**Title:** Assessing United States county-level exposure for research on tropical cyclones and human health

**Authors:** G. Brooke Anderson1,\*, Joshua Ferreri1,2, Mohammad Al-Hamdan3, William Crosson3, Andrea Schumacher4, Seth Guikema5, Steven Quiring6, Dirk Eddelbuettel7, Meilin Yan1, and Roger D. Peng8

**Affiliations:**

1Department of Environmental & Radiological Health Sciences, Colorado State University, Fort Collins, CO, 80523

2University of Colorado Anschutz Medical Campus, School of Medicine, Aurora, CO, 80045

3Universities Space Research Association, NASA Marshall Space Flight Center, Huntsville, AL, 35805

4Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, 80523

5Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, 48109

6Department of Geography, Ohio State University, Columbus, OH, 43210

7Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, IL, 61820

8Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, 21205

**Contact information for corresponding author:**

G. Brooke Anderson

Department of Environmental & Radiological Health Sciences

Colorado State University

1681 Campus Delivery  
Fort Collins, Colorado 80523

[brooke.anderson@colostate.edu](mailto:brooke.anderson@colostate.edu)

**Acknowledgements:** This work was supported in part by grants from the National Institute of Environmental Health Sciences (R00ES022631), the National Science Foundation (1331399), the Department of Energy (Grant No. DE-FG02-08ER64644), and a National Aeronautics and Space Administration Applied Sciences Program/Public Health Program Grant (NNX09AV81G). Rainfall data are based on data acquired as part of the mission of the National Aeronautics and Space Administration’s Earth Science Division and archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC).

**Declaration of competing financial interests:** The authors declare they have no actual or potential competing financial interests.

**ISC).Abstract**

**Background:** Tropical cyclone epidemiology can be advanced through county-level exposure assessment methods that are comprehensive and consistent across space and time, facilitating multi-year, multi-storm studies. Further, an understanding of patterns in and between exposure metrics that are based on specific hazards of the storm can help in designing tropical cyclone epidemiological studies.

**Objectives:** (1) Provide an open-source dataset for tropical cyclone exposure assessment for epidemiological research and (2) investigate patterns and agreement between county-level assessments of tropical cyclone exposure based on different storm hazards.

**Methods:** We created an open-source dataset with county-level data on exposure to four tropical cyclone hazards: peak sustained wind, rainfall, flooding, and tornadoes. The data cover all eastern United States (US) counties for all land-falling or near-land Atlantic basin storms, covering 1996–2011 for all metrics and up to 1988–2015 for specific metrics. We validated measurements against other data sources and investigated patterns and agreement among binary exposure classifications based on these metrics, as well as compared them to use of distance from the storm’s track as a proxy for exposure.

**Results:** We describe how we developed this dataset and explore how its measurements compare to other data that could be used to characterize tropical cyclone hazards for epidemiological research. Measurements in the dataset created here typically agreed with data from other sources, and we present and discuss areas of disagreement and other caveats for the data. Tropical cyclones typically brought different hazards to different counties, and so agreement was usually low between distance-based tropical cyclone exposure assessment and each hazard-specific metric, as well as between pairs of hazard-specific metrics.

**Discussion:** Our results provide a multi-hazard dataset that can be leveraged for epidemiologic research on tropical cyclones, as well as insights that can inform the design of tropical cyclone epidemiological studies.

**Introduction**

Hurricanes and other tropical cyclones can severely impact human health in United States (US) communities, as has been characterized in several long-term national studies (Rappaport 2000; Rappaport 2014; Rappaport and Blanchard 2016; Czajkowski and Kennedy 2010; Czajkowski et al. 2011; Moore and Dixon 2012). These studies characterize tropical cyclone health impacts using fatality data aggregated in large part from disaster-related mortality ascertainment and surveillance conducted by agencies like the US Centers for Disease Control and Prevention, the US National Weather Service, and local vital statistics departments. These fatality data are based on case-by-case ascertainment: identifying and characterizing deaths for which there is a clear link with a disaster, through death certificate coding or other indicators (Rocha et al. 2017).

For several types of climate-related disasters—including heat waves, floods, and wildfires—such research has been richly supplemented by studies that estimate community-wide excess mortality and morbidity associated with the disaster (e.g., Anderson and Bell 2010; Son et al. 2012; Haikerwal et al. 2015; Liu et al. 2017; Milojevic et al. 2017). These studies, which often use time series or case-crossover study designs, typically compare the rate of a health outcome in a community during a disaster to rates during comparable non-disaster periods. While such studies cannot attribute specific cases (e.g., a specific death) to a disaster, they can quantify the community-wide change in health risk during disasters and capture impacts that might be missed or underestimated with traditional disaster-related mortality ascertainment and surveillance. In many cases, these studies analyze multi-year, multi-community data, allowing them to estimate average associations over many disasters and to explore how a disaster’s characteristics, or the characteristics of the community it hits, modify associated health risks (e.g., Anderson and Bell 2010; Son et al. 2012; Liu et al. 2017). Some studies have begun to use this approach to study the health impacts of tropical cyclones, including several studies of Hurricane Maria (e.g., Santos-Lozada and Howard 2018; Santos-Burgoa et al. 2018), Hurricane Sandy (e.g., Kim et al. 2016; Mongin, Baron et al. 2017; Swerdel et al. 2014), and the 2004 hurricane season in Florida (McKinney, Houser, and Meyer-Arendt 2011). However, to expand this approach to longer time periods and large sets of communities, researchers must be able to assess exposure to tropical cyclones consistently and comparably across storms, years, and communities.

The National Hurricane Center (NHC) publishes a “Best Tracks” dataset that is considered the gold standard for Atlantic-basin tropical cyclone tracks. It records the central position of a tropical cyclone every six hours, as well as the storm’s minimum central pressure and maximum sustained surface wind. These data are openly available through the revised Atlantic hurricane database (HURDAT2), a post-storm assessment that incorporates data from several sources, including satellite data and, when available, aircraft reconnaissance data (Landsea and Franklin 2013; Jarvinen and Caso 1978). With these data, it is straightforward to measure whether a community was in a storm’s direct path or within a certain distance of its central track; some previous epidemiological studies have done this to assess exposure to a tropical storm (e.g., Currie and Rossin-Slater 2013; Kinney et al. 2008; Caillouët et al. 2008).

However, when epidemiological studies assess exposure to a tropical cyclone using, as a proxy, how close the storm came to a community, they may misclassify exposure (Grabich et al. 2015a). While a number of tropical cyclone hazards are strongly associated with distance from the tropical cyclone’s center (e.g., wind and, at the coast, storm surge and waves (Rappaport 2000; Kruk et al. 2010)), other hazards like heavy rainfall, floods, and tornadoes can occur well away from the tropical cyclone’s central track (Rappaport 2000; Atallah et al. 2007; Moore and Dixon 2012). For example, fatal tropical cyclone tornadoes, which were linked to over 300 deaths in the US between 1995 and 2009, most often occur 200–500 km from the tropical cyclone’s center (Moore and Dixon 2012). Further, when studies use distance from the storm’s track to assess exposure, they often use the same buffer constraints on each side of the storm track (e.g., Kinney et al. 2008; Currie and Rossin-Slater 2013). However, the forces of a tropical cyclone tend to be distributed around its center non-symmetrically. Extreme winds are more common to the track’s right, where counter-clockwise cyclonic winds move in concert with the tropical cyclone’s forward motion (Halverson 2015), and almost all of the fatal tornadoes associated with US tropical cyclones between 1995 and 2009 occurred to the right of the tropical cyclone’s track (Moore and Dixon 2012). Rain, conversely, is often heaviest to the left of the track, especially when the tropical cyclone interacts with other weather systems (Atallah and Bosart 2003; Atallah et al. 2007; Zhu and Quiring 2013) or undergoes an extratropical transition (Elsberry 2002). The multi-hazard nature of tropical cyclones therefore makes it hard to assess exposure based on how close the storm’s central track came to the community. While other approaches have been developed to incorporate storm hazards, particularly wind, into exposure assessment (e.g., Grabich et al. 2015a; Zandbergen 2009; Czajkowski et al. 2011), there is not yet a standard approach, and when different studies use different datasets or storm hazards when assessing storm exposure, it becomes difficult to compare and aggregate findings.

To help epidemiological researchers assess exposure to tropical cyclones, we developed an opensource dataset (Anderson et al. 2017a), which we present here. These data cover all counties in states in the eastern half of the US over multiple years and include five metrics characterizing exposure to tropical cyclones (Table 1): (1) closest distance the storm’s central track came to the county’s center (a proxy measurement of storm exposure used in some previous studies); (2) peak sustained surface wind at the county’s center over the course of the storm; (3) cumulative rainfall in the county over the course of the storm; (4) flooding in the county concurrent with the storm; and (5) tornadoes in the county concurrent with the storm. We aggregated these data at the county level since data for epidemiological studies are often available at county-level aggregations (e.g., direct hurricane-related deaths (Czajkowski et al. 2011); birth outcomes (Grabich et al. 2015a; Grabich et al. 2015b); autism prevalence (Kinney et al. 2008)) and since decisions and policies to prepare for, and respond to, tropical cyclones are often undertaken at the county level (Zandbergen 2009; Rappaport 2000). We also wrote functions to explore and map these data and to link it with county-level health data (Anderson et al. 2017b).

Here, we describe how we developed this dataset and explore how its measurements compare to other data that could be used to characterize tropical cyclone hazards for epidemiological research. Further, we expand on previous research on methods to measure exposure to tropical cyclones for epidemiological research (Grabich et al. 2015a). The previous study investigated differences in which communities were assessed as exposed to tropical cyclones during the 2004 hurricane season in Florida, depending on whether exposure assessment was based on distance to the storm’s track or a metric incorporating storm-related winds within the county during the storm. They found important differences across methods of assessing storm exposure, concluding that a study may be prone to bias from exposure misclassification if distance to the storm track is used as a proxy measurement for exposure assessment (Grabich et al. 2015a). Here we expand to investigate this question across a large set of counties and storm seasons. Further, we investigate patterns in exposure classification based on other metrics of storm-related hazards—rainfall, flooding, and tornadoes—which are important for inland health impacts of tropical cyclones (Czajkowski et al. 2011; Moore and Dixon 2012). These results can help epidemiologists design studies and plan statistical analysis for multi-year, multi-community studies that estimate excess mortality and morbidity exposure to the hazards brought by tropical cyclones.

# Methods and Materials

We aggregated these data for all counties in the eastern half of the US (Figure 1). We included all tropical cyclones in HURDAT2 (Landsea and Franklin 2013) that came within 250 km of at least one eastern US county between 1988 and 2015. This ensured we were including all land-falling or near-land Atlantic basin tropical cyclones in this period (the tracks of the study storms are shown in Figure 1).

## Distance-based exposure metric

We first measured how close each storm’s central track came to each county. We used tracking data from HURDAT2, which records the storm center’s position at the synoptic times of 6:00 am, 12:00 pm, 6:00 pm, and 12:00 am Coordinated Universal Time (UTC), and interpolated this position to 15-min intervals using natural cubic splines (Anderson, Yan, Ferreri, Crosson, Al-Hamdan, Schumacher, and Eddelbuettel 2017). At each 15-min interval, we measured the distance between the storm’s center and each county’s population mean center (Bivand, Pebesma, and Gomez-Rubio 2013; US Census Bureau 2015). We took the minimum distance for each county as a measure of how close the storm came to the county. We also recorded the time the storm came closest to each county, which we later used to link observed data on precipitation, flooding, and tornadoes to the storm. To allow matching with data recorded in local time (e.g., health data), we converted these times from UTC to local time (Anderson and Guo 2016).

## Rain-based exposure metric

We based storm rainfall measurements on precipitation data files from the North American Land Data Assimilation System Phase 2 (NLDAS-2) re-analysis dataset, which is available for the continental US (Rui and Mocko 2014). We used data that were previously aggregated from this re-analysis dataset to the county level by two of the coauthors (Al-Hamdan and Crosson) for the Centers for Disease Control and Prevention’s Wide-ranging Online Data for Epidemiological Research (WONDER) database (US CDC 2016; Al-Hamdan, Crosson, Economou, Estes Jr, Estes, Hemmings, Kent, Puckett, Quattrochi, Rickman, et al. 2014). To create this aggregated data, these coauthors took hourly precipitation measurements in the NLDAS-2 precipitation files, which are given at 1/8◦ grid points across the continential US, and summed them by day for each grid point. This created a daily precipitation total for each grid point; these were then averaged for all grid points within a county’s 1990 US Census boundaries (Al-Hamdan, Crosson, Economou, Estes Jr, Estes, Hemmings, Kent, Puckett, Quattrochi, Rickman, et al. 2014; US CDC 2016). This process generated daily county-level precipitation estimates for each continental US county for 1988–2011 (US CDC 2016). We matched these county-level daily measurements by date with storm tracks, using the date when the storm was closest to each county. Given the location of storm-affected counties and the typical timing of tropical cyclones, these precipitation measures primarily represent rainfall, although occasionally they may represent snowfall or other types of precipitation.

