Assessing exposure to hurricanes and other tropical storms for epidemiological research

NCAR Group Meeting Presentation

Brooke Anderson

: brooke.anderson@colostate.edu

: www.github.com/geanders

Department of Environmental & Radiological Health Sciences
Environmental Epidemiology Section
Colorado State University

Motivation

Epidemiologic research on tropical storms

Outcomes studied for U.S. tropical storms

- Mortality
  - Direct deaths
  - Indirect deaths
- Cardiovascular events
- Birth rates
- Birth outcomes

Focus of exposure assessment for this study

Multi-storm studies with aggregated daily counts of outcomes (for example, daily deaths by county).

Assessing exposure

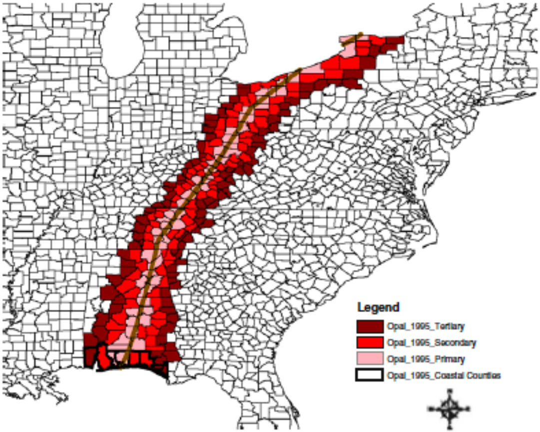
Challenge for epidemiological research

How can we determine whether a county was exposed to a tropical storm?

Previous approaches have varied but include:

- Distance from storm track
- Storm winds above a threshold
- Evacuation orders
- FEMA reports
- Combined metric (property damage, power outages, gas shortages, etc.)

Assessing exposure

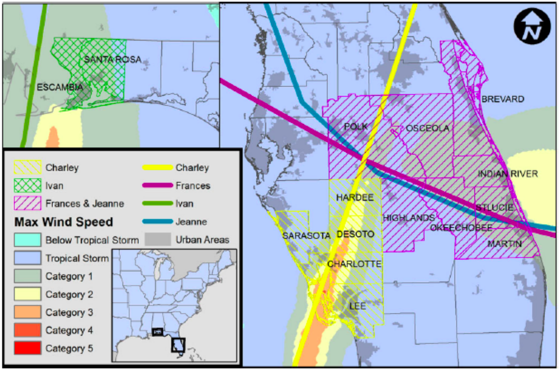


Czajkowski et al. 2011

Example exposure assessment

Czajkowski et al. (2011) classified counties based on distance to storm tracks to study mortality risks.

Assessing exposure



McKinney et al. 2011

Example exposure assessment

McKinney et al. (2011) classified counties based on distance to storm tracks, evacuations, and wind to study mortality risk.

Project aims

Project aims

- Develop exposure classifications of all U.S. Atlantic basin tropical storms, 1988–2011, based on reasonable measurements of tropical storm hazards
- Assess agreement between hazard-based classifications for (1) storm severity and (2) county-specific classification
- Make exposure assessments accessible to other researchers for epidemiological and other impact studies

Hazard-specific metrics

Tropical storm hazard metrics

- Distance from the storm
- High winds
- Rainfall
- Flood events
- Tornado events

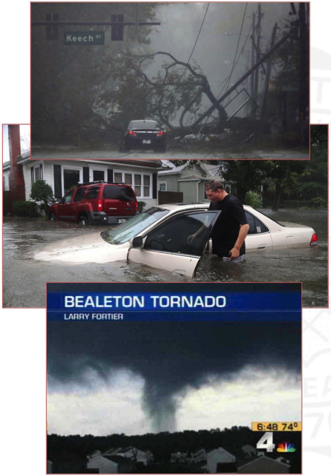
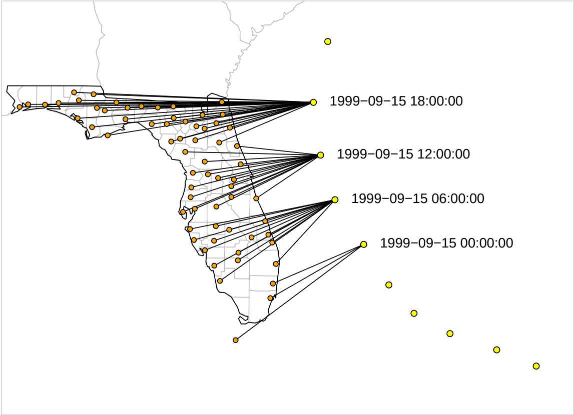


