

## **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 regimes).

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 length of record 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 analysis (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).

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 is 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, and a real-data-check (RDC) of quantiles, among others. Also, analyses such as a trend analysis of annual maximum series, a study of cross-correlation between stations, 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 5 miles in horizontal distance and 300 feet in elevation with non-concurrent records 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 to ensure the samples were from the same population and appropriate to be merged. Hourly stations did not meet these criteria and so were not merged.

**Discordancy.** The L-moment discordancy measure 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 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 n-minute, 1-hour, 24-hour, and some longer durations (typically the 10-day). Appendix A.7 which provides a list of stations used in the project also provides the discordancy measure for the 24-hour data or 60-minute data for each station in its region.

**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 assured 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 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 necessarily 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 20 years of data was considered. It was found that the degree of cross correlation between stations in the project area was very low. Only 7% of the data in the entire project area showed strong correlation (correlation coefficient  $\rho \geq 0.7$ ). The impact of cross correlation on the daily quantiles was very small. Relative errors were calculated between the results of an analysis using all stations versus an analysis using only stations that were not cross-correlated. The relative errors were minimal, 1.6% and 3.7% for 100-year and 1,000-year, respectively. Therefore, since the final quantiles were only minimally affected, it was concluded that it was not necessary to embed any measures to address dependence structures in the data.

#### **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 distributions 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 defined 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 designated regions were then confirmed by statistical homogeneity tests and other checks. In particular, 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. 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 regions for daily durations (24-hour through 60-day), Figure 4.4.1, were based on the 24-hour duration. 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 hourly regions, Figure 4.4.2, were based on the 60-minute data. The other short durations (2-hour through 24-hour) where H1 was greater than 2 were also closely examined to validate data quality. Given the lack of station density and the nature of precipitation in the semiarid southwest, it was particularly difficult to adhere to a threshold of 2, which was proposed as a conservative guideline, for the hourly data. In each case where the H1 measure was greater than 2, after validating data quality, tests were conducted where 1 to 3 stations were omitted. In each case, 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. Given the geographic locations of the stations and the validity of their data, the suspect stations were often retained in the region and the region was accepted as is, regardless of its high H1.

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, in this case the 100-year estimate. These stations were carefully investigated for data quality and appropriate regionalization. “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).

Overall, effort was made during the subdivision process to mitigate discrepancies that could be caused by (1) sample 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 optimal quantiles 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.8 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 Appendices A.9.

Figure 4.4.1. Regional groupings for daily data used to prepare NOAA Atlas 14 Volume 1.

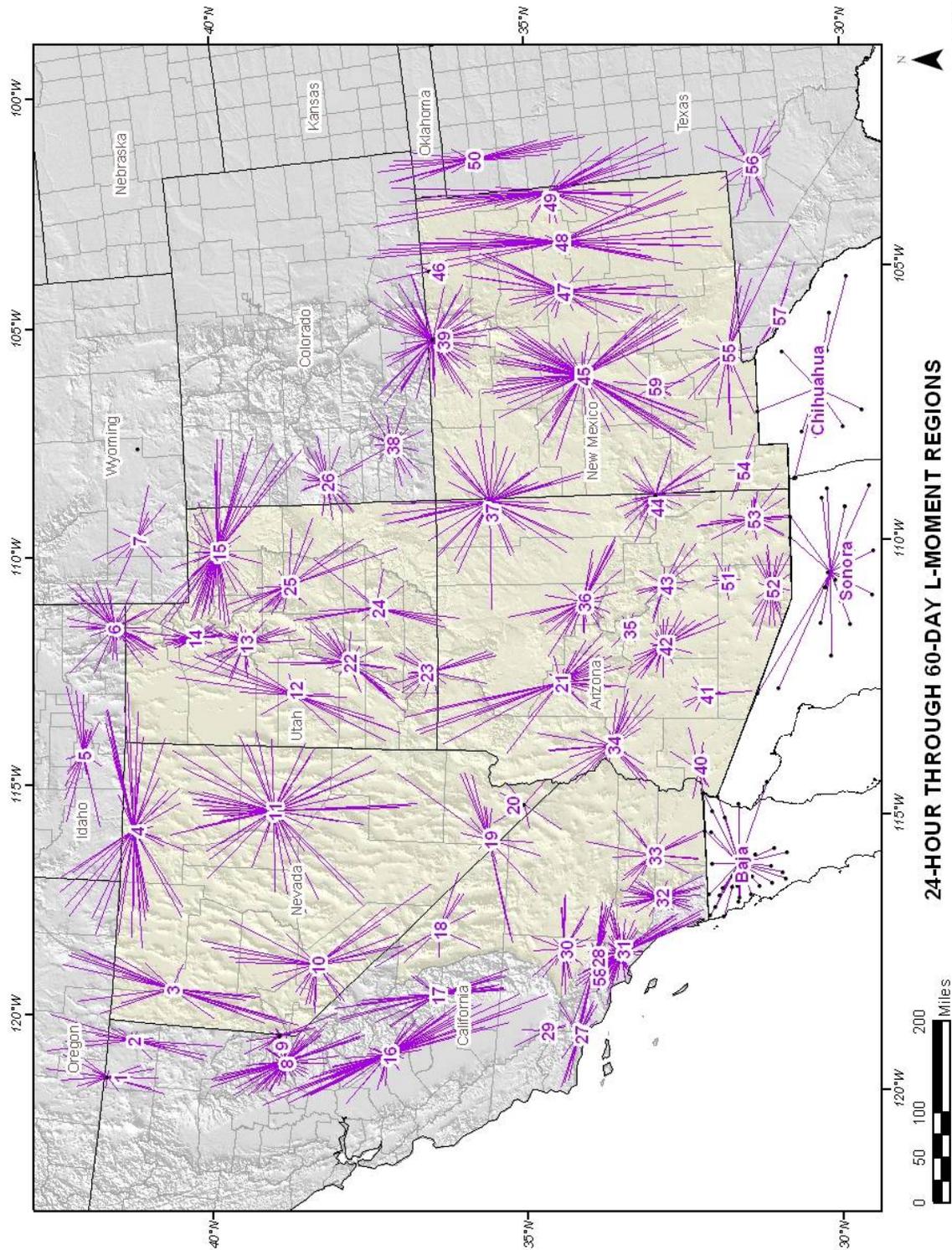
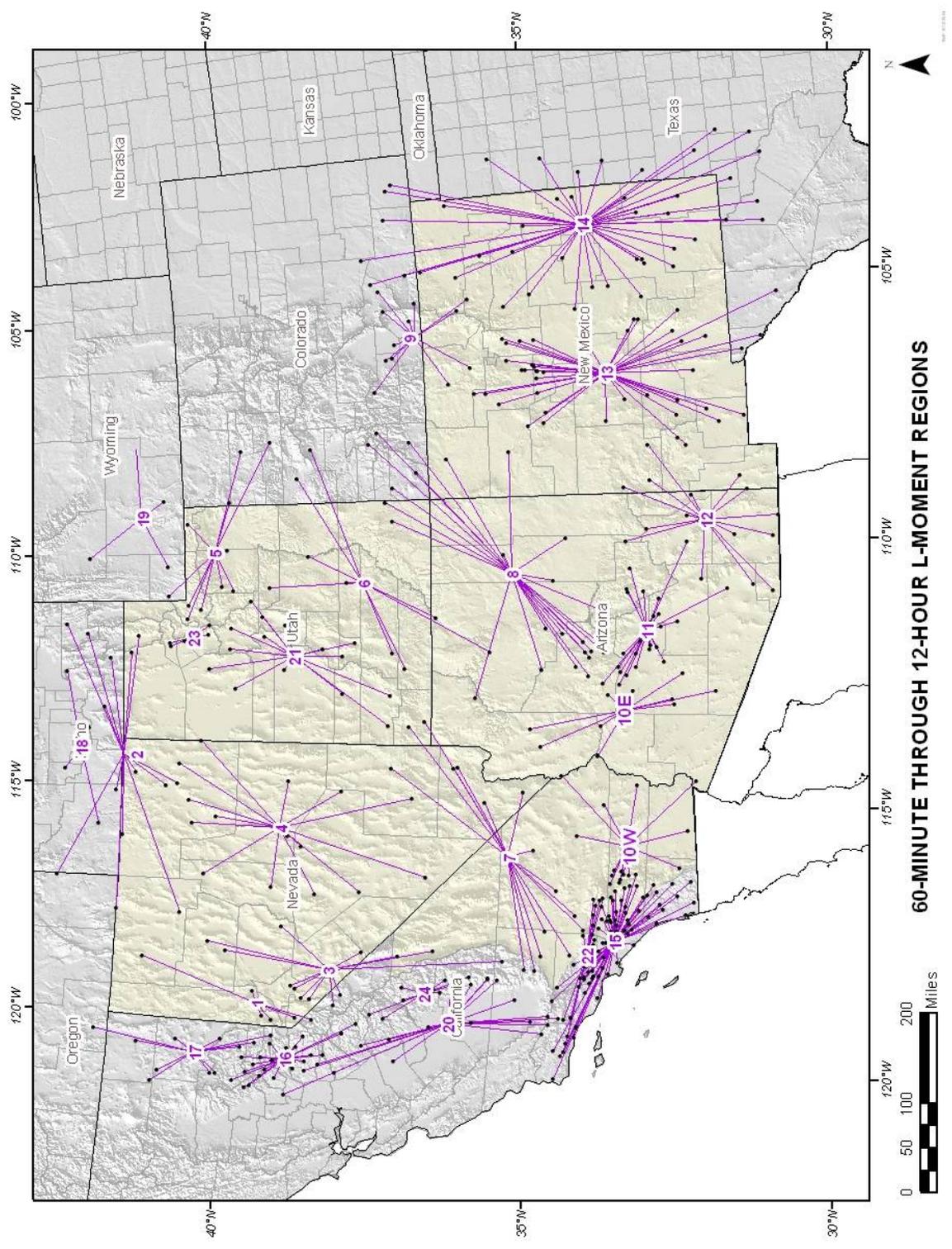


