the parish of interest. Storm vulnerability was a measure of how vulnerable the inhabitants of the parish were to the effects of a storm (higher socioeconomic neighborhoods and parishes have more resources to withstand and recover from a tropical storm for example).

Another obvious problem with categorizing continuous data is that the cutoff points are often arbitrary. In the case of dichotomization, the median is often used, but there is typically no reason to assume that the median

is a reasonable cutoff point. Because different samples will have different medians, this automatically makes many categorical bins difficult to compare across studies (Altman and Royston 2006). Further, choosing optimal cutoff points that give the smallest p-values can lead to spurious results (Altman and Royston 2006).

Not surprisingly, dichotomizing continuous variables can bias results. A study by Selvin showed that the odds ratios can be significantly different depending on the chosen cutoff that is implemented in a study (Van Walraven and Hart 2008). Categorical variables can also put otherwise similar observations into separate bins if they are close but on opposite sides of the cutoff (Altman and Royston 2006). Choosing a median as a cutoff is intended to delineate bins, but if the bins are a “high” and “low” group, two individual observations that may only be a fraction different, but on either sides of the mean, will be classified as high and low respectively, and give the false impression that they are significantly different.

While dichotomizing continuous variables is something that can be done for either the exposure or the outcome of interest in a study, for our purposes we are primarily interested in continuous *exposures*. This means that we are primarily interested in the effects of dichotomizing variables such as wind speed, rainfall, temperature, distance from storm center, and distance from coastline, among other factors. Many epidemiology studies will dichotomize continuous outcome variables such as blood pressure, body weight (BMI), and length of pregnancy in order to gauge medical concern and priorities, but because we are concerned with creating a data framework that makes storm exposure data accessible for epidemiologists, exposure scientists, economists, and other scientists to use, we have a priority to look at exposure variables.

Scales for Categorizing Wind Speeds There are several methods in existence for categorizing wind speed, one of the most frequently used variables for estimating exposure to hurricanes and tropical storms. The first is the Saffir-Simpson scale, which uses five different bins to classify varying levels of wind speed and determine the severity of a storm. The first level, Category 1 is designated for hurricanes and tropical storms with maximum wind speeds of between 64 - 82 knots and is generally considered dangerous to people, livestock, and pets from the hazard of flying and falling debris (Taylor et al. 2010). On the higher end of the scale, Category 5 designates hurricanes with maximum wind speeds above 137 knots and is considered to have catastrophic effect on damage and a high probability of injury or death to people, livestock, and pets even if they are sheltering indoors (Taylor et al. 2010).

Forecasters classify hurricanes into categories on the Saffir-Simpson scale based on maximum sustained surface wind speed. This is defined as the peak one minute wind speed at a height of 10 feet over an unobstructed exposure (Taylor et al. 2010). An important limitation of the Saffir-Simpson scale is that it doesn't account for other hurricane-related impact variables such as storm surges, flooding, and tornadoes (Taylor et al. 2010).

Another scale used to categorize wind speed is the Beaufort scale, created by Admiral Sir Francis Beaufort, used to classify wind speeds both over land and sea. While the Saffir-Simpson scale is only designated for wind speeds that are already at hurricane levels (greater than 64 knots), the Beaufort scale considers the wind speeds below this. The scale ranges from Force 0 (0-1 knots and calm) to Force 12 (64 to 71 knots and hurricane). Other interesting parts of the scale include Force 3 (4-6 knots) which is a gentle breeze, and Force 8 (34-40 knots) which is considered a gale.

Categorizing wind speeds presents researchers with some of the same problems mentioned above that happen when dealing with continuous data, but both scales are based off associations between winds at certain speeds and observed damage and health impacts to communities exposed to these wind speeds.

Aggregate Hurricane Exposure Metrics

Another method of assessing damage and impact of tropical storms and hurricanes is through a single aggregate exposure metric. While aggregate values often represent the mean of all the values recorded, weather data is typically assessed by the maximum value. This could be something like the maximum wind speed reached in a particular county or parish, or the total monetary cost in damage due to flooding in a metropolitan statistical area. The Saffir-Simpson scale is an example of how entire storms are often classified by their maximum wind speed.

