NOAA Atlas 14

Precipitation-Frequency Atlas of the United States

Volume 3 Version 4.0:
Puerto Rico and the U.S. Virgin Islands

Geoffrey M. Bonnin, Deborah Martin, Bingzhang Lin, Tye Parzybok, Michael Yekta, David Riley

U.S. Department of Commerce

National Oceanic and Atmospheric Administration

National Weather Service

Silver Spring, Maryland, 2006

NOAA Atlas 14

Precipitation-Frequency Atlas of the United States

Volume 3 Version 4.0:
Puerto Rico and the U.S. Virgin Islands

Geoffrey M. Bonnin, Deborah Martin, Bingzhang Lin, Tye Parzybok, Michael Yekta, David Riley

U.S. Department of Commerce

National Oceanic and Atmospheric Administration

National Weather Service

Silver Spring, Maryland, 2006

Library of Congress Classification Number
GC
1046
.C8
U6
no.14
v.3
(2006)

Table of Contents

1. Abstract

NOAA Atlas 14 contains precipitation frequency estimates with associated confidence limits for the United States and is accompanied by additional information such as temporal distributions and seasonality. The Atlas is divided into volumes based on geographic sections of the country. The Atlas is intended as the official documentation of precipitation frequency estimates and associated information for the United States. It includes discussion of the development methodology and intermediate results. The Precipitation Frequency Data Server (PFDS) was developed and published in tandem with this Atlas to allow delivery of the results and supporting information in multiple forms via the Internet.

2. Preface to Volume 3

NOAA Atlas 14 Volume 3 contains precipitation frequency estimates for Puerto Rico and the U.S. Virgin Islands. The Atlas supercedes precipitation frequency estimates contained in Technical Paper No. 42 "Generalized estimates of probable maximum precipitation and rainfall-frequency data for Puerto Rico and Virgin Islands" (U.S. Weather Bureau, 1961) and Technical Paper No. 53 "Two- to ten-day rainfall for return periods of 2 to 100 years in Puerto Rico and Virgin Islands" (Miller, 1965). The updates are based on more recent and extended data sets, currently accepted statistical approaches, and improved spatial interpolation and mapping techniques.

The work was performed by the Hydrometeorological Design Studies Center within the Office of Hydrologic Development of the National Oceanic and Atmospheric Administration's National Weather Service. Funding for the work was provided by NOAA National Weather Service, the U.S. Army Corps of Engineers, and the Commonwealth of Puerto Rico Department of Natural and Environmental Resources. 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: "Precipitation-Frequency Atlas of the United States" NOAA Atlas 14, Volume 3, Version 2.0, G. M. Bonnin, D. Martin, B. Lin, T. Parzybok, M. Yekta, and D. Riley, NOAA, National Weather Service, Silver Spring, Maryland, 2006.

The version number has the format P.S where:

P is an integer representing successive releases of primary information. Primary information is essentially the data – the values of precipitation frequencies (in ASCII grids of the precipitation frequency estimates and output from the PFDS), shapefiles, cartographic maps, temporal distributions, and seasonality.

S is an integer representing successive releases of secondary information. S reverts to zero (or nothing; i.e., Version 2 and Version 2.0 are equivalent) when P is incremented. Secondary information includes documentation and metadata.

When new information is completed and added, such as draft documentation, *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, the grids). An examination of any of the artifacts available through the Precipitation Frequency Data Server (PFDS) provides an immediate indication of the primary version number associated with all artifacts. All output from the

PFDS is stamped with the version number and date of download.

Several versions of the project have been released. Table 2.1 lists the version history associated with NOAA Atlas 14 Volume 3, the Puerto Rico and U.S. Virgin Islands precipitation frequency project and indicates the nature of changes made. If major discrepancies are observed or identified by users, a new release may be warranted.

Table 2.1. Version History of the NOAA Atlas 14 Volume 3.

Version no.	**Date**	**Notes**
Version 1	November 4, 2005	Draft data used in peer review
Version 2	June 28, 2006	Final released data
Version 3.0	October 26, 2006	Updated final data and final documentation released
Version 4.0	March 21, 2008	Updated final confidence limits for short durations released

3. Introduction

3.1. Objective

NOAA Atlas 14 Volume 3 provides precipitation frequency estimates for Puerto Rico and the U.S. Virgin Islands. Figures 4.1.1 and 4.1.2 show the project area where estimates are available and also include all stations used in the analysis. This Atlas provides precipitation frequency estimates for 5-minute through 60-day durations at average recurrence intervals of 1-year through 1,000-year. The estimates are based on the analysis of annual maximum series and then converted to partial duration series results. The information in NOAA Atlas 14 Volume 3 supercedes precipitation frequency estimates contained in Technical Paper No. 42 "Generalized estimates of probable maximum precipitation and rainfall-frequency data for Puerto Rico and Virgin Islands" (U.S. Weather Bureau, 1961) and Technical Paper No. 53 "Two- to ten-day rainfall for return periods of 2 to 100 years in Puerto Rico and Virgin Islands" (Miller, 1965). The results are provided at high spatial resolution and include confidence limits for the estimates. The Atlas includes temporal distributions designed for use with the precipitation frequency estimates (Appendix A.1) and seasonal information for heavy precipitation (Appendix A.2). In addition, the potential effects of climate change were examined (Appendix A.3).

The new estimates are based on improvements in three primary areas: denser data networks with a greater period of record, the application of regional frequency analysis using L-moments for selecting and parameterizing probability distributions and new techniques for spatial interpolation and mapping. The new techniques for spatial interpolation and mapping account for topography and have allowed significant improvements in areas of complex terrain.

NOAA Atlas 14 Volume 3 precipitation frequency estimates for Puerto Rico and the U.S. Virgin Islands are available via the Precipitation Frequency Data Server at http://hdsc.nws.noaa.gov/hdsc/pfds which provides the additional ability to download digital files. The types of results and information found there include:

- point estimates (via a point-and-click interface)
- ArcInfo© ASCII grids
- ESRI shapefiles
- color cartographic maps
- associated Federal Geographic Data Committee-compliant metadata
- data series used in the analyses: annual maximum series and partial duration series
- temporal distributions of heavy precipitation (1-hour, 6-hour, 12-hour, 24-hour and 96-hour)
- seasonal exceedance graphs: counts of events that exceed the 1 in 2, 5, 10, 25, 50 and 100 annual exceedance probabilities for the 60-minute, 24-hour, 48-hour, and 10-day durations.

As discussed in Sections 4.8.4 and 4.8.5, the color cartographic maps and ESRI shapefiles were created to serve as visual aids and, unlike Technical Papers 42 and 53, are not recommended for interpolating final point or area precipitation frequency estimates. Users are urged to take advantage of the Precipitation Frequency Data Server or the underlying ArcInfo© ASCII grids for obtaining estimates.

3.2. Terminology; Partial Duration and Annual Maximum Series

This publication adopts the terminology “average recurrence interval” (ARI) and “annual exceedance probability” (AEP) presented in Australian Rainfall and Runoff (Institute of Engineers, Australia, 1987) which in turn is based on Laurenson (1987). NOAA Atlas 14 is based on the analysis of annual maximum series data with the results converted to represent estimates based on partial duration series. The results for these two types of series differ at shorter average recurrence intervals and have different meanings. Factors for converting between these results are provided in Section 4.6.3.

An annual maximum series is constructed by taking the highest accumulated precipitation for a

particular duration in each successive year of record, whether the year is defined as a calendar year or using some other arbitrary boundary such as a water year. Calendar years are used in this Atlas. An annual maximum series inherently excludes other extreme cases that occur in the same year as a more extreme case. In other words, the second highest case on record at an observing station may occur in the same year as the highest case on record but will not be included in the annual maximum series. A partial duration series is constructed by taking all of the highest cases above a threshold regardless of the year in which the case occurred. In this Atlas, partial duration series consist of the N largest cases in the period of record, where N is the number of years in the period of record at the particular observing station.

Analysis of annual maximum series produces estimates of the average period between *years when a particular value is exceeded*. On the other hand, analysis of partial duration series gives the average period between *cases of a particular magnitude*. The two results are numerically similar at rarer average recurrence intervals but differ at shorter average recurrence intervals (below about 20 years). The difference can be important depending on the application.

Typically, the use of AEP and ARI reflects the analysis of the different series. However, in some cases, average recurrence interval is used as a general term for ease of reference.

3.3. Approach

The approach used in this project largely follows the regional frequency analysis using the method of L-moments described in Hosking and Wallis (1997). This section provides an overview of the approach. Greater detail on the approach is provided in Section 4.2.

NOAA Atlas 14 introduces a change from past NWS publications by its use of regional frequency analysis using L-moments for selecting and parameterizing probability distributions. Both annual maximum series and partial duration series were extracted at each observing station from quality controlled data sets. Because of the greater reliability of the analysis of annual maximum series, an average ratio of partial duration series to annual maximum series precipitation frequency estimates (quantiles) was computed and then applied to the annual maximum series quantiles to obtain the final equivalent partial duration series quantiles.

Quality control was performed on the initial observed data sets (see Section 4.3) and it continued throughout the process as an inherent result of the performance parameters of intermediate steps.

To support the regional approach, potential regions were initially determined based on climatology. They were then tested statistically for homogeneity. Individual stations in each region were also tested statistically for discordancy. Adjustments were made in the definition of regions based on underlying climatology in cases where homogeneity and discordancy criteria were not met.

A variety of probability distributions were examined and the most appropriate distribution for each region and duration was selected using several different performance measures. The final determination of the appropriate distributions for each region and duration was made based on sensitivity tests and a desire for a relatively smooth transition between distributions from region to region. Probability distributions selected for annual maximum series were not necessarily the same as those selected for partial duration series.

Quantiles at each station were determined based on the mean of the data series at the station and the regionally determined higher order moments of the selected probability distribution. Quantiles for durations below 60-minutes (n-minute durations) were computed using an average ratio between the n-minute and 60-minute quantiles due to the small number of stations recording data at less than 60-minute intervals.

For the first time, the National Weather Service is providing confidence limits for the precipitation frequency estimates in the area covered by NOAA Atlas 14. Monte Carlo Simulation was used to produce upper and lower bounds at the 90% confidence level.

In the regional approach, the second and higher order moments are constant for each region

resulting in a potential for discontinuities in the quantiles at regional boundaries. In order to avoid potential discontinuities and to achieve an effective spatial interpolation of quantiles between observing stations, the data series means at each station for each duration were spatially interpolated using PRISM technology by the PRISM Group, formerly Spatial Climate Analysis Service (SCAS), at Oregon State University (Appendix A.4). Because the mean was derived directly at each observing station from the data series and independently of the regional computations, it was not subject to the same discontinuities. The grid of quantiles for each successive average recurrence interval was then derived in an iterative process using a strong linear relationship between a particular duration and average recurrence interval and the next rarer average recurrence interval of the same duration (see Section 4.8.2). The resulting set of grids were tested and adjusted in cases where inconsistencies occurred between durations and frequencies.

Both the spatial interpolation and the point estimates were subject to external peer reviews (see Section 6 and Appendix A.5). Based on the results of the peer review, adjustments were made where necessary by the addition of new observations or removal of questionable ones. Adjustments were also made in the definition of regions.

Temporal precipitation patterns were extracted for use with the precipitation frequency estimates presented in the Atlas (Appendix A.1). The temporal patterns are presented in probabilistic terms and can be used in Monte Carlo development of ensembles of possible scenarios. They were specifically designed to be consistent with the definition of duration used for the precipitation frequency estimates.

The seasonality of heavy precipitation is represented in seasonal exceedance graphs that are available through the Precipitation Frequency Data Server. The graphs were developed for each region by tabulating the number of events exceeding the precipitation frequency estimate at each station for a given annual exceedance probability (Appendix A.2).

The 1-day annual maximum series were analyzed for linear trends in mean and variance and shifts in mean to determine whether climate change during the period of record was an issue in the production of this Atlas (Appendix A.3). The results showed little observable or geographically consistent impact of climate change on the annual maximum series during the period of record and so the entire period of record was used. The estimates presented in this Atlas make the necessary assumption that there is no effect of climate change in future years on precipitation frequency estimates. The estimates will need to be modified if that assumption proves quantifiably incorrect.

4. Method

4.1. Data

4.1.1. Properties

Sources. Daily, hourly, and n-minute (defined below) measurements of precipitation from various sources were used for this project (Table 4.1.1). Figure 4.1.1 shows the locations of daily stations in the project area. Figure 4.1.2 shows the hourly, n-minute and 15-minute stations. The 15-minute stations were primarily co-located with the hourly stations. They were used for the temporal distribution analysis and the n-minute ratio development.

The National Weather Service (NWS) Cooperative Observer Program's (COOP) daily and hourly stations were the primary source of precipitation gauge records. The following data sets of COOP data were obtained from National Oceanic and Atmospheric Administration's (NOAA) National Climatic Data Center (NCDC):

- Hourly data set: TD3240
- Daily data set: TD3200 and TD3206
- N-minute data set: TD9649, an additional NCDC dataset covering 1973-1979 and Automated Surface Observing System (ASOS) maintained by the Federal Aviation Administration
- 15-minute data set: TD3280

Additional 15-minute data were provided by the United States Geological Survey (USGS) but the period of record was too short to meet the needs of the project.

Table 4.1.1. Number of stations in the project area.

Location	Daily	Hourly	N-min	15-min
Puerto Rico	106	23	1	22
U.S. Virgin Islands	19	2	0	3
Total	125	25	1	25

Figure 4.1.1. Map of daily stations for NOAA Atlas 14 Volume 3.

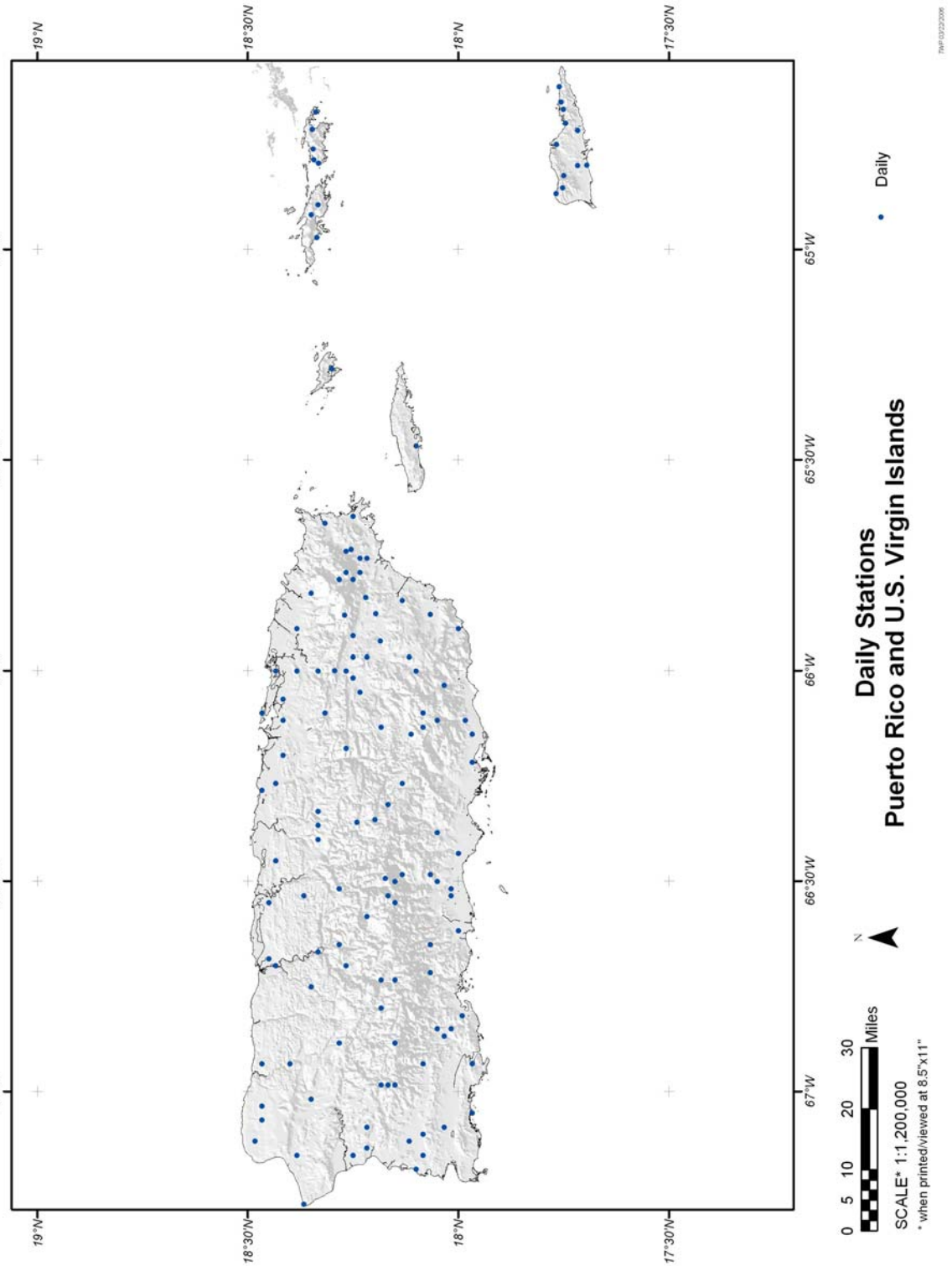
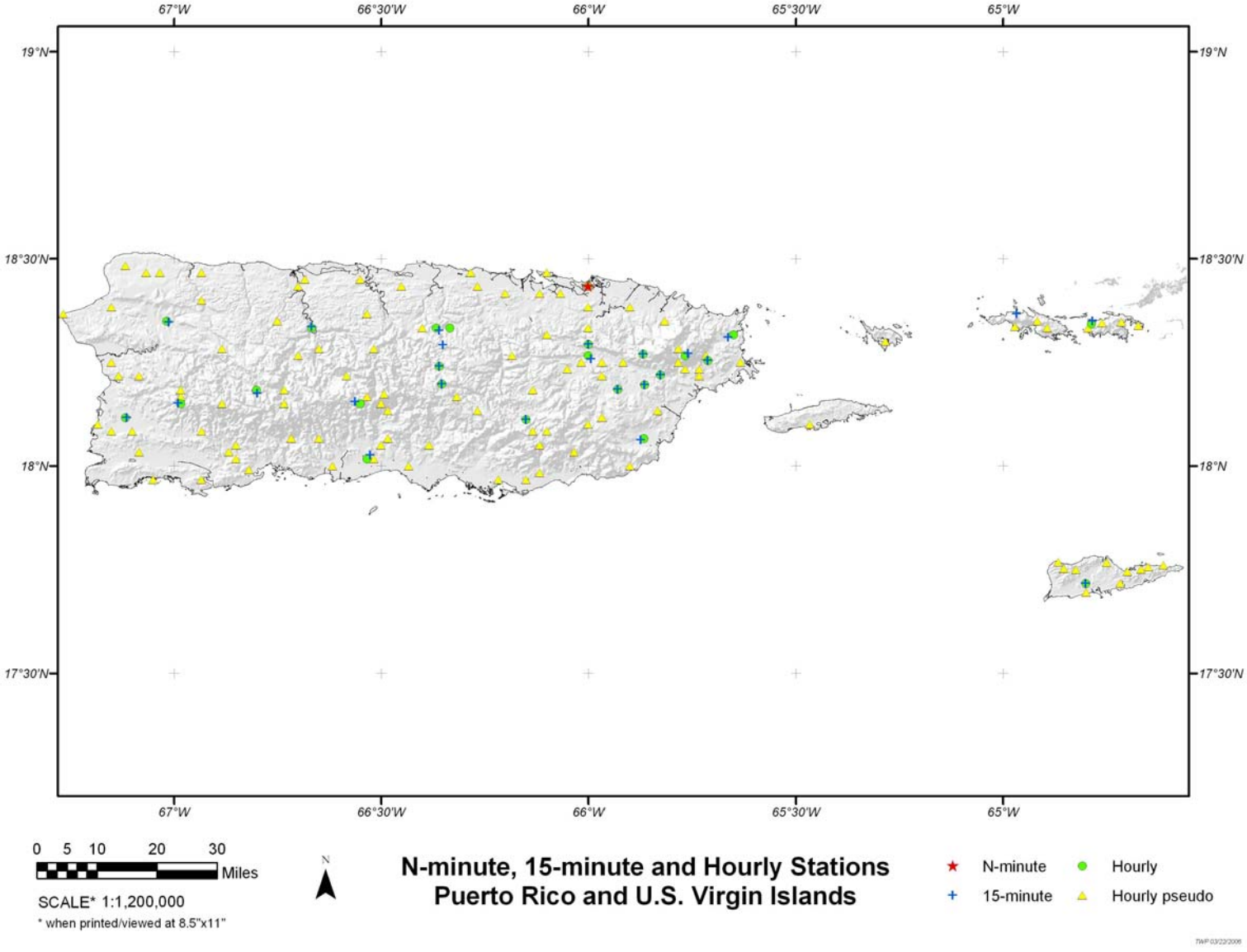


Figure 4.1.2. Map of hourly and n-minute stations for NOAA Atlas 14 Volume 3.



Record length. Record length may be characterized by the entire period of record or by the number of years of useable data within the total period of record (data years). For this project, only daily stations with 20 or more data years and hourly stations with 15 or more data years were used in the analysis. The records of the daily stations extend through December 2004 and average 54 data years in length (Table 4.1.2). The records of the hourly stations extend through December 2003 and average 29 data years. Although 15 data years was used as the minimum for the hourly stations, the shortest station used had 17 data years. Figures 4.1.3 and 4.1.4 show the number of data years by percent of stations for the daily and hourly data. The single n-minute record available for the analysis had 27 years of data extending through December 2004. The 15-minute records extended through December 2003 and average 28 data years in length with a distribution similar to the hourly stations depicted in Figure 4.1.4. (See Appendix A.6 for a complete list of stations or http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_data.html for downloadable comma-delimited station lists.)

Table 4.1.2. Information for daily data set through 12/2004, hourly dataset through 12/2003, n-minute dataset through 12/2004, and 15-minute data set through 12/2003.

	Daily	**Hourly**	**N-minute**	**15-minute**
No. of stations	126	25	1	25
Longest record length (data yrs) (Station ID)	104 (66-2801)	37 (66-8812)	27 (66-8812)	32 (66-2934)
Average record length (data yrs)	54	29	27	28

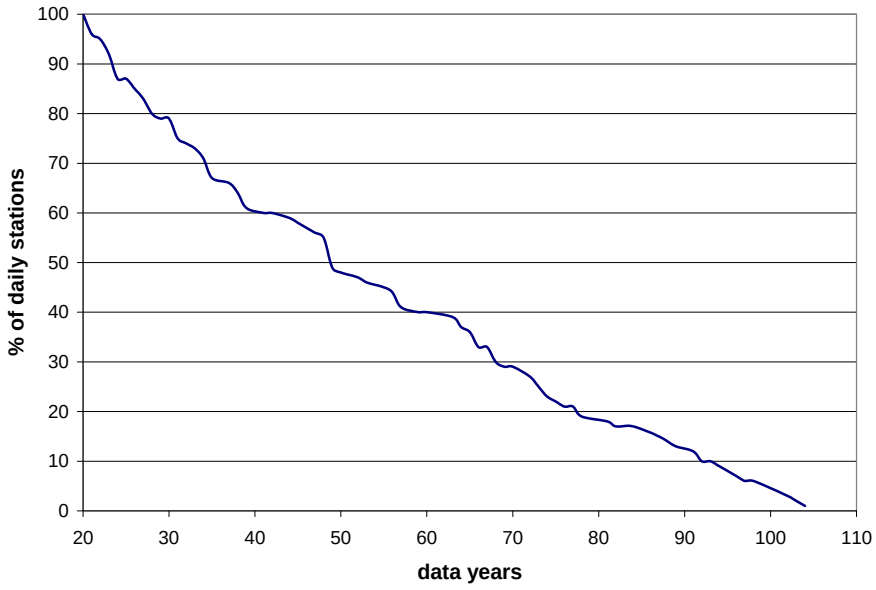


Figure 4.1.3. Plot of percentage of total number of daily stations used in NOAA Atlas 14 Volume 3 versus data years.

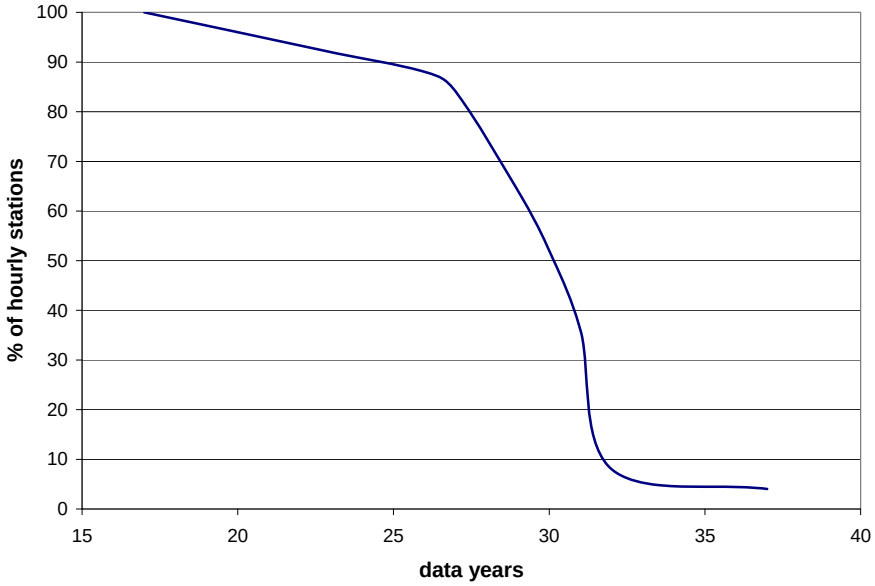


Figure 4.1.4. Plot of percentage of hourly stations used in NOAA Atlas 14 Volume 3 versus data years.

N-minute data. N-minute data are precipitation data measured at a temporal resolution of 5-minutes that can be summed to various "n-minute" durations (10-minute, 15-minute, 30-minute, and 60-minute). Because of the small number of n-minute data available, n-minute precipitation frequencies were estimated by applying a linear scaling to 60-minute data. The linear scaling factors were developed using ratios of n-minute quantiles to 60-minute quantiles from the single co-located n-minute and hourly station and from the twenty-five 15-minute stations with at least 15 years of data that were co-located with hourly stations. Figure 4.1.2 shows the locations of these stations. Because there were so few stations, the stations were analyzed as one large region. The ratios were calculated from quantiles computed using the Generalized Normal (GNO) distribution because it was best-fitting for the 60-minute data and the majority of n-minute durations and because there was little difference in the scaling factors obtained using GNO, Generalized Logistic (GLO) or Generalized Extreme Value (GEV).

Since the analysis using a single n-minute station was less reliable and less representative of the project area as a whole, the 15-minute stations were utilized to obtain a full set of n-minute over 60-minute ratios (5-, 10-, 15-, and 30-minute) through a combined approach. The 15-minute and 30-minute ratios were computed from the quantiles of the 15-minute and n-minute stations. 5-minute and 10-minute ratios were computed from the quantiles of the single n-minute station but used only to establish a linear trend from which to extrapolate the 5-minute and 10-minute ratios. The ratios were averaged over all return frequencies for each duration since there were so few data (26 stations with only 16 to 33 years of data). Table 4.1.3 shows the ratios used for this Volume, and includes those used in previous NOAA Atlas 14 Volumes (Bonnin et al., 2003 and 2004). The ratios developed for this Volume compare favorably with those developed for Volume 2 (south region) which has some climatological similarity.

Technical Paper 42 (U.S. Weather Bureau, 1961), which covered Puerto Rico and the U.S. Virgin Islands, provided for only the 30-minute duration using a ratio of 0.79 for all calculated return frequencies. The NOAA Atlas 14 ratios may be used with more confidence since they are based on considerably more data from within the project area, whereas the Technical Paper 42 ratio was

developed for other studies such as NOAA Atlas 2 (Miller et al., 1973) for the western United States and Technical Paper 40 (Hershfield, 1961) for the entire continental United States.

Table 4.1.3. N-minute ratios (5-, 10-, 15- and 30-minute to 60-minute) for NOAA Atlas 14 Volume 1 Semiarid Southwest U.S., Volume 2 Ohio River basin and surrounding states (ranges for northern and southern region) and Volume 3 Puerto Rico and U.S. Virgin Islands.

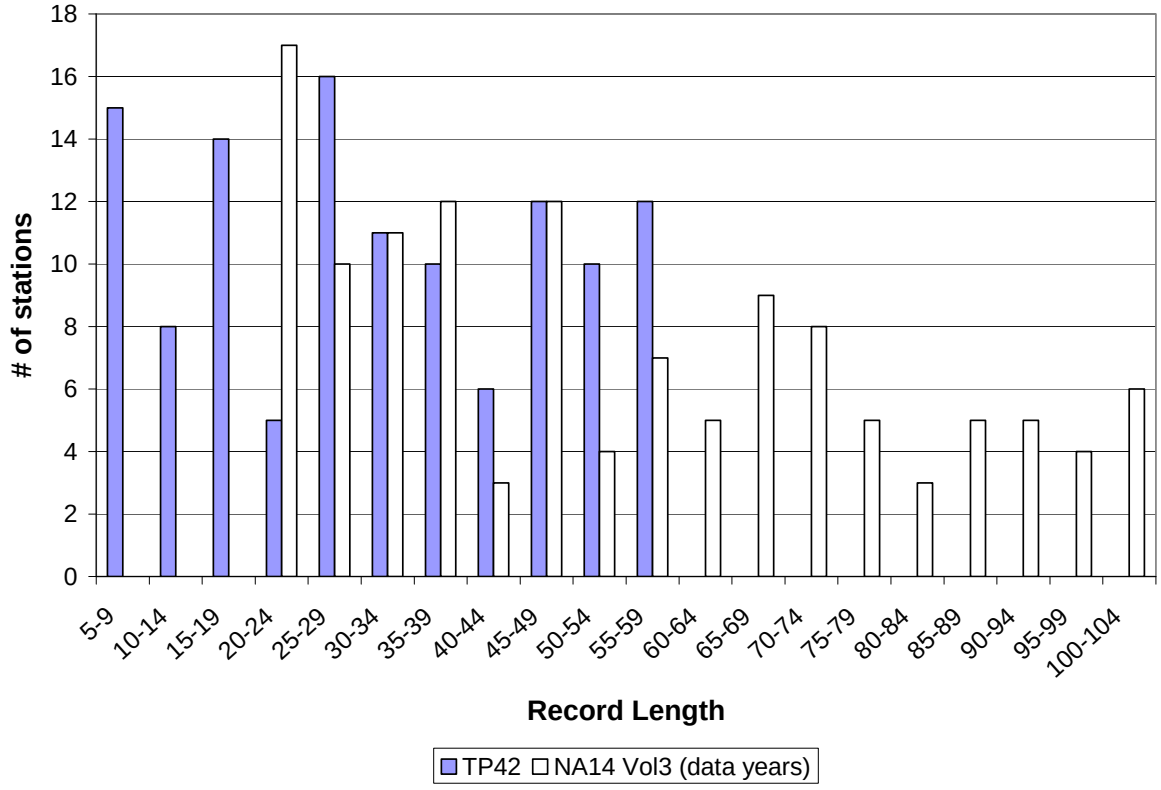
	5-min	**10-min**	**15-min**	**30-min**
NOAA Atlas 14 Volume 1	0.318	0.484	0.600	0.808
NOAA Atlas 14 Volume 2 (north region)	0.261-0.325	0.380-0.505	0.475-0.619	0.712-0.819
NOAA Atlas 14 Volume 2 (south region)	0.214-0.293	0.337-0.468	0.423-0.585	0.685-0.802
NOAA Atlas 14 Volume 3	**0.240**	**0.328**	**0.421**	**0.674**

Multi-day/hour durations. Maxima for durations greater than 24-hour were generated by accumulating daily data. The multi-day maxima, 2-day through 60-day, were extracted in an iterative process where 1-day observations were summed and compared with the value of the previous summation shifted by 1 day. Multi-hour durations, 2-hour through 48-hour, were generated by accumulating hourly data. (See Section 4.1.3 for additional details on the annual maximum series and partial duration series extraction process.)

Technical Paper 42 data comparison. Technical Paper 42 (U.S. Weather Bureau, 1961), herein after referred to simply as Technical Paper 42, which covered Puerto Rico and U.S. Virgin Islands was the most recent update of the precipitation frequencies for durations 30-minutes through 24-hours. Unlike NOAA Atlas 14, Technical Paper 42 utilized stations differently depending on their record length. Stations with longer records (greater than 19 years) were used to establish relationships between estimates for the rarer average recurrence intervals and the 2-year average recurrence interval. Stations with short record lengths were used to establish spatial patterns for the 2-year estimates only. However, in NOAA Atlas 14, all stations meeting the minimum requirements for number of years of data were used for all durations and recurrence intervals and the regional L-moment technique itself accounts for variations in record length.

A total of 102 daily stations in Puerto Rico and 17 daily stations in the U.S. Virgin Islands were used in Technical Paper 42. Since there were only three hourly stations, with two having records less than 8 years, Technical Paper 42 based 1-hour statistics on stations in the United States. NOAA Atlas 14 Volume 3 used a total of 126 daily stations and 25 hourly stations in the project area with considerably longer periods of record than were available to Technical Paper 42. The daily record length in Technical Paper 42 was 5 to 59 years, whereas 53 years was the average for NOAA Atlas 14 Volume 3 ranging from 20 to 104 years. Some stations available for NOAA Atlas 14 Volume 3 had up to 45 more years of record than those used in Technical Paper 42. This allowed for the exclusion of shorter, less reliable data records for all analyses in NOAA Atlas 14 Volume 3. Figure 4.1.5 shows the number of years of record for daily stations used in each Atlas.

Figure 4.1.5. Comparison of the years of record at stations used in Technical Paper 42 (TP42) and NOAA Atlas 14 Volume 3 (NA14 Vol3).



4.1.2. Conversions of data

Daily. Daily data have varying observation times. Maximum 24-hour amounts seldom fall within a single daily observation period. In order to make the daily and hourly data comparable, a conversion was necessary from 'observation day' (constrained observation) to 24 hours (unconstrained observation). NOAA Atlas 2 (Miller et al., 1973), Technical Paper 40 (Hershfield, 1961) and Technical Paper 42 used the empirically derived value of 1.13 to convert daily data to 24-hour data. The conversion factor for this project was computed using ratios of the 2-year quantiles computed from monthly maxima series at 18 co-located daily and hourly stations with at least 15 years of concurrent hourly and daily data in the project area. Monthly maxima time series for concurrent time periods were generated for 24-hour precipitation values summed from hourly observations and co-located daily precipitation observations. A total of 3,217 pairs of monthly maxima were extracted. Ten pairs where the daily observations were inconsistent with hourly accumulations were omitted from the analysis (Section 4.3).

The monthly maxima time series of 1-day and 24-hour data were analyzed separately using L-moments. Ratios of 2-year 24-hour to 2-year 1-day quantiles were then generated and averaged. The conversion factor, 1.21, was the same using different distributions (GNO, GEV, GLO). Similarly, a conversion factor, 1.13, for 2-day to 48-hours was calculated. These conversion factors were comparable to factors computed using the L-moment results of annual maxima series at these stations.

The conversion factors for this project were higher than the factors used in NOAA Atlas 14 Volumes 1 and 2 (see Table 4.1.7) and in Technical Paper 42 (U.S. Weather Bureau, 1961). However, even though the factors were higher, the ratio of the 1-day to 24-hour factor over the 2-day

to 48-hour factor (1.21/1.13 = 1.07) was consistent with that ratio for NOAA Atlas 14 Volume 1 (Semiarid Southwest), 1.11, and Volume 2 (Ohio River Basin), 1.09.

All daily and 2-day data used in the analyses were converted to equivalent 24-hour and 48-hour unconstrained values, respectively.

Hourly. In order to make hourly and 60-minute data comparable, a conversion was necessary from the constrained 'clock hour' to unconstrained 60-minute and from 2 hours to 120-minute. Conversion factors were computed using ratios of the 2-year quantiles computed from annual maxima series and monthly maxima series at one first order station with co-located hourly and n-minute stations (66-8812 San Juan) and twenty-three co-located hourly and 15-minute stations with at least 15 years of concurrent data (note: twenty-five 15-minute stations with at least 13 years of data were used in the monthly analysis). Time series from concurrent time periods were generated for 60-minute precipitation values summed from n-minute (or 15-minute) observations and for co-located hourly precipitation observations. The series were analyzed separately using L-moments and four distributions (GEV, GLO, GNO, and GPA). Ratios of 2-year 60-minute to 2-year 1-hour quantiles were generated and averaged. The conversion factor was the same regardless of distribution. It was further verified by a regression analysis of concurrent annual maxima pairs and monthly maxima pairs. There were 694 concurrent annual maxima data pairs from twenty-four co-located stations and 7,836 concurrent monthly maxima pairs from twenty-six co-located stations.

The resulting conversion factors were 1.13 for 1-hour to 60-minute and 1.04 for 2-hour to 120-minute. The 1-hour to 60-minute factor is in close agreement with Technical Paper 42 and with NOAA Atlas 14 Volumes 1 and 2 (see Table 4.1.4). No conversion was provided for 2-hour to 120-minutes in those studies except for NOAA Atlas 14 Volumes 1 and 2.

Table 4.1.4. Conversion factors for constrained to unconstrained observations.

	Conversion Factors			
Project	**1-day to 24-hour**	**2-day to 48-hour**	**1-hour to 60-minute**	**2-hour to 60-minute**
NOAA Atlas 14 Vol. 1 (Semiarid Southwestern United States)	1.14	1.03	1.12	1.03
NOAA Atlas 14 Vol. 2 (Ohio River Basin and Surrounding States)	1.13	1.04	1.16	1.05
NOAA Atlas 14 Vol. 3 (Puerto Rico and U.S. Virgin Islands)	**1.21**	**1.13**	**1.13**	**1.04**
Technical Paper 42	1.13	N/A	1.13	N/A

4.1.3. Extraction of series

Two methods were used for extracting series of data at a station for the analysis of precipitation frequency: **Annual Maximum Series (AMS)** and **Partial Duration Series (PDS)**.

The AMS method selected the largest single case that occurred in each calendar year of record. If a large case was not the largest in a particular year, it was not included in the series.

The PDS method recognized that more than one large case may occur during a single calendar year. For this Atlas, the largest N cases in the entire period of record, where N was the number of years of data, were selected to create the partial duration series. More than one case could be selected from any particular year and a large case that was not the largest in a particular year could appear in the series. Such a series is also called an annual exceedance series (AES) (Chow et al., 1988).

Differences in the meaning of the results of analysis using these two different types of series are discussed in Section 3.2. Average empirical conversion factors were developed to provide PDS-based results from the AMS-based results (see Section 4.6.4). The data series used in the analysis

(and associated documentation) are provided through the Precipitation Frequency Data Server which can be found at http://hdsc.nws.noaa.gov/hdsc.

The procedure for extracting maxima from the dataset used specific criteria. The criteria, described below, ensured that each year had a sufficient number of data, particularly in the assigned "wet season", to accurately extract statistically meaningful values. The "wet season" for each location was defined as the months in which extreme cases were mostly likely to occur and was assigned by assessing histograms of annual maximum precipitation for each homogeneous region (Tables 4.1.5 and 4.1.6). [The development and verification of the homogeneous regions are discussed in Section 4.4 and shown in Figures 4.4.1 and 4.4.2.]

Table 4.1.5. "Wet season" months for daily regions of NOAA Atlas 14 Volume 3.

Region	start month	end month
Daily Regions		
1	5	11
2	4	12
3	4	11
4	4	11
5	5	11
6	4	12
7	4	11
8	4	11
9	5	11

Table 4.1.6. "Wet season" months for hourly regions of NOAA Atlas 14 Volume 3.

Region	start month	end month
Hourly Regions		
1	4	11
2	4	11
3	4	11
4	1	12

Criteria for hourly annual maximum series. For all hourly durations (1-hour through 48-hours), the highest value in each year was extracted as the annual maximum for that particular year. Cases that spanned January 1st were assigned to the date on which the greatest hourly precipitation occurred during the corresponding duration.

A month was invalid and the maximum precipitation for that month was set to missing:

- if the hours of available data in a month were less than the duration hours
- if 240 hours or more in a month were missing and the maximum precipitation for the month <= 0.01 inches
- if 360 or more hours in a month were missing and the maximum precipitation for the month was less than 33% of the average precipitation for that month at that station
- if 50% or more hours (for a specific duration) were missing

Also, if more than 50% of the months in the wet season for a given region were missing, then the maximum precipitation for the year was set to missing.

Criteria for daily annual maximum series. An annual maximum was extracted for daily durations (1-day through 60-day), if at least 50% of the months in the assigned wet season and at least 50% of the data for the accumulated period were present. The highest value in each year was extracted as the annual maximum for that particular year. Cases that spanned January 1st were assigned to the date on which the greatest daily precipitation occurred during the corresponding duration.

In addition, the following criteria applied:

1-day:
If all the days in the month were missing, or if more than 10 days of the month were missing and the maximum precipitation for the month was 0.00", or if more than 15 days were missing and the maximum for the month was less than 30% of the average 1-day maximum precipitation for that month over the period of record at that station, then that month was set to missing.

2-day:
If there was only 1 day of data for the month and the rest of the days were missing, or if more than 10 days of the month were missing and the maximum precipitation for the month was 0.00", or if more than 15 days were missing and the maximum for the month was less than 30% of the average 2-day maximum precipitation for that month over the period of record at that station, then that month was set to missing.

4-day:
If more than 96% of the days in a given year were missing, or if 50% of the days of the year were missing and the maximum precipitation for the year was 0.3" or less, then that year was set to missing.

7-day:
If more than 93% of the days in a given year were missing, or if 50% of the days of the year were missing and the maximum precipitation for the year was 0.3" or less, then that year was set to missing.

10-day:
If more than 93% of the days in a given year were missing, or if 50% of the days of the year were missing and the maximum precipitation for the year was 0.35" or less, then that year was set to missing.

20-day:
If more than 88% of the days in a given year were missing, or if 50% of the days of the year were missing and the maximum precipitation for the year was 0.35" or less, then that year was set to missing.

30-day:
If more than 82% of the days in a given year were missing, or if 50% of the days of the year were missing and the maximum precipitation for the year was 0.45" or less, then that year was set to missing.

45-day:
If more than 73% of the days in a given year were missing, or if 50% of the days of the year were missing and the maximum precipitation for the year was 0.45" or less, then that year was set to missing.

60-day:
If more than 64% of the days in a given year were missing, or if 50% of the days of the year were missing and the maximum precipitation for the year was 0.45" or less, then that year was set to missing.

Criteria for partial duration series. The criteria listed above also apply for deciding whether a month or year has enough data to be included in the extraction process for a partial duration series. Cases that spanned January 1st were assigned to the date on which the greatest precipitation observation occurred during the corresponding duration.

Precipitation accumulations for each hourly duration were extracted and then sorted in descending order. The highest N accumulations for each duration were retained, where N was the number of actual data years for each station.

Precipitation *events* in the daily dataset were defined as being separated by at least one dry day to ensure independence. For 1-day and 2-day durations, one precipitation accumulation for each duration was extracted per *event*. For 4-day and longer durations, an accumulated period could include all of or parts of one or more event (and therefore may include dry days). However, each event was used only once (i.e., the same days were not used twice in partial duration maximums for a given duration). The accumulations were then sorted in descending order and the highest N accumulations for each duration were retained, where N was the number of actual data years for each station.

4.2. Regional approach based on L-moments

4.2.1. Overview

Hosking and Wallis (1997) describe regional frequency analysis using the method of L-moments. This approach, which stems from work in the early 1970s but which only began seeing full implementation in the 1990s, is now accepted as the state of the practice. The National Weather Service has used Hosking and Wallis, 1997, as its primary reference for the statistical method for this Atlas.

The method of L-moments (or linear combinations of probability weighted moments) provides great utility in choosing the most appropriate probability distribution to describe the precipitation frequency estimates. The method provides tools for estimating the shape of the distribution and the uncertainty associated with the estimates, as well as tools for assessing whether the data are likely to belong to a homogeneous region (e.g., climatic regime).

The regional approach employs data from many stations in a region to estimate frequency distribution curves for the underlying population at each station. The approach assumes that the frequency distributions of the data from many stations in a homogeneous region are identical apart from a site-specific scaling factor. This assumption allows estimation of shape parameters from the combination of data from all stations in a homogeneous region rather than from each station individually, vastly increasing the amount of information used to produce the estimate, and thereby increasing the accuracy. Weighted averages that are proportional to the number of data years at each station in the region are used in the analysis.

The regional frequency analysis using the method of L-moments assists in selecting the appropriate probability distribution and the shape of the distribution, but precipitation frequency estimates (quantiles) are estimated uniquely at each individual station by using a scaling factor, which, in this project, is the mean of the annual maximum series at each station. The resulting quantiles are more reliable than estimates obtained based on single at-site analyses (Hosking and Wallis, 1997).

4.2.2. L-moment description

Regional frequency analysis using the method of L-moments provided tools to test the quality of the dataset, test the assumptions of regional homogeneity, select a frequency distribution, estimate precipitation frequencies, and estimate confidence limits for this Atlas. Details and equations for the analysis may be found in other sources (Hosking and Wallis, 1997; Lin et al., 2004). What follows here is a brief description.

By necessity, precipitation frequency analysis employs a limited data sample to estimate the characteristics of the underlying population by selecting and parameterizing a probability distribution. The distribution is uniquely characterized by a finite set of parameters. In previous NWS publications such as NOAA Atlas 2, the parameters of a probability distribution have been estimated using the Moments of Product or the Conventional Moments Method (CMM). However, sample moment estimates based on the CMM have some undesirable properties. The higher order sample moments such as the third and fourth moments associated with skewness and kurtosis, respectively, can be severely biased by limited data length. The higher order sample moments also can be very sensitive or unstable to the presence of outliers in the data (Hosking and Wallis, 1997; Lin et al., 2004).

L-moments are expectations of certain linear combinations of order statistics (Hosking, 1989). They are expressed as linear functions of the data and hence are less affected by the sampling variability and, in particular, the presence of outliers in the data compared to CMM (Hosking and Wallis, 1997). The regional application of L-moments further increases the robustness of the estimates by deriving the shape parameters from all stations in a homogeneous region rather than from each station individually.

Probability distributions can be described using coefficient of L-variation, L-skewness, and L-kurtosis, which are analogous to their CMM counterparts. Coefficient of L-variation provides a measure of dispersion. L-skewness is a measure of symmetry. L-kurtosis is a measure of peakedness. L-moment ratios of these measures are normalized by the scale measure to estimate the parameters of the distribution shape independent of its scale. Unbiased estimators of L-moments were derived as described by Hosking and Wallis (1997).

Since these scale-free frequency distribution parameters are estimated from regionalized groups of observed data, the result is a dimensionless frequency distribution common to the N stations in the region. By applying the site-specific scaling factor (the mean) to the dimensionless distribution (regional growth factors), site-specific quantiles for each frequency and duration can be computed (Section 4.6.1).

Regional frequency analysis using the method of L-moments also provides tools for determining whether the data likely belong to similar homogeneous regions (e.g., climatic regimes) and for detecting potential problems in the quality of the data record. A measure of heterogeneity in a region, H1, uses coefficient of L-variation to test between-site variations in sample L-moments for a group of stations compared with what would be expected for a homogeneous region (Hosking and Wallis, 1997) (Section 4.4). A discordancy measure is used to determine if a station's data are consistent with the set of stations in a region based on coefficient of L-variation, L-skewness, and L-kurtosis (Section 4.3).