Figure 4.4.2. Regional groupings for hourly data used to prepare NOAA Atlas 14 Volume 1.



**At-site stations.** At some daily stations an at-site, instead of a regional, frequency analysis was a better approach to estimating the precipitation frequency quantiles. There were no hourly at-sites in the project. At-site stations were used because:

- They accounted for observed extreme precipitation regimes that the regional method could not resolve;
- They had enough data years to produce accurate estimates independent of a region;
- The spatial interpolation process was able to accommodate them;
- Error in the estimate was reduced compared to when included in a region.

Although at-sites have advantages in some cases, their use was considered a last-resort option because their precipitation frequency estimates sometimes caused irregularities in the spatial interpolation. All attempts to include a station in a region were considered before it was analyzed as an at-site. In fact, at-site stations had to meet at least 4 of the following criteria:

- Observed station data were markedly atypical and did not conform to adjacent regions;
- The at-site station caused adjacent regions that it would otherwise belong to be heterogeneous;
- The root mean-square-error (RMSE) for a region was less than if the station was included in the region;
- The at-site station was flagged during the discordancy check or the “real data check;”
- The at-site station had at least 50 data years (in most cases they actually had more than 80 data years);
- The absence of the at-site station in an adjacent region did not greatly impact final regional precipitation frequency estimates;
- There was a compelling local climatological or topographical reason to support an at-site analysis.

Empirical frequency plots provided a tool for assessing the accuracy of chosen distributions at a given station. In the case of at-sites, the difference between the empirical frequencies and the theoretical distribution precipitation frequency estimates, effectively the root-mean-square-error (RMSE), was much smaller from the at-site analysis than if the station was included in a region. For instance, figure 4.4.3 shows the empirical distribution for Bosque Del Apache, NM as an at-site.

Because at-site stations are often statistical exceptions and they ultimately influence the spatial pattern in an area, they were carefully investigated. However, the spatial impact of the at-site stations, if any, was mitigated by spatial smoothing. The smoothing helped to spatially blend the at-site precipitation frequency estimates with those derived from the regional-approach.

For NOAA Atlas 14 Volume 1, 5 daily stations and one small group of stations were analyzed using an at-site analysis (Table 4.4.1). They are labeled A1 through A6. A1 and A6 are outside of the core domain and therefore are not specifically addressed in this documentation.

Table 4.4.1. Stations analyzed using an at-site analysis.