Image sources: Los Angeles Times, NBC

Assessing exposure

Distance from storm

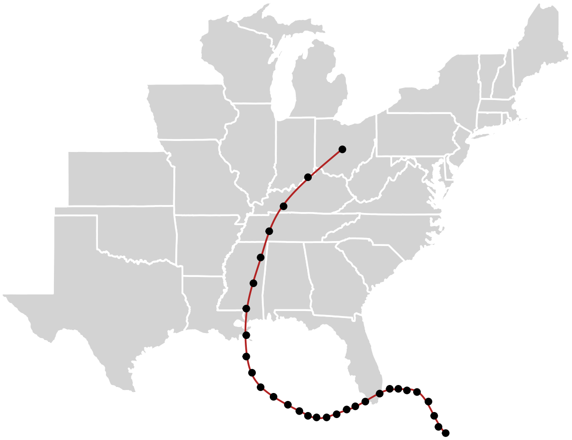
Tropical storm "Best Track" data



Distance metric

We matched storm tracks to county population mean centers to determine the closest approach and date of closest approach of each storm to each county.

Wind exposure



Wind metric

We modeled county winds with a wind model based on a Willoughby et al. paper. This model inputs storm location and maximum wind from best tracks data.

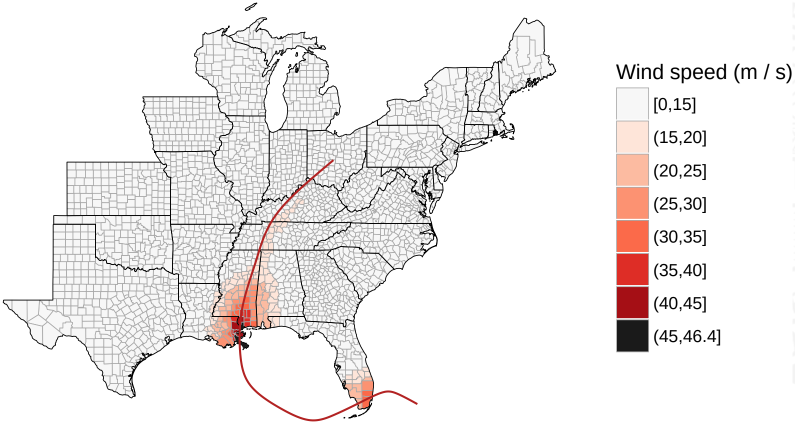
Wind exposure

Details of wind model

We have created an R package that implements this wind model. Full details on on the model implementation are available through one of the package vignettes at `https://cran.r-project.org/web/packages/stormwindmodel/vignettes/Details.html`.

Wind exposure

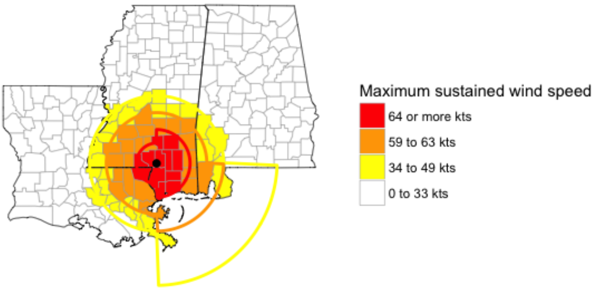
Modeled winds, Katrina, 2005



Wind exposure

Assessment

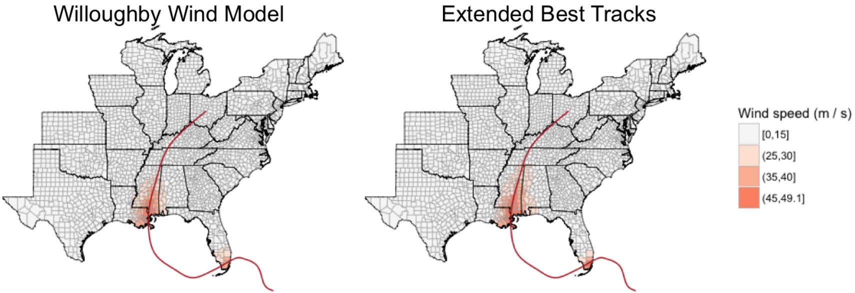
To assess results of the storm wind model, we compared modeled results with wind radii from the Extended Best Tracks for each storm.