Although using a single exposure value can simplify analysis and interpretation, particularly over an extended temporal scale, there are some obvious drawbacks to relying on one single aggregate value. For example, the

Saffir Simpson categories typically correspond only to the geographic point location where the maximum wind speed was observed (Taylor et al. 2010). Hurricane Wilma in 2005 for example, was a Category 3 hurricane when it made landfall on the southwest coast of Florida, but it created Category 1 and Category 2 conditions for the more populous Miami-Dade, Broward, and Palm Beach counties when it finally reached them (Taylor et al. 2010).

Single exposure metrics are often used after a storm event has happened. They are very common in assessing ecological damage after a large hurricane.

Ecological Bias/The Ecological Fallacy Because studying tropical storm and hurricane exposures requires us to look at different spatial scales, we run the risk of encountering the ecological bias when looking at larger spatial aggregations. Ecological bias occurs whenever the aggregate association between an exposure and an outcome does not properly reflect the association on the individual level (Greenland and Morgenstern 1989). Ecological studies themselves don't look at individuals, but rather at an aggregate value, usually within a defined geographic region. Looking at national levels of obesity, cancer, or life expectancy, and comparing countries with respect to these outcomes and some exposure is an example of what ecological studies aim to achieve.

An example of a study that could be prone to this kind of bias is (Kinney et al. 2008) where Louisiana parishes were considered vulnerable to hurricane exposure based on whether or not the storm center passed through that parish. It is possible that the cases considered exposed based on living in these parishes were not in fact exposed since the storm may have passed through only a certain part of the parish. Never the less, all cases in a parish are considered exposed or unexposed in the aggregate.

Common resolutions for human impacts outcome data

We conducted a literature review to investigate the different spatial and temporal scales most commonly used by researchers studying tropical storm and hurricane impacts on health, ecology, and economic systems in affected areas. Spatially, the most common units tended to be the county and state level, while the temporal scale ranged from days to weeks to cumulative measures across storm events.

Spatial Scales

Data measuring human impacts often represents events that happened to individuals, and often happened at a specific geographical location. In human impacts studies, such outcomes are often measured and recorded at the individual level, with a specific geographic point location. However, exposure data are often represented as aggregations, based on geopolitical boundaries. In this section, we describe some common spatial resolutions for outcome data from previous research on the human impacts of tropical cyclones in the United States.

Point Location

[BA: Let's think some about the order we want for these sections. We're making several good points / analysis here. First, we're defining what we mean by the resolution ("point location" here). We probably want to start with that. Then we have some examples for studies that have had outcome data at this resolution. Maybe that could go next, to help illustrate the definition we've given. We've got some information on *how* the data at this scale was collected (e.g., geocoding from addresses reported from the study subjects), which I think is really interesting. Finally, we've got some text that talks about how data at this resolution could be integrated with some main formats of exposure data. We might want to end with that (or maybe even, as we work on this draft, that might go into a different section of the paper).]

Point locations are the smallest resolution of spatial data used to assess the exposure to tropical storms and hurricanes, as they represent the specific location of individual, non-aggregated observations on the outcome of interest. In many cases, researchers collect information on the study subject's residential address through some sort of a survey to assess point location (Lieberman-Cribbin et al. 2017), (Jaycox et al. 2010), (Bayleyegn et al. 2006). These surveys are often designed to assess psychological needs of hurricane survivors, as well as medical, financial, and nutritional needs. For example in (??), New York City residents provided their

address in a self reported manner to look at associations between mental health outcomes and flooding data. This residential address served as a point location that could be mapped and was compared to flooding data maps created by FEMA. In other cases, a GPS device is used to record coordinates that mark a specific point locations. An example is (Hagy, Lehrter, and Murrell 2006), where specific point locations were used to take water samples were taken to measure parameters of water quality such as salinity, temperature, dissolved oxygen, and turbidity compared before and after Hurricane Ivan in Pensacola Bay, Florida. This is a common practice in ecological research because point locations distributed across a landscape can be used to observe patterns taking geography into account. Point locations are also advantageous when using satellite images in conjunction with analysis of hurricane impact as illustrated in (Bianchette et al. 2009), where Landsat 5 images were used to compare vegetation damage, by looking at specific trees at different elevations to assess the ecological impact of Hurricane Ivan.

The obvious advantage of a point location is that when mapped, it can be overlayed with physical exposure data on a storm or storms to gage a very accurate picture of exposure, taking full advantage of high resolution in the exposure data. Since storm tracks are often spatially represented by the path of the storm's center, having point locations for the exposed units of interest allows researchers to more accurately measure how close each observation was to the storm's central track, and make further conclusions on this. Similarly, point locations can be integrated in a straightforward way with gridded exposure data, as might result from re-analysis datasets or ... [check with James Done about this], as each point location can be assigned the exposure level of the closest gridded measurement.