4.3. Dataset preparation

Rigorous quality control is a major and integral part of dataset preparation. The methods used in this project for ensuring data quality included a check of extreme values above thresholds, L-moment discordancy tests, a real-data-check (RDC) of quantiles, and a spatially-based quality control tool (*QCseries*), among others. Also, analyses such as a trend analysis of annual maximum series, a study of cross-correlation between stations, careful consideration of data from tropical systems, and testing of data series with large gaps in record provided additional data quality assurance. An interesting and valuable aspect of the analysis process, including spatial interpolation, is that throughout the process there are interim results and measures which allow additional evaluation of data quality. At each step, these measures indicate whether the data conform to the procedural assumptions. Measures indicating a lack of conformance were used as flags for data quality.

Quality control and data assembly methods. Initial quality control included a check of extreme values above thresholds, merging appropriate nearby stations, and checking for large gaps in records. Erroneous observations were eliminated from the daily, hourly, and n-minute datasets through a check of extreme values above thresholds. The thresholds were established for 1-hour and 24-hour values based on climatological factors and previous precipitation frequency estimates in a given region. Observations above these thresholds were checked against nearby stations, original records and other climatological bulletins.

Daily stations in the project area within 1 mile in horizontal distance and 100 feet in elevation with records that contain an overlap of 5 years or less or a gap between records of 5 years or less were considered for merging to increase record length and reduce spatial overlaps. The 24-hour annual maximum series of candidate stations were tested using a statistical t-test (at the 90% confidence level) to ensure the samples were from the same population and appropriate to be merged. Nine pairs of daily stations were merged.

The quality of longer duration (24-hour through 60-day) data was ensured in several ways. First, all longer duration annual maxima that exceeded their 1,000-year estimate by more than 5% were investigated for data quality and appropriate regionalization. This process was termed the "real-data-

check" since it was comparing computed precipitation frequency estimates with observed ("real") data and is used again in identifying homogeneous regions (Section 4.4). "Real-data-check" is used to refer to any check or test that compares the real observations or empirical frequencies with the calculated quantiles. The term is also used regarding a test for best-fitting distributions (Section 4.5). Second, common errors that potentially impacted the accumulation of longer durations were identified and corrected if necessary. For example, raw daily data were screened for repeating values in a month that were erroneously recorded or monthly totals that were entered as having occurred in a single day.

Additionally for this project, the difficulties of collecting observations during tropical system events led to additional quality control. Observations were often the accumulation of several days. The accumulations of such events were identified and distributed objectively based on the temporal distributions of nearby stations during the given event. Tropical system events were defined as hurricanes, tropical storms, tropical depressions, tropical waves and extra-tropical storms. In the original data archives, an accumulated value has an unknown distribution over the days of accumulation. Therefore, this could lead to the loss of valuable extreme precipitation data at sub-storm durations. For example, station 66-9774 (Villalba 1 E) had a four day accumulation of 18.34 inches during Hurricane Georges ending September 21st, 1998 which could have been extracted as an equal distribution of 4.85" on each day. Observations at other stations during Hurricane Georges suggested that most of the precipitation actually fell in one day. To generate appropriate distributions of the accumulated data, ratios of each daily observation over the sum of total precipitation for the given days at stations with complete data were spatially distributed to the target station. The ratio for each day was then applied at the target station to generate appropriately distributed daily precipitation values for the accumulated period. Seventeen separate tropical system events were identified and a total of 139 separate cases of accumulated data were distributed at individual stations during these events. In the above example, station 66-9774 (Villalba 1 E), this procedure resulted in the four-day accumulated 18.34" being distributed as daily values of 0.36", 0.28", 1.59", and 16.11".

Finally, inconsistencies at co-located daily and hourly stations where the hourly 24-hour accumulated maximums were significantly higher than their concurrent 1-day counterpart were reconciled. In particular, ten cases where the 24-hour accumulations were greater than 10 inches and the percent differences between the concurrent maximums were more than 20% were identified. These cases were omitted from the calculation of the constrained 1-day to unconstrained 24-hour conversion factors (Section 4.1.2). Notably, these cases occurred during tropical system events with high sustained winds. Based on observations at nearby stations, storm reports and local expertise, it was suspected that the hourly data were more reliable. Therefore, the 1-day observations in these cases were replaced by the concurrent hourly accumulated 24-hour observation that was converted to a 1-day constrained observation using the inverse of the established conversion factor.

Discordancy. The L-moment discordancy measure (Hosking and Wallis, 1997) was used for data quality control. In evaluating regions, it was also used to determine if a station had been inappropriately assigned to a region. The measure is based on the coefficient of L-variation, L-skewness and L-kurtosis, which represent a point in 3-dimensional space for each station. Discordancy is a measure of the distance of each point from the cluster center of the points for all stations in a region. The cluster center is defined as the unweighted mean of the three L-moments for the stations within the region being tested. Stations at which the discordancy value was 3.0 or greater were scrutinized for suspicious or unusual data or to consider if they belonged in another region or as an at-site (Section 4.4). Some stations that captured a single high event or had a short data record were discordant but were accepted in a homogeneous region since no climatological or physical reason was found to justify their exclusion. Discordancy was checked at stations for all durations. Appendix A.6 which provides lists of stations used in the project also provides the L-statistics and discordancy measure for the 24-hour data or 60-minute data for each station in its region.

1-day annual maximum series screening. The 1-day annual maximum series (AMS) data were thoroughly scrutinized. For instance, large gaps (i.e., sequential missing years) in the annual maximum series of stations were screened since it was not possible to guarantee that the two given data segments were from the same population (i.e., same climatology, same rain gauge, same physical environment). The screening process ensured data series consistency before the data were used. Station records with large gaps were flagged and examined on a case-by-case basis. Nearby stations were inspected for concurrent data years to fill in the gap if they passed a statistical test for consistency. If there were a sufficient number of years (at least 10 years of data) in each data segment, a t-test (at the 90% confidence level) was conducted to assess the statistical integrity of the data record. To produce more congruent data records for analysis, station record lengths were adjusted where appropriate.

The 1-day AMS data were also checked for linear trends in mean, linear trends in variance, and shifts in mean. Overall, the data were statistically free from trends and shifts. See Appendix A.3 for more details.

And finally, the 1-day AMS data were investigated for cross correlation between stations to assess intersite dependence since it is assumed for precipitation frequency analysis that events are independent. Cases where annual maxima overlapped (+/- 1 day) at stations within 50 miles and with more than 30 years of data were analyzed using a t-test for correlation coefficients that were statistically significant at the 90% confidence level. It was found that the degree of cross correlation between stations in the project area was very low. Only 6% of the data in the entire project area showed significant correlation based on t-test results. The impact of this cross correlation on the daily quantiles was very minimal. Relative errors were calculated for each of the nine daily regions from results of analyses using all stations versus analyses using only stations that were not cross correlated. The average relative errors in quantile estimation for all nine regions were small, 0.4%, 0.4%, 1.3% and 2.2% for 2-year, 10-year, 100-year and 1,000-year, respectively. In addition for this project, cases in each of the nine daily regions where 10-day or 30-day annual maxima overlapped (+/- 5 or 10 days, respectively) at stations within 50 miles and with more than 20 years of concurrent data were analyzed with similar results. Therefore, since the final quantiles of three durations were only minimally affected, it was concluded that it was not necessary to embed any measures to address dependence structures in the data.

Additional data series quality control. The annual maximum series of all durations were carefully reviewed for errors and consistency among stations at a given station using a spatially-based quality control tool and checking suspiciously low annual maxima.

HDSC has developed a software tool called *QCseries* to provide an objective, spatially-based process for screening annual maximum series (AMS) and partial duration series (PDS) data. The tool identifies maximum precipitation values that are suspect relative to concurrent data at nearby stations. *QCseries* computes a quality control score ranging from 0 to 10, where a lower score indicates a more suspect value, for each of annual maximum and partial duration maximum. The score is primarily based on the precipitation at nearby stations for the given observation day (+/- 1 day) and the deviation of the maximum from spatially distributed values of percent of mean annual precipitation. Other indicators of consistency used to compute the score include station density, spatial variability, and the highest station precipitation reported within a given distance. Maxima with low scores were identified for further verification for the AMS data of all durations (1-hour through 60-day). A subset of durations (1-hour, 6-hour, 1-day, 2-day, 4-day, 10-day and 60-day) for PDS data were checked since no additional errors were being found.

In addition, the validity of the lowest annual maxima in the AMS data of all longer durations (1-day through 60-day) was ensured. Cases where the lowest annual maximum and the second lowest annual maximum were more than 35% different were checked for data quality. This afforded the

opportunity to find cases where a given station was missing data on days during which a high precipitation event occurred at other stations. In such cases a determination was made to set that given year to missing rather than retain a probably erroneous low annual maximum caused by data sampling.

Deleted stations. Four stations were deleted from the analysis after a thorough investigation of their data quality and since no local climatological or objective reason could be found for their inconsistent data relative to nearby stations. Station 66-6017 (Matrullas Dam) was deleted because its 24-hour mean annual maximum and maximum observations appeared anomalously low compared to nearby stations, it had suspicious accumulated data occurring in the later part of the record and there were at least three stations in the nearby vicinity to fill in information spatially. Station 66-0849 (Bayney) was deleted because the data record had 25 years of missing data and was inconsistent with nearby stations. Station 66-4115 (Guavate Camp) was identified as anomalously too low in a review of resulting spatial patterns and through regression techniques during the PRISM process (Appendix A.4). Since it had only 22 years of data and many questionable or missing data (more than 15% were missing), it was deleted. Finally, station 66-1142 (Cacaos-Orocovis) was deleted because it was anomalously high compared to nearby stations, peer reviewers found it unreasonable based on their local climate knowledge and site inspection testimony suggested poor rainfall collection procedures.

4.4. Development and verification of homogeneous regions

The underlying assumption of the regional approach is that stations can be grouped in sets or "regions" in which stations have similar frequency distribution statistics except for a site-specific scale factor. Regions which satisfy this assumption are referred to as "homogeneous." The key to the regional approach is to construct a set of homogeneous regions for the entire project area. Hosking and Wallis (1997) make the case that homogeneous regions should be identified based on factors other than the statistics used to test the assumption of homogeneity. Regions in this project were first delineated subjectively based on climate, season(s) of highest precipitation, type of precipitation (e.g., general storm, convective, tropical storms or hurricanes, or a combination), topography and the homogeneity of such characteristics in a given geographic area.

The regions were then investigated using statistical homogeneity tests and other checks. As suggested in Hosking and Wallis (1997), adjustments of regions, such as moving stations from one region to another or subdividing a region, were made to reduce heterogeneity. The heterogeneity measure, H1, tests between-site variations in sample L-moments for a group of sites with what would be expected for a homogeneous region based on coefficient of L-variation (Hosking and Wallis, 1997). Earlier studies (Hosking and Wallis, 1997; also, personal discussion with Hosking at NWS, 2001) indicated that a threshold of 2 is conservative and reasonable. Therefore, an H1 measure greater than 2 (H1>2) indicated heterogeneity and H1<2 indicated homogeneity.

The regions for daily durations (24-hour through 60-day), Figure 4.4.1, were based on the 24-hour data. Long duration (48-hour through 60-day) L-moment results where H1 was greater than 2 were closely examined to validate data quality. In most of these cases, one or several stations were driving the H1 measure due to the nature of their data sampling. Omitting the offending station(s) would decrease H1 significantly and the 100-year precipitation frequency estimates and regional growth factors would change by 5% or less. Once identified and checked, the high H1 values in these regions were sometimes accepted without modifying the regions themselves.

Similarly, the regions for hourly durations (60-minute through 48-hour), Figure 4.4.2, were based on the 60-minute data. The other short durations (2-hour through 48-hour) where H1 was greater than 2 were also closely examined to validate data quality. Just as with the daily dataset, in most of these cases, one or several stations were driving the H1 measure due to the nature of their data sampling

and these regions were sometimes accepted without modifying the regions themselves, particularly given the limited hourly dataset and geographic locations of the offending stations.

Ideally, coefficient of L-variation is sufficient to assess regional homogeneity. However, in practice, the National Weather Service found that sole use of H1 was not optimum for defining a homogenous region. The effect of L-skewness on the formation of a homogenous region was also considered, particularly since coefficient of L-variation and L-skewness do not necessarily correlate, and to take into account effects on longer average recurrence intervals (ARI). L-skewness and L-kurtosis were accounted for using a so-called "real-data-check" process. Real-data-check flags occurred where a maximum observation in the real (observed) data series at a station exceeded a given frequency estimate or confidence limit, in this case the 1,000-year estimate. These stations were carefully investigated for data quality and appropriate regionalization.

Overall, effort was made during the subdivision process to mitigate discrepancies that could be caused by (1) sampling error due to small sample sizes, or (2) regionalization that does not reflect a local situation. The purpose of the regionalization process was to obtain reliable quantiles at each station to reflect local conditions and reduce the relative error. The final groups of stations in the project area are illustrated in Figures 4.4.1 for daily regions and 4.4.2 for hourly regions. Appendix A.7 lists the H1 values and regionally-averaged L-moment statistics for all regions for the 24-hour and 60-minute durations. The heterogeneity measures (H1) for each region and all durations are provided in Appendix A.8.

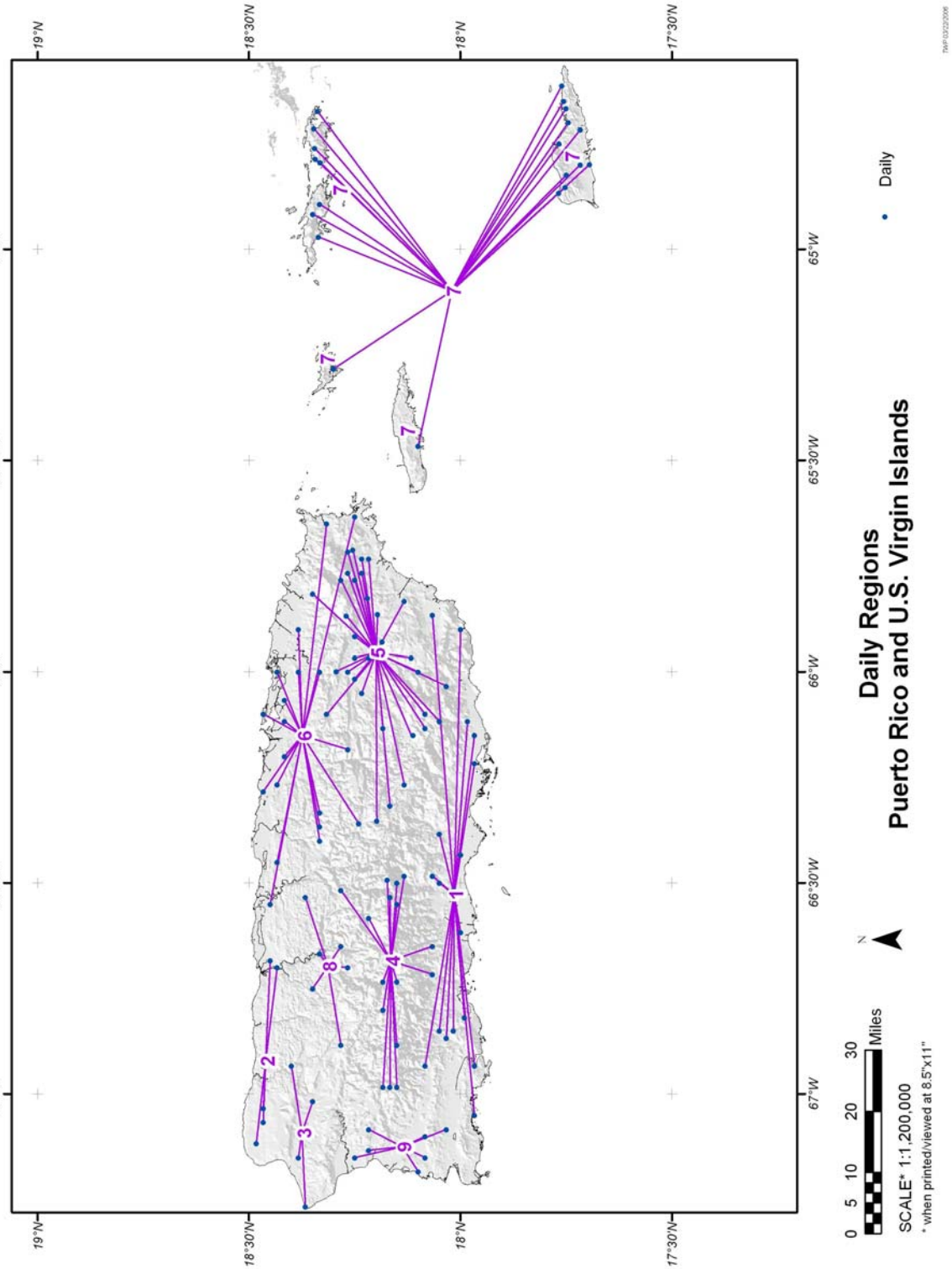
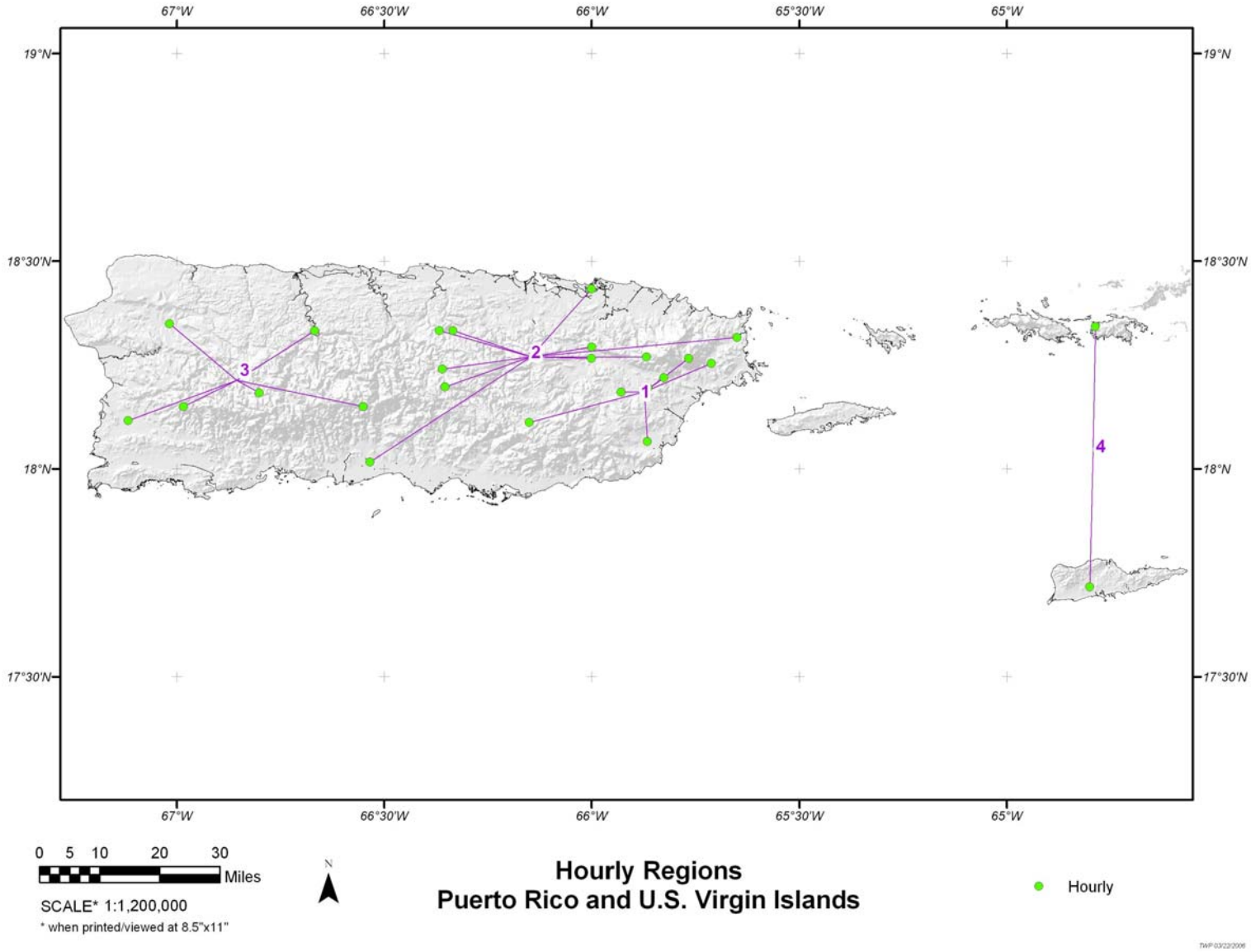


Figure 4.4.1. Regional groupings for daily data used to prepare NOAA Atlas 14 Volume 2.

Figure 4.4.2. Regional groupings for hourly data used to prepare NOAA Atlas 14 Volume 2.



At-site stations. In NOAA Atlas 14 Volumes 1 and 2, an at-site, instead of a regional, frequency analysis was a better approach to estimating the precipitation frequency quantiles at some stations. At-site analysis was used in those Volumes at stations where:

- It accounted for observed extreme precipitation regimes that the regional method could not resolve;
- The station had more than 50 data years to produce reasonable estimates independent of a region;
- The spatial interpolation process was able to accommodate it;
- Error in the estimate was reduced compared to when the station was included in a region.

No stations met these criteria in the development of this Volume.

4.5. Choice of frequency distribution

It was assumed that the stations within a region shared the same shape but not scale of their precipitation frequency distribution curves. It was not assumed that these factors or the distribution itself were common from region to region. In other words, a probability distribution was selected and its parameters were calculated for each region separately. Later during the sensitivity testing stage of the process, the selected distributions and their parameters were examined to ensure that they varied reasonably across the project domain. The goal was to select the distribution that best described the underlying precipitation frequencies. This goal was not necessarily achieved by a best fit to the sample data. Since a three-parameter distribution, which behaves both relatively reliably and flexibly, is more often selected to represent the underlying population, candidate theoretical distributions included: Generalized Logistic (GLO), Generalized Extreme Value (GEV), Generalized Normal (GNO), Generalized Pareto (GPA), and Pearson Type III (PE3). The five-parameter Wakeby distribution would have been considered only if the three-parameter distributions were found unsuitable for a region, but this did not happen. Three goodness-of-fit measures were used in this project to select the most appropriate distribution for the region. These were the Monte Carlo Simulation test, real-data-check test, and RMSE of the sample L-moments.

The Monte Carlo Simulation test. 1,000 synthetic data sets with the same record length and sample L-moments at each station in a region were generated using Monte Carlo simulation. Tests showed that 1,000 simulations were sufficient since means converged. Regional means of L-skewness and L-kurtosis were calculated for each simulation weighted by station data length. The regional means of all simulations were then calculated and plotted in an L-skewness versus L-kurtosis diagram and considered against candidate theoretical distributions (Figure 4.5.1). Assuming the distribution has L-skewness equal to the regional average L-skewness, the goodness-of-fit was then judged by the deviation from the simulated mean point to the theoretical distributions in the L-skewness dimension. To account for sampling variability, the deviation was standardized, (denoted as GZ) by assuming a Standardized Normal distribution Z. For the 90% confidence level, a distribution was acceptable if $|GZ| \leq 1.64$. Among accepted distributions, the distribution with the smallest GZ was identified as the most appropriate distribution according to this test (Hosking, 1991).

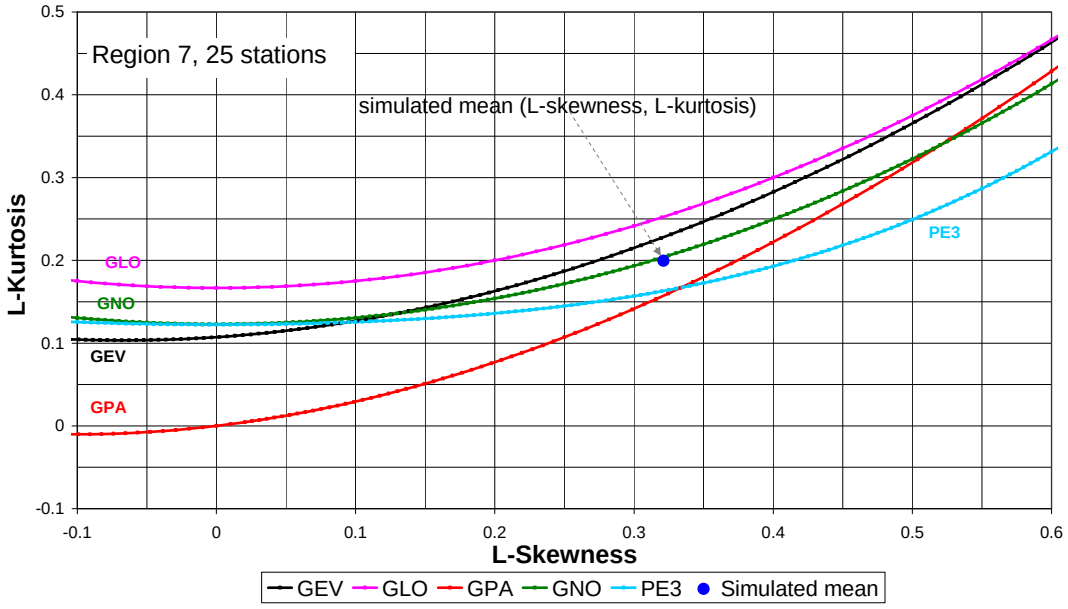


Figure 4.5.1. Plot of mean point from Monte Carlo simulations and theoretical distributions in L-skewness versus L-kurtosis diagram for daily region 7.

Real-data-check test. Similar to the practical application of a real-data-check in the construction of homogeneous regions, the real-data-check as a goodness-of-fit measure compared each theoretical distribution with empirical frequencies of the real (observed) data series at all stations in a region for recurrence intervals from 2-year to 100-year (Lin and Vogel, 1993). The relative error (or relative bias) of each distribution was calculated by comparing the quantiles that resulted from each fitted distribution to the empirical frequencies at each station. These were then averaged over all quantiles and stations in the region. This provided an indication of the degree of consistency between the empirical frequencies and the theoretical probabilities for the region. A smaller relative error indicated a better fit for that distribution. Although, relative error for a single station, or a few stations, is less meaningful in terms of goodness-of-fit due to sampling error, a relative error that is calculated over a number of stations to get a regional average is of statistical significance and was used as an index for the most appropriate distribution. For the ease of ranking distributions based on this test, the relative error was converted to an index in which the higher index indicated a smaller error.

RMSE of the sample L-moments. Unlike the Monte Carlo simulation test that emphasizes the effect of a simulated regional mean, the L-skewness and L-kurtosis of the real data were used in this test to assess the distribution. The deviation from the sample point (L- skewness, L- kurtosis) at each station against a given theoretical distribution in L- kurtosis scale was calculated. Then, the root-mean-square-error (RMSE) over the total set of deviations at all stations was obtained. The computation of the RMSE was done for each of the candidate distributions. The distribution with the smallest RMSE was identified as the most appropriate distribution based on this test.

Selecting the most appropriate distribution. A final decision of the most appropriate distribution for a region was primarily based upon a summary of the three tests. The goodness-of-fit tests were done on a region-by-region basis. Table 4.5.1 shows the results of the three tests for the 24-hour data in each of the nine daily regions. Table 4.5.2 shows the results for the 60-minute data in each of the 4

hourly regions. The results from the three tests provide a strong statistical basis for selecting the most appropriate distribution. However, the goodness-of-fit results were then weighed against climatologic and geographic consistency considerations. To reduce bull's eyes and/or gradients in precipitation frequency estimates between regions, the distribution identified by the three methods was sometimes changed during a review of results on a macro-scale. An effort was also made to maintain consistency of selected distribution from region to region. The use of an alternate distribution was supported with sensitivity testing to ensure that results using the selected distribution were acceptable (i.e., changes in 100-year quantiles were less than 5%). For example, in daily region 9, GNO was not ranked first statistically, but using the statistically best-fitting distribution, GEV, would have created a climatologically unreasonable steep spatial gradient between estimates in region 9 and region 1. Sensitivity tests showed that the 100-year 24-hour estimates in region 9 decreased by only 2.7% when using GNO rather than GEV. Therefore, GNO was selected for this region.

Based on the goodness-of-fit results, climatological considerations and sensitivity testing for all regions in the project area, GNO was selected to best represent the underlying distributions of the annual maximum data for all nine daily regions and all four hourly regions. GNO was also selected for the 5-, 10-, 15- and 30-minute annual maximum data that were used in the calculation of the n-minute ratios.

Table 4.5.1. Goodness-of-fit test results for 24-hour annual maximum series data in each daily region calculated for NOAA Atlas 14 Volume 3.

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
1	1st	GNO	-0.01	PE3	18.5	GNO	0.03343	GNO
	2nd	GEV	1.62	GNO	18.0	PE3	0.04572	
	3rd	PE3	-2.80	GEV	16.0	GEV	0.04657	
2	1st	PE3	-0.26	PE3	21.5	PE3	0.02644	GNO
	2nd	GNO	0.85	GNO	18.0	GNO	0.02912	
	3rd	GEV	1.43	GEV	16.0	GEV	0.03610	
3	1st	GLO	0.56	GEV	20.0	GNO	0.04267	GNO
	2nd	GEV	-0.61	GNO	17.5	GEV	0.04498	
	3rd	GNO	-1.10	PE3	15.0	GLO	0.05017	
4	1st	GEV	0.08	GNO	19.5	GNO	0.06918	GNO
	2nd	GLO	0.78	GPA	19.5	GEV	0.07128	
	3rd	GNO	-1.28	PE3	14.5	GLO	0.07759	
5	1st	GNO	0.46	PE3	17.5	GNO	0.04878	GNO
	2nd	GEV	2.22	GPA	17.5	GEV	0.05665	
	3rd	PE3	-2.57	GNO	17.0	GPA	0.05893	
6	1st	PE3	0.76	PE3	21.0	GPA	0.05362	GNO
	2nd	GNO	3.02	GPA	21.0	PE3	0.05670	
	3rd	GPA	-3.09	GNO	15.5	GNO	0.06471	
7	1st	GNO	-0.26	PE3	18.5	GNO	0.05464	GNO
	2nd	GEV	0.95	GNO	18.0	GEV	0.06269	
	3rd	GLO	2.20	GPA	15.5	GPA	0.06340	
8	1st	GLO	-1.10	GLO	20.0	GEV	0.04885	GNO
	2nd	GEV	-1.75	GEV	16.5	GLO	0.06080	

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
	3rd	GNO	-2.39	GNO	15.5	GNO	0.06476	
9	1st	GEV	-0.10	GLO	19.0	GEV	0.06768	GNO
	2nd	GLO	0.62	GEV	17.0	GPA	0.06891	
	3rd	GNO	-1.00	GNO	16.0	GNO	0.07027	

Table 4.5.2. Goodness-of-fit test results for 60-minute annual maximum series data in each hourly region calculated for NOAA Atlas 14 Volume 3.

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
1	1st	GEV	0.09	GPA	17.0	GEV	0.06733	GNO
	2nd	GNO	-0.34	PE3	16.5	GNO	0.07069	
	3rd	PE3	-1.16	GEV	15.5	PE3	0.08394	
2	1st	PE3	1.36	PE3	18.5	GEV	0.05590	GNO
	2nd	GEV	1.47	GNO	16.5	GNO	0.06264	
	3rd	GNO	1.64	GEV	16.5	PE3	0.06351	
3	1st	GNO	-0.37	PE3	18.5	GEV	0.05465	GNO
	2nd	PE3	-0.49	GNO	18.5	GNO	0.05484	
	3rd	GEV	-0.62	GEV	18.5	PE3	0.05717	
4	1st	GEV	0.00	PE3	16.0	GNO	0.01229	GNO
	2nd	PE3	-0.02	GNO	16.0	PE3	0.01291	
	3rd	GNO	0.06	GLO	16.0	GEV	0.01480	

4.6. Estimation of quantiles

4.6.1. Regional growth factors

In the index-flood based regional analysis approach, regional growth factors (RGFs) are defined as the quantiles of a regional dimensionless distribution. Regional growth factors are obtained by fitting the selected dimensionless distribution function with the weighted average L-moment ratios (or parameters) for a region that were computed using data re-scaled by the mean of the annual maximum series (Hosking and Wallis, 1997). Because the parameters are constant for each region, there is a single RGF for each region that varies only with frequency and duration. A table of RGFs for all durations for each region is provided in Appendix A.9. The RGFs are then multiplied by the site-specific scaling factor to produce the quantiles at each frequency and duration for each site. The site-specific scaling factor used in this project was the mean of the annual maximum series at each site. This scaling factor is often referred to as the "Index Flood" because the genesis of the statistical approach was in flood frequency analysis.

In this project, the scaling factors for each duration were first spatially interpolated to fine scale grids (Section 4.8.1) to take advantage of the RGFs at each frequency and obtain grids of the quantiles. A unique spatial interpolation procedure (Section 4.8.2) was developed to maintain differences between regions but generate spatially smooth quantiles across regional boundaries.

4.6.2. 1-year computation

The 1-year average recurrence interval (ARI) precipitation frequency estimates were computed for this project. ARI is the average period between exceedances (at a particular location and duration) and is associated with the partial duration series (PDS). Annual exceedance probability (AEP) is the probability that a particular level of rainfall will be exceeded in any particular year (at a particular location and duration) and is derived using the annual maximum series (AMS). An AEP depth or intensity may be exceeded once or more than once in a year. (Section 3.2 provides additional discussion on this topic.)

A 1-year AEP estimate, associated with AMS, has little meaning statistically or physically. However, the 1-year ARI, associated with PDS does have meaning and is used in several practical applications. The equation $T_{PDS} = [\ln(\frac{T_{AMS}}{T_{AMS}-1})]^{-1}$ (Chow et al., 1988), which is distribution free, provided a mathematical base for converting between frequencies for the AMS data and the PDS data. Here, T_{AMS} and T_{PDS} stand for the frequency associated with the AMS data and the frequency associated with the PDS data, respectively. The equation can be transformed into the following:

$$T_{AMS} = \frac{1}{1 - e^{-\frac{1}{T_{PDS}}}}$$

Therefore, T_{AMS} = 1.58-year when T_{PDS} = 1-year from the equation. This means that a PDS 1-year event is equivalent to an AMS 1.58-year event. This relationship was used to calculate the 1-year ARI from AMS data for this project. Appendix A.9 provides the regional growth factors computed for the 1.58-year AMS results. However, for all ARIs other than 1-year, the results were obtained by analyzing both AMS and PDS data separately, averaging ratios of PDS to AMS quantiles and then applying the average ratio to the AMS results (Section 4.6.4).

4.6.3. Practical consistency adjustments

In reality, data do not always behave ideally. Nor are datasets always collected perfectly through time or in dense spatial networks. Since quantiles for each duration and station in this project were computed independently, the practical adjustments described below were applied to produce realistic final results that are consistent in duration, frequency and space.

Annual maximum consistency adjustment. At some daily stations, there were inconsistencies in the annual maximum time series from one duration to the next. Specifically, a shorter duration observation in a given year may have sometimes been greater than the subsequent longer duration. Often this occurred because there were a significant number of missing data surrounding that particular case. A longer duration for the case could not be accumulated if the data immediately adjacent the relevant observations were not available. It also occurred in some cases when the average conversion factors that account for different sampling intervals were applied (e.g., 1-day data to 24-hour data; Section 4.1.2). If left unadjusted, these inconsistencies could result in a negative bias of longer duration precipitation frequency estimates relative to reality. Therefore, large inconsistencies in the annual maxima of a given year from one duration to the next were investigated and data added or corrected where possible. If missing data could not be found and/or the difference between the 2 durations was small (<10%), then the longer duration was set equal to the shorter duration. This adjustment ensured consistency from one duration to the next longer duration for each given year at a station.

Co-located hourly and daily station adjustment. Since hourly and daily durations were computed separately and from different datasets, it was necessary to explicitly ensure consistency of precipitation frequency estimates through the durations at co-located daily and hourly stations. At co-located daily and hourly stations the 24-hour estimates from the daily data were retained since they were based on more stations and generally had longer record lengths. The quantiles of co-located stations were adjusted for consistency particularly across the 12-hour and 24-hour durations where disparities could occur. There are a number of possible reasons for such disparities, such as gage differences or different recording periods. The adjustment preserved the daily 24-hour quantiles and the hourly distribution for the 120-minute (2-hour) through 12-hour quantiles at the given hourly station. The 24-hour through 2-hour quantiles for co-located hourly stations were adjusted using *station-specific* ratios of the station daily and hourly 24-hour means and ratios of the daily and hourly 24-hour regional growth factors (RGFs) at all frequencies (1.58-yr, 2-yr, 5-yr, …, 1,000-yr).

To ultimately avoid discontinuities at the 60-minute quantile relative to adjusted 2-hour through 24-hour quantiles and n-minute quantiles and reduce spatial bull's eyes in the final maps an additional criteria was developed for the adjustment to the 60-minute quantile. This better accommodated the different hourly and daily regions, close spatial proximity of stations, the average 1-hour to 60-minute conversion factor, and the application of n-minute ratios.

In some cases, the station-specific ratios of daily region versus hourly region RGFs at co-located stations were less than 1.0. This was not common but did occur. When the daily 100-year 24-hour RGF/hourly 100-year 24-hour RGF, which was used as an index, was *less* than 1.0, the station-specific adjustment ratios were applied from 24-hour through 60-minute to maintain consistency over all hourly durations and avoid over-adjusting. However, when the station-specific 100-year 24-hour RGF ratio was *greater* than 1.0, the 60-minute quantile was adjusted using regionally averaged RGF and 24-hour mean ratios calculated from all co-located stations in the hourly region to achieve a more reasonable and spatially consistent result.

The final result using the station-specific adjustment of the 60-minute quantile may not be as spatially smooth as the regionally averaged adjustment. However, the station-specific adjustment is more representative of the station data and mitigates the risk of over-adjusting.

In addition, the co-located adjustment was modified slightly for Volume 3 to accommodate a unique situation presented in hourly region 3 at 66-8881 (San Sebastian). The unique data characteristics at this station coupled with the relatively few hourly stations and different daily and hourly regional characteristics created discontinuities relative to nearby stations. At San Sebastian, the daily to hourly RGF ratios at each frequency were unusually low. The data of two or more hourly durations at the station shared the same annual maximum or had a very close values which created a very flat slope for quantiles from 5-year through 1,000-year. To ensure the consistency of precipitation frequency estimates in such a case, the regional RGF ratio and station-specific mean ratio were used to adjust the 60-minute duration at a station when the following criteria were met: (1) the station-specific daily/hourly 100-year RGF ratio was less than 1.0, and (2) the difference (range) of the 100-year RGF ratios of all hourly stations in the hourly region was greater than 0.2, and (3) the range divided by the lowest 100-year RGF ratio was equal to or greater than 0.4. These criteria were empirically determined and tested. The adjustment results in precipitation frequency estimates at such a co-located station that are more reasonable and consistent throughout the durations (60-minute through 24-hour) and with respect to other stations in that hourly region.

Hourly-only station consistency adjustment. To ensure that hourly-only stations were consistent with nearby co-located hourly/daily stations that occur in different regions and to reduce spatial bull's eyes observed in hourly results, an adjustment was applied to hourly-only stations. Specifically, the 48-hour through 60-minute quantiles for hourly-only stations were adjusted using a regionally averaged ratio of the daily and hourly 24-hour means and a set of regionally averaged RGF ratios at

all frequencies (1.58-yr, 2-yr, 5-yr, …, 1,000-yr) calculated from all co-located stations within the hourly region.

Internal consistency adjustment. Since the quantiles of each duration at a given station were calculated separately, inconsistencies could occur where a shorter duration had a quantile that was higher than the next longer duration at a given average recurrence interval. For example, it could happen that a 100-year 2-hour quantile was greater than a 100-year 3-hour quantile at a station. This result, although based on sound statistical analysis, is physically unreasonable. Such results primarily occurred where durations had similar mean annual maxima but the shorter duration had higher regional parameters, such as coefficient of L-variation and L-skewness that produced a quantile higher than the longer duration quantile. The underlying causes of such an anomaly were primarily discontinuities in selection and parameterization of distribution functions between durations, data sampling variability, and the application of average conversion factors to convert 1-hour data to 60-minute and to convert 1-day data to 24-hour.

Such inconsistencies were identified when the ratio of the longer duration to the next shorter duration quantiles was less than 1.0 for a given average recurrence interval. If the inconsistency occurred in the higher frequencies, it was mitigated by distributing the surplus of the ratio, which was greater than 1.0, of the previous frequency for those durations at a constant slope to the ratios of the inconsistent frequency and higher through 1,000-year, until it converged at 1.0 after 1,000-year (Table 4.6.1). If the inconsistency occurred in the lower frequencies, it was mitigated by distributing the surplus of the ratio, which was greater than 1.0, of the following frequency for those durations at a constant slope to the ratios of the inconsistent frequency and lower through 1.58-year, until it converged at 1.0 before 1.58-year. The adjusted ratios were then, appropriately, greater than or equal to 1.0. Table 4.6.1 shows an example from the Ohio River basin and surrounding states of the 3-hour to 2-hour ratios for average recurrence intervals from 1.58-year to 1,000-year at a station before and after the internal consistency adjustment. Figure 4.6.1 shows the associated 3-hour quantiles before and after adjustment.

In most cases, applying the adjustment from 1.58-year through 1,000-year was sufficient. However, in some cases where the inconsistency occurred only for some frequencies, such as between 50-year and 500-year only, adjustments were still required from 1.58-year through 1,000-year to ensure consistency without changing the existing compliant quantiles.

Table 4.6.1. Example of the internal consistency adjustment of quantiles showing the ratios of 3-hour to 2-hour quantiles for 1.58-year to 1,000-year at station 15-3709, Hazard, Kentucky.

3-hour to 2-hour ratios	1.58-yr	2-yr	5-yr	10-yr	25-yr	50-yr	100-yr	200-yr	500-yr	1,000-yr
Before adjustment	1.025	1.022	1.017	1.009	1.004	0.997	0.994	0.990	0.983	0.979
After adjustment	1.025	1.022	1.017	1.009	1.004	1.003	1.003	1.002	1.002	1.001

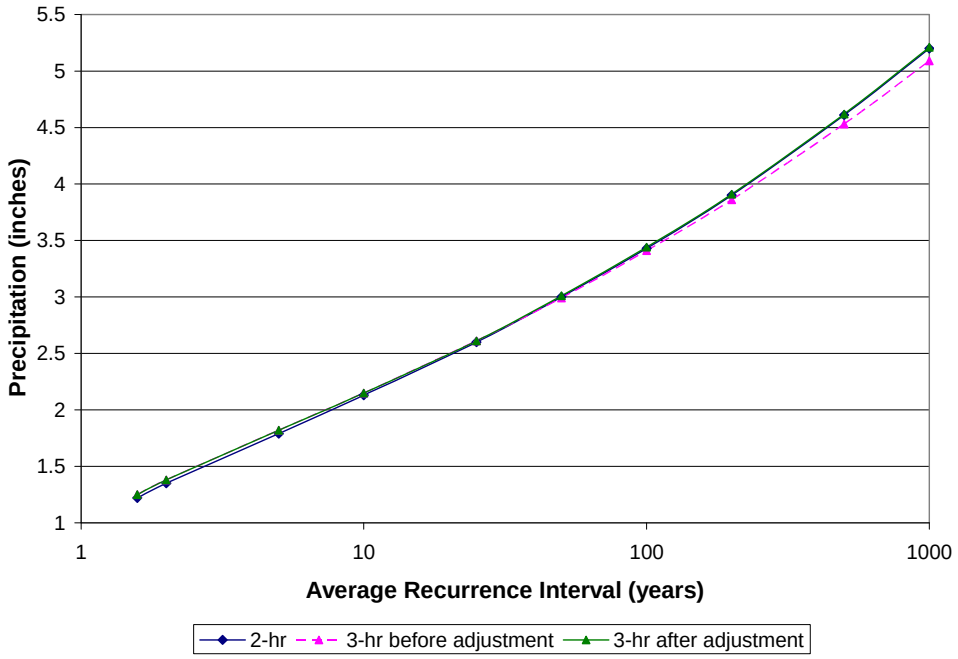


Figure 4.6.1. Example of internal consistency adjustment between the 3-hour and 2-hour quantiles at station 15-3709, Hazard, Kentucky.

4.6.4. Conversion factors for AMS to PDS

Annual maximum series (AMS) data consist of the largest case in each year, regardless of whether the second largest case in a year exceeds the largest cases of other years. In this project, the partial duration series (PDS) data is a subset of the complete data series where highest N cases are selected and N equals the number of years in the record. Such a series is also called an annual exceedance series (AES) (Chow et al., 1988). In this Atlas, the use of PDS refers to AES.

AMS data were used for all durations from 5-minute to 60-day and for annual exceedance probabilities of 1 in 2 to 1 in 1,000. The use of the AMS data is consistent with the concept of frequency analysis and the manipulation of annual probabilities of exceedance, and is consistent with the basis of development of the statistics used in this project. The statistical approach is less well demonstrated for PDS data. However, to remain consistent with the previous studies (e.g., NOAA Atlas 2) and to meet today's needs at lower return periods, NOAA Atlas 14 is also presented in terms of PDS results. The differences in meaning between AMS-based results and PDS-based results are discussed in Section 3.2.

PDS results were obtained by analyzing both AMS and PDS data separately, averaging ratios of PDS to AMS quantiles and then applying the average ratio to the AMS results. The PDS-AMS ratios were developed by independently fitting distributions to AMS and PDS data separately for each region before averaging. Figure 4.6.2 shows the average results of the PDS-AMS ratios for 24-hour data over the nine homogenous regions in the project area. Ratios computed using other durations were consistent with the 24-hour duration. To account for sampling variability and to generate a smooth consistent curve, an asymptote of 1.010 was applied for 100-year and above.

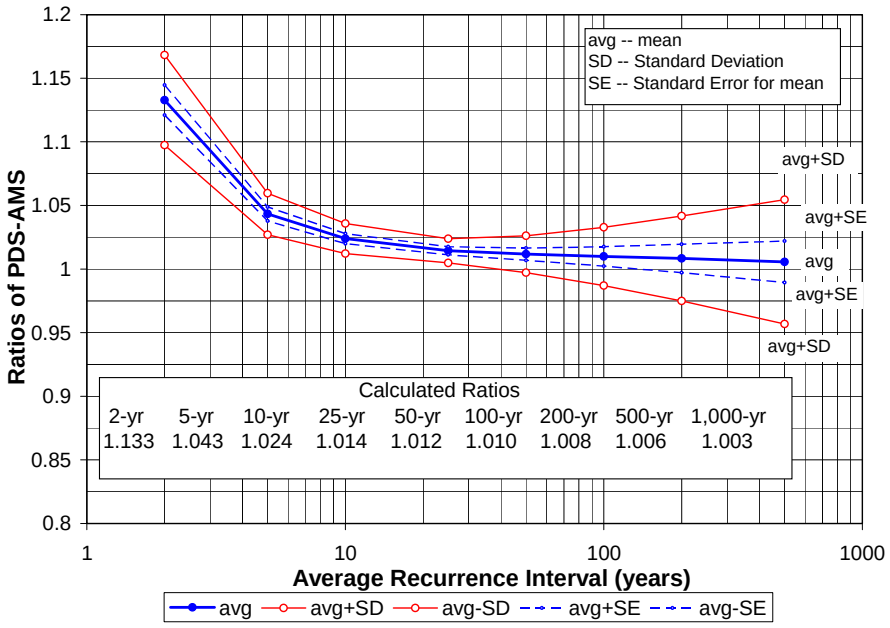


Figure 4.6.2. PDS-AMS ratio results for average recurrence intervals for the 24-hour duration over the nine homogeneous regions used to prepare NOAA Atlas 14 Volume 3.