Colorado State University

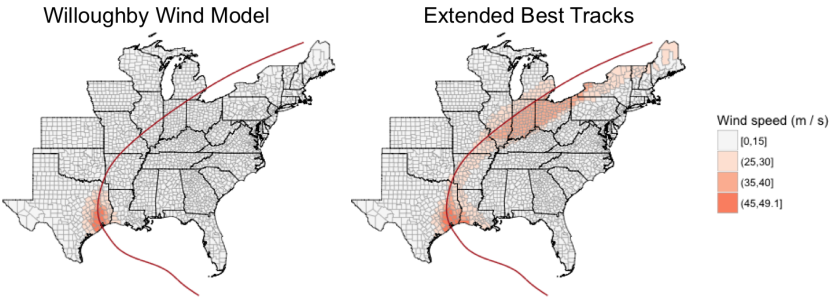
Wind exposure

Comparison of modeled wind versus wind radii, Katrina, 2005



Wind exposure

Comparison of modeled wind versus wind radii, Ike, 2008



Rain exposure

Rain during Tropical Storm Lee

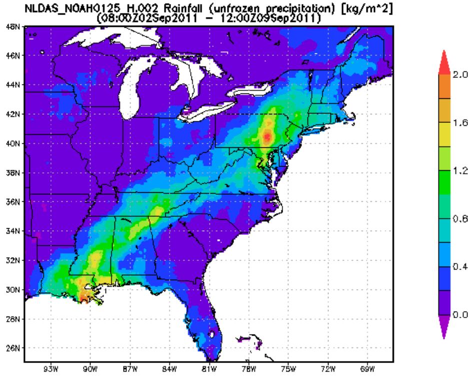
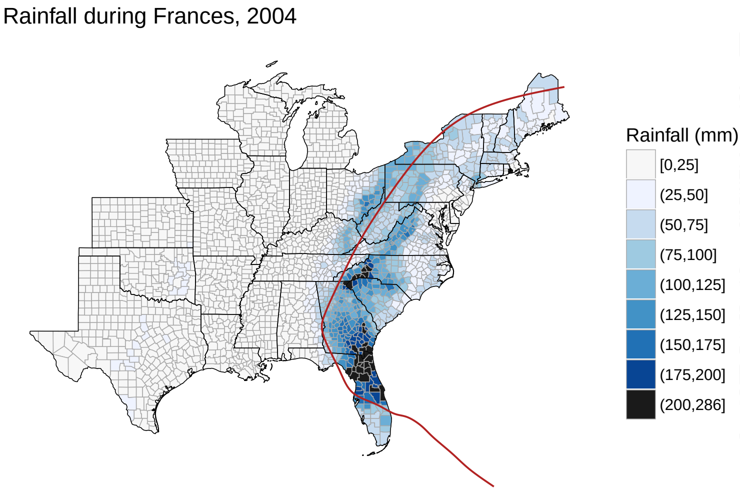


Image source: Goddard Earth Sciences DISC

Rain metric

We used NLDAS-2 precipitation data to assess county rainfall. We summed rain from two days before to one day after the storm. We include a distance threshold for the rain metric.

