Interim Report

Development of Wind-Driven Rain Climatology and Coincidental Wind Speed Return Period Maps for Florida and Surrounding Coastal Areas.

Prepared by:

Forrest J. Masters, Ph.D., P.E. (FL)
Prof. and Assoc. Dean for Research and Facilities
Herbert Wertheim College of Engineering
University of Florida
(Principal Investigator)

Arthur T. DeGaetano, Ph.D.
Professor Earth and Atmospheric Sciences
Director, NOAA Northeast Regional Climate Center
Cornell University

Brian N. Belcher
Research Support Specialist
NOAA Northeast Regional Climate Center
Cornell University

February 28, 2022

Table of Contents

1. Introduction

Rain water intrusion, in its various forms, persists as one of the most costly and prolific forms of damage to buildings in the United States. For exterior walls, rain water intrusion is closely related to the presence of wind to cause rain water impingement on walls. More importantly, the coincidental presence of wind also causes wind pressure differentials that force water behind claddings and into or through wall assemblies or components. This mechanism of rain water intrusion is prolific and is the cause of substantial economic impact, loss of building resiliency and useful life, and even has structural safety implications.

The development of wind-driven rain climatology and coincidental wind speed return period maps are ongoing for Florida and surrounding coastal areas of the Southeast United States since November 2021. This project will produce deliverables that are key to a better understanding of risk associated with coincident wind/rain events. Many of the tasks undertaken to meet project goals and objectives are performed in parallel when possible, and are therefore at different stages of completion. A summary of these tasks include:

- acquiring weather data and information at various timescales (one-minute and hourly),
- assessing the quality of the weather data and correcting if necessary,
- adjusting raw wind data to baseline design conditions,
- performing extreme value analysis to compute coincident wind speed/rainfall intensity occurrence probabilities,
- comparing results of extreme value analysis for one-minute and hourly timescales, and
- creating extreme wind speed risk maps for different rainfall intensity thresholds.

Details regarding completed and in-progress tasks are discussed in the appropriate sections that follow. In the final section, the status of all tasks are summarized and future tasks that are required for completing project goals and objectives are presented.

2. Data Acquisition

a. One-minute weather data

One-minute wind/precipitation data were obtained for 243 weather stations within Florida and other coastal and nearby states in the southeastern United States (Fig. 1) from the National Centers for Environmental Information (NCEI, https://www.ncei.noaa.gov/pub/data/asos-onemin/). Datasets for 3-second peak wind speed/direction (DSI-6405) and accumulated precipitation (DSI-6406) over each minute are available from as early as 2000.

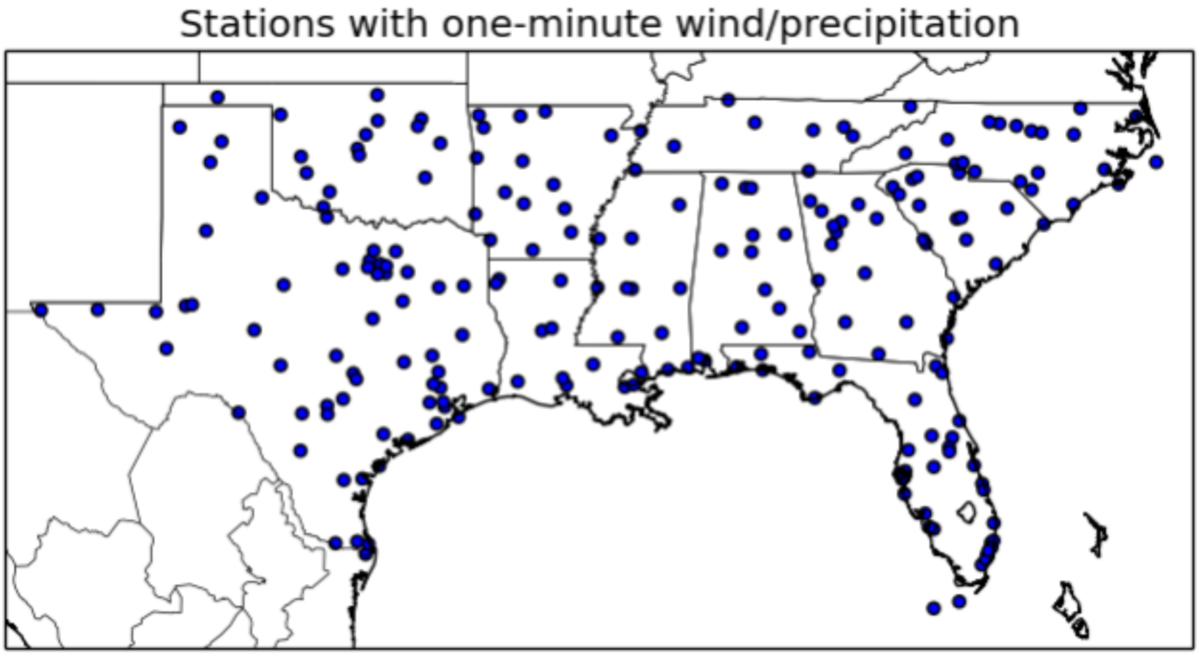

Figure 1. Region of interest and weather station locations with one-minute wind/precipitation data available.