<b>At-site</b>	<b>Station ID</b>	<b>Station Name</b>	<b>Data years</b>
A1	05-6524	Placerville, CO	53
A2	29-0818	Beaverhead, NM	56
A3	29-1138	Bosque del Apache, NM	102
A4	29-8535	State University, NM	109
A5	42-5733	Moab Radio, UT	108
A6	04-2504 & 04-2506	Doyle & Doyle 4 SSE, CA	74 & 44

The following is a brief discussion of the core area at-site stations:

- A2. Beaverhead, NM (29-0818):  
Observed precipitation at 29-0818 was not consistent with its vicinity. The heterogeneity was -0.06 for Region 44 without 29-0818, but worse (1.73) for Region 44 with 29-0818. The precipitation frequency estimates in Region 44 remained nearly the same with and without 29-0818. The empirical frequencies versus the theoretical probability precipitation frequency estimates suggested that an at-site resulted in reduced RMSE. And finally, the resulting spatial pattern when using an at-site analysis was consistent with the surrounding area at this location.
- A3. Bosque Del Apache, NM (29-1138):  
This at-site station was analyzed more than any other station in the project. Several attempts to include it in nearby regions, including region 59, failed. Climatological evidence suggests the area around Bosque Del Apache is prone to extreme events, with Bosque Del Apache being the epicenter of the risk. To mitigate the spatial bulls eye associated with the high  $\geq$  24-hour precipitation frequency estimates at Bosque Del Apache, region 59 was formed out of the stations around Bosque Del Apache. The at-site and region 59 are prone to two moisture sources which are consistent with Figure 7 in NOAA Atlas 2 and evaluation of synoptic maps during extreme events: Monsoonal flow from the south and Gulf of Mexico moisture from the southeast. Most of region 59 and Bosque Del Apache reside in the Jornada Del Muerto of New Mexico, which is a large, flat basin between two northeast-southwest oriented mountain ranges. The terrain is such that moisture is funneled into this area from the south or southeast, subjected to orographic lifting contributing to extreme precipitation and trapped by the higher terrain to the north. Regardless of the moisture source, the extreme precipitation events are primarily associated with localized thunderstorms. This unique climate and topography climatologically justified region 59 and the Bosque Del Apache at-site. The empirical frequencies versus the theoretical probability precipitation frequency estimates suggested that an at-site analysis resulted in lower RMSE. Figure 4.4.3 shows the empirical distribution for Bosque Del Apache, NM
- A4. State University, NM (29-8535):  
With 109 data years and unique precipitation characteristics, this station was analyzed as an at-site. One advantage of this at-site is that it accounts for the unique extreme precipitation data while conforming to a consistent spatial pattern. In other words, the at-site estimates are consistent with the surroundings.
- A5. Moab Radio, UT (42-5733):  
Moab, UT is an isolated valley at an elevation of around 4000 feet. Some of the surrounding mountains surpass 12,000 feet on its east and southeast sides. This relatively sheltered location creates the possibility for unique extreme precipitation climate conditions that are different from the surrounding region. Differential heating of mountain slopes leading to intense local convection, other orographic effects, and advection of Monsoon moisture into the Moab Valley all contribute to the enhancement of extreme precipitation at this location. Indeed, Moab has observed at least 3 cases of localized extreme precipitation causing high variation in the data at Moab. This unique climate and topography justified computing precipitation frequency estimates for the station in an at-site analysis.

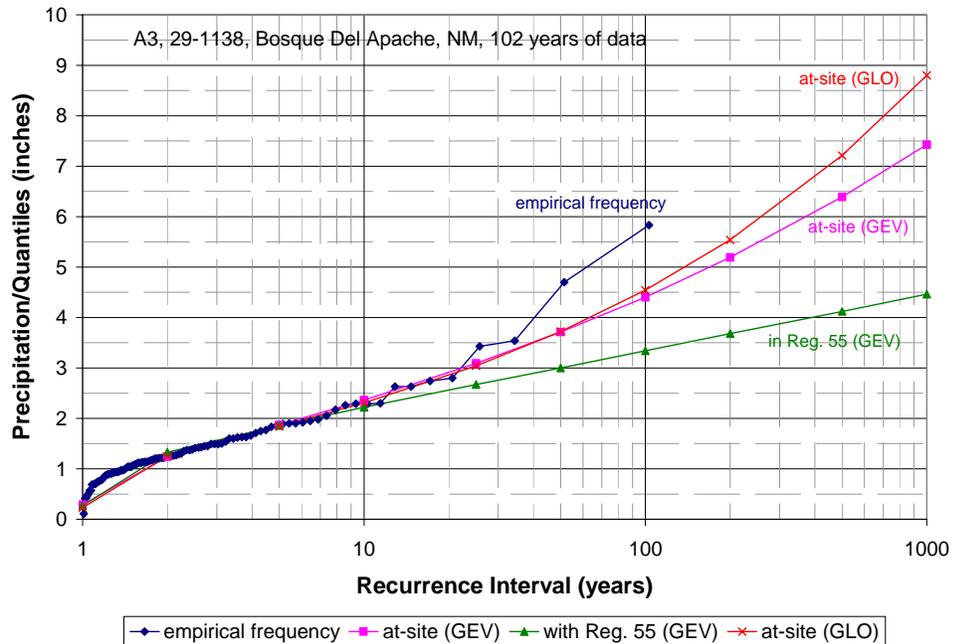


Figure 4.4.3. Empirical frequency plot of Bosque Del Apache, NM comparing at-site and regional analyses.

Since at-site stations accounted for localized 24-hour or longer duration extreme precipitation regimes, their precipitation frequency estimates sometimes did not relate well to the spatially interpolated hourly precipitation frequency estimates. In other words, the hourly interpolated estimates were lower than the at-site elevated estimates, therefore causing a “jump” from 12-hours to 24-hours. To make the precipitation frequency estimates temporally consistent, hourly pseudo data (Section 4.8.3) was created for Bosque Del Apache, NM; Moab Radio, UT and Doyle 4 SSE, CA.

#### 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 *between* regions. In other words, a probability distribution was selected and dimensionless 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 was considered only if the three-parameter distributions were unsuitable for a region, which rarely ever happened. 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, RMSE of the sample L-moments, and real-data-check test.

**The Monte Carlo Simulation test.** Using Monte Carlo simulation and the sample L-moment estimates for each station in a region, 1,000 synthetic data sets with the same record length and sample L-moments at each station were generated. 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 (Hosking, 1991).

