

## NOAA Atlas 14

# Precipitation-Frequency Atlas of the United States

Volume 9 Version 2.0: Southeastern States (Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi)

Sanja Perica, Deborah Martin, Sandra Pavlovic, Ishani Roy, Michael St. Laurent, Carl Trypaluk, Dale Unruh, Michael Yekta, Geoffrey Bonnin

U.S. Department of Commerce

National Oceanic and Atmospheric Administration

National Weather **Service** 

> Silver Spring, Maryland, 2013

NOAA Atlas 14

## Precipitation-Frequency Atlas of the United States

Volume 9 Version 2.0: Southeastern States (Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi)

Sanja Perica, Deborah Martin, Sandra Pavlovic, Ishani Roy, Michael St. Laurent, Carl Trypaluk, Dale Unruh, Michael Yekta, Geoffrey Bonnin

U.S. Department of Commerce

National Oceanic and Atmospheric Administration

National Weather Service

Silver Spring, Maryland, 2013

Library of Congress Classification Number G1046 .C8 U6 no.14 v.9 (2013)

## **Table of Contents**



#### **1. Abstract**

NOAA Atlas 14 contains precipitation frequency estimates for the United States and U.S. affiliated territories with associated 90% confidence intervals and supplementary information on temporal distribution of heavy precipitation, analysis of seasonality and trends in annual maximum series data, etc. It includes pertinent information on development methodologies and intermediate results. The results are published through the Precipitation Frequency Data Server [\(http://hdsc.nws.noaa.gov/hdsc/pfds\)](http://hdsc.nws.noaa.gov/hdsc/pfds).

The Atlas is divided into volumes based on geographic sections of the country. The Atlas is intended as the U.S. Government source of precipitation frequency estimates and associated information for the United States and U.S. affiliated territories.

#### **2. Preface to Volume 9**

NOAA Atlas 14 Volume 9 contains precipitation frequency estimates for selected durations and frequencies with 90% confidence intervals and supplementary information on temporal distribution of heavy precipitation, analysis of seasonality and trends in annual maximum series data, etc., for the six southeastern states of Alabama, Arkansas, Florida, Georgia, Louisiana, and Mississippi. The results are published through the Precipitation Frequency Data Server [\(http://hdsc.nws.noaa.gov/hdsc/pfds\)](http://hdsc.nws.noaa.gov/hdsc/pfds).

NOAA Atlas 14 Volume 9 was developed by the Hydrometeorological Design Studies Center within the Office of Hydrologic Development of the National Oceanic and Atmospheric Administration's National Weather Service. Any use of trade names in this publication is for descriptive purposes only and does not imply endorsement by the U.S. Government.

**Citation and version history.** This documentation and associated artifacts such as maps, grids, and point-and-click results from the PFDS are part of a whole with a single version number and can be referenced as:

Sanja Perica, Deborah Martin, Sandra Pavlovic, Ishani Roy, Michael St. Laurent, Carl Trypaluk, Dale Unruh, Michael Yekta, Geoffrey Bonnin (2013). NOAA Atlas 14 Volume 9 Version 2, *Precipitation-Frequency Atlas of the United States, Southeastern States*. NOAA, National Weather Service, Silver Spring, MD.

The version number has the format P.S where P is a primary version number representing a number of successive releases of primary information. Primary information is essentially the data. S is a secondary version number representing successive releases of secondary information. Secondary information includes documentation and metadata. S reverts to zero (or nothing; i.e., Version 2 and Version 2.0 are equivalent) when P is incremented. When documentation is completed and added without changing any prior information, the version number is not incremented.

The primary version number is stamped on the artifact or is included as part of the filename where the format does not allow for a version stamp (for example, files with gridded precipitation frequency estimates). All location-specific output from the PFDS is stamped with the version number and date of download.

Table 2.1 lists the version history associated with the NOAA Atlas 14 Volume 9 precipitation frequency project and indicates the nature of changes made.

Version no.	Date	<b>Notes</b>
Version 1.0	October 2012	Draft data used in peer review
Version 2.0	April 2013	Final data released

Table 2.1. Version history of NOAA Atlas 14, Volume 9.

#### **3. Introduction**

#### **3.1. Objective**

NOAA Atlas 14 Volume 9 provides precipitation frequency estimates for durations of 5-minutes through 60-days at average recurrence intervals of 1-year through 1,000-year for six southeastern states: Alabama, Arkansas, Florida, Georgia, Louisiana, and Mississippi. The estimates and associated bounds of 90% confidence intervals are provided at 30-arc seconds resolution. The Atlas also includes information on temporal distributions for heavy precipitation amounts for selected durations and seasonal information for annual maxima data used in the frequency analysis. In addition, the potential effects of climate change as trends in historic annual maximum series were examined.

The information in NOAA Atlas 14 Volume 9 supersedes precipitation frequency estimates for of Alabama, Arkansas, Georgia, Florida, Louisiana, and Mississippi contained in the following publications:

- a. Weather Bureau's Technical Paper No. 40*, Rainfall Frequency Atlas of the United States for Durations from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years* (Hershfield, 1961);
- b. Weather Bureau's Technical Paper No. 49*, Two- to Ten-Day Precipitation for Return Periods of 2 to 100 Years in the Contiguous United States* (Miller, 1964);
- c. NOAA Technical Memorandum NWS HYDRO-35*, Five- to 60-Minute Precipitation Frequency for the Eastern and Central United States* (Frederick et al., 1977).

#### **3.2. Approach and deliverables**

Precipitation frequency estimates have been computed for a range of frequencies and durations using a regional frequency analysis approach based on L-moment statistics calculated from annual maximum series. This section provides an overview of the approach; greater detail is provided in Section 4.

The annual maximum series were extracted from precipitation measurements recorded at variable or constant time increments from 1-minute to 1-day obtained from various sources. The table in Appendix A.1 gives detailed information on all stations whose data were used in the frequency analysis. The annual maximum series data were screened for data quality. The 1-day and 1-hour annual maximum series data were also analyzed for potential trends (Appendix A.2).

A region of influence approach was used for the regional L-moments computation at each station across all selected durations between 15-minute and 60-day. A variety of probability distribution functions were examined for each region and duration and the most suitable distribution was selected. Distribution parameters, and consequently precipitation frequency estimates, were determined based on the mean of the annual maximum series at the station and the regionally determined higher order Lmoments. Precipitation frequency estimates were smoothed across durations to ensure consistency. Partial duration series-based precipitation frequency estimates were calculated indirectly using Langbein's formula.

For areas where the snowfall contributes to the precipitation AMS, empirical equations were developed to estimate frequency estimates for rainfall (i.e., liquid precipitation only) from corresponding precipitation frequency estimates for selected durations up to 24-hours. In the NOAA Atlas Volume 9 project area, due to geo-climatic conditions, the contribution of snowfall to AMS is trivial, so no separate rainfall frequency analysis was needed.

A Monte-Carlo simulation approach was used to produce upper and lower bounds of the 90% confidence intervals for the precipitation frequency estimates. 5-minute and 10-minute precipitation

frequency estimates and confidence intervals were computed by applying scaling factors to corresponding 15-minute estimates.

Grids of precipitation frequency estimates and 90% confidence intervals were determined based on grids of mean annual maxima and at-station precipitation frequency estimates. The mean annual maxima grid for each duration was derived from at-station mean annual maxima using PRISM interpolation methodology (Appendix A.3). The grids of precipitation frequency estimates and confidence limits for all frequencies were then derived in an iterative process using the inherently strong linear relationship that exists between mean annual maxima and precipitation frequency estimates at the 2-year recurrence interval and between precipitation frequency estimates at consecutive frequencies for a given duration (Section 4.8.2). The resulting grids were examined and adjusted in cases where inconsistencies occurred between durations and frequencies. Both spatially interpolated and point estimates for selected durations and frequencies were subject to external peer review (Appendix A.4).

Climate regions were delineated based on characteristics of annual maxima data. The regions were used in the extraction of annual maximum series, calculations of temporal distributions of heavy precipitation, and in a seasonality analysis of annual maxima. Temporal distributions, expressed in probability terms as cumulative percentages of precipitation totals, were computed for precipitation magnitudes exceeding precipitation frequency estimates for the 2-year recurrence interval for selected durations (Appendix A.5). The seasonality analysis was done by tabulating the number of annual maxima exceeding precipitation frequency estimates for several selected threshold frequencies (Appendix A.6).

NOAA Atlas 14 Volume 9 precipitation frequency estimates for any location in the project area are available in a variety of formats through the Precipitation Frequency Data Server (PFDS) at <http://hdsc.nws.noaa.gov/hdsc/pfds> (via a point-and-click interface); more details are provided in Section 5. Additional results and information available there include:

- ASCII grids of partial duration series-based and annual maximum series-based precipitation frequency estimates and related confidence limits for a range of durations and frequencies with associated Federal Geographic Data Committee-compliant metadata;
- cartographic maps of partial duration series-based precipitation frequency estimates for selected frequencies and durations;
- final, quality controlled annual maximum series for all observing locations used in the analysis;
- temporal distributions;
- seasonality analysis of annual maxima.

Cartographic maps were created to serve as visual aids and are not recommended for estimating precipitation frequency estimates. Users are advised to take advantage of the PFDS interface or the downloadable underlying ASCII grids for obtaining precipitation frequency estimates.

Precipitation frequency estimates from this Atlas are estimates for a point location and are not directly applicable for an area. Precipitation frequency estimates for each volume of NOAA Atlas 14 were computed independently using all available data at the time. Some discrepancies between volumes at project boundaries are inevitable and they will generally be more pronounced for rarer frequencies.

#### **4. Frequency analysis**

#### **4.1. Project area**

The project area, shown in Figure 4.1.1, encompasses Alabama, Arkansas, Florida, Georgia, Louisiana, and Mississippi and covers 331,500 square miles (858,581 square kilometers). The terrain ranges from low inland mountains to swamps and everglades at the coast. Most of the project area is within the low-lying Gulf Coastal Plain that has very flat to rolling terrain. The primary exceptions are in the northern parts of the project area - the Ozark Plateau and Ouachita Mountains in Arkansas, the Cumberland Plateau and foothills in Alabama, and the Appalachian Mountains in Alabama and Georgia.

The Ozark Plateau and Ouachita Mountains in the northwest half of Arkansas are part of the Interior Highlands and reach summits over 2,500 feet (762 meters) with valleys as low as 500 feet (152 meters). Arkansas' highest point, Mount Magazine at 2,753 feet (839 meters) is in the Ouachita Mountains.

The Cumberland Plateau flanks the western side of the southern Appalachian Mountains and reaches into the northern part of Alabama. These highlands are comprised of parallel ridges and valleys. Relief differences here are about 400 feet (123 meters). The foothills of the Cumberland Plateau extend into Mississippi and include that state's highest point which is only 806 feet (246 meters), Woodall Mountain.

The Appalachian Mountains, specifically the Blue Ridge Mountains, reach into northern Georgia and parts of northeast Alabama. They include Cheaha Mountain which is Alabama's highest point at 2,407 feet (734 meters). The peaks in Georgia attain elevations of more than 3,500 feet (1,067 meters) with the highest, Brasstown Bald at 4,784 feet (1,458 meters).

South of these features, the relatively flat Gulf Coastal Plain stretches across southern Arkansas, Louisiana, Mississippi, Alabama, Georgia, and Florida. In Georgia and Florida, it merges with the Atlantic Coastal Plain. Overall elevation is very low and decreases gradually from the north to sea level at the Gulf Coast and Atlantic Ocean. Louisiana is one of the lowest areas in the United States; its highest point is Driskill Mountain which is upland in the Gulf Coastal Plain at only 535 feet (163 meters). As the Plains approach the coast, the terrain becomes dominated by swamps, marshlands and beaches.

Other features in the project area include the Mississippi Alluvial Plains that surround the Mississippi River. The Mississippi River forms the eastern border of Arkansas and Louisiana and western border of Mississippi before it cuts through southeastern Louisiana to the Gulf. Only Arkansas has mentionable terrain in these plains – Crowley's Ridge rises to about 500 feet (150 meters). Southern Louisiana and the Mississippi Delta south of New Orleans is primarily marshland.

Lastly, most of the state of Florida is at or near sea level and forms a peninsula between the Gulf of Mexico and the Atlantic Ocean. It has approximately 1,350 miles (2,170 km) of coastline. Central and northern Florida have hills that reach elevations up to about 250 feet (76 meters). The state's highest point, 345 feet (105 m), is not on the peninsula but in northwest Florida at Britton Hill.

Within the project area, the largest, most-populated urban areas include Atlanta (GA), Birmingham (AL), Jackson (MS), Jacksonville (FL), Little Rock (AR), Miami (FL), and New Orleans (LA). Other large cities in the project area include Fayetteville (AR), Huntsville (AL), Shreveport (LA), and the metropolitan area around Tampa Bay (FL).

Figure 4.1.1. Project area for NOAA Atlas 14 Volume 9. (The shaded relief was obtained from [USGS EROS Data Center.](http://eros.usgs.gov/))

**Climatology of heavy precipitation**. The climatology of heavy precipitation in the project area is strongly influenced by the warm, humid, subtropical air that is generated by the Gulf of Mexico and the Atlantic Ocean. Variations in mean annual precipitation and mean annual maxima depend mainly on the distance from the warm waters of the Gulf of Mexico. Higher magnitudes of heavy precipitation occur in the south along the coast and tend to decrease from south to north in the project area. The Ozark Plateau, the Ouachita Mountains and the southern end of the Appalachian Mountains are the main areas with large changes in elevation that result in stronger orographic variations in heavy precipitation.

Under certain large-scale pressure patterns, strong low-level southerly flow can transport warm moist subtropical air from the Gulf of Mexico over the southern states. This moist warm air meets with cold dry air from the west. The combination creates an environment of high instability and wind shear. Severe convective storms are triggered and intensified by the combination of this unstable atmosphere and dynamic forcing through the passing of an upper-level trough, convergence boundary, dry line, or cold front. These fronts tend to have a north-south alignment but can shift more east-west and become stationary producing heavy rain over one area for several days. Thunderstorms that develop in this region can become more organized, forming squall lines that may mature into larger mesoscale convective complexes (MCCs). MCCs can persist through the night producing heavy stratiform rainfall. Heavy rainfall can also result from training thunderstorms,

where consecutive storms follow the path of the preceding storm within a given system, which can lead to rainfall over one area for several hours.

In most of the southeastern states, these mechanisms can occur any time of year and can generate heavy precipitation (or annual maxima) for daily or longer durations. The exceptions are southeastern Georgia and the Florida peninsula where in the winter there is less influence due to the positioning of the subtropical jet stream that drives the major circulation patterns and keeps the systems from impacting the area during those months; for this area, daily maxima tend to occur in the spring through fall. For hourly durations, heavy precipitation is mostly likely to occur during the spring through fall throughout the project area with a maximum in the summer. Major systems and other dynamic forces are more prevalent in the spring through fall, although they can occur any time of year. During the summer months when there is weaker dynamic forcing, solar insolation and increased humidity tend to be the dominant factors for convective development of brief heavy storms. Solar insolation can also lead to the development of a sea-breeze front in Florida which often triggers strong afternoon thunderstorms.

During the summer and fall, other mechanisms to deliver heavy precipitation to the southeast are tropical cyclones (TCs), such as tropical depressions, tropical storms, and hurricanes. All the states in the project area are threatened with the potential of a landfalling TC. These tropical systems account for a majority of the extreme rainfall events along the coast. The amount of rain produced by these systems depends on their speed and size. As TCs move onshore they tend to slow down dispensing torrential amounts of rainfall over a few days.

Based on the climatology of heavy precipitation and precipitation mechanisms influencing the project area, two climate regions (shown in Figure 4.1.2) were delineated and used to assign a rainy season during the AMS extraction (Section 4.3), analysis of trends in AMS (Appendix A.2), analysis of temporal distributions of heavy precipitation (Appendix A.5), and in portraying the seasonality of annual maxima data (Appendix A.6). The Mississippi Valley region (region 1) also includes stations from the Mississippi Valley region (region 4) from NOAA Atlas 14 Volume 8.



NOAA Atlas 14 Volume 9 Version 2.0 7 Figure 4.1.2. Two climate regions delineated for NOAA Atlas 14 Volume 9.

#### **4.2. Precipitation data collection and formatting**

Precipitation measurements were obtained for 7,861 stations from a number of federal, state, and local agencies. The majority of the stations were from the NWS Cooperative Observer Program's database maintained by the NOAA's National Climatic Data Center (NCDC). In order to have a uniform system of numbering, each station was assigned a unique six-digit identification number (station ID) where the first two digits were common for all stations from the same data provider. Except for NCDC stations, assigned identification numbers do not match identification numbers assigned by agencies that provided the data. A list of all agencies that provided the data for this project together with agencies' abbreviated names used in this document and the first two digits of stations' identification numbers are shown in Table 4.2.1.

All data were formatted to a common format at one of three base durations that corresponded to the original reporting period: 15-minute, 1-hour, or 1-day. Data recorded at variable time steps, were formatted at 15-minute increments. Where available, records extended through October 2011 with some stations updated through August 2012. Table 4.2.2 lists the total number of stations that were obtained and formatted for each reporting interval.

In addition, monthly maxima for various n-minute durations (5-minute through 60-minute) were obtained for 93 NCDC stations to which any available data from the NWS and Federal Aviation Administration's Automated Surface Observing System (ASOS) network were added; they were used to develop scaling factors used for generation of precipitation frequency estimates grids at 5-minute and 10-minute durations (Section 4.8.2).



Table 4.2.1. Agencies that provided data for the project with their abbreviations, dataset names, data reporting interval, and assigned common first two digits of station identification numbers.