In the open-source data (Anderson, Schumacher, Crosson, Al-Hamdan, Yan, Ferreri, Chen, Quiring, and Guikema 2017), we provide precipitation values at a daily resolution for each county for the period from five days before to three days after each storm’s closest approach to the county.

In the software that we published in association with this data (Anderson, Yan, Ferreri, Crosson, Al-Hamdan, Schumacher, and Eddelbuettel 2017), we provide functionality to aggregate these daily values to create cumulative precipitation estimates for custom windows around the date of the storm’s closest approach to the county. For example, a user could determine precipitation only on the day the storm was closest to each county, but could also determine the cumulative rainfall for a more extended period, like from two days before to two days after the storm’s closest approach. For the analysis of binary exposure classifications in this study, we calculated storm-associated rainfall as the sum of precipitation from two days before the storm’s closest approach to the county to one day after.

To validate these precipitation estimates, we compared them with ground-based observations in a subset of study counties. We selected nine sample counties geographically spread across storm-prone regions of the eastern US and for which precipitation data were available from multiple ground-based stations within the Global Historical Climatology Network throughout the study period (Menne, Durre, Vose, Gleason, and Houston 2012; Chamberlain 2017; Severson and Anderson 2016). These sample counties were: Miami-Dade, FL; Harris County, TX; Mobile County, AL; Orleans Parish, LA; Fulton County, GA; Charleston County, SC; Wake County, NC; Baltimore County, MD; and Philadelphia County, PA. We summed daily station-specific measurements from two days before to one day after each storm’s closest approach and then averaged these cumulative station-based precipitation totals for each county to create a county-wide estimate of cumulative storm-related precipitation based on ground-based monitoring. We measured the rank correlation (Spearman’s *ρ*; Spearman (1904)) between storm-specific cumulative precipitation estimates for the two data sources (NLDAS-2 re-analysis data versus ground-based monitoring) within each sample county.

## Wind-based exposure metric

We created a dataset of county-level peak sustained surface wind during each storm. To do this, we first modeled each storm’s wind field to each county’s population mean center (US Census Bureau 2015) at 15-minute intervals throughout the course of the storm. We used a double exponential wind model based on results from Willoughby (Willoughby, Darling, and Rahn 2006) to model 1-minute surface wind (10 meters above ground) at each county center, based on inputs of the storm’s forward speed, direction, and maximum wind speed (Anderson, Schumacher, Guikema, Quiring, and Ferreri 2017). Our model incorporated the asymmetry in wind speeds around the tropical cyclone center that results from the storm’s forward movement (Phadke, Martino, Cheung, and Houston 2003). From these 15-min interval estimates, we identified the peak sustained surface wind in each county over the course of each storm.

To validate these modeled county-level peak wind estimates, we compared them to the postseason wind radii values, which have been routinely published since 2004 (Knaff, Slocum, Musgrave, Sampson, and Strahl 2016). These wind radii estimate the maximum distance from the storm’s center that winds of a certain intensity extend. They give separate estimates for four quadrants of the storm to capture asymmetry in the storm’s wind field and are based on a post-season re-analysis that incorporates available data (e.g., satellite data, aircraft reconnaissance and ground-based data if available) (Knaff, Slocum, Musgrave, Sampson, and Strahl 2016). Wind radii are estimated for three thresholds of peak sustained surface wind: 64, 50, and 34 kt. They therefore allow for the classification of counties into four categories of peak sustained surface wind: *<*34 kt; 34–49.9 kt; 50–63.9 kt; ≥64 kt.

We interpolated the wind radii data to 15-minute increments and classified a county as exposed to winds in a given wind speed category if its population mean center was within 85% of the maximum radius for that wind speed in that quadrant of the storm. The 85% adjustment is based on previous research that found that this ratio helps capture average wind extents within a quadrant from these maximum wind radii (Knaff, Slocum, Musgrave, Sampson, and Strahl 2016). We compared these wind radii–based estimates of county-level peak sustained surface wind during a storm with the modeled wind estimates, comparing all study storms since 2004 for which at least one study county experienced peak sustained winds of ≥34 kt (based on the HURDAT2 wind radii). For each of these storms, we calculated the percent of study counties that were classified in the same wind category based on both data sources.

## Flood- and tornado-based exposure metrics

To identify flood- and tornado-based tropical cyclone exposures in US counties, we matched storm tracks with event listings from the National Oceanic and Atmospheric Administration (NOAA)’s Storm Event Database (NOAA NCEI 2017). While this database has recorded storm data since 1950, its coverage changed substantially in 1996 to cover more types of storm events (NOAA NCEI 2017). We therefore only considered these metrics of hurricane exposure for tropical cyclones in 1996 and later.

For each tropical cyclone, we identified all events with event types related to flooding (“Flood,” “Flash Flood,” “Coastal Flood”) or tornadoes (“Tornado”) with a start date within a five-day window centered on the date of the tropical cyclone’s closest approach to the county (Anderson, Schumacher, Crosson, Al-Hamdan, Yan, Ferreri, Chen, Quiring, and Guikema 2017). To exclude events that started near in time to the storm but far from the storm’s track, and so were likely unrelated to the storm, we excluded any events that occurred ≥500 km from the tropical cyclone’s track. “Flood,” “Flash Flood,” and “Tornado” events in this database were reported by county Federal Information Processing Standard (FIPS) code and so could be directly linked to counties. “Coastal Flood” events were reported by forecast zone; for these, the event was matched to the appropriate county if possible using regular expression matching of listed county names (Anderson and Chen 2017).

The tornado observations from this dataset form a traditional tornado event database for the US, and so we did not further validate the tornado event data. It is difficult to characterize flooding at the county level because flooding can be very localized and can be triggered by a variety of causes. To investigate the extent to which the NOAA flood event listings capture extremes that might be identified with other flooding data sources, we investigated a sample of study counties, comparing the flood event data during tropical storms with streamflow measurements at US Geological Survey county streamflow gages (USGS 2016; Lammers and Anderson 2017; Hirsch and De Cicco 2015). We considered nine study counties, selecting counties geographically spread through storm-prone areas of the eastern US and with multiple streamflow gages reporting data during events. The sample counties were: Baltimore County, MD; Bergen County, NJ; Escambia County, FL; Fairfield County, CT; Fulton County, GA; Harris County, TX; Mobile County, AL; Montgomery County, PL; and Wake County, NC. For each county, we first identified all streamflow gages in the county with complete data for Jan. 1 1996–Dec. 31, 2015. If a storm did not come within 500 km of a county, it was excluded from this analysis, but all other study storms were considered. For each storm and county, we summed streamflow measurements across all county gages to generate daily totals for the five-day window around the storm’s closest approach. We took the maximum of these daily streamflow totals as a measure of the county’s maximum daily streamflow during that storm. We also calculated the percent of streamflow gages in the county with a daily streamflow that exceeded a threshold of flooding (the streamgage’s median value for annual peak flow) on any day during the five-day window. We investigated how these measurements varied between storms with associated flood events listings versus storms without an event listing for the county, to explore if storms with flood event listings tended to be associated with higher streamflow at gages within the county.

## Binary storm exposure classifications

In our open-source data, we provide continuous measurements of some of the exposure metrics: closest distance of each storm to each county, peak sustained surface wind at the county center, and cumulative rainfall. However, epidemiologic studies of tropical cyclones often use a binary exposure classification (“exposed” versus “unexposed”) to assess storm-related health risks (e.g., Grabich, Robinson, Engel, Konrad, Konrad, and Horney (2015), McKinney, Houser, and Meyer-Arendt (2011), and Caillouët, Michaels, Xiong, Foppa, and Wesson (2008)), and so we explored patterns in storm exposure based on binary clasifications of these exposure metrics.

Two of the exposure metrics (flood- and tornado-based) were inherently binary in our data, since these metrics were based on whether an event was listed in the NOAA Storm Events database. For other exposure metrics, each county was classified as exposed to a tropical cyclone based on whether the exposure metric exceeded a certain threshold (Table 1). We picked reasonable thresholds, but others could be used with the open-source data and its associated software.

For the rainfall metric, a distance constraint was also necessary, to ensure that rainfall unrelated and far from the storm track was not misattributed to a storm. Through exploratory analysis, we set this distance metric at 500 km (i.e., for a county to be classified as exposed based on rainfall, the cumulative rainfall in the county had to be over the 75 mm threshold and the storm must have passed within 500 km of the county; Table 1). This distance constraint was typically large enough to capture storm-related rain. However, data users should note that in rare cases—for example, exceptionally large storms (e.g., Hurricane Ike in 2008) or storms for which storm tracking was stopped at extratropical transition (e.g., Tropical Storm Lee in 2011)—some storm-related rain exposures may be missed because of this distance constraint (Figure S1). This distance constraint can be customized using the software published in association with the open-source data (Anderson, Yan, Ferreri, Crosson, Al-Hamdan, Schumacher, and Eddelbuettel 2017).

We characterized patterns in county-level exposure in the eastern US for exposure assessment based on each of the measured tropical cyclone hazards (cumulative precipitation, peak sustained winds, flooding, and tornadoes). Depending on available exposure data, this assessment included some or all of the period from 1988 to 2015 (Table 1). For each binary metric of exposure to a tropical cyclone hazard, we first summed the total number of county-level exposures over available years and mapped patterns in these exposures.

Finally, we investigated how well exposure assessment agreed across these metrics. For each storm we investigated the degree to which the set of counties assessed as exposed based on one metric overlapped with the set assessed as exposed based on each other metric, including a metric that used distance of a county from the storm track as a proxy for exposure to storm hazards. We calculated the within-storm Jaccard index (*Js*) (Jaccard 1901; Jaccard 1908) between each pair of exposure metrics. This measures similarity between two metrics (*X*1,*s* and *X*2,*s*) for tropical cyclone *s* as the proportion of counties in which both of the metrics classify the county as exposed out of all counties classified as exposed by at least one of the metrics:

*X*1,*s* ∩*X*2,*s*

*Js* = (1)

*X*1,*s* ∪*X*2,*s*

This metric can range from 0, in the case of no overlap between the counties classified as exposed based on the two metrics, to 1, in the case that the two metrics classify exactly the same set of counties as exposed to the tropical cyclone. We calculated these values for all study storms that affected ≥100 study counties (based on at least one exposure metric) during the years when all exposure data were available (1996–2011).

# Results

## Exposure assessment data, data validation, and software

We created and published this tropical cyclone exposure data as an R package (Anderson, Schumacher, Crosson, Al-Hamdan, Yan, Ferreri, Chen, Quiring, and Guikema 2017). The package size exceeded the recommended maximum size for the Comprehensive R Archive Network (CRAN), the standard repository for publishing R packages. Therefore, we used the *drat* framework to set up our own repository to host the package (Anderson and Eddelbuettel 2017). These data include continuous county-level measurements for the closest distance of each storm, cumulative rainfall, and peak sustained surface wind. These continuous metrics can be used for classifying counties as exposed or unexposed—using thresholds selected by the user—or can be extracted directly as continuous metrics. The dataset also include binary data on flood and tornado events associated with each storm in each county. To accompany these data, we published an additional R package with software tools to explore and map the data and to integrate it with human health datasets (Anderson, Yan, Ferreri, Crosson, Al-Hamdan, Schumacher, and Eddelbuettel 2017).

We explored potential limitations in these data by comparing them with data from other available sources. For estimates of storm-associated rainfall, we compared estimates included in the opensource data with ground-based observations in nine sample counties (Figure 2). Within these counties, storm-related rainfall measurements were well-correlated between the two data sources, with rank correlations (bottom right of each graph in Figure 2) between 0.87 and 0.98. There was some evidence that the primary rainfall metric may tend to underestimate rainfall totals in storms with extremely high rainfall, based on a few heavy-rainfall storms in Harris County, TX, Mobile County, AL, Charleston County, SC, and Wake County, NC (Figure 2). However, even in these cases, it was rare for a storm to be classified differently (under the precipitation threshold of 75 mm we used for binary exposure classification for later analysis) based on the source of precipitation data—horizontal and vertical lines in each small plot in Figure 2 show the threshold of 75 mm, so storms in the lower left and upper right quadrants would be classified the same (“exposed” or “unexposed”) regardless of the precipitation data source, while storms in the upper left and lower right quadrants would be classified differently. Such cases were rare.