The ratios for this Atlas (Table 4.6.2) are consistent with expectations, previous NOAA Atlas 14 Volumes and theoretical computations. For example, Chow (1988) proposed a mathematical relation in terms of recurrence interval (T) between PDS (or AES) and AMS:

$$T_{AES} = [\ln(\frac{T_{AMS}}{T_{AMS}-1})]^{-1}$$

According to this relation, a 2-year AMS value is equivalent to a 1.44-year AES (or PDS) value. Results are consistent with this relation. The ratios are also reasonably consistent with results from NOAA Atlas 14 Volumes 1 and 2 for the semiarid southwest and Ohio River basin and surrounding states precipitation frequency projects, respectively (Bonnin et al., 2003; Bonnin et al., 2004). The consistency of these PDS to AMS ratios with other derivations lends strong support to the validity of the results of this project because the PDS and AMS quantiles were derived independently using different probability distributions. To derive the PDS to AMS ratios, regional data were used. The best-fitting distributions for each individual region for the PDS and AMS computations were used.

Table 4.6.2. NOAA Atlas 14 Volumes 1,2, and 3 PDS to AMS ratios for all durations with asymptotes applied.

NOAA Atlas 14	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
Vol. 1 Semiarid southwest	1.113	1.029	1.013	1.006	1.004	1.004	1.004	1.004	1.004
Vol. 2 Ohio River basin and surrounding states	1.086	1.023	1.010	1.004	1.004	1.004	1.004	1.004	1.004
Vol. 3 Puerto Rico and U.S. Virgin Islands	**1.133**	**1.043**	**1.024**	**1.014**	**1.012**	**1.010**	**1.010**	**1.010**	**1.010**

4.7. Estimation of confidence limits

The National Weather Service is providing confidence limits for the estimates to quantify uncertainty. This will allow users a greater understanding of the uncertainty and will thus improve the utility of the estimates in engineering and environmental design practice. The quantiles per se are statistical variables that vary within an unknown range following an unknown distribution. To quantitatively assess the uncertainty, a Monte Carlo simulation technique was used to generate 1,000 synthetic data sets having the same statistical features.

Upper and lower confidence limits at the 90% confidence level were computed for each station's precipitation frequency estimate using Monte Carlo simulations coupled with the regional L-moments method, as suggested by Hosking and Wallis (1997). The sample parameters at each station were used in 1,000 Monte Carlo simulations to produce 1,000 samples with the same data length and same average regional parameters as the actual data. 1,000 quantiles were calculated for each station and then the upper 5% and lower 5% were delineated to produce the upper and lower confidence bounds. For n-minute data, the n-minute ratios (Section 4.1.1, N-minute data) were applied to the 60-minute upper/lower grids to compute the upper and lower bounds for n-minute estimates.

Confidence limits were adjusted to be consistent with their corresponding quantiles by applying ratios of the unadjusted quantiles and the adjusted quantiles in a manner comparable to the co-located hourly and daily station and hourly-only station consistency adjustments. 24-hour confidence limits at co-located or daily-only stations were derived from the station in the daily region analysis.

The estimation of confidence limits provides error bounds on the quantiles themselves under the assumption that the data have been well quality controlled and does not include error associated with rainfall measurement or the spatial interpolation procedure.

4.8. Spatial interpolation

4.8.1. Mean annual maximum (or "Index flood") grids

As explained in Section 4.6.1, mean annual maximum values were used as the site-specific scaling factor to generate precipitation frequency estimates from regional growth factors (RGFs). The station mean annual maximum values were spatially interpolated to produce mean annual maximum, or "index flood", grids using technology developed by Oregon State University's PRISM Group, formerly Spatial Climate Analysis Service (SCAS). The PRISM Group has developed PRISM (Parameter-elevation Regressions on Independent Slopes Model), a hybrid statistical-geographic approach to mapping climate data (Daly and Neilson, 1992; Daly et al., 1994; Daly et al., 1997; Daly et al., 2002). PRISM spatially interpolated the HDSC-calculated mean annual maximum values for all durations by using a naturally strong relationship with mean annual precipitation. Because of the limited hourly (≤ 12-hour) data for this project, additional effort was made to bring the hourly station density up to that of the daily (≥ 24-hour) stations by objectively developing ≤ 12-hour mean annual maximum data for daily-only stations (≥ 24 hour only) during the PRISM modeling of ≤ 12-hour durations (details may be found in Appendix A.4).

The PRISM Group adapted PRISM to use their existing mean annual precipitation grids (USDA-NRCS, 1998), transformed using the square-root, as the predictor grid for interpolating mean annual maximum precipitation to a uniformly spaced grid. Mean annual precipitation was used as the predictor because it is based on a large data set, accounts for spatial variation of climatic information and is consistent with methods used in previous projects, including NOAA Atlas 2 (Miller et al., 1973). PRISM uses a unique regression function for each target grid cell and has the ability to account for: user knowledge, the distance of an observing station to the target cell, if the station is in a cluster of stations grouped together, the difference between station and target cell mean annual precipitation, topographic facet, and coastal proximity. Other parameters include radius of influence, minimum number of stations on a facet, and total number of stations required for the regression to estimate the mean annual maximum precipitation at a given grid cell. PRISM cross-validation statistics were computed where each observing station was deleted from the data set one at a time and a prediction made in its absence. Results indicated that any overall bias was less than 2 percent and mean standard error was less than 10 percent for this Atlas. Appendix A.4 provides additional information regarding the details of the work done by the PRISM Group for HDSC.

Table 4.8.1 lists the mean annual maximum (a.k.a. "index flood") grids, one for each duration of the project, that were interpolated by PRISM. The resulting high-resolution 3-seconds (about 80 meters x 80 meters or 262 feet x 262 feet) mean annual maximum grids then served as the basis for deriving precipitation frequency estimates at different recurrence intervals using a unique HDSC-developed spatial interpolation procedure, the Cascade, Residual Add-Back (CRAB) derivation procedure (described in detail in Section 4.8.2).

Deviations may occur between the observed point mean annual maximum values in the HDSC database and the resulting grid cell value due to spatial interpolating and smoothing techniques employed by PRISM. The "HDSC database" consists of precipitation frequency estimates, mean annual maximum values and metadata (longitude, latitude, period of record, etc.) for each station. These deviations occur because PRISM produces interpolated values that mitigate differences between the observed point estimates and surrounding stations with similar climate, mean annual precipitation, elevation, aspect, distance from large water bodies and rain-shadow influences. See Appendix A.4 for more details.

Table 4.8.1. Mean annual maximum grids interpolated by PRISM.

Duration
60-minute
120-minute
3-hour
6-hour
12-hour
24-hour
48-hour
4-day
7-day
10-day
20-day
30-day
45-day
60-day
Total 14

4.8.2. Derivation of precipitation frequency grids

The Cascade, Residual Add-Back (CRAB) grid derivation procedure is a unique spatial interpolation technique, developed by HDSC, to convert mean annual maximum grids into grids of precipitation frequency estimates (see Figure 4.8.1). The CRAB philosophy was first used in the derivation of several of the National Climatic Data Center's Climate Atlas of the United States maps (Plantico et al., 2000).

CRAB accommodates spatial smoothing and interpolating across "region" boundaries to eliminate potential discontinuities due to different RGFs as a result of the regional L-moment analysis. The CRAB process, as the term cascade implies, uses the previously derived grid to derive the next grid in a cascading fashion. The technique derives grids along the frequency dimension with quantile estimates for different durations being separately interpolated. Hence, duration-dependent spatial patterns evolve independently of other durations. The CRAB process utilizes the inherently strong relationship between different frequencies for the same duration. In reality, this linear relationship is equivalent to the ratio of RGFs (e.g., 100-year 24-hour RGF over the 50-year 24-hour RGF) and is a constant for each region. CRAB initially makes a generalization that all regions have the same RGF ratios, thereby causing the linearly-predicted precipitation frequency estimates in some regions to be over predicted, while others under predicted. To account for these regional differences, CRAB utilizes residuals – the differences between the precipitation frequency estimates from the generalized all-region RGF ratios and the actual precipitation frequency estimates at each station. As a by-product of the generalization, the residuals (at each station) within each individual region are either all positive, negative or close to zero thereby supporting spatial autocorrelation and skill in interpolating the residuals. This combined with the inherently strong linear predictability from one frequency to the next makes CRAB an effective and accurate method for deriving the suite of precipitation frequency grids.

As mentioned above, the CRAB derivation process utilizes the strong, linear relationship between a particular duration and frequency, the *predictor* estimates, and the next rarer frequency of the same duration. Figure 4.8.2 shows the relationship between the *predictor* precipitation frequency estimates, 50-year 24-hour in this example, and the subsequent precipitation frequency estimates, 100-year 24-hour. The R-squared value here of 0.9936 is very close to 1.0 which was common

throughout all of the regressions. Since this was calculated using all stations in the project area, the slope of this relationship (1.2788) can be thought of as an average domain-wide RGF ratio. Regional differences are then accounted for using residuals.

A summary of the complete CRAB derivation procedure is illustrated in Figure 4.8.1 and can be summarized in a series of steps. In this description, the term *predictor* refers to the previous grid upon which the subsequent grid is based.

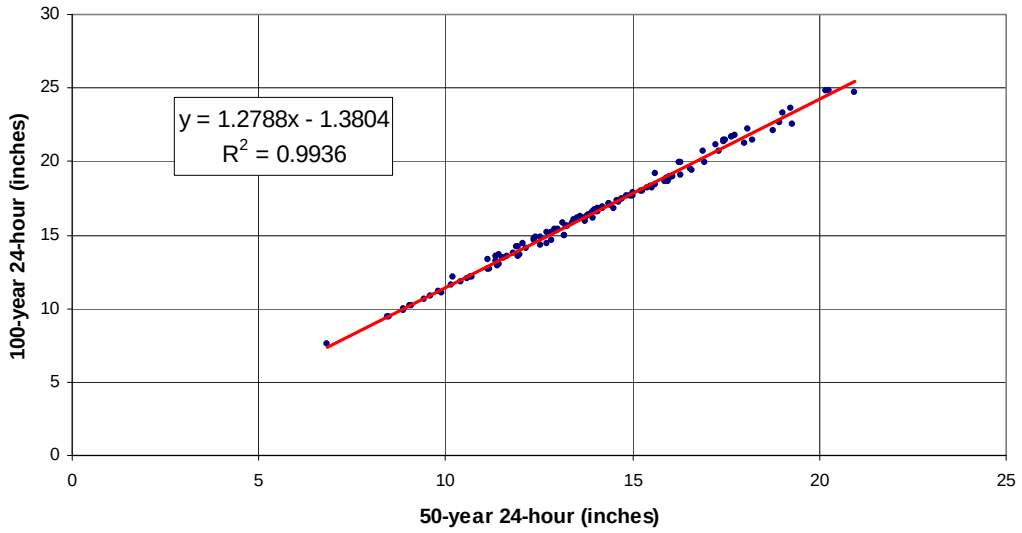


Figure 4.8.2. A scatter plot of 100-year 24-hour vs. 50-year 24-hour precipitation frequency estimates and the linear regression line from NOAA Atlas 14 Volume 3.

Step 1: Development of regression. The cascade began with the mean annual maximum grid derived by the PRISM Group using PRISM for a given duration as the initial *predictor grid* (e.g., 24-hour mean annual maximum) and the 2-year frequency as the subsequent grid (e.g., 2-year 24-hour). All precipitation frequency estimates in the HDSC database were adjusted to accommodate the spatial smoothing of the PRISM mean annual maximum grids. An adjustment factor was calculated based on the difference between the mean annual maximum PRISM grid cell value and the point mean annual maximum as computed from observed data as listed in the HDSC database. The adjustment factor was a station-unique value applied to the precipitation frequency estimates and was independent of frequency. For example, a station has an observed mean annual maximum 60-minute value (from the database) of 0.82 inches, but the PRISM grid cell at this station has a value of 0.861 inches. This results is an adjustment factor of 1.05 which is applied to each of the 60-minute precipitation frequency estimates (2-years through 1,000-years) before constructing the regression equation. These adjusted precipitation frequency estimates are equivalent to the *actual estimates*. In most cases, this adjustment was ±5% (See Appendix A.4 for more details). A global (all-region) relationship for each duration/frequency pair was developed at the beginning of each iteration based on station precipitation frequency estimates, adjusted for spatial smoothing, at all stations.

To develop the global relationship, an x-y data file was built where initially x was the mean annual maximum for a given duration and y the 2-year precipitation frequency estimate for that duration for each observing station. The slope and y-intercept of a least-square fit linear regression line using x and y for all stations in the domain was calculated. For each individual region, the slope of such a line is equivalent to the 2-year RGF in the initial run and equivalent to the RGF ratio in subsequent runs.

Figure 4.8.1. Flowchart of the cascade residual add-back (CRAB) grid derivation procedure beginning with the mean annual maximum grid of the x-duration and deriving the 2-year x-duration grid as an example.

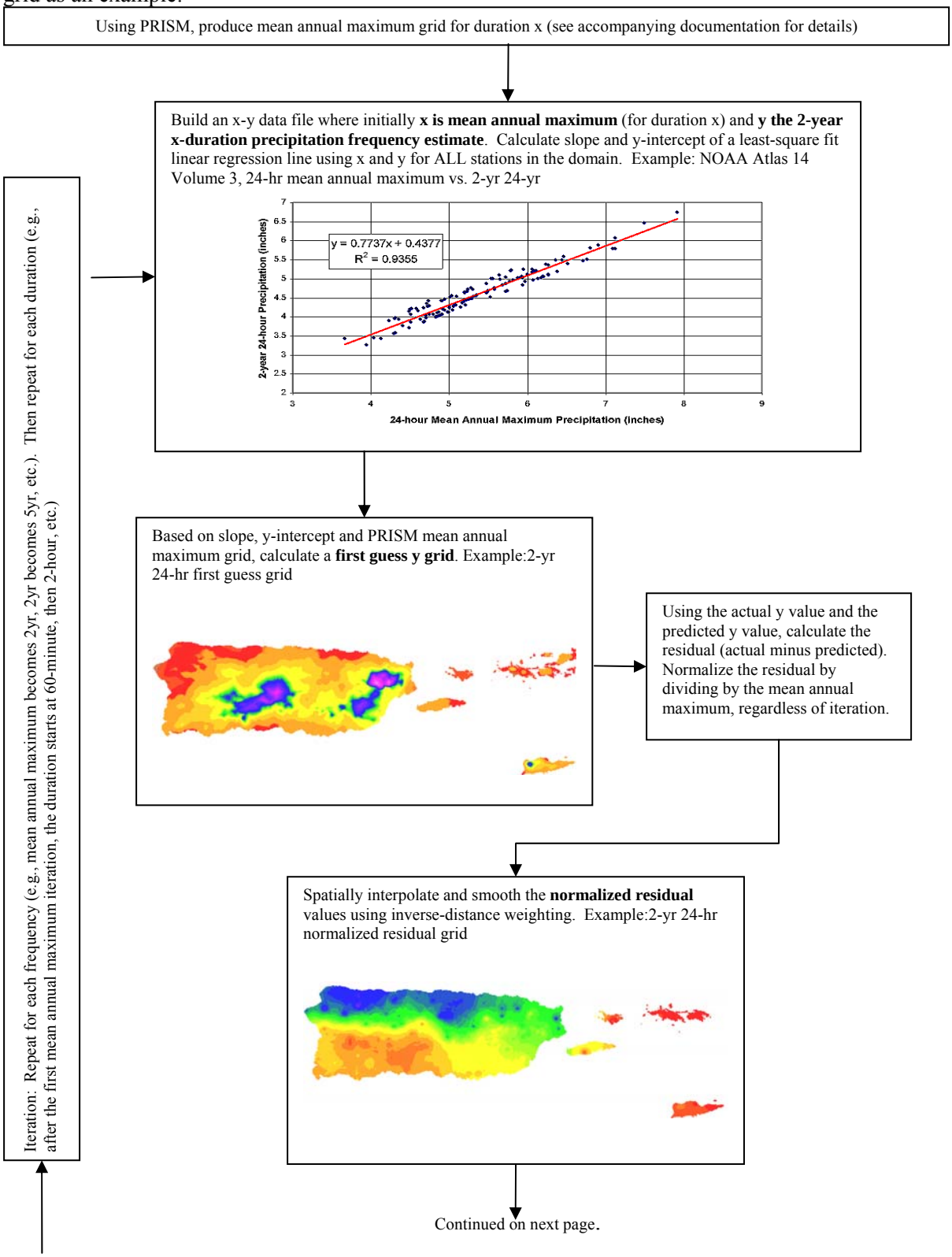
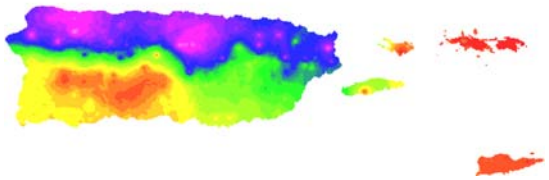
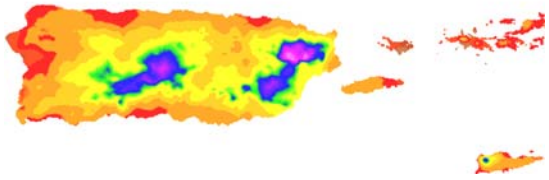


Figure 4.8.1. cont'd

Multiply the **normalized residual grid** by the mean annual maximum grid, regardless of iteration, to obtain a grid of **actual residuals**. Example:2-yr 24-hr actual residual grid



Add the actual residual grid to the first guess grid to obtain **pre-final grid**. Example:2-yr 24-hr grid



The unmasked, unfiltered, unadjusted (for internal consistency (IC) violations) grid is saved and used as the next **predictor grid**. Apply small (3x3 grid cell) center-weighted, block filter to the entire pre-final filtered grid to promote smooth contours.

In cases where x is not mean annual maximum (*Note: mean annual maximum is only used in the first iteration)*, check the **pre-final, filtered grid** for duration-based IC violations by making sure y is greater than the next lower duration final grid at the same frequency (e.g., 5y24h>5y12h). For grid cells that violate this rule, adjust y by setting the grid cell equal to the next lower duration plus 1%.

In cases where x is not mean annual maximum (*Note: mean annual maximum is only used in the first iteration.)*, check the **pre-final, filtered grid** for frequency-based IC violations by making sure y is greater than x (the next higher frequency final grid, e.g. 5y24h>2y24h). For grid cells that violate this rule, adjust y by setting the grid cell equal to the next higher frequency plus 1%. The result is the **final grid**.

If x & y represent upper and/or lower precipitation frequency bounds, then subject the grids to additional IC checks (e.g., make sure 5y24h upper > 5y24h).

If duration equals 60-minutes, calculate the n-minute (5-, 10-, 15- and 30-min) grids by applying the domain-wide 60-min to n-min ratios to the final grid.

Step 2: Development of *first guess grids*. The global linear regression relationship was then applied, using a Geographic Information System (GIS), to the *predictor grid* (e.g., 24-hour mean annual maximum) to establish a *first guess grid* (e.g., 2-year 24-hour) that was not necessarily equivalent to the *actual* estimates which were based on the unique RGF for each region.

Step 3: Development of spatially interpolated *residual grids*. To account for the regional differences, residuals (*actual* estimates minus *predicted* estimates) at each station were calculated. Here, *predicted* estimates (e.g., 2-year 24-hour) were those derived in the *first guess grid*. The residuals were normalized by the mean annual maximum to facilitate the interpolation of residuals to ungauged locations.

The normalized residuals at each station were then spatially interpolated to a grid using a modified version of the Geographic Resources Analysis Support System or GRASS© (GRASS, 2002) GIS inverse-distance-weighting (IDW) algorithm to produce a *normalized residual grid*. The IDW method assumes the value at an unsampled point can be estimated as a weighted average of points within a certain distance or from a given number of m closest points; CRAB used the 12 closest points (i.e., m = 12). Weights are inversely proportional to the power of the distance in meters which at an unsampled point r = (x,y) is:

$$F(r) = \frac{\sum_{i=1}^{m} z(r_i) / |r - r_i|^p}{\sum_{j=1}^{m} 1 / |r - r_j|^p}$$ (E.8, Neteler and Mitasova, 2002)

where

F(r) = interpolated precipitation at unsampled grid cell
z = precipitation at sample point
m = 12
p = 2
$r_{i,j}$ = location of sample point
r = location of unsampled grid cell

The IDW was conducted in a geographic (i.e., latitude-longitude) projection with the distance between *r* and $r_{i,j}$ being computed in true distance (meters) units. IDW was used because by definition it is an exact interpolator and remained faithful to the *normalized residuals* at stations; this is important so that when the *normalized residuals* were converted back to *actual residuals* they were equal to the original *actual residual* at each station. Since there is a great deal of spatial autocorrelation of the *normalized residuals*, i.e. the *normalized residuals* tend to be spatially consistent within the regions, IDW was an adequate and appropriate interpolation scheme (see embedded map of normalized residuals in Figure 4.8.1).

Prior to the peer review, an internal review of preliminary 100-year 60-minute spatially-interpolated estimates showed several bull's eyes and contours closely drawn to stations reflecting a strong station influence rather than a reasonable climatological pattern. It was clear that the limited hourly dataset was not sufficient to accurately resolve important short duration patterns at a high spatial resolution (3-seconds). To help resolve this at the shorter (≤ 12-hour) durations, increased smoothing was applied to the *normalized residuals*. This increased smoothing was achieved by decreasing the spatial resolution from 3-seconds to 9-seconds before spatially interpolating the *normalized residuals* with the IDW algorithm. Sensitivity tests were conducted to determine the

optimum resolution to use to avoid "over-smoothing" the estimates which would cause the maps to deviate significantly from the quantile estimates achieved through the L-moment analysis. The results were then re-sampled to the 3-second resolution before the next step.

The *normalized residual grid* was then de-normalized by multiplying it by the original spatially interpolated mean annual maximum grid to obtain a spatially interpolated grid of *actual residuals* for the entire project area. Figure 4.8.3 shows the relationship between the 100-year 24-hour *actual residuals* and the 24-hour mean annual maximum estimates. Each linear cluster shown on this scatter plot represents stations within the same region that have varying 100-year 24-hour precipitation depths.

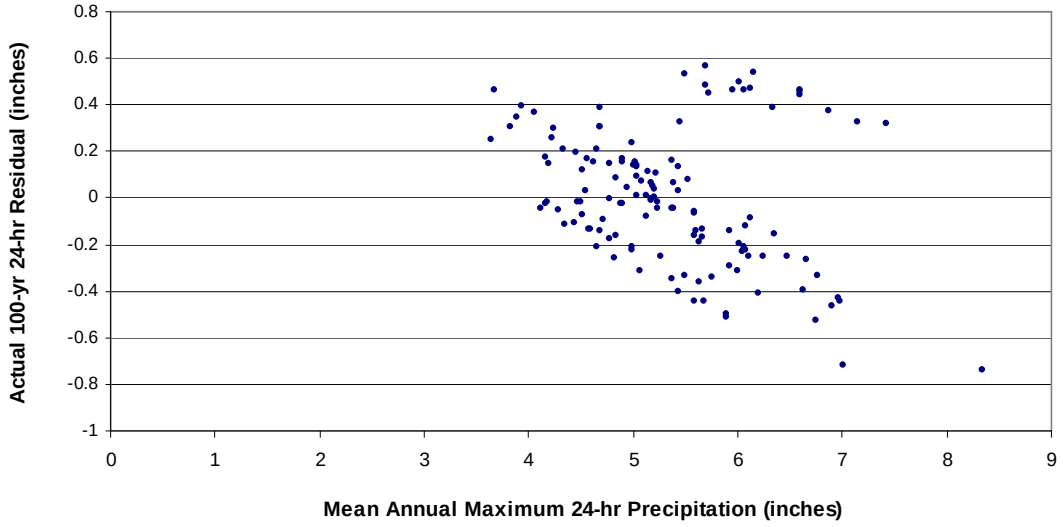


Figure 4.8.3. The relationship between the 100-year 24-hour *actual residuals* and the mean annual maximum precipitation from NOAA Atlas 14 Volume 3.

Step 4: Development of *pre-final grids*. The spatially interpolated grid of *actual residuals* was added to the *first guess grid* to create a spatially interpolated *pre-final* grid (e.g., 2-year 24-hour). To prevent error propagation potentially introduced in the internal consistency adjustment steps (described in Step 5), the *pre-final grid* was archived before being smoothed and became the *predictor grid* for the next precipitation frequency grid derivation. For example, the *pre-final* 2-year 24-hour grid was used as the predictor for the 5-year 24-hour grid rather than the *final* 2-year 24-hour grid to remain faithful to the data and allow patterns to develop without any differences that may be introduced by adjustments and filters.

To promote smooth contours and remove unnatural variability in the spatially distributed precipitation frequency estimates the *pre-final grid* was smoothed using a small (3x3 grid cell) center-weighted block filter, thus creating the *smoothed pre-final grid*.

Step 5: Internal consistency check. To ensure internal consistency in the *smoothed pre-final* grid cell values, duration-based and frequency-based internal consistency checks were conducted. Frequency-based internal consistency violations (e.g., 100-year < 50-year) were very rare and when they did exist, they were small violations relative to the precipitation frequency estimates involved. Duration-based internal consistency violations (e.g., 24-hour < 12-hour) were more common, particularly between 120-minute and 3-hour, but again were small violations relative to the magnitude of precipitation frequency estimates. To mitigate internal consistency violations, the longer duration

or rarer frequency grid cell value was adjusted by multiplying the shorter duration or lower frequency grid cell value by 1.01 to provide a 1% difference between the grid cells. One percent was chosen over a fixed factor to allow the difference to change according to the grid cell magnitudes while at the same time providing a minimal, but sufficient, adjustment without changing otherwise compliant data in the process. The duration-based check and adjustment was conducted first, resulting in a new pre-final grid, which was then subjected to the frequency-based check and adjustment. The resulting grid became the *final grid* for the particular frequency and duration (e.g., 2-year 24-hour).

Development of n-minute grids. Because of the small number of n-minute data available, durations shorter than 60-minute (i.e., n-minute precipitation frequency estimates) were calculated using linear scaling factors applied to *final grids* of spatially interpolated 60-minute precipitation frequency estimates. The linear scaling factors were developed using ratios of n-minute quantiles to 60-minute quantiles from co-located n-minute and hourly stations (Section 4.1.1, Data Properties, N-minute data). Table 4.8.2 show the n-minute ratios for all durations and annual exceedance probabilities used in Volume 3. Unlike NOAA Atlas 14 Volume 2, the ratios here are valid for all annual exceedance probabilities since there were so few data (only 16-33 years) at the 15-minute and n-minute stations. The appropriate ratios were multiplied by the same-frequency 60-minute grid to create the *final* n-minute precipitation frequency grids. These ratio grids were also used for both the n-minute upper- and lower- confidence limit grids.

Table 4.8.2. N-minute ratios for NOAA Atlas Volume 3: 5-, 10-, 15- and 30-minute to 60-minute.

Annual Exceedance Probability	5-min	10-min	15-min	30-min
All	0.240	0.328	0.421	0.674

Validation. Initial draft mean annual maximum, "index flood", grids for this Atlas, as well as the CRAB-derived 100-year 24-hour and 100-year 60-minute precipitation frequency grids were subjected to a peer-review (Appendix A.5). After considering and resolving all reviewer comments, final mean annual maximum grids were created by PRISM and the CRAB procedure re-run.

In addition, jackknife cross-validation allowed further, objective evaluation and validation of the precipitation frequency grids. The jackknife cross-validation exercise entailed running the CRAB procedure with a station in the dataset, storing the target grid cell value (at the station), then running CRAB without the station and comparing the target grid cell values. 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 CRAB procedure based on the same mean annual maximum grids. The comparison was used to test the robustness and accuracy of the CRAB interpolation using the 100-year 60-minute estimates since it required the most interpolation to ungauged locations because of the lower number of hourly stations. A perfect validation would result in equal values (0% difference) – with and without the station. 100-year 60-minute results (Figure 4.8.5) indicated that the CRAB process performed well. The primary message that Figure 4.8.5 conveys is the fact that, overall, CRAB did a good job reproducing the values in the absence of station data. The figure also indicates that there was a greater tendency for CRAB to slightly under-predict the precipitation frequency value at a location in a station's absence.

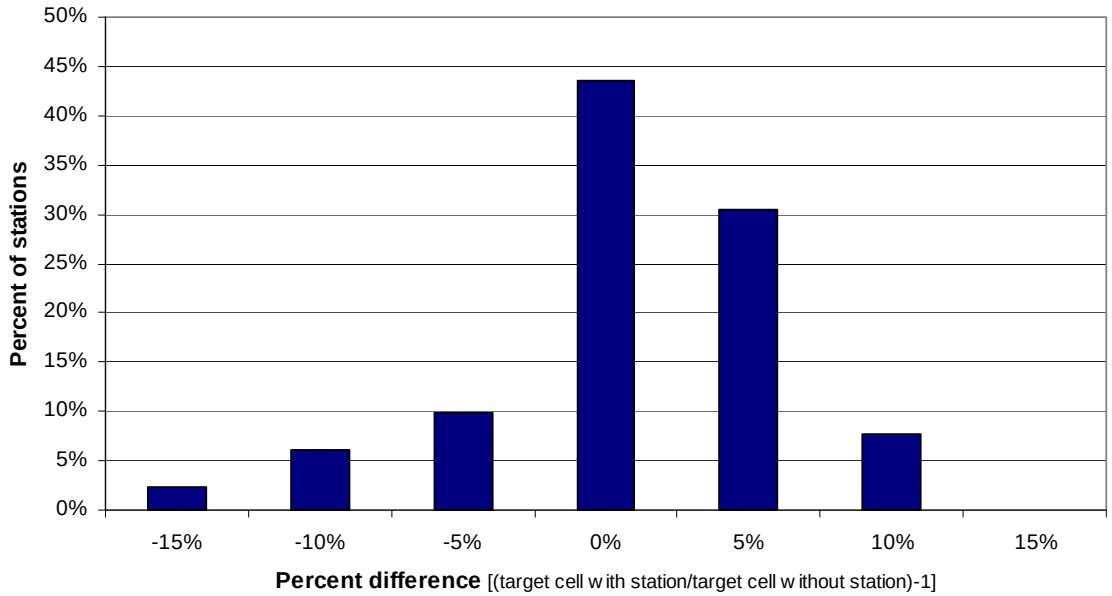


Figure 4.8.5. NOAA Atlas 14 Volume 3 100-year 60-minute jackknife cross-validation results.

Derivation of upper/lower limit precipitation frequency grids. The upper and lower limit precipitation frequency grids were also derived using the CRAB procedure. Testing suggested that the best method by which to derive the upper/lower limit grids was to use the preceding upper (or lower) grid as the *predictor grid* and *normalizing grid* for the upper/lower limit grid being derived, as opposed to using the corresponding mean precipitation frequency grid. Although the upper (lower) limit precipitation frequency estimates were slightly less stable than the mean grids, they still exhibited strong linear relationships with the previous (*predictor*) grid. The appropriate (i.e., same duration) mean annual maximum grid (PRISM-produced "index flood") was used as the initial *predictor grid* for the 2-year upper and lower limit precipitation frequency estimate grids. Figure 4.8.6 shows a scatter plot of the 24-hour mean values versus the 2-year 24-hour upper limit precipitation frequency estimates.

Similar to the precipitation frequency estimate grids, the upper and lower limit grids were evaluated and adjusted for internal consistency. Although very rare, duration-based adjustments were made to ensure the upper (lower) limit grid cell values were larger (smaller) than the mean values. In the event of a violation (e.g., 100-year 60-minute < 100-year 60-minute lower limit) the upper (lower) limit grid was adjusted up (down) by 1% of the mean grid. Like the precipitation grids, frequency-based or duration-based adjustments were made when needed. To mitigate any internal consistency violations, the longer duration or rarer frequency grid cell value was adjusted by multiplying the shorter duration or lower frequency grid cell value by 1.01 to provide a 1% difference between the grid cells.

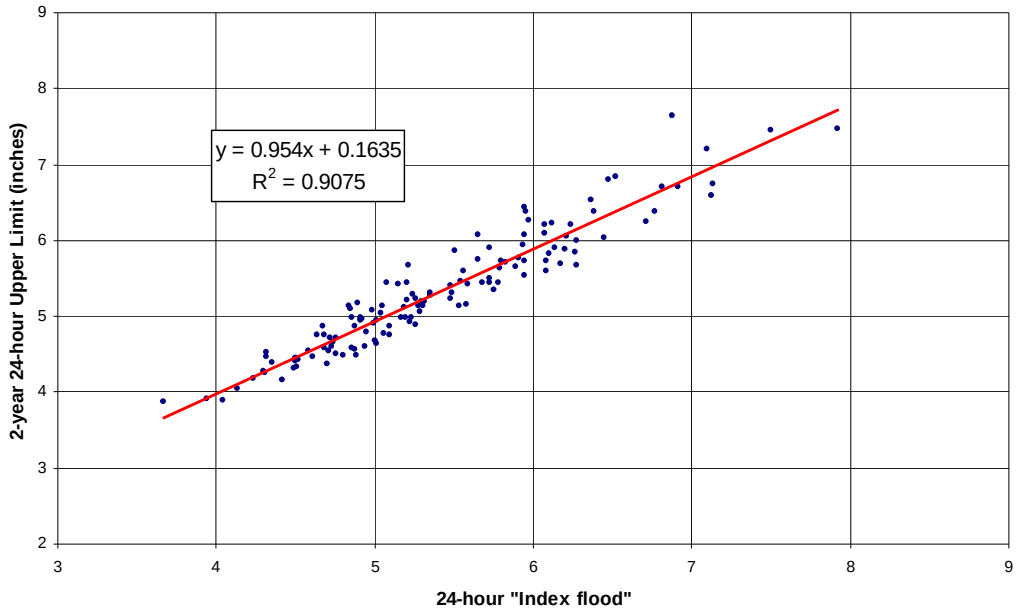


Figure 4.8.6. Scatter plot of the 24-hour mean precipitation frequency estimates vs. the 2-year 24-hour upper limit showing a coefficient of determination of 0.9075 in NOAA Atlas 14 Volume 3.

4.8.3. Pseudo data

The limited shorter (<24-hour) duration dataset was not sufficient to accurately resolve patterns at the final high spatial resolution (3-sec), therefore so called hourly "pseudo data" were objectively generated at all but one daily-only station to create a more coherent spatial pattern in the hourly durations. Daily-only station 66-2330 (Cerro Gordo Ciales) was excluded from the pseudo data because it was unnecessarily increasing estimates in that area. Table 4.8.4 lists the hourly pseudo stations generated for this Atlas, while Figure 4.1.2 shows their locations. This dramatically increased the shorter duration dataset by 108 stations (from 26 to 133 stations) thereby providing the station density necessary to accurately resolve important spatial patterns that would have otherwise been undetected. Adding such data reduces uncertainty in areas otherwise void of hourly data.

The creation of pseudo hourly precipitation frequency estimates was similar to the approach used to alleviate 12-hour to 24-hour inconsistencies at co-located stations (Section 4.6.3). The pseudo precipitation frequency estimates were generated by applying a ratio of x-hour estimates to 24-hour estimates that was spatially interpolated using GRASS©'s inverse-distance-weighting algorithm (GRASS, 2002), which is shown in Section 4.8.2, based on only co-located daily/hourly stations. The ratio at each co-located station was calculated using the daily station's 24-hour precipitation frequency estimate to its x-hour precipitation frequency estimate. The interpolated ratio was then applied to the daily-only 24-hour precipitation frequency estimates to generate the pseudo hourly data at that station location. The mitigation provided a smoother, more meteorologically-sound transition from hourly to daily precipitation frequency estimates.

Tests showed that creating pseudo hourly data for daily-only stations that did not exhibit a large difference from 12-hour to 24-hour resulted in nearly identical precipitation frequency estimates before and after the inclusion of pseudo data. Pseudo data were not added to stations that did not need it or at ungauged locations.

Since each duration was computed independently, it was possible, although rare, for inconsistencies from duration to duration at a given location to occur in the spatial interpolation. However, the use of pseudo hourly data helped to mitigate these by instilling a relationship between the daily quantiles and pseudo hourly quantiles before the spatial interpolation took place.

Table 4.8.4. 108 hourly pseudo stations used in the preparation of NOAA Atlas 14 Volume 3.

Station ID	Station Name	State
66-0040	ACEITUNA	PR
66-0053	ADJUNTAS 1 NW	PR
66-0152	AGUIRRE	PR
66-0158	AIBONITO 1 S	PR
66-0410	ARECIBO 3 ESE	PR
66-0426	ARECIBO OBSERVATORY	PR
66-0662	BARCELONETA 2	PR
66-0736	BARRANQUITAS	PR
66-0842	CANDELARIA TOA BAJA	PR
66-0948	BOCA	PR
66-1123	CABO ROJO	PR
66-1301	CAGUAS 2 ENE	PR
66-1309	CAGUAS 1 W	PR
66-1345	CALERO CAMP	PR
66-1590	CANOVANAS	PR
66-1623	CAONILLAS UTUADO	PR
66-1701	CARITE DAM	PR
66-1712	CARITE PLANT 1	PR
66-1845	CATANO	PR
66-2634	CIDRA 1 E	PR
66-2723	COAMO 2 SW	PR
66-2801	COLOSO	PR
66-2825	COMERIO FALLS PLANT 2	PR
66-3023	CORRAL VIEJO	PR
66-3145	CULEBRA ISLAND	PR
66-3409	DORADO 2 WNW	PR
66-3532	ENSENADA 1 W	PR
66-3871	GARZAS	PR
66-3904	GUAJATACA DAM	PR
66-4126	GUAYABAL	PR
66-4193	GUAYAMA 2 E	PR
66-4260	GUINEO RESERVOIR	PR
66-4271	GURABO	PR
66-4330	HACIENDA CONSTANZA	PR
66-4613	HUMACAO 2 SSE	PR
66-4677	INDIERA ALTA	PR
66-4702	ISABELA SUBSTATION	PR
66-4867	JAJOME ALTO	PR
66-4910	JAYUYA	PR
66-4976	JOSEFA	PR
66-5020	JUANA DIAZ CAMP	PR

66-5064	JUNCOS 1 SE	PR
66-5075	LA FE	PR
66-5097	LAJAS SUBSTATION	PR
66-5123	LA MUDA CAGUAS	PR
66-5175	LARES	PR
66-5474	LOS CANOS	PR
66-5693	MAGUEYES ISLAND	PR
66-5807	MANATI 2 E	PR
66-5906	MARICAO	PR
66-5911	MARICAO FISH HATCHERY	PR
66-6050	MAUNABO	PR
66-6073	MAYAGUEZ CITY	PR
66-6083	MAYAGUEZ AIRPORT	PR
66-6255	MONA ISLAND	PR
66-6258	MONA ISLAND 2	PR
66-6270	MONTE BELLO MANATI	PR
66-6361	MORA CAMP	PR
66-6390	MOROVIS 1 N	PR
66-6432	NAGUABO 6 W	PR
66-6805	PARAISO	PR
66-6904	PATILLAS DAM	PR
66-6983	PENUELAS 1 NE	PR
66-7295	PONCE CITY	PR
66-7348	POTALA	PR
66-7492	PUERTO REAL	PR
66-7843	QUEBRADILLAS	PR
66-8126	RINCON	PR
66-8144	RIO BLANCO LOWER	PR
66-8155	RIO BLANCO UPPER	PR
66-8245	RIO GRANDE EL VERDE	PR
66-8306	RIO PIEDRAS EXP STA	PR
66-8412	ROOSEVELT ROADS	PR
66-8536	SABANA GRANDE 2 ENE	PR
66-8634	ST JUST	PR
66-8745	SAN CRISTOBAL	PR
66-8757	SAN GERMAN 4 W	PR
66-8808	SAN JUAN CITY	PR
66-8815	SAN LORENZO 3 S	PR
66-8817	SAN LORENZO ESPINO	PR
66-8822	SAN LORENZO FARM 2 NW	PR
66-8940	SANTA ISABEL 2 ENE	PR
66-8955	SANTA RITA	PR
66-9421	TOA BAJA 1 SSW	PR
66-9432	TORO NEGRO FOREST	PR
66-9521	TRUJILLO ALTO 2 SSW	PR
66-9608	UTUADO	PR
66-9763	VIEQUES ISLAND	PR
66-9774	VILLALBA 1 E	PR

66-9860	YAUCO 1 NW	PR
66-9862	YAUCO 1 S	PR
67-0198	ALEX HAMILTON FLD FAA	VI
67-0240	ANNALY	VI
67-1348	CATHERINEBURG	VI
67-1740	CHRISTIANSTED FORT	VI
67-1790	CORAL BAY	VI
67-1810	COTTON VALLEY 2	VI
67-1980	CRUZ BAY	VI
67-2551	EAST END	VI
67-2560	EAST HILL	VI
67-2823	ESTATE FORT MYLNER	VI
67-2870	ESTATE THE SIGHT	VI
67-3150	FOUNTAIN	VI
67-3677	GRANARD	VI
67-3880	HAM BLUFF LIGHTHOUSE	VI
67-4900	MONTPELLIER	VI
67-8905	TRUMAN FIELD AIRPORT	VI
67-9450	WINTBERG	VI

4.8.4. Derivation of isohyetals of precipitation frequency estimates

Isohyetal (contour) GIS files were created from the grids of partial duration series based precipitation frequency estimates for users with geographical information systems (GISs). The isohyetals are provided as Environmental Systems Research Institute, Inc. line shapefiles (ESRI, 2003). The isohyetals were created by contouring the grid files with GRASS©'s r.contour command (GRASS, 2002). The resulting files were when exported as shapefiles with GRASS©'s v.out.shapefile command (GRASS, 2002). In order to keep the isohyets and grids consistent, no line generalization or smoothing was conducted. The precision, resolution and smoothness of the grids were sufficiently high to result in smooth contour lines.

The choice of contour intervals was determined by an algorithm which used the maximum, minimum and range of grid cell values. The number of individual contour intervals was constrained between 10 and 30; however, some of the n-minute grids did not exhibit the range necessary to meet the 10 interval threshold and therefore have fewer than 10. All of the intervals are evenly divisible by 0.10 inches – the finest interval. A script that computed the appropriate contour intervals and shapefiles also generated Federal Geographic Data Committee compliant metadata for the shapefiles and a "fact" file. The HTML-formatted fact file provides details of the shapefile and also includes a list of the contour intervals. To simplify the downloading of the isohyetal shapefiles from the Precipitation Frequency Data Server (PFDS), all of the shapefile components (*.shp, *.dbf, and *.shx, *.prj), metadata and fact file were compiled and compressed into a single archive file containing many files (*.tar). For projection, resolution and other details of the shapefiles, please refer to the metadata and/or fact file.

The isohyetal shapefiles were created to serve as visual aids and are not recommended for interpolating final point or area precipitation frequency estimates for design criteria. Users are urged to take advantage of the grids or the Precipitation Frequency Data Server user interface for accessing final estimates.

4.8.5. Creation of color cartographic maps

The isohyetal shapefiles were used to create color cartographic maps of partial duration series-based precipitation frequency grids. The maps were created using Environmental Systems Research Institute, ArcGIS© 9.1 software, in particular ArcMap© (ESRI, 2003). Although in appearance the cartographic maps look to be comprised of polygons, enclosed two-dimensional cells, they are not. Instead, color shading of the grids combined with the line shapefiles provides the clean look of polygons. The cartographic maps are provided in an Adobe Portable Document format (PDF) format for easy viewing and printing. The scale of the maps is 1:600,000 when printed in their native size, 17" x 22" (ANSI C), however the maps can be printed at any size. Users should be mindful that future maps and/or other projects may be in different scales or print sizes.

The color cartographic maps were created to serve as visual aids and, unlike Technical Paper 42 and 53, are not recommended for interpolating final point or area precipitation frequency estimates for design criteria. Users are urged to take advantage of the Precipitation Frequency Data Server graphical user interface for accessing estimates.

5. Precipitation Frequency Data Server

5.1. Introduction

NWS precipitation frequency estimates have traditionally been delivered in the form of Weather Bureau Technical Papers and Memoranda as well as NOAA Atlases. These are hard copy (i.e., paper) documents.

NOAA Atlas 14 precipitation frequency estimates are now delivered entirely in digital form rather than hard copy form in order to make the estimates more widely available and to provide the data in a broader and more accessible range of formats. The National Weather Service specifically developed the Precipitation Frequency Data Server (PFDS) as the primary web portal for precipitation frequency estimates and associated information (Parzybok and Yekta, 2003). The PFDS is an easy to use, point-and-click interface for official NOAA/NWS precipitation frequency estimates and intensities. It is based on work done for the Alabama Rainfall Atlas (Durrans and Brown, 2002). The PFDS can be found at http://hdsc.nws.noaa.gov/hdsc/pfds/.

5.2. Underlying data

The PFDS operates from a large set of ASCII grids of precipitation frequency estimates. There are a total of 540 grids for NOAA Atlas 14 Volume 3: 180 for the precipitation frequency estimates and 180 each for the lower and upper bounds of the 90% confidence limits of the estimates. Table 5.2.1 shows the complete table of average recurrence intervals (1-year to 1,000-year) and durations (5-minutes to 60-days) available from the PFDS for any particular location.

Table 5.2.1. Average recurrence intervals (ARI) (1-year to 1,000-year) and durations (5-minutes to 60-days) available from the PFDS for any particular location for estimated precipitation frequency estimates as well as upper (and lower) limit precipitation frequency estimates.

ARI / Duration	1-yr	2-yr	5-yr	10-yr	25-yr	50-yr	100-yr	200-yr	500-yr	1,000-yr
5-minute	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
10-minute	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
15-minute	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
30-minute	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
60-minute	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
120-minute	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
3-hour	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
6-hour	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
12-hour	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
24-hour	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
48-hour	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
4-day	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
7-day	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
10-day	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
20-day	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
30-day	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
45-day	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
60-day	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

The PFDS operates directly from ArcInfo© (ESRI, 2003) ASCII Grids. The same grids can be downloaded from the website and imported into a Geographical Information System (GIS). The

ASCII grids, which represent the official precipitation frequency estimates, have the following pertinent metadata:

- Resolution: 3-seconds (about 80 meters x 80 meters or 262 feet x 262 feet)
- Units: inches*1000 (integer)
- Projection: Geographic (longitude/latitude)
- Datum: WGS 1972
- No data value: -9999

The PFDS operates with conventional web-tools, including cgi-bin (Perl) scripts, JAVAScript and one C program. The main cgi-bin script is activated when a user selects a location, either by manually entering a longitude/latitude coordinates, selecting a station, or clicking a location on the state map. The cgi-bin script develops a comprehensive output web page on the fly.

5.3. Methods (point and areal)

Since the PFDS is not an Internet Map Server (IMS), a so-called "information" function had to be coded so that when provided a longitude/latitude coordinate, the PFDS could return the appropriate precipitation frequency estimates. "Getcell," a fast-running, C-compiled program, was written to accomplish this simple, yet crucial function. "Getcell" uses the header information provided in the ArcInfo© ASCII Grid and the supplied longitude-latitude coordinate to calculate the x and y location of the desired grid cell within the grid matrix. The output is the grid cell value.