\*NCDC IDs by state: 01 (Alabama), 03 (Arkansas), 08 (Florida), 09 (Georgia), 14 (Kansas), 15 (Kentucky), 16 (Louisiana), 22 (Mississippi), 23 (Missouri), 31 (North Carolina), 34 (Oklahoma), 38 (South Carolina), 40 (Tennessee), 41 (Texas)

Table 4.2.2. The number of stations that were obtained per reporting interval.

Data reporting interval	Number of stations
1-day	4,743
1-hour	1,573
15-minute or variable	1,545

#### **4.3. Annual maximum series extraction**

The precipitation frequency analysis approach used in this project is based on analysis of annual maximum series (AMS) across a range of durations. AMS for each station were obtained by extracting the highest precipitation amount for a particular duration in each successive calendar year. Calendar year was used in this project area, rather than a standard water year (October - September), based on the distribution of heavy precipitation events so that a year begins and ends during a relatively dry season. AMS at stations were extracted for all durations equal to and longer than the base duration (or reporting interval) up to 60 days. AMS for the 1-day through 60-day durations were compiled from daily, hourly, and 15-minute records. To accomplish this, 15-minute and hourly data were first aggregated to constrained 1-day (hours 0 to 24) values before extracting 1-day and longer duration annual maxima. Hourly and 15-minute data were used to compile AMS for 1-hour through 12-hour durations, where, 15-minute data were aggregated first to constrained 1-hour (0 to 60 minutes) values before extracting AMS. 15-minute data were also used to compile AMS for 15 minute and 30-minute durations.

The procedure for developing an AMS from a precipitation dataset used specific criteria designed to extract only reasonable maxima if a year was incomplete or had accumulated data. Accumulated data occurred in some records where observations were not taken regularly, so recorded numbers

represent accumulated amounts over extended periods of time. Since the precipitation distribution over the period is unknown, the total amount was distributed uniformly across the whole period. All annual maxima that resulted from accumulated data were flagged and went through screening to ensure that the incomplete data did not result in erroneously low maxima (Section 4.5.1).

The criteria for AMS extraction were designed to exclude maxima if there were too many missing or accumulated data during the year and more specifically during critical months when precipitation maxima were most likely to occur ("wet season"). Wet seasons were resolved by assessing the periods in which two-thirds of annual maxima occurred at each station and by inspecting histograms of annual maxima for the 1-day and 1-hour durations in a region. The final wet season months were determined using the climate regions as depicted in Figure 4.1.2. The assigned wet season months are shown in Table 4.3.1.

	Wet season months				
Region	<b>Daily durations</b>	<b>Sub-daily durations</b>			
Mississippi Valley (1)	$January - December$	April – October			
Southeast $(2)$	March – October	$May - October$			

Table 4.3.1. Wet season months for each region for daily and sub-daily durations.

The flowchart in Figure 4.3.1 depicts the AMS extraction criteria for all durations. Various thresholds for acceptable amounts of missing or accumulated data were applied to the year and wet season. The extracted maximum value of a given duration for a given year had to pass through all of the criteria in the flowchart to be accepted. Various codes were assigned to both accepted and rejected maxima based on the amount of missing and accumulated data in each year (see Figure 4.3.1) to assist in further quality control of AMS as described in Section 4.5.1.

 For example, in a year with less than 20% of the measurements missing in the whole year and during the assigned wet season, if more than 66% of the measurements were accumulated, then the maxima for that year was (conditionally) rejected, and assigned code 130. If the year had between 33% and 66% accumulated data, then it was further screened by assessing the lengths of the accumulation periods. If the lengths of the accumulation periods for more than 33% of the accumulated data were equal to or longer than threshold accumulation period lengths  $(D_{\text{thresh}})$ , then the maximum for that year was (conditionally) rejected (code 140). Threshold accumulation period lengths were defined as matching the selected duration for durations less than 2 days, as equal to half of duration period for durations between 2 days and 20 days, and as equal to 15 days for durations equal to or longer than 30 days. If the year had less than 33% accumulated data, the extracted maximum was passed to another set of criteria for accumulations during its wet season, etc.

If a rejected annual maximum was higher than 95% of the accepted maxima at that station, then it was kept in the series (code 30). Also, if a rejected 1-day annual maximum was higher than any accumulated amount in a year, then it was kept in the series and assigned code 40. Years in which a maximum was rejected were marked as missing in the series.



Figure 4.3.1. Criteria used to extract annual maxima. Data quality codes were assigned based on acceptance and rejection; D<sub>thresh</sub> depends on duration.

### **4.4. Station screening**

Station screening was done in the following order: a) examination of geospatial data, b) screening for duplicate records at co-located daily, hourly, and/or 15-minute stations and extending records using data from co-located stations, c) screening nearby stations for potentially merging records or removing shorter, less reliable records in station dense areas, and d) screening for sufficient number of years with usable data.

**Geospatial data**. Latitude, longitude, and elevation data for all stations were screened for errors. Several stations had to be re-located because they plotted in a different state or were clearly misplaced based on inspection of satellite images and maps. Misplacement was typically the result of no seconds recorded in latitude and longitude data. There were also several stations with no elevation data; for those stations, elevation was estimated from high-resolution digital elevation model (DEM) grids. Several corrections to metadata were also made based on input received during the peer review (see Appendix A.4)

**Co-located stations**. Co-located stations were defined as stations that have the same geospatial data, but report precipitation amounts at different time intervals. The screening of co-located stations was done as follows:

- If co-located 15-minute and hourly stations provided data for the same period and there were no differences in AMS for constrained 1-hour maxima (15-minute data aggregated on the clock hour), only the 15-minute station was retained and used to extract AMS for all longer durations.
- If a 15-minute or hourly station provided data for the same period as a co-located daily station and there were no differences in AMS for constrained 1-day maxima (15-minute or 1-hour data aggregated from 0 to 24 hours), only the 15-minute or hourly station was retained and used to extract AMS for all longer durations.
- If periods of record at co-located stations were consistent but did not completely overlap, aggregated data from the station with the shorter reporting interval were used to extend the record of the station with the longer reporting interval.
- If the station with the longer reporting interval had a longer period of record, then it was retained in the dataset in addition to the co-located station with the shorter reporting interval.

AMS data consistency across durations was ensured in later quality control procedures (Section 4.5.3).

**Nearby stations**. Nearby stations were defined as stations located within three miles with consideration to elevation differences. However, in areas of flat terrain, stations up to five miles apart or farther may have been considered. The records of nearby stations were considered for merging to increase record lengths. In station-dense areas, such as in Miami in Florida, some stations were removed from the analysis if a nearby station had a longer overlapping record or better quality data.

**Record length**. Record length was characterized by the number of years for which annual maxima could be extracted (i.e., data years) rather than the entire period of record. Only stations with at least 30 data years were considered for frequency analysis. Allowances were made for isolated stations or stations recording at very short intervals. A minimum of 20 data years was used for hourly stations.

Figure 4.4.1 shows histograms for the number of data years of stations available for frequency analysis across daily, hourly, and sub-hourly durations after all the screenings were done. The average and median record lengths as well as corresponding ranges of record lengths are given in Table 4.4.1.



Figure 4.4.1. Number of stations used for precipitation frequency analysis grouped by record length for daily, hourly and sub-hourly durations.

Duration $(D)$	Number of	Record length (data years)			
	stations	average	median	range	
Daily (1-day $\leq D \leq 60$ -day)	1.450	66	62	$19 - 153$	
Hourly (1-hr $\leq D < 24$ -hr)	478	41	41	$20 - 108$	
Sub-hourly (15-min $\leq D < 60$ -min)	255	27	26	$20 - 48$	

Table 4.4.1. Record length statistics for stations used in frequency analysis for different durations.

Locations of stations recording precipitation data at 1-day intervals that were used in the frequency analysis are shown in Figure 4.4.2 and locations of stations recording at 1-hour and subhourly intervals are shown in Figure 4.4.3. More detailed information on each station whose data were used to calculate precipitation frequency estimates is given in three tables in Appendix A.1. The first table in the appendix lists stations in the core states of Alabama, Arkansas, Florida, Georgia, Louisiana and Mississippi. The second table lists stations in the approximately 1 degree buffer surrounding the core states. Those stations were used in the regionalization task (Section 4.6.2) and to assist with interpolation of at-station estimates (Section 4.8). The third table lists n-minute stations that were not directly used in frequency analysis but assisted in development of precipitation frequency estimates at 5-minute and 10-minute durations (Section 4.8.2). Information provided for each station includes: source, name, identification number and data reporting interval, as well as latitude, longitude, elevation, and period of record. All adjusted geospatial data are shown in bold font in the latitude, longitude, and/or elevation columns. Bold font in the period of record column was used to indicate stations whose records were extended with the data from co-located stations or whose records were lengthened by merging with another station. The metadata from the station listed as the 'Post-merge station ID' was retained in the dataset for the merged record; the metadata for this station will reflect the combined periods of records in bold text. If an hourly and a daily station with different IDs were co-located, then the metadata, including ID, of the daily station shown in the 'Colocated station ID' column of the table should be used to locate the hourly (or 15-minute) station on the PFDS web page.

Figure 4.4.2. Map of stations recording at 1-day intervals used in frequency analysis.

Figure 4.4.3. Map of stations recording at 1-hour (green circles) and 15-minute (or variable intervals and formatted to 15-minute; red circles) used in the analysis. Also, shown n-minute stations (yellow circles) used in the analysis.

## **4.5. AMS screening and quality control**

## **4.5.1. Outliers**

For this project, outliers are defined as annual maxima which depart significantly from the trend of the corresponding remaining maxima. Since data at both high and low extremities can considerably affect precipitation frequency estimates, they have to be carefully investigated and either corrected or removed from the AMS if due to measurement errors. The high and low outliers thresholds from the Grubbs-Beck statistical test (Interagency Advisory Committee on Water Data, 1982) and the median +/- two standard deviations thresholds were used to identify low and high outliers for all durations. Low outliers, which frequently came from years with missing and/or accumulated data, were typically removed from the annual maximum series. All values identified as high outliers were mapped with concurrent measurements at nearby stations. Questionable values that could not be confirmed were investigated further using climatological observation forms, monthly storm data reports and other historical weather event publications. Depending on the outcome of each investigation, values were either kept as is, corrected, or removed from the datasets. An example of outlier examination is shown in Figure 4.5.1: statistical tests indicated that a 24-hour amount of 10 inches recorded on July 9, 1916 at Adairsville (09-0041) in Georgia was an outlier. Further investigation of the original observation form for that date showed that the recorded value was actually accumulated over four consecutive days, and so the value was edited in the dataset.



Figure 4.5.1. Outlier tests for 24-hour AMS at station 09-0041. Data quality codes were assigned to annual maxima during the extraction process (Section 4.3).

## **4.5.2. Correction for constrained observations**

**Daily durations.** The majority of daily AMS data used in this project came from daily stations at which readings were taken once every day at fixed times (constrained observations). Due to the fixed beginning and ending of observation times at daily stations, it is to be expected that extracted (constrained) annual maxima were lower than the true (unconstrained) maxima, especially for shorter daily durations. To account for the likely failure of capturing the true-interval maxima, correction

factors were applied to constrained AMS. The correction factor for each daily duration was estimated as the coefficient of a zero-intercept regression model using concurrent (occurring within  $+/-1$  day) constrained and unconstrained annual maxima from hourly stations as independent and dependent model variables, respectively. Correction factors for all daily durations are given in Table 4.5.1.

Table 4.5.1. Correction factors applied to constrained AMS data across daily durations.

<b>Duration (days)</b>				
<b>Correction factor</b>   $1.12$ $1.04$ $1.03$		1.02	1.01	1.00

**Hourly durations.** Similar adjustments were needed on hourly AMS data to account for the effects of constrained 'clock hour' on observations. The correction factors for hourly AMS were developed using co-located hourly (constrained) and 15-minute (unconstrained) concurrent (occurring within +/- 1 hour) annual maxima; they are shown in Table 4.5.2.

Table 4.5.2. Correction factors applied to constrained AMS data across hourly durations*.*

<b>Duration (hours)</b>				
<b>Correction factor</b> $\vert$ 1.09 1.04		1.02	- 1.01	1.OO

**Sub-hourly durations.** No correction factors were applied to durations under 1-hour.

## **4.5.3. Inconsistencies across durations**

At co-located stations, it was not unusual that corresponding annual maxima differed for some years during their overlapping periods of record. Related 1-day AMS at co-located daily and hourly stations were compared and each pair of significantly different estimates was investigated. Effort was made to identify the source of the error and to correct erroneous observations across all durations that were affected.

Annual maxima at each station were also compared across all durations in each year to ensure that the extracted amount for a longer duration was at least equal to the corresponding amount for the successive shorter duration. Inconsistencies of this type occurred at stations with a significant number of missing and/or accumulated data and resulted from different AMS extraction rules applied for different durations (Section 4.3), or from the correction for constrained observations (Section 4.5.2). In those cases, shorter duration annual maxima were used to replace annual maxima extracted for longer durations. Typically, adjustments of this type were very small.

## **4.5.4. Trend analysis**

Precipitation frequency analysis methods used in NOAA Atlas 14 volumes are based on the assumption of a stationary climate over the period of observation (and application). Statistical tests for trends in AMS and the main findings for this project area are described in more detail in Appendix A.2. Briefly, the stationarity assumption was tested by applying a parametric *t*-test and nonparametric Mann-Kendal test for trends in means and Levene's test for trends in variance in the 1-day and 1-hour annual maximum series data at the 5% significance level. For the 1-day duration, testing was done on stations with at least 70 years of data; for the 1-hour duration, the minimum number of data years was lowered to 40 to increase sample size. Overall, the Mann-Kendall test detected slightly more positive trends in the means than the *t*-test: for the 1-hour duration, no trends were detected at about 87% and 90% of the stations, respectively; for the 1-day duration, no trends were detected at about 83% and 87% of the stations, respectively. Levene's test did not detect trends in

variance at any station at the 1-hour duration and in about 95% of stations at 1-day. Spatial maps did not reveal any spatial coherence in trend results.

The relative magnitude of any trend in the AMS means was also assessed for both climate regions (see Figure 4.1.2). AMS were rescaled by corresponding mean values and then regressed against time. The regression results were tested as a set against a null hypothesis of zero serial correlation. The null hypothesis of no trends in AMS data could not be rejected at 5% significance level.

Therefore, the assumption of stationary AMS was accepted for this project area and no adjustment of AMS magnitudes was made.

#### **4.6. Precipitation frequency estimates with confidence limits at stations**

#### **4.6.1. Overview of methodology and related terminology**

Precipitation magnitude-frequency relationships at individual stations have been computed using a regional frequency analysis approach based on L-moment statistics. Frequency analyses were carried out on annual maximum series (AMS) for the following seventeen durations: 15-minute, 30-minute, 1-hour, 2-hour, 3-hour, 6-hour, 12-hour, 1-day, 2-day, 3-day, 4-day, 7-day, 10-day, 20-day, 30-day, 45-day and 60-day. Frequency estimates based on partial duration series (PDS), which include all amounts for a specified duration at a given station above a pre-defined threshold regardless of year, were developed from AMS data using a formula that allows for conversion between AMS and PDS frequencies. Precipitation frequency estimates at 5-minute and 10-minute durations were derived from corresponding 15-minute estimates. To assess the uncertainty in estimates, 90% confidence intervals were constructed on both AMS and PDS frequency curves.

Frequency analysis involves fitting an assumed distribution function to the data. The following distribution functions were analyzed in this project with the aim to identify a distribution that provides the best precipitation frequency estimates for the project area across all frequencies and durations: 3-parameter Generalized Extreme Value (GEV), Generalized Normal, Generalized Pareto, Generalized Logistic and Pearson Type III distributions; 4-parameter Kappa distribution; and 5 parameter Wakeby distribution.

When fitting a distribution to a precipitation annual maximum series extracted at a given location (and selected duration), the result is a frequency distribution relating precipitation magnitude to its annual exceedance probability (AEP). The inverse of the AEP is frequently referred to as the average recurrence interval (ARI), also known as return period. When used with the AMS-based frequency analysis, ARI does not represent the "true" average period between exceedance of a given precipitation magnitude, but the average period between years in which a given precipitation magnitude is exceeded at least once. Those two average periods can be considerably different for more frequent events. The "true" average recurrence interval (ARI) between exceedance of a particular magnitude can be obtained through frequency analysis of PDS.

Differences in magnitudes of corresponding frequency estimates (i.e., quantiles) from the two series are negligible for ARIs greater than about 15 years, but notable at smaller ARIs (especially for  $ARI \leq 5$  years). Because the PDS can include more than one event in any particular year, the results from a PDS analysis are considered to be more reliable for designs based on frequent events (e.g., Laurenson, 1987). To avoid confusion, herein the term AEP is used with AMS frequency analysis and ARI with PDS frequency analysis. The term "frequency" is interchangeably used to specify the ARI and AEP.

L-moments (Hosking and Wallis, 1997) provide an alternative way of describing frequency distributions to traditional product moments (conventional moments) or maximum likelihood approach. Since sample estimators of L-moments are linear combinations of ranked observations, they are less susceptible to the presence of outliers in the data than conventional moments and are well suited for the analysis of data that exhibit significant skewness. L-moments typically used to

calculate parameters of various frequency distributions include  $1<sup>st</sup>$  and  $2<sup>nd</sup>$  order L-moments: Llocation (λ<sub>1</sub>) and L-scale (λ<sub>2</sub>), and the following L-moment ratios: L-CV (τ), L-skewness (τ<sub>3</sub>), and Lkurtosis ( $\tau_4$ ). L-CV, which stands for "coefficient of L-variation", is calculated as the ratio of L-scale to L-location ( $\lambda_2/\lambda_1$ ). L-skewness and L-kurtosis represent ratios of the 3<sup>rd</sup> order ( $\lambda_3$ ) and 4<sup>th</sup> order  $(\lambda_4)$  L-moments to the  $2^{nd}$  order  $(\lambda_2)$  L-moment, respectively, and thus are independent of scale.