These one-minute data are originally in plain text format, contain gaps in files for missing data, and have changes or errors in column formatting throughout the periods of record. All of these dataset characteristics make scientific analysis on these data extremely difficult. Instead of working with the data in its original format, the text files were converted to HDF5 format and gaps in data were filled with missing data identifiers. Errors in text file formatting were also identified during this process, and associated data were set to missing. These changes in data formats allows for more efficient and reliable data access during subsequent analyses.

b. Hourly weather data

Hourly wind/precipitation data were obtained for each of the 243 weather stations from the Applied Climate Information System (ACIS, https://www.rcc-acis.org/). At this temporal resolution, periods of record can start as early as 1948 for many stations.

Like the one-minute data, these hourly data will be used to calculate coincident extreme wind speed / rainfall intensity recurrence probabilities, and allows for direct comparison of magnitudes between the two timescales. These hourly data also assist during the quality control of one-minute data.

c. Data used to assist in quality control

Some quality control methods rely on comparison of questionable values to other datasets when identifying erroneous data. The following data were acquired and used for this purpose during automated QC methods:

- ASOS hourly reports (available from ACIS)
- Radar-guided daily precipitation (available from ACIS)

When manual QC methods were necessary, the following data assisted our assessment:

- Historical METARs from https://mesonet.agron.iastate.edu/request/download.phtml;
- Adjacent one-minute data from NCEI DSI-6405/6406 (+/- 5 minutes before and after data in question).

3. Quality Control Methods

a. One-minute wind QC methodology

Aside from the Automated Surface Observing System (ASOS) processing algorithms, the one-minute data do not undergo further quality control before being archived at NCEI. Therefore, it was necessary to assess the quality of one-minute wind data, and remove errors when appropriate. The following automated checks were performed and data replaced with missing identifiers if:

- values were physically impossible (e.g. direction not in 0-360, negative wind speeds);
- one-minute peak wind speeds exceed max gust for the day by more than 5 knots;
- one-minute peak wind speeds exceed 30 kts, but max daily gusts are not reported for that day (derived from hourly reports). This represents an inconsistency between the one-minute data and the higher-quality hourly reports.

Following the above automated checks, all remaining one-minute peak wind speeds that exceed 80 kts were verified manually. Consistency of weather conditions during adjacent minutes were used during this verification process. Sometimes erroneous data that made it through automated checks are found here – for instance, when both one-minute and hourly reports contain the same error.

b. One-minute precipitation QC methodology

The quality of one-minute precipitation data was also assessed, and errors removed. Automated checks assisted in finding errors and replacing them with missing identifiers if:

- values were physically impossible (negative precipitation amounts)

- one-minute values exceed observed (radar-guided) daily precipitation by more than 0.05 inches.

After the automated checks, all remaining one-minute precipitation amounts that exceed 0.30 inches were verified manually. Like the wind speed checks, weather conditions during adjacent minutes were used for verification of these extreme values. These manual checks are required, for example, on days when observed precipitation grids are not available for one reason or another.

c. Hourly weather QC methodology

The hourly weather data is of higher quality than one-minute data. However, it was clear that errors exist in the hourly data when they were used to assist in the quality assessment of one-minute data. Sometimes the same errors occurred in both one-minute and hourly data, forcing manual checks of those extreme data.

While the hourly data has been acquired, quality assessment has not yet been performed on the hourly data directly. The same QC methods used on one-minute data will be applied to hourly data, when possible, before the hourly data is used in extreme value analysis.

4. Weather stations available for analysis

After applying the quality control methods to the one-minute data, there was a better understanding of the amount of useable data at each station. The valid periods of record ranged from a couple of years, to over 20 years (Fig. 2).

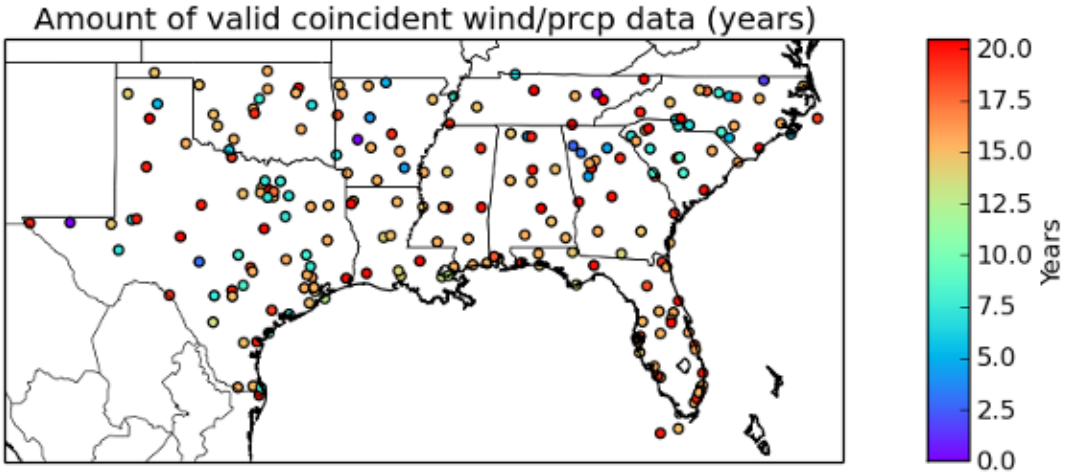

Figure 2. Amount of valid coincident wind/precipitation data after QC of one-minute data.