For peak sustained surface wind estimates, we found that the primary wind exposure metric in the open-source data generally agreed well with the wind radii reported in HURDAT2. For most storms, ≥95% of counties were assigned the same category of wind speed (*<*34 kt; 34–49.9 kt; 50– 63.9 kt; ≥64 kt) by both data sources (Figure 3). Disagreement was limited to a few storms (e.g., Hurricanes Sandy in 2012 and Ike in 2008). For these two storms, the modeled wind speed values used in the open-source data somewhat overestimated the extent of the storm’s most severe winds at landfall but then underestimated the extent of 34–49.9 kt storm winds further inland (Figure S2). For epidemiology researchers who would like to conduct sensitivity analysis between these two sets of wind data, we have included estimates from the HURDAT2 wind radii as a secondary measure of county-level peak sustained wind in the open-source dataset (Anderson, Schumacher, Crosson, Al-Hamdan, Yan, Ferreri, Chen, Quiring, and Guikema 2017).

For flood data, we compared the flood status values included in the open-source data to streamflow measurements at US Geological Survey gages within nine sample counties (Figure 4). Across the sampled counties, streamgage data generally indicated higher discharge during periods identified as flood events based on the NOAA Storm Events database, particularly in counties with more reporting streamgages (e.g., Harris County, TX, with 10 reporting stream gages; Figure 4). There were some cases, however, where the two flooding data sources were somewhat inconsistent. For example, there were one or two tropical cyclones in several of the counties (Mobile County, AL, Escambia County, FL, and Montgomery County, PA) with associated flood event listings but for which the total discharge across county streamflow gages was relatively low. For storms without a flood listing for the county, in most cases the total streamflow discharge in the county was relatively low, and, in all but one of these cases all streamgage flows were below the flooding threshold. The exception was for Hurricane Ida in 2009 in Fulton County, GA.

## Patterns in tropical cyclone exposures

We used this storm exposure data to classify counties as exposed or unexposed to four different hazards for each tropical cyclone and then explored patterns over years with available data. Across the four storm hazards considered, there was wide variation in the average number of county exposures per year (Table 2). For tropical cyclone tornadoes, there were on average about 50 county exposures per year within our study. County exposures were more frequent for tropical cyclone wind exposures (*>*140 / yr on average), even more frequent for tropical cyclone flood exposures (*>*200 / yr on average), and most frequent for tropical cyclone rain exposure (*>*400 / yr on average). For every hazard except tornadoes, we identified at least one tropical cyclone that exposed over 250 counties (Table 2). However, the largest-extent tropical cyclone varied across hazards: Frances in 2004 exposed the most counties based on rain, Ike in 2008 based on wind, and Ivan in 2004 based on flooding and tornadoes (Table 2).

When we calculated and mapped the average number of exposures per decade in each county for single-hazard exposures (Figure 5), strong geographical patterns were clear. Peak sustained wind exposure had a strong coastal pattern, with almost all exposures in counties within about 200 km (124 mi) of the coastline. While tropical cyclone rain exposure was also more frequent in coastal areas compared to inland areas, there were also inland rain exposures in counties that were never classified as exposed based on wind. Flood-based exposures were frequently in the Mid-Atlantic region, with a pattern that skewed north compared with other exposures. Rain and, to some extent, flood exposures were characterized by a pattern defined by the Appalachian Mountains, with fewer exposures west of the mountain range than to the east. Almost all tornado-based exposures were in coastal states, with many in Florida, and almost none north of Maryland. Patterns were similar when analysis was restricted to years with exposure data for all four hazards available (1996–2011; Figure S3).

## Agreement across exposure metrics

Finally, we measured the agreement between exposure classifications for each pair of hazards, as well as with a proxy measurement based on distance from the storm’s track. We found that these exposure classifications typically did not agree strongly between pairs of metrics—the set of counties identified as “exposed” based on one metric often overlapped little with the set identified as “exposed” by another metric, as storms frequently brought different hazards to different locations.

Figure 6 shows as an example Hurricane Ivan in 2004. For the distance-based metric, the counties assessed as exposed follow the tropical cyclone’s track. For the wind-based metric, only counties near the tropical cyclone’s first landfall were assessed as exposed. For rain- and flood-based metrics, however, exposure extended to the left of the track, including counties as far north as New York and Connecticut, while for the tornado metric, exposed counties tended to be to the right of the track and included several counties in central North Carolina, South Carolina, and Georgia that were not identified as exposed to Ivan based on any other metric. Figure S4 provides similar maps for four other example tropical cyclones (selected because they exposed many US counties based on at least one metric).

We drew similar conclusions when we investigated all 46 tropical cyclones between 1996 and 2011

(when data for all five metrics were available) for which 100 or more counties were exposed based on at least one metric (Figure 7). In this figure, each row provides results for one tropical cyclone, and each box in that row shows the Jaccard coefficient for a pair of metrics. For almost all tropical cyclones, Jaccard coefficients were ≤ 0.6 between the distance-based metric and any of the four hazard-based metrics (i.e., out of the counties assessed as exposed by at least one of the two metrics in the pair, 60% or fewer were assessed the same under both metrics). For many tropical cyclones, Jaccard coefficients were ≤ 0.5, indicating at best moderate agreement. Among the four hazard-based metrics, the tornado-based metric had universally poor agreement with other metrics in county-level classification across the tropical cyclones considered. Agreement between other pairs of metrics were better, but still typically at best moderate (Jaccard coefficients usually ≤ 0.5).

There were, however, exceptions—tropical cyclones in which exposure assessment agreed well across several of the metrics considered. For Floyd in 1999 (Figure S4) and Irene in 2011, for example, county-level classification agreed moderately to well for all pairs of exposure metrics except those involving the tornado-based metric (Figure 7). For another set of tropical cyclones (e.g., Ernesto in 2006, and Bertha in 1996, Isabel in 2003), there was moderate to good agreement for all pairwise combinations of distance, rain, and wind, but poor agreement for other combinations of metrics, while for another set of storms (e.g., Matthew in 2004 and Katrina in 2005), there was moderate to good agreement between distance and rain.

# Discussion

Epidemiological studies can help characterize which health risks are elevated during disasters, to what degree, and to whom (Ibrahim 2005; Noji 2005). As a result, these studies help improve disaster preparedness and response (Noji 2005). However, tropical cyclones are multi-hazard events, making it complicated to measure exposure and so to conduct multi-community, multi-year studies leveraging large administrative datasets. Here, we provide open-source county-level data for several tropical cyclone exposures and explore limitations in that data. Further, we explore patterns in storm exposure classifications based on different metrics, and we find that county-level tropical cyclone exposure assessments vary substantially when using different metrics, based either on storm hazards or on how close the storm’s central track came to the community. Our results can inform exposure assessment for future county-level studies of the health risks and impacts associated with tropical cyclones exposure as well as insights to inform epidemiological study design.

## Exposure assessment data and software

The open-source data generally correspond well with data from other potential sources (Figures 2–4 and S2), but there are some caveats. The rainfall data are generally well-correlated with ground-based observations, but may sometimes oversmooth very high rainfall values (Figure 2). When rainfall data is used to create binary exposure classifications, this disagreement is unlikely to influence results, as both data sources agree in identifying these as storms with high rainfall, but would be important to consider for cases that include rainfall as a continuous measurement.

The peak sustained wind estimates are based on modeled, rather than observed, values, and while the modeled wind data generally agree well with post-analysis maximum wind radii from

HURDAT2 (Landsea and Franklin 2013) (Figure 3), there were a few storms with some discrepancies. These storms—for example, Hurricane Sandy in 2012 and Hurricane Ike in 2008—were unusually large systems for which high winds persisted well inland from landfall (Figures 3 and S2).

For the flooding data, we found that flood event status, as determined based on the NOAA Storm Events listings, typically agreed with measurements from US Geographical Survey (USGS) streamgages, with a flood event more likely to be listed if a storm elevated streamflow at streamgages across the county (Figure 4). However, there are differences between the two flooding datasets, and these highlight both the difficulty of measuring flood exposure at the county level and inherent challenges in using data from a storm event database for epidemiologic exposure assessment. For example, there was one storm in Fulton County, GA, for which there was high streamflow but not an associated flood event listing (Hurricane Ida, 2009). This storm occurred in November 2009, following a month with historic rainfall and flooding in Georgia (Shepherd, Mote, Dowd, Roden, Knox, McCutcheon, and Nelson 2011). In this case, the flooding associated with Ida was incorporated into an ongoing flood event listing, with a start date well before the five-day window we used to temporally match storm event listings with tropical cyclone tracks.

This disagreement highlights the difficulty of large-scale pairing of storm tracks with storm event listings—without criteria for temporally matching event start dates to storm dates, many false positives would be captured, for which the occurrance of a storm in the midst of an ongoing storm event might be improperly attributed to the storm. However, distance and time restrictions, like those we used in matching NOAA Storm Event listings with tropical cyclone locations and dates, do create the opportunity for occasional false negatives, as for Hurricane Ida in Fulton County, GA, where a storm contributes meaningfully to an ongoing event, but the event is not captured for the storm in the exposure data.

There are further limitations for the flood and tornado data—these data came from the National Oceanic and Atmospheric Administration (NOAA) Storm Events database, which, while a widely-used database of events maintained by NOAA, is based on reports, and so may be prone to underreporting (Ashley and Ashley 2008; Curran, Holle, and López 2000), especially in sparselypopulated areas (Witt, Eilts, Stumpf, Johnson, Mitchell, and Thomas 1998; Ashley 2007), as well as to other reporting errors.

While these are important caveats for the data, we selected these data sources as among the best currently available for measuring each of these hazards consistently and comprehensively at a multicounty, multi-year scale. In addition to providing tropical cyclone exposure metrics for individual hazards, this dataset and its associated software allow users not only to access measurements for single hazards, but also to create tropical cyclone exposure profiles based on multiple hazards or to craft exposure indices that combine hazard metrics (Chakraborty, Tobin, and Montz 2005; Peduzzi, Dao, Herold, and Mouton 2009). This functionality can be critical, as different hazards of tropical cyclones can act synergistically in causing impacts (Smith and Petley 2009).

## Patterns in tropical cyclone exposures

We found average exposures to different tropical cyclone hazards differed geographically (Figure 5). These patterns were not unexpected based on what is known about tropical cyclone hazards, but still highlight variations that are critical to consider in designing studies and statistical analysis for tropical cyclone epidemiology. Further, they demonstrate the need for multi-hazard exposure datasets for tropical cyclone epidemiology, especially in capturing inland risks.

Tropical cyclone wind exposures had a strong coastal pattern, which is consistent with the dramatic decrease in wind intensity that typically characterizes the landfall of tropical cyclones. Tropical cyclone rain exposures tended to extended further inland compared to wind exposures, up to the Appalachian mountains. This agrees with previous research indicating that the Appalachian mountains’ topography both enhances precipitation during tropical cyclones and provides hydrological conditions for severe flooding (Sturdevant–Rees, Smith, Morrison, and Baeck 2001). Almost all tropical cyclone tornado exposures were in southern coastal states, consistent with previous evidence that most noteworthy tropical cyclone-related tornadoes occur on the right-side of the tropical cyclone track in Atlantic-basin US storms (Moore and Dixon 2012). The exposure averages we calculated may be limited as estimates of long-term frequencies, as tropical cyclones follow decadal patterns (Kossin and Vimont 2007) likely not adequately captured in the available data, which covered less than 20 years.

## Agreement between exposure metrics

We found that tropical cyclones tended to bring different hazards to different counties, and so agreement was typically low between distance-based tropical cyclone exposure assessment and each of the single hazard-specific exposures, as well as between pairs of hazard-specific metrics (Figures 6– 7 and S4). These findings align with previous results from atmospheric science and related fields on the characteristics of tropical cyclones. While tropical cyclone rainfall and windspeed can be wellcorrelated when the tropical cyclone is over water (Cerveny and Newman 2000), this relationship often weakens once the hurricane has made landfall (Jiang, Halverson, Simpson, and Zipser 2008). Fast-moving tropical cyclones bring higher risks of dangerous winds inland (Kruk, Gibney, Levinson, and Squires 2010), while slow-moving tropical cyclones are likely to bring more rain (Rappaport 2000) and cause more damage because of sustained hazardous conditions (Rezapour and Baldock 2014). Further, while the likelihood and extent of flooding during a tropical cyclone is related to the tropical cyclone’s rainfall, it is also driven by factors like top soil saturation and the structure of the water basin’s drainage network (Chen, Kumar, and McGlynn 2015; Sturdevant–Rees, Smith, Morrison, and Baeck 2001).