One of the primary problems in precipitation frequency analysis is the need to provide estimates for average recurrence intervals that are significantly longer than available records. Regional approaches, which use data from stations that are expected to have similar frequency distributions, have been shown to yield more accurate estimates of extreme quantiles than approaches that use only data from a single station. The number of stations used to define a region should be large enough to smooth variability in at-station estimates, but also small enough that regional estimates still adequately represent local conditions. The region of influence approach (Burn, 1990) used in this volume defines regions such that each station has its own region with a potentially unique combination of nearby stations. Stations are selected based on the maximum allowable distance from the target station that is defined in a geographic space and in a space of selected statistical attribute variables. Like with other regionalization approaches, there is level of subjectivity involved in the process, for example, in choosing attribute variables, selecting the maximum allowable distance as well as attributes' weights and transformations for similarity distance algorithms. One of the advantages of the region of influence approach is that it results in a smooth transition in estimates across regional boundaries, which is relevant for the mapping of precipitation frequency estimates.

A frequency curve that is calculated from sample data represents some average estimate of the population frequency curve, but there is a high probability that the true value actually lies above or below the sample estimate. Confidence limits provide a measure of the uncertainty. They represent values between which one would expect the true value to lie with a certain confidence; they are not necessarily equidistant from the estimates. The width of a confidence interval between the upper and lower confidence limits is affected by a number of factors, such as the degree of confidence, sample size, exceedance probability, and so on. In this volume, simulation-based procedures were used to estimate confidence limits of a 90% confidence interval.

Precipitation frequency estimates from NOAA Atlas 14 are point estimates, and are not directly applicable to an area. The conversion of a point to an areal estimate is usually done by applying an appropriate areal reduction factor to the average of the point estimates within the subject area. Areal reduction factors are generally a function of the size of an area and the duration of the precipitation. The depth-area-duration curves from the Technical Paper No. 29 (U.S. Weather Bureau, 1960) developed for the contiguous United States, can be used for this purpose.

Precipitation frequency estimates for each NOAA Atlas 14 volume were computed independently using all available data at the time. Some discrepancies between volumes at project boundaries are inevitable and they will generally be more pronounced for more rare frequencies.

#### **4.6.2. Regionalization**

Initial regions for each station were created by grouping the closest 10 stations. Stations were then added to or removed from regions based on examination of their distance from a target station, elevation difference, difference in MAMs at various durations, inspection of their locations with respect to mountain ridges, etc. (see an example in Figure 4.6.1) and assessment of similarities/dissimilarities in the progression of relevant L-moment statistics across durations compared with other stations in the region (see Figure 4.6.2). Typically, final regions included between 8 and 16 stations with a cumulative number of data years between 600 and 1,100 for daily durations and 100 and 250 for hourly durations. However, in some areas of low station density, final numbers of data years for some regions were as low as 400 for daily durations and 50 for hourly durations.

**Regional L-moments calculation.** For a given duration, regional estimates of L-moment ratios (L-CV, L-skewness and L-kurtosis) were obtained by averaging corresponding station-specific estimates weighted by record lengths. Regional L-moment ratios were then used to estimate higher order Lmoments at each station.





Figure 4.6.1. An example of spatial plot with accompanying table used in an interactive process for adding or removing stations from a region for station Falkville 1 E, AL (01-2840).

 $\pmb{\dot{\textbf{1}}}$ 

18

17

16

15

 ${\bf 14}$ 

11



Figure 4.6.2. An example of plots of L-moments (left panels),  $MAM/MAM_{24hr}$  and L-moment ratios (right panels) across hourly and daily durations for a region. Thick red lines show statistics for the target station (daily station 01-2840); thin colored lines show statistics for other stations in the region; thick dashed black lines show corresponding regional estimates.

**Station dependence.** Since stations were selected based on geographic proximity to a target station, it was likely that some of the extracted annual maxima at nearby stations came from the same storm events. Dependence in AMS data for stations within a region was analyzed using a *t*-test for the significance of a correlation coefficient at the 5% level. Analysis indicated that cross-correlation among stations was often statistically significant in areas with a dense network of rain gauges and that the number of dependent station pairs increased with duration length. The impact of station dependence on precipitation frequency estimates is considered to be minimal (e.g., Hosking and Wallis, 1997), so it was not addressed in the calculation of precipitation frequency estimates.

However, it was accounted for during the construction of confidence intervals on estimates where it could have noticeable influence (see Section 4.6.5).

## **4.6.3. AMS-based estimates**

**Choice of distribution.** A goodness-of-fit test based on L-moment statistics for 3-parameter distributions, as suggested by Hosking and Wallis (1997), was used to assess which of the five 3 parameter distributions listed in Section 4.6.1 provide acceptable fit to the AMS data. Results of  $\chi^2$ and Kolmogorov-Smirnov tests and visual inspection of probability plots for all seven distributions for 1-hour, 1-day and 10-day durations, like the one shown in Figure 4.6.3, were considered during distribution selection. The GEV distribution was adopted across all stations and for all durations for several reasons. GEV is a distribution generally recommended for analysis of extreme events. Based on the test results, the GEV distribution provided an acceptable fit to data more frequently than any other distribution. Finally, although it is not required to use the same type of distribution across all durations and/or regions, changes in distribution type for different durations or regions often lead to considerable discontinuities in frequency estimates across durations or between nearby locations, particularly at more rare frequencies.



Figure 4.6.3. Probability plots for selected distributions for 1-day AMS at station Vero Beach 4 SE (08-9219) in Florida.

**Frequency estimates for hourly and daily durations.** For each station and for each hourly and daily duration, L-moment statistics were used to calculate the parameters of the GEV distribution and to produce precipitation frequency estimates for the following annual exceedance probabilities (AEPs): 1/2 (50%), 1/5, 1/10, 1/25, 1/50, 1/100, 1/200, 1/500 and 1/1000. This calculation was repeated for all durations and for all stations. Since L-moments, and consequently, precipitation frequency estimates, were calculated independently for each duration, the resulting depth-durationfrequency (DDF) curves did not always look smooth. Smoothing of quantiles by cubic spline

functions improved the shape of DDF curves. Figure 4.6.4 illustrates precipitation depth-durationfrequency curves before and after smoothing for Augusta, AR (03-0326).



Figure 4.6.4. Precipitation frequency estimates for a range of durations for selected AEPs for station Augusta, AR (03-0326). Blue lines represent original estimates; black lines represent estimates obtained after quantiles were smoothed across durations.

**Frequency estimates for sub-hourly durations.** The shortest duration at which AMS data were extracted was 15 minutes. L-moments were calculated for the 15-minute and 30-minute durations at stations that had 15-minute AMS data available for at least one station assigned to their region. Lmoments were then used to produce precipitation frequency estimates in the same manner as for hourly and daily durations. However, in a number of cases, it was observed that resulting precipitation frequency estimates were implausible, especially for AEPs of 1/100 (1%) or less. The primary cause of this was the sample size, as very few stations with measurements at sub-hourly durations were available, and when they were available, they typically had short periods of record. This resulted in unreliable moments (especially higher-order moments), and consequently, unreliable precipitation frequency estimates.  $\lambda_1$  moments (i.e., mean annual maxima) were less sensitive to a sample size and were generally in line with corresponding estimates at nearby stations.  $\lambda_1$  moments were also, for the most part, consistent with the expected progression across hourly and daily durations (see top left panel of Figure 4.6.2). For that reason, mean annual maxima at 15-minute and 30-minute durations were retained for derivation of MAM grids (see Section 4.8.1). At-station quantiles, which were assessed as unreliable, were not interpolated to create precipitation frequency grids; an alternative approach, described in Section 4.8.2 was used for that purpose.

Similarly, for the 5-minute and 10-minute durations, very few n-minute stations were available to compute precipitation frequency estimates using regional L-moments or to develop MAM grids. Therefore, an alternative approach described in Section 4.8.2 was used to develop these estimates, as well.

#### **4.6.4. PDS-based estimates**

PDS-based precipitation frequency estimates were calculated indirectly from Langbein's formula (Langbein, 1949) which transforms a PDS-based average recurrence interval (ARI) to an annual exceedance probability (AEP):

$$
AEP = 1 - \exp\left(-\frac{1}{ARI}\right).
$$

PDS-based frequency estimates were calculated for the same durations as AMS-based estimates for 1-, 2-, 5-, 10-, 25-, 50-, 100-, 200-, 500- and 1,000-year ARIs. Selected ARIs were first converted to AEPs using the above formula and then precipitation frequency estimates were calculated for those AEPs following the same approach that was used in the AMS analysis.

#### **4.6.5. Confidence limits**

A Monte Carlo simulation procedure that accounts for inter-station dependence, as described in Hosking and Wallis (1997), was used to construct 90% confidence intervals (i.e., 5% and 95% confidence limits) on both AMS-based and PDS-based precipitation frequency curves. It should be noted that confidence intervals constructed through this approach account for uncertainties in distribution parameters, but not for other sources of uncertainties (for example, distribution selection), that could also significantly impact the total error, particularly at more rare frequencies.

Since the station dependence analysis (Section 4.6.2) indicated that for regions with a more dense station network, AMS data from different stations could be dependent (especially for longer durations), the simulation algorithm that accounts for inter-station correlation was used. At each station, 1,000 simulated data sets per duration were used to generate precipitation quantiles. Estimates were sorted from smallest to largest and the  $50<sup>th</sup>$  value was selected as the lower confidence limit and the  $950<sup>th</sup>$  value was selected as the upper confidence limit.

Due to differences in record lengths across hourly and daily durations, confidence intervals for hourly durations were wider than corresponding intervals at daily durations for some stations; therefore, they were restricted by the corresponding values at 24-hour duration. Confidence limits for sub-hourly durations were calculated using similar approaches that were used to calculate frequency estimates. Since confidence limits were derived for each duration independently, like precipitation frequency estimates, confidence limits could fluctuate from duration to duration; they were smoothed across durations using cubic spline functions.

## **4.7. Rainfall frequency estimates with confidence limits at stations**

#### **4.7.1. Background**

Precipitation frequency estimates from Section 4.6 represent precipitation magnitudes regardless of the type of precipitation. For some applications it may be important to know frequency estimates from liquid precipitation (i.e., rainfall) only. For example, rainfall is treated differently from snowfall in watershed modeling because of different runoff producing mechanisms. While the rainfall generates runoff almost immediately, snowfall generally goes into storage until it melts and produces runoff at a later time.

For some areas, particularly for high elevation areas, the contribution of snowfall to the total yearly precipitation amount is significant and may translate to its significant participation in precipitation annual maximum series (AMS). For areas where the snowfall contributes to the precipitation AMS, a separate rainfall frequency analysis was done for durations up to 24 hours, which are of most interest to design projects relying on peak flows. In the NOAA Atlas Volume 9 project area, due to geo-climatic conditions, the contribution of snowfall to AMS is trivial, so no

separate rainfall frequency analysis was done. This section was retained for consistency with the documentation of Volume 8.

## **4.8. Derivation of grids**

## **4.8.1. Mean annual maximum precipitation**

Grids of mean annual maxima (MAM) served as the basis for deriving gridded precipitation frequency estimates at different frequencies and durations. The station mean annual maximum values for the 17 selected durations between 15 minutes and 60 days were spatially interpolated to produce corresponding mean annual maximum grids at 30 arc-seconds resolution using a hybrid statisticalgeographic approach for mapping climate data named Parameter-elevation Regressions on Independent Slopes Model (PRISM) developed by Oregon State University's PRISM Climate Group (e.g., Daly et al., 2002). The MAM grids were developed at the same time for both Volume 8 and Volume 9.

Several iterations with the PRISM Climate Group were made to ensure satisfactory MAM patterns. In particular, gauged locations where interpolated MAMs for selected base durations (15 minute, 1-hour, 1-day, 10-day) were more than 10% different (determined by jackknife analysis) than the expected at-station MAMs were carefully re-examined. As a result of those reviews, some MAM estimates were adjusted. MAMs were also estimated for a couple of locations to better anchor the spatial interpolation in areas of varied terrain and/or where the lack of stations with sufficiently long records unduly influenced expected spatial patterns, particularly at hourly durations. Two notable changes to the MAM dataset to improve patterns were:

- 1) daily-only stations with less than 50 years of data in areas of flat terrain and/or areas with a high density of stations were excluded from the MAM interpolation to reduce a number of station-driven contours in MAM maps;
- 2) MAMs were estimated at sub-daily durations for two daily stations in Florida and Georgia and across all durations at several ungauged locations along the coast to anchor the interpolation and improve spatial patterns.

Appendix A.3 provides detailed information on the PRISM-based methodology for creating the mean annual maximum grids. In summary, a unique regression function was developed for each target grid cell to derive mean annual maximum values for each duration that accounted for the difference between an observing station's and the target cell's mean annual precipitation, topographic facet, coastal proximity, the distance of an observing station to the target cell, etc. Jacknife crossvalidation indicated that overall bias for project areas in Volumes 8 and 9 combined was less than one percent for all durations except for 15-minute which had a bias of -1.8%. The mean absolute error was less than 5 percent across all durations.

## **4.8.2. Precipitation frequency estimates with confidence limits**

**Estimates for 60-minute through 60-day durations.** The spatial interpolation technique used in this volume developed grids of AMS-based and PDS-based precipitation frequency estimates along the frequency dimension for a given duration. Hence, the evolution of frequency-dependent spatial patterns for a given duration was independent of other durations. The technique utilizes the inherently strong linear relationship that was found to exist between precipitation frequency estimates for consecutive frequencies, as well as mean annual maxima and 2-year precipitation frequency estimates. For example, Figure 4.8.1 shows the relationship between the 50-year and 100-year estimates for the 24-hour duration for this project area together with regression lines for a linear model and zero-intercept model. The  $R^2$  values of 0.996 and 0.995, respectively, are very close to 1.0, which was common for all relationships. Another common occurrence was a negligible intercept coefficient in the linear model regression equations, so a zero-intercept model was adopted for all

frequencies and durations. The slope coefficient of the zero-intercept model represents an average domain-wide ratio between consecutive quantiles; in this case, 1.1419 is an average ratio between 100-year and 50-year quantiles for the 24-hour duration for the whole project area. Although the correlation coefficients were very high, when plotted on a map, at-station ratios showed some regional features (as shown in Figure 4.8.2 for the same example); this finding was used in the grid generation process.



Figure 4.8.1. Scatter plot of 100-year versus 50-year precipitation frequency estimates based on 24 hour annual maximum series. Linear model and zero-intercept model regression lines are also shown.



Figure 4.8.2. Spatially interpolated ratios used to calculate 24-hour 100-year precipitation frequency grid from the 24-hour 50-year grid.

NOAA Atlas 14 Volume 9 Version 2.0 26

For each duration, the calculation began with the PRISM-derived mean annual maximum (MAM) grid as the initial predictor grid and the grid of 2-year precipitation frequency estimates as the resulting subsequent grid. At-station ratios between the 2-year estimates and corresponding MAM estimates were spatially interpolated to a grid using a natural neighbor interpolation method, which is based on construction of Thiessen polygons from the Delauney triangulation of irregularly spaced gauged locations. The advantage of this method is that it remains true to the at-station estimates; the resulting function is continuous everywhere within the project area and also has a continuous first derivative everywhere except at the data points themselves. Gridded MAM estimates were then multiplied by corresponding gridded ratios to create a grid of 2-year precipitation frequency estimates. In the subsequent run, ratios between the 5-year and 2-year estimates were interpolated and used to calculate 5-year precipitation grid from the 2-year grid, and so forth. The grid of 2-year precipitation frequency estimates was also used to create a grid of 1-year estimates. The same process was repeated for all hourly and daily durations.

During the review process, station-driven contour lines that were showing up in cartographic maps in flat terrain areas. The majority of these was driven by small differences in MAM estimates at nearby stations and selected mapping contour intervals, but to reduce a number of station-driven contours in the final cartographic maps, a dynamic filter was applied to the precipitation frequency grids. Parameters of the filter, which controlled the amount of smoothing, were a function of elevation gradients and proximity to the coastline. Parameters were selected such that no smoothing was applied at the coastline or in the mountains, maximum smoothing was applied in flat terrain, and the transition from one to another was gradual. The resulting smoothed grid then served in the subsequent run as the basis for the derivation of the next grid.

To ensure consistency in grid cell values across all durations and frequencies (e.g., 24-hour estimate has to be at least equal to 12-hour estimate), duration-based internal consistency checks were conducted. For inconsistent cases, the longer duration grid cell value was adjusted by multiplying the shorter duration grid cell value by 1.01 to provide a 1% difference between the values. After grid cell consistency was ensured across durations, it was performed across frequencies to ensure that there were no frequency-based inconsistencies caused by the adjustment across durations.

A jacknife cross-validation technique (Shao and Tu, 1995) was used to evaluate the spatial interpolation technique's performance for interpolating precipitation frequency estimates. It was cost prohibitive to re-create the PRISM mean annual maximum grids for each cross-validation iteration. For this reason, the cross-validation results reflect the accuracy of the interpolation procedure based on the same mean annual maximum grids. Figure 4.8.3 shows validation results for 100-year estimates for the 1-hour and 24-hour durations as histograms showing the distribution of differences in estimates with and without each station (errors). Overall, the spatial interpolation technique adequately reproduced values. For the 1-hour duration, differences were less than ±5% at 97% of the stations; for the 24-hour duration, differences were less than  $\pm 5\%$  at 99% of the stations.



Figure 4.8.3. NOAA Atlas 14 Volume 9 jackknife cross-validation results for: a) 100-year 1-hour estimates, and b) 100-year 24-hour estimates.