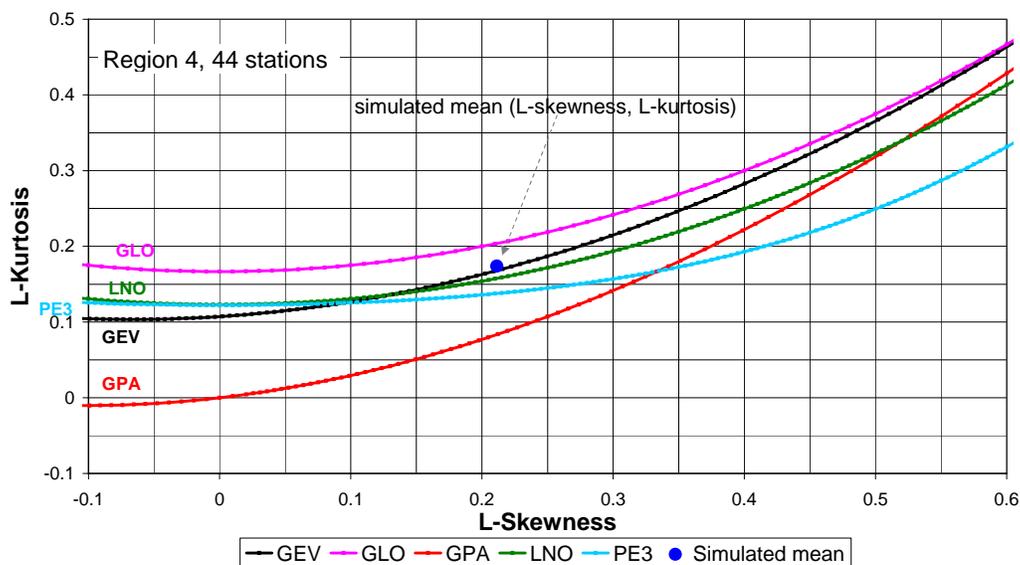


Figure 4.5.1. Plot of mean point from Monte Carlo simulations and theoretical distributions in L-skewness versus L-kurtosis diagram.

**RMSE of the sample L-moments.** Unlike the Monte Carlo simulation test that emphasizes the effect of 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.

**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.

A final decision of the most appropriate distribution for a region was aided by 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 59 daily regions and 6 at-sites. Table 4.5.2 shows the results for the 60-minute data in each of the 25 hourly regions. The results from the three tests provide a strong statistical basis for selecting the most appropriate distribution. Based on the goodness-of-fit results for all regions in the project area, GEV was selected to best represent the underlying distributions of all daily and hourly annual maximum data. GEV was also selected for the 5-, 10-, and 15-minute data and GNO was selected for the 30-minute annual maximum data that were used in the calculation of the n-minute ratios.

The at-site stations were extensively tested for the most appropriate distribution for all durations, since by their nature they are not consistent with the regional approach and required special treatment. It was found that for one at-site station within the core project area, A3, different distributions were most appropriate for different durations. GLO was selected for the 24-hour through 30-day durations for at-site A3 (29-1138) and GEV was selected for 45-day and 60-day.

Table 4.5.1. Goodness-of-fit test results for 24-hour data in each daily region used to prepare NOAA Atlas 14 Volume 1.

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
1	1st	GEV	-0.42	GLO	22.5	GEV	0.12795	GEV
	2nd	GNO	-0.92	GEV	18.0	GNO	0.13153	
	3rd	GUM	-1.16	GNO	16.0	GLO	0.13598	
2	1st	GLO	0.90	GEV	19.0	GEV	0.13807	GEV
	2nd	GEV	-1.13	GNO	17.0	GLO	0.13956	
	3rd	GNO	-2.11	GLO	16.0	GNO	0.14005	
3	1st	GEV	-0.33	GEV	21.5	GNO	0.10771	GEV
	2nd	GNO	-1.09	GNO	20.5	GEV	0.10842	
	3rd	GUM	-1.63	PE3	13.0	PE3	0.11205	
4	1st	GEV	-1.02	GNO	18.5	GEV	0.09502	GEV
	2nd	GNO	-1.97	GEV	18.5	GNO	0.09689	
	3rd	GUM	-2.46	PE3	17.0	GLO	0.10194	
5	1st	GEV	-0.85	GEV	21.0	GEV	0.11629	GEV
	2nd	GNO	-1.67	GNO	20.5	GNO	0.11698	
	3rd	GUM	-2.12	PE3	16.0	PE3	0.12256	
6	1st	GEV	-1.93	GEV	20.5	GLO	0.10816	GEV
	2nd	GLO	1.93	GLO	18.5	GEV	0.10836	
	3rd	GNO	-2.91	GNO	17.5	GNO	0.11044	

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
7	1st	GNO	-0.21	PE3	17.5	GNO	0.17183	GEV
	2nd	GEV	0.42	GNO	16.5	GEV	0.17281	
	3rd	GUM	-0.81	GEV	14.5	PE3	0.17348	
8	1st	GEV	0.09	PE3	20.0	GNO	0.08923	GEV
	2nd	GUM	-0.80	GEV	18.5	GEV	0.08975	
	3rd	GNO	-0.92	GNO	17.5	PE3	0.09234	
9	1st	GEV	-0.22	GEV	20.5	GNO	0.12301	GEV
	2nd	GNO	-0.98	GNO	18.5	GEV	0.12350	
	3rd	GUM	-1.73	GLO	17.0	PE3	0.12672	
10	1st	GEV	-1.54	GEV	20.0	GEV	0.08236	GEV
	2nd	GLO	1.73	GNO	19.0	GNO	0.08428	
	3rd	GNO	-2.33	GLO	16.0	GLO	0.08663	
11	1st	GEV	-1.24	GEV	22.0	GEV	0.08419	GEV
	2nd	GNO	-2.42	GNO	18.0	GNO	0.08519	
	3rd	GUM	-3.18	GLO	16.0	PE3	0.09176	
12	1st	GEV	-1.01	PE3	18.0	GEV	0.14403	GEV
	2nd	GUM	-1.32	GEV	17.5	GNO	0.14504	
	3rd	GNO	-1.47	GNO	16.0	GLO	0.14907	
13	1st	GLO	1.67	GLO	22.5	GEV	0.06946	GEV
	2nd	GEV	-2.48	GEV	20.0	GLO	0.07001	
	3rd	GUM	-2.56	GNO	17.0	GNO	0.07188	
14	1st	GEV	0.08	GNO	19.5	GEV	0.08189	GEV
	2nd	GUM	-0.62	PE3	19.0	GNO	0.08267	
	3rd	GNO	-0.64	GEV	15.5	PE3	0.08631	
15	1st	GEV	-1.27	GEV	21.0	GEV	0.06844	GEV
	2nd	GNO	-2.63	GNO	20.0	GNO	0.07128	
	3rd	GLO	2.64	PE3	16.0	GLO	0.07612	
16	1st	GEV	-2.52	GEV	24.5	GEV	0.06716	GEV
	2nd	GLO	3.17	GNO	18.0	GNO	0.07304	
	3rd	GNO	-3.62	GLO	15.0	GLO	0.07467	
17	1st	GLO	0.81	GNO	18.5	GEV	0.09861	GEV
	2nd	GEV	-1.86	GEV	18.5	GLO	0.09909	
	3rd	GNO	-3.17	GLO	18.0	GNO	0.10213	
18	1st	GNO	0.02	PE3	19.5	GNO	0.15977	GEV
	2nd	GEV	0.82	GNO	19.5	PE3	0.16119	
	3rd	GUM	-1.17	GPA	14.5	GPA	0.16197	
19	1st	GUM	-0.75	GLO	19.5	GEV	0.08115	GEV
	2nd	GEV	-1.51	GNO	16.5	GNO	0.08257	
	3rd	GNO	-1.60	GEV	16.5	GLO	0.08803	
20	1st	GLO	0.02	GNO	19.5	GEV	0.19198	GEV
	2nd	GEV	-0.99	GEV	19.5	GLO	0.19285	