Based on our results, the use of a distance-based metric to assess exposure to any of these hazards, or the use of measurements from one hazard as a proxy for exposure to any of the other hazards considered, would often introduce exposure misclassification. This conclusion reinforces similar findings from a study of Florida’s 2004 storm season (Grabich, Horney, Konrad, and Lobdell 2015). For some studies, such exposure misclassification might plausibly be differential. For example, tropical cyclone wind exposures tend to be concentrated in counties near the coast, while tropical cyclone wind exposures sometimes extend well inland. If the etiologically-relevant exposure for a health outcome is extreme rainfall but exposure is classified based on measurements of wind, the probability of being misclassified as unexposed would be higher in inland counties, while the probability of being misclassified as exposed would be higher in coastal counties. If coastal counties differ from inland counties in either the outcome of interest or in factors associated with risk of that outcome, exposure misclassification would be differential (Savitz and Wellenius 2016). Such differential exposure misclassification could bias estimates of tropical cyclone effects either towards the null (estimating a lower or null association compared to the true association that exists) or away from the null (estimating a larger association than actually exists) (Savitz and Wellenius 2016; Armstrong 1998).

We did, however, find a small set of tropical cyclones for which for which agreement was high across several single-hazard exposure assessments (e.g., Floyd in 1999, Irene in 2011, Hannah in 2008, Bertha in 1996; Ernesto in 2006 (Figure 7)). Hurricanes Floyd in 1999 and Irene in 2011 both made their first US landfall in North Carolina at minor hurricane windspeeds (Category 2 and 1, respectively) and then skimmed the eastern coast of the US north through New England, bringing substantial rainfall to much of the eastern coast from North Carolina north and causing extensive inland flooding in North Carolina (Floyd) and New England (Irene) (Avila and Stewart 2013; Lawrence, Avila, Beven, Franklin, Guiney, and Pasch 2000). Hurricanes Hannah in 2008, Bertha in 1996, and Ernesto in 2006 also followed the eastern coastline. For these storms, the tropical cyclones’ persistent proximity to water may have helped maintain wind speeds in similar patterns to rain and distancebased exposure assessments, resulting in more similarities across exposure assessments compared to other tropical cyclones. For these storms, it may be possible to assess exposure to multiple hazards of the storm using a single metric, in some cases even a proxy like the distance between the county and the storm’s track. However, for these storms it may be difficult to untangle the contribution of each hazard to the overall effect of the storm, given that several hazards have similar geographical patterns.

## Limitations

The dataset presented here does have several limitations, in addition to the caveats previously discussed. First, the dataset is not comprehensive of all tropical cyclone hazards. For example, coastal counties can experience dangerous storm surge, which is not specifically covered in this dataset (although some resulting coastal flooding is captured). We are exploring ways to include this in future versions of the dataset, but to date we have focused on exposures that could affect any county, whether inland or coastal. Second, this data is aggregated to the county level. Many health outcome datasets are aggregated at this level, but some may be aggregated at a finer spatial resolution (e.g., Census tract or ZIP code) or unaggregated (i.e., point locations for each outcome). We have published the wind model used to create this dataset as its own open-source R package (Anderson, Schumacher, Guikema, Quiring, and Ferreri 2017), and it can be used to model storm-associated winds at a finer spatial resolution; however, measurements of other hazards cannot similarly be rescaled through tools we provide. Next, we provide these data and associated software tools through R packages, and so some experience in the R programming language is required to make full use of them. However, R is currently a popular programming language for environmental epidemiology, allowing the data to reach a large audience, and we are exploring options to create a web application using the Shiny platform to allow broader web-based access of the data (Chang, Cheng, Allaire, Xie, and McPherson 2019). Finally, to assess patterns and agreement for binary exposure classifications, we have chosen one set of sensible thresholds for binary classifications based on continuous metrics (rainfall, maximum sustained surface wind, and distance from the storm’s track). Results and conclusions would differ somewhat with other threshold choices. We have published code for this analysis online, allowing other researchers to explore other threshold choices for these analyses.

## Conclusions

To conduct tropical cyclone epidemiological studies that span multiple communities and storms, it is critical to have consistent and comprehensive measurements of exposure to storm hazards. Here we have created and shared a dataset that provides these data for counties in the United States over multiple years. Despite some limitations in these data, they provide a powerful tool for expanding tropical cyclone epidemiology studies to more extensively leverage existing administrative health data, allowing researchers to investigate how these storms affect county-wide health risk. Further, this dataset provides hazard measurements that are comparable across communities and storms, allowing epidemiological researchers to design studies to explore how health risks are modified by characteristics of both the storms and the communities that are hit. The data are given in an opensource format, along with associated software tools, which allows them to be freely used and for others to explore all associated code and to contribute additions through platforms like GitHub.

Based on our analysis in this paper, these data are typically in agreement with measurements from other sources of data available to characterize storm-associated hazards (e.g., ground-based monitors, streamgages, post-storm wind radii estimates). However, researchers who are planning to use the data should explore the analyses presented in this paper to understand the strengths and weaknesses of the data. Further, our results indicate that county-level storm exposure is not wellcharacterized by the closest distance that a storm’s central track came to a county, and that exposure to one storm hazard within a county (e.g., severe winds) does not imply exposure to other hazards (e.g., excessive rainfall, flooding). As a result, it is critical that researchers consider which storm hazards are likely on the causal pathway for the outcomes they are studying, and to characterize storm exposure in a way that captures those specific hazards, to avoid exposure misclassification.

# References

Al-Hamdan, Mohammad Z, William L Crosson, Sigrid A Economou, Maurice G Estes Jr, Sue M Estes, Sarah N Hemmings, Shia T Kent, Mark Puckett, Dale A Quattrochi, Douglas L Rickman, et al. (2014). “Environmental public health applications using remotely sensed data”. *Geocarto International* 29 (1), 85–98.

Anderson, G Brooke and Michelle L Bell (2010). “Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 US communities”. *Environmental Health Perspectives* 119 (2), 210–218.

Anderson, G Brooke and Dirk Eddelbuettel (2017). “Hosting data packages via drat: A case study with hurricane exposure data”. *The R Journal* 9 (1).

Anderson, G. Brooke and Ziyu Chen (2017). *noaastormevents: Explore NOAA Storm Events Database*.

R package version 0.1.0. url: <https://github.com/zailchen/noaastormevents>.

Anderson, G. Brooke and Zhiming Guo (2016). *countytimezones: Convert from UTC to Local Time for United States Counties*. R package version 1.0.0. url: [https://github.com/geanders/ countytimezones](https://github.com/geanders/countytimezones).

Anderson, G. Brooke, Andrea Schumacher, William Crosson, Mohammad Al-Hamdan, Meilin Yan, Joshua Ferreri, Ziyu Chen, Steven Quiring, and Seth Guikema (2017). *hurricaneexposuredata: Data Characterizing Exposure to Hurricanes in United States Counties*. R package version 0.0.2.

url: <https://github.com/geanders/hurricaneexposuredata>.

Anderson, G. Brooke, Andrea Schumacher, Seth Guikema, Steven Quiring, and Joshua Ferreri (2017). *stormwindmodel: Model Tropical Cyclone Wind Speeds*. R package version 0.1.0. url: [https: //CRAN.R-project.org/package=stormwindmodel](https://CRAN.R-project.org/package=stormwindmodel).

Anderson, G. Brooke, Meilin Yan, Joshua Ferreri, William Crosson, Mohammad Al-Hamdan, Andrea Schumacher, and Dirk Eddelbuettel (2017). *hurricaneexposure: Explore and Map County-Level Hurricane Exposure in the United States*. R package version 0.0.1. url: [https://cran.rproject.org/package=hurricaneexposure](https://cran.r-project.org/package=hurricaneexposure).

Armstrong, Ben G (1998). “Effect of measurement error on epidemiological studies of environmental and occupational exposures”. *Occupational and Environmental Medicine* 55 (10), 651–656.

Ashley, Sharon T and Walker S Ashley (2008). “Flood fatalities in the United States”. *Journal of Applied Meteorology and Climatology* 47 (3), 805–818.

Ashley, Walker S (2007). “Spatial and temporal analysis of tornado fatalities in the United States: 1880–2005”. *Weather and Forecasting* 22 (6), 1214–1228.

Atallah, Eyad H. and Lance F. Bosart (2003). “The extratropical transition and precipitation distribution of Hurricane Floyd (1999)”. *Monthly Weather Review* 131, 1063–1081.

Atallah, Eyad, Lance F Bosart, and Anantha R Aiyyer (2007). “Precipitation distribution associated with landfalling tropical cyclones over the eastern United States”. *Monthly Weather Review* 135 (6), 2185–2206.

Avila, Lixion A and Stacy R Stewart (2013). “Atlantic hurricane season of 2011”. *Monthly Weather Review* 141 (8), 2577–2596.

Bivand, Roger S., Edzer Pebesma, and Virgilio Gomez-Rubio (2013). *Applied Spatial Data Analysis with R*. NY: Springer.

Caillouët, Kevin A, Sarah R Michaels, Xu Xiong, Ivo Foppa, and Dawn M Wesson (2008). “Increase in West Nile neuroinvasive disease after Hurricane Katrina”. *Emerging Infectious Diseases* 14 (5), 804–808.

Cerveny, Randall S. and Lynn E. Newman (2000). “Climatological relationships between tropical cyclones and rainfall”. *Monthly Weather Review* 128, 3329–3336.

Chakraborty, Jayajit, Graham A Tobin, and Burrell E Montz (2005). “Population evacuation: assessing spatial variability in geophysical risk and social vulnerability to natural hazards”. *Natural Hazards Review* 6 (1), 23–33.

Chamberlain, Scott (2017). *rnoaa: ’NOAA’ Weather Data from R*. R package version 0.7.0. url: <https://CRAN.R-project.org/package=rnoaa>.

Chang, Winston, Joe Cheng, JJ Allaire, Yihui Xie, and Jonathan McPherson (2019). *shiny: Web Application Framework for R*. R package version 1.3.2. url: [https://CRAN.R-project.org/ package=shiny.](https://CRAN.R-project.org/package=shiny)

Chen, Xing, Mukesh Kumar, and Brian L. McGlynn (2015). “Variations in streamflow response to large hurricane-season storms in a southeastern US watershed”. *Journal of Hydrometeorology* 16, 55–69.

Curran, E Brian, Ronald L Holle, and Raúl E López (2000). “Lightning casualties and damages in the United States from 1959 to 1994”. *Journal of Climate* 13 (19), 3448–3464.

Currie, Janet and Maya Rossin-Slater (2013). “Weathering the storm: hurricanes and birth outcomes”. *Journal of Health Economics* 32, 487–503.

Czajkowski, Jeffrey and Emily Kennedy (2010). “Fatal tradeoff? Toward a better understanding of the costs of not evacuating from a hurricane in landfall counties”. *Population and Environment* 31 (1-3), 121–149.

Czajkowski, Jeffrey, Kevin Simmons, and Daniel Sutter (2011). “An analysis of coastal and inland fatalities in landfalling US hurricanes”. *Natural Hazards* 59, 1513–1531.

Elsberry, Russell L. (2002). “Predicting hurricane landfall precipitation: optimistic and pessimistic views from the symposium on precipitation extremes”. *Bulletin of the American Meteorological Society* 83 (9), 1333–1339.

Grabich, Shannon C, Jennifer Horney, Charles Konrad, and Danelle T Lobdell (2015). “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, C Konrad, and Jennifer Horney (2015). “County-level hurricane exposure and birth rates: application of differencein-differences analysis for confounding control”. *Emerging Themes in Epidemiology* 12, 19.

Haikerwal, Anjali, Muhammad Akram, Anthony Del Monaco, Karen Smith, Malcolm R Sim, Mick Meyer, Andrew M Tonkin, Michael J Abramson, and Martine Dennekamp (2015). “Impact of fine particulate matter (PM 2.5) exposure during wildfires on cardiovascular health outcomes”. *Journal of the American Heart Association* 4 (7), e001653.