Stations with over 15 years of coincident wind/precipitation data are deemed 'primary' stations for this project, while those with less than 15 years are considered 'secondary' stations. There are

137 stations that qualify as primary stations, with good spatial coverage across our region of interest (Fig. 3). The primary stations serve as stations included in subsequent extreme value analyses, while the secondary stations are retained in case those data are needed for any reason during future analyses.

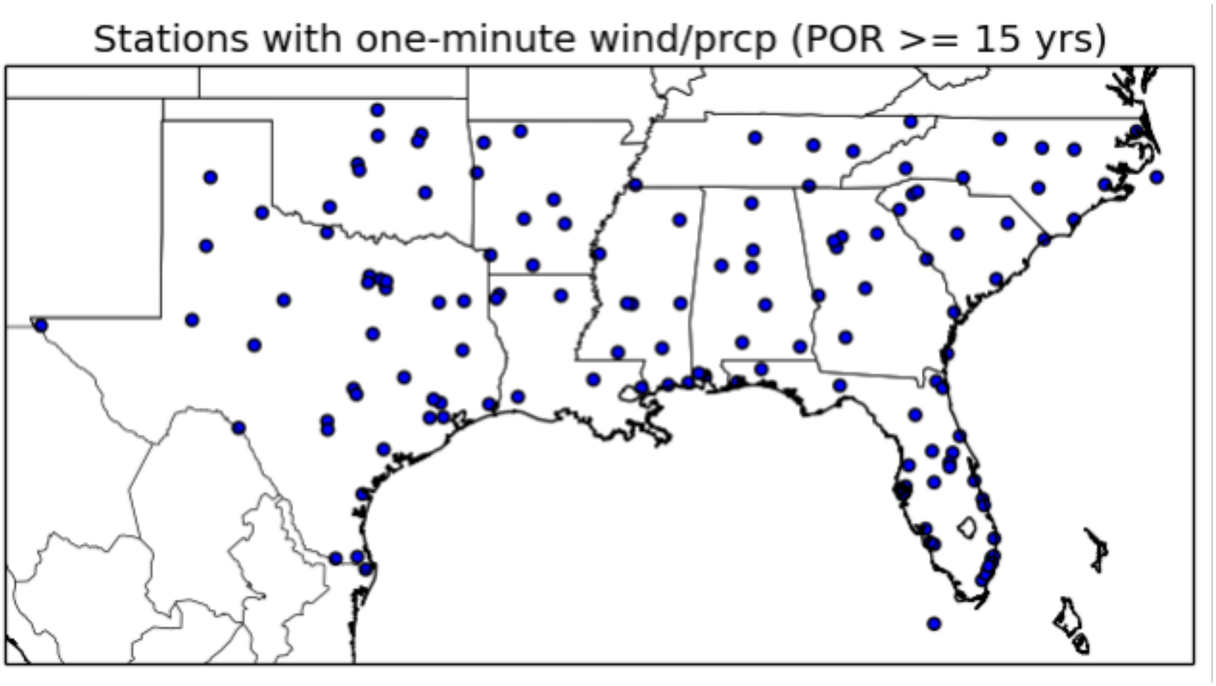

Figure 3. Primary stations used in project analyses. These stations have at least 15 years of valid coincident wind/precipitation data after quality control.

5. Siting information and wind adjustment factors

In contrast to the precipitation data, the wind speed data collected by the ASOS platform must be pre-processed to standardize the reported values to a common metadata format, i.e., gust duration, height, and terrain exposure condition. In the early 2000s, ASOS platforms operated cup anemometers that reported 5 s block average gusts. Between 2003-2009, the National Weather Service replaced the cup anemometers at most sites with ultrasonic anemometers that report 3 s moving average gusts. A cup anemometer behaves like a mechanical filter, i.e., it does not instantly respond to changes in wind speed like its ultrasonic counterpart (which has no moving parts). Thus, an appropriate adjustment had to be made to correct for the instrument response characteristics. Further, the observation height of the anemometer can deviate from the 10 m standard height, and the surface roughness of the upwind fetch (terrain) varies significantly. Collectively, these variations can cause reported gust values to underreport the surface wind field intensity by as much as 40%.