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
	3rd	GNO	-1.62	PE3	14.5	GNO	0.19447	
21	1st	GEV	-1.49	GEV	19.5	GEV	0.06105	GEV
	2nd	GNO	-2.30	GNO	17.5	GNO	0.06697	
	3rd	GUM	-2.31	GLO	17.5	GLO	0.07256	
22	1st	GLO	1.77	PE3	19.0	GEV	0.05648	GEV
	2nd	GEV	-1.89	GNO	19.0	GNO	0.05958	
	3rd	GNO	-2.71	GEV	18.0	GLO	0.06004	
23	1st	GEV	-0.08	GEV	20.5	GEV	0.12292	GEV
	2nd	GNO	-0.78	GNO	18.5	GNO	0.12502	
	3rd	GUM	-1.15	GLO	16.0	GLO	0.12971	
24	1st	GEV	-0.80	GEV	20.0	GEV	0.15892	GEV
	2nd	GLO	1.23	GLO	17.0	GNO	0.16155	
	3rd	GNO	-1.50	GNO	16.5	GLO	0.16249	
25	1st	GEV	0.10	GNO	18.0	GEV	0.09318	GEV
	2nd	GUM	0.17	PE3	17.0	GNO	0.09472	
	3rd	GNO	-0.24	GEV	17.0	PE3	0.10036	
26	1st	GEV	-0.79	GEV	19.5	GEV	0.10688	GEV
	2nd	GNO	-1.61	GLO	18.0	GNO	0.10735	
	3rd	GLO	1.87	GNO	17.0	PE3	0.11193	
27	1st	PE3	0.36	PE3	22.0	PE3	0.11129	GEV
	2nd	NOR	-0.45	GNO	18.0	GNO	0.11312	
	3rd	GNO	1.54	GEV	17.0	GEV	0.11405	
28	1st	GEV	-0.41	GNO	21.0	GEV	0.09215	GEV
	2nd	GNO	-1.59	GEV	21.0	GNO	0.09349	
	3rd	GUM	-2.79	PE3	14.0	PE3	0.10130	
29	1st	GLO	-0.06	GEV	19.5	GEV	0.17908	GEV
	2nd	GEV	-1.40	GNO	18.5	GLO	0.18107	
	3rd	GNO	-2.02	PE3	15.0	GNO	0.18131	
30	1st	PE3	-0.53	PE3	22.0	PE3	0.09627	GEV
	2nd	GUM	0.85	GNO	17.0	GNO	0.09635	
	3rd	GNO	1.04	GPA	13.0	GEV	0.09733	
31	1st	PE3	0.05	PE3	18.5	PE3	0.06318	GEV
	2nd	NOR	-2.43	GNO	16.5	GNO	0.06446	
	3rd	GNO	3.51	GEV	16.0	GEV	0.06612	
32	1st	GNO	0.09	PE3	20.5	GNO	0.08061	GEV
	2nd	GUM	-1.46	GNO	18.5	GEV	0.08254	
	3rd	GEV	1.47	GEV	15.5	PE3	0.08370	
33	1st	GLO	0.59	GEV	20.0	GLO	0.11650	GEV
	2nd	GEV	-1.65	GLO	19.5	GEV	0.11676	
	3rd	GNO	-2.73	GNO	16.5	GNO	0.11872	
34	1st	GLO	0.97	GEV	22.0	GEV	0.11298	GEV

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
	2nd	GEV	-1.68	GLO	18.0	GLO	0.11380	
	3rd	GNO	-2.49	GNO	16.5	GNO	0.11702	
35	1st	GEV	0.01	GNO	18.5	GNO	0.21691	GEV
	2nd	GNO	-0.53	GEV	18.5	GEV	0.21803	
	3rd	GLO	1.18	GLO	17.0	PE3	0.21869	
36	1st	GEV	-0.32	GEV	20.5	GEV	0.09814	GEV
	2nd	GNO	-1.22	GLO	18.5	GNO	0.10126	
	3rd	GUM	-1.57	GNO	17.5	GLO	0.10746	
37	1st	GEV	-0.63	GLO	18.5	GEV	0.09089	GEV
	2nd	GNO	-1.84	GEV	18.0	GNO	0.09312	
	3rd	GUM	-2.77	GNO	16.0	PE3	0.09986	
38	1st	GEV	-0.78	GEV	20.0	GEV	0.11280	GEV
	2nd	GUM	-1.10	GLO	17.5	GNO	0.11447	
	3rd	GNO	-1.30	GNO	16.5	GLO	0.11937	
39	1st	GEV	-0.33	GNO	19.0	GEV	0.07051	GEV
	2nd	GNO	-1.66	PE3	18.0	GNO	0.07422	
	3rd	GUM	-2.50	GEV	17.0	PE3	0.08397	
40	1st	GEV	0.04	GEV	19.0	GEV	0.14036	GEV
	2nd	GNO	-0.73	GNO	18.0	GNO	0.14086	
	3rd	GLO	1.63	GLO	15.5	PE3	0.14594	
41	1st	GLO	0.47	GLO	22.5	GEV	0.16359	GEV
	2nd	GEV	-1.21	GEV	18.0	GLO	0.16651	
	3rd	GNO	-1.61	GNO	14.5	GNO	0.16677	
42	1st	GNO	-0.36	GEV	20.5	GEV	0.08687	GEV
	2nd	GEV	0.75	GNO	17.5	GNO	0.08715	
	3rd	GUM	-1.01	PE3	16.0	PE3	0.09243	
43	1st	GEV	-0.55	GNO	21.0	GEV	0.10722	GEV
	2nd	GNO	-1.22	GEV	19.0	GNO	0.10763	
	3rd	GUM	-1.55	PE3	17.0	PE3	0.11140	
44	1st	GEV	-1.03	GEV	23.5	GEV	0.09660	GEV
	2nd	GUM	-1.46	GNO	18.5	GNO	0.09779	
	3rd	GNO	-1.61	GLO	15.0	GLO	0.10121	
45	1st	GEV	-2.20	GEV	21.5	GEV	0.07639	GEV
	2nd	GUM	-2.88	GNO	18.5	GNO	0.07899	
	3rd	GNO	-3.15	GLO	18.0	GLO	0.08315	
46	1st	GEV	0.01	GLO	19.5	GEV	0.23419	GEV
	2nd	GNO	-0.66	GNO	16.5	GNO	0.23598	
	3rd	GLO	0.86	GEV	16.0	GLO	0.23700	
47	1st	GEV	-1.17	GEV	23.0	GEV	0.08716	GEV
	2nd	GLO	2.15	GNO	19.0	GNO	0.08908	
	3rd	GNO	-2.24	GLO	15.0	GLO	0.09399	