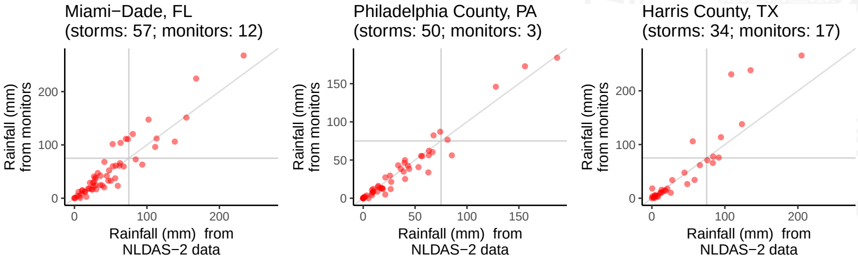
Rain exposure



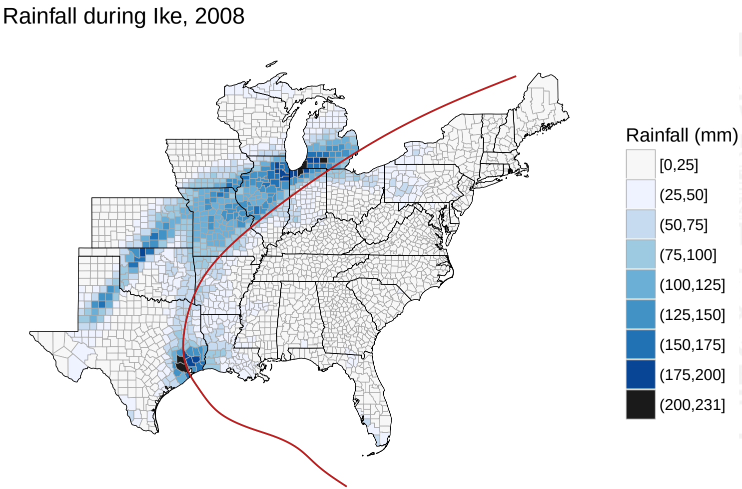
Rain exposure

Assessment

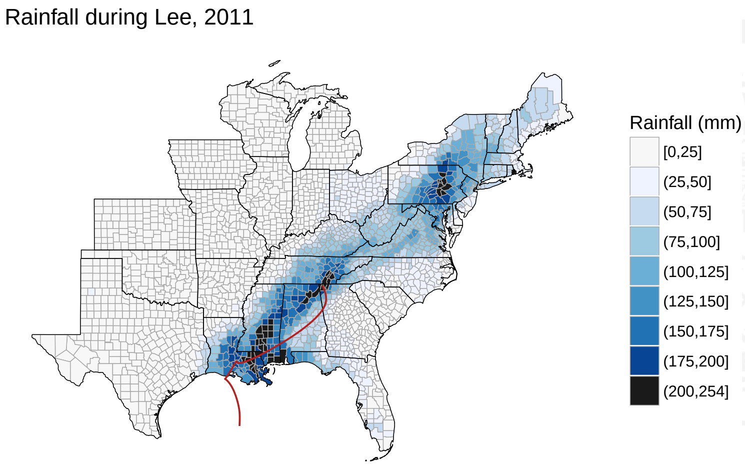
To assess this rain metric, we compared it to rainfall measured at weather stations. X-axis: Rainfall summed for days near storm; y-axis: average of summed rain at each monitor for the same days.