The method described in Masters et al. (2010) was applied to perform this conversion, using the equivalent of ASCE 7-22 (2022) basic wind speed conditions (10 m height, 3 s gust, open exposure conditions). Given the duration of the study (up to 20 years), year-specific conversion factors are being developed for 16 wind directions to account for the era of the anemometer, terrestrial growth (e.g., tree canopy growth), new construction (e.g., new terminals or densification of building stock outside the airport), and other changes to the upwind fetch that can modify the surface roughness.

Progress to date includes:

- Acquisition of one-minute wind speed data between 2000-2021 for 246 stations (NCEI Data Set 6405), which amounts to 235 GB of raw data
- Computation of directionally dependent mean gust factors (ratios of peak gusts to the correspond mean of the record containing the gust) for all years when only one instrument was in service, and the intervals in which the cup or ultrasonic anemometer operated during a “changeover year”
- Computation of wind speed conversion factors based on the mean gust factor, accounting for station’s observation height and the observed wind speed

The investigators are now developing a heuristic approach to estimate the mean gust factor during years when data are scarce or not available, with the goal of testing the conversion algorithm on the entire dataset next. (Note: the figures that follow show the raw data.)

6. Extreme value analysis

a. Method to calculate wind speed return periods for coincident rainfall intensities

The process of calculating wind speed return periods conditional on rainfall intensities begins with grouping the one-minute data into bins based on observed rainfall amount exceeding a given threshold. For instance, the first bin includes all minutes that observe precipitation amounts greater than or equal to 0.01 inch. The second bin includes all minutes that observe precipitation amounts greater than or equal to 0.02 inch. Multiple sets of data continue to be constructed based on incrementally higher thresholds until the maximum observed one-minute precipitation amount is reached (i.e. if the maximum one-minute precipitation is 0.23 inches, this would results in 23 separate sets of minutes to be analyzed for this station).

Then within each of these bins, the maximum wind speed is identified and retained. Data within an 8-day window centered on this observation are eliminated. The highest wind speed in the remaining data series is retained and data within an 8-day window centered on this value omitted. This process repeats until all available observations are either retained or excluded. The resulting data series contains the highest wind speeds within independent 8-day windows over the period of record at the station associated with precipitation equal to or exceeding a given amount. Although all wind events within a particular precipitation bin are independent (i.e. are separated by at least 4-days), there is not requirement of independence between events in different precipitation bins. Thus, it is possible that the same wind speed observation is included in multiple cumulative bins and separate bins could include wind speeds separated by less than 4-days.

The set of wind speeds within each cumulative bin is then fit with a Gumbel distribution (chosen based on previous pilot studies). The probability density function of the Gumbel distribution is given by two parameters, shape and scale such that

$$f(x) = \frac{1}{\beta} e^{\frac{x-\mu}{\beta}} e^{-e^{\frac{x-\mu}{\beta}}}$$

where μ is the location parameter and β is the scale parameter. For each station-bin combination the values of μ and β are fit to the data using the maximum likelihood method via the scipy stats gumbel_r.fit algorithm. Once fit, the wind speeds corresponding to recurrence intervals of 3-months, 6-months, and 1-, 5-, 10-, 25-, 50-, 100-, and 500-years were calculated using the scipy gumbel_r.sf algorithm.

b. Preliminary results used for data verification and development (non-adjusted winds)

The developed method was applied using raw (non-adjusted) wind speed data at all primary stations. At this point during the research process it was advantageous to view individual contour plots for each station (Fig. 4) and wind speed maps that focus on individual station results (Fig. 5) before applying any wind adjustment factors, smoothing techniques, or spatial interpolation required for the final regional contour maps. From these intermediate plots and maps we were able to verify our application of methods to all stations, and begin preparing software that will be used in final regional contour maps. This process also allowed us to identify any locations that may appear as outliers compared to others in the region, and double-check the recently quality-controlled data for those locations. Finally, this intermediate step will also allow for comparison of results before and after applying any wind adjustment factors.