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
48	1st	GUM	-0.31	GNO	22.5	GNO	0.08534	GEV
	2nd	GNO	1.00	PE3	19.0	GEV	0.08638	
	3rd	PE3	-1.54	GEV	18.5	PE3	0.08843	
49	1st	GNO	-0.68	GEV	20.5	GNO	0.08092	GEV
	2nd	GEV	0.74	GNO	19.5	GEV	0.08095	
	3rd	GUM	-2.07	PE3	16.5	PE3	0.08639	
50	1st	GLO	0.88	GEV	19.0	GEV	0.09805	GEV
	2nd	GEV	-1.66	GNO	17.0	GLO	0.10052	
	3rd	GNO	-2.50	GLO	17.0	GNO	0.10112	
51	1st	GLO	0.66	GLO	18.0	GEV	0.14030	GEV
	2nd	GEV	-1.62	GEV	18.0	GLO	0.14130	
	3rd	GNO	-2.18	GNO	16.5	GNO	0.14219	
52	1st	GNO	0.44	GNO	21.5	GEV	0.10154	GEV
	2nd	GUM	0.56	GEV	20.5	GNO	0.10186	
	3rd	PE3	-1.04	PE3	18.0	PE3	0.10338	
53	1st	GNO	0.50	PE3	24.0	GNO	0.08635	GEV
	2nd	PE3	-0.81	GNO	18.5	GEV	0.08681	
	3rd	GEV	0.93	GEV	14.5	PE3	0.08700	
54	1st	GNO	-0.36	GLO	19.5	GNO	0.20462	GEV
	2nd	GEV	0.43	GEV	18.0	GEV	0.20488	
	3rd	GLO	1.69	GNO	16.5	PE3	0.20908	
55	1st	PE3	-0.13	PE3	21.0	GEV	0.11877	GEV
	2nd	NOR	-0.97	GNO	20.5	GNO	0.11969	
	3rd	GNO	1.09	GEV	14.0	PE3	0.12137	
56	1st	GLO	0.37	GEV	19.5	GEV	0.10186	GEV
	2nd	GEV	-1.82	GLO	18.0	GLO	0.10287	
	3rd	GNO	-2.80	GNO	16.5	GNO	0.10627	
57	1st	GEV	-0.32	GEV	17.0	GNO	0.15977	GEV
	2nd	GLO	0.76	GNO	16.5	GEV	0.16049	
	3rd	GNO	-0.94	PE3	15.0	GLO	0.16424	
58	1st	NOR	0.38	GPA	21.0	GPA	0.21246	GEV
	2nd	GPA	-0.61	PE3	18.0	PE3	0.21341	
	3rd	PE3	1.16	GNO	14.5	GNO	0.21686	
59	1st	GEV	-0.56	PE3	17.5	GEV	0.14145	GLO
	2nd	GUM	-0.70	GLO	16.5	GNO	0.14312	
	3rd	GNO	-0.88	GEV	16.5	GLO	0.14701	
A1	1st	GEV	-0.09	PE3	18.5	GEV	0.58763	GEV
	2nd	GLO	0.15	GPA	18.5	GNO	0.58811	
	3rd	GNO	-0.39	GNO	13.5	GLO	0.58817	
A2	1st	GNO	0.08	GNO	18.5	GNO	0.36387	GNO
	2nd	GEV	0.47	GEV	18.5	PE3	0.36430	

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
	3rd	PE3	-0.60	GLO	13.5	GEV	0.36499	
A3	1st	GLO	-1.00	GLO	18.0	GLO	0.53849	GLO
	2nd	GEV	-1.40	GEV	17.0	GEV	0.54077	
	3rd	GNO	-1.68	GNO	15.0	GNO	0.54337	
A4	1st	GLO	0.19	GNO	19.5	GEV	0.55521	GEV
	2nd	GEV	-0.22	PE3	16.0	GLO	0.55544	
	3rd	GNO	-0.57	GEV	15.5	GNO	0.55600	
A5	1st	GLO	-0.63	GNO	18.5	GLO	0.58584	GEV
	2nd	GEV	-0.91	GEV	17.5	GEV	0.58701	
	3rd	GNO	-1.24	GLO	14.0	GNO	0.58958	
A6	1st	GNO	0.25	GPA	19.0	GNO	0.39431	GNO
	2nd	GUM	-0.42	PE3	18.0	PE3	0.39445	
	3rd	PE3	-0.47	GNO	15.5	GEV	0.39518	

Table 4.5.2. Goodness-of-fit test results for 60-minute data in each hourly region used to prepare NOAA Alas 14 Volume 1.