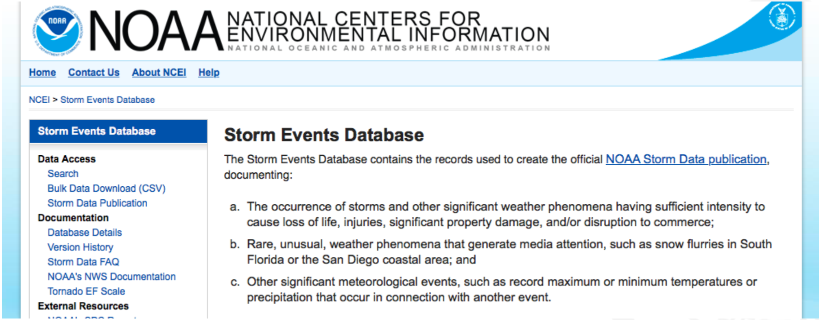
Rain exposure



Rain exposure



Flood and tornado events

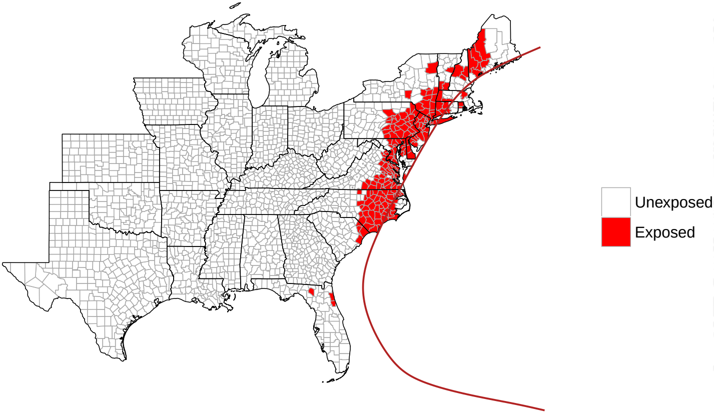


Website: `https://www.ncdc.noaa.gov/stormevents/`

Colorado State University

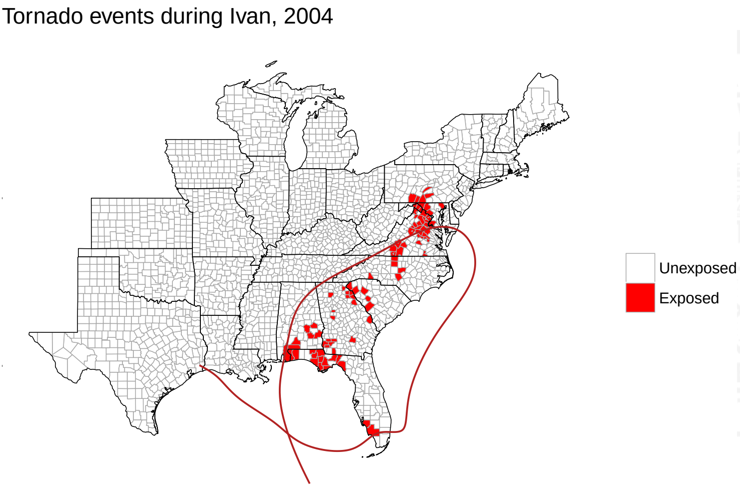
Flood and tornado events

Flood events during Floyd, 1999



Colorado State University

Flood and tornado events



Agreement between exposure metrics

Storm exposure

Exposure metric	Criterial for exposure
Distance	County population mean center within 100 km of storm track
Rain	County received 75 mm or more rain over the period from two days before to one day after the storm's closest approach and the storm passed within 500 km of the county
Wind	Modeled wind speed at county's population mean center met or exceeded 15 m / s during the storm
Flood	Flood event listed with a start date within two days of the storm's closest approach and county within 500 km of storm track
Tornado	Tornado event listed with a start date within two days of the storm's closest approach and county within 500 km of storm track

Storm exposure

Exposure metric	Median number of exposed counties (IQR)	Storm with most counties exposed
Distance	62 (12, 156)	Beryl, 1994 (330)
Rain	32 (4, 133)	Frances, 2004 (464)
Wind	26 (3, 65)	Ike, 2008 (355)
Flood	9 (0, 39)	Ivan, 2004 (317)
Tornado	1 (0, 9)	Ivan, 2004 (91)

*Note: Flood and Tornado events only include storms in 1996–2011. All other event listings cover storms in 1988–2011.

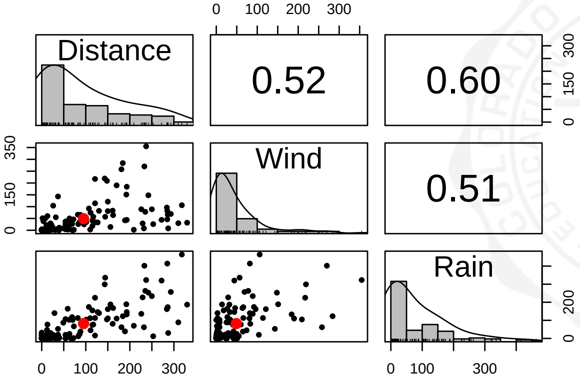
Storm-specific severity

Assessing agreement in storm severity across metrics

- For this assessment, we measured **storm-specific severity** as the number of counties classified as exposed by a given metric for a given storm
- We measured this storm-specific severity for every storm and every metric (1988–2011 for distance, rain, wind; 1996–2011 for flood and tornado)
- We measured the rank correlation (Kendall's τ) between storm severity for each pair of exposure metrics

Storm-specific severity

Agreement in storm severity for distance, wind, and rain metrics



Storm-specific severity

Agreement in storm severity for exposure metrics

	Distance	Rain	Wind	Flood	Tornado
Distance	-	-	-	-	-
Rain	0.60	-	-	-	-
Wind	0.52	0.51	-	-	-
Flood	0.35	0.43	0.32	-	-
Tornado	0.32	0.38	0.34	0.62	-

The table gives Kendall's τ for each pair of exposure metrics. All comparisons that include flood or tornado metrics are limited to storms since 1996.