Halverson, Jeffrey B. (2015). “Second wind: The deadly and destructive inland phase of East Coast hurricanes”. *Weatherwise* 68 (2), 20–27.

Hirsch, Robert M. and Laura A. De Cicco (2015). “User guide to Exploration and Graphics for RivEr

Trends (EGRET) and dataRetrieval: R packages for hydrologic data”. *Techniques and Methods*.

Reston, VA: U.S. Geological Survey. Chap. A10. url: <https://pubs.usgs.gov/tm/04/a10/>. Ibrahim, Michel A (2005). “Unfortunate, but Timely”. *Epidemiologic Reviews* 27, 1–2.

Jaccard, Paul (1901). “Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines”. *Bulletin de la Société Vaudoise des Sciences Naturelles* 37, 241–272.

— (1908). “Nouvelles recherches sur la distribution floral”. *Bulletin de la Société Vaudoise des Sciences Naturelles* 44, 223–270.

Jarvinen, Brian R. and Eduardo L. Caso (1978). “A tropical cyclone data tape for the North Atlantic basin, 1886–1977: Contents, limitations, and uses”. *NOAA Technical Memorandum NWS NHC*

*6*.

Jiang, Haiyan, Jeffrey B. Halverson, Joanne Simpson, and Edward J. Zipser (2008). “Hurricane

’rainfall potential’ derived from satellite observations aids overland rainfall prediction”. *Journal of Applied Meteorology and Climatology* 47, 944–959.

Kim, Hyun, Rebecca M Schwartz, Jerrold Hirsch, Robert Silverman, Bian Liu, and Emanuela Taioli

(2016). “Effect of Hurricane Sandy on Long Island emergency departments visits”. *Disaster Medicine and Public Health Preparedness* 10 (3), 344–350.

Kinney, Dennis K, Andrea M Miller, David J Crowley, and Erika Gerber (2008). “Autism prevalence following prenatal exposure to hurricanes and tropical storms in Louisiana”. *Journal of Autism and Developmental Disorders* 38 (3), 481–488.

Knaff, John A, Christopher J Slocum, Kate D Musgrave, Charles R Sampson, and Brian R Strahl (2016). “Using routinely available information to estimate tropical cyclone wind structure”. *Monthly Weather Review* 144, 1233–1247.

Kossin, James P and Daniel J Vimont (2007). “A more general framework for understanding Atlantic hurricane variability and trends”. *Bulletin of the American Meteorological Society* 88 (11), 1767– 1781.

Kruk, Michael C., Ethan J. Gibney, David H. Levinson, and Michael Squires (2010). “A climatology of inland winds from tropical cyclones for the Eastern United States”. *Journal of Applied Meteorology and Climatology* 49, 1538–1547.

Lammers, Rod and G. Brooke Anderson (2017). *countyfloods: Quantify United States CountyLevel Flood Measurements*. R package version 0.0.2. url: [https://CRAN.R-project.org/ package=countyfloods](https://CRAN.R-project.org/package=countyfloods).

Landsea, Christopher W. and James L. Franklin (2013). “Atlantic hurricane database uncertainty and presentation of a new database format”. *Monthly Weather Review* 141, 3576–3592.

Lawrence, Miles, Lixion Avila, Jack Beven, James Franklin, John Guiney, and Richard Pasch (2000). “Atlantic hurricanes: Floyd’s fury highlighted another above-average year”. *Weatherwise* 53 (2), 39–43.

Liu, Jia Coco, Ander Wilson, Loretta J Mickley, Francesca Dominici, Keita Ebisu, Yun Wang, Melissa P Sulprizio, Roger D Peng, Xu Yue, Ji-Young Son, et al. (2017). “Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties”. *Epidemiology* 28 (1), 77–85.

McKinney, Nathan, Chris Houser, and Klaus Meyer-Arendt (2011). “Direct and indirect mortality in Florida during the 2004 hurricane season”. *International Journal of Biometeorology* 55 (4), 533–546.

Menne, Matthew J, Imke Durre, Russell S Vose, Byron E Gleason, and Tamara G Houston (2012).

“An overview of the Global Historical Climatology Network—Daily database”. *Journal of Atmospheric and Oceanic Technology* 29 (7), 897–910.

Milojevic, Ai, Ben Armstrong, and Paul Wilkinson (2017). “Mental health impacts of flooding: a controlled interrupted time series analysis of prescribing data in England”. *Journal of Epidemiology and Community Health* 71 (10), 970–973.

Mongin, Steven J, Sherry L Baron, Rebecca M Schwartz, Bian Liu, Emanuela Taioli, and Hyun Kim

(2017). “Measuring the Impact of Disasters Using Publicly Available Data: Application to 2012 Hurricane Sandy”. *American Journal of Epidemiology* 186 (11), 1290–1299.

Moore, Todd W. and Richard W. Dixon (2012). “Tropical cyclone–tornado casualties”. *Natural Hazards* 61, 621–634.

National Oceanic and Atmospheric Administration’s National Centers for Environmental Information (2017). *Storm Events Database*. <https://www.ncdc.noaa.gov/stormevents/>. Accessed: 2017-03-23.

Noji, Eric K (2005). “Disasters: Introduction and state of the art”. *Epidemiologic Reviews* 27 (1), 3–8.

Peduzzi, Pascal, Hy Dao, Christian Herold, and Frederic Mouton (2009). “Assessing global exposure and vulnerability towards natural hazards: the Disaster Risk Index”. *Natural Hazards and Earth System Sciences* 9 (4), 1149–1159.

Phadke, Amal C, Christopher D Martino, Kwok Fai Cheung, and Samuel H Houston (2003). “Modeling of tropical cyclone winds and waves for emergency management”. *Ocean Engineering* 30 (4), 553–578.

Rappaport, Edward N (2014). “Fatalities in the United States from Atlantic tropical cyclones: New data and interpretation”. *Bulletin of the American Meteorological Society* 95 (3), 341–346.

Rappaport, Edward N. (2000). “Loss of life in the United States associated with recent Atlantic tropical cyclones”. *Bulletin of the American Meteorological Society* 81 (9), 2065–2073.

Rappaport, Edward N and B Wayne Blanchard (2016). “Fatalities in the United States indirectly associated with Atlantic tropical cyclones”. *Bulletin of the American Meteorological Society* 97 (7), 1139–1148.

Rezapour, Mehdi and Tom E. Baldock (2014). “Classification of hurricane hazards: the importance of rainfall”. *Weather and Forcasting* 29, 1319–1331.

Rocha, Luciana A, Catharine Q Fromknecht, Sarah Davis Redman, Joanne E Brady, Sarah E Hodge, and Rebecca S Noe (2017). “Medicolegal death scene investigations after natural disaster-and

weather-related events: a review of the literature”. *Academic Forensic Pathology* 7 (2), 221–239.

Rui, Hualan and David Mocko (2014). “README Document for North America Land Data Assimilation System Phase 2 (NLDAS-2) Products”. *Goddard Earth Sciences Data and Information Services Center*.

Santos-Burgoa, Carlos, John Sandberg, Erick Suárez, Ann Goldman-Hawes, Scott Zeger, Alejandra Garcia-Meza, Cynthia M Pérez, Noel Estrada-Merly, Uriyoan Colón-Ramos, Cruz Marı́a Nazario, et al. (2018). “Differential and persistent risk of excess mortality from Hurricane Maria in Puerto Rico: a time-series analysis”. *The Lancet Planetary Health* 2 (11), e478–e488.

Santos-Lozada, Alexis R and Jeffrey T Howard (2018). “Use of death counts from vital statistics to calculate excess deaths in Puerto Rico following Hurricane Maria”. *Journal of the American Medical Association* 320 (14), 1491–1493.

Savitz, David A and Gregory A Wellenius (2016). *Interpreting Epidemiologic Evidence: Connecting Research to Applications*. Oxford University Press.

Severson, Rachel and G. Brooke Anderson (2016). *countyweather: Compiles Meterological Data for U.S. Counties*. R package version 0.1.0.9000.

Shepherd, Marshall, Thomas Mote, John Dowd, Mike Roden, Pamela Knox, Steven C. McCutcheon, and Steven E. Nelson (2011). “An overview of synoptic and mesoscale factors contributing to the disastrous Atlanta flood of 2009”. *Bulletin of the American Meteorological Society* 92 (7), 861–870.

Smith, Keith and David N. Petley (2009). *Environmental Hazards: Assessing Risk and Reducing Disaster*. Taylor and Francis.

Son, Ji-Young, Jong-Tae Lee, G Brooke Anderson, and Michelle L Bell (2012). “The impact of heat waves on mortality in seven major cities in Korea”. *Environmental Health Perspectives* 120 (4), 566–571.

Spearman, Charles (1904). “The proof and measurement of association between two things”. *The American Journal of Psychology* 15 (1), 72–101.

Sturdevant–Rees, Paula, James A. Smith, Julia Morrison, and Mary Lynn Baeck (2001). “Tropical storms and the flood hydrology of the central Appalachians”. *Water Resources Research* 37 (8), 2143–2168.

Swerdel, Joel N, Teresa M Janevic, Nora M Cosgrove, John B Kostis, et al. (2014). “The effect of Hurricane Sandy on cardiovascular events in New Jersey”. *Journal of the American Heart Association* 3 (6), e001354.

United States Census Bureau (2015). *Centers of Population*. [https://www.census.gov/geo/ reference/centersofpop.html](https://www.census.gov/geo/reference/centersofpop.html). Accessed: 2017-03-21.

United States Centers for Disease Control (2016). *North America Land Data Assimilation System (NLDAS) Daily Precipitation years 1979–2011 on CDC WONDER Online Database*. [http://www.](http://www.nhc.noaa.gov/aboutnames_history.shtml)

[nhc.noaa.gov/aboutnames\_history.shtml](http://www.nhc.noaa.gov/aboutnames_history.shtml). http://wonder.cdc.gov/wonder/help/Precipitation.html.

United States Geological Survey (2016). *National Water Information System data available on the World Wide Web (USGS Water Data for the Nation)*. <https://waterdata.usgs.gov/nwis/>. Accessed: 2018-02-26.

Willoughby, H E, R W R Darling, and M E Rahn (2006). “Parametric representation of the primary hurricane vortex. Part II: A new family of sectionally continuous profiles”. *Monthly Weather Review* 134 (4), 1102–1120.

Witt, Arthur, Michael D Eilts, Gregory J Stumpf, JT Johnson, E De Wayne Mitchell, and Kevin

W Thomas (1998). “An enhanced hail detection algorithm for the WSR-88D”. *Weather and Forecasting* 13 (2), 286–303.

Zandbergen, Paul A. (2009). “Exposure of US counties to Atlantic tropical storms and hurricanes, 1851–2003”. *Natural Hazards* 48, 83–99.

Zhu, Laiyin and Steven M Quiring (2013). “Variations in tropical cyclone precipitation in Texas (1950 to 2009)”. *Journal of Geophysical Research: Atmospheres* 118 (8), 3085–3096.