The 3-seconds (about 80 meters x 80 meters or 262 feet x 262 feet) resolution of the underlying grids is more than adequate to provide accurate point estimates. Using the PFDS, a point can be selected by a number of ways:

- Clicking on the map
- Manually entering a longitude/latitude coordinate
- Selecting a station from a pull-down list
- Entering an area.

"Getcell" also has the capability of calculating the mean, maximum, and minimum of a set of grid cells (i.e., area) represented by a list of user-input longitude/latitude coordinates. This powerful capability makes it possible for the PFDS to provide areal precipitation frequency estimates. In order to obtain areal precipitation frequency estimates from the PFDS, the user must enter at least 3 and as many as 12 longitude/latitude coordinates that represent the perimeter of the area of interest. The PFDS then determines which grid cells fall within the area and calculates an arithmetic mean of the grid cells. The mean is then subjected to an areal-reduction factor to compute the areal precipitation frequency estimate for each frequency. At the time of this publication, there is a companion project to update previously developed areal reduction factors. Areal estimates will be available for:

- Areas of 10 to 500 square-miles, and
- Durations of 1-hour through 24-hours.

5.4. Output

After the web server has successfully extracted all 540 precipitation frequency and confidence limit estimates from the underlying grids, an output web page is built and displayed on-the-fly. There are two basic types of output: Depth-Duration-Frequency (DDF) and the Intensity-Duration-Frequency (IDF). Both outputs are based on the same data, but presented differently. The PFDS provides DDF graphs in two formats to provide a complete perspective of the data (Figure 5.4.1, 5.4.2). An example of the classic IDF graph, which is widely used in engineering applications, is shown in Figure 5.4.3.

The output pages also consist of data tables of the precipitation frequency depths (or intensities) and tables of the lower and upper bounds of the 90% confidence limits. These can also be

downloaded as text via a button on the output page. In addition, location maps and helpful links are provided. Embedded maps on the output page are provided by a hyperlink to the U.S. Census Bureau Mapping and Cartographic Resources Tiger Map Server (http://tiger.census.gov/cgi-bin/mapbrowse-tbl). The graphs (portable network graphics or ".png" format) are produced using gnuplot (http://www.gnuplot.info), while the remainder of the page is basic HTML.

Additionally, seasonal exceedance graphs are provided via a button on the output page. Exceedance graphs indicate the percentage of events exceeding the corresponding annual exceedance probability for the specified duration (Appendix A.2). The purpose of the graphs is to portray the monthly seasonality of extreme precipitation events. See Figure 5.4.4 for an example. The percentages are based on regional statistics and the seasonal graphs are unique for each region. The number of stations and cumulative years of record are provided in the graph title to provide the user a sense of the amount of data used and therefore the reliability of the results. Durations include:

- 60-minute
- 24-hour
- 48-hour
- 10-day

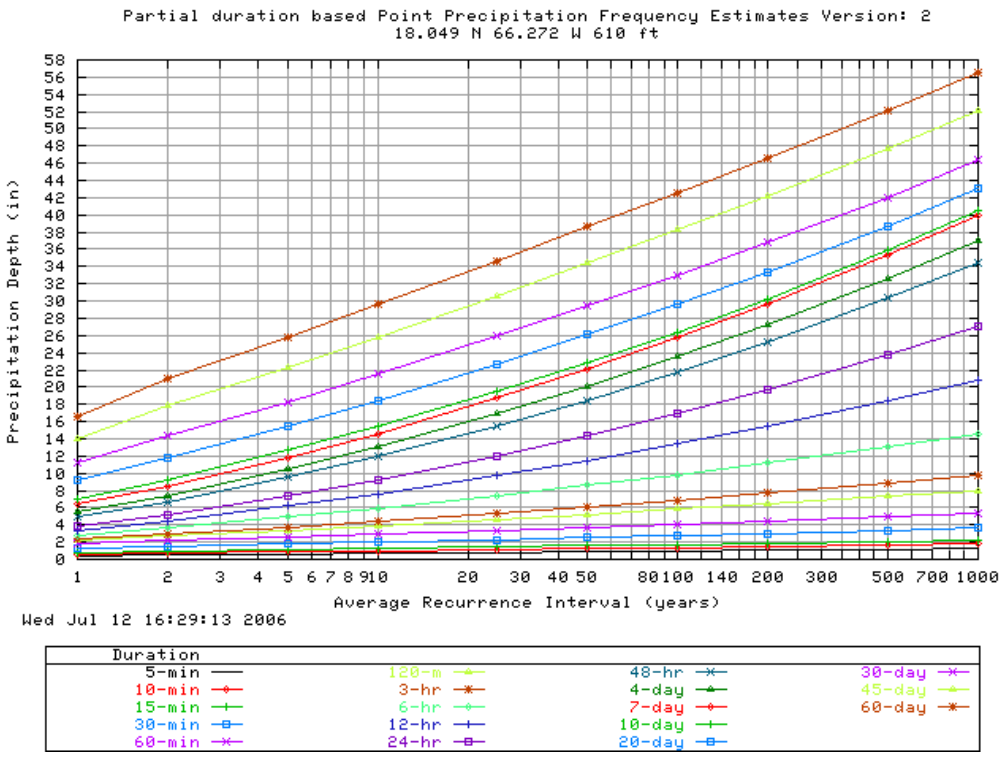


Figure 5.4.1. Sample depth-duration frequency plot with average recurrence interval on the x-axis.

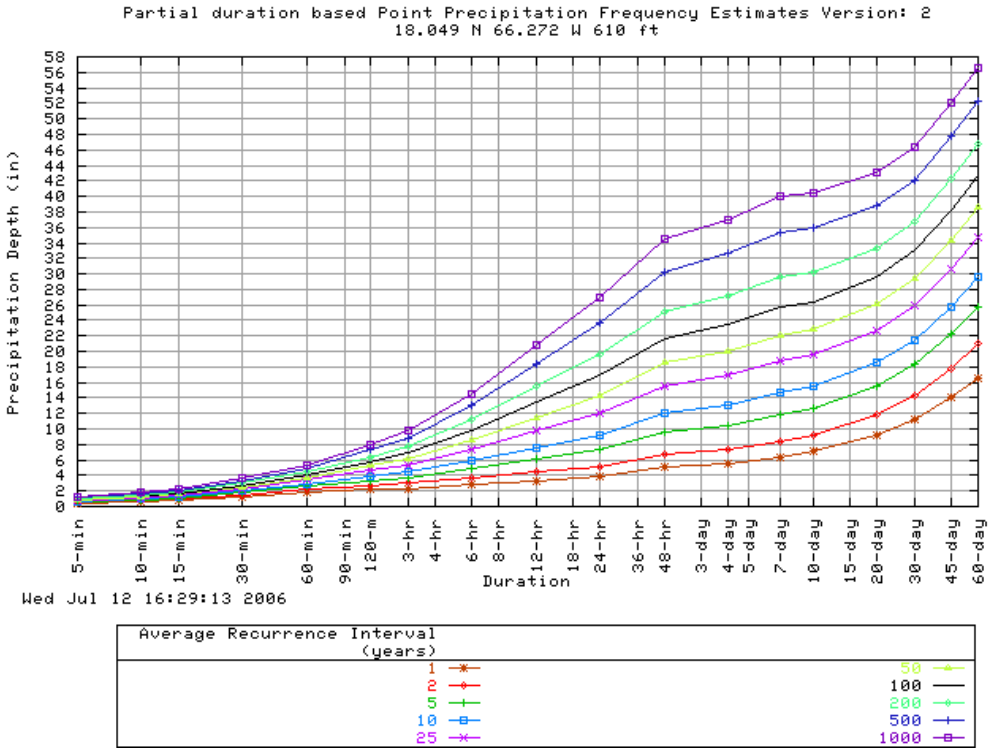


Figure 5.4.2. Sample depth-duration frequency plot with duration on the x-axis.

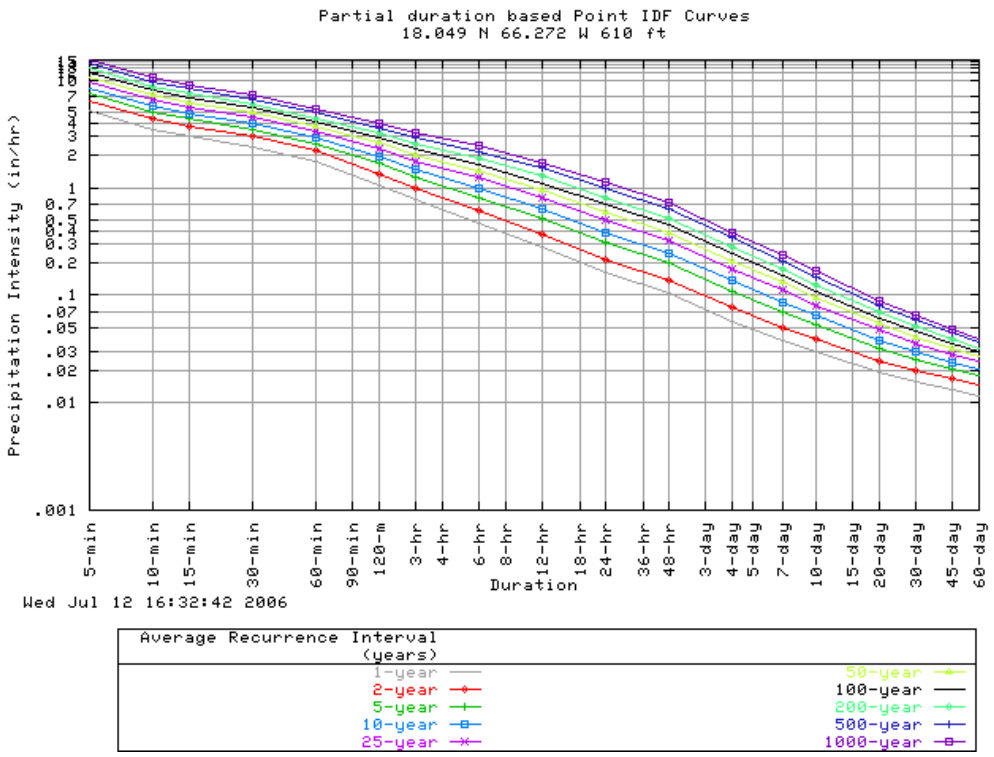


Figure 5.4.3. Sample intensity-duration-frequency (IDF) plot.

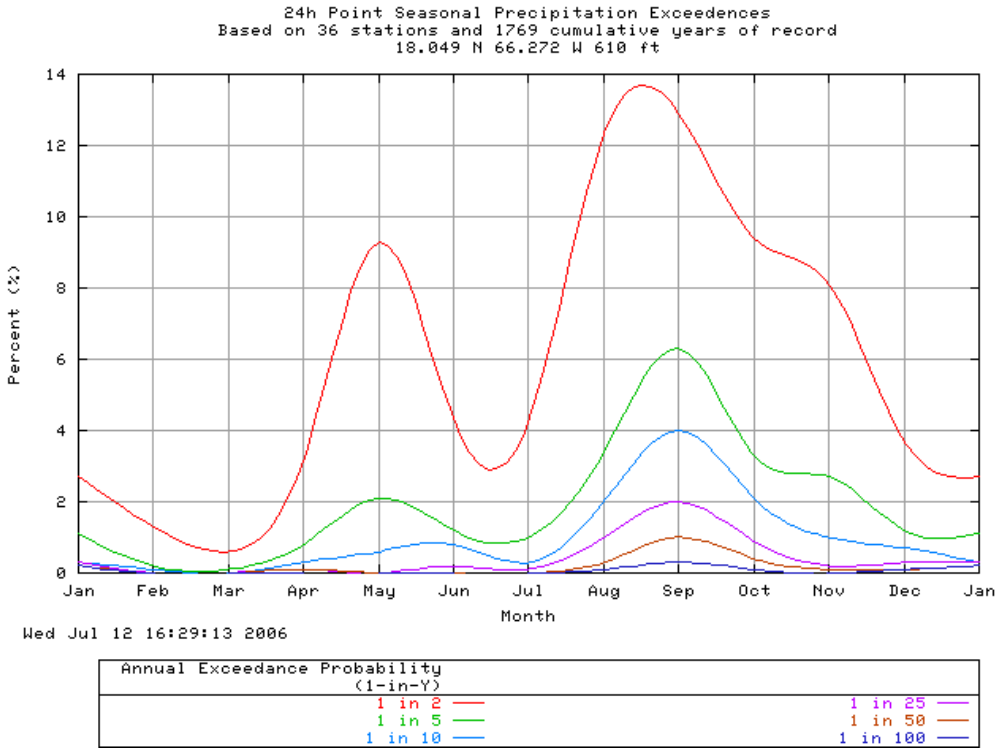


Figure 5.4.4. Sample 24-hour seasonal exceedance graph.

5.5. Using the Precipitation Frequency Data Server

The PFDS homepage (http://hdsc.nws.noaa.gov/hdsc/pfds/) has a clickable map of the United States. States with available precipitation frequency updates are indicated. Upon clicking on a state, a state-specific web page appears. From this page the user selects the desired location, units, and output via a web form.

The PFDS is also the portal for all NOAA Atlas 14 data formats, including:

- **Cartographic maps**
  These color maps were created to serve as visual aids and are not recommended for interpolating final point or area precipitation frequency estimates for design criteria. It is strongly recommended that point and areal values be obtained from the PFDS interface which accesses its data directly from the grids. These maps were based on contour lines (available as shapefiles from the PFDS) created from the final precipitation frequency estimate grids (Section 4.8.5). Figure 5.5.1 shows an excerpt from a cartographic map.

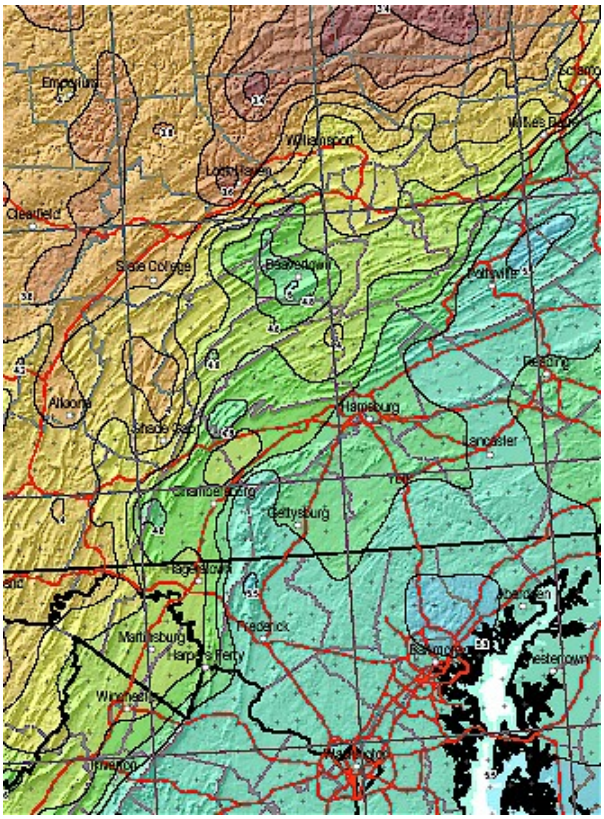


Figure 5.5.1. An excerpt of a cartographic map available from the PFDS.

- **GIS data**
  - Shapefiles (lines, vectors)
    These are the same files used to create the cartographic maps and are recommended for visual aids only.

  - ArcInfo© ASCII grids
    These grids represent the highest resolution precipitation frequency estimates from which all other formats are derived.

- **Time Series (annual maximum and partial duration series)**
  The final high quality annual maximum series (AMS) and partial duration series (PDS) datasets used in the preparation of NOAA Atlas 14 Volume 3 are available at http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_series.html. Information regarding their extraction can be found in Section 4.1.3.

- **Temporal distributions of heavy rainfall**
  This report is available via a link (http://hdsc.nws.noaa.gov/hdsc/pfds/docs/NA14Vol3_A1.pdf) on the PFDS web site and in Appendix A.1 of this document. The report provides information about the temporal distribution of heavy precipitation for use with precipitation frequency estimates in NOAA Atlas 14 Volume 3. Temporal distributions for 1-, 6-, 12-, 24- and 96-hour durations are available.

- **Documentation**
  The complete NOAA Atlas 14 Volume 3 documentation is available via links on the PFDS web site.

It is strongly advised that users review the Federal Geographic Data Committee (FGDC) compliant metadata before using any of the GIS datasets. On-line help and frequently-asked questions (FAQ) are also available via links on the PFDS web site.

Questions regarding the use of the PFDS or its data can be addressed by emailing HDSC.Questions@noaa.gov or visiting the inquiry web site, http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_contact.html.

6. Peer Review

A peer review was conducted for the preliminary point precipitation frequency estimates and preliminary spatially interpolated estimates. Nearly 120 users, project sponsors and other interested parties were contacted via email for the review, which occurred from November 3, 2005 through January 11, 2006. The reviews provided critical feedback that HDSC used to create a better product.

The point precipitation frequency estimates and spatial distribution, which focused on a subset of maps, were reviewed during a 10 week period. For review purposes, draft 60-minute and 24-hour mean annual maximum grids were produced using PRISM. CRAB was then used to derive 100-year 60-minute and 100-year 24-hour grids from the PRISM grids. Both sets of grids were converted into cartographic maps in a PDF format for review. Also included in the review were the temporal distributions of heavy rainfall (Appendix A.1) and the statistical trend analysis of 1-day annual maximum series (Appendix A.3).

HDSC received responses from 14 individuals or groups that were divided into 47 unique comments. Similar reviewer issues/comments were grouped together and addressed in a single HDSC response. Reviewer comments and HDSC responses can be found in Appendix A.5. Reviewers were asked to address comparisons to current design thresholds, cartographic elements, reasonableness of estimates and patterns when compared to local or regional knowledge, station locations, confidence limits, and potentially bad data. Further investigation and modification occurred subsequent to the initial HDSC responses.

Reviewer comments regarding data quality and expected spatial patterns generated further verification and/or modification of various geographic areas, such as the deletion of station 66-1142 (Cacaos-Orocovis). Further consideration of reviewer concerns was made during the PRISM process (Appendix A.4) to produce more spatially, temporally and climatologically consistent results, such as adjusting the “hill” northeast of Guayabal, PR (66-4162), which was previously and erroneously an area of maximum precipitation, to better reflect the expected precipitation regime in the area.

7. Interpretation

Point and areal estimates. The precipitation frequency estimates in this Atlas are point estimates, that is, estimates of precipitation frequency at a point location, not for an area. The conversion of point to areal estimates must take into account that, all other things being equal, as the area increases, the intensity decreases. This is done by applying an areal reduction factor (ARF) to the point estimates that are provided in this Atlas. Precipitation frequency estimates for areas can be computed by obtaining an average of the point values at all locations within the subject area and then multiplying that average by the appropriate areal reduction factor. Areal reduction factors have been published in previous publications: Technical Report 24 (Meyers and Zehr, 1980), Technical Memorandum HYDRO-40 (Zehr and Meyers, 1984), NOAA Atlas 2 (Miller et al., 1973), etc. At the time of this publication there is a companion project to update previously developed areal reduction factors.

Independence. Precipitation is highly variable both spatially and temporally, however within any particular storm event, point observations have a degree of correlation. The methods used to develop the point precipitation frequency estimates for this Atlas assume independence between the annual maxima analyzed and so the individual estimates in this Atlas express independent, point probabilities. That a point within a particular watershed may receive an amount equal to or greater than its 1 in 50 or 1 in 100 values at a particular time does not affect probabilities for any other point within that watershed.

Annual Exceedance Probability (AEP) and Average Recurrence Interval (ARI). As discussed in Section 3.2 and throughout this document, AEP is the probability that a particular level of rainfall will be exceeded in any particular year (at a particular location and duration) and is derived using the annual maximum series. An AEP depth or intensity may be exceeded once or more than once in a year. ARI is the average period between each exceedance and is derived for the partial duration series. As a result, the inverse of AEP is not ARI as is commonly assumed. Rather, the inverse of AEP is the average period between years with exceedances (Laurenson, 1987). One can convert between annual maximum and partial duration series results by using the ratio between partial duration and annual maximum results discussed in Section 4.6.3. This ratio approaches 1.0 for ARIs greater than about 25 years and so becomes significant only for values with ARIs less than about 25 years.

Exceedances. A certain number of exceedances can be statistically expected at a given station. For example, a rainfall with an AEP of 1 in 100 has a 1% chance of being exceeded approximately once in any given year at a particular station. When considering multiple stations that are sufficiently far apart to satisfy independence, the chance of observing such an event is directly proportional to the number of stations. For example, in the case of the 1 in 100 rainfall one can expect to observe approximately 10 such events each year in a network of 1,000 independent observing stations.

Use of confidence limits. Confidence limits provide users with an estimate of the uncertainty or potential error associated the precipitation frequency estimates. The error bounds about the precipitation frequency estimates and the probabilistic temporal distributions (Appendix A.1) enable designers to include estimations of error in the calculations by using Monte Carlo based ensemble modeling to estimate flow, rather than just applying a single value estimate.

Spatially interpolated confidence limits are provided with this Atlas. They were derived using the CRAB spatial derivation procedure (Section 4.8.2). The confidence limits are a function only of the error associated with the point precipitation frequency estimation and do not include error that may be associated with the spatial interpolation process.

Climate change. The current practice of precipitation (and river height and flow) frequency analysis makes the implicit assumption that past is prologue for the future. Rainfall frequency distribution characteristics are extracted from the historical record and the estimates are applied in the design of future projects assuming the climate will remain the same as it was during the period of the analyzed record. If the climate changed in the past, then the characteristics extracted are an "average" for the analyzed period, not specifically representing the period before the change or after the change. Furthermore, if the climate changes in the future, there is no guarantee that the characteristics extracted are suitable for representing climate during the future lifecycle of projects being designed. There has been considerable research done regarding climate change and precipitation. NOAA's National Weather Service conducted an analysis of shifts and trends in the NOAA Atlas 14 Volume 3 1-day annual maximum series data (Appendix A.3). Results suggested little consistent observable effects of climate change on the annual maximum series and therefore on parameters used for this Atlas. As such, NOAA's National Weather Service has assumed that the full period of the available historical record derived from rain gauges was suitable for use in this analysis even though there were some local instances of linear trends and shifts in mean in the data.

Comparison with Technical Paper 42. In general, reasons for differences between the NOAA Atlas 14 precipitation frequency estimates and Technical Paper 42 estimates include longer records of data, more stations and greater effectiveness of new statistical procedures, including an objective spatial analysis. Figure 7.1 and 7.2 show the percent differences between NOAA Atlas 14 and Technical Paper 42 for 100-year 24-hour and 100-year 60-minute estimates, respectively.

Differences between NOAA Atlas 14 Volume 3 and Technical Paper 42 results have been carefully considered. Areas of large differences were considered and found justified by the increased data availability, sound regionalization, more robust statistical procedures used in the current analysis and more advanced spatial interpolation process including higher resolution. "Differences" in this context refers to differences in the mean of the estimates. Because NOAA Atlas 14 is the first NWS publication to include confidence limits, a comparison of the confidence limits with previous publications was not possible. It should be noted from the width of the confidence limits that the errors associated with the estimates are not insignificant. It should also be noted that the confidence limits associated with NOAA Atlas 14 estimates are likely much narrower than in previous publications because of improvements in estimating techniques.

Some 1,000-year estimates exceed Probable Maximum Precipitation (PMP) values published in Technical Paper 42 published in 1961. All cases where 1,000-year 24-hour estimates exceeded PMP were carefully reviewed and we have much greater confidence in the NOAA Atlas 14 precipitation frequency estimates than we have in the Technical Paper No. 42 PMP estimates. The PMP estimates in Technical Paper No. 42 are likely to be low and an update to those estimates should be undertaken. Technical Paper No. 42 itself notes that its estimates of PMP have a higher level of uncertainty, particularly with respect to hurricane data, because of the limited record of data available when the study was being undertaken.

Estimates were peer reviewed and careful consideration was given to reviewer comments. Often the analysis was modified to accommodate reviewer suggestions or additionally provided data. Appendix A.5 provides reviewer comments and NWS initial responses to those comments. Further investigation was conducted subsequent to the initial responses to satisfactorily resolve reviewer concerns.

Figure 7.1. Differences between NOAA Atlas 14 Volume 3 and Technical Paper 42 100-year 24-hour estimates.

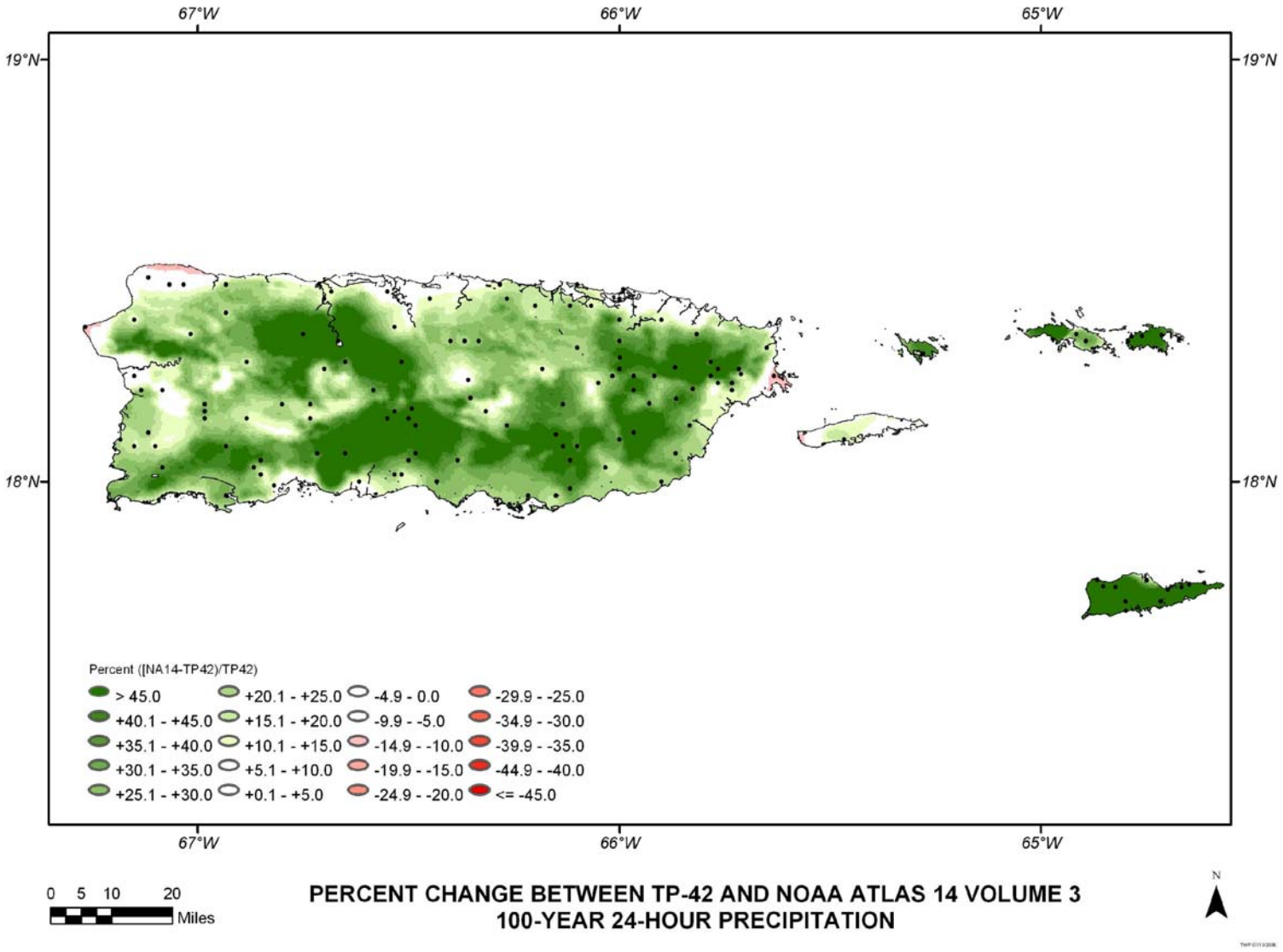
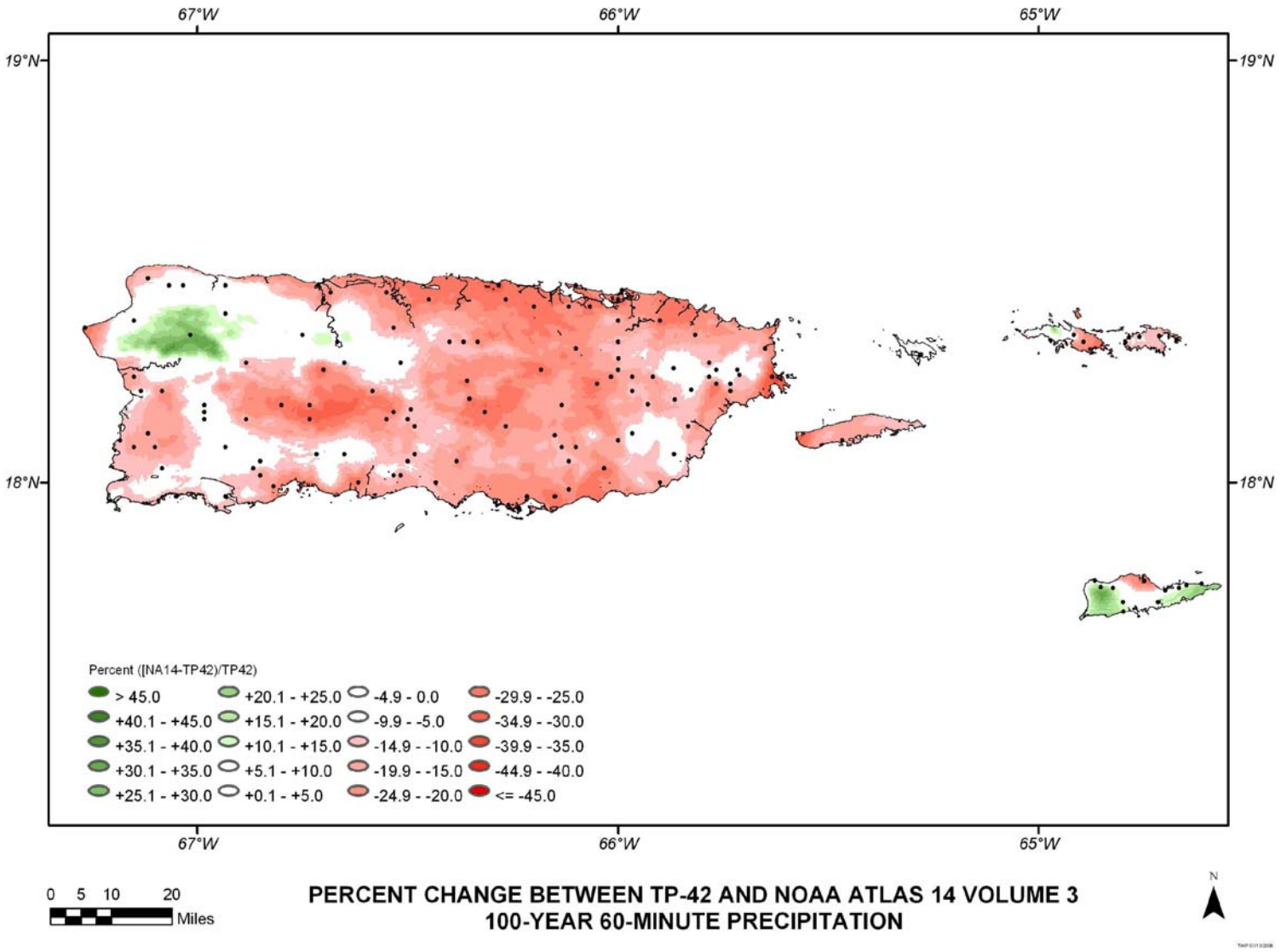


Figure 7.2. Differences between NOAA Atlas 14 Volume 3 and Technical Paper 42 100-year 60-minute estimates.



Appendix A.1. Temporal distributions of heavy precipitation associated with NOAA Atlas 14 Volume 3

1. Introduction

Temporal distributions of heavy precipitation are provided for use with precipitation frequency estimates from NOAA Atlas 14 Volume 3 for 1-, 6-, 12-, 24- and 96-hour durations covering Puerto Rico and the U.S. Virgin Islands. The temporal distributions are expressed in probabilistic terms as cumulative percentages of precipitation and duration at various percentiles. The starting time of precipitation accumulation was defined in the same fashion as it was for precipitation frequency estimates for consistency.

Temporal distributions for each duration are presented in Figure A.1.1. The data were also subdivided into quartiles based on where in the distribution the most precipitation occurred in order to provide more specific information on the varying distributions that were observed. Figures A.1.2 through A.1.6 depict temporal distributions for each quartile for the five durations. Digital data to generate the temporal distributions are available at http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_temporal.html. Table A.1.1 lists the number and proportion of cases in each quartile for each duration.

2. Methodology

This project largely followed the methodology used by the Illinois State Water Survey (Huff, 1990) except in the definition of the precipitation accumulation. This project computed precipitation accumulations for specific (1-, 6-, 12-, 24- and 96-hour) time periods as opposed to single events or storms in order to be consistent with the way duration was defined in the associated precipitation frequency project. As a result, the accumulation cases may contain parts of one, or more than one precipitation event. Accumulation computations were made moving from earlier to later in time resulting in an expected bias towards front loaded distributions when compared with distributions for single storm events.

For every precipitation observing station in the project area that recorded precipitation at least once an hour, the three largest precipitation accumulations were selected for each month in the entire period of record for the 12-, 24- and 96-hour durations. For the 1- and 6-hour durations, 15-minute data were accumulated in the same manner. Therefore, the 1-hour distribution contains only four data points while the 6-hour distribution contains twenty-four points. A minimum threshold was applied to make sure only heavier precipitation cases were being captured. Precipitation with an average recurrence interval (ARI) of 2 years at each observing station for each duration was used as the minimum threshold at that station.

A minimum threshold of 25-year ARI was tested. It was found to produce results similar to using a 2-year ARI minimum threshold. The 25-year ARI threshold was rejected because it reduced the number of samples sufficiently to cause concern for the stability of the distributions.

To determine whether distributions varied appreciably across the project area, temporal distributions based on data from each hourly region were computed separately, and compared to the distributions computed for the project area as a whole. The distributions were nearly identical, although there was more noise in the distributions from the separate regions due to smaller sample size. As a result the temporal distributions presented here were based on the entire project area because of the larger sample size and because the distributions varied so little by region.

Each of the accumulations was converted into a ratio of the cumulative hourly (or 15-minute) precipitation to the total precipitation for that duration, and a ratio of the cumulative time to the total time. Thus, the last value of the summation ratios always had a value of 100%. The data were combined, cumulative deciles of precipitation were computed at each time step, and then results were plotted to provide the graphs presented in Figure A.1.1. The data were also separated into categories by the quartile in which the greatest percentage of the total precipitation occurred and the procedure

was repeated for each quartile category to produce the graphs shown in Figures A.1.2 through A.1.6. A moving window weighted average smoothing technique was performed on each curve.

Consideration was given to a temporal distribution in which a majority of the rain fell symmetrically about the center hours of the distribution. However, after dividing the distributions into smaller time segments and reviewing the distributions in more temporal detail, most cases did not occur around the center regardless of duration or average recurrence interval. Therefore, we have not produced a center-loaded time distribution.

3. Interpreting the Results

Figure A.1.1 presents cumulative probability plots of temporal distributions for the 1-, 6-, 12-, 24- and 96-hour durations for the project area. Figures A.1.2 through A.1.6 present the same information but for categories based on the quartile of most precipitation. The x-axis is the cumulative percentage of the time period. The y-axis is the cumulative percentage of total precipitation.

The data on the graph represent the average of many events illustrating the cumulative probability of occurrence at 10% increments. For example, the 30% of cases in which precipitation is concentrated closest to the beginning of the time period will have distributions that fall above and to the left of the 30% curve. At the other end of the spectrum, only 10% of cases are likely to have a temporal distribution falling to the right and below the 90% curve. In these latter cases the bulk of the precipitation falls toward the end of the time period. The 50% curve represents the median temporal distribution on each graph.

First-quartile graphs consist of cases where the greatest percentage of the total precipitation fell during the first quarter of the time period, i.e., the first 90 minutes of a 6-hour period, the first 3 hours of a 12-hour period, etc. The second, third and fourth quartile plots, similarly are for cases where the most precipitation fell in the second, third or fourth quarter of the time period.

The time distributions consistently show a greater spread, and therefore greater variation, between the 10% and 90% probabilities as the duration increases. Longer durations are more likely to have captured more than one event separated by drier periods; however, this has not been objectively tested as the cause of the greater variation at longer durations.

The following is an example of how to interpret the results using Figure A.1.5a and Table A.1.1. Of the 392 cases of the 24-hour duration, 126 of them were first-quartile events:

- In 10% of these cases, 50% of the total rainfall (y-axis) fell in the first 1.8 hours of event time (7.5% on the x-axis). By the 9th hour (37% on the x-axis), all of the precipitation (100% on the y-axis) had fallen and it was dry for the rest of the 24-hour period.
- A median case of this type will drop half of its total rain (50% on the y-axis) in 5.3 hours (22% on the x-axis).
- In 90 percent of these cases, 50% of the total precipitation fell by 9.6 hours (40% on the x-axis).

4. Application of Results

Care should be taken in the use of these data. The data are presented in order to show the range of possibilities and to show that the range can be broad. The data should be used in a way that reflects the goals of the user. For example while all cases represented in the data will preserve volume, there will be a broad range of peak flow that could be computed. In those instances where peak flow is a critical design criterion, users should consider temporal distributions likely to produce higher peaks rather than the 50^{th} percentile or median cases, for example. In addition, users should consider whether using results from one of the quartiles rather than from the "all cases" sample might achieve more appropriate results for their situation.

5. Summary and General Findings

The results presented here can be used for determining temporal distributions of heavy precipitation at particular durations and at particular levels of probability. The results are designed for use with precipitation frequency estimates and may not be the same as the temporal distributions of single storms or single precipitation events. The time distributions show a greater spread between the percentiles with increasing duration. At the 6-, 12-, and 24-hour durations a majority of the cases analyzed were first quartile with fewer cases falling into each subsequent quartile category. At the 96-hour duration, however, the number of cases was nearly evenly distributed between all four quartile categories. The majority of the cases by far at the 1-hour duration were second quartile.

Table A.1.1. Numbers and proportion of cases in each quartile for each duration and temporal distribution associated with NOAA Atlas 14 Volume 3.

	1st Quartile	**2nd Quartile**	**3rd Quartile**	**4th Quartile**	**Total number of cases**
1-hour	81 (18%)	186 (42%)	123 (29%)	51 (11%)	441
6-hour	159 (40%)	107 (27%)	85 (21%)	49 (12%)	400
12-hour	138 (34%)	132 (33%)	79 (20%)	54 (13%)	403
24-hour	126 (32%)	106 (27%)	85 (22%)	75 (19%)	392
96-hour	112 (27%)	92 (22%)	108 (25%)	109 (26%)	421