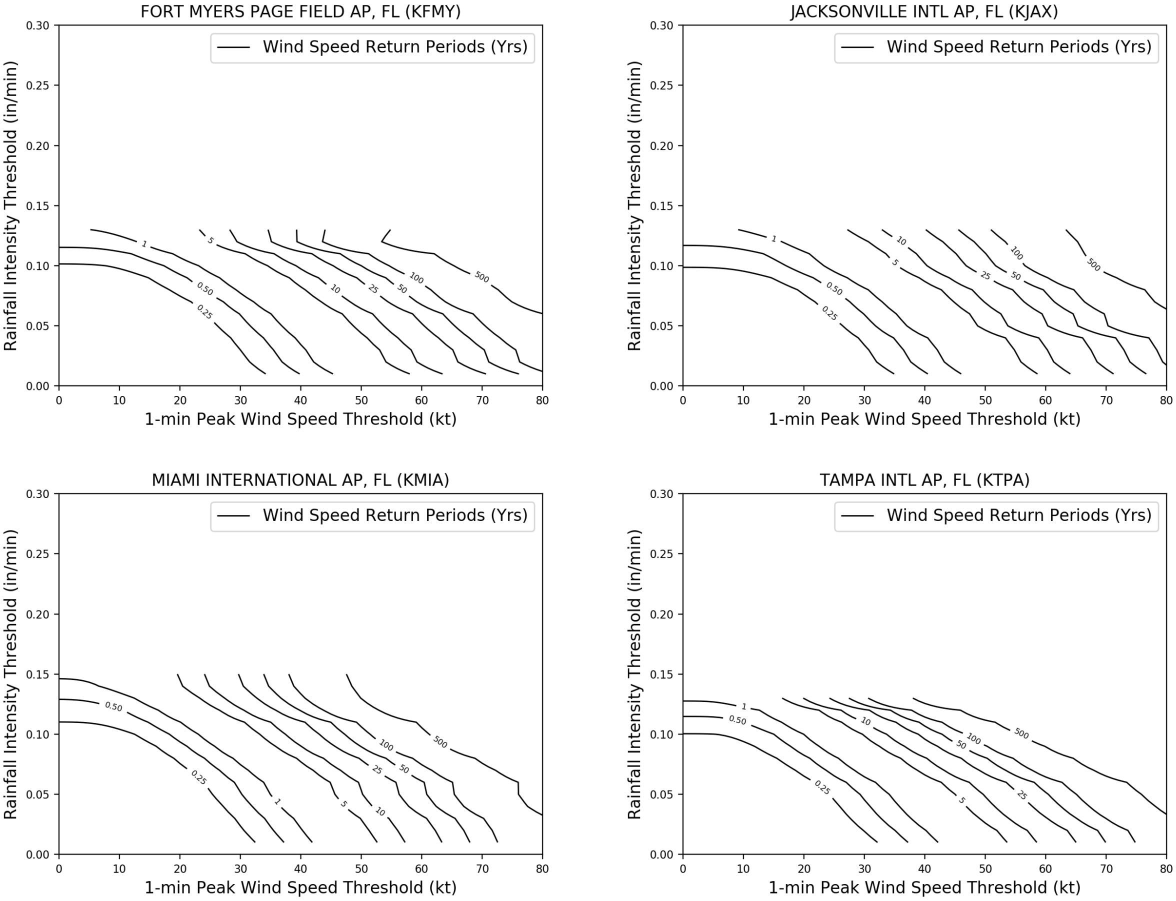

Figure 4. Wind speed return periods associated with rainfall intensity thresholds at four stations in FL. Contour lines end at 0.13 - 0.15 inch min^{-1} due to low number of independent events associated with the most extreme rainfall thresholds. Results shown are unsmoothed, with sometimes jagged contour lines resulting from sampling variability between precipitation bins.