Table 1: Exposure metrics considered to assess county-level exposure to tropical cyclones

|  |  |  |
| --- | --- | --- |
| Metric | Available | Criteria for exposure |
| Distance | 1988–2015 | Storm track came within 100 km of county population mean center |
| Rain | 1988–2011 | County cumulative rainfall of ≥75 mm over the period from two days before to one day after the storm’s closest approach and the storm track came within 500 km of the county population mean center |
| Wind | 1988–2015 | Modeled storm-associated peak sustained surface wind of |
| Flood | 1996–2015 | ≥34 kts at the county’s population mean center  Flood event listed in the NOAA Storm Events database for the county with a start date within two days of the storm’s closest approach and the storm track came within 500 km of the county population mean center |
| Tornado | 1996–2015 | Tornado event listed in the NOAA Storm Events database for the county with a start date within two days of the storm’s closest approach and the storm track came within 500 km of the county population mean center |

Table 2: Summary statistics for the number of county tropical cyclone exposures under each metric.

|  |  |  |
| --- | --- | --- |
| Metric (years available) | Mean (interquartile range) of county exposures per year | Tropical cyclone with most counties exposures (# exposed counties) |
| Distance (1988–2015) | 433 (224, 503) | Beryl, 1994 (330) |
| Rain (1988–2015) | 401 (152, 553) | Frances, 2004 (464) |
| Wind (1988–2011) | 143 (51, 232) | Ike, 2008 (251) |
| Flood (1996–2015) | 213 (84, 248) | Ivan, 2004 (318) |
| Tornado (1996–2015) | 53 (16, 43) | Ivan, 2004 (91) |

**Figures**

# List of Figures

1. Study area and storms considered in this study. All counties in the states shown in this map were investigated. The lines show the paths of the study storms, which included all tracked storms in 1988–2015 that are recorded in HURDAT2 and that came within 250 km of at least one US county. Thicker lines show the tracks of storms whose names have been retired, indicating that the storm was particularly severe or had notable impacts.
2. Comparison in nine sample counties of two sources of storm-associated rainfall estimates: (1) county-level estimates derived from a re-analysis dataset and (2) countylevel estimates based on ground-based observations in nine sample counties. For both, estimates include rainfall from two days before to one day after the storm’s closest approach to the county. Each small graph shows data for one sample county, and each point shows one tropical storm. The number of storms within each county and the number of ground-based monitors reporting precipitation during the county’s storms are given above each plot. Horizontal and vertical lines in each plot show the threshold of 75 mm used to classify a storm as “exposed” for the binary classification considered for this metric in further analysis (Table 1). Note that ranges of the xand y-axes differ across counties.
3. Comparison of two sources of wind exposure estimates: (1) modeled peak sustained surface wind and (2) estimates based on HURDAT2’s wind speed radii. Each point represents a storm, with the x axis giving the percent of counties classified in the same category of peak sustained surface wind (*<*34 kt; 34–49.9 kt; 50–63.9 kt; ≥64 kt) by both sources of data. The color of each point gives the number of study counties that were exposed to peak sustained surface wind of at least 34 kt (based on modeled wind). Estimates are shown for study storms since 2004, the earliest year for which post-storm reanalysis wind speed radii estimates are routinely available in HURDAT2, and for which at least one study county had a peak sustained wind of ≥34 kn based on the HURDAT2 wind radii.
4. Comparison in nine sample study counties of flood status based on two data sources: (1) NOAA Storm Events listings matched to HURDAT2 storm track data and (2) county streamflow gages. Each small plot shows results for one of the sample counties. Each point represents a single tropical storm, and the point’s position along the xaxis shows the highest daily total streamflow (cubic feet per second) during the five-day window surrounding the storm, summed across all identified streamgages in the county. The y-axis separates storms for which a flood event was reported in NOAA’s Storm Events database in the county with a start date within the five-day window of the storm’s closest approach. The color of each point gives the percent of streamflow gages in the county with a daily streamflow that exceeded a threshold of flooding (the streamgage’s median value for annual peak flow) on any day during the five-day window. The number of streamflow gages used in analysis are given in parentheses beside the county’s name in the panel title. Note that the x-axis scales differ by county, depending on the number of streamflow gages and typical flow rates for each gage, and are on a log-10 scale.
5. Average number of county-level storm exposures per decade for binary classifications based on each single-hazard exposure metric (Table 1). The years used to estimate these averages are based on years of available exposure data (rain: 1988–2011; wind: 1988–2015; flood and tornado events: 1996–2015).
6. Counties classified as exposed to Hurricane Ivan in 2004 under each exposure metric considered (Table 1). The red line shows the track of Hurricane Ivan based on HURDAT2. Similar maps for other large-extent storms are given in Figure S4.
7. Agreement between exposure classifications based on different exposure metrics for all storms between 1996 and 2011 for which at least 100 counties were exposed based on at least one metric. Each row shows one storm, and the color of each cell shows the measured Jaccard index for each pair of exposure metrics (proportion of counties classified as exposed by both metrics out of counties classified as exposed by either metric). For Grace in 2003 and Ida in 2009, there were no county exposures for either the tornado-based metric or the wind-based metric (indicated by gray squares). Colors to the right of the heatmap show the number of exposed counties based on any of the metrics. Storms are displayed within clusters that have similar patterns based on hierarchical clustering.

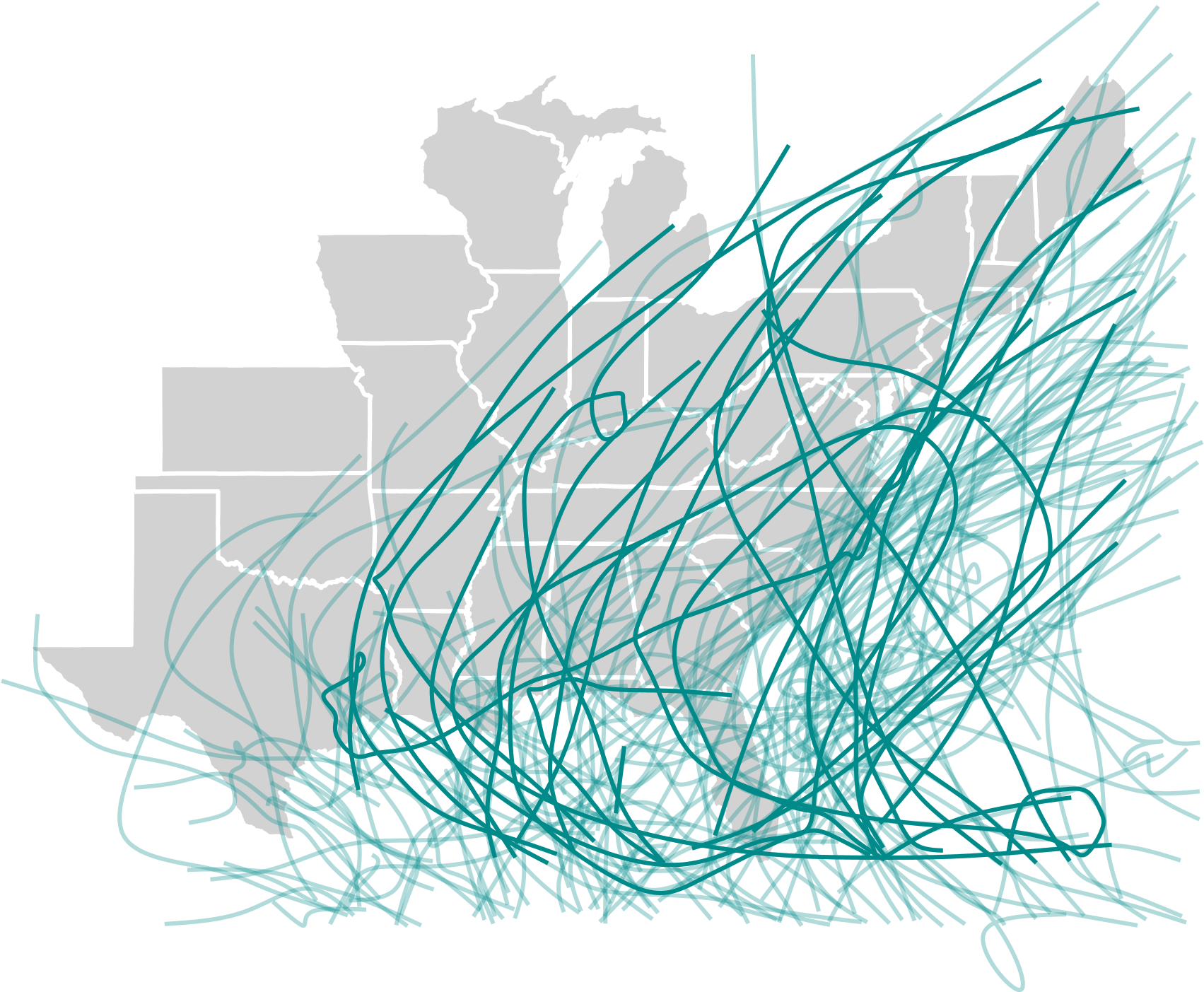
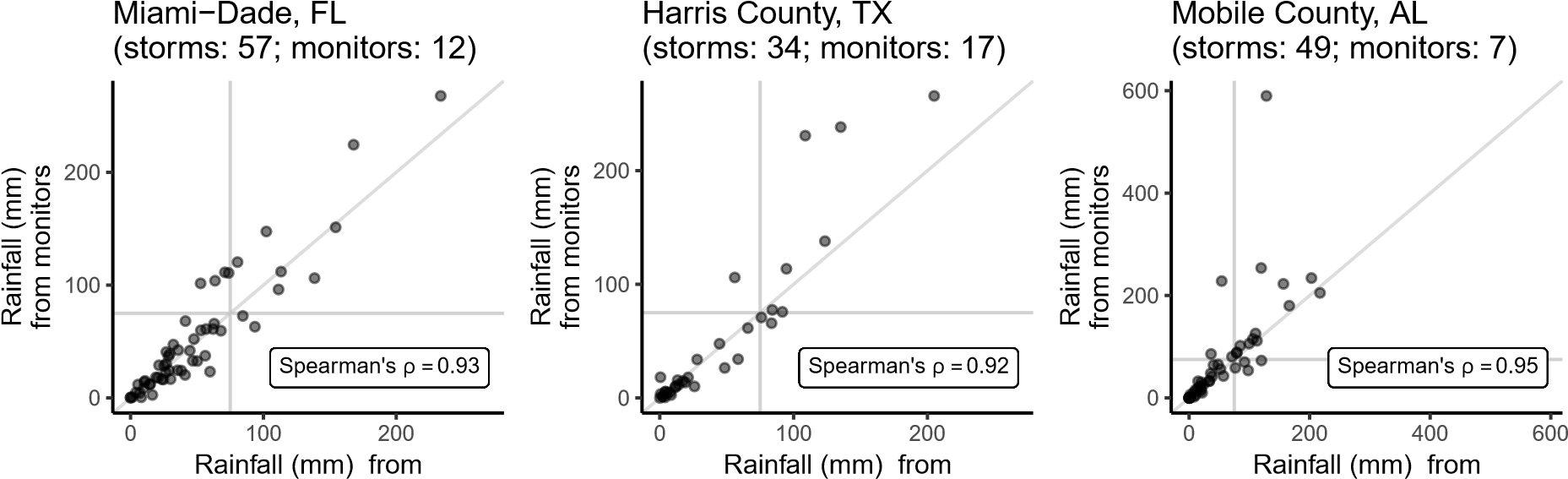
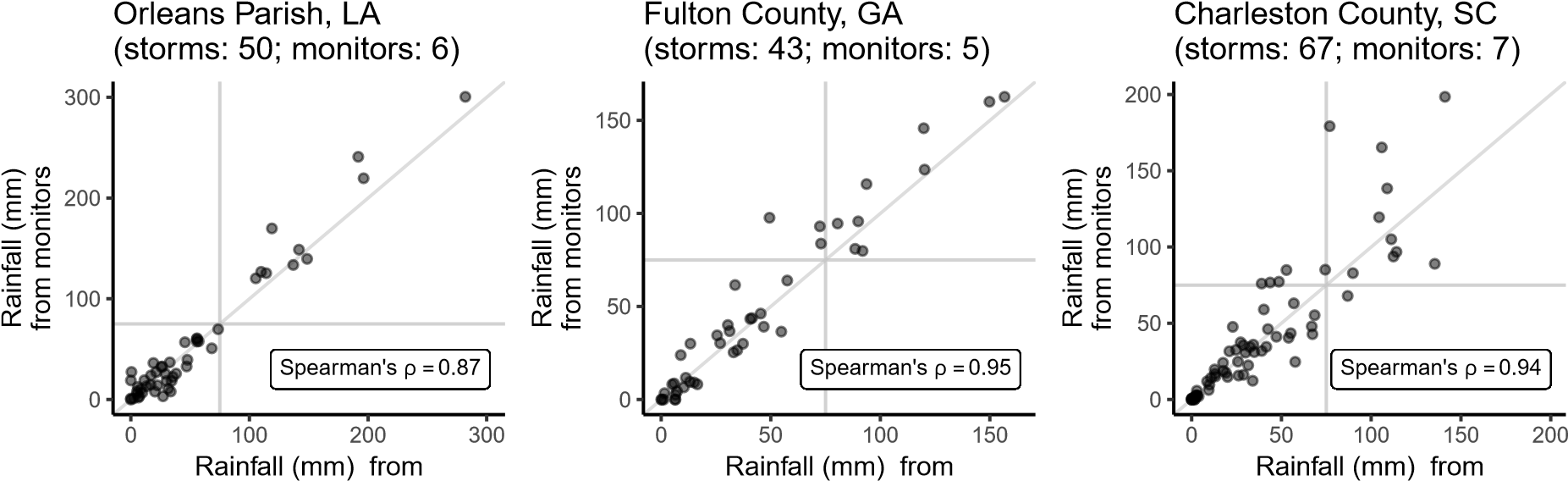