FIGURE A.1.1
TEMPORAL DISTRIBUTION: ALL CASES

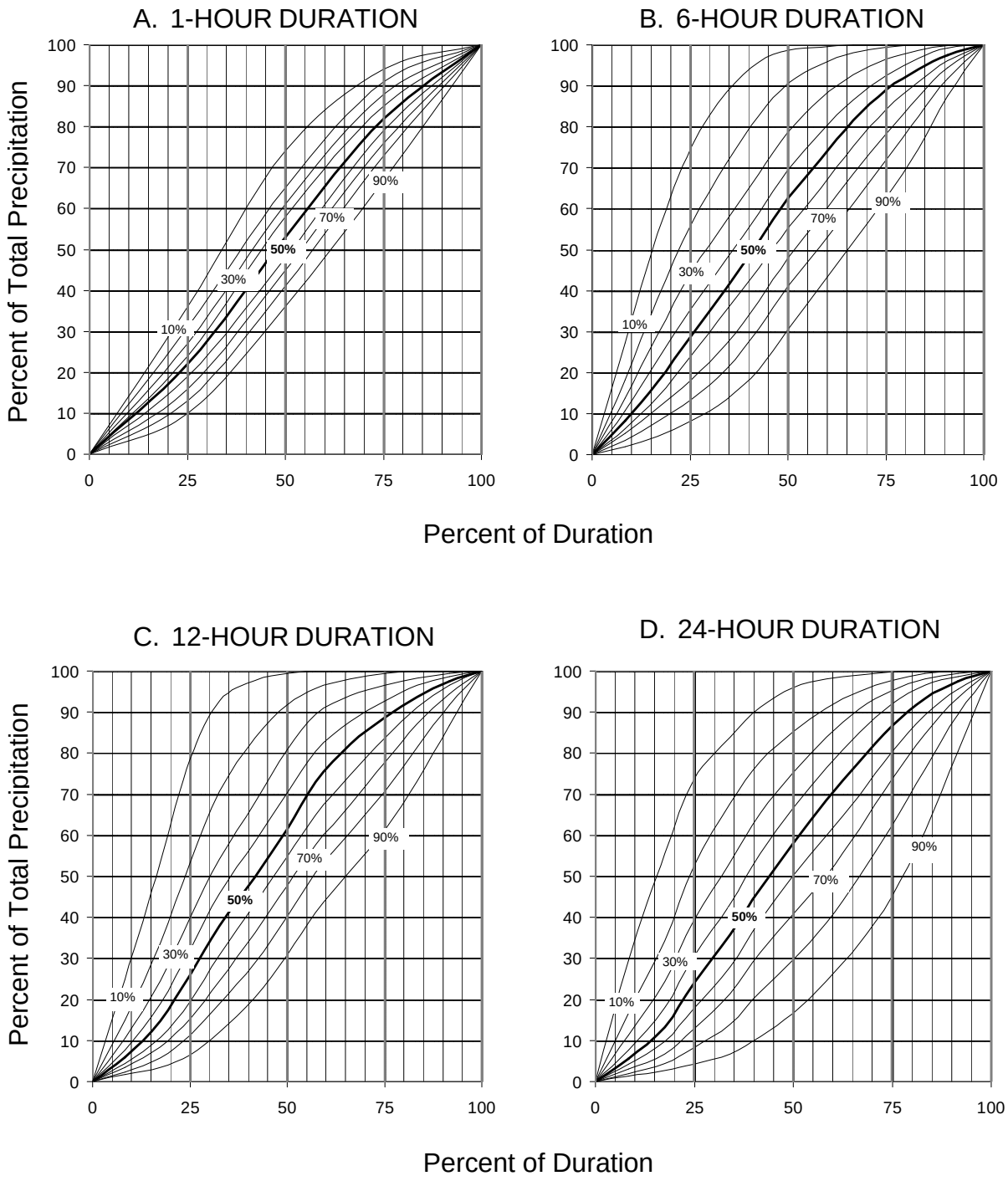


FIGURE A.1.1 (CONTINUED)
TEMPORAL DISTRIBUTION: ALL CASES

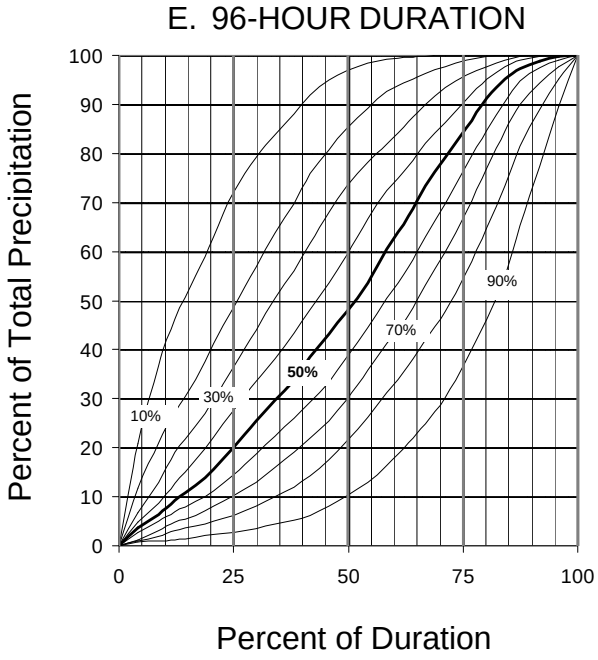


FIGURE A.1.2
TEMPORAL DISTRIBUTION: 1-HOUR DURATION

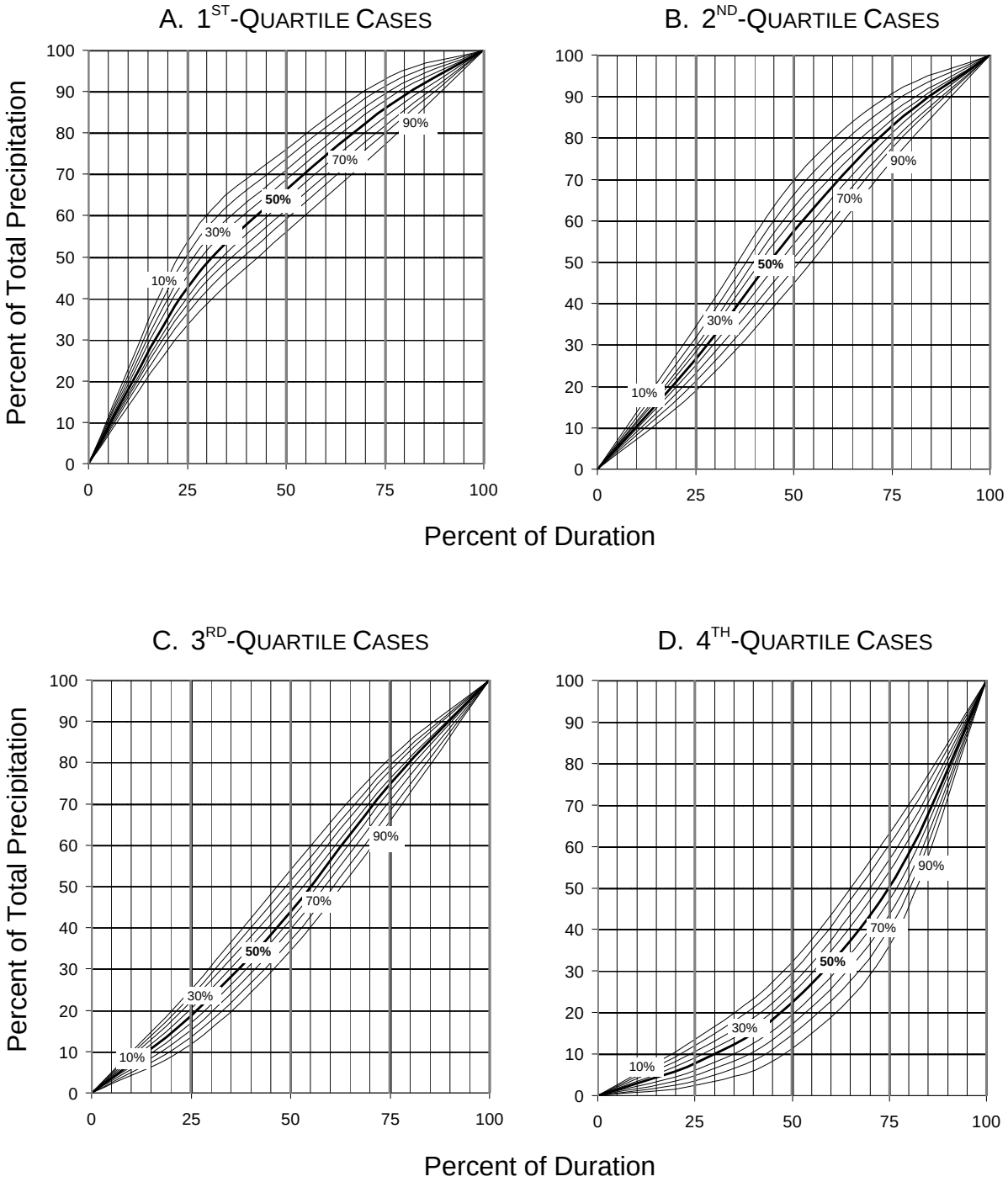


FIGURE A.1.3
TEMPORAL DISTRIBUTION: 6-HOUR DURATION

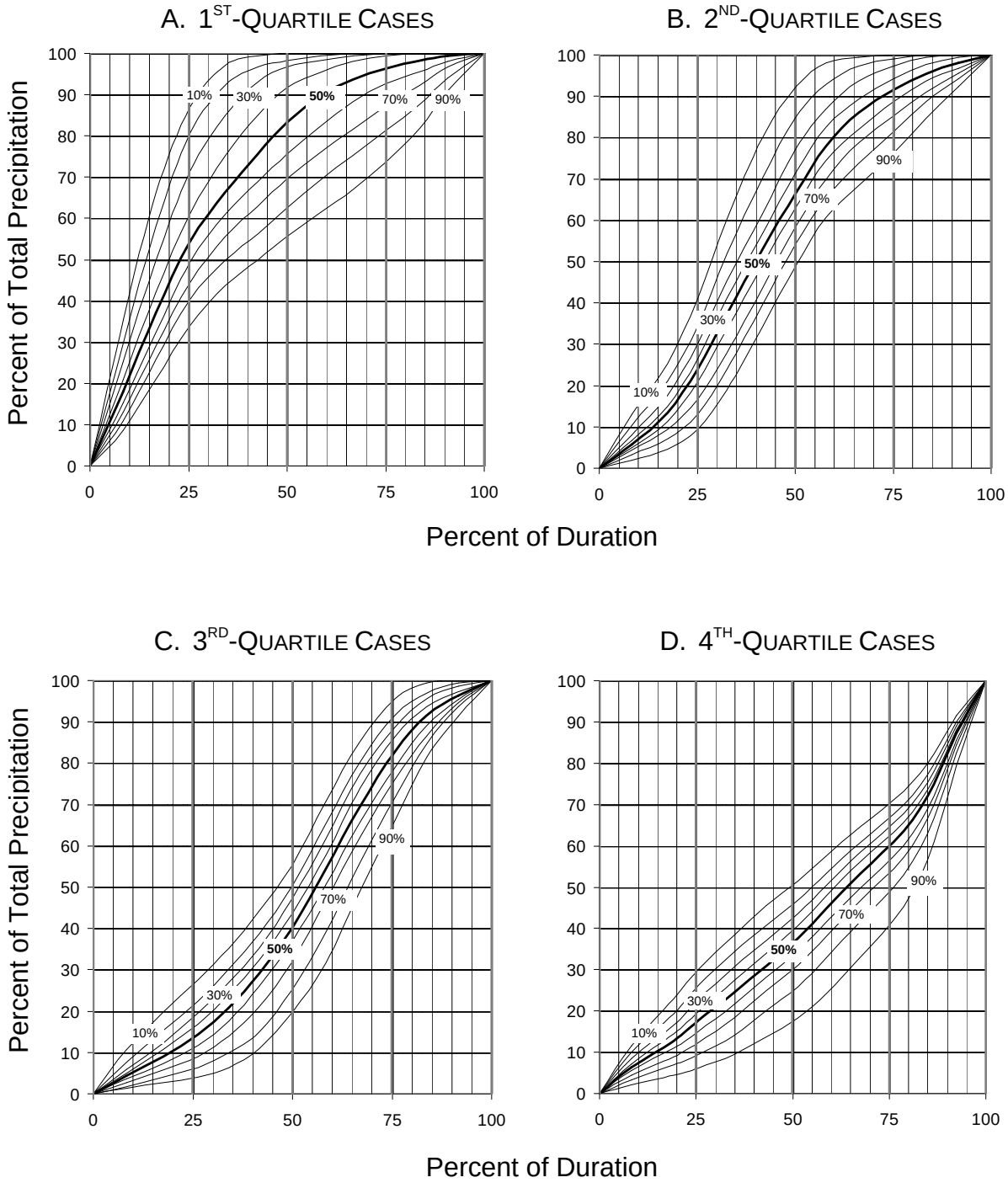


FIGURE A.1.4
TEMPORAL DISTRIBUTION: 12-HOUR DURATION

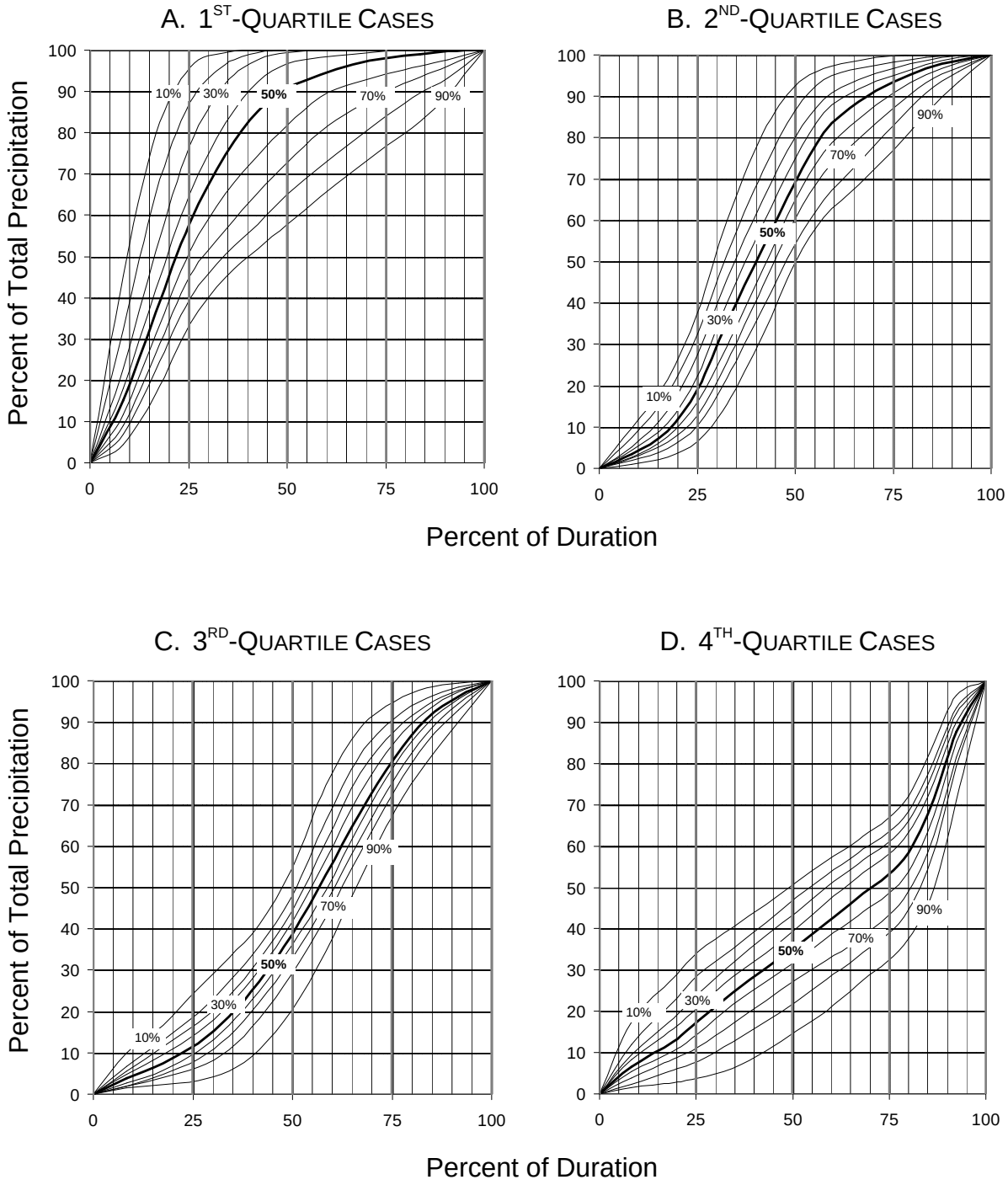


FIGURE A.1.5
TEMPORAL DISTRIBUTION: 24-HOUR DURATION

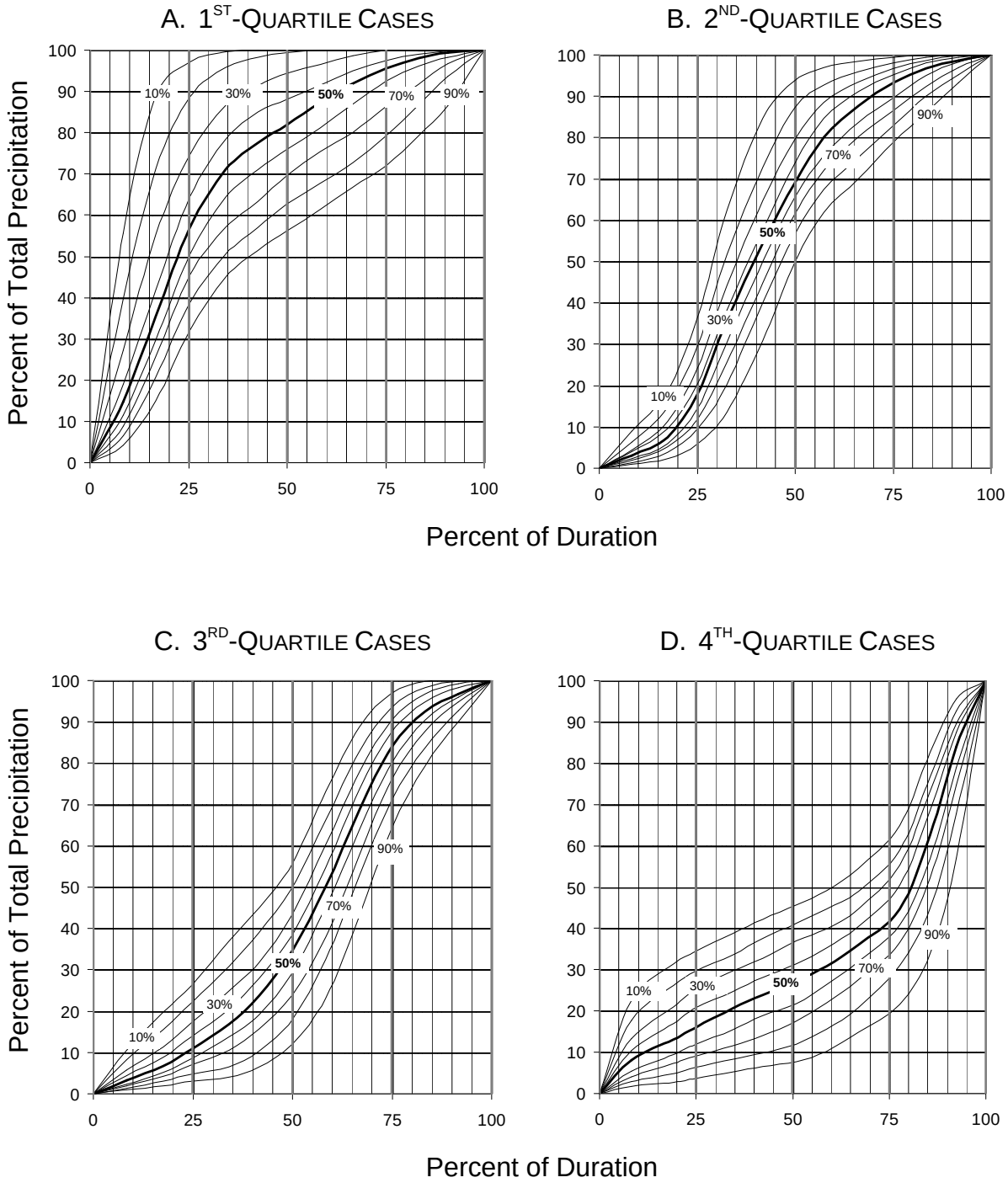
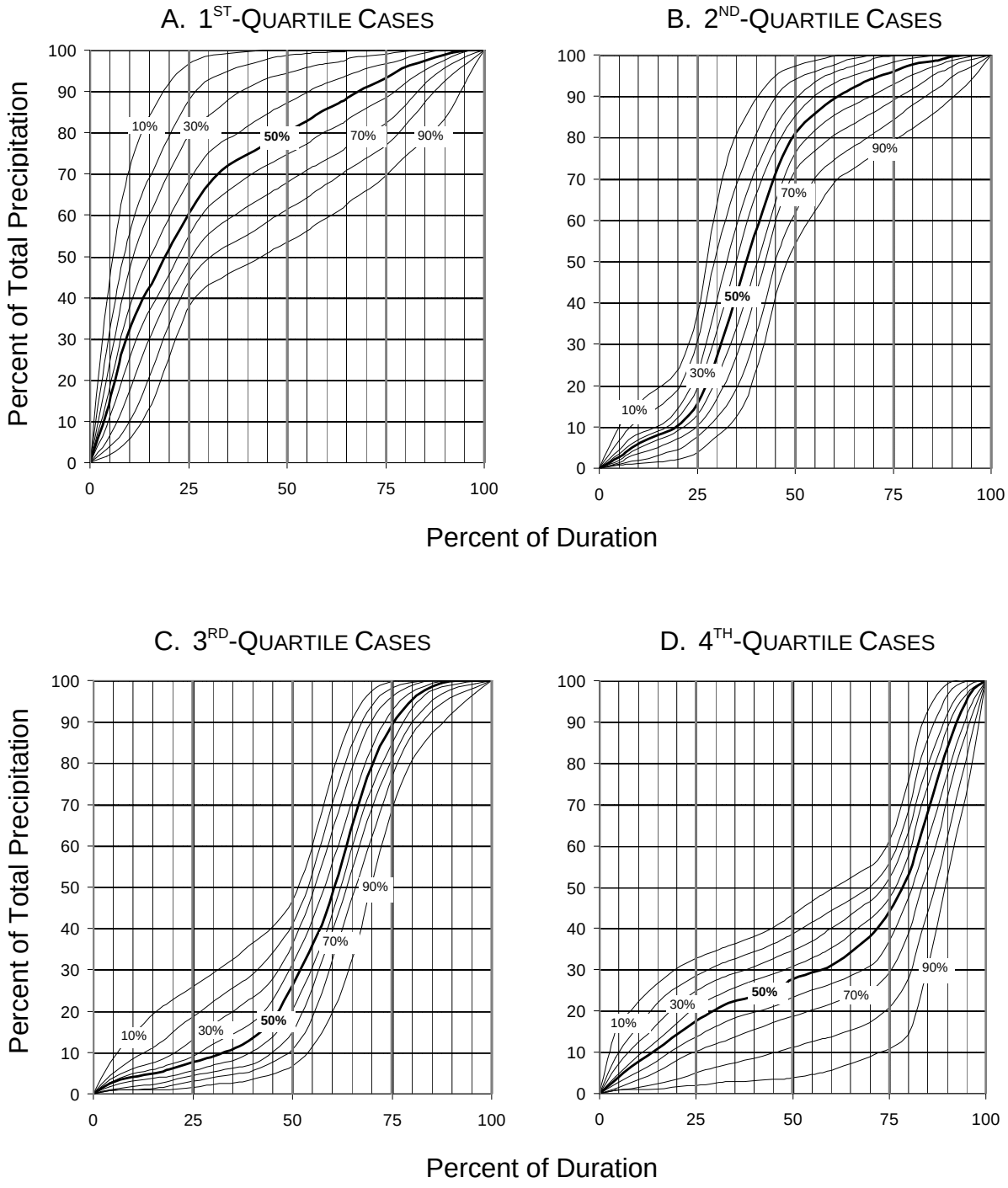


FIGURE A.1.6
TEMPORAL DISTRIBUTION: 96-HOUR DURATION



Appendix A.2. Seasonality

1. Introduction

Extreme precipitation over the Puerto Rico and U.S. Virgin Islands project area varies seasonally and regionally. The rainy season is typically from May to November as easterly waves embedded in the trade winds are the main mechanism to generate substantial rainfall. Orography enhances (reduces) precipitation on the windward (leeward) sides of the mountains. Orography has a greater influence on Puerto Rico than the smaller Virgin Islands. Rainfall from tropical low pressure systems and hurricanes is responsible for much of the extreme precipitation in the late summer and early fall (July through November). Localized convective showers and thunderstorms occur mainly during the warm season (April to October) and produce short to medium duration (5-minute to 48-hour) annual maximum precipitation. Sea breeze interactions enhance precipitation near the coast. During the winter months fronts from the north may reach the islands and produce extreme precipitation, especially long duration (greater than 24-hour) annual maximum precipitation.

To portray the seasonality of extreme precipitation throughout the project area, precipitation observations that exceeded given annual exceedance probabilities were examined for each region used in the analysis (Figures 4.4.1 and 4.4.2). Exceedance graphs showing this information on a monthly basis are provided as part of the Precipitation Frequency Data Server (PFDS).

2. Method

Exceedance graphs were prepared showing the percentage of events that exceeded selected annual exceedance probabilities (AEPs) in each month for each region. The quantiles were derived from annual maximum series at each station in the region as described in Section 4.2, Regional approach based on L-moments. Each graph shows the exceedances of the 1 in 2, 5, 10, 25, 50 and 100 AEPs.

Results for the 60-minute, 24-hour, 48-hour and 10-day durations are each provided in separate graphs. The results were compiled for each hourly region for the 60-minute (Figure 4.4.2) and each daily region for the 24-hour, 48-hour and 10-day (Figure 4.4.1).

To prepare the graphs, the number of events exceeding the precipitation frequency estimate at a station for a given AEP was tabulated for the selected durations. Cases were extracted in the same manner as for the generation of the annual maximum series (Section 4.1.3). The output for all stations in a given region was then combined, sorted by month, normalized by the total number of data years in the region and plotted via the PFDS.

3. Results

Seasonal exceedance graphs are available via the PFDS (http://hdsc.nws.noaa.gov/hdsc/pfds/). When a point is selected, a user can view the seasonal exceedance graphs by clicking the “Seasonality” button. The exceedance graphs (see Figure A.2.1 for an example) indicate a measure of events exceeding the corresponding AEP for the specified duration. The percentages are based on regional statistics. The total number of stations and the total number of cumulative data years for a given region are provided in the graph title.

The AEPs represent the probability of an event occurring that exceeds the quantile in any given year (i.e., 1 in 100 or 0.01 probability). Theoretically, 50% of the total number of events could exceed the 1 in 2 AEP, 4% could exceed the 1 in 25 AEP, 2% could exceed the 1 in 50 AEP and only 1% could exceed the 1 in 100 AEP. In other words, the sum of the 1 in 2 AEP percentages for each month in the graph roughly equals 50%.

The graphs also show how the seasonality of precipitation may differ between shorter duration and longer duration events in a region.

Seasonal precipitation frequency estimates cannot be derived from the graphs.

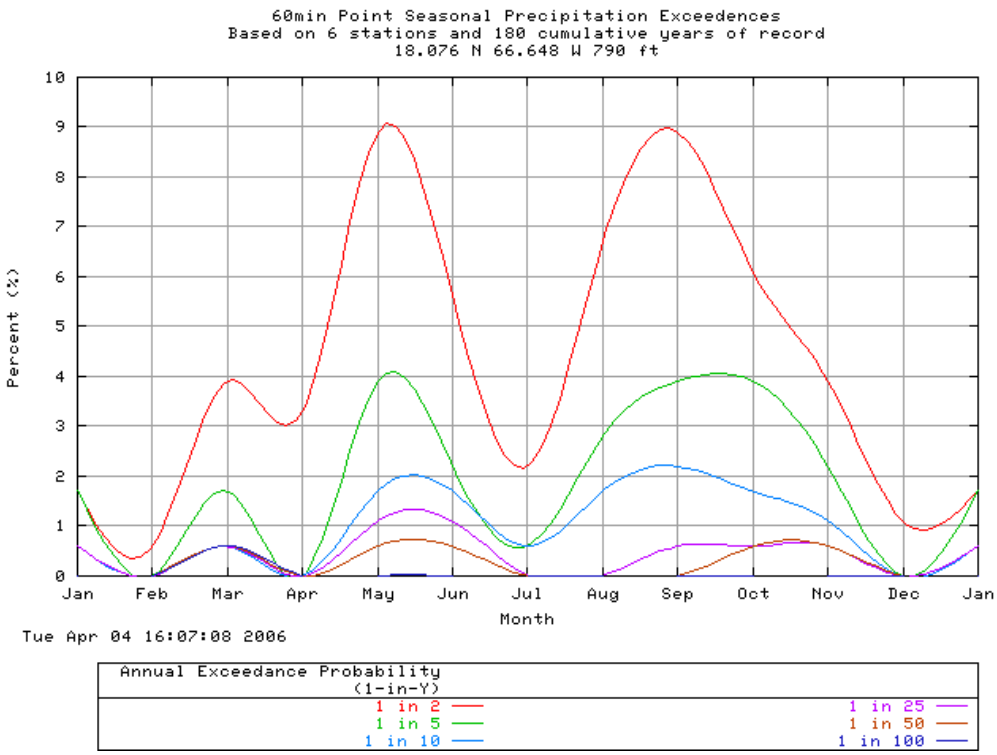


Figure A.2.1. Example of seasonal exceedance graph for the 60-minute duration.

Appendix A.3. Time series trend analysis associated with NOAA Atlas 14 Volume 3

1. Introduction

Precipitation frequency studies make the implicit assumption that the past is prologue for the future, i.e. that climate is stationary. Tests for linear trends in means and variance and shifts in mean were conducted on the 1-day annual maximum time series to verify the suitability of the data for this Atlas. The results of each test are provided and two specific examples of stations with linear trends and shifts are presented here. It was concluded that while there are some local instances of linear trends and shifts in mean in the data, it could be assumed that there was no consistent observed impact of climate change on the annual maximum series used for this Atlas. In particular, the impact upon the L-moment statistics and results of this Atlas would be small. Therefore, since it is beneficial to retain as much data as possible and thereby increase the robustness of the results, the entire period of record was used.

2. Linear Trend Tests

2.1. Methods

Linear trend tests were conducted to determine if there were any general increasing or decreasing patterns in the 1-day annual maximum series at a station through time. Data were tested for a linear trend in annual maximum series using the linear regression model and t-test of the correlation coefficient (Maidment, 1993, p17.30) at the 90% confidence level. Linear trends in variance were also tested by constructing a variance-related variable, an index of the square of deviation, or $v_i = (x_i - \bar{x})^2$ where, x_i is the annual maximum series data for i = 1, 2, …, n - the data year at a station, and $\bar{x}$ is the mean of the 1-day annual maximum series data. The index was then applied as a simple variable in the linear trend model. It was necessary for there to be a continuous time series to be eligible for the linear trend test. A minimum length of 50 years was chosen because it was sufficient to give reliable results and was close to the average data length of available stations.

Stations with gaps in the data record (i.e., sequential years of missing data) were evaluated and the following additional criteria were applied to maximize the use of limited data while still maintaining the integrity of the time series for the tests.

- Stations with gaps greater than or equal to 10 years were not used.
- Stations with a 5-9 year gap but with at least 6 years of data on both sides of the gap were retained.
- Stations were truncated where appropriate to eliminate gaps and still retain a record of 50 years or more. For instance, stations with a 5-9 year gap and less than 6 years of data at the beginning or end of a time series were truncated.

2.2. Linear Trend Results

Of 128 stations in the project area, 55 (or 43.0%) were eligible for the test. Of those tested stations, 23.6% exhibited a linear trend in their annual maximum series (23.6% in a positive direction, none in a negative direction). It is notable that all stations exhibiting a linear trend were in the positive direction, however there were relatively few stations available for testing. Table A.3.1 lists the linear trend results in the project area. Figure A.3.1 shows the spatial distribution of stations with linear trends.

Table A.3.1. Number of stations tested and linear trend test results in the Puerto Rico project area.

	# Tested	# No Trend	# Trend	# Pos. Trend	# Neg. Trend	% tested with Trend
Total	**55**	**42**	**13**	**13**	**0**	**23.6**

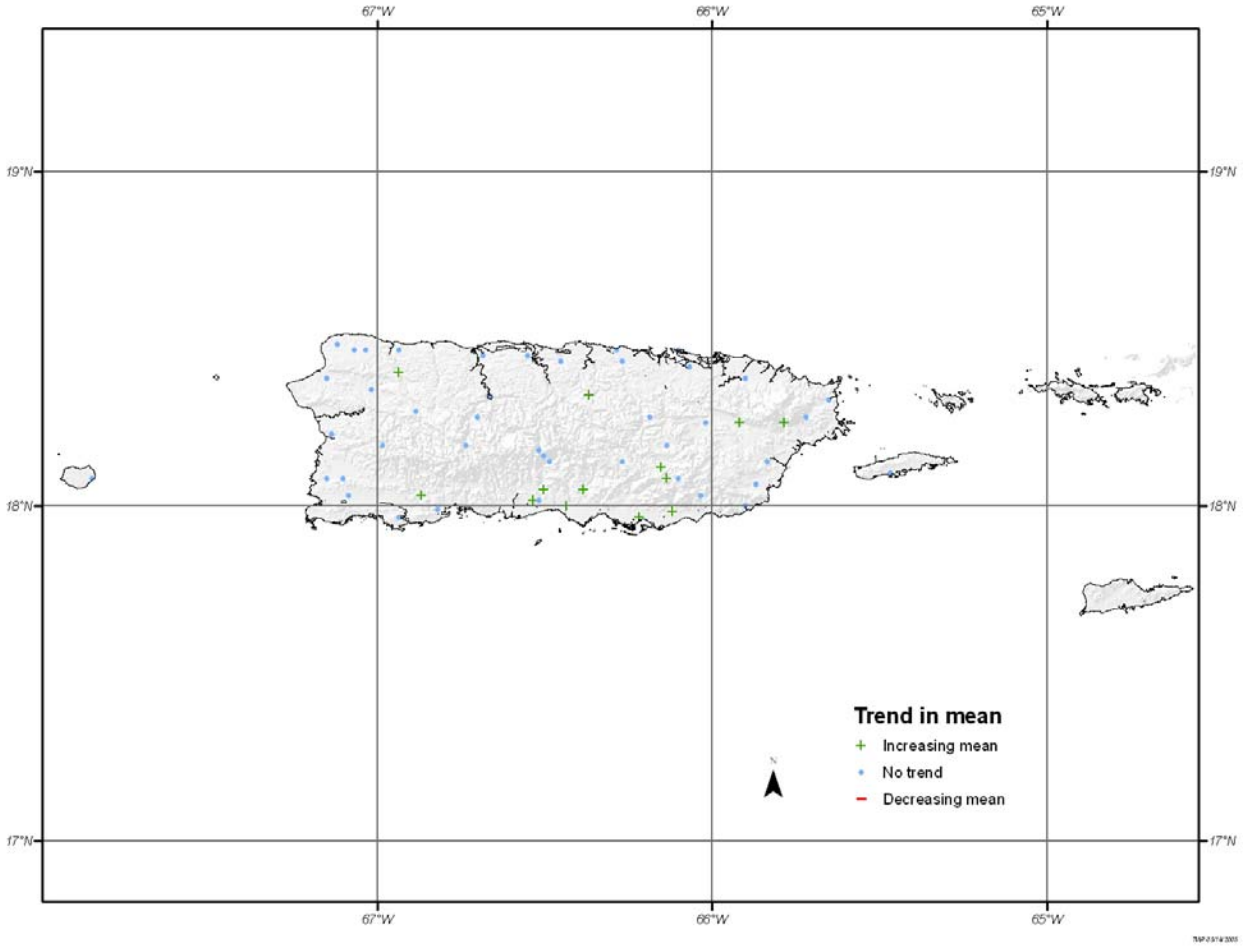


Figure A.3.1. Spatial distribution of linear trend results, where "+" indicates a station with a positive trend and "-" indicates a negative trend.

Geographically, most of the more upward trending stations occurred along the southern coastal and interior areas. This coast typically receives the highest maximum observations but the lowest mean annual precipitation. However, the majority of stations exhibited no trend.

Overall, there appeared to be no definitive linear trend in the tested annual maximum time series. However, in the stations that did test positive for a trend in mean, there was a preference for an increase in mean particularly in the southern areas.

2.3. Linear Trend in Variance Results

Of the 55 stations tested, 27.3% exhibited a trend in the variance of annual maximums (20.0% in a positive direction, 7.3% in a negative direction). In other words, 20.0% of the stations that exhibited

such a trend showed an increase in variance. Table A.3.2 lists the trend in variance results in the project area. Figure A.3.2 shows the spatial distribution of those stations that had a trend in variance.

Table A.3.2. Number of stations tested and linear trend in variance test results in the Puerto Rico project area.

	# Tested	# No Trend	# Trend	# Pos. Trend	# Neg. Trend	% tested with Trend
Total	**55**	**40**	**15**	**11**	**4**	**27.3**

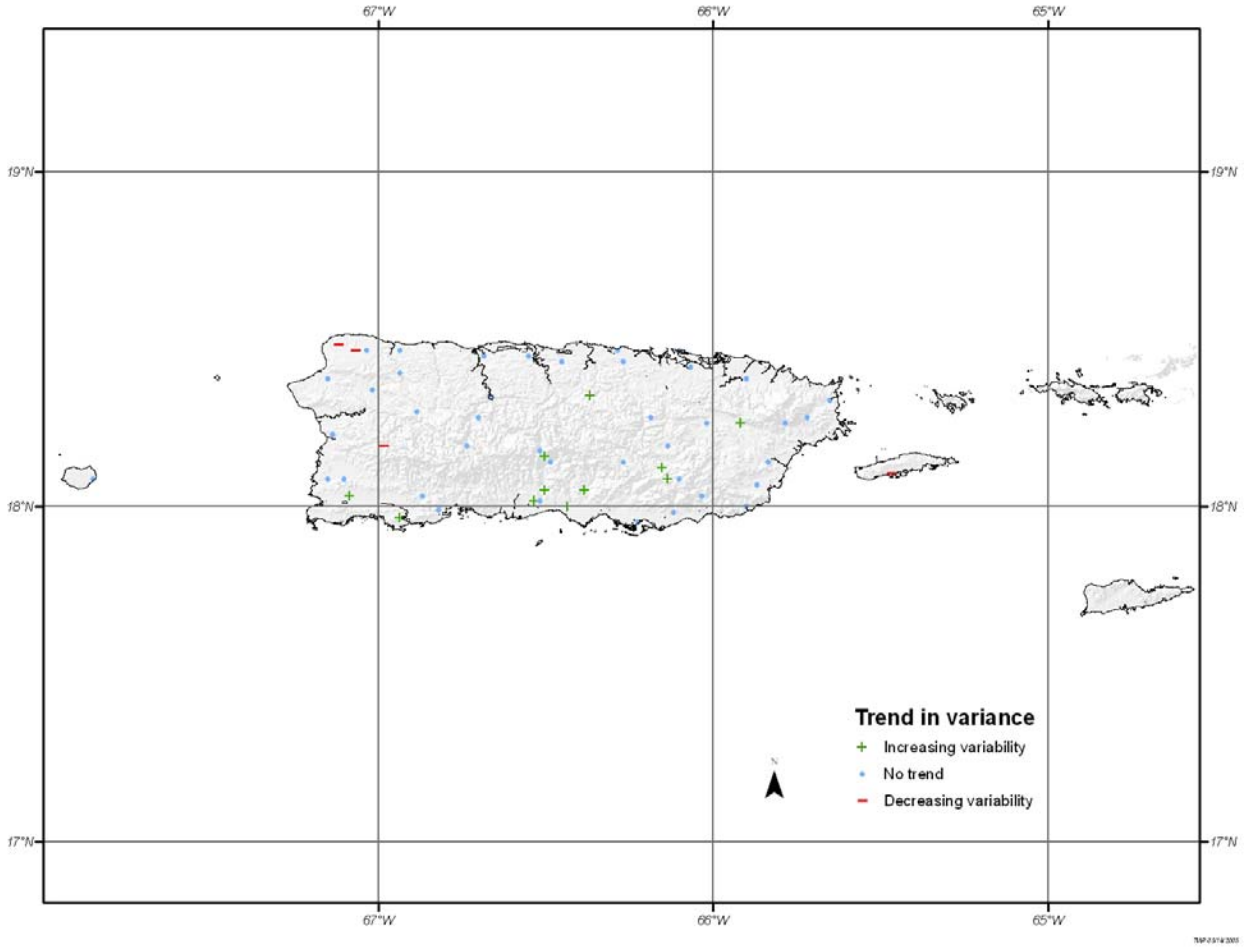


Figure A.3.2. Spatial distribution of trend in variance results, where "+" indicates a station with a positive trend and "-" indicates a negative trend.

Again, geographically, there were more upward trending stations along the southern coastal and interior areas. This coast typically receives the highest maximum observations but the lowest mean annual precipitation. However, the majority of stations exhibited no trend in variance.

Overall, there appeared to be no definitive linear trend in variance in the tested annual maximum. However, in the stations that did test positive for a trend in variance, there was a preference for an increase in variance particularly in the southern areas.

3. Shift in Mean Tests

3.1. Methods

A shift test was conducted to compare the means of 1-day annual maximum series for two consecutive time periods at a station. The data were tested for shifts in mean using Mann Whitney non-parametric test (Newbold, 1988, p403) and the t-test (Lin, 1980, p160) at the 90% confidence levels. The Mann Whitney is a qualitative test that indicated if a shift occurred but not the direction of the shift. The t-test provided a quantitative measurement of the percentage that the mean shifted from one time period to the next. Both tests give consistent results suggesting that the parametric t-test results can be used with assurance to assign quantitative values to observed shifts. The year 1958 was used to divide the time periods because 1958 was roughly the final year for which Technical Paper 42 (U.S. Weather Bureau, 1961) had data. The results would indicate whether a shift has occurred since the publication of earlier precipitation frequency estimates. A minimum of 30 years of data in each data segment were required at a station to test for shifts in mean.

Since the Mann Whitney test uses ranks, it was better to have similar sizes between the two subsamples. A threshold of 30 years difference in the lengths of the subsamples was set based on testing and used to screen the stations eligible for that test. However, since the t-test is a parametric test following the t-distribution or Normal distribution, the test is less sensitive to the difference between the sample sizes. However, in this project, all stations eligible for the t-test were also eligible for the Mann Whitney test.

3.2. Shift in mean results

The results when using 1958 as the division were:

- T-test: 22 of 128 (17.2%) were eligible. 36.4% of those tested had a shift in mean (36.4% increased in mean, 0.0% decreased in mean).
- Mann Whitney test: 22 of 128 (17.2%) were eligible. 36.4% of those tested had a shift in mean.

Table A.3.3 lists the shift in mean results comparing pre-1958 data and post-1958 data in the project area. The last column in the table shows the average percent change in mean for the project area. Overall, the majority of stations tested did not exhibit a shift in mean. Where shifts did occur, the shifts in mean showed a clear preference toward increasing shifts.

Table A.3.3. Number of stations tested and test for shift in mean results (1958 split) in the Puerto Rico project area.

	# Tested	# No Shift	# Shift	# Pos. Shift	# Neg. Shift	% Change in Mean
Total	**22**	**14**	**8**	**8**	**0**	**36.4 (avg)**

Figure A.3.3 shows the spatial distribution of the stations that have a shift in mean. The numbers plotted above the station location indicate the percentage of change in mean at each station. The shift in mean was dominantly upward and nearly all occurrences were found along that southern coastal and interior areas. However, the majority of the stations exhibited no trend. The shift results are consistent with the results of the linear trend results.

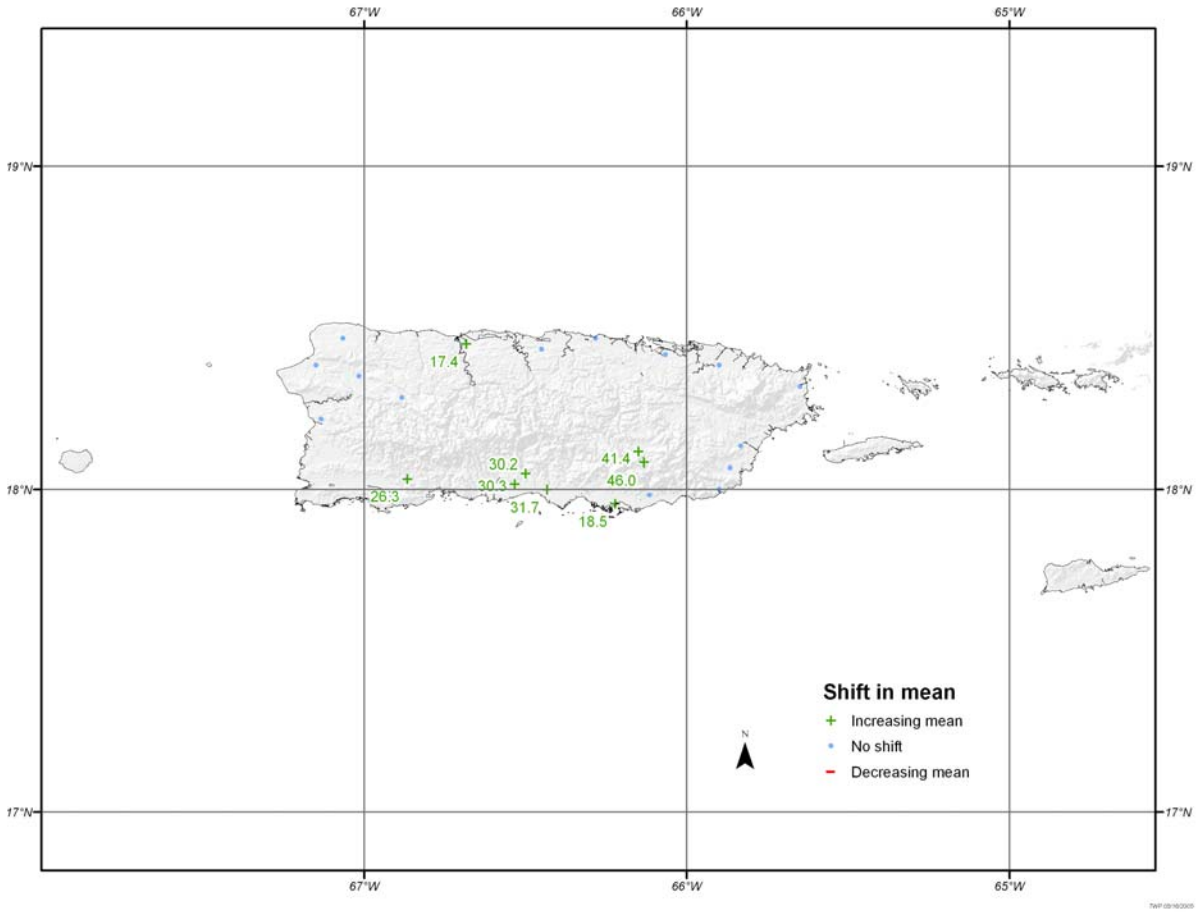


Figure A.3.3. Spatial distribution of shift in mean results, where "+" indicates a station with a positive trend, "-" indicates a negative trend and the number indicates the percentage of change (1958 split).

4. Specific Examples

In many cases, stations that showed a linear increase had a similar shift in mean. Figure A.3.4 shows a combined upward linear trend with an upward shift in mean (where the subsamples are divided at the year 1958) at Juana Diaz Camp, PR (66-5020), which is located in the south-central area of the island. The time series for the station (1901 - 2004) is plotted and a solid straight diagonal line represents the linear trend. There was an accompanying increasing shift in mean (+30.2%) from the 1901 -1958 time period (3.76") to the 1959-2004 time period (4.90"). The means of each time period are represented as separate horizontal lines. The linear trend in variance was also increasing through time. This indicates that there were more extreme events with time. The increase in variance is shown in the Figure by the dashed lines outward of the linear trend line. Sensitivity testing showed that the trend at this station was driven by the highest three events that happened to be hurricane-related (12.78" on October 6, 1985 – extra-tropical storm; 12.60" on September 22, 1998 – Hurricane Georges; and 11.13" on September 18, 2000 – Tropical Storm Helene). The linear trend in mean, linear trend in variance or shift in mean are removed when these three events are omitted from the time series.

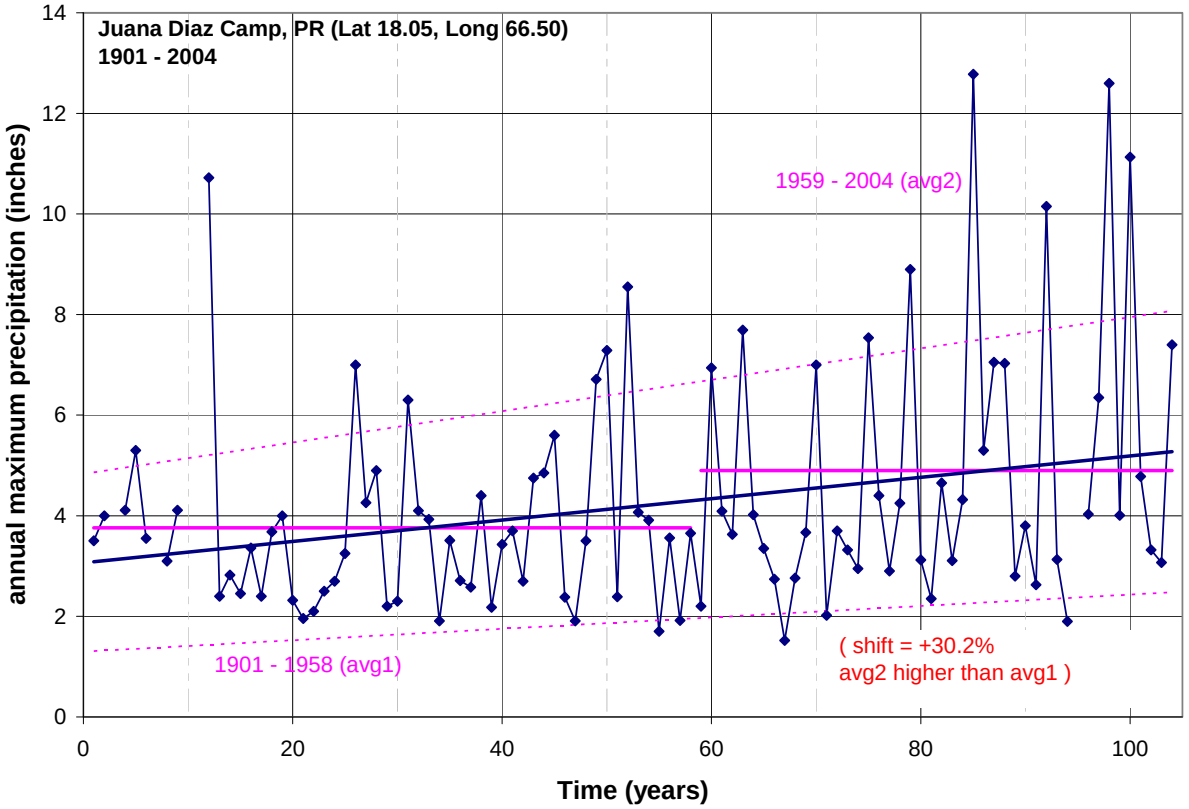


Figure A.3.4. Plot of increasing linear trend and shift tests and increasing linear variance for annual maximum time series at Juana Diaz Camp, PR (66-5020).

Figure A.3.5 shows another example of a station, Ponce 4 E, PR (66-7292), which is located along the south-central coast, with an upward linear trend and an upward shift in mean (where the subsamples are divided at the year 1958). The time series for the station (1900 - 2004) is plotted and

a solid straight diagonal line represents the linear trend. There was an accompanying increasing shift in mean (+30.3%) from the 1901 -1958 time period (3.66") to the 1959-2004 time period (4.77"). The linear trend in variance was also increasing through time. This indicates that there were more extreme events with time. Sensitivity testing showed that the trend in mean was *not* driven by the highest four events (18.20" on October 7, 1985 – extra-tropical storm; 12.87" on September 18, 2000 – Tropical Storm Helene; 9.10" on September 22, 1998 – Hurricane Georges; and 8.99" on August 23, 2001 – Hurricane Dean). After omitting these four events, the data still exhibited a trend in mean, however, the trend in variance and shift in mean were no longer present.

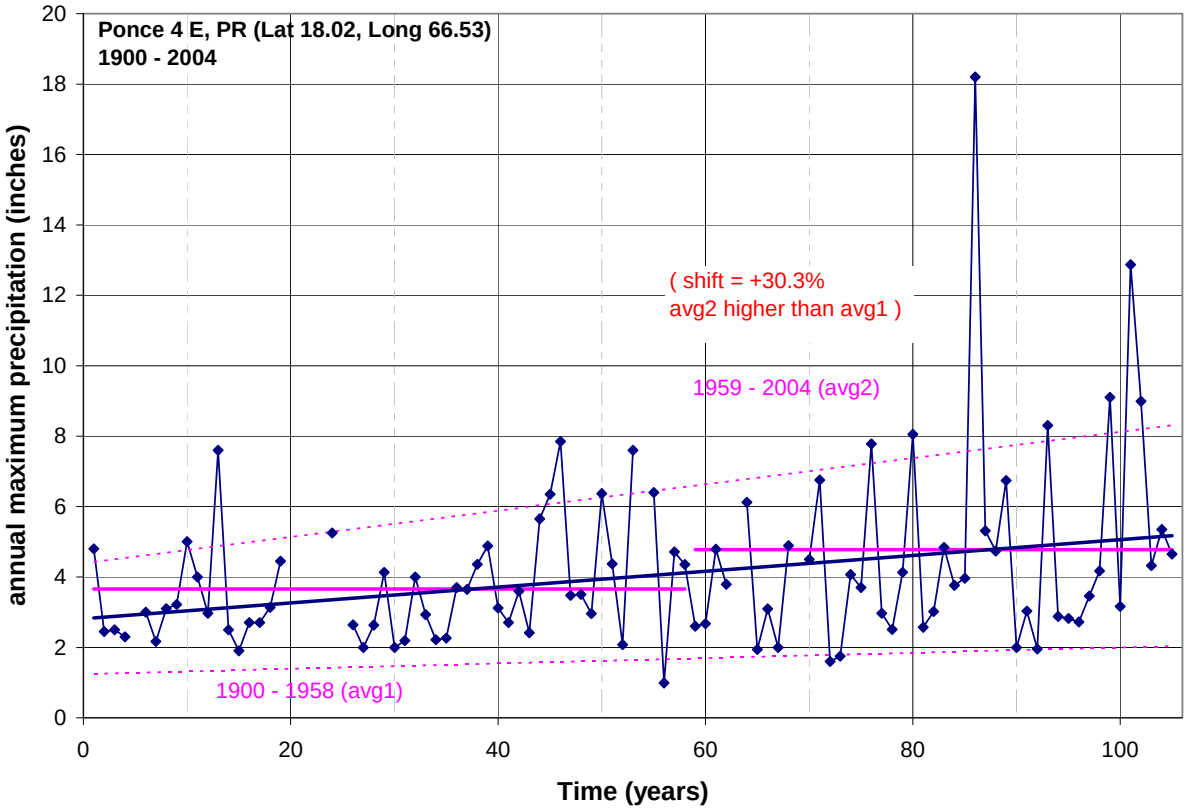


Figure A.3.5. Plot of increasing linear trend and shift tests for annual maximum time series at Ponce 4 E, PR (66-7292).