**Estimates for 5-minute through 30-minute durations.** A similar approach to the one used to derive grids of precipitation frequency estimates for hourly and daily durations was used to derive gridded estimates for the 15-minute and 30-minute durations. For 15-minute, a grid of 2-year precipitation frequency estimates was calculated by multiplying the 15-minute MAM grid with a grid of ratios between the 2-year estimates and corresponding MAM estimates. In the subsequent run, a grid of ratios between the 5-year and 2-year estimates was used to calculate 5-year grid from the 2-year grid, and so forth. The main difference is that, due to concerns about the soundness of at-station precipitation frequency estimates computed directly from AMS for sub-hourly durations, instead of interpolating gridded ratios from sub-hourly estimates, corresponding 60-minute ratio grids were assumed to characterize 15-minute ratio grids. The same process was used for 30-minute duration, as well.

Precipitation frequency grids for 5-minute and 10-minute durations were derived by multiplying the 15-minute precipitation frequency grids by scaling factors. Scaling factors were obtained from nminute stations; they were calculated as average ratios of 5-minute and 10-minute annual maxima to corresponding 15-minute annual maxima. Given that relatively few n-minute stations were available and that at-station scaling factors varied little across the project area, they were assumed to be uniform for the whole area: 0.57 for 5-minute duration and 0.82 for 10-minute duration. The scaling factors were applied to the 15-minute precipitation frequency grids for all frequencies to create matching 5-minute and 10-minute grids.

**Confidence limits.** Grids of upper and lower limits of the 90% confidence interval for the precipitation frequency estimates between 5-minutes and 60-day durations were derived using same procedures that were used to create grids of precipitation frequency estimates.

#### **5. Precipitation Frequency Data Server**

NOAA Atlas 14 precipitation frequency estimates are delivered entirely in digital form in order to make the estimates more widely available and to provide them in various formats. The Precipitation Frequency Data Server - PFDS [\(http://hdsc.nws.noaa.gov/hdsc/pfds/\)](http://hdsc.nws.noaa.gov/hdsc/pfds/) provides a point-and-click web portal for precipitation frequency estimates and associated information.

In early 2011a major redesign of the PFDS web interface was done to make PFDS pages interactive. Since then, PFDS pages were enhanced on several occasions to improve the usability and readability of PFDS website's content, to increase data download speeds and to provide additional information. In order to keep this section of the documentation up-to-date for all volumes, the PFDS section is offered as a separate document. This document is updated as needed and is available for download from here: [https://www.weather.gov/media/owp/hdsc\\_documents/NA14\\_Sec5\\_PFDS.pdf](https://www.weather.gov/media/owp/hdsc_documents/NA14_Sec5_PFDS.pdf).

#### **6. Peer review**

A peer review of preliminary results for the NOAA Atlas 14 Volume 9 precipitation frequency project was carried out during a five week period starting on October 15, 2012. The request for review was sent via email to the members of the HDSC list-server from all over the United States and other interested parties. Potential reviewers were asked to evaluate the reasonableness of point precipitation frequency estimates as well as their spatial patterns. The review included the following items:

- a. Metadata for stations whose data were used to prepare mean annual maximum precipitation maps and/or in precipitation frequency analysis. The table included information on station name, state, source of data, assigned station ID, latitude, longitude, elevation, and period of record. It also showed if the station was merged with another station, if the station was colocated with another station with a different ID, and if metadata at the station were changed. (Station IDs were assigned by HDSC and do not match station IDs assigned by the agency that provided the data, except for National Climatic Data Center.)
- b. Metadata for stations whose data were collected, but not used in the analysis. The table contained metadata for stations that were examined, but not used, with brief comments on why the data were not used. Generally, stations were not used because there was another station with a longer period of record nearby, station data were assessed as not reliable for this specific purpose, or the station's period of record was not long enough and it was not a candidate for merging with any nearby station.
- c. At-station depth-duration-frequency (DDF) curves for 60-minute to 10-day durations and for 2-year to 100-year ARIs.
- d. Maps of spatially-interpolated estimates of mean annual maximum precipitation for 60 minute, 24-hour and 10-day durations.
- e. Maps of spatially-interpolated precipitation frequency estimates for 60-minute, 24-hour and 10-day durations and for 2-year and 100-year average recurrence intervals.

Comments were received from eight individuals or offices and agencies including Water Management District Offices and Weather Forecast Offices. The reviews provided critical feedback that improved the estimates. The reviews provided critical feedback that improved the estimates. Reviewers' comments regarding station metadata, at-station precipitation frequency estimates and their spatial patterns, and supplemental information along with HDSC responses can be found in Appendix A.4.

## **7. Comparison with previous NOAA publications**

The precipitation frequency estimates in NOAA Atlas 14 Volume 9 supersede the estimates published in the following publications:

- a. NOAA Technical Memorandum NWS HYDRO-35, *Five- to 60-Minute Precipitation Frequency for the Eastern and Central United States* (Frederick et al., 1977) for 5-minute to 60-minute durations;
- b. Weather Bureau Technical Paper No. 40, *Rainfall Frequency Atlas of the United States for Durations from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years* (Hershfield, 1961) for 2-hour to 24-hour durations;
- c. Weather Bureau Technical Paper No. 49, *Two- to Ten-Day Precipitation for Return Periods of 2 to 100 Years in the Contiguous United States* (Miller, 1964) for 2-day to 10-day durations.

NOAA Atlas 14 Volume 9 Version 2.0 30 Precipitation frequency estimates at the 100-year average recurrence interval from NOAA Atlas 14 were examined in relation to corresponding estimates from NOAA Technical Memorandum NWS HYDRO-35 (HYDRO35) for the 60-minute duration and the Weather Bureau's Technical Paper No.

40 (TP40) for the 24-hour duration. Corresponding grids from HYDRO35 and TP40, which were used in the comparison, were obtained by interpolating digitized isopluvials from paper cartographic maps using the standard spatial interpolation tools available in ArcGIS.

The maps in Figures 7.1 and 7.2 illustrate the differences between NA14 and HYDRO35 100 year 60-minute estimates in inches and in percentages, respectively. The contour lines superimposed on the maps represent isopluvials from HYDRO35. On average, 100-year 60-minute precipitation frequency estimates across the project area decreased only 0.2 inches (less than 6%), but at specific locations estimates changed between -2.02 and 1.04 inches or from -45% to 22%. The maximum increase of 1.04 inches (22%) occurred in the extreme southeast Louisiana and the largest decrease of 2.02 inches (-45%) occurred in the mountains of northern Georgia. Increases in magnitudes between 0.5 and 1.0 inches were observed in central Louisiana, portions of the Ouachita mountains in Arkansas, along the Florida panhandle just south of Tallahassee, and along the southeast coast in Florida north from Miami.

The differences in estimates between the two publications are attributed to a number of factors. Firstly, differences in data quality control procedures and frequency analysis approaches (distribution selection, parameter estimation method, regional versus at-station methods) affect estimates, especially at higher ARIs. Section 4 of this document describes methods used in NA14 and their advantages. Secondly, differences in spatial interpolation techniques impact estimates at ungauged locations. Isopluvials in HYDRO35 were based solely on station data without incorporating topographic features; NA14 estimates were based on PRISM products that integrate topography (see Section 4.8 for more details). Finally, the increase in the amount of available data from HYDRO35 to NA14, both in the number of stations and their record lengths, has a considerable effect on estimates. HYDRO35 was published in 1977, so potentially more than 35 additional years of data at existing stations were available for the NA14 analyses. Also, many stations that were not suitable for frequency analysis in HYDRO35 due to short records could be included in NA14. A detailed comparison of the numbers of stations and record lengths available to each of the two projects could not be provided since the HYDRO35 project covered a significantly larger area and the necessary information was not available in the HYDRO35 document.

The maps in Figure 7.3 and 7.4 illustrate the differences between NA14 and TP40 100-year 24 hour estimates in inches and in percentages, respectively. The contour lines superimposed on the maps represent isopluvials from TP40. On average, for the whole project area estimates increased about 0.66 inches (6%); with differences ranging from -2.77 to 4.73 inches and from -29% to 46%. Some of the largest differences in precipitation frequency estimates are in areas where TP40 did not account for orographic influence, such as in the mountains of north Georgia, the Ouachita Mountains in Arkansas (southwest of Little Rock), as well as in the area that lies in the "rain shadow" of mountains north of the Arkansas River (northwest of Little Rock). Estimates decreased as much as 2.77 inches (-29%) in the mountains of north Georgia, increased as much as 2.85 inches (30%) in the Ouachita Mountains in Arkansas, and decreased up to 1.5 inches in the mountains north of the Arkansas River. Other locations with significant increases of up to 4.00 inches (on average around 30%) occurred in central and southwestern Louisiana, southern Alabama, and along Florida's southeast coast from Miami to Palm Beach.

Differences in estimates can be attributed to similar factors as for the 60-minute duration: different data quality control techniques and frequency analysis approaches; different spatial interpolation techniques; and an increase in a number of available stations and record lengths for NA14 relative to TP40. Since TP40 was published in 1961, potentially more than 50 additional data years were available for the NA14 analyses. A more detailed comparison of the numbers of stations and their record lengths between two projects could not be provided since the necessary information was not available in the TP40 document.


Figure 7.1. Map showing differences in 100-year 60-minute estimates (in inches) between NOAA Atlas 14 Volume 9 and HYDRO35. Superimposed on the map are isopluvials (blue lines) from HYDRO35.

NOAA Atlas 14 Volume 9 Version 2.0 32



Figure 7.2. Map showing percent differences in 100-year 60-minute estimates between NOAA Atlas 14 Volume 9 and HYDRO35. Superimposed on the map are isopluvials (blue lines) from HYDRO35.



Figure 7.3. Map showing differences in 100-year 24-hour estimates (in inches) between NOAA Atlas



Figure 7.4. Map showing percent differences in 100-year 24-hour estimates between NOAA Atlas 14

## **Acknowledgments**

This work was funded by the following state agencies through the Federal Highway Administration's Transportation Pooled Fund Program: Alabama Department of Transportation, Florida Department of Transportation, Southwest Florida Water Management District, Georgia Department of Transportation, and Mississippi Department of Transportation. We also thank Kornel Korenyi and Cynthia Nurmi for conducting the activities of the Federal Highway Administration's Transportation Pooled Fund Program.

We acknowledge the many colleagues who provided data for this project beyond what was available from NOAA's National Climatic Data Center, including: Jim Asby of the Western Region Climate Center; Kristopher Barrios of the Northwest Florida Water Management District; Scott A. Braun of the National Aeronautics and Space Administration (NASA), TRMM Satellite Validation Office; Daniel Chan of the Georgia Forestry Commission; Maggie Dunklee of the Natural Resources Conservation Service; Bob Larson of the Illinois State Water Survey, National Atmospheric Deposition Program; Bach McClure of the Natural Resources Management Office, Brevard County, Florida; William (Rick) Lusher of the University of Florida, Institute of Food and Agricultural Sciences; Glenda McDermont of the St. Johns River Water Management District; Bill Messersmith of the City of Vero Beach, Florida; Chandra S. Pathak of the South Florida Water Management District; Ken Romie of the Southwest Florida Water Management District; Nancy Westcott of the Midwestern Region Climate Center; and Joyce Williamson of the U.S. Geological Survey.

We also acknowledge colleagues who provided feedback to improve the final product, including: Tim Cera, David Clapp, Michael Cullum, Dale Smith, and David Watt of the St. Johns River Water Management District, Florida; Lee Crowley and colleagues at the Arkansas-Red River Forecast Center; Harry Downing, Paul O'Neil, and Dawn Turner of the Southwest Florida Water Management District; Eugene L. Jacobi of the NWS Mobile/Pensacola Weather Forecast Office, Alabama; Jack Settelmaier of the NWS Southern Region Headquarters; and Nick Wooten of the Northwest Florida Water Management District.

Lastly, we'd like to acknowledge the contributions from ZhiXin Li of the NWS Office of Hydrologic Development and former members of HDSC including: Li-Chuan Chen, Sarah Dietz, Sarah Heim, Lillian Hiner, Kazungu Maitaria, Fenglin Yan, and Tan Zhao.

## **Appendix A.1 List of stations used to prepare precipitation frequency estimates**

Table A.1.1. List of stations in the states of Alabama, Arkansas, Florida, Georgia, Loiusiana, and Mississippi used in the analysis showing station name, station ID, post-merge station ID, co-located daily station ID, base duration, source of data, latitude, longitude, elevation, and period of record. Bold font in the latitude, longitude, and elevation fields indicates information that has been adjusted. Bold font in the 'Period of record' field indicates that the station data was extended using data from station that has the same ID in 'Post-merge station ID' column. For an hourly station co-located with a daily station with a different ID, the daily station's ID shown in the 'Co-located station ID' column should be used to locate the hourly station on the PFDS web page.





NOAA Atlas 14 Volume 9 Version 2.0 NOAA Atlas 14 Volume 9 Version 2.0



NOAA Atlas 14 Volume 9 Version 2.0 NOAA Atlas 14 Volume 9 Version 2.0









































NOAA Atlas 14 Volume 9 Version 2.0 NOAA Atlas 14 Volume 9 Version 2.0














































































Table A.1.2. List of stations used in the analysis in the buffer zone in the U.S. states of Kansas, Kentucky, Missouri, North Carolina, lahoma, South Carolina, Tennessee, and Texas. The table shows station name, station ID, post-merge station ID, co-located daily station base duration, source of data, latitude, longitude, elevation, and period of record. Bold font in the latitude, longitude, and elevation fields icates information that has been adjusted. Bold font in the 'Period of record' field indicates that the station data was extended using data n station that has the same ID in 'Post-merge station ID' column. For an hourly station co-located with a daily station with a different ID, the daily station's ID shown in the 'Co-located station ID' column should be used to locate the hourly station on the PFDS web page.




































	station iD, source of data, familiac, folightace, cicvation, and perfod of record.										
	<b>State Station name</b>	<b>Station ID</b>	Source of data				Latitude   Longitude   Elevation (ft)   Period of Record				
AL	ALABASTER SHELBY CO AP	01-0116	<b>NCDC</b>	33.1783	$-86.7817$	565	3/2005-7/2009				
AL	ANNISTON METRO AP	01-0272	<b>NCDC</b>	33.5872	$-85.8556$	594	1/1984-7/2009				
AL	<b>BIRMINGHAM WSFO</b>	01-0829	<b>NCDC</b>	33.4667	$-86.8333$	744	1/1984-4/1991				
AL	<b>BIRMINGHAM AP ASOS</b>	01-0831	<b>NCDC</b>	33.5656	$-86.7450$	615	1/1973-7/2009				
AL	<b>CENTREVILLE 6 SW</b>	01-1525	<b>NCDC</b>	32.8661	$-87.2383$	450	1/1984-5/1994				
AL	HUNTSVILLE INTNL AP	01-4064	<b>NCDC</b>	34.6439	$-86.7861$	624	1/1973-7/2009				
AL	<b>MOBILE RGNL AP</b>	01-5478	<b>NCDC</b>	30.6883	$-88.2456$	215	1/1973-7/2009				
AL	MONTGOMERY AP ASOS	01-5550	<b>NCDC</b>	32.2997	$-86.4075$	202	1/1973-7/2009				
AL	<b>MUSCLE SHOALS AP</b>	01-5749	<b>NCDC</b>	34.7442	$-87.5997$	540	1/1984-7/2009				
AL	<b>TUSCALOOSA ACFD</b>	01-8380	<b>NCDC</b>	33.2119	$-87.6161$	169	1/1984-4/1994				
AR	DEQUEEN SEVIER CO AP	03-1953	<b>NCDC</b>	34.0500	$-94.4006$	355	3/2005-7/2009				
AR	EL DORADO GOODWIN FLD	03-2300	<b>NCDC</b>	33.2208	$-92.8142$	252	1/1984-7/2009				
AR	FT SMITH RGNL AP	03-2574	<b>NCDC</b>	35.3331	$-94.3625$	449	1/1973-7/2009				
AR	HARRISON BOONE CO AP	03-3165	<b>NCDC</b>	36.2667	$-93.1567$	1374	1/1984-7/2009				
AR	LITTLE ROCK ADAMS FLD	03-4248	<b>NCDC</b>	34.7272	-92.2389	258	1/1973-7/2009				
AR	MONTICELLO MUNI AP	03-4900	<b>NCDC</b>	33.6361	$-91.7556$	290	3/2005-7/2009				
AR	<b>NORTH LITTLE ROCK WFO</b>	03-5320	<b>NCDC</b>	34.8353	$-92.2597$	563	1/1978-5/1997				
AR	PINE BLUFF GRIDER FLD	03-5756	<b>NCDC</b>	34.1667	-91.9333	$-999$	3/2005-7/2009				
AR	<b>TEXARKANA WEBB FLD</b>	03-7048	<b>NCDC</b>	33.4536	$-94.0075$	361	1/1984-7/2009				
FL	APALACHICOLA AP	08-0211	<b>NCDC</b>	29.7258	$-85.0206$	20	1/1973-7/2009				
FL	<b>CRESTVIEW BOB SIKES AP</b>	08-1986	<b>NCDC</b>	30.7797	$-86.5225$	190	3/2005-7/2009				
FL	DAYTONA BEACH INTL AP	08-2158	<b>NCDC</b>	29.1828	$-81.0483$	31	1/1973-7/2009				
FL	FT LAUDERDALE INTL AP	08-3165	<b>NCDC</b>	26.0719	$-80.1536$	11	3/2005-7/2009				
FL	FT MYERS PAGE FLD AP	08-3186	<b>NCDC</b>	26.5850	$-81.8614$	15	1/1973-7/2009				
FL	<b>GAINESVILLE RGNL AP</b>	08-3326	<b>NCDC</b>	29.6919	$-82.2756$	123	1/1984-7/2009				
FL	<b>JACKSONVILLE INTL AP</b>	08-4358	<b>NCDC</b>	30.4950	$-81.6936$	26	1/1973-7/2009				
FL	JACKSONVILLE CRAIG AP	08-4370	<b>NCDC</b>	30.3361	$-81.5147$	41	3/2005-7/2009				
FL	<b>KEY WEST INTL AIRPORT</b>	08-4570	<b>NCDC</b>	24.5550	$-81.7522$	$\overline{4}$	1/1973-7/2009				
FL	<b>LAKELAND</b>	08-4797	<b>NCDC</b>	28.0206	$-81.9219$	145	1/1976-8/1978				
FL	<b>MAYPORT PILOT STN</b>	08-5549	$\rm NCDC$	30.4000	$-81.4167$	16	1/2007-7/2009				
FL	<b>MELBOURNE WFO</b>	08-5612	<b>NCDC</b>	28.0958	$-80.6308$	35	3/2005-7/2009				
FL	<b>MIAMI BEACH</b>	08-5658	<b>NCDC</b>	25.8064	$-80.1336$	$\mathbf{1}$	1/1973-12/1974				

Table A.1.3. List of stations used in the analysis for n-minute scaling factors (see Section 4.6.3) showing state, station name, station ID, source of data, latitude, longitude, elevation, and period of record.