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
1	1st	GLO	-0.42	GPA	17.5	GEV	0.26435	GEV
	2nd	GEV	-0.89	PE3	15.5	GLO	0.26451	
	3rd	GNO	-1.51	GNO	15.5	GNO	0.26654	
2	1st	GNO	0.07	PE3	22.0	GNO	0.14997	GEV
	2nd	GEV	1.13	GPA	22.0	PE3	0.15201	
	3rd	PE3	-1.75	GNO	14.5	GEV	0.15353	
3	1st	GLO	0.09	GNO	19.0	GEV	0.13290	GEV
	2nd	GEV	-0.82	GEV	17.0	GNO	0.13398	
	3rd	GNO	-1.68	PE3	13.5	GLO	0.13657	
4	1st	GLO	0.29	GPA	16.5	GEV	0.09848	GEV
	2nd	GEV	-0.70	PE3	15.5	GNO	0.10085	
	3rd	GNO	-1.60	GNO	15.0	GLO	0.10527	
5	1st	GEV	-0.17	GEV	17.0	GEV	0.19905	GEV
	2nd	GLO	0.70	GNO	16.5	GPA	0.20164	
	3rd	GNO	-1.00	GLO	16.5	GNO	0.20212	
6	1st	GEV	-0.12	GEV	19.0	GEV	0.12998	GEV
	2nd	GLO	0.73	GLO	18.0	GNO	0.13211	
	3rd	GNO	-0.92	GNO	16.0	GPA	0.13467	
7	1st	GEV	0.48	PE3	18.5	GNO	0.17481	GEV
	2nd	GNO	-0.51	GEV	18.5	GEV	0.17537	

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
	3rd	GLO	1.85	GNO	17.5	PE3	0.18091	
8	1st	GEV	0.42	GNO	19.0	GEV	0.10531	GEV
	2nd	GNO	-0.49	PE3	18.0	GNO	0.10642	
	3rd	GUM	-1.70	GEV	14.0	PE3	0.11274	
9	1st	GEV	0.10	GEV	17.0	GNO	0.15911	GEV
	2nd	GNO	-0.58	PE3	16.5	GEV	0.15918	
	3rd	GUM	-1.52	GNO	16.0	PE3	0.16294	
10E	1st	NOR	0.43	PE3	17.5	GPA	0.13788	GEV
	2nd	PE3	0.89	GPA	17.0	PE3	0.14033	
	3rd	GNO	1.55	GNO	14.0	GNO	0.14229	
10W	1st	GNO	-0.25	GNO	19.0	GNO	0.18046	GEV
	2nd	GEV	0.83	GEV	17.5	PE3	0.18299	
	3rd	GLO	1.98	PE3	15.0	GEV	0.18374	
11	1st	GEV	-0.18	GLO	22.5	GEV	0.11029	GEV
	2nd	GNO	-0.75	GEV	18.0	GNO	0.11413	
	3rd	GUM	-1.20	GNO	16.0	GLO	0.12386	
12	1st	GNO	0.04	GLO	18.5	GEV	0.11718	GEV
	2nd	GUM	0.24	GEV	18.0	GNO	0.11769	
	3rd	GEV	0.37	PE3	17.5	PE3	0.12000	
13	1st	PE3	0.39	PE3	18.5	PE3	0.07046	GEV
	2nd	NOR	-0.73	GNO	18.0	GNO	0.07110	
	3rd	GNO	1.98	GEV	17.0	GEV	0.07178	
14	1st	GUM	-0.26	GEV	23.0	GNO	0.10488	GEV
	2nd	GEV	0.33	GNO	20.0	PE3	0.10526	
	3rd	GNO	-0.38	PE3	14.0	GEV	0.10668	
15	1st	GLO	1.31	GEV	20.0	GEV	0.09094	GEV
	2nd	GEV	-2.60	GNO	18.0	GNO	0.09615	
	3rd	GNO	-4.11	GLO	16.0	GLO	0.09653	
16	1st	GLO	-1.67	GLO	20.0	GEV	0.13951	GEV
	2nd	GEV	-2.64	GEV	20.0	GLO	0.14120	
	3rd	GNO	-3.73	GNO	18.0	GNO	0.14662	
17	1st	GLO	0.09	GNO	18.0	GLO	0.18737	GEV
	2nd	GEV	-0.48	GEV	18.0	GEV	0.18750	
	3rd	GNO	-1.49	GPA	15.0	GNO	0.18944	
18	1st	PE3	-0.08	GNO	17.5	GPA	0.34593	GEV
	2nd	GUM	-0.17	GEV	17.5	PE3	0.34705	
	3rd	GPA	-0.68	PE3	15.0	GNO	0.35029	
19	1st	GNO	-0.13	PE3	17.0	PE3	0.21012	GEV
	2nd	GEV	0.30	GLO	15.5	GNO	0.21183	
	3rd	GUM	-0.76	GEV	15.5	GEV	0.21522	
20	1st	GEV	-0.47	GNO	18.0	GEV	0.14207	GEV

region	rank	Monte Carlo Simulation		Real-data-check test		RMSE test		selected
		distribution	test value	distribution	test value	distribution	RMSE	
	2nd	GLO	1.23	GEV	17.5	GNO	0.14653	
	3rd	GNO	-1.49	PE3	15.0	GLO	0.14908	
21	1st	GEV	0.32	GLO	19.5	GEV	0.16714	GEV
	2nd	GNO	-0.69	GEV	17.0	GNO	0.16908	
	3rd	GLO	1.55	GNO	14.5	GPA	0.17343	
22	1st	GEV	-0.38	GNO	17.0	GNO	0.12547	GEV
	2nd	GNO	-1.04	GEV	16.0	GEV	0.12816	
	3rd	GLO	1.58	PE3	15.5	PE3	0.12854	
23	1st	GNO	-0.11	GNO	20.5	GNO	0.22287	GEV
	2nd	GEV	0.64	GLO	15.0	GEV	0.22490	
	3rd	GLO	1.36	GEV	13.5	GPA	0.22550	
24	1st	GLO	-0.38	GEV	19.5	GEV	0.20494	GEV
	2nd	GEV	-1.27	GLO	18.5	GNO	0.20698	
	3rd	GNO	-2.06	GNO	17.0	GLO	0.20763	

## 4.6. Estimation of quantiles

### 4.6.1. Regional growth factors

Regional growth factors (RGFs) are computed by applying appropriate higher order moments of the selected probability distributions for a region. They are multiplied by the site-specific scaling factor to produce the quantiles at each frequency and duration. Because the higher order moments 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.10. The site-specific scaling factor used in this project was the mean of the annual maximum series at each observing station. 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. 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, practical adjustments 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 data sets, it was necessary to 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 estimate from the daily data was retained since it was based on more stations, generally had longer record lengths, and were less prone to under catch precipitation. 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 60-minute through 12-hour quantiles at the given hourly station. It adjusted the quantiles using ratios of the 24-hour mean annual maxima and the 100-year 24-hour regional growth factors (RGFs) of both stations.