Overall, sensitivity testing of the drivers of trends in the series were inconclusive. Hurricane events drove trends at some stations but not at others. There was no clear indication that trends in annual maxima were related to any such trend in hurricane-related events at stations in the project area. Also, an extreme event on January 6, 1992 that was the result of a stationary cold front was identified as contributing to upward trends.

5. Conclusions

1-day precipitation annual maximum series for stations used in NOAA Atlas 14 Volume 3 were examined for linear trends in mean, linear trends in variance, and shifts in mean. The following conclusions about the stations tested can be made:

1. Overall, the annual maximum time series were free from linear trends and from shifts in mean for most of the stations in the project area.
2. The southern coastal and interior island was qualitatively shown to be the geographical preference for upward trending or shifting annual maximum series data.

Therefore, since the results showed little observable or geographically consistent impact of change in the statistics used to estimate precipitation frequency, the entire historical time series was used in this Atlas.

Appendix A.4 (report was formatted by HDSC)

Final Report
Production of Rainfall Frequency Grids for Puerto Rico and the U.S. Virgin Islands Using a Specifically Optimized PRISM System

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

Prepared by
Christopher Daly
PRISM Group
Oregon State University

June 2006

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 Puerto Rico and the U.S. Virgin Islands (PRVI). In order to complete the spatial interpolation of point estimates, HDSC requires spatially interpolated grids of mean annual maximum precipitation. To that end, the PRISM Group (formerly known as the Spatial Climate Analysis Service) at Oregon State University, produced a series of grids for rainfall frequency estimation using an optimized system based on the Parameter-elevation Regressions on Independent Slopes Model (PRISM) and HDSC-calculated point estimates for the Puerto Rico and Virgin Islands study domains.

Background

HDSC used the "Index Flood" approach as described by Hosking and Wallis in "Regional Frequency Analysis; An Approach Based on L-Moments", 1997, to estimate rainfall frequencies. In this approach, the mean of the underlying rainfall frequency distribution is estimated at point locations with a sufficient history of observations. This mean is referred to as the "Index Flood" because early applications of the method were used to analyze flood data in hydrology. The form of the distribution its parameters are estimated regionally. Once the form of the distribution has been selected and its parameters have been estimated, rainfall frequency estimates can be computed from grids of the Index Flood. The grids that are the subject of this report are spatially interpolated grids of the point estimates of the Index Flood for various precipitation durations. The point estimates of the Index Flood were provided by HDSC. HDSC selected an appropriate rainfall frequency distribution along with regionally estimated parameters and used this information with the grids of the Index Flood to derive grids of rainfall frequency estimates.

The PRISM Group has previously performed similar work to produce spatially interpolated Index Flood grids for updates of precipitation frequency estimates in the Semiarid Southwest United States and the Ohio River Basin and Surrounding States study areas.

This Report

This report describes tasks performed to produce final index flood grids for 14 precipitation durations, ranging from 60 minutes to 60 days, for PRVI. These tasks were not necessarily performed in this order, nor were they performed just once. The process was dynamic and had numerous feedbacks.

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 study area. 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) 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).

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 1961-1990 mean precipitation maps (USDA-NRCS 1998). 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 rainfall stations only. This meant that mapping precipitation frequency using HDSC stations would sacrifice a significant amount of the spatial detail present in the 1961-1990 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 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 much of the incorporation of the effects of various physiographic features on mean precipitation patterns had already been accomplished with the creation of 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. Square-root 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, previously-developed grids of MAP for Puerto Rico (1963-1995 averaging period) and the Virgin Islands (1971-2000) were used. Both of these grids were developed under funding from the International Institute of Tropical Forestry, US Forest Service (Daly et al., 2003). MAP grids for Puerto Rico and the Virgin Islands (US and British) are shown in Figures 1 and 2.

PRISM Configuration and Operation for PRVI

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). 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 index flood vs. MAP took the form

$$\text{Index flood value} = \beta_1 * \text{sqrt(MAP)} + \beta_0 \qquad (1)$$

where β_1 is the slope and β_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. In the general PRISM formulation for precipitation, the combined weight of a station can be a function of distance, elevation, cluster, vertical layer, topographic facet, coastal proximity, and effective terrain weights, respectively. A full discussion of the station weighting functions is available in Daly (2002) and Daly et al. (2002).

A subset of these functions was used for this study. For PR, the combined weight of a station was a function of distance, MAP, cluster, topographic facet, and coastal proximity, respectively. For VI, only distance and cluster weighting were used, due to lack of station data, and the small size of the islands. Distance, MAP, and cluster weighting are relatively straightforward in concept. A station is down-weighted when it is relatively distant or has a much different MAP value than the target grid

cell, or when it is clustered with other stations (which can lead to over-representation). Facet weighting effectively groups stations into individual hillslopes (or facets), at a variety of scales, to account for sharp changes in climate regime that can occur across facet boundaries. Coastal proximity weighting is used to define gradients in precipitation that may occur due to proximity to large water bodies (Daly et al., 2003).

The moving-window regression function was populated by station data provided by the HDSC. Locations of these stations are shown in Figure 3. A PRISM GUI snapshot of the moving-window relationship between MAP and 24-hour index flood in the Cordillera Central (Central Mountains) is shown in Figure 4a.

As exemplified by the 60-minute duration station map in Figure 3a, there were little station data available for durations of 12 hours or less from which to perform the interpolation. In addition, it became clear that, at least in Puerto Rico, the spatial patterns of durations of 12 hours or less were very different than those of durations of 24 hours or more. In an effort to bring the ≤ 12-hour station density up to that for ≥ 24 hours, the following procedure was developed:

(1) Convert available ≤ 12-hour station values to an index flood/24-hr index flood ratio (termed R24) by dividing by the 24-hour values;

(2) using the station R24 data in (1), interpolate R24 values for each ≤ 12-hour duration (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 ≥ 24-hour durations only;

(4) multiply the estimated R24 values from (3) by the 24-hour value at each ≥ 24-hour station to obtain estimated ≤ 12-hour values;

(5) append the estimated stations from (4) to the ≤ 12-hour station list to generate a station list that matches the density of that for ≥ 24 hours; and

(6) interpolate index flood values for ≤ 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 were strongly affected by MAP, coastal proximity, topographic facets, or other factors. Therefore, the slope of the moving-window regression function for R24 vs. MAP of the form

$$R24 = \beta_1 * \text{sqrt}(MAP) + \beta_0 \qquad (2)$$

was forced to zero everywhere. This meant that, for both PR and VI, the interpolated value of R24 was a function of distance and cluster weighting only (essentially inverse-distance weighting). A PRISM GUI snapshot, shown in Figure 4b, of the relationship between MAP and 60-minute R24 in the Cordillera Central (Central Mountains) shows a rather weak scatter plot, with no obvious relationship between 60-minute R24 and MAP.

Relevant PRISM parameters for applications to 60-minute R24 and 24-hour index flood statistics are listed in Tables 1 and 2, respectively. Further explanations of these parameters and associated

equations are available in Daly (2002) and Daly et al. (2002). The difference to note between the parameter set in Tables 1 and 2 and that in Daly et al. (2002) is that the elevation weighting parameters in Daly et al. (2002) are now referred to here as MAP weighting parameters. This is because MAP, rather than elevation, was used as the predictor variable. Input parameters used for the 60-minute R24 application were generally applied to all durations for which it was applied (less than or equal to 12 hours). The 24-hour index flood input parameters were generally applied to all durations.

The values of radius of influence (*R)*, the minimum number of on-facet (s_f) and total (s_t) 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).

Input parameters that changed readily among the various durations were the maximum allowable slope (β_{1x}) and default slope (β_{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). Slope bounds and 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. The upper and lower bounds were set to approximately the 95th and 5th percentiles of the distribution of slopes, respectively, because many of the slopes outside this range are typically found to be questionable. For these applications, slope bounds typically increased with increasing duration (Table 3). In general, the longer the duration, the larger the maximum allowable slope. This is primarily a result of higher precipitation amounts at the longer durations, and the tendency for longer-duration index flood statistics to bear a stronger and steeper relationship with MAP than shorter-duration statistics.

Review of Draft Grids

Initial draft grids of 1- and 24-hour index flood statistics for the PR and VI regions were produced by running PRISM directly on the index flood station data for all durations (the ratio method for durations of 12 hours or less was not developed until after the external review). HDSC derived 100-year return period maps from the initial drafts, and also made these available for review. Comments received and our responses are presented in Appendix A.

The external review resulted in two major changes to the mapping methodology:

(1) Development and implementation of the ratio (R24) method for durations of 12 hours and less (discussed earlier): and
(2) A revision of the MAP grid for PR that lowers the precipitation on the hill discussed in Question 5, Appendix A. MAP grid revision required that the PRISM model be re-applied to PR after adding an estimated station on the questionable hill with a MAP of 1564 mm.

Results

PRISM grids of 60-minute/24-hour index flood intensity ratio (R24), 60-minute index flood intensity, and 24-hour index flood intensity are shown in Figures 5, 6, and 7, respectively. In inverse-distance-weighting mode, PRISM produces a R24 map that has no appreciable physiographic features (Figure 5). Current data are insufficient to provide further insight into the causes and controls of these spatial patterns. However, major patterns across the island are clearly evident. The western side of the island has substantially higher R24 values than the eastern side. The northwestern corner of PR had 60-minute index flood values that were over 60 percent of the 24-hour values. On the eastern side, values were as low 30 percent. This spatial discrepancy in R24 is the source of the differences in spatial patterns between the 60-minute index flood (Figure 6) and the 24-hour index flood maps (Figure 7). Sixty-minute values are highest in the northwest and in the western portion of the Cordillera Central (Central Mountains), with smaller maxima in the eastern mountains, especially the Luquillos. In contrast, the 24-hour map bears a much greater resemblance to the MAP map (Figure 1), with the highest values in the central and eastern Cordillera Central, and the Luquillo Mountains.

PRISM cross-validation statistics for applications 1- and 24-hour applications to the PRVI region were compiled and summarized in Tables 4 and 5. Overall bias was less than 2 percent and the mean absolute error was less than 10 percent for the 60-minute/24-hour index flood ratio, and the 60-minute and 24-hour index flood intensities. Errors for 2- to 12-hour durations were similar to those for the 60-minute duration, and errors for 2 to 60-day durations were similar to those for the 24-hour duration. Given the lack of data, one would have expected the 1 to 12-hour index flood errors to be somewhat higher than those for the 24-hour to 60-day index floods. A likely reason for this is 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 1–hour index flood are higher than those shown in Table 4.

Deliverables

A full set of maps for all index flood durations were produced for the PRVI region, including 60 minutes, 2, 3, 6, 12, and 24 hours; and 2, 4, 7, 10, 20, 30, 45, and 60 days. The maps were subjected to pixel-by-pixel tests to ensure that shorter duration values did not exceed those of longer duration values. PRISM modeling was performed at 15-second (~450-m) resolution in PR and at 3-second (~90-m) resolution in VI. The PR grids were subsequently filtered to a 3-second resolution to match that of VI using a modified Gaussian filter (Barnes, 1964), and the PR and VI grids merged into single PRVI grids. These grids were delivered electronically to HDSC via ftp.

References

Barnes, S.L. 1964. A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology,* 3:396-409.

Daly, C. 2002. Variable influence of terrain on precipitation patterns: Delineation and use of effective terrain height in PRISM. World Wide Web document. http://www.ocs.orst.edu/pub/prism/docs/effectiveterrain-daly.pdf

_____. 2006. Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology,* Vol 26: 707-721. http://www.ocs.orst.edu/pub/prism/docs/intjclim06-guidelines-daly.pdf

_____, E.H. Helmer, and M. Quinones. 2003. Mapping the climate of Puerto Rico, Vieques, and Culebra. *International Journal of Climatology*, 23: 1359-1381. http://www.ocs.orst.edu/pub/prism/docs/jclim03-map_climate_PR.pdf

_____, 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://www.ocs.orst.edu/pub/prism/docs/jappclim94-modeling_mountain_precip-daly.pdf

_____, 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://www.ocs.orst.edu/pub/prism/docs/asae00-spatial_climate_datasets-daly.pdf

_____, 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://www.ocs.orst.edu/pub/prism/docs/climres02-kb_approach_statistical_mapping-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.

Table 1. Values of relevant PRISM parameters for interpolation of 60-minute/24-hour index flood ratio (60-minute R24) for PR (Puerto Rico) and VI (US Virgin Islands). See Daly et al. (2002) for details on PRISM parameters.

Name	Description	PR/VI Values
Regression Function		
R	Radius of influence	1.7/0.4 km*
s_f	Minimum number of on-facet stations desired in regression	4/4 stations
s_t	Minimum number of total stations desired in regression	12/12 stations
β_{1m}	Minimum valid regression slope	0.0/0.0[+]
β_{1x}	Maximum valid regression slope	0.0/0.0[+]
β_{1d}	Default valid regression slope	0.0/0.0[+]
Distance Weighting		
A	Distance weighting exponent	2.0/2.0
F_d	Importance factor for distance weighting	1.0/1.0
D_m	Minimum allowable distance	0.0/0.0 km
MAP Weighting**		
B	MAP weighting exponent	NA/NA
F_z	Importance factor for MAP weighting	NA/NA
Δz_m	Minimum station-grid cell MAP difference below which MAP weighting is maximum	NA/NA
Δz_x	Maximum station-grid cell MAP difference above which MAP weight is zero	NA/NA
Facet Weighting		
C	Facet weighting exponent	NA/NA
g_m	Minimum inter-cell elevation gradient, below which a cell is flat	NA/NA
λ_x	Maximum DEM filtering wavelength for topographic facet determination	NA/NA
Coastal Proximity Weighting		
v	Coastal proximity weighting exponent	NA/NA

* 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].

** Normally referred to as elevation weighting.

Table 2. Values of relevant PRISM parameters for modeling of 24-hour index flood statistics for PR (Puerto Rico) and VI (US Virgin Islands). See Daly et al. (2002) for details on PRISM parameters.

Name	Description	PR/VI Values
Regression Function		
R	Radius of influence	1.0/0.4 km*
s_f	Minimum number of on-facet stations desired in regression	10/6 stations
s_t	Minimum number of total stations desired in regression	10/15 stations
β_{1m}	Minimum valid regression slope	$0.0/0.0^+$
β_{1x}	Maximum valid regression slope	$6.1/10.0^+$
β_{1d}	Default valid regression slope	$1.7/5.0^+$
Distance Weighting		
A	Distance weighting exponent	2.0/2.0
F_d	Importance factor for distance weighting	0.5/0.5
D_m	Minimum allowable distance	0.0/2.7 km
MAP Weighting**		
B	MAP weighting exponent	1.0/0.0
F_z	Importance factor for MAP weighting	0.5/0.5
Δz_m	Minimum station-grid cell MAP difference below which MAP weighting is maximum	50/50%
Δz_x	Maximum station-grid cell MAP difference above which MAP weight is zero	500/500%
Facet Weighting		
C	Facet weighting exponent	1.5/NA
g_m	Minimum inter-cell elevation gradient, below which a cell is flat	1 m/NA
λ_x	Maximum DEM filtering wavelength for topographic facet determination	17 km/NA
Coastal Proximity Weighting		
v	Coastal proximity weighting exponent	1.0/NA

* 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].

** Normally referred to as elevation weighting

Table 3. Values of PRISM slope parameters for modeling of index flood statistics for PR (Puerto Rico) and VI (US Virgin Islands) for all durations. For durations of 12 hours and below, station data were expressed as the ratio of the given duration's index flood value to the 24-hour index flood value, and interpolated; this was followed by an interpolation of the actual index flood values. See text for details. See Table 1 for definitions of parameters.

	Puerto Rico			Virgin Islands		
Duration	β_{1m}	β_{1x}	β_{1d}	β_{1m}	β_{1x}	β_{1d}
60m/24h ratio	0.0	0.0	0.0	0.0	0.0	0.0
2h/24h ratio	0.0	0.0	0.0	0.0	0.0	0.0
3h/24h ratio	0.0	0.0	0.0	0.0	0.0	0.0
6h/24h ratio	0.0	0.0	0.0	0.0	0.0	0.0
12h/24h ratio	0.0	0.0	0.0	0.0	0.0	0.0
60 minute index flood	0.0	5.0	1.4	0.0	5.0	10.0
2 hour index flood	0.0	5.0	1.4	0.0	5.0	10.0
3 hour index flood	0.0	5.0	1.4	0.0	5.0	10.0
6 hour index flood	0.0	5.0	1.4	0.0	5.0	10.0
12 hour index flood	0.0	5.0	1.4	0.0	5.0	10.0
24 hour index flood	0.0	6.1	1.7	0.0	5.0	10.0
48 hour index flood	0.0	6.3	1.7	0.0	5.0	10.0
4 day index flood	0.0	6.3	1.9	0.0	5.0	10.0
7 day index flood	0.0	6.5	2.4	0.0	5.0	10.0
10 day index flood	0.0	6.7	2.7	0.0	5.0	10.0
20 day index flood	0.0	6.9	3.3	0.0	5.0	10.0
30 day index flood	0.0	7.0	3.7	0.0	5.0	10.0
45 day index flood	0.0	8.0	3.8	0.0	5.0	10.0
60 day index flood	0.0	10.0	4.0	0.0	5.0	10.0

Table 4. PRISM cross-validation errors for 60-minute/24-hour index flood ratio and 24-hour index flood applications to PR (Puerto Rico). Since the 60-minute/24-hour index flood 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).

Statistic	N	% Bias	% MAE
60-min/24-hr index flood ratio	23	0.64	5.78
60-minute index flood	113	1.77	8.38
24-hour index flood	113	1.83	8.82

Table 5. PRISM cross-validation errors for 60-minute/24-hour index flood ratio and 24-hour index flood applications to VI (US Virgin Islands). Since the 60-minute/24-hour index flood 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).

Statistic	N	% Bias	% MAE
60-min/24-hr index flood ratio	2	0.0	2.00
60-minute index flood	19	2.32	8.88
24-hour index flood	19	2.32	8.88

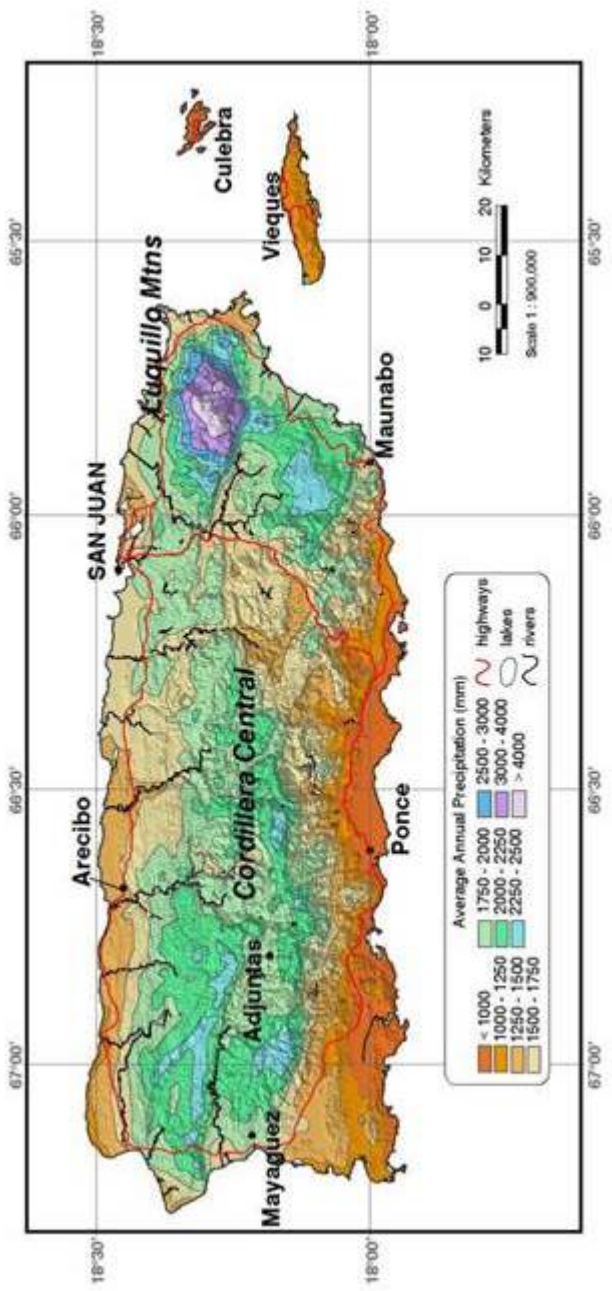


Figure 1. 1963-1995 mean annual precipitation (MAP) grid for Puerto Rico (Source: Daly et al., 2003).

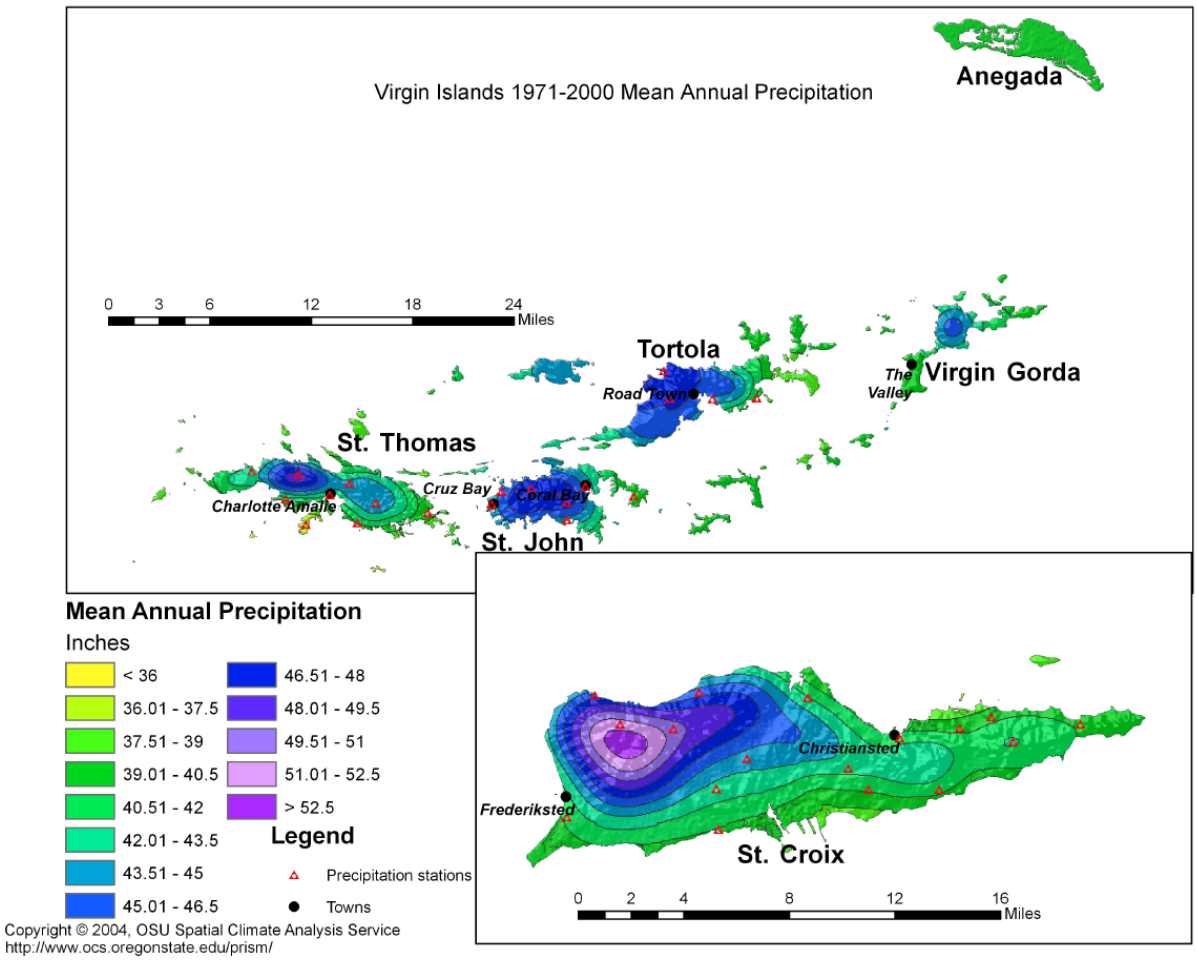
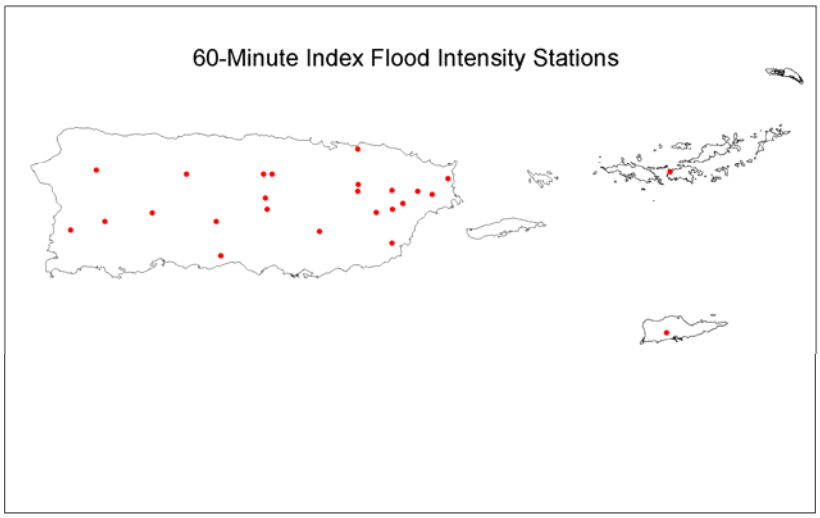


Figure 2. 1971-2000 mean annual precipitation (MAP) grid for the Virgin Islands.

(a)\



(b)

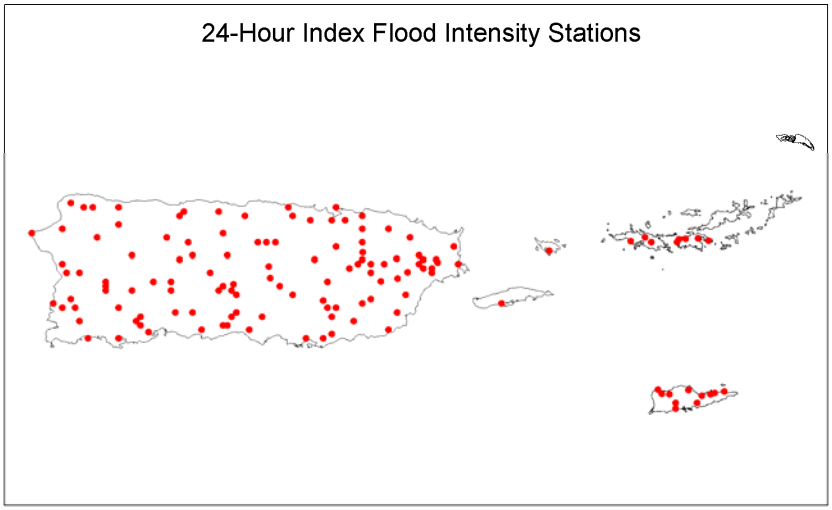
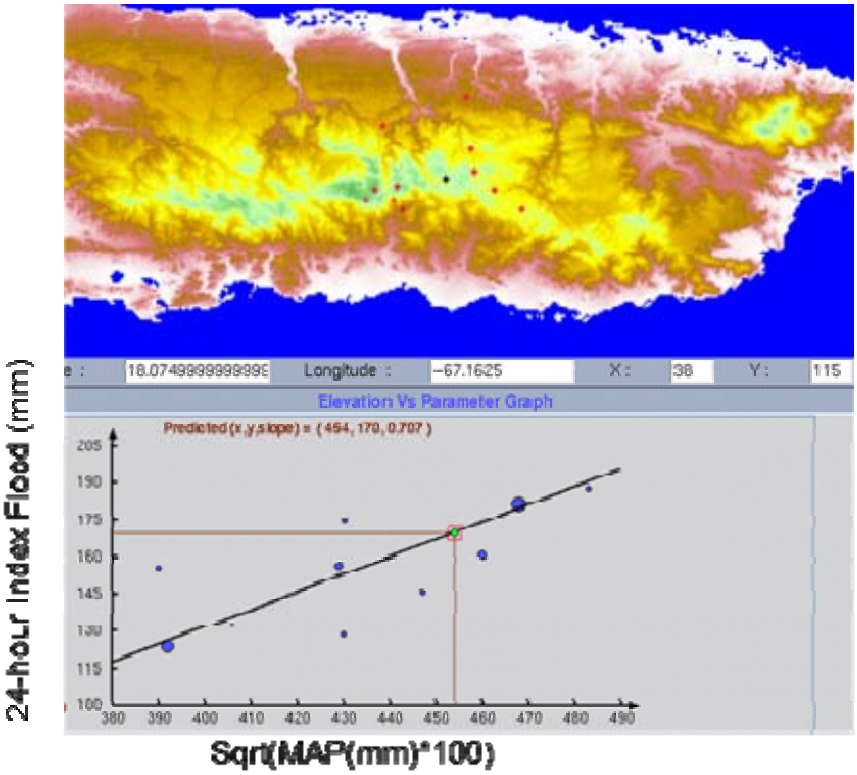


Figure 3. Distribution of station data in the PRVI region: (a) 60-minute; and (b) 24-hour index flood intensities.

(a)



(b)

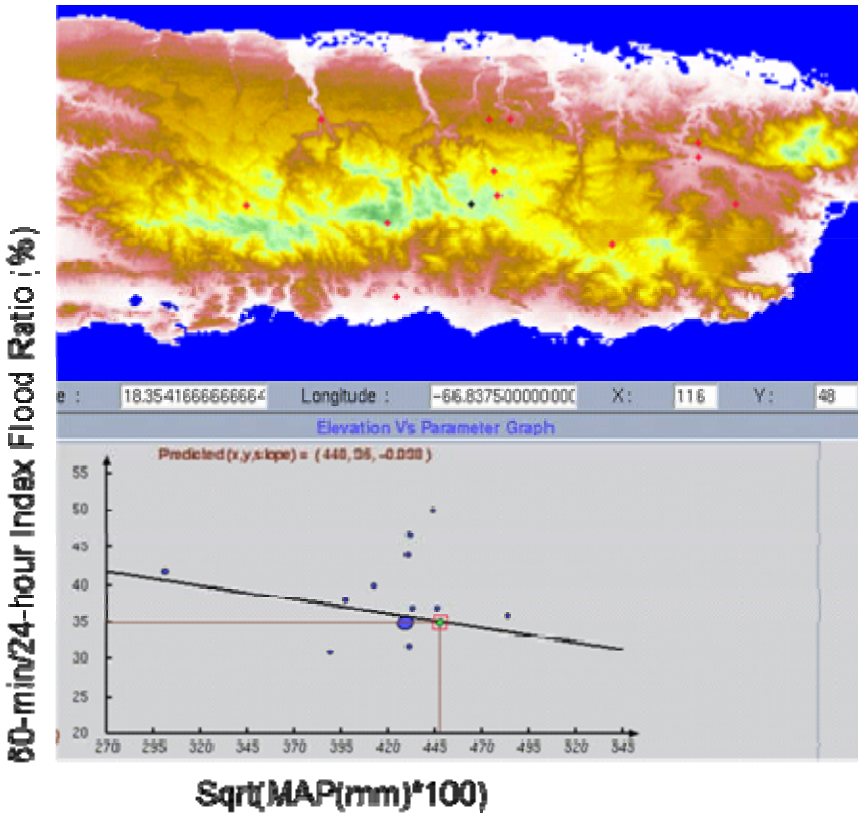


Figure 4. PRISM GUI snapshot of the moving-window relationship between: (a) MAP and 24-hour index flood; and (b) MAP and the 60-minute/24-hour ratio (R24), in the Cordillera Central (Central Mountains) of Puerto Rico.

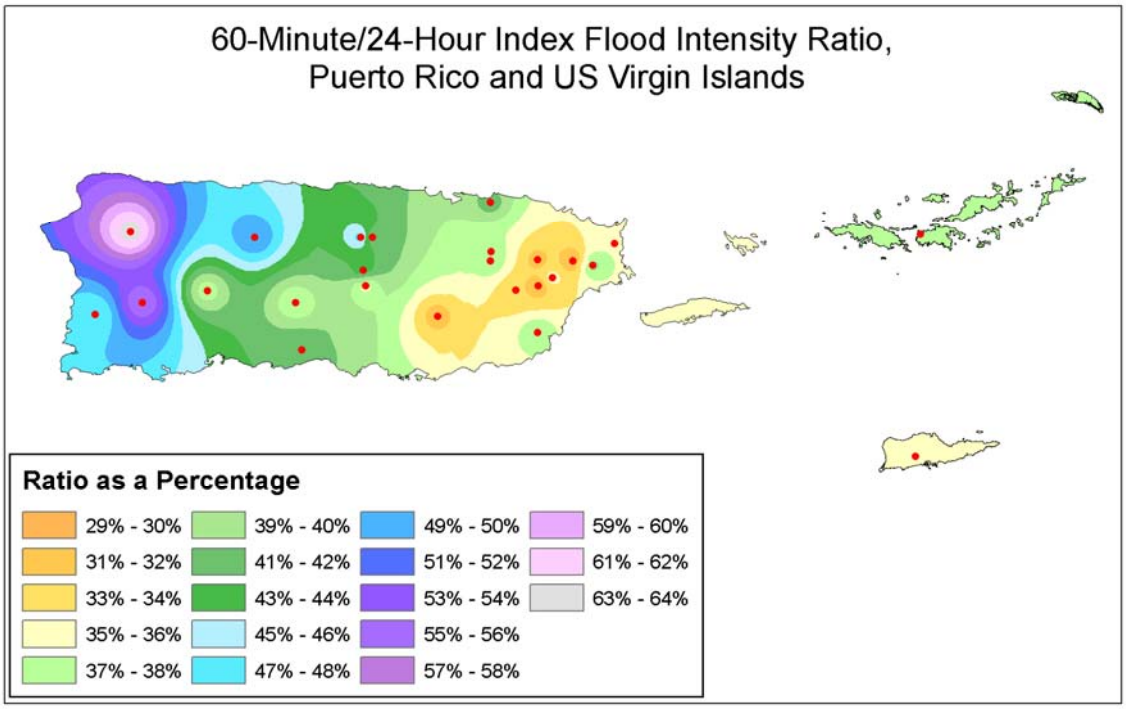


Figure 5. PRISM grid of 60-minute/24-hour index flood intensity ratio (R24) for the PRVI region.

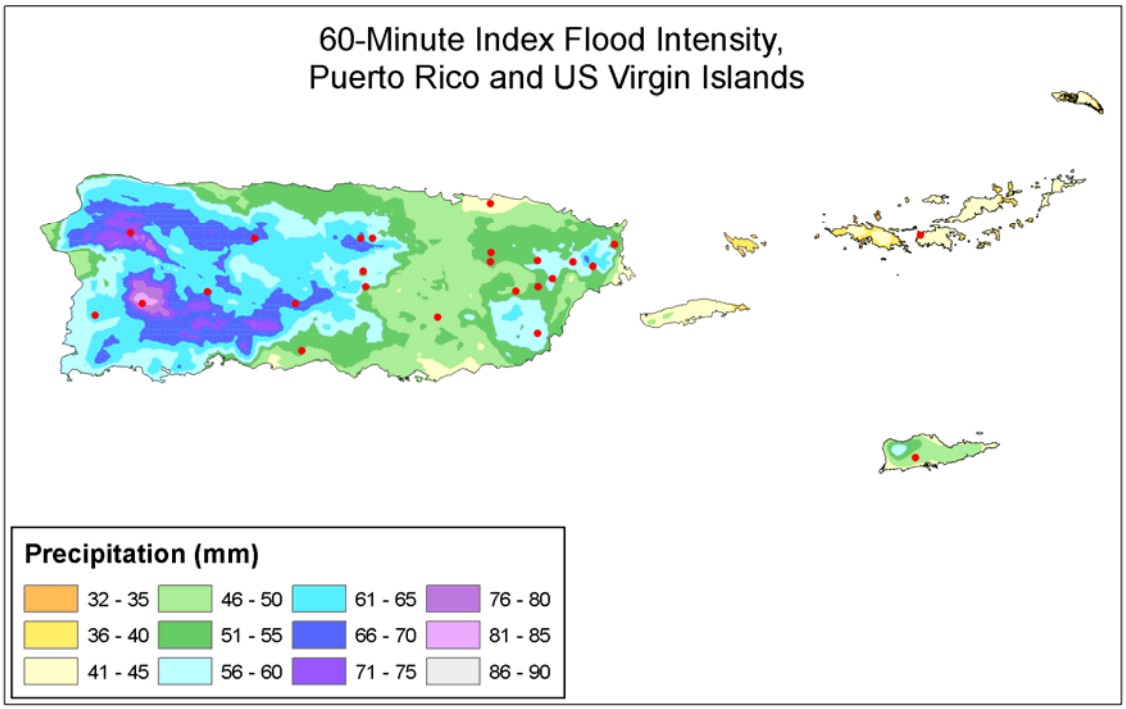


Figure 6. PRISM grid of 60-minute index flood intensity for the PRVI region.

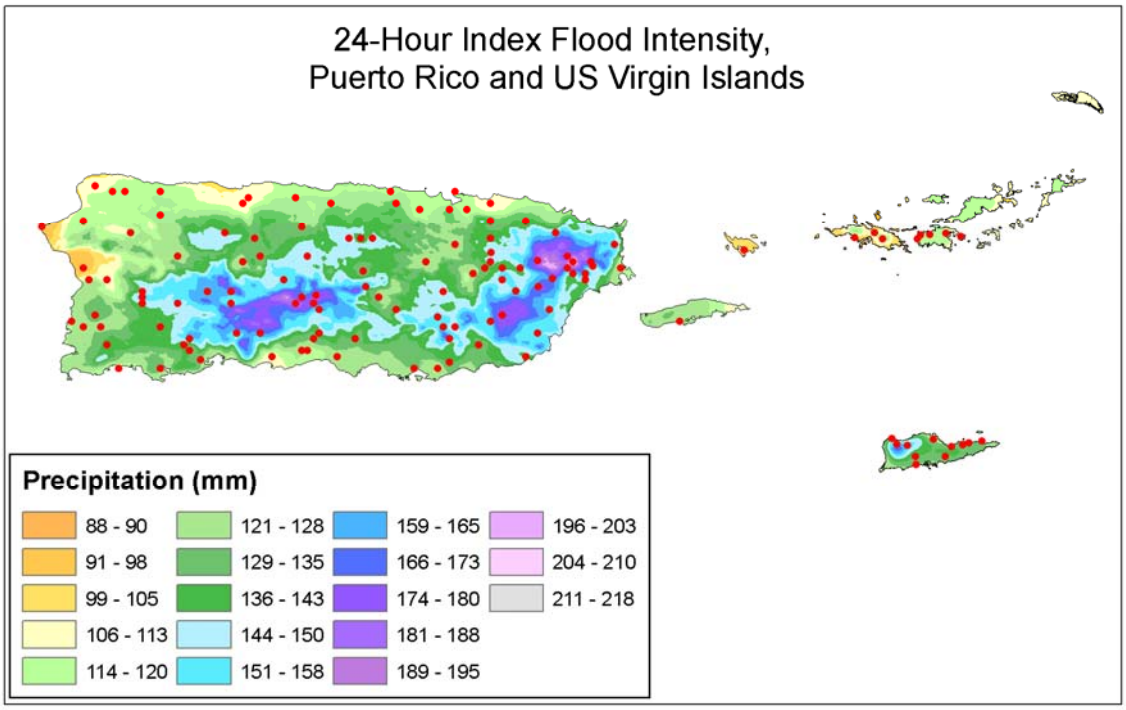
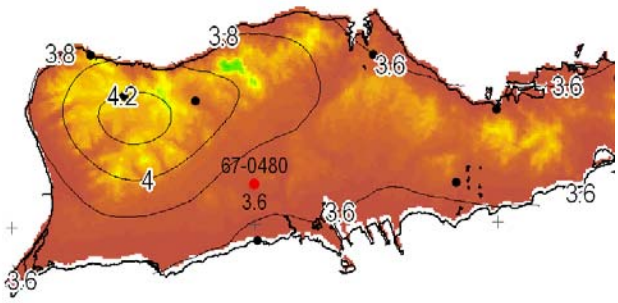


Figure 7. PRISM grid of 24-hour index flood intensity for the PRVI region.

APPENDIX A
External Review Comments and Responses

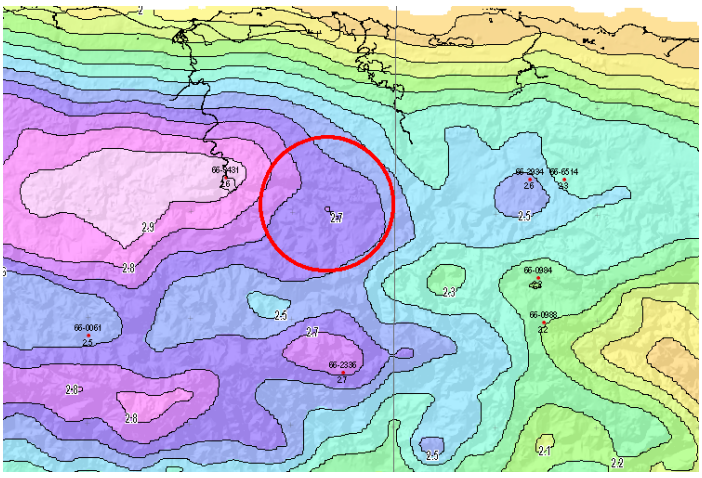
1. Question: Are we correctly capturing all of the high precip (at 60m & 24h) areas on St Croix? It appears as though the maximum should maybe extend slightly more eastward to pick up additional high terrain.

   DEM overlaid by 100-yr 60-min:



   Response by Chris Daly: That is difficult to say, due to lack of data. The PRISM MAP map shows a relatively sharp extension of maximum values over the terrain question, and the 24-hour index flood map shows a similar extension. There is only one 60-minute precipitation station on St, Croix, giving us little to go on at that duration. Recall that the MAP map was generated by PRISM using a DEM that was filtered to a 5-minute effective wavelength, which does produce relatively smooth patterns. Typically, for topical islands, I have been using a DEM filtered to a wavelength of 2.5 to 5 minutes. The precipitation patterns to not appear to respond to small-scale terrain variations, but we frankly do not have the high-quality, high-resolution precipitation station data to be able to reproduce the actual patterns with a high degree of confidence.

2. Question: What is driving up the 60-min means in this area? We expect lower values since it's on a down slope side and at a lower elevation. 66-3431 station value is 2.60 and PRISM interpolated 2.95. To be exact, PRISM interpolates it to be 2.93, but still that is a 12.8% difference; this is the biggest deviation in the dataset. The average difference is 1.01%. Perhaps Chris can force this to match a little better and not gloss over these valleys.



Response by Chris Daly: This question is not easily answered, and may lead to a change in our interpolation methodology for durations of 12 hours and less.

3. Question: How are the Islands with no hourlies interpolated? It does not seem consistent or relies too much on mean annual? For instance, the Vieques value for the entire Island seems to be 2.1 while Culebra is 1.4 for the entire Island. How can we justify it with no stations?

   Response by Tye Parzybok: PRISM interpolated the mean annual precipitation on these islands, and for our work, PRISM is simply using the relationships from nearby islands (with hourly data) and PR to interpolate the means on these islands. Are we OK with this, well, good question. I checked the PRISM values against TP-42 and they are comparable. PRISM has Culebra at about 2.7" and TP-42 (2-yr) has Culebra at 2.3-2.8. PRISM has Vieques at about 1.5-1.6", TP-42 (2-yr) has Vieques at about 1.9". So Vieques might be too dry, but Culebra is about right. Regardless, I will add this to my list of things to consider in the final run.

   Response by Chris Daly: Tye is correct, the values here are dependent on the relationships between 60-minute precipitation and MAP from stations on Puerto Rico. Lacking corroborating data, it is unknown whether this is a good idea. However, there are 24-hour observations on Culebra and Vieques. Comparing the relationships between the 24-hour index flood and MAP on each island, we see that the ratio is exactly the same: 0.12. If I use the PRISM GUI and pick off the 60-minute values at these station locations, we again get identical ratios, this time 0.05. This gives us some confidence that at least the relationships are consistent.

	MAP	*24 hour (obs)/MAP ratio*	*60-minute (estimated)/MAP ratio*
Vieques	*1051mm*	*124/1051=0.12*	*52/1051=0.05*
Culebra	*842mm*	*100/842=0.12*	*41/842=0.05*

4. Question: On Vieques 9793 is 4.9” but the east end of the Island drops way off to 4.3” based on no stations. Is PRISM handling this OK? Seems like too much of a drop off to me.

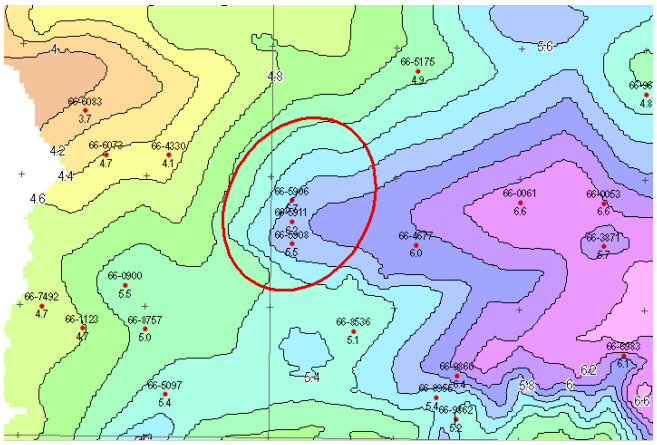
   Response by Tye Parzybok: The TP-42 maps hint at a gradient like this (wetter west, drier east), but yes, perhaps not as steep as PRISM has it. We, nor did TP-42 have any data for the eastern part of Vieques, so we're all guessing to some extent. I wonder if Culebra's dryness is influencing this gradient. Again, I will add this to my list of things to look at again.

   Response by Chris Daly: This appears to be driven largely by a drop in mapped MAP from about 1050 mm on the west side of Vieques to about 950 mm on the east side. The lower values appear to be influenced slightly by lower values on Culebra. The10% drop in MAP results in a 12% drop in 24-hour precipitation, which, again, is probably well within the error margin of the observations.

5. Question: Local experts (peer reviewers) believe this area on the 24-hour mean map is too high. Can you reduce this interpolated/extrapolated maximum a little?

Response by Chris Daly: Yes, I am aware of this hill, and agree it does seem a little high. This hill is on the southern edge of a steep rain shadow transition from the relatively wet Cordillera Central to the relatively dry southern coast. Part of the reason for the relatively high value is station 66-4126, just to the southwest of the hill in question. If we compare this station to 66-1634, (a station used in the MAP mapping only, a little southeast of 66-9774), further to the north and also just to the southwest of a hill of lower elevation, it has a very similar MAP value (66-4126=1268 mm, 66-4126=1224 mm. PRISM sees the hill in question as being in a similar precipitation regime as the hill to the north. However, the hill in question is physically removed from the main mountain chain, and should, therefore, have a little less increase in precipitation per unit elevation than further to the north, due to less influence of blow-over precipitation. There is no easy way to force the model to do this, so MAP will have to be remapped. I will attempt to do that.