Levene's test (Levene, 1960) was used to test for homogeneity of variance in the AMS data. The test has been proven to be less sensitive to non-normality in data than some other commonly used tests (such as the Barlett test). The test statistic, *W*, is defined as follows:

$$
W = \frac{(N - k) \sum_{i=1}^{k} N_i (Z_i - Z_i)^2}{(k - 1) \sum_{i=1}^{k} \sum_{j=1}^{N_i} N_i (Z_{ij} - Z_i)^2}
$$

where k is the number of sub-groups, N is the sample size,  $N_i$  is the sample size of the i<sup>th</sup> subgroup,  $Y_{ij}$  is the value of the j<sup>th</sup> sample from the i<sup>th</sup> subgroup, and  $Z_{ij}$  is the absolute deviation of  $Y_{ij}$  from the mean of the i<sup>th</sup> subgroup. Levene's test rejects the hypothesis that the variances are equal if

$$
W\!>F_{a,\,k\text{-}1,\,N\text{-}k}
$$

where *F<sup>α</sup>*, *k*-1, *N-k* is the upper critical value of the *F* distribution with *k*-1 and *N*-*k* degrees of freedom at a significance level of *α*.

At-station trends in AMS means were inspected using the parametric *t*-test and non-parametric Mann-Kendall test (e.g., Maidment, 1993). Both tests are extensively used for trend analysis in environmental sciences and are appropriate for records that have undergone a gradual change. The tests are fairly robust, readily available, and easy to use and interpret. Since each test is based on different assumptions and different test statistics, the rationale was that if both tests have similar outcomes there can be more confidence about the results, and if the outcomes are different, it would provide an opportunity to investigate reasons for discrepancies.

Parametric tests in general have been shown to be more powerful than non-parametric tests when the data are approximately normally distributed and when the assumption of homoscedasticity (homogeneous variance) holds (Hirsch et al., 1991), but are less reliable when those assumptions do not hold. The parametric *t*-test for trend detection is based on linear regression, and therefore checks only for a linear trend in data. A linear trend assumption seemed adequate here, since, time series plots indicated, if any, monotonic, linear changes in AMS. The Pearson correlation coefficient (*r*) was used as a measure of linear association between annual maximum series data and time for the *t*test. The hypothesis that the data are not dependent on time (and also that they are independent and normally distributed values) was tested using the *t*-statistic that follows Student's distribution defined as:

$$
t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}
$$

where *n* is the record length of the AMS. The hypothesis is rejected when the absolute value of the computed *t-*statistic is greater than the critical value obtained from Student's distribution with (*n* - 2) degrees of freedom and exceedance probability of  $\alpha/2$  %, where  $\alpha$  is the significance level. The sign of the *t*-statistic indicates the direction of the trend, positive or negative.

Non-parametric tests have advantages over parametric tests since they make no assumption of probability distribution and are performed without specifying whether trend is linear or nonlinear. They are also more resilient to outliers in data because they do not operate on data directly. One of the disadvantages of non-parametric tests is that they do not account for the magnitude of the data. The Mann-Kendall test was selected among various non-parametric tests because it can accommodate missing values in a time series, which was a frequent occurrence in the AMS data. The Mann-Kendall test compares the relative magnitudes of annual maximum data. If annual maximum values are indexed based on time, and  $x<sub>i</sub>$  is the annual maximum value that corresponds to year  $t<sub>i</sub>$ , then the Mann-Kendall statistic is given by:

$$
S = \sum_{k=1}^{n-1} \sum_{i=k+1}^{n} sign(x_i - x_k)
$$

The test statistic *Z* is then computed using a normal approximation and standardization of the statistic *S*. The null hypothesis that there is no trend in the data is rejected at significance level  $\alpha$  if the computed *Z* value is greater, in absolute terms, than the critical value obtained from a standard normal distribution that has probability of exceedance of  $\alpha/2$  %. The sign of the statistic indicates the direction of the trend, positive or negative.

In addition to an at-station trend analysis, the relative magnitude of any trend in AMS for each of four climate regions (see Figure 4.1.2) as a whole was assessed by linear regression techniques. 1 hour and 1-day station-specific AMS for stations with at least 70 years of data for the 1-day duration and with at least 40 years of data for the 1-hour duration were rescaled by corresponding mean annual maximum values and then regressed against time, where time was defined as year of occurrence minus 1900. The regression results from all stations were tested against a null hypothesis of zero serial correlation (zero regression slopes).

#### **2. Trend analysis results and conclusion**

The stationarity assumption was tested by applying a parametric *t*-test and non-parametric Mann-Kendal test for trends in means and the Levene's test for trends in variance in the 1-day and 1-hour AMS data at 5% significance level. For the 1-day duration, testing was done on stations with at least 70 years of data; for the 1-hour duration, the minimum number of data years was lowered to 40 to increase the sample size. 205 and 407 stations satisfied the record length criterion for the 1-hour duration and 1-day duration, respectively. For 1-hour, based on the Levene's test using two subgroups of equal length, the hypothesis that variance did not change could not be rejected for any of the stations. The *t*-test and Mann-Kendall test indicated no statistically significant trends in the mean at about 89% and 87% of stations, respectively. In the 1-day dataset, Levene's test indicated non-homogeneous variance in less than 5% of stations. Based on *t*-test and Mann-Kendall test results, respectively, no trends were detected at about 87% and 83% of stations, respectively. More details are provided in Table A.2.1. The spatial distribution of the results for all three tests for 1-hour and 1-day AMS are shown in Figures A.2.1 and A.2.2, respectively. Small clusters of stations where

tests indicated positive trends are often due to AMS data sampled from the same storm events at several nearby locations.

	1-hour			1-day			
Number of stations	t-test	<b>Mann-</b> Kendall test	Levene's test	t-test	<b>Mann-</b> Kendall test	Levene's test	
no trend	183	179	205	353	336	387	
positive trend	19	21	$\theta$	46	63	20	
negative trend	3	5		8	8		
Total	205	205	205	407	407	407	

Table A.3.1. Trend analysis results for 1-hour and 1-day AMS data.

Results from the regional trend analysis also indicated that the null hypothesis, that there are no trends in AMS, could not be rejected at the 5% significance level for either climate region for the 1 hour and 1-day durations.

Because tests at both, the 1-hour and 1-day durations indicated no statistically significant trends in the data, the assumption of stationary AMS was accepted for this project area and no adjustment to AMS data was recommended.



Figure A.2.1. Spatial distribution of results of *t*-, Mann-Kendall, and Levene's tests for 1-hour AMS. Circles were used to present *t*-test results and plus signs were used to present Mann-Kendal test results. Red color indicates positive trends, green no trend, and blue negative trends. There were no stations where Levene's test detected non-homogeneous variance.



Figure A.2.2. Same as in Figure A.2.1, but for 1-day duration. Yellow circles show locations where Levene's test detected changes in variance.

NOAA Atlas 14 Volume 9 Version 2.0 A.2-5

#### **Appendix A.3 PRISM report**

*[The results shown in this Appendix apply for Volumes 8 and 9. This report was formatted by HDSC.]*

# **Final Report Production of Mean Annual Maximum Grids for the Midwestern and Southeastern Regions Using a Specifically Optimized PRISM System**

**Prepared for** National Weather Service, Hydrologic Design Service Center Silver Spring, Maryland

> **Prepared by** Christopher Daly PRISM Climate Group Oregon State University

> > February 2013

## **1. Project Goal**

The Hydrometeorological Design Studies Center (HDSC) within the Office of Hydrologic Development of NOAA's National Weather Service is updating precipitation frequency estimates for the Midwest and Southeast regions (hereafter referred to as MWSE). In order to complete the spatial interpolation of point estimates, HDSC requires spatially interpolated grids of MAM (Mean Annual Maximum) precipitation. The contractor, the PRISM Climate Group at Oregon State University (OSU), was tasked with producing a series of grids for rainfall frequency estimation using an optimized system based on the Parameterelevation Regressions on Independent Slopes Model (PRISM) and HDSC-calculated point estimates for the MWSE.

#### **2. Background**

HDSC used L-moment based regional frequency analysis approach to estimate precipitation frequencies. In this approach, the mean of the underlying precipitation frequency distribution is estimated at point locations with a sufficient history of observations. The form of the distribution and its parameters are estimated regionally. Once the form of the distribution has been selected and its parameters have been estimated, precipitation frequency estimates can be computed from grids of the MAM. The grids that are the subject of this report are spatially interpolated grids of the point estimates of the MAM for various precipitation durations. The point estimates of the MAM were provided by HDSC. HDSC selected an appropriate precipitation frequency distribution along with regionally estimated parameters and used this information with the grids of the MAM to derive grids of precipitation frequency estimates.

The PRISM Climate Group has performed similar work previously to produce spatially interpolated MAM grids for updates of precipitation frequency estimates in the Semiarid Southwest United States, the Ohio River Basin and Surrounding States, Puerto Rico/US Virgin Islands, Hawaiian Islands, California, and Alaska study areas.

## **3. Report**

This report describes tasks performed to produce mean annual maximum (MAM) grids for 17 precipitation durations: 15 and 30 minutes; 1, 2, 3, 6, and 12 hours; and 1, 2, 3, 4, 7, 10, 20, 30, 45, and 60 days for the MWSE. The tasks described were not necessarily performed in the order described, nor were they performed just once. The process was dynamic and had numerous feedbacks.

#### **3.1. Adapting the PRISM system**

The PRISM modeling system was adapted for use in this project after a small investigation was performed for the Semiarid Southwest United States, and subsequently used in the Ohio River Basin and Surrounding States, Puerto Rico/Virgin Islands, Hawaiian Islands, California, and Alaska study areas. This investigation and adaptation procedure is summarized below.

PRISM is a knowledge-based system that uses point data, a digital elevation model (DEM), and many other geographic data sets to generate gridded estimates of climatic parameters (Daly et al. 1994, 2002, 2003, 2006, 2008) at monthly to daily time scales. Originally developed for precipitation estimation, PRISM has been generalized and applied successfully to temperature, among other parameters. PRISM has been used extensively to map precipitation, dew point, and minimum and maximum temperature over the United States, Canada, China, and other countries. Details on PRISM formulation can be found in Daly et al. (2002, 2003, 2008), which are available from http://prism.oregonstate.edu/docs/.

Adapting the PRISM system for mapping precipitation frequencies required an approach slightly different than the standard modeling procedure. The amount of station data available to HDSC for precipitation frequency was much less than that available for high-quality precipitation maps, such as the peer-reviewed PRISM 1971-2000 mean precipitation maps (Daly et al. 2008). Data sources suitable for long-term mean precipitation but not for precipitation frequency included snow courses, short-term COOP stations, remote storage gauges, and others. In addition, data for precipitation durations of less than 24 hours were available from hourly precipitation stations only. This meant that mapping precipitation frequency using HDSC stations would sacrifice a significant amount of the spatial detail present in the 1971-2000 mean precipitation maps.

A pilot project to identify ways of capturing more spatial detail in the precipitation frequency maps was undertaken. Early tests showed that mean annual precipitation (MAP) was an excellent predictor of precipitation frequency in a local area, much better than elevation, which is typically used as the underlying, gridded predictor variable in PRISM applications. In these initial tests, the DEM, the predictor grid in PRISM, was replaced by the official USDA digital map of MAP for the lower 48 states (USDA-NRCS 1998, Daly et al. 2000). Detailed information on the creation of the USDA PRISM precipitation grids is available from Daly and Johnson (1999). MAP was found to have superior predictive capability over the DEM for locations in the southwestern US. The relationships between MAP and precipitation frequency were strong because many of the effects of various physiographic features on mean precipitation patterns had already been incorporated into the MAP grid from PRISM. Preliminary PRISM maps of 2-year and 100-year, 24-hour precipitation were made for the Semiarid Southwest and compared to hand-drawn HDSC maps of the same statistics. Differences were minimal, and mostly related to differences in station data used.

Further investigation found that the square-root transformation of MAP produced somewhat more linear, tighter and cleaner regression functions, and hence, more stable predictions, than the untransformed values; this transformation was incorporated into subsequent model applications. Squareroot MAP was a good local predictor of not only longer-duration precipitation frequency statistics, but for short-duration statistics, as well. Therefore, it was determined that a modified PRISM system that used square-root MAP as the predictive grid was suitable for producing high-quality precipitation frequency maps for this project.

For this study, an official USDA grid of MAP for the study region (1981-2010 average) was used (Figure 1). This grid was developed under funding from the USDA Natural Resources Conservation Service, and is an update to the 1971-2000 grids described in Daly et al. (2008).

#### **3.2. PRISM configuration and operation for the MWSE**

In general, PRISM interpolation consists of a local moving-window regression function between a predictor grid and station values of the element to be interpolated. The regression function is guided by an encoded knowledge base and inference engine (Daly et al., 2002, 2008). This knowledge base/inference engine is a series of rules, decisions and calculations that set weights for the station data points entering the regression function. In general, a weighting function contains knowledge about an important relationship between the climate field and a geographic or meteorological factor. The inference engine sets values for input parameters by using default values, or it may use the regression function to infer grid cell-specific parameter settings for the situation at hand. PRISM acquires knowledge through assimilation of station data, spatial data sets such as MAP and others, and a control file containing parameter settings.

The other center of knowledge and inference is that of the user. The user accesses literature, previously published maps, spatial data sets, and a graphical user interface to guide the model application. One of the most important roles of the user is to form expectations for the modeled climatic patterns, i.e., what is deemed "reasonable." Based on knowledgeable expectations, the user selects the station weighting algorithms to be used and determines whether any parameters should be changed from their default values. Through the graphical user interface, the user can click on any grid cell, run the model with a given set of algorithms and parameter settings, view the results graphically, and access a traceback of the decisions and calculations leading to the model prediction.

For each grid cell, the moving-window regression function for MAM vs. MAP took the form

$$
MAM value = \beta_1 * sqrt(MAP) + \beta_0
$$
 (1)

where  $\beta_l$  is the slope and  $\beta_0$  is the intercept of the regression equation, and MAP is the grid cell value of mean annual precipitation.

Upon entering the regression function, each station was assigned a weight that is based on several factors. For PRISM MAP mapping (used as the predictor grid in this study), the combined weight of a station was a function of distance, elevation, cluster, vertical layer, topographic facet, coastal proximity, and effective terrain weights, respectively. A full discussion of the general PRISM station weighting functions is available from Daly et al. (2008).

Given that the MAP grid incorporated detailed information about the complex spatial patterns of precipitation, only a subset of these weighting functions was needed for this study. For the MWSE, the combined weight of a station was a function of distance and clustering, respectively. A station is downweighted when it is relatively distant from the target grid cell, or when it is clustered with other stations (which can lead to over-representation).

The moving-window regression function was populated by station data provided by the HDSC. A PRISM GUI snapshot of the moving-window relationship between sqrt(MAP) and 24-hour MAM in south-central Colorado is shown in Figure 2.

There were relatively few stations with data for durations of 12 hours or less from which to perform the interpolation. In addition, it was clear that the spatial patterns of durations of 12 hours or less could be very different than those of durations of 24 hours or more. This issue was encountered in a previous study for Puerto Rico. During that study the following procedure was developed, and adopted here:

- (1) Convert available  $\leq 12$ -hour station values to an MAM/24-hr MAM ratio (termed R24) by dividing by the 24-hour values;
- (2) using the station R24 data in (1), interpolate R24 values for each  $\leq$  12-hour duration (15, 30, and 60 minutes; and 2, 3, 6, and 12 hours) using PRISM in inverse-distance weighting mode;
- (3) using bi-linear interpolation from the cells in the R24 grids from (2), estimate R24 at the location of each station having data for  $\geq$  24-hour durations only;
- (4) multiply the estimated R24 values from (3) by the 24-hour value at each  $\geq$  24-hour station to obtain estimated ≤ 12-hour values;
- (5) append the estimated stations from (4) to the  $\leq$  12-hour station list to generate a station list that matches the density of that for  $\geq$  24 hours; and
- (6) interpolate MAM values for  $\leq 12$ -hour durations with PRISM, using MAP as the predictor grid.

Investigation of the little available data failed to provide convincing evidence that the spatial patterns of R24 values in the MWSE were strongly affected by coastal proximity, topographic facets, or other factors. Therefore, the slope of the moving-window regression function for R24 vs. MAP of the form

$$
R24 = \beta_I * \sqrt{MAP} + \beta_0
$$
 (2)

was forced to zero everywhere. This meant that the interpolated value of R24 was a function of distance and cluster weighting only (essentially inverse-distance weighting).