County-specific classification

Assessing agreement in county classifications

For each storm and each pair of metrics, we measured the probability a county was classified as exposed by one of the metrics conditional on it being classified as exposed by the other metric:

$$Pr(Exposure_Y = 1 | Exposure_X = 1) \tag{1}$$

County-specific classification– Floyd, 1999

	Distance	Rain	Wind	Flood	Tornado
Distance		0.93	0.99	0.72	0.08
Rain	0.44		0.65	0.69	0.04
Wind	0.64	0.88		0.68	0.05
Flood	0.47	0.94	0.68		0.05
Tornado	0.92	0.92	0.92	0.83	

County agreement

Above the diagonal shows the probability that, given the county was exposed based on the metric for that row, it was also exposed based on the metric for that column. Below the diagonal shows the probability that, given the county was exposed based on the metric for that column, it was also exposed based on the metric for that row.

County-specific classification– All storms

	Distance	Rain	Wind	Flood	Tornado
Distance		0.41	0.33	0.18	0.04
Rain	0.47		0.3	0.34	0.07
Wind	0.65	0.5		0.23	0.06
Flood	0.4	0.69	0.27		0.07
Tornado	0.37	0.53	0.28	0.27	

County agreement

Above the diagonal shows the probability that, given the county was exposed based on the metric for that row, it was also exposed based on the metric for that column. Below the diagonal shows the probability that, given the county was exposed based on the metric for that column, it was also exposed based on the metric for that row.

Software

Project software

‘hurricaneexposure‘

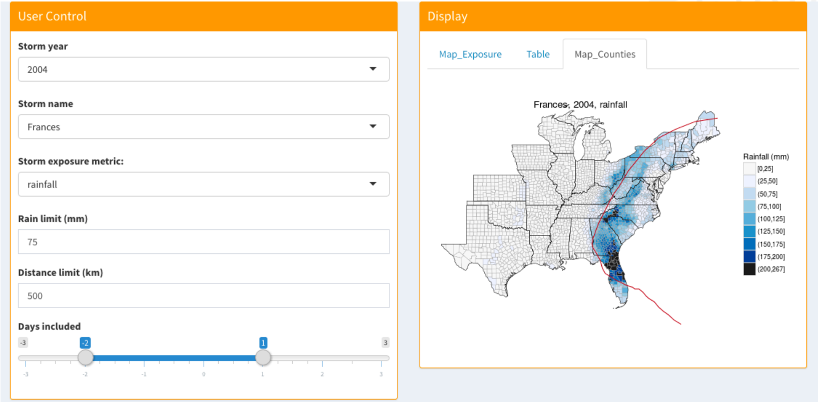
Create county-level exposure time series for tropical storms in U.S. counties. Exposure can be determined based on several hazards (e.g., distance, wind, rain), with user-specified thresholds.
https://github.com/geanders/hurricaneexposure

```
county_rain(counties = c("22071", "51700"), rain_limit = 100,
            start_year = 1995, end_year = 2005, dist_limit = 100,
            days_included = c(-1, 0, 1))
```

```
##        storm_id  fips closest_date storm_dist tot_precip
## 1     Bill-2003 22071   2003-06-30   38.78412      141.1
## 2  Charley-2004 51700   2004-08-14   43.01152      136.2
## 3    Cindy-2005 22071   2005-07-06   32.21758      113.2
## 4    Floyd-1999 51700   1999-09-16   46.50729      207.5
## 5  Isidore-2002 22071   2002-09-26    6.37844      249.0
## 6  Katrina-2005 22071   2005-08-29   36.88933      196.2
```

Project software

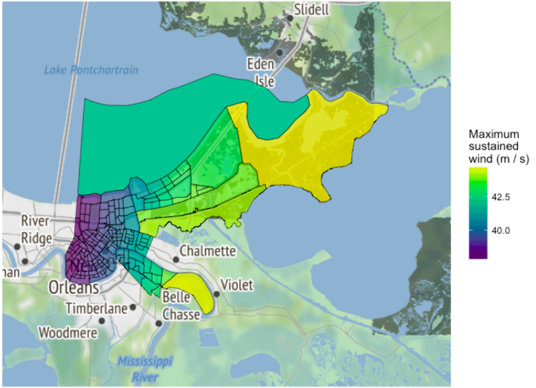
Web application interface to ‘hurricaneexposure‘



Project software

'stormwindmodel'

Model storm winds from Best Tracks data at U.S. locations. Includes modeling sustained and gust winds, as well as duration of sustained and gust winds above a specified threshold. On CRAN.