6. Question: Local experts (peer reviewers) believe this area on the 24-hour mean map isn't high enough:



Response by Chris Daly: This is an area where there are three stations in close proximity, with different 24-hour index flood values: 5.7", 6.2", and 5.5". The modeled values are approximately 5.6", 5.7", and 5.6", respectively. This translates into differences of approximately 2%, 8%, and 2%, respectively. These are not large differences.

Appendix A.5

Peer Review Comments and Responses
Puerto Rico and the U.S. Virgin Islands

NOAA's National Weather Service
Office of Hydrologic Development
Hydrometeorological Design Studies Center
Silver Spring, Maryland

15 February 2006

Introduction

The Hydrometeorological Design Studies Center (HDSC) conducted a comprehensive peer review of the Puerto Rico and U.S. Virgin Island precipitation frequency project during the period November 3, 2005 to January 11, 2006. The review included the following draft items:

1. Point precipitation frequency estimates
2. Mean annual maximum 60-minute map
3. Mean annual maximum 24-hour map
4. 100-year 60-minute map
5. 100-year 24-hour map
6. Statistical Trend Analysis report
7. Temporal Analysis report

This document presents a consolidation of all the review comments collected during the 10-week review period and HDSC's response. We have used the original wording of the comments to make sure the meaning of the comment/question was not misconstrued and so that individual reviewers can identify their comments. HDSC requested comments from approximately 115 individuals and we received comments from 14 individuals. Some of the responses represented feedback from their staff. After parsing all of the comments, we found 47 unique comments which are included in this document.

Similar issues/comments were grouped together and are accompanied by a single response. The comments and their respective responses have been divided into seven categories:

1. **Comments on point precipitation frequency estimates**
2. **Spatial interpolation comments**
3. **Comments pertaining to statistical trend analysis report**
4. **Comments pertaining to temporal analysis report**
5. **General questions, comments and typographical errors**

1 Comments pertaining to point precipitation frequency estimates

1.1 *For daily stations (24-hour plus data only, e.g., Canovanas, Mayaguez AP, Humacao) you may not want to extend duration curves in precipitation-duration graphs below 24 hours*

where there is no data to support a line being drawn. Similarly (e.g., Cubuy, Botijas 1 Orocovis) do not extend the curves at the durations beyond 48 hr. where no data exists.

It seems like this area would have small streams that are responsive to short duration rainfall events. The tables I looked at had as the shortest duration 24-hr durations, which seems long for this type of area. I expect that you were limited by data availability, but is there no way to use creative science to extrapolate to shorter durations? If you, who have the scientific tools and knowledge to do so, do not generate shorter duration values, those who need the values will have to try to guess at some values... likely with less success than a scientific approach.

Are there any other frequencies besides the 100 year and mean values [maps]*?*

Response: The preliminary information available for the peer review is a subset of the final information to be published resulting in gaps when displayed on the Precipitation Frequency Data Server. The final NOAA Atlas 14 Volume 3 results will provide frequencies of 1-year to 1000-year and durations of 5-minutes to 60-days for all locations in a point-and-click interface and as grids and maps.

The point estimates under review were directly computed from observed data available. Hourly stations had data for 5-minutes through 48-hours. Daily stations had data for 24-hour through 60-days. Stations with complete data (5-min through 60-day) had co-located gauges and we designate them as co-located hourly-daily stations. Once the spatial analysis is complete, all durations/frequencies will be available for all locations; there will be no gaps in the data (i.e. missing durations) at any point in the project area.

All of the input data into the spatial interpolation scheme is based on the duration bounds of the station data (i.e., daily, hourly, co-located daily and hourly). Any duration not represented by station data will be automatically interpolated. The spatial analysis will be described fully in the final documentation.

1.2 *I pulled off this type of preliminary information from the [Precipitation Frequency] Data Server for some 12 to 15 stations. Some of the precipitation depth vs annual exceedance probability plots look great whereas others indicate a large amount of durational convergence/divergence. I take it that these plotted relations will be smoothed in the final analysis?*

1.3 *Is there reason for or previous studies that showed similar flat slopes after 48-hours, for instance at Cerro Maravilla (66-2336)?*

Response: Yes, the plots you saw during the review represented the raw results from the L-moment software. The spatial interpolation scheme will smooth the raw results both temporally and spatially as to make them consistent with neighboring stations and other durations. Even so, some of the general patterns observed in the curves may remain after the smoothing. We considered these patterns carefully and found climatological justification for them. We included this climatological summary as part of the peer review at http://hdsc.nws.noaa.gov/hdsc/pfds/pr/pr_IDF_climatology_final.pdf.

1.4 *It seems reasonable to believe that Cacaos (for whatever reason) is an outlier and we tend to agree that it should be discarded. I really hate to discard a station with such data but we do not find a justification for this station collecting a significant amount of rainfall higher that its neighbor stations.*

Response: The observations and precipitation frequency estimates at this station were high, did not seem consistent with nearby stations, and was therefore on our suspect list. After receiving this comment and learning that the observer at the Cacaos station once had a tarp that drained toward the rain gauge, we deleted the station from the analysis.

1.5 *It is interesting and should be investigated, why we do not have a bulls eye maxima in the central northwest near the Maricao Fish hatchery station (66-5911). It always shows up in the climate data analysis and maps. It is a mountainous area. Can you check?*

Response: We investigated the Maricao Fish hatchery station, as well as the two nearby stations: Maricao (66-5906) and Maricao 2 SSW (66-5908). Although the means and precipitation frequency estimates at the Maricao Fish hatchery are slightly lower (hence preventing a maximum on the maps) , they are within acceptable limits as compared to the other two stations. Given the consistency in the observations and results at these three stations, we will not make changes to the regionalization or station data. However, we will be mindful of this region during the final spatial interpolation.

2 Comments pertaining to spatial interpolation and maps

2.1 *USVI – St. Thomas: Wintberg (elev. 600 ft) has 2-yr 24-hr value nearly 25% lower than Ft. Mylner (167 ft) (2.78" vs 3.65"). Without access to a high-resolution topographic map the rationale for such a difference escapes me. Wintberg also seems quite low compared with all stations on the other USVI, Culebra and Vieques, unless this is a very "shadowed" stations from the prevailing trades and with respect to hurricane winds which would likely have an easterly component as well in the heaviest rain scenarios.*

Response: The data at these stations supports the estimates. Nearly all of the annual maximums are higher at Estate Fort Mylner (67-2823) than Wintburg (67-9450) during their concurrent period of record from 1972 to 1995. The data at 67-9450 extends further through 2003 and is similar to its earlier record. The highest observed annual maximum for the period of record at 67-2823 was 15.00 inches on 4/18/1983 while 67-9450 observed 9.11 inches on 4/19/1983. However, the 2-day totals from this event were closer in magnitude with 17.50 inches at 67-2823 and 15.01 inches at 67-9450. With the relatively short period of record at these two stations differences such as observation time make a larger difference in the estimates than if the stations have longer periods of record. The records at these stations meet our criteria and do not require a change in our approach. In addition, maps of elevation reveal that 67-9450 is on the western side of an elevation barrier of almost 1,000 feet (305 m) while 67-2823 is on the east (upslope) side of this elevation barrier. This supports the existence of a rain shadow at 67-9450 with an easterly wind. Therefore, we are confident in the estimates at these two stations.

2.2 *Surprisingly large differences in pf values even at low elevations (Mayaguez Airport at 32 feet vs. Mayaguez City at 203 feet) where over 20% variation in 2-yr 24-hr. value (2.76" vs. 3.52"). Hacienda Constanza (elev. 925 feet), which might be presumed to have more*

topographic influence falls midway between these values at 3.05". It is unlikely that there are substantial moisture availability variations in these situations. However, there may be small-scale phenomena which account for these "discrepancies".

Response: Mayaguez City (66-6073), with 104 years of data, has twice the period of record of Hacienda Constanza (66-4300) and Mayaguez Airport (66-6083). However, all stations meet our minimum requirement of at least 20 years of data to be included in the analysis. During the concurrent period of record many of the annual maximums were higher at 66-6073 than 66-6083 from the same event. For example, at 66-6073, 21.30 inches was measured on 9/22/1998 while just 12.00 inches was measured at 66-6083. Therefore, since 66-6073 and 66-6083 are clearly measuring different rainfall, we prefer not to make any changes to the regionalization or station data.

Even so, the spatial interpolation scheme appears to be accommodating this variation through the use of PRISM and smoothing algorithms. PRISM (and our unique use of it for precipitation frequency estimates) was specifically designed to incorporate climatological knowledge into its process. You'll notice that the contours generated by the spatial interpolation scheme are more in line with your expectations than the point data, keeping the highest values near Hacienda Constanza (66-4300) and the lowest near the Mayaguez Airport (66-6083). This scenario illustrates how using the regional L-moments method in conjunction with our spatial interpolation scheme produces reliable results that are consistent with the climatology.

2.3 *The map annual maximum 60-minute rainfall shows a broad maxima of 2.8" plus in the northwest quadrant of the island that is not reflected in the 60-minute exceedance map which even depicts a slight minimum (<4 inches) southwest of Dos Bocas (66-3431). Not sure if this is a result of the PRISM analysis, but appears slightly anomalous.*

Response: This was an observation we made as well. The smoothed spatial interpolation by PRISM at 66-3431 is actually the highest deviation from the observed data. We are going to review this area carefully in our final PRISM runs. However, we feel that the resulting patterns are likely reproducing the local climate regimes. The northwest quadrant of Puerto Rico is susceptible to frequent moderate rainfall events from thunderstorms hence the mean annual maximum is relatively high in this area. But rarely do these thunderstorms drop extremely heavy rainfall, which is reflected in the relatively low 100-year estimates.

2.4 *The 24-hour 100-year current analysis looks much like the general pattern indicated by the annual maximum 24-hr precipitation that was provided. However, this is not the case when comparing similar figures for the 60-minute duration especially for western PR. Is there a reason why the 24-hour looks so reasonably similar but there are significant differences at 60-minutes?*

Response: We carefully reviewed this issue and feel the 60-minute patterns are consistent with the climate regimes (see Response 2.3). During the initial internal evaluation of spatial results, it was clear that the spatially limited shorter duration dataset (hourly) was not sufficient to accurately resolve patterns at the final high spatial resolution (3-sec). Therefore, so-called hourly "pseudo data" were objectively generated at all daily-only stations to create a more coherent spatial pattern in the hourly durations over such complex

terrain. The general 100-year 60-minute patterns did not change with the addition of the pseudo data, with one exception – in the middle of the island, estimates in/around 66-9432 (Toro Negro Forest) increased 18.2%. This increase occurs in an area that lacks hourly observing gages. The resulting spatial pattern is more consistent with climatological expectations and with Technical Paper 42. Additional information can be found in the 22nd Progress Report Section 3.6 Spatial Interpolation, Spatial Smoothing (http://www.nws.noaa.gov/ohd/hdsc/current-projects/PuertoRicoPR22.pdf). To further verify that the pseudo data were not changing the overall pattern, we compared the 60-minute mean annual maximum map that was generated without any pseudo data with the 2-year 60-minute map, which is statistically closest to the mean, and that was generated with pseudo data. We observed that the pseudo data did not change the overall pattern between the compared maps. We have concluded that the use of hourly pseudo data in the quantile maps adds climatologically sound definition to areas of complex terrain that are devoid of hourly observations.

2.5 *The exceedance probability curves do seem to reflect nicely some of the complex climatological patterns, such as the continued increases at longer durations in the El Yunque (Pico Del Este) area versus the central or western mountains (Adjuntas, Cerro Maravilla). I do question the fairly low short-duration values at El Yunque stations compared to the central mountain (2-yr. 60 min. value of 1.98" at Pico vs. 2.33" at Cerro, 2.44" at San Sebastain). Convective rainfall at such short-time scales would seem have more to do with storm-scale factors rather than mesoscale or terrain-induced phenomena. The paucity of hourly data leads to some questions at least about the physical mechanisms responsible for short duration (< 60-mins.) rainfall. In contrast to the longer duration events, elevation does not have a dominant role in short-duration rainfall as shown by the fact that San Sebastian (elev. 173) has higher 5 to 60 minute values than Cerro Marravilla (elev. 3684). It is perhaps beyond the purview of this report to address the causative factors here, but the isopluvial patterns drawn on the 60-minute maps may reflect the few stations available far more than actual variations in short-duration rainfall. The 19th Progress report discussed the reasons for the elimination of the USGS short-term data. That data may have still been useful in depicting the spatial distribution of short-duration rainfall in this complex terrain, which is so dependent on a few stations. Perhaps archived radar data from the San Juan WSR-88D would also be useful in depicting the rainfall distribution in this area, although I realize that is another entire project.*

Response: We agree short duration rainfall is influenced more by storm-scale factors rather than mesoscale or terrain-induced phenomena. In fact, the smooth interpolation of the mean annual 60-minute maximums is consistent with this theory. Although we are using a limited hourly dataset, it may be sufficient for resolving the smooth spatial patterns in the means because fewer stations are necessary to support such patterns. Additionally, the interpolation of the 60-minute means indirectly utilizes a much denser network of stations through the use of the PRISM-created mean annual precipitation, which is based on a larger dataset. As described in the response to 2.4, pseudo hourly data were developed at all daily-only stations to supplement the spatial density of the hourly quantiles for interpolation. Unfortunately, a radar study is beyond the scope and schedule of this project, but past investigations into radar use for our purposes has been inconclusive due to the relatively short periods of record and the absolute and spatial errors in radar fields, particularly in cases of extreme precipitation such as are not the subject of this project.

2.6 *For the 100-year 60-minute & 24-hour analysis for the island of St Croix, I note a rather large increase in the magnitude of precipitation in the NW portion of the island. Does one have station values to support this large increase? Your topo analysis provided does not give me the impression that the orographic enhanced precipitation should be that great.*

Response: There are several things influencing the high values on the NW portion of St Croix. (1) There are two stations (Annaly, 700', 27 years; Fountain, 250', 20 years) supporting the high values, (2) the elevation in this area reaches 1,000+ feet and does provide some orographic enhancement, (3) the PRISM-created mean annual precipitation map, from which the precipitation frequency maps are indirectly derived, indicates elevated values in this area. In fact, one may argue that the high values ought to extend slightly further east-northeast to pick up additional high terrain. Regardless, we will be mindful of this area during the final spatial interpolation.

2.7 [The area just northeast of] *station 66-4126 (Guayabal) to the southwest of Cacaos station has a maxima of 22 which also seems to be to high. There is a hill in that area but should not be that significant, could you also check on this one..*

Response: The lack of data in this area forces the estimates to be extrapolated up the terrain by the spatial interpolation scheme. The relationship between the mean annual maximum 24-hr precipitation and the mean annual precipitation suggests this area experiences relatively heavy precipitation. However, based on your concern, we will look at this area again during the final interpolation.

3 Comments pertaining to statistical trend analysis report

3.1 *The report does not mention the Atlantic Multidecadal Oscillation (AMO) and its relationship, if any, to rainfall in Puerto Rico and the USVI. Perhaps, describing such atmospheric phenomenon compared to the period of record of the gages utilized in this report may further explain the cause of the positive trend as shown in the report.*

Response: That is a good suggestion, but beyond the scope of this project. The purpose of the trend analysis is as a quality control measure to demonstrate that the annual maximum series data is independent, generally free of trends, and therefore appropriate for use in the precipitation frequency analysis.

3.2 *The title on figures should be posted at the bottom of the figure and not at the top. The maps included in figure A.3.1 and A.3.2 are very small and is difficult to identify the location of the stations.*

Response: Agreed, the titles have been moved below the figures. We will consider ways to make the maps clearer.

3.3 *In the second row of the introduction the appropriate word is stationary instead of the word constant. If the climate is stationary implies that the first and the second moments of the time series do not change with time, i.e., the autocorrelation function is independent of time.*

Response: Thank you we have changed it in the text.

3.4 *I'm not entirely certain what the value of the Statistical Trend Analysis is. It does not appear to definitively answer any question- the statistics are (as far as I can tell) profoundly inconclusive on the subject of trends. I suppose that I would encourage some discussion of what this is intended to convey. A small number of stations indicate some possibility of a positive trend over the period of record. I saw no discussion of whether or not that could have been attributable to measurement bias or systematic error (I would suspect that given no other evidence). I also saw no discussion of how the period of record might exert an impact on that. Unless I am mistaken, there were periods of distinct difference in the number of tropical cyclonic events over the twentieth century, particularly what appeared to be a lull during the latter half. How might things like that impact the statistics? Should they be considered as anomalies in a stable long-term process, as indicators that there is no overall stability in climate, or (as the press is likely to do) indicators of some anthropogenic impact on climate. I don't expect anyone to know the*

answer to these questions, but if you analyze for trends and publish the results of that analysis, shouldn't you at least discuss what lead you to do so? As it is, I saw the trend analysis as not contributing to the overall atlas in a meaningful way. I suppose that what I am saying is that the Introduction and/or conclusions sections should be fleshed out somewhat.

Response: The current practice of precipitation (and river height and flow) frequency analysis makes the implicit assumption that past is prologue for the future. The purpose of our trend analysis is as a quality control measure to demonstrate that the annual maximum series data is independent, generally free of trends, and appropriate for use in the precipitation frequency analysis. Therefore, based on the results we were able to assume that the full period of the available historical record derived from rain gauges was suitable for use in this analysis even though there were some local instances of linear trends and shifts in mean in the data. This point is actually discussed in various other sections of the final documentation (which were not part of the review) and directs the user to refer to Appendix A.3 (the report reviewed here) for additional details. However, you make a good point and perhaps it would be good to include additional text covering this further in the Appendix as well.

Discussion or research into the causes of any observed trends is beyond the scope of this project.

3.5 *Regarding the time series trend analysis it would be useful to see a table showing the trend data at all the 55 available stations. This would not be too onerous a task if the station, percentage change and years of data were all that was listed. This would permit researchers to assess whether changes in station location, equipment, or observer were also partially responsible for trends and changes. Figures A.3.4 and A.3.5 were very interesting and it would be nice (albeit unlikely) to see more of these. The 30% increases at these stations seem very high since the most (42 of 55 showed) no trend at all and many if not most stations would have experienced the "outlier" hurricane events.*

Response: A list of the stations could certainly be provided upon request but will not be included in the final documentation to maintain consistency between the available documentation for all Volumes of NOAA Atlas 14. As you say, Figures A.3.4 and A.3.5 are provided as interesting examples of the types of trends/shifts observed in the data. Time and resources preclude us from creating any additional figures. The time series data will be available on the Precipitation Frequency Data Server and trend/shift statistics can be provided if one would like to conduct a more detailed review of additional stations in Puerto Rico.

3.6 *In the paragraph 2.1 Methods in row 7 the variable should be clearly defined. This variable is either the mean of the total rainfall observations or the mean of the only 1-day maximum precipitation time series.*

Response: The variable is the mean of the 1-day annual maximum series data. It has been clarified in the text.

3.7 *Section 2.2, first paragraph. Length of record is based on the 50 year criterion + additional criteria, right? Same section, last paragraph. This is a repetition of the first*

paragraph of this section.

Response: Right. The stations used in the analysis were required to have at least 50 years of data and meet various criteria if there were any gaps in the record. There are similar wordings in the results descriptions, but they vary depending on the particular test.

3.8 *Section 2.1 Methods, second paragraph. (Start with Stations with gaps….) add "the following" before "additional criteria"*

Response: Thank you for the suggestion. The text has been added.

3.9 *Section 3.1 Methods. It is not clear what a "division of 1958" means.*

Response: The text was clarified to indicate that the year 1958 was used to divide the time periods that were tested for a shift. A division at 1958 was chosen because that was the end date for the previous estimates published in 1961. It allows us to test whether there are differences between the data used in the previous publication and the additional data we have used in this project.

3.10 *Page A.3-5, table A.3.3 and figure A.3.3. The table indicates there are eight stations that show a positive shift in the mean results yet on figure A.3.3 there are only seven such stations plotted. Am I missing part of the figure or has one station been left out on the figure*?

Response: Although all stations were plotted, the label for one was covered up by another label. We will correct the document.

3.11 *Page A.3-4, section 3.1, last sentence. It is stated " In this project, there were 20 stations that were screened out (not eligible) for the Mann Whitney test that were included for the t-test." Yet in section 3.2 (see bullets), you indicate that for both the t-test and the Mann Whitney test there are a total of 22 stations eligible. Where is this difference of 20 stations. Either the t-test should show 42 stations available and 22 for the Mann Whitney test OR there are 22 stations available for the t-test and only 2 stations available for the Mann Whitney test. Please clarify text*.

Response: You are correct, all stations eligible for the t-test were also eligible for the Mann Whitney test. The erroneous statement was actually a remnant of a previous trend analysis report written for NOAA Atlas 14 Volume 2. The text has been corrected.

3.12 *The south bias on the analysis looks different than I expected, but as you explained in the review materials it is probably because of single events such as 1985 and Georges. I would like to know if you have had that comment.*

Response: We presume you are referring to the geographical preference for upward trending stations in the south. The extent of our research and knowledge about this is contained in the report. And as we've said, unfortunately, we do not have time or resources to investigate causes of the trends further.

3.13 *In the paragraph 3.1 Shift in the mean results in row 7. There may be another justification of change in the mean of the process. The global warming may be a better justification to determine the shift in the mean. The limitation of having at least 30 observations on each side of the change it is not required all the times; it depends on the method of analysis. For instance the exponential weighted moving average method requires having at least 30 observations on the left side and 1 to 10 observations in the other side. Ramirez and Julca (2006) introduced an algorithm to detect local climate change. The algorithm is vary simple and is based on the fingerprint of an autoregressive process. Some applications are presented in this manuscript.*

Response: We appreciate your suggestion and will review the paper which you reference. However, given time and resource constraints it may be difficult to implement a change in the methods for this project. We will give it consideration and may be able to incorporate modifications in our future projects.

3.14 *In the paragraph 4. Specific Examples at row 11. Please replace 1996 by 1998, because hurricane Gorges occurred on September 22, 1998. This mistake was repeated on the next paragraph.*

The following document needs to accurately show the year of Hurricane Georges, which was 1998, not 1996. The error is in two instances. Statistical trend analysis of 1-day annual maximum series - via the review page or directly at http://hdsc.nws.noaa.gov/hdsc/pfds/pr/NA14Vol3_A3.pdf.

Response: Thank you for catching the typographical mistake. It has been corrected.

3.15 *Section 4. Specific examples. Second paragraph (Starts with "Figure A.3.5…" I don't understand how the data still exhibits a trend in the mean, but not a shift in the mean.*

Response: This station, Ponce 4 E (66-7292) does indeed exhibit a shift in mean. This will be made more clear in the text and figure caption.

3.16 *Page A.3-7, title of figure A.3.5. I do not understand that this plot indicates a decreasing linear trend. Looks to me like the trend is increasing - check the title of this figure.*

Response: Yes, that was a typographical error. It has been fixed to indicate an increasing linear trend.

3.17 *The legend on Figure A.3.4 (Failed to pass …) should be removed from the figure.*

Response: Agreed. It does not add information to the plot. It will be removed from Figure A.3.4 and A.3.5.

3.18 *Page A.3-6, section 4. It would be helpful if one would include a general descriptive location of the two stations provided as examples as part of the text.*

Response: Good suggestion. Text describing the general location has been added.

3.19 *The x's axis of Figure A.3.5. is shifted one unit. The year 2000 should coincide with the number 100. Apparently year 2000 is located in the number 101 in the x's axes.*

Response: This is the result of the convention used during the plotting. The first year of record is plotted as year 1, not year 0. Therefore, given a record starting in 1900, the year 2000 would be located at year 101.

3.20 *Figure A.3.5. does not reveal the effects of hurricane Hortense. This hurricane made a landfall in the south part of Puerto Rico in September 9, 1996 and caused flooding in Ponce, PR. Figure A.3.5 shows approximately 3.5 inches, which looks very small. I would recommend checking this particular value.*

Response: We reviewed this particular value and could not find error in it. Our data for this event is verified through the original surface observation form for September 1996 at Ponce 4E (66-7292) obtained through NOAA's password-protected "Web Search Store Retrieve Display" (WSSRD) (http://noaa.imcwv.com/) which shows a total of 5.22" fell during Hurricane Hortense (3.46" + 1.76"). This also agrees with the National Hurricane Center's preliminary report for Hurricane Hortense (http://www.nhc.noaa.gov/1996hortense.html). Without additional information, we are inclined to believe the station observations. The flooding you mentioned was perhaps a direct result of higher precipitation amounts (more than 16 inches) that fell in the mountains to the north during the event (see Figure 3.3.1 from the National Hurricane Center's report).

Figure 3.3.1. National Hurricane Center's map of rainfall totals for Puerto Rico during Hurricane Hortense.

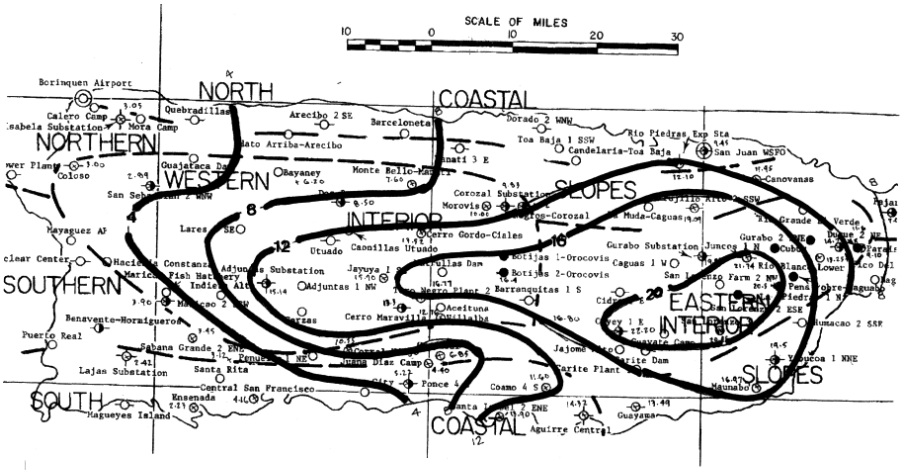


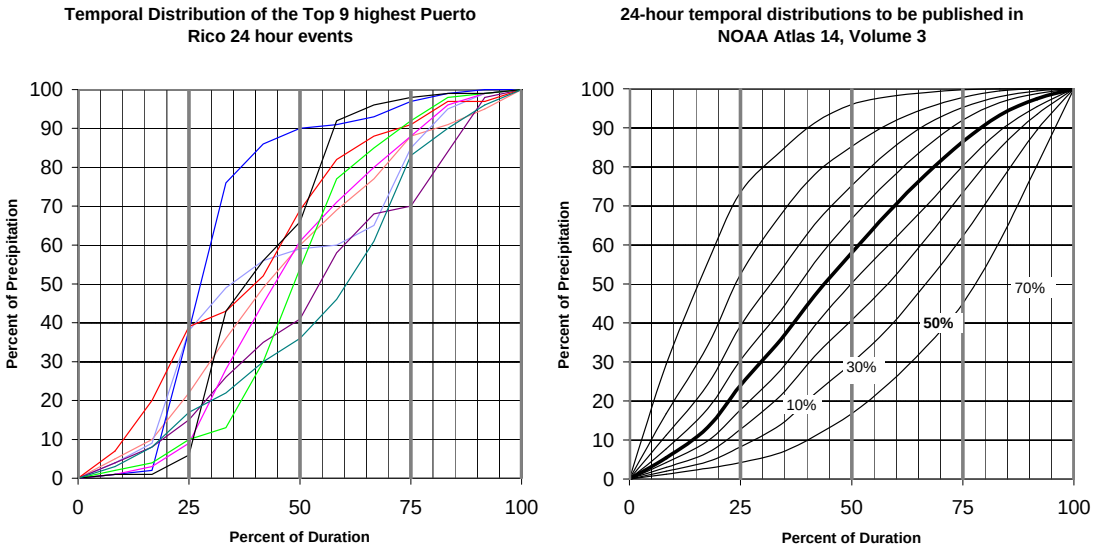
Fig. 5 Hurricane rainfall totals (inches) over Puerto Rico during 9-10 September 1996.

4 Comments pertaining to temporal analysis report

4.1 Regarding the temporal distributions of heavy precipitation, the methodology is sound but it would be of interest to see at least some data from more extreme events rather than just the 2-year ARI. Perhaps using a 10-yr ARI if that sample was stable and for just a single duration. The interpretation section (3) is useful but it might be useful to provide a "real-

world" example of a storm distribution. Using a single extreme event would demonstrate how an actual storm distribution might vary (or conform nicely) with the theoretical graphs presented.

Response: The temporal distributions are based on the summary of all types of precipitation events and meant to be used with the precipitation frequency estimates. The temporal distributions were created based on events greater than the 2-year ARI, which includes events up to and even greater than the 100-year ARI. Comparison between distributions based on the 2-year ARI and greater versus 25-year ARI and greater showed no noteworthy difference except that the reduced number of samples resulted in less smooth curves. The "real world" examples are variable and mostly fall between the 10 and 90 percent curves. For instance, the nine highest 24-hour events from all hourly stations in the project area are shown in the figure below on the left. The event totals in these nine cases ranged from 22.20 to 17.00 inches. Notice the variability and lack of smoothness; yet the curves are generally within the 10 and 90 percent curves shown on the right which are the distributions to be published in NOAA Atlas 14, Volume 3. Because of the lack of stability and smoothness, we do not plan to publish this example.



4.2 *The temporal distribution information appears quite consistent with what I've seen in earlier volumes of Atlas 14 and in other recent research. The Huff-type curves are very useful if you understand what they represent. The use of such curves gives a common ground for comparison among such work worldwide. There appears to be adequate documentation and information for any purpose I can conceive at this time. The percentile ranking allows for the consideration of uncertainty in this arena, also.*

Response: Thank you for the comment.

4.3 *Section 2, Methodology, first paragraph. Add "storm" before duration, so it will read: "storms in order to be consistent with the way storm duration was defined…"*

Response: There is a difference between "storm" duration and the specific case durations used in this study. The cases have a specific definition here. The starting time of the duration accumulation was defined in the same way as it was for the precipitation frequency estimates. The start and end times were set to capture the highest accumulation for a duration by a moving time window technique. As a result, the accumulation cases may contain parts of one, or more than one, precipitation event.

4.4 *Page A.1.1, section 2, 4th paragraph, first sentence. It is stated that some regional analysis was made of the station precipitation on a temporal basis and no significant variation was indicated. I would like to be made aware of the sub-regional breakdown that was used for the analysis. Did you break regions by coastal/nonorographic vs interior/orographic or what exactly was done for this analysis. A brief explanation in the text or figure would be helpful.*

Response: The regions used were the same hourly regions used for the precipitation frequency analysis. We will make this clear in the final documentation.

5 General questions, comments and typographical errors

5.1 *The easterly winds during winter time play an important role bringing humidity into the island and stimulate orographical rainfall in the mountain region of Puerto Rico. Cold fronts cover large areas of cloud and precipitation during winter times.*

The easterly waves occur during the rainy season (May to October) and generate large amounts of rainfall in the Caribbean basin. The easterly waves generate low pressure systems in the east part of Puerto Rico creating several days of rainfall, usually 2 to 3 days, due to instability and thermal convection.

Reference: Ramirez-Beltran, N.D., and Jualca, O. Detection of Local Climate Change. 18th Conference on Climate Variability and Change, 86th Annual AMS Meeting. Georgia Atlanta, January 29-February 2, 2006

Response: Thank you for these valuable comments and the reference.

5.2 *Would like to see some discussion about differences in final values between NOAA Atlas14 and the soon-to-be displaced Technical Paper 42 for PR/USVI. The time series trend analysis did discuss the change in mean trends between pre- and post 1958 but I didn't see where or how this might have impacted changes in the precipitation frequency values [between NOAA Atlas 14 and Technical Paper 42].*

Response: In the final documentation, we will include a brief discussion of the differences in the estimates of NOAA Atlas 14 and Technical Paper 42 as well as a map showing the differences between the 100-year 24-hour estimates. It will be found in Section 7, Interpretation.

5.3 *I can't conclude without a jab at pure web-based publishing. I am an old-fashioned person who believes in things I can touch. I am not certain from what I've seen so far whether or not there will be a hard-copy "master document" somewhere, but I hope so. In the end, electronic media are volatile. Over my engineering career, I've seen many examples of electronic record "perpetuation" that ultimately failed due to money or manpower constraints that could not be foreseen at the time of origin. I am skeptical- but I realize I am in the minority and a dinosaur. This is great work, and I'd hate to see it lost in some cataclysmic virtual equivalent to the burning of the Great Library at Alexandria.*

Response: We recognize your concern, but in accordance with the E-Government Act of 2002, NOAA Atlas 14 is an on-line document. The documentation and the maps will be made available in PDF format so that a user can print them if they wish. Some of the other electronic artifacts are also printable. A limited budget is also a contributing factor to any hard-copy versions of the atlas and electronic publication allows us to publish several orders of magnitude more information than we would be able to in hard copy.

5.4 *For the majority of issues, it is obvious that the group performing this work has become very comfortable and capable with the Hosking and Wallis methodology and with the tasks involved in doing extensive analyses like this. Considering the magnitude of effort that must go into this, I must say that I consider this work to be fantastic. I did not try every station or possible interpolation- not being familiar with the area I probably would not have recognized a major blunder anyway.*

Response: We appreciate your comments and thank you for taking a look.

5.5 *I have no personal frame of reference whatsoever regarding the magnitude and reasonability of the actual numbers that come from the Precipitation Frequency Data Server- I must assume that they are reasonable. The web services appear to work reliably and give consistent information in several forms. I am happy to see that the date and time of retrieval are documented on the reports- I think that is an essential detail. Spatial interpolation also appears to work consistently and reliably. I really appreciate the confidence intervals and other representations of uncertainty- that information was sorely lacking in previous work. Its presence gives the ability to enhance the credibility of work based on this information tremendously.*

Response: Thank you for your comments, and we agree, the confidence intervals add a tremendous amount of value and information to NOAA Atlas 14.

5.6 *It would be good if the map showing the rainfall stations also included the watershed boundaries so it could be easily seen which stations corresponded to which watersheds.*

Response: We will consider this as an enhancement in the future. With the final release, we will provide GIS compatible grids of the estimates so a user can easily download the data and overlay such boundaries.

5.7 *I noticed also that your precip duration/frequency tables were in the intensity units of in/hr, which I agree is intensity values. It seems like many or most users though are more likely to have precip accumulation amounts for various durations (e.g. 6-hr, 24-hr, storm total, etc). Thus it might be nice to at least offer tables in accumulated formats as an*

optional product that the user can select. Perhaps it was there and I missed it in my less than thorough review.

Response: The PFDS does in fact offer two types of output -- intensity (in/hr or mm/hr) or depth (in or mm).

5.8 *I welcome that the methodoloogy is no longer tied to the Gumbel Distribution but has moved to far more appropriate, particularly the GEV, and Peaks over Thereshold.*

Response: Thank you for the comment. We agree that the current state of the science (i.e., a regional approach using L-moments) is a welcome advancement over previous techniques. We actually found in Puerto Rico that, although GEV and other distributions were tested, the generalized normal (GNO) distribution was most appropriate to fit the data. Details will be forthcoming in our final documentation.

5.9 *Regarding L-Moments estimation, I prefer to perform the analyses within and between models, through a Bayesian Approach. There is a tendency to under-estimate high floods, doing the estimation via "plug-in" estimates, and that may be serious specially for very large values, far in the tail.*

Response: Thank you for your comment and your mention of using the Bayesian approach for frequency analysis. This provides an opportunity to briefly discuss the approach that HDSC is using for the updates by understanding the difference between the two approaches.

First of all, it could be useful for us if you would provide details or examples of your analyses, particularly where high floods might be under-estimated through “plug-in” estimates. However, regional L-moments, as we have used here, have shown advantages in modeling extremes, especially for very large events far in the tail producing more robust and reliable quantile estimates. Below is a brief comparison of the classic frequency analysis and the Bayesian analysis approaches with respect to modeling hydrologic extremes.

Currently in statistics, there are two major approaches: Frequency or Classic Approach and Bayesian Approach. In frequency analysis of hydrometeorological extremes, the Frequency Approach considers a sample series x_i (i = 1, 2, …, n) as a realization randomly drawn from an unknown population X that is characterized by a distribution with a set of fixed parameters describing the properties of the population. The distribution and its parameters uniquely determine the population and can be estimated based on the sample data. There is no way to obtain the population distribution or parameters through a theoretical analysis. Due to sampling error and non-linearity, the parameter estimates are unstable based on single-station analyses or Conventional Moments Method. In NOAA Atlas 14, HDSC has demonstrated advantages, using real data, of the regional analysis using the L-moments Method to provide stable quantile estimates in terms of robustness and reliability. The method is described in detail in “Regional frequency analysis, an approach using L-moments” by Hosking and Wallis (1997).

The Bayesian Approach also considers a parametric distribution to model the sample data. However, unlike the Frequency Approach, the Bayesian Approach views the parameters of

the distribution as random variables that can be estimated through the use of some initial, or prior, information. Then, the prior information is combined with the sample data together to transform into a posterior distribution. The key of the transformation is the adjustment of the prior probability to the posterior probability. The quantiles then are estimated based only on the posterior probability. The Bayesian approach is based on the Bayes's theorem, which can be developed from the concept of conditional probability as written below:

$$P\{B/A\} = \frac{P\{A/B\}P\{B\}}{P\{A\}} = P\{B\} * \frac{P\{A/B\}}{P\{A\}} = \frac{P\{A/B\}P\{B\}}{P\{A/B\}P\{B\} + P(A/\overline{B}\}P\{\overline{B}\}} = \frac{1}{1 + \frac{P\{A/\overline{B}\}P\{\overline{B}\}}{P\{A/B\}P\{B\}}}$$

where P{B} is called a prior probability and P{B/A} is called posterior probability, as A is known to have happened and the relevant probability for B is now the conditional probability of B given A. Here, $\overline{B}$ is the opposite event of B. Bayes's theorem can be thought of a mechanism for updating a prior probability to a posterior probability when the additional information that event A has occurred becomes available. From the equation above, it is clear that the updating is accomplished through multiplication of the prior probability P{B} by P{A/B}/P{A}. Here, P{A/B} and/or $P\{A/\overline{B}\}$ is called the likelihood function of B.

So far, the Bayesian Approach in appearance seems better because it uses more information (observations plus the initial information about the parameters) than the Frequency Approach (observations only). But, the major reservation regarding the application of the Bayesian Approach to frequency analysis is the selection of prior information, such as elevation and mean precipitation over a certain period of time. The determination of a prior probability is very difficult in the real world of modeling extremes and sometimes it becomes a subjective process.

Generally speaking, we have reservations about the application of the Bayesian Approach to frequency analysis of hydrologic extremes at least for now when data are too limited in the determination of valuable prior events, although it is seeing wider application in other fields.

5.10 *Please provide information on the inter-event dry period used for the analyses.*

Response: The estimates are based on the analysis of annual maximum series and then converted to partial duration series results. An annual maximum series is constructed by taking the highest accumulated precipitation for a particular duration in each successive year of record. As such the maximums are inherently independent of one another and no inter-event dry period is necessary. A partial duration series is constructed by taking all of the highest cases above a threshold regardless of the year in which the case occurred. In this Atlas, partial duration series consist of the N largest cases in the period of record, where N is the number of years in the period of record at the particular observing station. Independence of the partial duration maximums was assured by selecting events separated

by at least one dry day. The final documentation will provide the details of the maximum extraction procedure.

5.11 *The topo-map elevations are incorrect. The highest point in Puerto Rico is Cerro del Punta at 4,389 ft. (1,338m), not 3693m (13,000ft) as indicated in the legend.*

Response: Thank you, we will resolve this oversight.

5.12 *I reviewed the subject information and was familiarized with the dataset subject to frequency and trend analysis. The maps and tables shown coincide with work produced by the Office of Water Plan of Puerto Rico, with the exception of the use of 128 National Oceanographic and Atmospheric (NOAA) rainfall stations. The Office of the Water Plan evaluated frequency and trends for 134 stations. Obviously the methodology used for screening of stations was also different. In your analysis a shift in the estimate of the mean was used while for the purpose of our evaluation we used the duration or extend for the available information or period of data. I used those stations in which the data period extends for more than 20 years. However, both analyses are very similar when comparing the results in magnitude and frequency.*

Of importance was the correction of 30-year statistics of the old technical paper 42 which are currently showing recurrence intervals of about 10-years. I also suggest the evaluation of extreme events rainfall, more specifically the pass of a high-convection low pressure system over the island of Puerto Rico at a speed of less than 5 miles per hour. Typically, hurricanes that hit the island have a horizontal movement of 8 to 15 miles per hour reducing the amount of rainfall over the land mass. Hurricane Mitch produced substantial amounts of rainfall over Honduras when it became stationary during days beginning October 28, 1998.

Response: Thank you for your comment and information. It is good to know that there is consistency between your analysis and ours. In our final documentation we will provide a brief comparison of our estimates with those from Technical Paper 42. We should note that our estimates are based on an analysis of all observations in the period of record and so incorporate all observed extremes within the limitations of the distribution and reliability of rainfall gages and the construction of annual maximum and partial duration series. We have not screened out or given unequal weight to any particular events. We assume that the period of record we use represents the range of variability in the underlying population fairly. By doing this we ensure that our results represent the variability of the extreme rainfall climatology that must be accounted for in engineering design.

Appendix A.6. Daily and hourly station lists for NOAA Atlas 14 Volume 3 showing station ID, station name and state, daily region in which the station resides, longitude, latitude, elevation (feet), begin date of record, end date of record, number of data years (i.e., years for which a reliable annual maximum was extracted), station coefficient of L-variation (L-CV), L-skewness (L-CS), L-kurtosis (L-CK), and discordancy of the station within its region (Disc.).