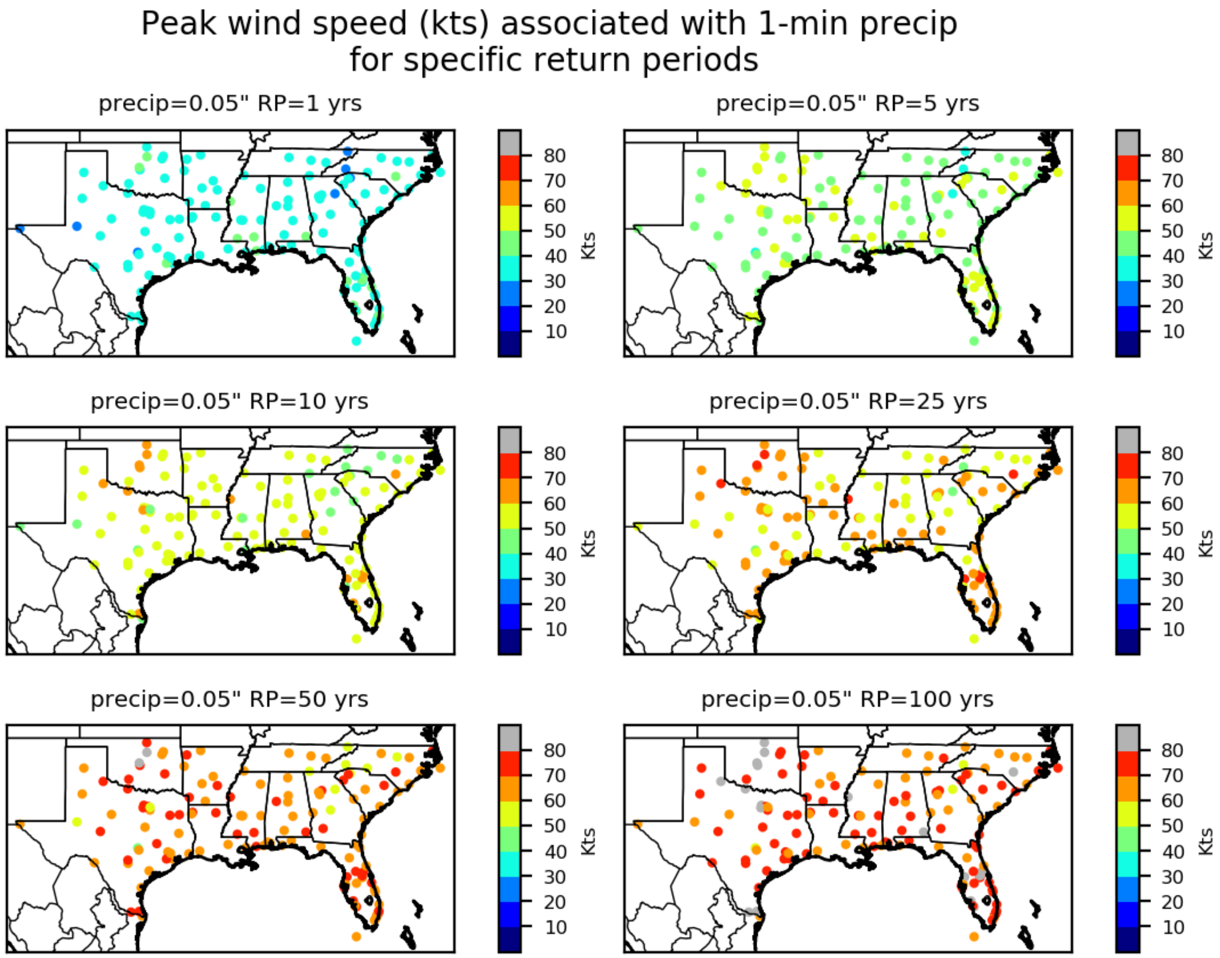

Figure 5. Peak wind speeds (kts) having return periods (RP) of 1, 5, 10, 25, 50 and 100 years, when rainfall intensities exceed 0.05 inch min^{-1} . Results are provided for each primary station location without applying wind adjustment factors or spatial interpolation.

c. Estimating wind speed return periods for extreme rainfall intensities

The number of independent events observed for the most extreme rainfall intensities are often too low to reliably fit theoretical distributions and calculate wind speed return periods. For stations in Florida and the surrounding region, there is often less than 15 independent events associated with extreme rainfall intensities (rates that exceed 0.12 - 0.15 inches min^{-1}). As a result, wind speed return periods become unavailable above these rainfall intensities using current methodology (Fig. 4). Other methods are required to estimate wind speed return periods associated with the most extreme rainfall intensities.

Preliminary results on a few stations from our region of interest (Miami, Tampa, Oklahoma City) suggested a potential relationship between fitted distribution parameters and rainfall intensity. Such a relationship would allow for extrapolation of distribution parameter estimates for extreme rainfall intensities. However after viewing results from methodology applied to all primary stations, a lack of consistency in the distribution parameter / rainfall intensity relationship across all stations convinced us to explore other options to solve this problem.

Currently, we are investigating a new approach to fill these missing estimates. This approach starts with performing the original analysis of calculating wind speed return periods conditional on rainfall intensity thresholds. Then, a separate second analysis is performed to calculate rainfall intensity return periods conditional on wind speed thresholds. This separate analysis is similar to the first, except the one-minute data are binned by wind speed and distributions are fit to independent rainfall intensity events in order to calculate conditional rainfall intensity return periods. Our initial approach uses the Gumbel distribution to fit precipitation data when calculating rainfall intensity return periods.

When overlaying results from the two separate analyses, we see consistency between the wind speed return periods and rainfall intensity return periods (Fig. 6). The final step of the process involves fitting a function to the combined sets of common return periods between the two analyses (Fig. 7). This approach looks promising and makes climatological sense. We are continuing to investigate the best distribution to fit to the rainfall intensity during the additional analysis. Additionally, we are investigating the most appropriate function(s) to fit to the common return period data in the final step, which will serve to both fill missing data and smooth results.