Relevant PRISM parameters for applications to 60-minute R24 and 24-hour MAM statistics are listed in Tables 1 and 2, respectively. Further explanations of these parameters and associated equations are available in Daly et al. (2002, 2008).

The values of radius of influence  $(R)$ , the minimum number of total  $(s<sub>i</sub>)$  stations required in the regression were based on information from user assessment via the PRISM graphical user interface, and on a jackknife cross-validation exercise, in which each station was deleted from the data set one at a time, a prediction made in its absence, and mean absolute error statistics compiled (see Results section).

The input parameter that changed readily among the various durations was the default slope ( $\beta_{1d}$ ) of the regression function. Slopes are expressed in units that are normalized by the average observed value of the precipitation in the regression data set for the target cell. Evidence gathered during PRISM model development indicates that this method of expression is relatively stable in both space and time (Daly et al. 1994).

Bounds are put on the slopes to minimize unreasonable slopes that might occasionally be generated due to local station data patterns; if the slope is out of bounds and cannot be brought within bounds by the PRISM outlier deletion algorithm, the default slope is invoked (Daly et al., 2002). The maximum slope bound was set to a uniformly high value of 30.0, to accommodate a large range of valid slopes; lower values were not needed to handle extreme values, because all values were within reasonable ranges. Slope default values were based on PRISM diagnostics that provided information on the distribution of slopes across the modeling region. The default value was set to approximate the average regression slope calculated by PRISM. For these applications, default slopes typically increased with increasing duration (Table 3). In general, the longer the duration, the larger the slope. This is primarily a result of higher precipitation amounts at the longer durations, and the tendency for longer-duration MAM statistics to bear a stronger and steeper relationship with MAP than shorter-duration statistics.

#### **3.3. Preparation and review of draft grids**

Draft grids for the 60-minute, 24-hour and 10-day durations were produced and made available to HDSC for evaluation. All of the necessary station data were provided by HDSC. The process began with a careful scrutiny of the station data and PRISM behavior. A version of PRISM which predicts for stations locations in the absence of each station (termed jackknifing) was run, and stations that were difficult for PRISM to predict for were identified, and sent to HDSC for review. HDSC removed the stations, modified their values, or determined that the stations were accurate as-is. This process was performed iteratively, until an acceptable station data set was produced. The draft PRISM grids were subsequently completed and submitted to HDSC for review. HDSC submitted the draft PRISM grids for external review, and revised the station data accordingly.

## **3.4. Final grids**

Having found the revised draft grids acceptable, HDSC requested that grids for all durations be completed. Before delivering the final grids to HDSC, the PRISM Climate Group checked them for internal consistency. In other words, the value of the MAM at each grid point for each duration must have been greater than the value for shorter durations at the same grid point. If an inconsistency of this nature occurred, the convention was to start with the 24 duration as a baseline, and set longer durations to slightly higher values and shorter durations to slightly lower values.

The final delivered grids inherited the spatial resolution of the latest 1981-2010 PRISM mean annual precipitation grids for MWSE, which is 30 arc-seconds (~800 meters). The grid cell units are in mm\*100. Final MAM grids delivered to HDSC are as follows:

> 15-minute 30-minute 60-minute 2-hour 3-hour 6-hour 12-hour 24-hour 48-hour 3-day 4-day 7-day 10-day 20-day 30-day 45-day 60-day Total: 17

# **3.5. Performance evaluation**

PRISM cross-validation statistics for 60-minute/24-hour MAM ratio and the 60-minute and 24-hour MAM intensities were compiled and summarized in Table 4. These errors were estimated using an omitone jackknife method, where each station is omitted from the data set, estimated in its absence, then replaced. Since the 60-minute/24-hour MAM ratio was expressed as a percent, the percent bias and mean absolute error are the given as the bias and MAE in the original percent units (not as a percentage of the percent).

For the 60-minute/24-hour MAM ratio, the overall bias was near zero and the mean absolute error (MAE) about 3 percent. For the 60-minute, 24-hour, and 10-day MAM intensities, biases were less than 0.25 percent, and the MAE varied around the 4 percent mark. Biases for the 15- and 30-minute durations were slightly negative (-1.8 and -0.19 percent, respectively), while those for the other durations were slightly positive (ranging between 0.08 and 0.25 percent). MAEs for all durations were less than 5 percent, with most less than 4 percent. Given the lack of independent data at durations of less than 24 hours, one would have expected the 15-minute to 12-hour MAM errors to be substantially higher than those for the 24-hour to 60-day MAMs. A likely reason why this was not the case was that the addition of many synthesized stations, derived from a PRISM interpolation of R24 values, resulted in a station data set that was spatially consistent, and thus, somewhat easier to interpolate with each station deleted from the data set. Therefore, there is little doubt that the true interpolation errors for the 60-minute MAM are higher than those shown in Table 4.

Table 1. Values of relevant PRISM parameters for interpolation of 60-minute/24-hour mean annual maximum ratio (60-minute R24) for the MWSE. See Daly et al. (2002) for details on PRISM parameters.



 $*$  Expands to encompass minimum number of total stations desired in regression  $(s_t)$ .

Slopes are expressed in units that are normalized by the average observed value of the precipitation in the regression data set for the target cell. Units here are 1/[sqrt(MAP(mm))\*1000].



Table 2. Values of relevant PRISM parameters for modeling of 24-hour mean annual maximum statistics for the MWSE. See Daly et al. (2002) for details on PRISM parameters.

\* Expands to encompass minimum number of total stations desired in regression (*st*). <sup>+</sup> Slopes are expressed in units that are normalized by the average observed value of the precipitation in the regression data set for the target cell. Units here are 1/[sqrt(MAP(mm))\*1000].

Table 3. Values of PRISM slope parameters for modeling of MAM statistics for the MWSE for all durations. For durations of 12 hours and below, station data were expressed as the ratio of the given duration's MAM value to the 24-hour MAM value, and interpolated; this was followed by an interpolation of the actual MAM values. See text for details. See Table 1 for definitions of parameters.



Table 4. PRISM cross-validation errors for 60-minute/24-hour MAM ratio and 24-hour MAM applications to the MWSE. Since the 60-minute/24-hour MAM ratio was expressed as a percent, the percent bias and mean absolute error are the given as the bias and MAE in the original percent units (not as a percentage of the percent).



Figure 1. PRISM 1981-2010 mean annual precipitation (MAP) grid for the MWSE region.

Figure 2. PRISM GUI snapshot of the moving-window weighted regression between the square root of mean annual precipitation and 24-hour mean annual maximum precipitation (MAM) in south-central Colorado.

# **4. References**

- Barnes, S. L. 1964. A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology,* 3:396-409.
- Daly, C. 2006. Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology,* 26: 707-721. [http://prism.oregonstate.edu/pub/prism/docs/intjclim06-guidelines](http://prism.oregonstate.edu/pub/prism/docs/intjclim06-guidelines-daly.pdf)[daly.pdf](http://prism.oregonstate.edu/pub/prism/docs/intjclim06-guidelines-daly.pdf)
- Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris. 2002. A knowledge-based approach to the statistical mapping of climate. *Climate Research,* 22: 99-113. [http://prism.oregonstate.edu/pub/prism/docs/climres02-kb\\_approach\\_statistical\\_mapping-daly.pdf](http://prism.oregonstate.edu/pub/prism/docs/climres02-kb_approach_statistical_mapping-daly.pdf)
- Daly, C., Halbleib, M., Smith J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., and Pasteris, P.A. 2008. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States*. International Journal of Climatology*, 28: 2031-2064. [http://prism.oregonstate.edu/pub/prism/docs/intjclim08-physiographic\\_mapping-daly.pdf](http://prism.oregonstate.edu/pub/prism/docs/intjclim08-physiographic_mapping-daly.pdf)
- Daly, C., E. H. Helmer, and M. Quinones. 2003. Mapping the climate of Puerto Rico, Vieques, and Culebra. *International Journal of Climatology*, 23: 1359-1381. [http://prism.oregonstate.edu/pub/prism/docs/jclim03-map\\_climate\\_PR.pdf](http://prism.oregonstate.edu/pub/prism/docs/jclim03-map_climate_PR.pdf)
- Daly, C., R. P. Neilson, and D. L. Phillips. 1994. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology,* 33: 140-158. [http://prism.oregonstate.edu/pub/prism/docs/jappclim94-modeling\\_mountain\\_precip-daly.pdf](http://prism.oregonstate.edu/pub/prism/docs/jappclim94-modeling_mountain_precip-daly.pdf)
- Daly, C., J.I. Smith, M. D. Halbleib, W.P. Gibson, and P. Sousanes. 2009. 1971-2000 mean monthly and annual precipitation spatial climate data set for the State of Alaska. USDOI National Park Service. Accessible at [http://irma.nps.gov/,](http://irma.nps.gov/) search term: "mean precipitation".
- Daly, C., G. H. Taylor, W. P. Gibson, T. W. Parzybok, G. L. Johnson, P. Pasteris. 2000. High-quality spatial climate data sets for the United States and beyond. *Transactions of the American Society of Agricultural Engineers* 43: 1957-1962. [http://prism.oregonstate.edu/pub/prism/docs/asae00](http://prism.oregonstate.edu/pub/prism/docs/asae00-spatial_climate_datasets-daly.pdf) spatial climate datasets-daly.pdf
- USDA-NRCS, 1998. *PRISM Climate Mapping Project--Precipitation. Mean monthly and annual precipitation digital files for the continental U.S*. USDA-NRCS National Cartography and Geospatial Center, Ft. Worth TX. December, CD-ROM.

#### **Appendix A.4 Peer review comments and responses**

A peer review of preliminary results for the Volume 9 precipitation frequency project was carried out during a five week period starting on October 15, 2012. The request for review was sent via email to the over 700 members of the HDSC list-server from all over the United States and other interested parties. Potential reviewers were asked to evaluate the reasonableness of point precipitation frequency estimates as well as their spatial patterns. The review included the following items:

- a. Metadata for stations whose data were used to prepare mean annual maximum precipitation maps and/or in precipitation frequency analysis. The table included information on station name, state, source of data, assigned station ID, latitude, longitude, elevation, and period of record. It also showed if the station was merged with another station, if the station was colocated with another station with a different ID, and if metadata at the station were changed. (Station IDs were assigned by HDSC and do not match station IDs assigned by the agency that provided the data, except for National Climatic Data Center.)
- b. Metadata for stations whose data were collected, but not used in the analysis. The table contained metadata for stations that were examined, but not used, with brief comments on why the data were not used. Generally, stations were not used either because there was another station with a longer period of record nearby, station data were assessed as not reliable for this specific purpose, or the station's period of record was not long enough and it was not a candidate for merging with any nearby station.
- c. At-station depth-duration-frequency (DDF) curves for 60-minute to 10-day durations and for 2-year to 100-year ARIs.
- d. Maps of spatially-interpolated estimates of mean annual maxima for 60-minute, 24-hour and 10-day durations.
- e. Maps of spatially-interpolated precipitation frequency estimates for 60-minute, 24-hour and 10-day durations and for 2-year and 100-year average recurrence intervals (ARIs).

Comments were received from eight individuals or offices and agencies including Water Management District Offices and Weather Forecast Offices. All reviewers' comments and HDSC's responses (in red) are shown below. The comments and their respective HDSC responses have been separated into four categories:

- 1. Station metadata
- 2. At-station precipitation frequency estimates
- 3. Precipitation frequency grids and cartographic maps
- 4. Miscellaneous.

# **1. Station metadata**

1.1 [Alabama, Florida, Mississippi] After reviewing the Metadata spreadsheet I would suggest the following lat/lon updates for the sites listed. The lat/lon data provided are from the most recent B-44 forms on file at NCDC.



NOAA Atlas 14 Volume 9 Version 2.0 A.4-1

Greenville, AL..31.7901/-86.6087 Highland Home, AL.................................31.8814/-86.2503.....Also note that Highland Home closed in May 2012 and was reestablished in July 2012 as Highland Home #2 at 31.9477/-

86.3131 Buckatunna 1NE, MS...............................31.5476/-88.5172 Wiggins, MS...30.8713/-89.1215 Wiggins Ranger Station, MS...................30.8502/-89.1571

Niceville, FL...30.5316/-86.4928

*We implemented recommended coordinates for: Claiborne Lock and Dam, AL (01-1690); Dauphin Island, AL (01-2172); Evergreen, AL (01-2758); Greenville, AL (01-3519); Highland Home, AL (01-3816); Buckatunna 1NE, MS (22-1174); Wiggins, MS (22-9639); Niceville, FL (08-6240). Wiggins Ranger Station (22-9648 which had hourly and 15-minute data) was colocated with Wiggins (22-9639 which had daily data) and is plotted at that location in the final output.* 

- 1.2 [Florida] Period of Record: The SJRWMD used the longest period of record available for a primary station between 1914 and 2008. We merged together stations from nearby (including some SJRWMD rain gauges) to develop the longest period of record possible. Only stations with greater than 50 years of record were used. It appears that NOAA/NWS merged some sites together but it was difficult to determine from the station metadata. The station period of records listed in the NOAA/NWS station metadata had wide variation in start and end dates when, for example, the currently active stations would be expected to have the same end date. Different periods of record for individual stations can influence an analysis. Current NOAA/NWS documentation states that a weighting between stations was used to compare different periods of records, but the actual weighting process was not described. *For the analysis, we collected data from various sources, so, the currently active stations from different data providers may have different end dates. Since precipitation frequency analysis is a statistical approach, sample size is an important factor to consider. Statistics calculated from a small sample could be greatly skewed by a single estimate. That is especially true for higher order moments, and consequently, estimates for more rare frequencies (50-year ARI or more). Therefore, we strived to have longer records at stations by merging nearby stations and also grouped stations using a regional approach when computing frequency estimates. During our initial station screening we looked at nearby stations (usually within 3 miles, but up to 5 miles in flat terrain, with consideration to distance from coast, climatological characteristics of extreme precipitation, etc.) to see if they could be merged to form a single longer record. Generally, for record length, 30 years of data (after merging) is considered an acceptable minimum for a meaningful statistical analysis. For sub-daily durations, we lowered that threshold to 20 years to retain as many stations as possible. We also increased sample size, and ultimately the reliability of the estimates, by using a regional frequency analysis approach, where, except for the first L-moment which is station specific, moments are regional estimates obtained by averaging corresponding station-specific estimates from all stations in a region, weighted by record lengths. A complete description of the NOAA Atlas 14 methods is available in Section 4.*
- 1.3 [Florida] Data Records: SJRWMD used the published data for each individual station and infilling missing data from the average of the three nearest stations. We did not attempt to do quality assurance of the outliers. NOAA/NWS states that a quality assurance (verification) process was used to identify outliers and remove if necessary. NOAA/NWS does not mention anything about infilling missing records so assume this was not done.

NOAA Atlas 14 Volume 9 Version 2.0 A.4-2

*We do not infill missing data at a station unless it was a significant amount recorded at colocated station. Estimating missing values from surrounding stations provides added weight to those surrounding stations without providing additional raw information for the statistical analysis. However, we carefully investigated AMS data at both high and low extremities, as they can considerably affect precipitation frequency estimates. We corrected or removed outliers from the AMS if due to measurement errors (see Section 4.5 for more details). Additionally, since we use a regional frequency analysis approach, if a given station missed a particular event, it would be included in the analysis as part of AMS at nearby station(s) that captured the event, and therefore would contribute to target station's estimates.*

## **2. At-station precipitation frequency estimates**

2.1 [Florida] Estimation of 24-hour Events from Daily Data: SJRWMD used a method of adding 60 percent of the largest surrounding daily precipitation to a daily maximum total to determine the 24-hour event total (Hershfield, 1963). NOAA/NWS used values from a table in the July quarterly report to multiply by the daily maximum total to determine the 24-hourly event total. This was 1.12 for the 24-hour event. We developed an analysis based on using hourly data for the Orlando International Airport gauge to calculate observation day and running 24-hour storm totals and aligning the peak values from each dataset. This analysis showed that the 1.12 factor was a good model to predict the 24-hour rain total from the observation day total (Figure 1 and Figure 2). Hershfield (1963) used 50 percent of the maximum adjacent daily precipitation, but an early analysis by SJRWMD indicated that 60 percent is a better value for Northeast Florida. We redid this analysis and it is indeed the case, at least for the Orlando International Airport gauge. The modified Hershfield (1963) model used by SJRWMD is a better predictor of 24 hour events from observation day values (Figure 3) than the 1.12 factor approach. See attached spreadsheets, 'Orlando\_obsday\_vs\_24hr\_storm.xlsx' and 'Orlando\_maxadj\_vs\_24hr\_storm.xlsx' for additional detail.



**Figure 1: Simple factor determination to estimate 24-hour storm totals from observation day totals.**



**Figure 2: Application of 1.12 factor to estimate 24-hour storm totals from observation day totals.**



## **Figure 3: Application of adding 0.6 of maximum adjacent observation day value to estimate 24 hour storm totals from observation day totals.**

*Based on our investigation, the suggested approach worked well for some stations, but there were about the same number of cases where it performed worse than our current approach. Also, we looked at spatial patterns in the at-station 24-hour to 1-day average ratios and did not see any spatial coherence in northeast Florida or in the entire project area. Therefore, we decided to use one average factor for the project area.* 

2.2 [Florida] Frequency Analysis: The SJRWMD used a Log Pierson Type III Distribution for the 2-year, 5-year, 10-year, and most 25-year storm events. A manual adjustment using Weibull plotting positions was used to evaluate the 200-year, 100-year, 50-year, and some 25-year storm events since Log Pierson distributions often crossed during these higher storm events. NOAA/NWS used a Generalized Extreme Value Distribution.