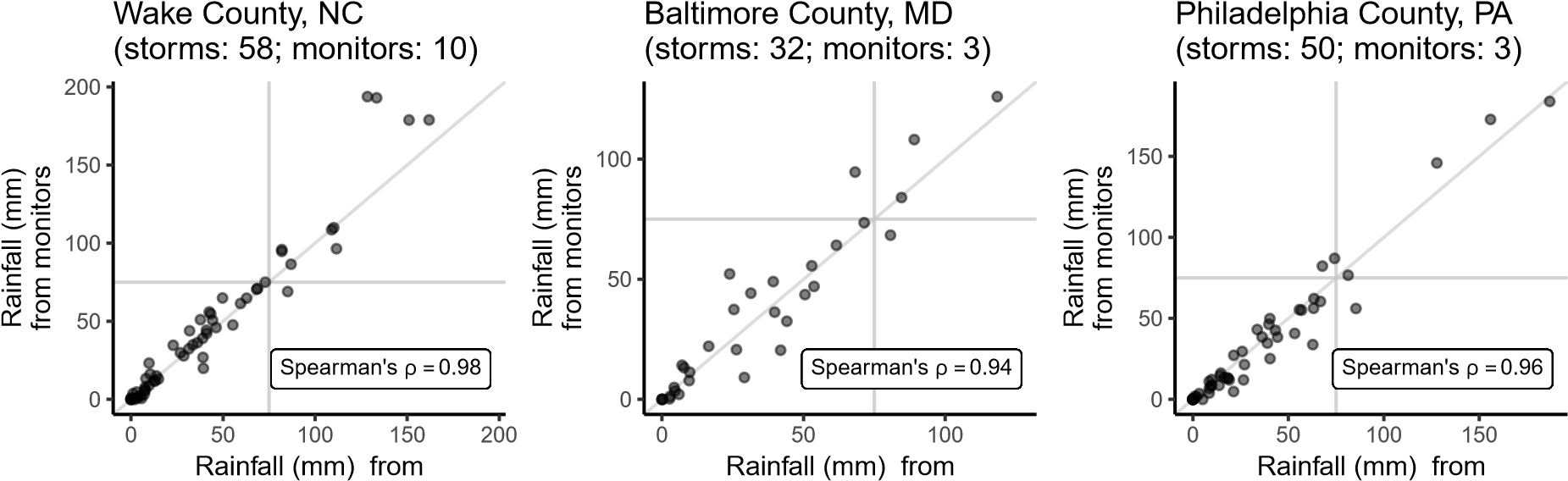
Figure 1: Study area and storms considered in this study. All counties in the states shown in this map were investigated. The lines show the paths of the study storms, which included all tracked storms in 1988–2015 that are recorded in HURDAT2 and that came within 250 km of at least one US county. Thicker lines show the tracks of storms whose names have been retired, indicating that the storm was particularly severe or had notable impacts.



NLDAS−2 data NLDAS−2 data NLDAS−2 data

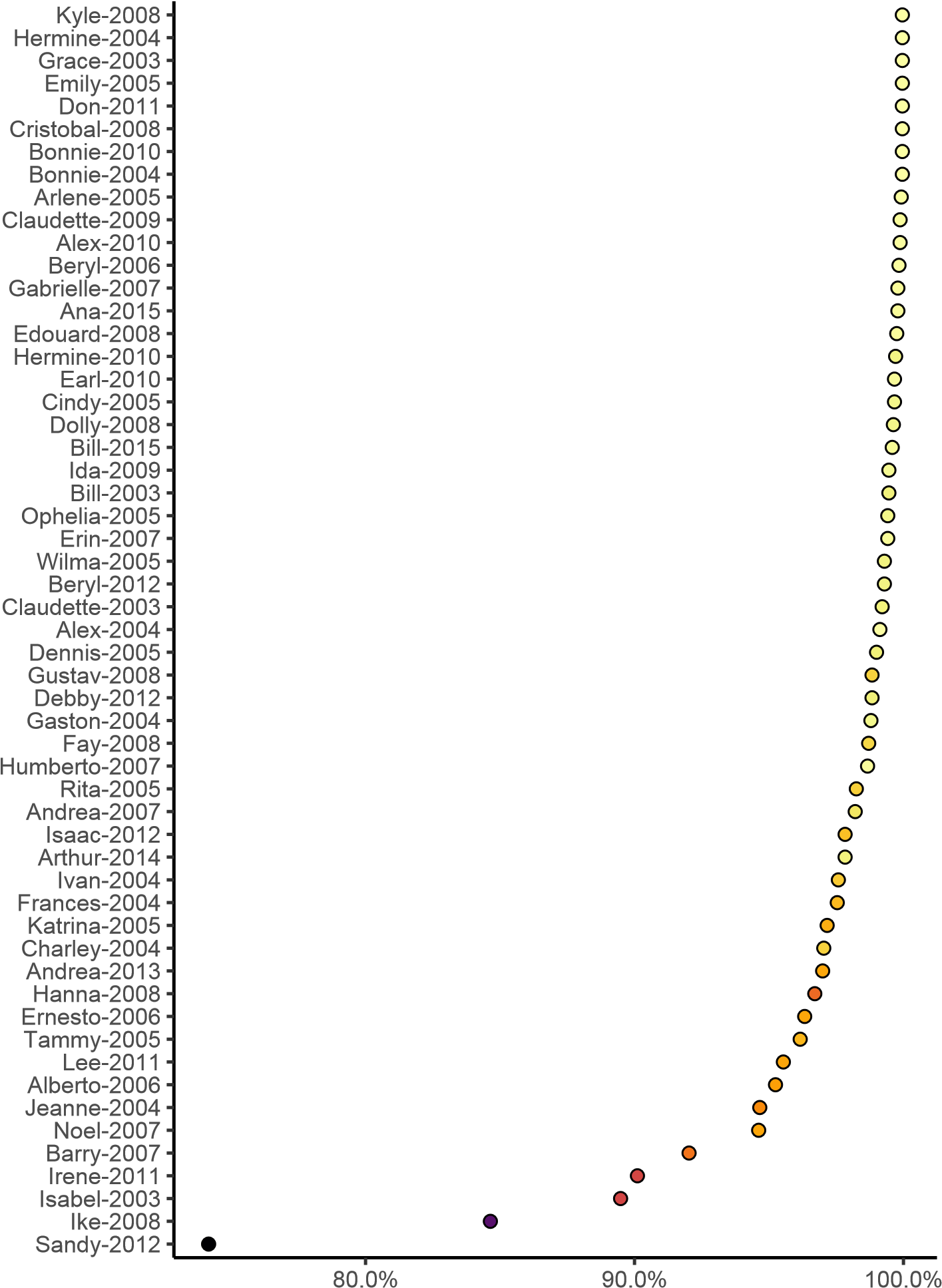


NLDAS−2 data NLDAS−2 data NLDAS−2 data



NLDAS−2 data NLDAS−2 data NLDAS−2 data

Figure 2: Comparison in nine sample counties of two sources of storm-associated rainfall estimates: (1) county-level estimates derived from a re-analysis dataset and (2) county-level estimates based on ground-based observations in nine sample counties. For both, estimates include rainfall from two days before to one day after the storm’s closest approach to the county. Each small graph shows data for one sample county, and each point shows one tropical storm. The number of storms within each county and the number of ground-based monitors reporting precipitation during the county’s storms are given above each plot. Horizontal and vertical lines in each plot show the threshold of 75 mm used to classify a storm as “exposed” for the binary classification considered for this metric in further analysis (Table 1). Note that ranges of the x- and y-axes differ across counties.

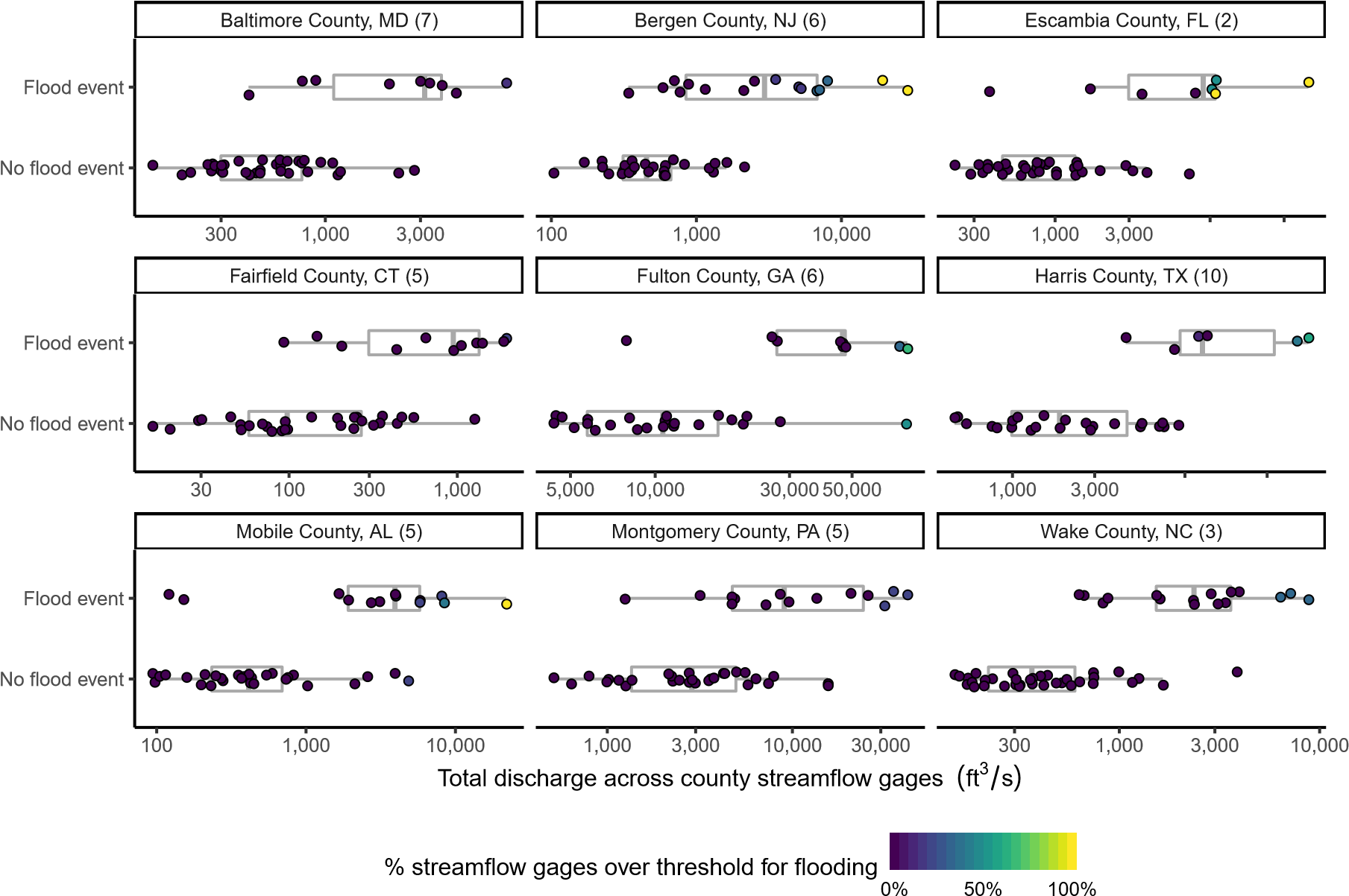


% of counties classified in same wind category by modeled storm wind and Extended Best Tracks

Number of counties with modeled 

sustained winds of ≥ 34 knots 200 400 600

Figure 3: Comparison of two sources of wind exposure estimates: (1) modeled peak sustained surface wind and (2) estimates based on HURDAT2’s wind speed radii. Each point represents a storm, with the x axis giving the percent of counties classified in the same category of peak sustained surface wind (*<*34 kt; 34–49.9 kt; 50–63.9 kt; ≥64 kt) by both sources of data. The color of each point gives the number of study counties that were exposed to peak sustained surface wind of at least 34 kt (based on modeled wind). Estimates are shown for study storms since 2004, the earliest year for which post-storm reanalysis wind speed radii estimates are routinely available in HURDAT2, and for which at least one study county had a peak sustained wind of ≥34 kn based on the HURDAT2 wind radii.