**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 increased the quantile above the longer duration. 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 shorter duration quantiles was less than one for a given average recurrence interval. They were mitigated by distributing the surplus of the ratio, which was greater than 1.0, of the previous frequency at a constant slope to the inconsistent frequency and higher through 1,000-year, until it converged at 1.0 after 1,000-year. The adjusted ratios were then, appropriately, greater than or equal to 1.0. Table 4.6.1 shows an example of the 3-hour to 2-hour ratios for average recurrence intervals from 2-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.

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

<b>2-hour to 3-hour ratios</b>	2-yr	5-yr	10-yr	25-yr	50-yr	100-yr	200-yr	500-yr	1,000-yr
Before adjustment	1.022	1.011	1.009	1.004	1.000	0.994	0.992	0.984	0.979
After adjustment	1.022	1.011	1.009	1.004	1.003	1.003	1.002	1.002	1.001

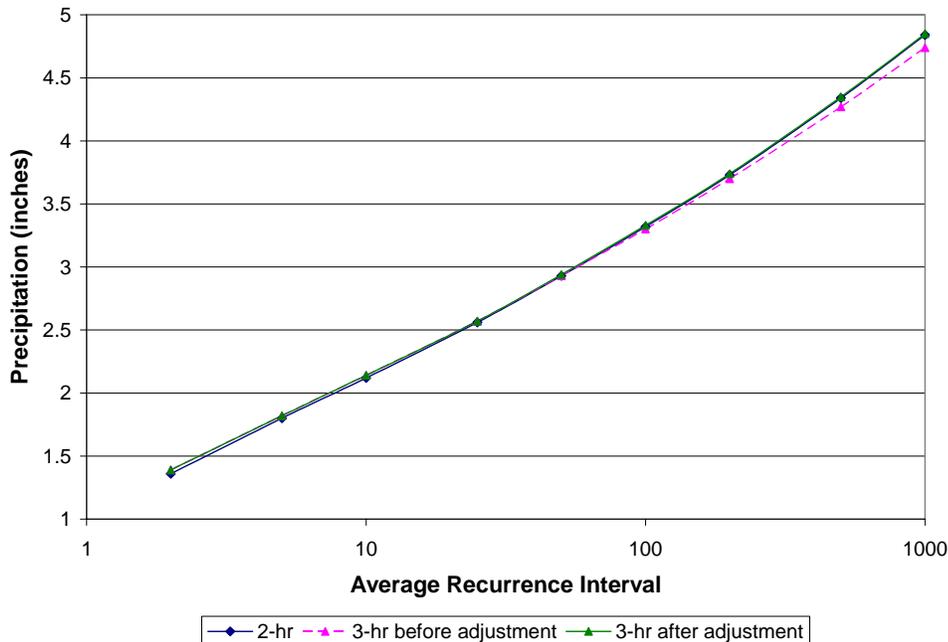


Figure 4.6.1. Example of internal consistency between the 3-hour and 2-hour quantiles at station 15-3709, Hazard, Kentucky.

In most cases, applying the adjustment from 2-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 2-year through 1,000-year to ensure consistency without changing the existing compliant quantiles.

#### 4.6.3. Conversion factors for AMS to PDS

Annual maximum series (AMS) data consist of the largest event in each year, regardless of whether the second largest event in a year exceeds the largest events of other years. In this project, the partial duration series (PDS) data is a subset of the complete data series where highest N events 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 occurrence, 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 59 homogenous regions in the project area. To account for sampling variability and to generate a smooth consistent curve, an asymptote of 1.004 was applied for 50-year and above.

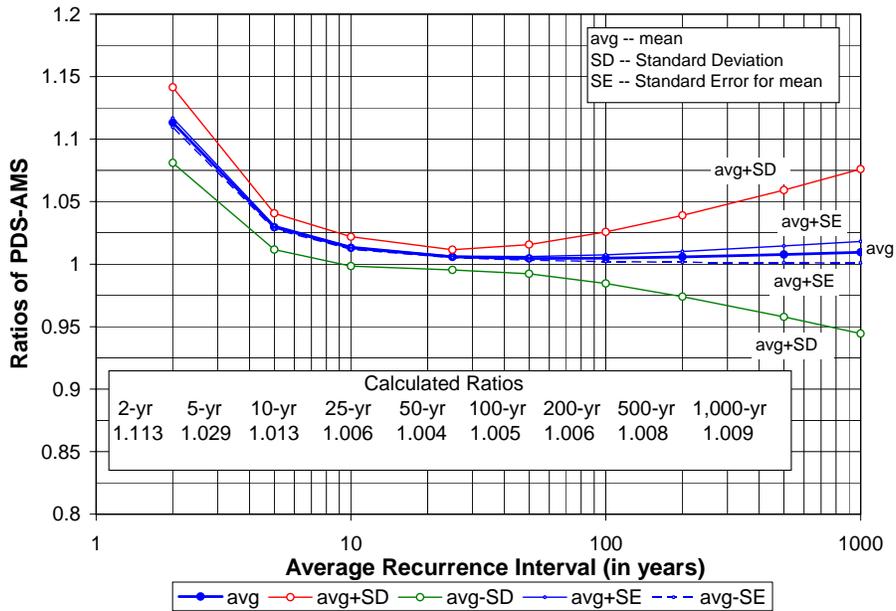


Figure 4.6.2. PDS-AMS ratio results for average recurrence intervals for the 24-hour duration over the 59 homogeneous regions used to prepare NOAA Atlas 14 Volume 1.

The ratios for this Atlas (Table 4.6.2) are consistent with NOAA Atlas 2 and theoretical computations. They are also consistent with results from the recently released Ohio River Basin and surrounding states precipitation frequency project (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, excluding at-site stations were used. Generalized Pareto (GPA) was selected as the most appropriate distribution for the PDS data in all but 9 regions. For regions 9, 24, 29, 33, 35, 50, 55, 56 and 59, Generalized Normal (GNO) was the best-fitting distribution.

Table 4.6.2. NOAA Atlas 14 Volume 1 PDS to AMS ratios for all durations with asymptote applied after 50-year.

2-year	5-year	10-year	25-year	50-year	100-year	200-year	500-year	1,000-year
1.113	1.029	1.013	1.006	1.004	1.004	1.004	1.004	1.004

#### **4.7. Estimation of confidence limits**

For the first time, 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 (n-minute to 60-minute mean precipitation frequency estimates) 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. Then, to maintain consistency within the confidence limits themselves, the confidence limits were adjusted where appropriate using the internal consistency check described in Section 4.6.2.