Project software

‘noaastormevents‘

Download and explore listings from the NOAA Storm Events database. Includes the ability to pull events based on a tropical storm, using events listed close in time and distance to the storm’s tracks.
`https://github.com/zailchen/noaastormevents`

‘countytimezones‘

Convert time-stamps from UTC to local time zones for U.S. counties based on county FIPs. Facilitates merging weather observations with locally measured data, including health outcomes. On CRAN.

‘countyweather‘

Download weather monitor data through NOAA API by U.S. county. Includes functions to map available monitors for each county. On CRAN.

Future work

Future work

- Epidemiological study of mortality risk and tropical storm hazards
- Exploring alternative measures of flood and tornado exposure
- Improving wind model for extra-tropical storms
- Incorporating metrics of infrastructure damage

R tools for open science

Open data / reproducible research

AJE
American Journal of Epidemiology

Vol. 163, No. 9
DOI: 10.1093/aje/kwj093
Advance Access publication March 1, 2006

Commentary

Reproducible Epidemiologic Research

OPEN ACCESS Freely available online
PLoS BIOLOGY

Community Page

The Open Knowledge Foundation: Open Data Means Better Science

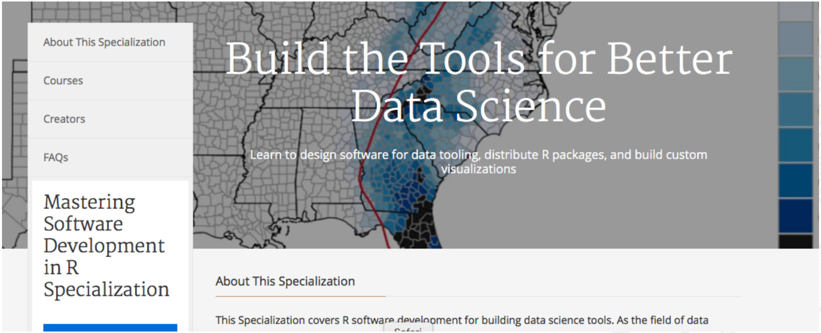
Jennifer C. Molloy*
Department of Zoology, University of Oxford, Oxford, United Kingdom

Annals of Internal Medicine | ACADEMIA AND CLINIC

Reproducible Research: Moving toward Research the Public Can Really Trust

Christine Laine, MD, MPH; Steven N. Goodman, MD, PhD, MHS; Michael E. Griswold, PhD; and Harold C. Sox, MD

Software as a research product



Courses: `https://www.coursera.org/specializations/r`
Course book: `https://bookdown.org/rdpeng/RProgDA/`

R package repositories

Sharing R packages

R packages can be shared with others through:

- By sharing a `.tar.gz` file
- Through the Comprehensive R Archive Network (CRAN)
- Through GitHub (packages can be installed from GitHub with the `install_github` function from the `devtools` package)
- Through a `drat` repository, hosted locally or through GitHub pages

Software vignettes

Documenting R Software

R packages can include **vignettes**, which provide tutorials or documentation for users. These vignettes can be written in RMarkdown, which allows a mix of executable R code and text, figures, tables, etc.

Example vignettes

The `stormwindmodel` package has two vignettes:

- "Overview" vignette: `https://cran.r-project.org/web/packages/stormwindmodel/vignettes/Overview.html`
- "Details" vignette: `https://cran.r-project.org/web/packages/stormwindmodel/vignettes/Details.html`

Shiny web applications

Shiny web applications

Web applications that run R code can be created using Shiny. You can use these to "wrap" an R package and make it accessible to non-R users. Shiny web apps can:

- Be freely hosted through `https://www.shinyapps.io`
- Allow users to upload and download data and create figures
- Be extensively customized

Help with Shiny

- Lots of tutorials are available through RStudio: `http://shiny.rstudio.com/`
- The Shiny Gallery provides simple to advanced examples, with code: `https://shiny.rstudio.com/gallery/`

Peer-review options

Peer-review for R software

R packages do not undergo peer review for CRAN and most other repositories. However, there are a few avenues for partial or full peer review of packages:

- *ROpenSci*: Includes the option for community-contributed packages. These undergo peer review through GitHub. The full source code is reviewed and comments must be addressed before the package is added to the ROpenSci repository
- *The Journal of Open Source Software*: Very short articles, peer review of code
- *The R Journal*, *Journal of Statistical Software*: Peer-review of articles, but may include some suggestsion for software