10,000 30,000

10,000 30,000

Figure 4: Comparison in nine sample study counties of flood status based on two data sources: (1) NOAA Storm Events listings matched to HURDAT2 storm track data and (2) county streamflow gages. Each small plot shows results for one of the sample counties. Each point represents a single tropical storm, and the point’s position along the x-axis shows the highest daily total streamflow (cubic feet per second) during the five-day window surrounding the storm, summed across all identified streamgages in the county. The y-axis separates storms for which a flood event was reported in NOAA’s Storm Events database in the county with a start date within the five-day window of the storm’s closest approach. The color of each point gives the percent of streamflow gages in the county with a daily streamflow that exceeded a threshold of flooding (the streamgage’s median value for annual peak flow) on any day during the five-day window. The number of streamflow gages used in analysis are given in parentheses beside the county’s name in the panel title. Note that the x-axis scales differ by county, depending on the number of streamflow gages and typical flow rates for each gage, and are on a log-10 scale.

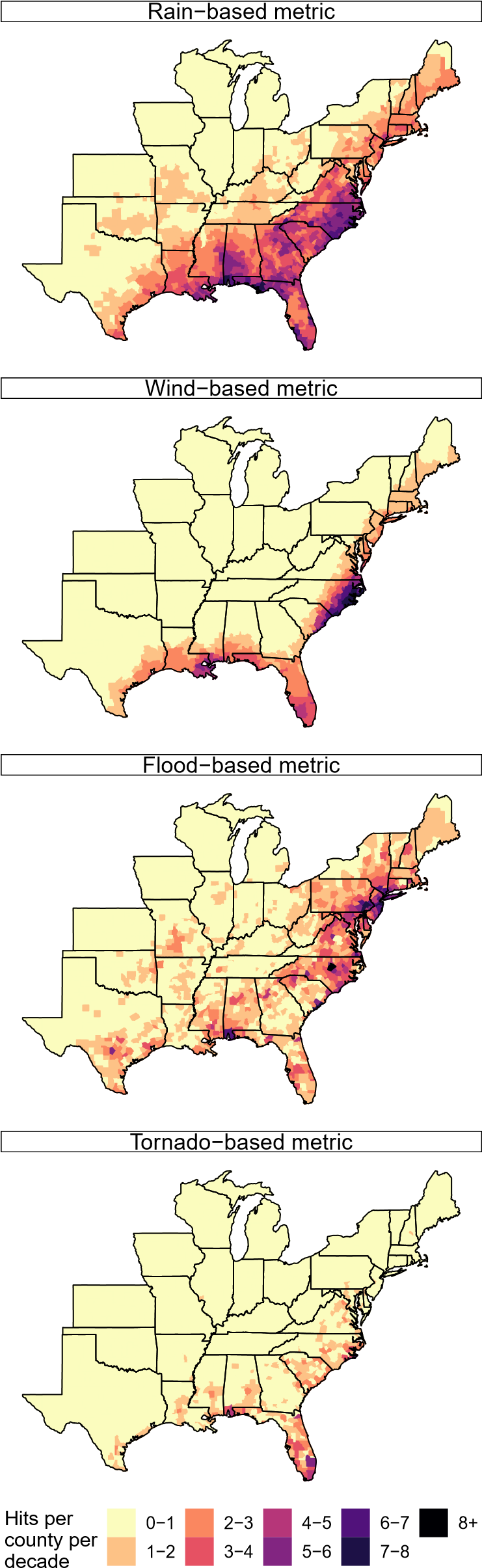
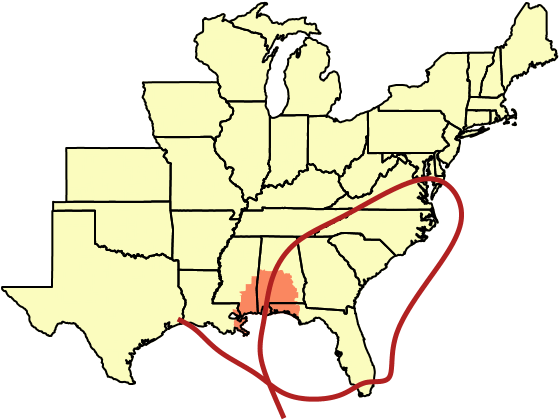
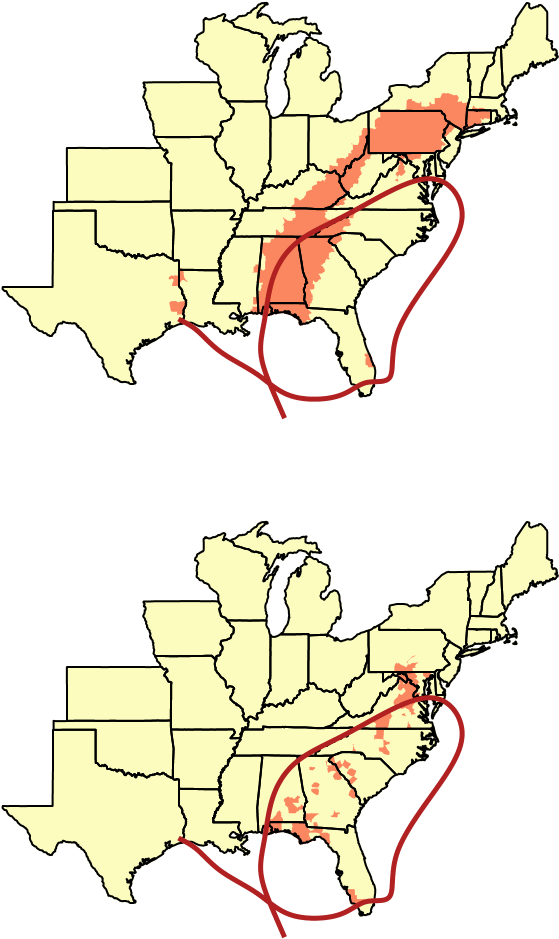
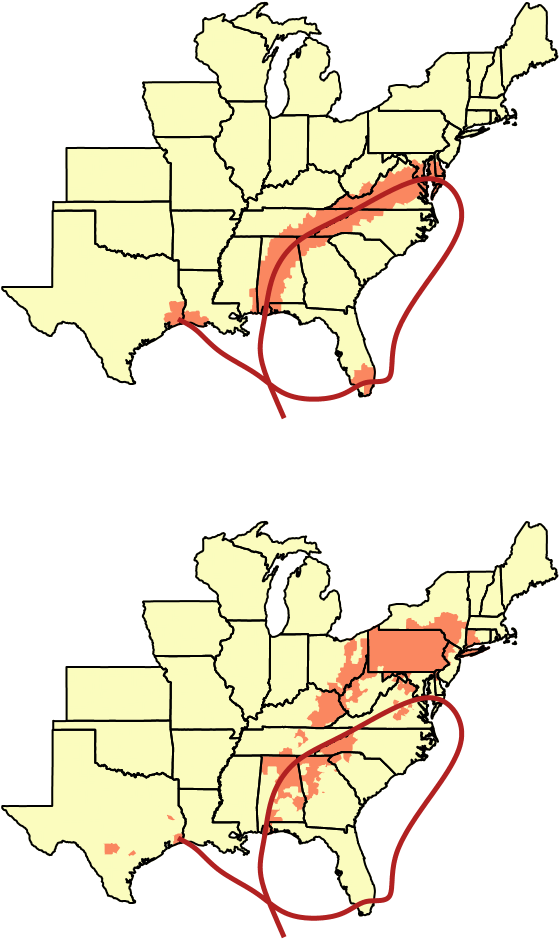


Figure 5: Average number of county-level storm exposures per decade for binary classifications based on each single-hazard exposure metric (Table 1). The years used to estimate these averages are based on years of available exposure data (rain: 1988–2011; wind: 1988–2015; flood and tornado events: 1996–2015).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Distance−based metric |  | Rain−based metric |  | Wind−based metric |

|  |  |  |
| --- | --- | --- |
| Flood−based metric |  | Tornado−based metric |

|  |
| --- |
|  |
|  |

Unexposed

Exposed

Figure 6: Counties classified as exposed to Hurricane Ivan in 2004 under each exposure metric considered (Table 1). The red line shows the track of Hurricane Ivan based on HURDAT2. Similar maps for other large-extent storms are given in Figure S4.

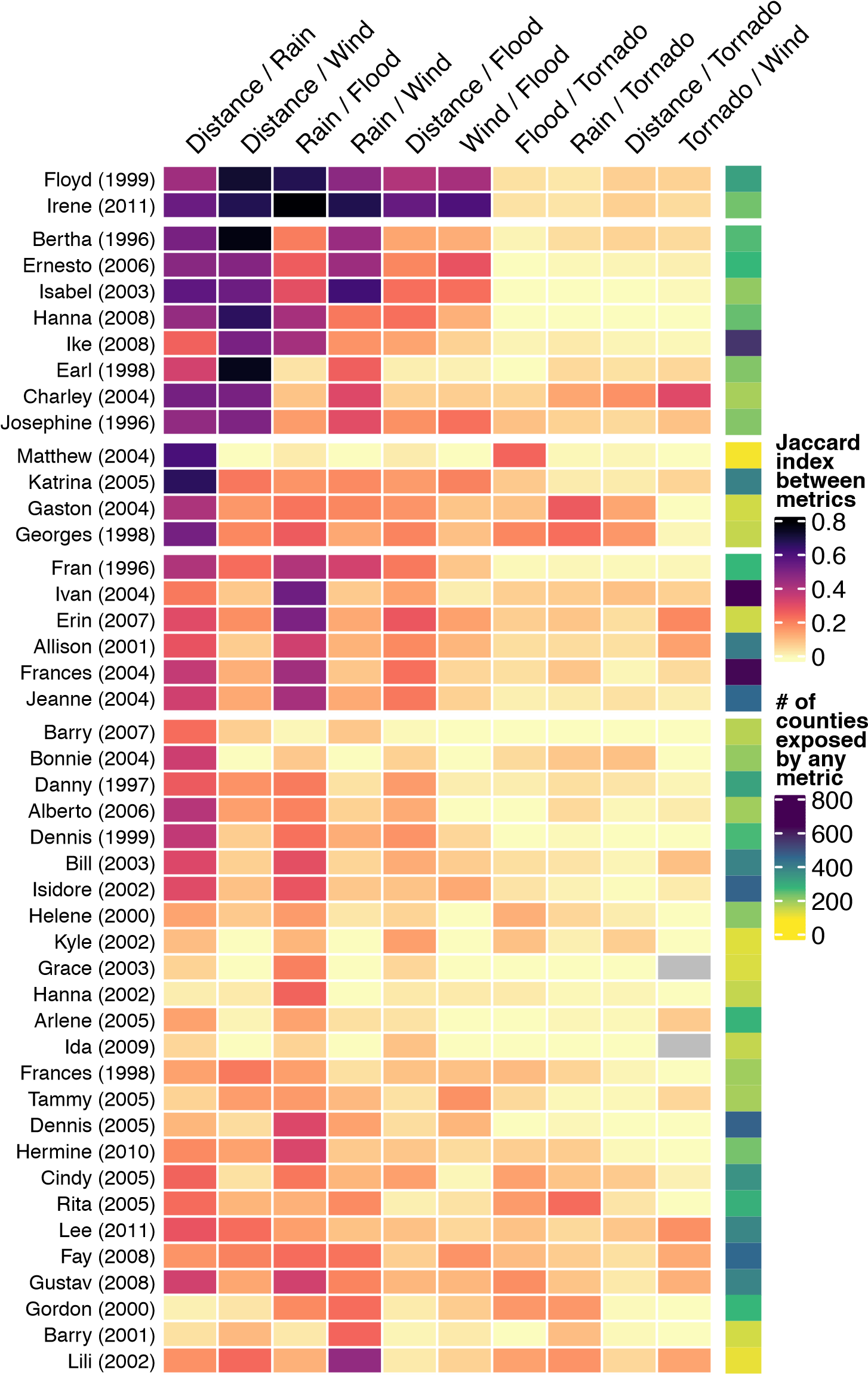


Figure 7: Agreement between exposure classifications based on different exposure metrics for all storms between 1996 and 2011 for which at least 100 counties were exposed based on at least one metric. Each row shows one storm, and the color of each cell shows the measured Jaccard index for each pair of exposure metrics (proportion of counties classified as exposed by both metrics out of counties classified as exposed by either metric). For Grace in 2003 and Ida in 2009, there were no county exposures for either the tornado-based metric or the wind-based metric (indicated by gray squares). Colors to the right of the heatmap show the number of exposed counties based on any of the metrics. Storms are displayed within clusters that have similar patterns based on hierarchical clustering.