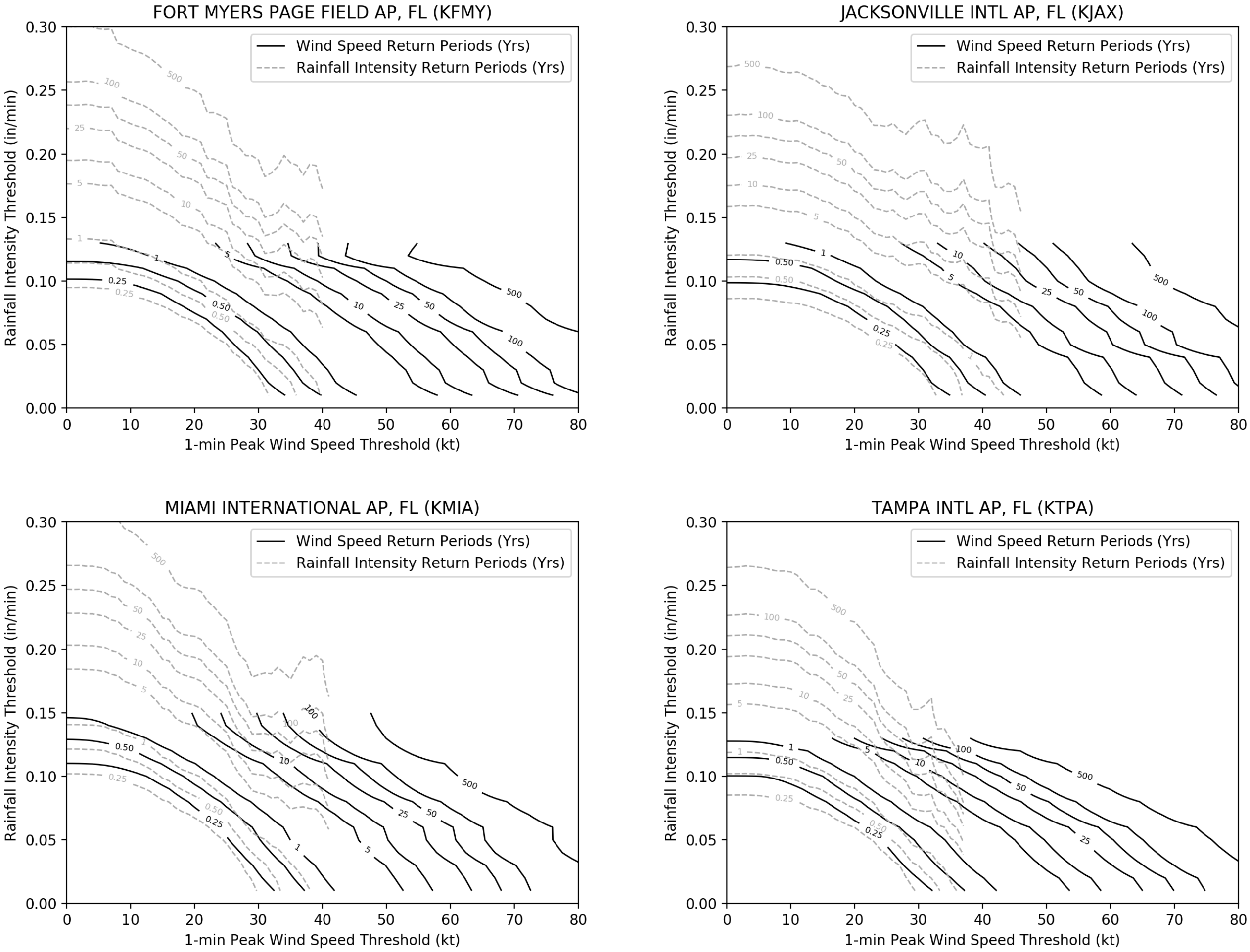

Figure 6. Wind speed return periods (solid) associated with rainfall intensity thresholds (y-axis) and rainfall intensity return periods (dashed) associated with peak wind speed thresholds (x-axis) at four stations in FL.