*We tested seven distributions (3-parameter Generalized Extreme Value - GEV, Generalized Normal, Generalized Pareto, Generalized Logistic and Pearson Type III distributions; 4 parameter Kappa distribution; and 5-parameter Wakeby distribution) using the Hosking and Wallis test based on L-moment statistics (for 3-parameter distributions), Kolmogorov*‐*Smirnov test and χ2-test. We tested across all durations from 15-minute to 60-day. We also inspected* 

NOAA Atlas 14 Volume 9 Version 2.0 A.4-4
*probability plots of all selected distributions for three base durations (1-hour, 1day and 10 day). Based on the results, the GEV distribution provided an acceptable fit more times than any other distribution across all durations. Also, GEV is a distribution generally recommended for analysis of extreme events. For consistency, we used the same distribution for all durations and frequencies. Please see Section 4.6.3 for more information.*

2.3 [Florida] Regionalization: For consistency, the SJRWMD looked at surrounding stations, however this was primarily when looking at large storm events. The evaluation was to establish how much weight to put on outliers and between coastal stations. NOAA/NWS used a 10 station regionalization analysis to evaluate nearby stations. A better regionalization approach in Florida would be to group stations according to distance from the coast rather than simple proximity to each other. It is not clear from the available NOAA/NWS documentation exactly how the 10 stations were chosen.

*The regionalization process is described fully in Section 4.6.2. We began by grouping the target station with the nearest 10 stations, but then stations were added to or removed from the target station's region based on examination of their distance, elevation differences, inspection of their locations with respect to the coast, and assessment of similarities/ dissimilarities in the progression of relevant L-moment statistics across durations compared with other stations in the region. We also strived to maintain an acceptable amount of daily and hourly data years in each region. Based on your comment we checked some regions in Florida particularly along the coast to ensure proper regionalization.*

2.4 [Florida] Why do the NOAA/NWS 2-year storm events have a much better correlation to the SJRWMD's Mean Annual Maximum event than the SJRWMD 2-year storm event? This could be explained by the different observation total to storm total models used, where for smaller storm events it is less likely that the maximum adjacent model would alter the daily value, whereas the factor model applies regardless of the adjacent values.

*The most likely reason for this difference is that, in the peer review, we presented estimates for the 2-year average recurrence interval, which means that these are PDS-based results (AMSbased results are presented as annual exceedance probabilities). PDS-based precipitation frequency estimates were calculated indirectly from the Langbein's formula that transforms a PDS-based average recurrence interval (ARI) to an annual exceedance probability (AEP). PDS-based 2-year estimates are approximately equal to the 2.54-year AMS-based estimates. Please see Section 4.1 for more details.* 

2.5 [Florida] To review the NOAA/NWS precipitation frequency analysis we initially compared the Depth Duration Frequency (DDF) curves to our updated analysis [based on methodologies from Technical Publications SJ-86-3 (SJRWMD, 1986a) and SJ88-3 (SJRWMD, 1988)] for stations which were analyzed in common. The spreadsheet with this comparison is 'NOAA\_Rain\_Frequency\_Comparisons\_Final.xlsx' and is attached to this e-mail. The numbers used for the NOAA/NWS analysis come from approximating precipitation from the NOAA/NWS DDF curves. These may differ slightly from the actual numbers used to develop these curves. In general we found the following differences: 1.The NOAA/NWS 2-year event data for both the 24-hour and 10 day storm events were 0.4 and 0.6 inches higher respectively than for the SJRWMD analysis. The NOAA/NWS 2-year event matched much better with what we consider the Mean Annual Maximum precipitation event. The SJRWMD calculates the Mean Annual Maximum precipitation event by averaging

the annual maximum precipitation of that storm length for the period of record at a rainfall

station. An analysis by SJRWMD staff indicated that the Mean Annual Maximum precipitation event approximates the 2.33 frequency precipitation event.

2.The largest average differences occurred for the 100-year 24 hour storm event. The NOAA/NWS analysis average was 0.9 inches lower than the SJRWMD analysis but these differences varied widely between individual stations. Some of these large differences appear regionally along the coastal Northeast Florida from Fernandina Beach to St. Augustine. 3.For the 100-year, 10 day storm event the NOAA/NWS analysis average difference was 0.5 inches lower than the SJRWMD analysis. The differences between individual stations had an even wider variation than for the 100-year, 24 hour storm event.

*For the first comment, please see our response to comment 2.4 regarding the 2-year estimates. For the 24-hour and 10-day 100-year estimates, we investigated cases with the largest differences between our peer reviewed estimates and the SJRWMD estimates. Since the peer review, we re-evaluated the regionalization to ensure consistency between nearby stations, particularly along the coast, and to ensure that nearby stations with high maximum events were included as appropriate. As a result, overall patterns improved and some 24-hour 100-year estimates increased. For example, the estimate at Bithlo (08-1565) increased from 9.9" to 11.6", Daytona Beach (08-2150) estimate increased from 11.3" to 12.5", and St. Augustine (08-7826) estimate increased from 10.8" to 12.4". Similar changes were seen at 10-day, with Melbourne (08-5612) increasing from 16.7" to 17.6" and St. Augustine from 17.2" to 18.6", for example.*

*However, there are still locations where estimates from two projects will be considerably different. For example, SJRWMD 24-hour 100-year estimates for Jacksonville Airport (08- 4358) and Beach (08-4366), which are only separated by approximately 23 miles, are 11.0" and 15.8" respectively (in contrast to our estimates of 12.2" and 12.6" respectively). We could not find any justification in the data for such significant differences. Our data show that these two stations have very similar AMS.*

2.6 Is there additional documentation from NOAA/NWS that would help the SJRWMD staff understand the differences outlined earlier?

1.Need a description of how NOAA/NWS evaluated stations with different period of records and used them to develop consistent DDF curves for each station.

2.Need to know what data NOAA/NWS removed or modified and why they made these decisions for each station.

3.What stations in Northeast Florida does NOAA/NWS have grouped for their regionalization technique? Do these regions adequately represent differences that often occur for coastal stations compared to inland stations in Northeast Florida?

*Full descriptions of our methods are provided in the accompanying documentation. However, we'll briefly address each concern here.*

*1. Please see our response to comment 1.2.* 

*2. We quality controlled low and high outliers in the AMS data. High outliers were only removed if they could be proven erroneous through inspection of nearby data and published data records. Low outliers, which usually occurred in years with missing or accumulated data, were typically removed from the data set. We do not provide the details of this analysis for each station, but we do provide the final quality-controlled AMS data through links on the PFDS web page.*

*3.We have ensured that coastal stations were grouped together to properly reflect the coastal influence relative to inland stations.* 

2.7 [Florida] Evaluation of the Mean Annual Maximum between agencies would indicate differences in the raw data used at each station. Included in the 'NOAA\_Rain\_Frequency\_Comparisons\_Final.xlsx' spreadsheet in columns 'C' and 'U' on the 'Data' tab are the period of record maximum events from SJRWMD data for 24-Hr storms and 10-day storms at each station. *Since SJRWMD infilled missing data by using an average of the three closest stations, in addition to using the published data, we expect to see some differences in the raw data for stations where any infilling was done. We checked the annual maxima series for stations where differences were significant and noticed some inconsistencies. For example, at Clermont (08- 1641) the highest 24-hour annual maximum in the SJRWMD dataset is 14.8" versus our 8.2" (6.6" difference). However, there was not enough information to ascertain the reason for this difference, particularly since this station is only missing 1% of its data in our record. At Federal Point (08-2915), there is also a large difference of 4.7", where we have 12.6" and* 

*SJRWMD has 7.9". At St. Augustine, the largest 24-hour annual maximum, based on our data, is 12.6" compared to SJRWMD's 14.3". This station has a long reliable record of over 100 years of data and a review of the published NCDC data confirmed the observation at this station. We had no reason to question the quality of the NCDC data and did not make additional edits.*

## **3. Precipitation frequency grids and cartographic maps**

3.1 Regarding the display of the final data (isopluvials) and areas where interpolation may have to be used, do you have any recommendations?

*In the final deliverable, our estimates have been interpolated to grids with a spatial resolution of 30 arc-seconds so that values do not need to be interpolated from isopluvials (contour lines) on cartographic maps. We recommend using the [Precipitation Frequency Data Server](http://hdsc.nws.noaa.gov/hdsc/pfds/index.html) (PFDS) for obtaining estimates for a location rather than the cartographic maps. This can be done by clicking on a location on a map-based interface, entering coordinates, selecting a station from the drop-down list, or using the high resolution grids of precipitation frequency estimates provided there.*

3.2 [Florida] I have reviewed the rainfall data provided on the NOAA website and from spreadsheets that were furnished to the Southwest Florida Water Management District. Comparing the NOAA information to that previously generated by University of Central Florida who used only hourly and data, there are areas of significantly higher rainfall depths for all durations and return frequencies for the NOAA station data . Based on my review of the NCDC data, areas where significant departures between NOAA and UCF results appear to occur in areas where significant one-hour rainfall data was missing or where stations were moved from one location to another with only short-term record available. I agree that the infilling of the data is necessary to ensure that the rainfall is not under predicted. However, from a water management perspective where the rainfall depths are used for design, it will be important that the procedures used for infilling of missing data or extending magnitude rainfall be sufficiently addressed. Significant changes in the design storm depths will occur if the NOAA rainfall data is adopted. If you have any documentation on the record infilling or record extension that would be most helpful. *We did not infill missing data at a station unless it was a significant amount recorded at co-*

*located station. However, we carefully investigated AMS data at both high and low extremities, as they can considerably affect precipitation frequency estimate. We corrected or removed*

NOAA Atlas 14 Volume 9 Version 2.0 A.4-7

*outliers from the AMS if due to measurement errors. During our initial station screening we also looked at nearby stations (usually within 3 miles, with consideration to elevation differences, climatological characteristics of extreme precipitation, etc.) to see if they could be merged to form a single longer record. We explain the procedures used to quality control data and resulting estimates in Section 4. Differences between the UCF and NOAA Atlas 14 estimates may be caused by a combination of factors such as differences in approach and in the data used. We used regional approach that pools the information from all stations that were assigned to a target station's region to compute its estimates (see Section 4.6) versus the singlestation approach used in the UCF study. In addition to hourly stations, as used in the UCF study, we used daily stations that typically had much longer periods of record.* 

3.3 [Florida] I have extensively reviewed the one-hour data provided by the NCDC whereby there are significant missing data which when analyzed seems to bias the rainfall depths down at certain stations. For example the Lakeland gauge missed rainfall for several hurricanes. Based on the California Map Atlas procedures it appeared that correlations between stations and other station data was used to extend, merge or ensure complete records at a location. I have no issue with the procedures NOAA has implemented as long as they can be statistically and scientifically justified. The issue is more on the side of the District in justifying the increases in rainfall depths in Manatee, Sarasota, and Charlotte Counties along the coast. From the data it appears that the NOAA 100-year, 24-hour rainfall depths will increase between 1-2.5 inches for these areas which can have a pronounced effects on project design and FEMA floodplains in these areas. It is more the political implications of adopting new rainfall depths. Just want to make sure that the increases can be scientifically justified and documented, which from my perspective they can be.

*The differences between the new NA14 estimates and other estimates could be attributed to a number of factors, including, the difference in the number of stations used in analysis and their periods of record, quality control and data derivation procedures, and differences in frequency analysis approaches and interpolation techniques. We strive to use the latest techniques and the most complete data available. Section 7 provides a comparison of the updated estimates with previous NWS publications.* 

# **4. Miscellaneous**

- 4.1 Consider making the hypothetical rainfall grids available for download. *Final precipitation frequency grids for all durations and frequencies are available for download from the [Precipitation Frequency Data Server \(PFDS\).](http://hdsc.nws.noaa.gov/hdsc/pfds/index.html)*
- 4.2 I see you offer GIS data download options. Do you have plans to offer services of the data, so that they might be used in various displays? Not requiring me to download, process, and create my own services just to show this data side-by-side other relevant data (e.g. UDSM, forecast rainfall) would be tremendous.

*We do not have current plans to offer services of the data other than what is currently provided via the PFDS. We are open to suggestions for improvement of all aspects of our product.* 

4.3 I would very much like to have access (download) to the draft Vol 8 and 9 data, but I don't see any download options for the Vol 8 and 9 data I presume is fairly far along in maturity. Would it be possible for you to give me access to, say, the Arkansas data, if not the entire Southeast, to create a demo service, so that I might work it into a display? What I have planned is a ESRI

NOAA Atlas 14 Volume 9 Version 2.0 A.4-8

storytelling interface that would show, on one side, the record rainfall amounts for a period, and, on the other side, show a forecast rainfall display. Using the Swipe bar, one could easily assess if a record is forecast to be exceeded, provided I can time-match the record data with the forecast data length of time. In this example (slow to load, I'm guessing), I compare the current US Drought Monitor status, with HPC's 5-day forecast rainfall guidance: [http://www.srh.noaa.gov/rtimages/srh/stsd/Jack/swipe/DroughtRainfall.html.](http://www.srh.noaa.gov/rtimages/srh/stsd/Jack/swipe/DroughtRainfall.html) *This is a great example of a use of the product we provide; we'll be happy to assist as we can.* 

*The gridded data are now available for download from PFDS.*

- 4.4 I see, for the present state's data you offer from 5-minute to 60-day data, including 3-, 4-, and 7 day data, but you skip a 5-day duration. Might you consider 5-day durations, so that one could easily compare that data with HPC's readily-available 5-day QPF guidance, which has been made available to create WMS layers from GIS data (unfortunately, they don't have their own geospatial services of their data, but also offer shapefiles from which we have created WMS services). <http://www.hpc.ncep.noaa.gov/qpf/p120i12.gif> *We are currently not able to accommodate this request but may be able to add the 5-day duration in the future.*
- 4.5 Will there be a GIS tool developed that would allow the overlay of watershed boundary that would result in computation of an "average" rainfall over the basin or even an exact representation that could be imported into a rainfall-runoff model or other analysis tool? I think that could be done with the proper spatial query as long as the map data were digital and not simply an image, similar to what you can do with gridded or non-gridded DEM's. I'm not sure exactly how this would work, but it would be useful.

*We have been looking to provide something similar to what you are requesting and may be able to add such functionality in the near future.* 

4.6 It would be very beneficial if NOAA would prepare and provide a single document describing the methodology and approach. Currently, readers have to search for the information in numerous progress reports.

*We create a single document describing the data and methodology for each Volume of NOAA Atlas 14. The document is available online [\(http://www.nws.noaa.gov/oh/hdsc/currentpf.htm\)](http://www.nws.noaa.gov/oh/hdsc/currentpf.htm).* 

4.7 It would be helpful if NOAA would provide a Webinar for users of the new Atlas 14 to addressed what has changed (and if possible an explanation) and how to access and use the results. I believe that organizations who contributed financially to your efforts to complete the new Atlas 14 for this region would appreciate some sort of outreach. *We agree and are looking into setting up webinars (with the American Society of Civil Engineers) for this purpose. Please join our list server to receive any future announcements regarding this (see [http://www.nws.noaa.gov/ohd/hdsc/listserver.html\)](http://www.nws.noaa.gov/ohd/hdsc/listserver.html).* 

#### **Appendix A.5 Temporal distributions of heavy precipitation**

#### **1. Introduction**

Temporal distributions of precipitation amounts exceeding precipitation frequency estimates for the 2-year recurrence interval are provided for 6-, 12-, 24-, and 96-hour durations. The temporal distributions are expressed in probability terms as cumulative percentages of precipitation totals at various time steps. To provide detailed information on the varying temporal distributions, separate temporal distributions were also derived for four precipitation cases defined by the duration quartile in which the greatest percentage of the total precipitation occurred.

Stations were grouped into two climate regions, shown in Figure 4.1.1, and separate temporal distributions were derived for each climate region. The Mississippi Valley region (region 1) also includes stations from the Mississippi Valley region (region 4) from NOAA Atlas 14 Volume 8 (see Figure 4.1.1 in Volume 8). Regions were delineated based on extreme precipitation characteristics expressed through 24 hour mean annual maximum (MAM) estimates, mean annual precipitation, elevation and latitude.

#### **2. Methodology and results**

The methodology used to produce the temporal distributions is similar to the one developed by Huff (1967) except in the definition of precipitation cases. In accordance with the way a precipitation case ("event") was defined for the precipitation frequency analysis, a precipitation case for the temporal distribution analysis was computed as the total accumulation over a specific duration (6-, 12-, 24-, or 96 hours). As a result, it may contain parts of one or more storms. Because of that, temporal distribution curves presented here may be different from corresponding temporal distribution curves obtained from the analysis of single storms. Also, precipitation cases for this project always start with precipitation but do not necessarily end with precipitation, resulting in potentially more front-loaded cases when compared with distributions derived from the single storm approach. Cases were selected from all events of a given duration that exceeded the 2-year average recurrence interval at each station. Table A.5.1 shows the total number of precipitation cases and number of cases in each quartile for each region and duration.

For each precipitation case, cumulative precipitation amounts were converted into percentages of the total precipitation amount at one hour time increments. All cases for a specific duration were then combined and probabilities of occurrence of precipitation totals were computed at each hour. The temporal distribution curves for nine deciles (10% to 90%) were smoothed using a linear programming method (Bonta and Rao, 1988) and plotted in the same graph. Figure A.5.1 shows, as an example, temporal distribution curves computed from all cases for the four selected durations for the Mississippi Valley region (region 1); time steps were converted into percentages of durations for easier comparison.