Table A.6.1. Daily stations (statistical values for the 24-hour duration)

ID	Name	ST	Daily Region	LON	LAT	Elev (ft)	Begin	End	Data yrs	L-CV	L-CS	L-CK	Disc.
66-0040	ACEITUNA	PR	4	-66.5000	18.1500	2140	01/1949	11/2004	54	0.2648	0.3762	0.1933	0.75
66-0053	ADJUNTAS 1 NW	PR	4	-66.7333	18.1833	1500	05/1899	12/2004	83	0.3182	0.4078	0.2315	0.30
66-0061	ADJUNTAS SUBSTATION	PR	4	-66.8000	18.1833	1830	01/1970	12/2004	35	0.3296	0.4384	0.2343	0.36
66-0152	AGUIRRE	PR	1	-66.2167	17.9667	25	07/1899	12/2004	104	0.2597	0.3111	0.2017	0.25
66-0158	AIBONITO 1 S	PR	5	-66.2667	18.1333	2370	01/1906	12/2004	75	0.3422	0.4137	0.2606	1.44
66-0410	ARECIBO 3 ESE	PR	2	-66.6833	18.4500	10	02/1900	01/1999	94	0.2139	0.2352	0.1265	1.18
66-0426	ARECIBO OBSERVATORY	PR	8	-66.7500	18.3500	1060	02/1980	12/2004	25	0.1615	0.1854	0.1203	0.90
66-0662	BARCELONETA 2	PR	6	-66.5500	18.4500	10	04/1915	05/1990	65	0.2439	0.2627	0.1552	0.08
66-0736	BARRANQUITAS	PR	5	-66.3167	18.1667	2060	01/1943	12/1991	39	0.2870	0.3317	0.2002	0.05
66-0842	CANDELARIA TOA BAJA	PR	6	-66.2000	18.4167	150	05/1899	05/1973	63	0.1994	0.1657	0.0804	1.03
66-0948	BOCA	PR	1	-66.8189	17.9906	30	09/1928	12/2004	74	0.3271	0.3732	0.2067	0.60
66-1123	CABO ROJO	PR	9	-67.1500	18.0833	249	01/1909	08/1969	58	0.2292	0.3428	0.3109	1.09
66-1301	CAGUAS 2 ENE	PR	5	-66.0167	18.2500	177	05/1899	05/1967	64	0.2781	0.2736	0.1902	0.50
66-1309	CAGUAS 1 W	PR	5	-66.0500	18.2333	260	03/1970	03/1995	24	0.3016	0.4337	0.1720	4.13
66-1345	CALERO CAMP	PR	2	-67.1167	18.4833	279	09/1928	12/2004	74	0.1920	0.1706	0.1340	1.32
66-1590	CANOVANAS	PR	6	-65.9000	18.3833	30	01/1901	12/2004	102	0.2389	0.2389	0.1855	0.71
66-1623	CAONILLAS UTUADO	PR	8	-66.6500	18.2833	854	01/1949	11/1987	36	0.1565	0.1438	0.0846	0.73
66-1701	CARITE DAM	PR	5	-66.1000	18.0833	1808	06/1911	04/1980	66	0.2431	0.2665	0.1855	0.25
66-1712	CARITE PLANT 1	PR	5	-66.1167	18.0500	968	01/1949	03/1980	28	0.2810	0.3337	0.2744	1.20
66-1845	CATANO	PR	6	-66.1167	18.4167	20	01/1942	05/1976	31	0.2127	0.1785	0.0498	0.97
66-1901	CAYEY 1 E	PR	5	-66.1500	18.1167	1370	01/1899	06/2001	95	0.3414	0.4189	0.2576	1.35
66-2330	CERRO GORDO CIALES	PR	4	-66.5167	18.2833	955	10/1969	09/1997	27	0.1607	0.2804	0.3683	3.80
66-2336	CERRO MARAVILLA	PR	4	-66.5500	18.1500	4002	04/1969	12/2004	36	0.3751	0.4422	0.2168	1.39
66-2634	CIDRA 1 E	PR	5	-66.1333	18.1833	1400	07/1899	06/1994	75	0.3003	0.3309	0.1845	0.19
66-2723	COAMO 2 SW	PR	1	-66.3833	18.0500	365	01/1921	12/2003	79	0.3209	0.3055	0.1709	0.79
66-2801	COLOSO	PR	3	-67.1500	18.3833	40	07/1899	12/2004	105	0.1828	0.3099	0.2040	0.96

ID	Name	ST	Daily Region	LON	LAT	Elev (ft)	Begin	End	Data yrs	L-CV	L-CS	L-CK	Disc.
66-2825	COMERIO FALLS PLANT 2	PR	6	-66.1833	18.2667	367	01/1908	05/1974	64	0.2372	0.2167	0.1758	1.00
66-2934	COROZAL SUBSTATION	PR	6	-66.3667	18.3333	650	01/1931	12/2004	70	0.2323	0.2338	0.0993	0.16
66-3023	CORRAL VIEJO	PR	4	-66.6500	18.0667	574	04/1970	12/2004	35	0.3060	0.3442	0.1545	0.73
66-3145	CULEBRA ISLAND	PR	7	-65.2833	18.3000	49	01/1947	07/1975	26	0.2564	0.3084	0.1444	1.58
66-3409	DORADO 2 WNW	PR	6	-66.2833	18.4667	5	01/1908	12/2004	96	0.2394	0.2811	0.1447	0.20
66-3431	DOS BOCAS	PR	8	-66.6667	18.3333	200	01/1937	12/2004	67	0.2457	0.4654	0.3819	0.46
66-3532	ENSENADA 1 W	PR	1	-66.9333	17.9667	213	01/1923	12/2004	80	0.2965	0.3303	0.1917	0.07
66-3657	FAJARDO	PR	6	-65.6500	18.3167	30	01/1899	01/1996	96	0.2510	0.1833	0.0726	0.70
66-3871	GARZAS	PR	4	-66.7333	18.1500	2359	01/1939	01/1981	39	0.2157	0.3127	0.2007	0.83
66-3904	GUAJATACA DAM	PR	3	-66.9333	18.4000	663	01/1929	12/2004	74	0.1528	0.1653	0.1869	0.57
66-4126	GUAYABAL	PR	1	-66.4833	18.0667	370	01/1912	12/2004	76	0.2874	0.4384	0.3134	2.30
66-4193	GUAYAMA 2 E	PR	1	-66.1167	17.9833	72	01/1902	12/2004	103	0.2684	0.3068	0.1713	0.22
66-4260	GUINEO RESERVOIR	PR	4	-66.5333	18.1667	2999	01/1929	03/1969	37	0.2326	0.3093	0.1709	0.89
66-4271	GURABO	PR	5	-65.9667	18.2500	249	01/1944	05/1967	21	0.2456	0.4877	0.3633	3.99
66-4276	GURABO SUBSTATION	PR	5	-66.0000	18.2667	160	03/1956	12/2004	49	0.2804	0.2568	0.1438	0.35
66-4330	HACIENDA CONSTANZA	PR	9	-67.0833	18.2167	480	10/1969	12/2004	35	0.1954	0.2860	0.1821	0.52
66-4613	HUMACAO 2 SSE	PR	5	-65.8333	18.1333	131	01/1900	01/1996	93	0.2784	0.2737	0.1544	0.16
66-4677	INDIERA ALTA	PR	4	-66.8833	18.1500	2600	10/1962	06/1990	27	0.2926	0.4794	0.3888	1.07
66-4702	ISABELA SUBSTATION	PR	2	-67.0667	18.4667	420	01/1899	10/2004	105	0.2073	0.2150	0.1493	0.07
66-4867	JAJOME ALTO	PR	5	-66.1333	18.0833	2385	01/1914	12/2004	89	0.3080	0.3817	0.2490	0.50
66-4910	JAYUYA	PR	4	-66.5833	18.2167	1540	04/1909	08/2002	55	0.3053	0.3951	0.2819	1.00
66-4976	JOSEFA	PR	1	-66.1500	17.9667	30	01/1911	01/1969	49	0.2277	0.2186	0.1612	1.06
66-5020	JUANA DIAZ CAMP	PR	1	-66.5000	18.0500	262	01/1901	11/2004	99	0.2775	0.3397	0.2135	0.14
66-5064	JUNCOS 1 SE	PR	5	-65.9167	18.2500	213	01/1931	12/2004	69	0.2870	0.3209	0.1762	0.10
66-5075	LA FE	PR	5	-65.7667	18.2333	148	01/1925	03/1969	39	0.2678	0.2932	0.1641	0.08
66-5097	LAJAS SUBSTATION	PR	9	-67.0833	18.0333	90	01/1947	12/2004	57	0.3156	0.3662	0.1688	1.81
66-5123	LA MUDA CAGUAS	PR	5	-66.1000	18.3167	290	09/1971	06/1994	22	0.2422	0.3142	0.2733	1.06
66-5175	LARES	PR	8	-66.8833	18.2833	1480	06/1903	12/1991	85	0.1711	0.3497	0.3424	1.32
66-5474	LOS CANOS	PR	2	-66.7000	18.4333	30	01/1950	08/1973	21	0.2064	0.2241	0.1537	0.32
66-5693	MAGUEYES ISLAND	PR	1	-67.0500	17.9667	12	01/1959	12/2004	46	0.3417	0.4031	0.2241	1.03
66-5807	MANATI 2 E	PR	6	-66.4500	18.4333	250	01/1900	12/2004	98	0.2229	0.2371	0.1467	0.11
66-5906	MARICAO	PR	4	-66.9833	18.1833	1499	01/1908	04/1969	53	0.2479	0.4272	0.3509	0.37

ID	Name	ST	Daily Region	LON	LAT	Elev (ft)	Begin	End	Data yrs	L-CV	L-CS	L-CK	Disc.
66-5908	MARICAO 2 SSW	PR	4	-66.9833	18.1500	2832	05/1969	12/2004	35	0.2015	0.4143	0.3628	1.05
66-5911	MARICAO FISH HATCHERY	PR	4	-66.9833	18.1667	1500	01/1955	12/2004	47	0.2321	0.4881	0.3796	2.05
66-6050	MAUNABO	PR	1	-65.9000	18.0000	49	05/1899	04/2003	88	0.2488	0.2407	0.1776	0.68
66-6073	MAYAGUEZ CITY	PR	9	-67.1333	18.2167	74	06/1899	11/2004	104	0.2718	0.4929	0.3888	1.06
66-6083	MAYAGUEZ AIRPORT	PR	9	-67.1500	18.2500	38	01/1957	12/2004	47	0.2011	0.4075	0.2897	1.05
66-6255	MONA ISLAND	PR	7	-67.8500	18.0833	167	02/1918	08/1974	51	0.2901	0.3603	0.2240	0.55
66-6258	MONA ISLAND 2	PR	7	-67.9333	18.1000	7	02/1980	11/2004	24	0.2661	0.2428	0.2346	1.44
66-6270	MONTE BELLO MANATI	PR	8	-66.5333	18.3667	640	10/1969	09/2001	32	0.2554	0.3584	0.2232	1.02
66-6361	MORA CAMP	PR	2	-67.0333	18.4667	410	01/1931	12/2004	72	0.2224	0.2048	0.1029	1.59
66-6390	MOROVIS 1 N	PR	6	-66.4000	18.3333	600	02/1956	12/2004	48	0.2102	0.1660	0.1245	1.02
66-6432	NAGUABO 6 W	PR	5	-65.7333	18.2333	98	01/1909	05/1967	34	0.2968	0.4136	0.3366	1.86
66-6514	NEGRO-COROZAL	PR	6	-66.3333	18.3333	1710	01/1976	12/2004	29	0.3158	0.2935	0.0203	4.16
66-6805	PARAISO	PR	5	-65.7167	18.2667	360	01/1926	12/2004	73	0.2444	0.2277	0.1507	0.46
66-6904	PATILLAS DAM	PR	5	-66.0333	18.0333	240	01/1912	01/1969	57	0.2320	0.2425	0.1515	0.47
66-6983	PENUELAS 1 NE	PR	4	-66.7167	18.0667	510	01/1920	12/2003	49	0.3088	0.3726	0.1679	0.51
66-6992	PICO DEL ESTE	PR	5	-65.7667	18.2667	3448	10/1969	12/2004	35	0.1872	0.1018	0.0600	2.93
66-7292	PONCE 4 E	PR	1	-66.5333	18.0167	70	09/1899	12/2004	96	0.2827	0.3382	0.2275	0.20
66-7295	PONCE CITY	PR	1	-66.6167	18.0000	10	04/1900	08/1998	69	0.2303	0.3180	0.1332	3.44
66-7348	POTALA	PR	1	-66.5167	18.0167	49	01/1909	02/1969	52	0.2365	0.1389	0.0860	2.31
66-7492	PUERTO REAL	PR	9	-67.1833	18.1000	33	01/1944	08/2001	49	0.2510	0.2546	0.1841	0.68
66-7843	QUEBRADILLAS	PR	2	-66.9333	18.4667	372	04/1924	09/2000	74	0.2254	0.2592	0.1968	1.52
66-8126	RINCON	PR	3	-67.2667	18.3667	10	06/1968	11/2004	29	0.1928	0.2542	0.1169	1.32
66-8144	RIO BLANCO LOWER	PR	5	-65.7833	18.2500	130	01/1941	12/2004	54	0.2558	0.3532	0.2087	0.75
66-8155	RIO BLANCO UPPER	PR	5	-65.7833	18.2833	1440	01/1904	03/1974	51	0.2318	0.2361	0.1553	0.47
66-8245	RIO GRANDE EL VERDE	PR	5	-65.8167	18.3500	600	01/1912	12/1987	59	0.3140	0.2586	0.0523	2.05
66-8306	RIO PIEDRAS EXP STA	PR	6	-66.0667	18.4167	92	01/1911	11/2004	93	0.2177	0.1791	0.0773	0.54
66-8412	ROOSEVELT ROADS	PR	6	-65.6333	18.2500	38	10/1942	03/2004	49	0.2453	0.2089	0.0950	0.24
66-8536	SABANA GRANDE 2 ENE	PR	1	-66.9333	18.0833	850	06/1928	10/2004	63	0.2880	0.4198	0.2951	1.63
66-8634	ST JUST	PR	6	-66.0000	18.3833	98	03/1943	12/1966	22	0.1863	0.2124	0.0538	2.94
66-8745	SAN CRISTOBAL	PR	5	-65.7333	18.2167	79	01/1925	03/1972	42	0.2579	0.2108	0.0759	0.97
66-8757	SAN GERMAN 4 W	PR	9	-67.1000	18.0833	89	01/1900	07/1973	67	0.2179	0.2560	0.1074	0.61
66-8808	SAN JUAN CITY	PR	6	-66.1000	18.4667	20	01/1899	05/1977	77	0.2198	0.2831	0.1863	0.45

ID	Name	ST	Daily Region	LON	LAT	Elev (ft)	Begin	End	Data yrs	L-CV	L-CS	L-CK	Disc.
66-8812	SAN JUAN INTL AP	PR	6	-66.0000	18.4333	9	01/1956	12/2004	49	0.2534	0.2879	0.2151	0.94
66-8815	SAN LORENZO 3 S	PR	5	-65.9667	18.1167	510	03/1966	12/2004	38	0.2858	0.2438	0.1382	0.68
66-8817	SAN LORENZO ESPINO	PR	5	-66.0000	18.1000	1270	01/1900	06/1959	40	0.2282	0.2870	0.1587	1.11
66-8822	SAN LORENZO FARM 2 NW	PR	5	-65.9667	18.2167	240	01/1925	09/1988	47	0.3011	0.3682	0.1793	0.75
66-8881	SAN SEBASTIAN 2 WNW	PR	3	-67.0167	18.3500	170	01/1908	10/1997	70	0.1318	0.2848	0.2301	1.26
66-8940	SANTA ISABEL 2 ENE	PR	1	-66.4333	18.0000	30	06/1901	12/2004	93	0.2960	0.2952	0.2467	1.99
66-8955	SANTA RITA	PR	1	-66.8667	18.0333	75	01/1903	11/2004	93	0.3365	0.3603	0.1998	0.87
66-9421	TOA BAJA 1 SSW	PR	6	-66.2667	18.4333	26	04/1926	08/1994	64	0.2581	0.2701	0.1354	0.21
66-9432	TORO NEGRO FOREST	PR	4	-66.4928	18.1731	2848	06/1911	12/2004	90	0.2988	0.4417	0.2877	0.13
66-9521	TRUJILLO ALTO 2 SSW	PR	6	-66.0000	18.3333	115	02/1957	12/2004	39	0.2093	0.3209	0.2521	1.63
66-9608	UTUADO	PR	8	-66.7000	18.2667	520	06/1920	07/1998	68	0.1640	0.1683	0.1559	1.65
66-9763	VIEQUES ISLAND	PR	7	-65.4667	18.1000	50	05/1899	12/1975	70	0.2567	0.2702	0.2249	0.93
66-9774	VILLALBA 1 E	PR	4	-66.4833	18.1333	430	01/1941	11/2004	62	0.2398	0.4134	0.3399	0.32
66-9829	YABUCOA 1 NNE	PR	1	-65.8667	18.0667	30	01/1905	03/1995	84	0.2552	0.2625	0.1821	0.36
66-9860	YAUCO 1 NW	PR	1	-66.8500	18.0500	180	12/1981	12/2004	23	0.3385	0.3570	0.1653	1.53
66-9862	YAUCO 1 S	PR	1	-66.8500	18.0167	30	01/1900	06/1969	42	0.2773	0.2984	0.1512	0.39
67-0198	ALEX HAMILTON FLD FAA	VI	7	-64.7989	17.6946	44	03/1951	12/2003	44	0.3377	0.3880	0.2708	0.46
67-0240	ANNALY	VI	7	-64.8529	17.7521	700	01/1972	02/2003	27	0.3639	0.3840	0.2038	0.53
67-0480	BETH UPPER NEW WORKS	VI	7	-64.8000	17.7167	110	01/1972	05/2003	31	0.3976	0.4133	0.2245	1.25
67-1316	CANEEL BAY PLANTATION	VI	7	-64.7863	18.3429	60	01/1972	02/1998	25	0.3207	0.4055	0.3642	2.09
67-1348	CATHERINEBURG	VI	7	-64.7606	18.3453	845	01/1972	11/1996	25	0.3026	0.2364	0.1814	0.76
67-1740	CHRISTIANSTED FORT	VI	7	-64.6996	17.7446	30	01/1921	07/1995	38	0.3818	0.3855	0.1670	1.15
67-1790	CORAL BAY	VI	7	-64.7146	18.3471	30	01/1972	12/2003	25	0.3214	0.1795	0.0248	2.23
67-1810	COTTON VALLEY 2	VI	7	-64.6128	17.7603	140	02/1982	08/2001	20	0.3291	0.1794	0.1515	2.63
67-1980	CRUZ BAY	VI	7	-64.7946	18.3321	8	01/1972	12/2003	30	0.2567	0.2104	0.0459	2.29
67-2551	EAST END	VI	7	-64.6729	18.3388	150	01/1972	12/2003	26	0.3073	0.3558	0.2410	0.23
67-2560	EAST HILL	VI	7	-64.6494	17.7561	120	01/1972	12/2003	30	0.3428	0.3253	0.2093	0.16
67-2823	ESTATE FORT MYLNER	VI	7	-64.8938	18.3329	200	01/1972	08/1995	24	0.3097	0.3056	0.2685	0.91
67-2870	ESTATE THE SIGHT	VI	7	-64.6667	17.7500	130	11/1979	06/2002	22	0.3305	0.3524	0.1795	0.29
67-3150	FOUNTAIN	VI	7	-64.8238	17.7496	250	01/1972	04/1992	20	0.3389	0.2846	0.1830	0.35
67-3677	GRANARD	VI	7	-64.7167	17.7164	65	01/1972	06/2003	31	0.3962	0.3549	0.1758	1.12
67-3880	HAM BLUFF LIGHTHOUSE	VI	7	-64.8667	17.7679	80	01/1972	10/1992	21	0.3182	0.3923	0.2506	0.50

ID	Name	ST	Daily Region	LON	LAT	Elev (ft)	Begin	End	Data yrs	L-CV	L-CS	L-CK	Disc.
67-4900	MONTPELLIER	VI	7	-64.7500	17.7667	200	11/1979	12/2003	24	0.3276	0.3331	0.1937	0.03
67-8905	TRUMAN FIELD AIRPORT	VI	7	-64.9713	18.3363	20	09/1972	12/2003	28	0.2626	0.2526	0.1341	0.82
67-9450	WINTBERG	VI	7	-64.9167	18.3500	645	01/1972	12/2003	30	0.2536	0.3724	0.2513	1.73

Appendix A.6. Daily and hourly station lists for NOAA Atlas 14 Volume 3 showing station ID, station name and state, daily region in which the station resides, longitude, latitude, elevation (feet), begin date of record, end date of record, number of data years (i.e., years for which a reliable annual maximum was extracted), station coefficient of L-variation (L-CV), L-skewness (L-CS), L-kurtosis (L-CK), and discordancy of the station within its region (Disc.).

Table A.6.2. Hourly stations (statistical values for the 60-minute duration)

ID	Name	ST	Hourly Region	LON	LAT	Elev (ft)	Begin	End	Data yrs	L-CV	L-CS	L-CK	Disc.
66-0061	ADJUNTAS SUBSTN	PR	3	-66.8000	18.1833	1830	01/1971	12/2003	33	0.1142	0.2019	0.1623	1.25
66-0900	BENAVENTE-HORMIGUEROS	PR	3	-67.1167	18.1167	45	01/1973	12/2001	20	0.1324	0.0605	0.2080	1.27
66-0984	BOTIJAS 1 OROCOVIS	PR	2	-66.3589	18.2408	1840	01/1973	12/2003	30	0.1718	0.1758	0.1148	1.76
66-0988	BOTIJAS 2 OROCOVIS	PR	2	-66.3531	18.1981	2230	01/1973	12/2003	31	0.1620	-0.0069	0.0458	0.90
66-1901	CAYEY 1 E	PR	1	-66.1500	18.1117	1370	01/1971	12/2003	32	0.2230	0.2884	0.2135	0.98
66-2336	CERRO MARAVILLA	PR	3	-66.5500	18.1500	4002	01/1971	12/2003	29	0.1498	0.1012	0.0667	1.57
66-2934	COROZAL SUBSTN	PR	2	-66.3667	18.3333	650	01/1971	12/2003	33	0.1539	0.2139	0.2045	1.56
66-3113	CUBUY	PR	1	-65.8681	18.2706	1970	01/1973	12/2003	30	0.1482	0.0025	0.1564	2.18
66-3431	DOS BOCAS	PR	3	-66.6667	18.3333	200	01/1971	12/2003	33	0.1039	-0.0289	0.0912	0.87
66-3480	OUQUE 2 NE	PR	1	-65.7119	18.2550	450	01/1973	12/2001	27	0.2017	0.2577	0.0891	1.06
66-3657	FAJARDO	PR	1	-65.6500	18.3167	30	01/1971	12/2003	32	0.2148	0.2265	0.0671	0.71
66-4272	GURABO 2 NNE	PR	2	-65.9989	18.2936	945	01/1973	12/2003	31	0.1621	0.1883	0.1709	0.47
66-4276	GURABO SUBSTN	PR	2	-66.0000	18.2667	160	01/1971	12/2003	33	0.1661	0.1280	0.1162	0.11
66-5258	LAS PIEDRAS 1 N	PR	1	-65.8650	18.1964	270	01/1973	12/2003	31	0.2030	0.1242	0.0740	0.34
66-5908	MARICAO 2 SSW	PR	3	-66.9833	18.1500	2832	01/1971	12/2003	33	0.1140	0.1288	0.2117	0.56
66-6514	NEGRO-COROZAL	PR	2	-66.3333	18.3333	1710	01/1973	12/2003	31	0.1131	0.0937	0.0042	2.29
66-6942	PENA POBRE NAGUABO	PR	1	-65.8261	18.2197	330	01/1973	12/2003	31	0.2057	0.0673	0.0344	0.80
66-6992	PICO DEL ESTE	PR	1	-65.7667	18.2667	3448	01/1973	12/2003	29	0.2051	0.3308	0.2777	0.68
66-7292	PONCE 4 E	PR	2	-66.5333	18.0167	70	01/1971	12/2003	33	0.1724	0.0525	0.0883	0.33
66-8812	SAN JUAN WSFO	PR	2	-66.0000	18.4333	9	01/1967	12/2003	37	0.1713	0.0167	0.0564	0.58
66-8816	SAN LORENZO 2 ESE	PR	1	-65.9294	18.1850	460	01/1973	12/2003	31	0.1982	0.3350	0.2864	0.59
66-8881	SAN SEBASTIAN 2 WNW	PR	3	-67.0167	18.3500	170	01/1971	12/2003	32	0.1081	-0.0016	0.1059	0.47
66-9829	YABUCOA 1 NNE	PR	1	-65.8667	18.0667	30	01/1971	12/2003	33	0.1668	0.3559	0.2871	1.66
67-0480	BETH UPPER NEW WORKS	VI	4	-64.8000	17.7167	110	01/1978	05/2003	24	0.2422	0.1110	0.1218	1.00
67-1316	CANEEL BAY PLANTATION	VI	4	-64.7863	18.3429	60	01/1978	12/2002	17	0.2015	0.0864	0.1424	1.00

Appendix A.7. Average L-moment statistics and heterogeneity measures for regions used to prepare NOAA Atlas 14 Volume 3.

Table A.7.1. Number of daily and hourly stations, H1 statistic, mean number of data years, and weighted L-statistics of 24-hour data for each daily region and at-site.

region	# daily stations	# hourly stations	H1	Mean number of data years*	Coeff. of L-variation Weighted Mean	L-Skewness Weighted Mean	L-Kurtosis Weighted Mean
1	19	2	1.17	70	0.282	0.317	0.198
2	6	0	-0.83	73	0.212	0.218	0.143
3	4	1	1.51	62	0.159	0.242	0.195
4	16	3	0.62	45	0.277	0.402	0.265
5	26	10	1.45	47	0.278	0.309	0.188
6	18	5	0.60	57	0.236	0.239	0.134
7	23	2	0.17	29	0.316	0.321	0.199
8	6	1	1.40	49	0.200	0.322	0.266
9	7	1	0.67	55	0.246	0.351	0.239
total	**125**	**25**		**51**			

*includes both daily and hourly stations

Table A.7.2. Number of hourly stations, H1 statistic, mean number of data years, and weighted L-statistics of 60-minute data for each hourly region.

region	# hourly stations	H1	Mean number of data years	Coeff. of L-variation Weighted Mean	L-Skewness Weighted Mean	L-Kurtosis Weighted Mean
1	9	-0.46	31	0.196	0.222	0.166
2	8	0.46	32	0.159	0.106	0.100
3	6	0.37	30	0.119	0.078	0.138
4	2	-0.05	21	0.225	0.101	0.130
total	**25**		**29**			

Appendix A.8. Heterogeneity statistic, H1, for regions and durations used in NOAA Atlas 14 Volume 3.

Table A.8.1. H1 for daily regions (1-84) for durations 24-hour through 60-day.

Region	24-hr	2-day	4-day	7-day	10-day	20-day	30-day	45-day	60-day
1	1.17	0.67	0.79	0.21	0.90	2.36	2.50	2.08	2.93
2	-0.83	0.79	0.35	0.45	0.53	0.63	0.17	0.40	1.49
3	1.51	0.26	0.60	1.63	1.18	-0.27	0.21	0.54	0.32
4	0.62	0.67	0.66	0.18	0.43	0.45	1.00	1.31	2.82
5	1.45	0.97	0.86	0.73	0.82	0.80	1.80	2.53	3.12
6	0.60	0.28	0.24	0.47	1.78	3.23	4.17	2.97	2.91
7	0.17	0.01	0.94	1.11	-1.2	-1.04	-1.16	-0.04	0.46
8	1.40	0.69	1.22	2.56	2.83	3.56	2.79	5.09	4.54
9	0.67	1.49	1.58	1.92	2.26	1.64	1.92	1.15	0.58

Table A.8.2. Heterogeneity statistic, H1, for hourly regions (1-26) for durations 60-minute through 48-hour.

Region	60-min	2-hour	3-hour	6-hour	12-hour	24-hour	48-hour
1	0.46	0.76	1.41	0.23	1.61	1.05	1.83
2	0.46	0.75	0.05	1.49	0.28	-0.67	-0.78
3	0.37	3.63	4.04	4.16	3.11	2.57	2.06
4	0.05	0.03	0.36	0.90	-1.18	-1.28	-0.98

Appendix A.9. Regional growth factors for regions used in NOAA Atlas 14 Volume 3.

Table A.9.1. Regional growth factors for daily regions and at-site analyses for each duration 24-hour to 60-day for the annual maximum series results. *Note that the 1.58-year was computed to equate the 1-year average recurrence interval (ARI) for partial duration series results (see Section 4.6.2) and the 1.58 year results were not released as annual exceedance probabilities (AEP).

24-hour										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.720	0.845	1.314	1.686	2.223	2.670	3.158	3.688	4.461	5.104
2	0.810	0.918	1.271	1.516	1.837	2.084	2.336	2.596	2.953	3.234
3	0.854	0.932	1.198	1.388	1.643	1.843	2.051	2.268	2.570	2.810
4	0.708	0.813	1.258	1.654	2.279	2.838	3.481	4.217	5.348	6.336
5	0.726	0.851	1.314	1.677	2.197	2.628	3.095	3.600	4.334	4.941
6	0.783	0.901	1.295	1.576	1.952	2.245	2.549	2.866	3.306	3.657
7	0.685	0.825	1.349	1.767	2.372	2.878	3.430	4.033	4.912	5.644
8	0.801	0.889	1.221	1.486	1.870	2.192	2.543	2.927	3.487	3.955
9	0.749	0.853	1.257	1.593	2.094	2.524	3.002	3.533	4.322	4.991

48-hour										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.711	0.838	1.317	1.701	2.259	2.727	3.239	3.799	4.618	5.302
2	0.827	0.930	1.259	1.482	1.767	1.983	2.202	2.424	2.726	2.961
3	0.841	0.921	1.200	1.409	1.696	1.927	2.170	2.429	2.795	3.091
4	0.710	0.812	1.248	1.643	2.272	2.840	3.499	4.257	5.430	6.461
5	0.724	0.851	1.318	1.683	2.205	2.635	3.101	3.605	4.334	4.936
6	0.785	0.900	1.289	1.569	1.945	2.240	2.547	2.867	3.314	3.670
7	0.693	0.837	1.362	1.767	2.340	2.810	3.316	3.860	4.643	5.287
8	0.809	0.894	1.213	1.467	1.835	2.143	2.480	2.846	3.382	3.828
9	0.767	0.864	1.239	1.551	2.015	2.412	2.853	3.343	4.070	4.686

4-day										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.721	0.850	1.326	1.695	2.219	2.651	3.117	3.619	4.344	4.942
2	0.829	0.928	1.247	1.468	1.755	1.975	2.200	2.430	2.745	2.992
3	0.853	0.929	1.191	1.383	1.645	1.852	2.070	2.299	2.620	2.879
4	0.740	0.842	1.251	1.601	2.135	2.601	3.126	3.716	4.606	5.369
5	0.730	0.854	1.312	1.670	2.180	2.600	3.055	3.547	4.258	4.846
6	0.786	0.899	1.285	1.563	1.939	2.234	2.542	2.865	3.316	3.676
7	0.698	0.844	1.367	1.765	2.321	2.772	3.253	3.768	4.503	5.104
8	0.829	0.913	1.212	1.436	1.748	1.999	2.265	2.549	2.952	3.280
9	0.783	0.879	1.240	1.528	1.947	2.297	2.680	3.097	3.708	4.217

7-day										
region	*1.58-year	2-year	5-year	10-year	25-year	50-year	100-year	200-year	500-year	1,000-year
1	0.730	0.858	1.322	1.678	2.180	2.589	3.028	3.499	4.174	4.729
2	0.830	0.927	1.243	1.463	1.751	1.972	2.198	2.431	2.751	3.003
3	0.877	0.949	1.182	1.341	1.546	1.702	1.860	2.022	2.242	2.414
4	0.777	0.874	1.243	1.540	1.973	2.338	2.738	3.176	3.819	4.358
5	0.743	0.864	1.302	1.641	2.120	2.514	2.937	3.393	4.050	4.590
6	0.801	0.909	1.271	1.529	1.874	2.143	2.422	2.712	3.115	3.436
7	0.693	0.837	1.362	1.766	2.340	2.809	3.315	3.859	4.642	5.286
8	0.852	0.934	1.206	1.398	1.652	1.849	2.052	2.263	2.554	2.784
9	0.812	0.902	1.226	1.474	1.822	2.107	2.411	2.737	3.204	3.587

10-day										
region	*1.58-year	2-year	5-year	10-year	25-year	50-year	100-year	200-year	500-year	1,000-year
1	0.745	0.870	1.314	1.649	2.116	2.494	2.895	3.323	3.933	4.430
2	0.856	0.952	1.239	1.423	1.649	1.813	1.976	2.137	2.350	2.512
3	0.884	0.953	1.173	1.322	1.514	1.660	1.807	1.958	2.162	2.321
4	0.800	0.895	1.236	1.499	1.871	2.176	2.503	2.855	3.362	3.778
5	0.766	0.882	1.290	1.597	2.022	2.364	2.728	3.114	3.664	4.110
6	0.818	0.921	1.258	1.493	1.802	2.040	2.284	2.535	2.881	3.154
7	0.708	0.850	1.355	1.739	2.275	2.710	3.174	3.669	4.377	4.954
8	0.870	0.950	1.202	1.369	1.580	1.738	1.897	2.057	2.272	2.439
9	0.826	0.915	1.225	1.452	1.762	2.009	2.267	2.540	2.924	3.233

20-day										
region	*1.58-year	2-year	5-year	10-year	25-year	50-year	100-year	200-year	500-year	1,000-year
1	0.775	0.890	1.287	1.581	1.984	2.305	2.644	3.001	3.506	3.914
2	0.876	0.966	1.226	1.383	1.571	1.704	1.833	1.958	2.120	2.240
3	0.903	0.966	1.158	1.283	1.437	1.551	1.664	1.776	1.925	2.040
4	0.842	0.928	1.216	1.421	1.694	1.908	2.129	2.360	2.680	2.934
5	0.802	0.907	1.265	1.523	1.871	2.144	2.428	2.726	3.141	3.472
6	0.845	0.939	1.236	1.435	1.689	1.879	2.071	2.265	2.527	2.731
7	0.753	0.885	1.333	1.655	2.086	2.424	2.776	3.142	3.653	4.061
8	0.890	0.970	1.200	1.339	1.505	1.623	1.736	1.846	1.989	2.095
9	0.854	0.939	1.212	1.400	1.644	1.831	2.020	2.215	2.481	2.689

30-day										
region	*1.58-year	2-year	5-year	10-year	25-year	50-year	100-year	200-year	500-year	1,000-year
1	0.791	0.904	1.284	1.556	1.919	2.202	2.497	2.803	3.229	3.568

30-day										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
2	0.884	0.971	1.219	1.367	1.541	1.663	1.780	1.893	2.038	2.145
3	0.909	0.969	1.149	1.265	1.409	1.514	1.618	1.721	1.859	1.963
4	0.865	0.950	1.211	1.384	1.601	1.763	1.925	2.088	2.307	2.476
5	0.821	0.921	1.253	1.484	1.789	2.025	2.266	2.516	2.860	3.132
6	0.864	0.954	1.226	1.400	1.615	1.771	1.925	2.078	2.281	2.435
7	0.790	0.916	1.318	1.588	1.935	2.196	2.460	2.729	3.092	3.375
8	0.892	0.968	1.192	1.329	1.495	1.613	1.728	1.840	1.987	2.096
9	0.859	0.944	1.212	1.393	1.625	1.801	1.978	2.158	2.402	2.592

45-day										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.810	0.919	1.273	1.519	1.839	2.084	2.334	2.592	2.944	3.222
2	0.893	0.977	1.210	1.347	1.504	1.613	1.716	1.814	1.940	2.032
3	0.925	0.983	1.146	1.242	1.353	1.429	1.502	1.572	1.661	1.727
4	0.877	0.958	1.205	1.362	1.556	1.698	1.837	1.976	2.160	2.300
5	0.834	0.931	1.242	1.456	1.733	1.943	2.158	2.377	2.677	2.911
6	0.870	0.957	1.217	1.383	1.587	1.735	1.881	2.027	2.218	2.364
7	0.815	0.934	1.297	1.534	1.829	2.047	2.264	2.481	2.771	2.993
8	0.899	0.974	1.187	1.315	1.467	1.573	1.675	1.774	1.902	1.997
9	0.877	0.960	1.209	1.367	1.559	1.699	1.835	1.971	2.149	2.284

60-day										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.818	0.926	1.272	1.507	1.810	2.038	2.270	2.507	2.828	3.078
2	0.904	0.986	1.204	1.327	1.465	1.558	1.645	1.727	1.829	1.903
3	0.930	0.988	1.144	1.233	1.334	1.403	1.467	1.528	1.605	1.661
4	0.884	0.965	1.203	1.351	1.530	1.659	1.784	1.907	2.068	2.189
5	0.850	0.943	1.234	1.427	1.672	1.854	2.037	2.221	2.469	2.660
6	0.880	0.962	1.205	1.358	1.543	1.677	1.808	1.937	2.107	2.235
7	0.822	0.935	1.282	1.510	1.795	2.006	2.217	2.428	2.711	2.928
8	0.908	0.981	1.182	1.299	1.434	1.527	1.616	1.700	1.807	1.886
9	0.885	0.968	1.207	1.353	1.527	1.651	1.770	1.887	2.037	2.150

Table A.9.2. Regional growth factors for hourly regions analyses for each duration 60-minute to 24-hour for the annual maximum series results. *Note that the 1.58-year was computed to equate the 1-year average recurrence interval (ARI) for partial duration series results (see Section 4.6.2) and the 1.58 year results were not released as annual exceedance probabilities (AEP).

60-minute										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.823	0.923	1.250	1.479	1.779	2.010	2.248	2.493	2.830	3.096
2	0.879	0.970	1.225	1.379	1.560	1.687	1.808	1.926	2.078	2.190
3	0.915	0.983	1.171	1.280	1.405	1.490	1.571	1.648	1.746	1.817
4	0.831	0.959	1.320	1.535	1.787	1.963	2.131	2.294	2.502	2.657

120-minute										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.796	0.909	1.285	1.550	1.900	2.172	2.451	2.741	3.140	3.457
2	0.861	0.955	1.236	1.413	1.630	1.786	1.940	2.091	2.291	2.442
3	0.900	0.973	1.183	1.310	1.461	1.567	1.670	1.769	1.898	1.994
4	0.793	0.921	1.322	1.588	1.925	2.177	2.429	2.683	3.026	3.290

3-hour										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.778	0.895	1.293	1.582	1.973	2.282	2.604	2.942	3.415	3.794
2	0.847	0.945	1.247	1.443	1.688	1.868	2.047	2.226	2.464	2.647
3	0.880	0.952	1.181	1.335	1.532	1.681	1.831	1.984	2.191	2.352
4	0.799	0.930	1.329	1.586	1.905	2.139	2.370	2.601	2.907	3.141

6-hour										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.754	0.877	1.308	1.630	2.075	2.432	2.811	3.212	3.782	4.244
2	0.808	0.921	1.283	1.530	1.849	2.091	2.338	2.590	2.933	3.201
3	0.841	0.925	1.211	1.417	1.696	1.915	2.144	2.383	2.718	2.986
4	0.732	0.873	1.355	1.705	2.178	2.550	2.940	3.348	3.920	4.378

12-hour										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.752	0.876	1.312	1.637	2.086	2.446	2.827	3.232	3.805	4.270
2	0.761	0.880	1.300	1.613	2.045	2.392	2.759	3.149	3.700	4.148
3	0.787	0.882	1.237	1.520	1.929	2.270	2.643	3.049	3.641	4.135
4	0.671	0.826	1.390	1.824	2.437	2.938	3.476	4.055	4.886	5.569

24-hour										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.741	0.862	1.304	1.646	2.132	2.531	2.961	3.424	4.091	4.641
2	0.743	0.867	1.311	1.649	2.123	2.507	2.918	3.358	3.986	4.500
3	0.748	0.854	1.265	1.601	2.099	2.523	2.991	3.509	4.273	4.918
4	0.643	0.796	1.382	1.857	2.556	3.146	3.797	4.512	5.564	6.448

48-hour										
region	***1.58-year**	**2-year**	**5-year**	**10-year**	**25-year**	**50-year**	**100-year**	**200-year**	**500-year**	**1,000-year**
1	0.753	0.874	1.306	1.631	2.081	2.445	2.831	3.242	3.827	4.303
2	0.765	0.890	1.313	1.618	2.030	2.353	2.690	3.042	3.535	3.928
3	0.763	0.860	1.240	1.557	2.032	2.441	2.898	3.405	4.162	4.806
4	0.668	0.826	1.396	1.833	2.449	2.951	3.489	4.066	4.895	5.574

Glossary

annual exceedance probability (AEP) – The probability associated with exceeding a given amount in any given year; the inverse of AEP (1/AEP) provides a measure of the average time between *years in which a particular value is exceeded at least once*; the term is associated with analysis of annual maximum series.

annual maximum series (AMS) – Time series created by the extraction of the largest single case in each calendar year of record.

ArcInfo© ASCII grid – Also known as an ESRI© ASCII grid, a very simple 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-hand corner of the grid.

average recurrence interval (ARI) – Average time between *cases of a particular magnitude*; the term is associated with the analysis of partial duration series.

Cascade, Residual Add-Back (CRAB) – HDSC-developed spatial interpolation procedure for deriving grids of precipitation frequency estimates from mean annual maximum grids of different annual exceedance probability.

data years – Number of years in which enough data existed to extract maxima in a station's period of record.

depth-duration-frequency plot (DDF) - Graphical depiction of precipitation frequency estimates in terms of depth (y-axis) and duration (x-axis)

Discordancy – Measure based on coefficient-of-L-variation, L-skewness and L-kurtosis of a station's data, which represents a point in 3-dimensional space. Discordancy is a measure of the distance of each point from the cluster center of the points for all stations in a region. The cluster center is defined as the unweighted mean of the three L-moments for the stations within the region being tested. It is used for data quality control and to determine if a station is consistent with other stations in a region.

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."

GEV - Generalized Extreme Value – A 3-parameter theoretical probability distribution function.

GLO – Generalized Logistic – A 3-parameter theoretical probability distribution function.

GNO – Generalized Normal – A 3-parameter theoretical probability distribution function.

GPA – Generalized Pareto – A 3-parameter theoretical probability distribution function.

heterogeneity measure, H1 – Measure that uses coefficient of L-variation to compare between-site variations in sample L-moments for a group of stations in a region with expectations for a

homogeneous region. The H1 measure was used to assess regional homogeneity, or lack thereof.

"Index Flood" – The mean of the annual maximum series, also known as the scaling factor, at each observing station that is multiplied by the regional growth factor to produce precipitation frequency estimates. It is often referred to as the "Index Flood" because of the genesis of the statistical approach in flood frequency analysis.

intensity-duration-frequency curve (IDF) - A log-log graphical depiction of precipitation frequency estimates in terms of intensity (y-axis) and duration (x-axis).

internal consistency – Term used to describe the required behavior of the precipitation frequency estimates from one duration or frequency to the next. For instance, it is required that the 100-year 3-hour depth estimates be greater than the 100-year 120-minute depth estimates.

L-moments – Linear combinations of probability weighted moments that provide great utility in choosing the most appropriate probability distribution to describe the precipitation frequency estimates.

mean annual precipitation – The climatological average total annual precipitation. For the spatial interpolation of NOAA Atlas 14 Volume 1, the mean annual precipitation for the climatological period 1961-90 was used as a predictor grid for interpolating mean annual maximum precipitation to a uniformly spaced grid.

Monte Carlo simulation – Simulation technique used to randomly generate 1,000 synthetic data sets for each station in a region to determine sample L-moment estimates and test the fitting of theoretical distributions. The technique was also used to quantitatively assess confidence bounds.

n-minute – Precipitation data measured at a temporal resolution of 5-minutes that can be summed to various "n-minute" durations (10-minute, 15-minute, 30-minute, and 60-minute).

partial duration series (PDS) – Time series created by the extraction of all large events in which more than one large event may occur during a single calendar year. For this Atlas, the annual exceedance series (AES) consisting of the largest N events in the entire period of record, where N is the number of years of data, was used.

PE3 – Pearson Type III – A 3-parameter theoretical probability distribution function.

precipitation frequency – General term for specifying the average recurrence interval or annual exceedance probability associated with specific depths for a given duration.

Precipitation Frequency Data Server (PFDS) – The on-line portal for all NOAA Atlas 14 deliverables, documentation and information. Link to it via the HDSC home page at: http://www.nws.noaa.gov/ohd/hdsc/.

PRISM – Parameter-elevation Regressions on Independent Slopes Model – a hybrid statistical-geographic approach to mapping climate data developed by Oregon State University's Spatial Climate Analysis Service.

probability distribution – Mathematical description of a random variable, precipitation in this case, in terms of the chance of exceedance associated with each value.

pseudo data –Precipitation frequency estimates for stations that did not have observed data at a given duration. The estimates were based on ratios derived from nearby co-located stations and applied to actual observed data at the station.

quantile – Generic term to indicate the precipitation frequency estimates associated with ARIs and AEPs.

regional growth factor (RGF) – Dimensionless factors that are a function of appropriate higher order moments for a region; used to develop the site-specific quantiles for each region by multiplying by the site-specific scaling factor to produce the quantiles at each frequency and duration; there is a single RGF for each region that varies only with frequency and duration

root-mean-square-error (RMSE) – The positive square root of the mean-square-error (MSE). MSE is the mean square of any residual. RMSE is the also called the standard error of estimate.

shapefile – An ESRI© vector file format for displaying non-topological geometry and attribute information for use with Geographical Information Systems (GIS). The shapefile has the .shp extension, and comes with other associated files which can include, .shx, sbx, .sbn and .dbf.

temporal distribution – Temporal patterns in probalistic terms specifically designed to be consistent with the definition of duration used in this Atlas and for use with the precipitation frequency estimates. They are expressed as cumulative percentages of precipitation and duration at various percentiles for 1-, 6-, 12-, 24- and 96-hour durations.

t-test – for testing whether a difference between means of two samples is significant:

$t = \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{n_1 s_1^2 + n_2 s_2^2}}$, following a Student's t distribution with (n_1+n_2-2) degree of freedoms, where, $\bar{x}_1$ and $\bar{x}_2$ are the means for sample 1 and sample 2, respectively. s_1^2 and s_2^2 are sample variances. n_1 and n_2 are sample sizes. At 90% confidence level (or significance level α = 10%), reject H_0: the means have no significant difference if $| t | > t_{n_1+n_2-2,\alpha/2}$.

– for testing for population correlation: $t = \left| \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \right|$, following a Student's t distribution with (n-2) degrees of freedom. At 90% confidence level (or significance level α = 10%), reject H_0: there is no correlation or the correlation is not significant at significance level of 10% if $| t | > t_{n-2,\alpha/2}$.

Wakeby distribution – A 5-parameter theoretical probability distribution function.

References

Arkell, R.E., and F. Richards, 1986: Short duration rainfall relations for the western United States. *Conference on Climate and Water Management-A Critical Era and Conference on the Human Consequences of 1985's Climate, August 4-7, 1986,* Asheville, North Carolina.

Bonnin, G., D. Martin, T. Parzybok, B. Lin, D. Riley, and M. Yekta, 20063: Precipitation frequency atlas of the United States. *NOAA Atlas 14 Volume 1, Version 4.0*, National Weather Service, Silver Spring, Maryland.

Bonnin, G., D. Martin, T. Parzybok, B. Lin, D. Riley, and M. Yekta, 2006: Precipitation frequency atlas of the United States. *NOAA Atlas 14 Volume 2, Version 3.0*, National Weather Service, Silver Spring, Maryland.

Chow, V.T., D.R. Maidment, and L.W. Mays, 1988: *Applied Hydrology*. McGraw-Hill International Editions, 572 pp.

Daly, C., and R.P. Neilson. 1992: A digital topographic approach to modeling the distribution of precipitation in mountainous terrain. *Interdisciplinary Approaches in Hydrology and Hydrogeology*, American Institute of Hydrology, 437-454.

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.

Daly, C., R.P. Neilson, and D.L. Phillips, 1994: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal Applied. Meteorology*, **33**, 140-158.

Daly, C., G. Taylor, and W. Gibson, 1997: The PRISM Approach to Mapping Precipitation and Temperature, *10th Conf. on Applied Climatology*, American Meteorology Society, 10-12, Reno, Nevada. http://www.ocs.orst.edu/pub/prism/docs/appclim97-prismapproach-daly.pdf

Durrans, S.R., and P.A. Brown, 2002: Development of an Internet-Based Rainfall Atlas for Alabama. *Water Science and Technology*, **45**, no. 2, 11-17.

Environmental Systems Research Institute, Inc. (ESRI), 2003: ArcMap, ArcGIS version 8.3. Redlands, California.

Frederick, R.H. and J.F. Miller, 1979: Short Duration Rainfall Frequency Relations for California. *Third Conference on Hydrometeorology, August 20-24, 1979,* Bogata, Columbia.

GRASS Development Team, 2002: Geographic Resources Analysis Support System (GRASS), Grass version 5.0.

Hershfield, D.M., 1961: Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years. *Weather Bureau Technical Paper No. 40*, U.S. Weather Bureau, Washington, D.C., 115 pp.

Hosking, J.R.M. and J.R. Wallis, 1997: *Regional frequency analysis, an approach based on L-moments*. Cambridge University Press, 224 pp.

Hosking, J.R.M., and J.R. Wallis, 1991: Some statistics useful in regional frequency analysis. *Research Report RC 17096 (#75863) 8/12/1991*, Mathematics, IBM Research Division, T.J. Watson Research Center, Yorktown Heights, New York.

Huff, F. A., 1990: Time Distributions of Heavy Rainstorms in Illinois. *Illinois State Water Survey*, Champaign, **173**, 17 pp.

Institution of Engineers, Australia, 1987: *Australian Rainfall and Runoff, A Guide to Flood Estimation.* The Institution of Engineers, D.H. Pilgrim, ed., Canberra Australia.

Laurenson, E.M., 1987: Back to basics on flood frequency analysis. *Civil Engineers Trans.*, Institution of Engineers, Australia, **CE29**, 47-53.

Lin, B., G. Bonnin, D. Todd, T. Parzybok, M. Yekta, and D. Riley, 2004: Regional frequency studies of annual extreme precipitation in the United States using regional L-moments analysis. *International Ocean-Atmosphere Conference,* Chinese-American Oceanic and Atmospheric Association (COAA), Beijing, China, June 27-30, 2004.

Lin, B., and J.L. Vogel, 1993: A comparison of L-moments with method of moments. *Proceedings of the International Symposium on Engineering Hydrology*, American Society of Civil Engineers, July 1993, San Francisco, California.

Lin, Shao-Gong, 1980: *Basic Probability and Statistics*. People's Education Publisher, Beijing, China, 162 pp.

Maidment, D. R., 1993: *Handbook of Hydrology*. McGraw-Hill Publishing, 29.47 pp.

Miller, J.F., 1964: Two- to ten-day precipitation for return periods of 2 to 100 years in the contiguous United States. *Technical Paper No. 49*, U.S. Weather Bureau and U.S. Department of Agriculture, 29 pp.

Miller, J.F., 1965: Two- to Ten-Day Rainfall for Return Periods of 2 to 100 Years in Puerto Rico and the Virgin Islands, *Technical Paper No. 53*, U.S. Weather Bureau and U.S. Department of Agriculture, 35 pp.

Miller, J.F., R.H. Frederick and R.J. Tracy, 1973: Precipitation-frequency atlas of the western United States. *NOAA Atlas 2*, 11 vols., National Weather Service, Silver Spring, Maryland.

Myers, V.A. and R.M. Zehr, 1980: A Methodology for Point-to-Area Rainfall Frequency Ratios. *NOAA Technical Report NWS 24*, Office of Hydrology, National Weather Service, Silver Spring, Maryland.

Neteler, Markus and Helena Mitasova. 2002: *Open Source GIS: A GRASS GIS Approach*. Kluwer Academic Publishers, Boston.

Newbold, P., 1988: *Statistics for Business and Economics*. Prentice Hall, 866 pp.

Parzybok, T. and M. Yekta, 2003: NOAA/NWS precipitation frequency data server. *19th International Conference on Interactive Information Processing Systems (IIPS) for Meteorology*,

Oceanography, and Hydrology, 83rd American Meteorological Society Annual Meeting, Long Beach, California.

Plantico, M.S., L.A. Goss, C. Daly, and G. Taylor. 2000: A new U.S. climate atlas. In: *Proc., 12th AMS Conf. on Applied Climatology,* American Meteorological Society, May 8-11, 247-248, Asheville, North Carolina.

U.S. Weather Bureau, 1961: Generalized Estimates of Probable Maximum Precipitation and Rainfall-Frequency Data for Puerto Rico and Virgin Islands for Areas to 400 Square Miles, Durations to 24 Hours and Return Periods from 1 to 100 Years, *Technical Paper 42*, Washington, DC, 94 pp.

Zehr, R.M., and V.A. Myers, 1984: Depth-area ratios in the semi-arid southwest United States. *NOAA Technical Memorandum NWS HYDRO-40*, Office of Hydrology, National Weather Service, Silver Spring, Maryland.