Zip Code/County/Parish

While point locations are very useful, many of the papers cited used larger geographic areas to denote spatial exposure to storms. Zip codes (Bevilacqua et al. 2020), (Lane et al. 2013), are often used to aggregate groups of people living in a given area. Counties are at a higher aggregation level than zip codes (Kinney et al. 2008), (S. C. Grabich et al. 2016), (S. Grabich et al. 2016), (Schwartz et al. 2018), (Harville et al. 2010). Often these levels seem to be used when a specific metropolitan area is being looked at, such as New York City after Hurricane Sandy (Lane et al. 2013), and Houston after Hurricane Harvey (Schwartz et al. 2018). The county level is a convenient method to use the storm path of the hurricane to quickly categorize exposed areas as in (S. Grabich et al. 2016).

There are several disadvantages and pitfalls to using this spatial level. For one, not all counties and zip codes (which are called parishes in Louisiana) are the same size or have the same population, so they may not be immediately comparable. Using the county/parish or zip code makes it easier for researchers to misclassify exposure. There are many ways that this can occur in a study on tropical storms; one common example is that counties selected as exposed are those that had the center of the storm pass through their county's physical boundaries. However it is very possible that some individuals lived in a county classified as exposed based on this criteria, but were in a region of the county far enough away from the storm center that they were not severely impacted. These individuals would be classified as exposed when they really were not and it could bias an apparent association towards the null. Alternatively, individuals who lived in a unexposed county, but were near the border of an exposed county could be incorrectly categorized as being unexposed even if they actually experienced many of the effects of the storm.

State/Metropolitan Region

Many studies used the spatial level of entire states or specific metropolitan areas to gather information on those who were exposed. (Harville et al. 2010) is an interesting paper because it looks at the state level as well as the regional and parish level. In this paper researchers observed birth outcomes in response to Hurricane Katrina in the state of Louisiana as a whole, the New Orleans metropolitan area, and Orleans parish, which is the heart of New Orleans. Looking at these three levels is a way to compare different incident rates and other measures of associations across different spatial scales.

The state or national level is the spatial level of an ecological study and can be useful to compare the emergency preparedness and policies of different states. The potential for the ecological bias is of course

present when looking at this spatial scale however, which occurs when the outcomes on the population level (typically an average), do not represent the individual outcomes very well.

Temporal Scales

Week

Week is a very common unit of time used to ascertain exposure, particularly for studies that are concerned with birth outcomes and gestation during hurricane exposure (Kinney et al. 2008), (S. C. Grabich et al. 2016), (S. Grabich et al. 2016). When the week of gestation is known, the timing that the hurricane makes landfall, or has its storm center pass through a county can be matched up to this week of gestation to identify possible “critical periods” of exposure during development.

Cumulative Measures of Time

Many of the studies looked at didn’t assess exposure at the moment of the storm, but rather after it had done damage.

Terms

These are terms we’re using right now that we might want to iterate on, in conjunction with our colleagues on the project, to make sure we have terms that are precise and consistent across the document:

- **physical exposure data:** By this, we mean things that are measured about the storm like wind speed, rainfall, measures of flooding, and other things that might be considered more in the realm of what an atmospheric scientist or engineer might measure about the storm. We’re contrasting this with data that for human impacts studies on outcomes among humans (e.g., pregnancy outcomes, economic outcomes like unemployment)
- **resolution:** We’re using this right now to talk about spatial and temporal levels of aggregation. Sometimes, we’re using “scales” instead, I think.