The cases were further divided into four categories by the quartile in which the greatest percentage of the total precipitation occurred. Table A.5.1 shows the numbers and proportion of precipitation cases used to derive the temporal distributions in each quartile. Unlike the cases of 12-, 24-, and 96-hour durations in which the number of data points can be equally divided by four, the cases of 6-hour duration contain only six data points and they cannot be evenly distributed into four quartiles. Therefore, in this analysis, for the 6-hour duration, the first quartile contains precipitation cases where the most precipitation occurred in the first hour, the second quartile contains precipitation cases where the most precipitation occurred in the second and third hours, the third quartile contains precipitation cases where the most precipitation occurred in the fourth hour, and the fourth quartile contains precipitation cases where the most precipitation occurred in the fifth and sixth hours. This uneven distribution affects the number of cases contained in each quartile for the 6-hour duration. Figures A.5.2 through A.5.5 show the Mississippi Valley region's temporal

distribution curves for the four quartile cases for 6-hour, 12-hour, 24-hour and 96-hour durations, respectively.



Table A.5.1. Total number of precipitation cases and number (and percent) of cases in each quartile for selected durations for each climate region: Mississippi Valley region (1) and Southeast region (2). Region 1 in this volume includes stations from region 4 of Volume 8.

From the Precipitation Frequency Data Server, regional temporal distribution data are available in a tabular form for a selected location under the "Supplementary information" tab or through the temporal distribution web page (http://hdsc.nws.noaa.gov/hdsc/pfds/pfds\_temporal.html). For 6-, 12- and 24-hour durations, temporal distribution data are provided in 0.5-hour increments and for 96-hour duration in hourly increments.

### **3. Interpretation**

Figure A.5.1 shows as an example the temporal distribution curves of all precipitation cases in the Mississippi Valley region for the 6-, 12-, 24-, and 96-hour durations. For these plots, time steps were converted into percentages of total durations for easier comparison. Figures A.5.2 through A.5.5 show temporal distribution curves for the first-, second-, third-, and fourth-quartile cases for 6-hour, 12-hour, 24 hour and 96-hour durations, respectively. First-quartile plots show temporal distribution curves for cases where the greatest percentage of the total precipitation fell during the first quarter of the duration (e.g., the first 3 hours of a 12-hour duration). The second, third, and fourth quartile plots are similarly for cases where the most precipitation fell in the second, third, or fourth quarter of the duration.

The temporal distribution curves represent averages of many cases and illustrate the temporal distribution patterns with 10% to 90% occurrence probabilities in 10% increments. For example, the 10% curve in any figure indicates that 10% of the corresponding precipitation cases had distributions that fell above and to the left of the curve. Similarly, 10% of the cases had temporal distribution falling to the right and below the 90% curve. The 50% curve represents the median temporal distribution.

The following is an example of how to interpret the results using the figure in the upper left panel of Figure A.5.4 for 24-hour first-quartile cases in the Mississippi Valley region.

- In 10% of the first-quartile cases, 50% of the total precipitation fell in the first 2 hours and 90% of the total precipitation fell by 5.6 hours.
- A median case of this type will drop half of the precipitation (50% on the y-axis) in approximately 5.1 hours.
- In 90% of the cases, 50% of the total precipitation fell by 10.1 hours and 90% of precipitation fell by 22.1 hours.

Temporal distribution curves are provided in order to show the range of possibilities. Care should be taken in the interpretation and use of temporal distribution curves. For example, the use of different temporal distribution data in hydrologic models may result in very different peak flow estimates. Therefore, they should be selected and used in a way to reflect users' objectives.



Figure A.5.1. Temporal distribution curves for the Mississippi Valley region (region 1) all cases for: a) 6-hour, b) 12-hour, c) 24-hour, and d) 96-hour durations.



Figure A.5.2. 6-hour temporal distribution curves for the Mississippi Valley region (region 1)**:** a) first-quartile, b) second-quartile, c) third-quartile, and d) fourth-quartile cases.



Figure A.5.3. 12-hour temporal distribution curves for the Mississippi Valley region (region 1): a) first-quartile, b) second-quartile, c) third-quartile, and d) fourth-quartile cases.



Figure A.5.4. 24-hour temporal distribution curves for the Mississippi Valley region (region 1): a) first-quartile, b) second-quartile, c) third-quartile, and d) fourth-quartile cases.



Figure A.5.5. 96-hour temporal distribution curves for the Mississippi Valley region (region 1): a) first-quartile b) second-quartile, c) third-quartile, and d) fourth-quartile cases.

### **Appendix A.6 Seasonality**

#### **1. Introduction**

To portray the seasonality of extreme precipitation throughout the project area, annual maxima that exceeded precipitation frequency estimates (quantiles) with selected annual exceedance probabilities (AEPs) for chosen durations were examined for the two climate regions described in Section 4.1. Graphs showing the monthly variation of the exceedances for a region are provided for each location in the project area via the [Precipitation Frequency Data Server \(PFDS\).](http://hdsc.nws.noaa.gov/hdsc/pfds/) For a selected location, seasonal exceedance graphs can be viewed by selecting "V. Seasonality analysis" of the "Supplementary information" tab on the output page.

### **2. Method**

Separate seasonal exceedance graphs were created for the Mississippi Valley region and the Southeast region (regions 1 and 2, respectively) shown in Figure 4.1.1. Note that the Mississippi Valley region (region 1) also includes stations from the Mississippi Valley region (region 4) from NOAA Atlas 14 Volume 8 (see Figure 4.1.1 in Volume 8). They show the percentage of annual maxima for a given duration from all stations in a region that exceeded corresponding precipitation frequency estimates at selected AEP levels in each month. Results are provided for unconstrained 60-minute, 24-hour, 2 day, and 10-day durations and for AEPs of 1/2, 1/5, 1/10, 1/25, 1/50, and 1/100.

To prepare the graphs, first, the number of annual maxima exceeding the precipitation frequency estimate at a station for a given AEP was tabulated for each duration. Those numbers were then combined for all stations in a given region, sorted by month, normalized by the total number of data years in the region, and finally plotted via the PFDS.

### **3. Results**

The exceedance graphs for a selected location (see an example for a location in the Mississippi Valley region in Figure A.6.1) indicate percent of annual maxima exceeding the quantiles with selected AEPs for various durations. The percentages are based on regional statistics. On average, 1% of annual maxima for a given duration in a year (i.e., the sum of percentages of all twelve months) are expected to exceed the 1/100 AEP quantile, 4% is expected to exceed the 1/25 AEP quantile, etc.

Note that seasonality graphs are not intended to be used to derive seasonal precipitation frequency estimates.



Figure A.6.1. Example of seasonal exceedance graphs for the Mississippi Valley climate region (region 1) for the: a) 60-minute, b) 24-hour, c) 2-day, and d) 10-day durations.

## **Glossary**

*(All definitions are given relative to precipitation frequency analyses in NOAA Atlas 14 Volume 9.)*

- ANNUAL EXCEEDANCE PROBABILITY (AEP) The probability associated with exceeding a given amount in any given year once or more than once; the inverse of AEP provides a measure of the average time between years (and not events) in which a particular value is exceeded at least once; the term is associated with analysis of annual maximum series (see also AVERAGE RECCURENCE INTERVAL).
- ANNUAL MAXIMUM SERIES (AMS) Time series of the largest precipitation amounts in a continuous 12-month period (calendar or water year) for a specified duration at a given station.
- ASCII GRID Grid format with a 6-line header, which provides location and size of the grid and precedes the actual grid data. The grid is written as a series of rows, which contain one ASCII integer or floating point value per column in the grid. The first element of the grid corresponds to the upper-left corner of the grid.
- AVERAGE RECURRENCE INTERVAL (ARI; a.k.a. RETURN PERIOD, AVERAGE RETURN PERIOD) – Average time between *cases of a particular precipitation magnitude* for a specified duration and at a given location; the term is associated with the analysis of partial duration series. However, ARI is frequently calculated as the inverse of AEP for the annual maximum series; in this case it represents the average period between years in which a given precipitation magnitude is exceeded at least once.
- CONSTRAINED OBSERVATION A precipitation measurement or observation bound by clock hours and occurring in regular intervals. This observation requires conversion to an unconstrained value (see UNCONSTRAINED OBSERVATION) because maximum 60-minute or 24-hour amounts seldom fall within a single hourly or daily observation period.
- DATA YEARS See RECORD LENGTH.
- DEPTH-DURATION-FREQUENCY (DDF) CURVE Graphical depiction of precipitation frequency estimates in terms of depth, duration and frequency (ARI or AEP).
- DISTRIBUTION FUNCTION (CUMULATIVE DISTRIBUTION FUNCTION) Mathematical description that completely describes frequency distribution of a random variable, here precipitation. Distribution functions commonly used to describe precipitation data include 3 parameter distributions such as Generalized Extreme Value (GEV), Generalized Normal, Generalized Pareto, Generalized Logistic and Pearson type III, the 4-parameter Kappa distribution, and the 5-parameter Wakeby distribution.
- FEDERAL GEOGRAPHIC DATA COMMITTEE (FGDC) COMPLIANT METADATA A document that describes the content, quality, condition, and other characteristics of data and follows the guidelines set forth by the FGDC; metadata is "data about data."
- FREQUENCY General term for specifying the average recurrence interval or annual exceedance probability associated with specific precipitation magnitude for a given duration.
- FREQUENCY ANALYSIS Process of derivation of a mathematical model that represents the relationship between precipitation magnitudes and their frequencies.
- FREQUENCY ESTIMATE Precipitation magnitude associated with specific average recurrence interval or annual exceedance probability for a given duration.
- INTENSITY-DURATION-FREQUENCY (IDF) CURVE Graphical depiction of precipitation frequency estimates in terms of intensity, duration and frequency.
- INTERNAL CONSISTENCY Term used to describe the required behavior of the precipitation frequency estimates from one duration to the next or from one frequency to the next. For instance, it is required that the 100-year 3-hour precipitation frequency estimates be greater than (or at least equal to) corresponding 100-year 2-hour estimates.
- L-MOMENTS L-moments are summary statistics for probability distributions and data samples. They are analogous to ordinary moments, providing measures of location, dispersion, skewness, kurtosis, and other aspects of the shape of probability distributions or data samples, but are computed from linear combinations of the ordered data values (hence the prefix L).
- MEAN ANNUAL PRECIPITATION (MAP) The average precipitation for a year (usually calendar) based on the whole period of record or for a selected period (usually 30 year period such as 1971-2000).
- PARTIAL DURATION SERIES (PDS) Time series that includes all precipitation amounts for a specified duration at a given station above a pre-defined threshold regardless of year; it can include more than one event in any particular year.
- PRECIPITATION FREQUENCY DATA SERVER (PFDS) The on-line portal for all NOAA Atlas 14 deliverables, documentation, and information[; http://hdsc.nws.noaa.gov/hdsc/pfds/.](http://hdsc.nws.noaa.gov/hdsc/pfds/)
- PARAMETER-ELEVATION REGRESSIONS ON INDEPENDENT SLOPES MODEL (PRISM) Hybrid statistical-geographic approach to mapping climate data developed by Oregon State University's PRISM Climate Group.
- QUANTILE Generic term to indicate the precipitation frequency estimate associated with either ARI or AEP.
- RECORD LENGTH Number of years in which enough precipitation data existed to extract meaningful annual maxima in a station's period of record (or data years).
- UNCONSTRAINED OBSERVATION A precipitation measurement or observation for a defined duration. However the observation is not made at a specific repeating time, rather the duration is a moveable window through time.
- WATER YEAR Any 12-month period, usually selected to begin and end during a relatively dry season. In NOAA Atlas 14 Volume 9, it is defined as the calendar year (January 1 to December 31).

#### **References**

#### **NOAA Atlas 14 documents**

- Bonnin, G., D. Martin, B. Lin, T. Parzybok, M. Yekta, and D. Riley (2004). NOAA Atlas 14 Volume 1, Precipitation-Frequency Atlas of the United States, Semiarid Southwest. NOAA, National Weather Service, Silver Spring, MD.
- Bonnin, G., D. Martin, B. Lin, T. Parzybok, M. Yekta, and D. Riley (2006). NOAA Atlas 14 Volume 2, Precipitation-Frequency Atlas of the United States, Delaware, District of Columbia, Illinois, Indiana, Kentucky, Maryland, New Jersey, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Virginia, West Virginia. NOAA, National Weather Service, Silver Spring, MD.
- Bonnin, G., D. Martin, B. Lin, T. Parzybok, M. Yekta, and D. Riley (2006). NOAA Atlas 14 Volume 3, Precipitation-Frequency Atlas of the United States, Puerto Rico and the U.S. Virgin Islands. NOAA, National Weather Service, Silver Spring, MD.
- Perica, S., D. Martin, B. Lin, T. Parzybok, D. Riley, M. Yekta, L. Hiner, L.-C. Chen, D. Brewer, F. Yan, K. Maitaria, C. Trypaluk, G. Bonnin (2009). NOAA Atlas 14 Volume 4, Precipitation-Frequency Atlas of the United States, Hawaiian Islands. NOAA, National Weather Service, Silver Spring, MD.
- Perica, S., B. Lin, D. Martin, F. Yan, D. Brewer, C. Trypaluk, M. Yekta, L. Hiner, S. Heim, S. Dietz, T. Parzybok, L.-C. Chen, K. Maitaria, R. Chen, I. Roy, D. Unruh, T. Zhao, J. Yarchoan, G. Bonnin (2009). NOAA Atlas 14 Volume 5, Precipitation-Frequency Atlas of the United States, Selected Pacific Islands. NOAA, National Weather Service, Silver Spring, MD.
- Perica, S., S. Dietz, S. Heim, L. Hiner, K. Maitaria, D. Martin, S. Pavlovic, I. Roy, C. Trypaluk, D. Unruh, F. Yan, M. Yekta, T. Zhao, G. Bonnin, D. Brewer, L.-C. Chen, T. Parzybok, J. Yarchoan (2011). NOAA Atlas 14 Volume 6, Precipitation-Frequency Atlas of the United States, California. NOAA, National Weather Service, Silver Spring, MD.
- Perica, S., D. Kane, S. Dietz, K. Maitaria, D. Martin, S. Pavlovic, I. Roy, S. Stuefer, A. Tidwell, C. Trypaluk, D. Unruh, M. Yekta, E. Betts, G. Bonnin, S. Heim, L. Hiner, E. Lilly, J. Narayanan, F. Yan, T. Zhao (2012). NOAA Atlas 14 Volume 7, Precipitation-Frequency Atlas of the United States, Alaska. NOAA, National Weather Service, Silver Spring, MD.
- Perica, S., D. Martin, S. Pavlovic, I. Roy, M. St. Laurent, C. Trypaluk, D. Unruh, M. Yekta, G. Bonnin (2013). NOAA Atlas 14 Volume 8, Precipitation-Frequency Atlas of the United States, Midwestern States. NOAA, National Weather Service, Silver Spring, MD.
- Perica, S., D. Martin, S. Pavlovic, I. Roy, M. St. Laurent, C. Trypaluk, D. Unruh, M. Yekta, G. Bonnin (2013). NOAA Atlas 14 Volume 9, Precipitation-Frequency Atlas of the United States, Southeastern States. NOAA, National Weather Service, Silver Spring, MD.

### **Other references**

- Bonta, J. V., and A. R. Rao (1988). Fitting Equations to Families of Dimensionless Cumulative Hyetographs. Transactions of the ASAE 31(3), 756-760.
- Burn, D. H. (1990). Evaluation of Regional Flood Frequency Analysis With a Region of Influence Approach. Water Resources Research 26(10), 2257-2265.

NOAA Atlas 14 Volume 9 Version 2.0 references-1

- Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris (2002). A Knowledge-Based Approach to the Statistical Mapping of Climate. Climate Research 23, 99-113.
- Frederick, R.H., V. A. Myers, E. P. Auciello (1977). NOAA Technical Memorandum NWS Hydro-35, Five- to 60-Minute Precipitation Frequency for the Eastern and Central United States, National Weather Service, Silver Spring, MD.
- Hershfield, D.M. (1961). Technical Paper No. 40, Rainfall Frequency Atlas of the United States for durations from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years, U.S. Weather Bureau, Washington D.C.
- Hirsch, R. M., R. B. Alexander, and R. A. Smith (1991). Selection of Methods for the Detection and Estimation of Trends in Water Quality. Water Resources Research 27, 803-814.
- Hosking, J. R. M. and J. R. Wallis (1997). Regional Frequency Analysis, an Approach Based on L-Moments. Cambridge University Press.
- Huff, F. A. (1967). Time Distribution of Rainfall in Heavy Storms. Water Resources Research 3(4), 1007-1019.
- Interagency Advisory Committee on Water Data (1982). Guidelines for Determining Flood Flow Frequency. Bulletin 17B of the Hydrology Subcommittee, Office of Water Data Coordination, U.S. Geological Survey, Reston, VA.
- Langbein, W. B. (1949). Annual Floods and the Partial-Duration Flood Series. Transactions American Geophysical Union 30, 879-881.
- Laurenson, E. M. (1987). Back to Basics on Flood Frequency Analysis. Civil Engineers Transactions, Institution of Engineers, Australia, CE29, 47-53.
- Levene, H. (1960). In I. Olkin et al. eds., Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press, 278-292.
- Maidment, D. R. (1993). Handbook of Hydrology. McGraw-Hill Publishing.
- Miller, J. F. (1964). Technical Paper No. 49, Two- to Ten-Day Precipitation for Return Periods of 2 to 100 Years in the Contiguous United States. U.S. Weather Bureau and U.S. Department of Agriculture.
- Miller, J. F., R. H. Frederick and R. J. Tracy (1973). NOAA Atlas 2 Volume 3, Precipitation-Frequency Atlas of the Western United States, Colorado. National Weather Service, Silver Spring, MD.
- Shao, J. and D. Tu (1995). The Jackknife and Bootstrap. Springer-Verlag, Inc.
- U.S. Weather Bureau (1960). Technical Paper No. 29, Rainfall Intensity-Frequency Regime. U.S. Dept. of Commerce, Washington, D.C.