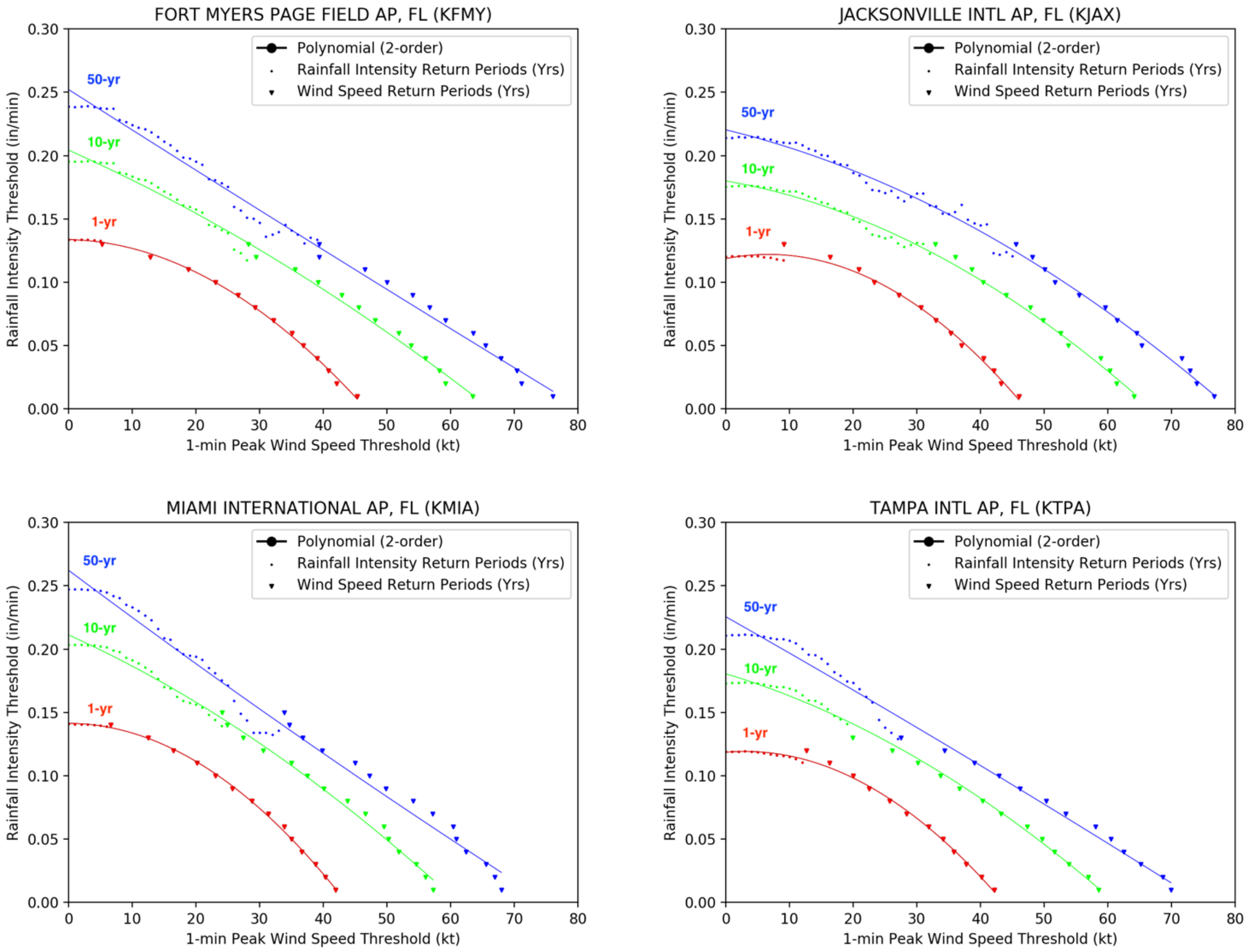

Figure 7. Polynomial functions (2nd order) fit to 1, 10 and 50-year return periods when combining two separate analyses at four stations in FL. Wind speed return periods are shown from the original analysis (triangles) and rainfall intensity return periods are shown from a second analysis (circles).

7. Summary and future tasks required to complete project goals

Here we summarize the status of various tasks discussed above and future tasks required to complete the projects goals and objectives.

Data acquisition

- Obtain one-minute wind and precipitation data from all available Automated Surface Observing System sites in Florida and adjoining states. **(Complete)**
- Obtain hourly data from same sites for comparison. **(Complete)**

Data quality control and archiving for analysis

- Convert format of one-minute data from text to HDF5 **(Complete)**
- Quality control and removal of bad data from one-minute dataset **(Complete)**
- Quality control and removal of bad data from hourly dataset **(In-progress)**
  Note: Hourly data is obtained, and quality control of these data are in progress.

Siting information and wind adjustment factors

- Acquire NCEI Data Set 6405 data between 2000-2021 **(Complete)**
- Compute directionally dependent mean gust factors for all years when only one instrument was in service, and the intervals in which the cup or ultrasonic anemometer operated during a "changeover year" **(Complete)**
- Compute wind speed conversion factors based on the mean gust factor, accounting for station's observation height and the observed wind speed **(Complete)**
- Develop a heuristic approach to estimate the mean gust factor during years where data are scarce or not available **(In-progress)**
- Compute the wind speed conversion factors based on the mean gust factor, accounting for station's observation height and the observed wind speed
- Test the conversion algorithm on the entire dataset, comparing the rainfall intensity threshold maps referenced to raw and standardized datasets

Extreme value analysis

- Develop coincident wind-rain partial duration series from one-minute data series and fit appropriate extreme value distribution to data from each station. **(In-progress)**
- Compute coincident rain intensity wind speed occurrence probabilities. **(In-progress)**
  Note: Methods developed for low-to-moderate rain intensity amounts are complete, as these contain enough wind speed values to reliably fit distributions. Final data series and distribution fitting will occur after wind adjustment factors are applied. We are currently developing methods to estimate wind speed occurrence probabilities associated with the more extreme rainfall intensity amounts.

Comparison with hourly data

- Compare magnitudes of coincident extreme wind speed rainfall intensity recurrence probabilities based on hourly and one-minute observations. **(Not started)**
- Determine trade-offs between the longer period of record available for the hourly samples and the higher sampling frequency (but shorter record) of the one-min data. **(Not started)**
  Note: These tasks will be addressed following quality control of hourly data and completion of the one-minute extreme value analysis.

Develop risk maps

- Mapping software will be used to produce extreme wind speed risk maps for different rainfall intensity thresholds. Station specific values will be spatially interpolated and contoured. **(In-progress)**
  Note: Regional maps (focusing on individual station results) were produced in an intermediate format to view preliminary (non-adjusted wind) results. We intend to update this software to modify these regional maps by including spatial interpolation and contouring in the final versions.

8. References

ASCE 7-22, 2022. Minimum Design Loads for Buildings and Other Structures. American Society of Civil Engineers, Reston, Virginia.

Masters, F.J., P.J. Vickery, P. Bacon, and E.N. Rappaport, 2010: Toward objective, standardized intensity estimates from surface wind speed observations. Bulletin of the American Meteorological Society, 91, 1665-1682.