References

- Altman, Douglas G, and Patrick Royston. 2006. “The Cost of Dichotomising Continuous Variables.” *Bmj* 332 (7549): 1080.
- Aschengrau, Ann, and George R Seage. 2013. *Essentials of Epidemiology in Public Health*. Jones & Bartlett Publishers.
- Bayleyegn, Tesfaye, Amy Wolkin, Kathleen Oberst, Stacy Young, Carlos Sanchez, Annette Phelps, Joann Schulte, Carol Rubin, and Dahna Batts. 2006. “Rapid Assessment of the Needs and Health Status in Santa Rosa and Escambia Counties, Florida, After Hurricane Ivan, September 2004.” *Disaster Management & Response* 4 (1): 12–18.
- Bevilacqua, Kristin, Rehana Rasul, Samantha Schneider, Maria Guzman, Vishnu Nepal, Deborah Banerjee, Joann Schulte, and Rebecca M Schwartz. 2020. “Understanding Associations Between Hurricane Harvey Exposure and Mental Health Symptoms Among Greater Houston-Area Residents.” *Disaster Medicine and Public Health Preparedness*, 1–8.
- Bianchette, TA, K-B Liu, NS-N Lam, and LM Kiage. 2009. “Ecological Impacts of Hurricane Ivan on the Gulf Coast of Alabama: A Remote Sensing Study.” *Journal of Coastal Research*, 1622–6.
- Grabich, SC, J Horney, C Konrad, and DT Lobdell. 2016. “Measuring the Storm: Methods of Quantifying Hurricane Exposure with Pregnancy Outcomes.” *Natural Hazards Review* 17 (1): 06015002.

- Grabich, Shannon C, Whitney R Robinson, Stephanie M Engel, Charles E Konrad, David B Richardson, and Jennifer A Horney. 2016. "Hurricane Charley Exposure and Hazard of Preterm Delivery, Florida 2004." *Maternal and Child Health Journal* 20 (12): 2474–82.
- Greenland, Sander, and Hal Morgenstern. 1989. "Ecological Bias, Confounding, and Effect Modification." *International Journal of Epidemiology* 18 (1): 269–74.
- Hagy, James D, John C Lehrter, and Michael C Murrell. 2006. "Effects of Hurricane Ivan on Water Quality in Pensacola Bay, Florida." *Estuaries and Coasts* 29 (6): 919–25.
- Harville, Emily W, Tri Tran, Xu Xiong, and Pierre Buekens. 2010. "Population Changes, Racial/Ethnic Disparities, and Birth Outcomes in Louisiana After Hurricane Katrina." *Disaster Medicine and Public Health Preparedness* 4 (S1): S39–S45.
- Jaycox, Lisa H, Judith A Cohen, Anthony P Mannarino, Douglas W Walker, Audra K Langley, Kate L Gegenheimer, Molly Scott, and Matthias Schonlau. 2010. "Children's Mental Health Care Following Hurricane Katrina: A Field Trial of Trauma-Focused Psychotherapies." *Journal of Traumatic Stress: Official Publication of the International Society for Traumatic Stress Studies* 23 (2): 223–31.
- Kinney, Dennis K, Andrea M Miller, David J Crowley, Emerald Huang, and Erika Gerber. 2008. "Autism Prevalence Following Prenatal Exposure to Hurricanes and Tropical Storms in Louisiana." *Journal of Autism and Developmental Disorders* 38 (3): 481–88.
- Lane, Kathryn, Kizzy Charles-Guzman, Katherine Wheeler, Zaynah Abid, Nathan Graber, and Thomas Matte. 2013. "Health Effects of Coastal Storms and Flooding in Urban Areas: A Review and Vulnerability Assessment." *Journal of Environmental and Public Health* 2013.
- Lieberman-Cribbin, Wil, Bian Liu, Samantha Schneider, Rebecca Schwartz, and Emanuela Taioli. 2017. "Self-Reported and Fema Flood Exposure Assessment After Hurricane Sandy: Association with Mental Health Outcomes." *PLoS One* 12 (1): e0170965.
- Naggara, O, J Raymond, F Guilbert, D Roy, A Weill, and Douglas G Altman. 2011. "Analysis by Categorizing or Dichotomizing Continuous Variables Is Inadvisable: An Example from the Natural History of Unruptured Aneurysms." *American Journal of Neuroradiology* 32 (3): 437–40.
- Schwartz, Rebecca M, Stephanie Tuminello, Samantha M Kerath, Janelle Rios, Wil Lieberman-Cribbin, and Emanuela Taioli. 2018. "Preliminary Assessment of Hurricane Harvey Exposures and Mental Health Impact." *International Journal of Environmental Research and Public Health* 15 (5): 974.
- Taylor, Harvey Thurm, Bill Ward, Mark Willis, and Walt Zaleski. 2010. "The Saffir-Simpson Hurricane Wind Scale." *Atmospheric Administration: Washington, DC, USA*.
- Van Walraven, Carl, and Robert G Hart. 2008. "Leave'em Alone-Why Continuous Variables Should Be Analyzed as Such." *Neuroepidemiology* 30 (3): 138.