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 Spatial Climate Analysis Service (SCAS). SCAS 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 by using a naturally strong relationship with mean annual precipitation.

SCAS 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 about 10 percent for this Atlas. Appendix A.4 provides additional information regarding the details of the work done by SCAS 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 (30-second, or about 0.5 mile x 0.5 mile) 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 0.9986 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 (1.1658) 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.

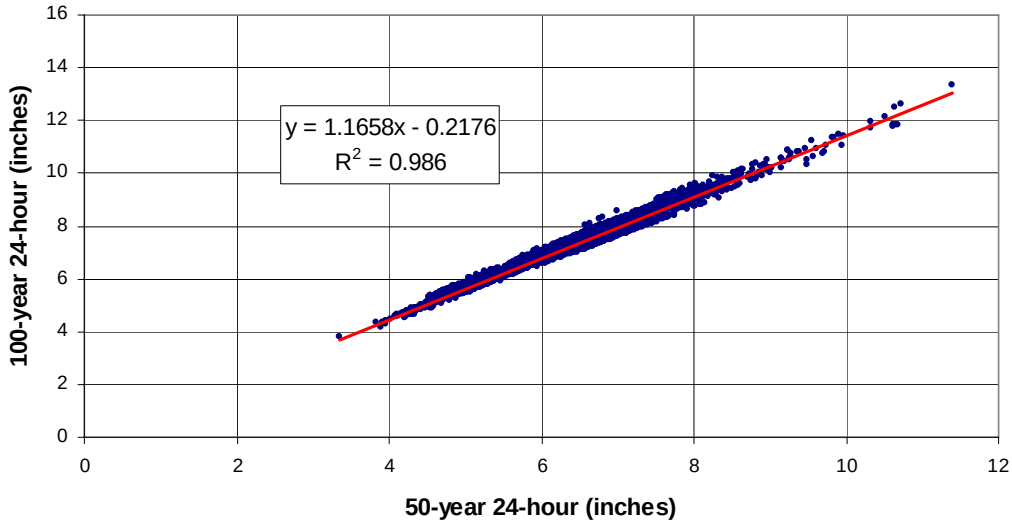


Figure 4.8.2. A scatter plot of 100-year 24-hour vs. 50-year 24-hour precipitation frequency estimates and the linear regression line from NOAA Atlas 14 Volume 2.

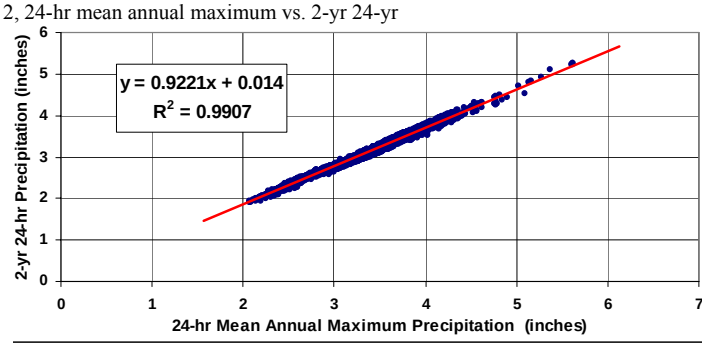
Step 1: Development of regression. The cascade began with the mean annual maximum grid derived by SCAS using PRISM for a given duration as the initial *predictor grid* (e.g., 24-hour mean annual maximum) and the 2-year frequency as the subsequent grid (e.g., 2-year 24-hour). All precipitation frequency 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 in an adjustment factor of 1.05 which is applied to each of the 60-minute precipitation frequencies (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.

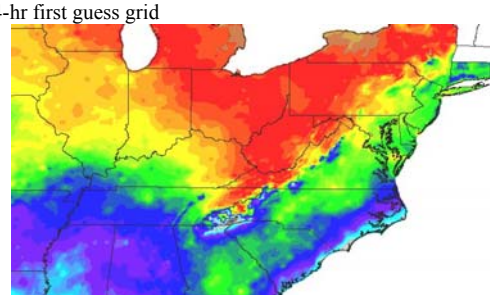
Figure 4.8.1. Flowchart of the cascade residual add-back (CRAB) grid derivation procedure beginning with the mean annual maximum grid of the x-duration and deriving the 2-year x-duration grid as an example.

Using PRISM, produce mean annual maximum grid for duration x (see accompanying documentation for details)

Build an x-y data file where initially **x is mean annual maximum** (for duration x) and **y the 2-year x-duration precipitation frequency estimate**. Calculate slope and y-intercept of a least-square fit linear regression line using x and y for ALL stations in the domain. Example: NOAA Atlas 14 Volume 2, 24-hr mean annual maximum vs. 2-yr 24-yr

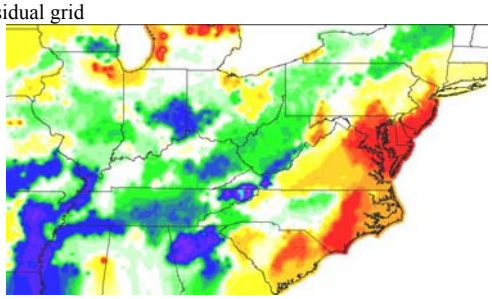


Based on slope, y-intercept and PRISM mean annual maximum grid, calculate a **first guess y grid**. Example:2-yr 24-hr first guess grid



Using the actual y value and the predicted y value, calculate the residual (actual minus predicted). Normalize the residual by dividing by the mean annual maximum, regardless of iteration.

Spatially interpolate the **normalized residual** values using inverse-distance weighting. Example:2-yr 24-hr normalized residual grid

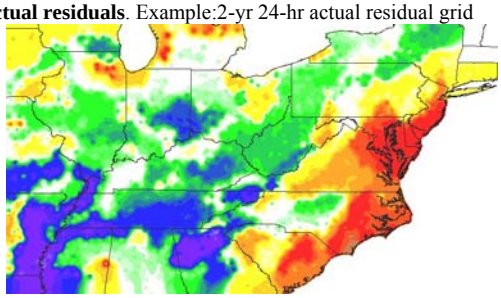


Iteration: Repeat for each frequency (e.g., mean annual maximum becomes 2yr, 2yr becomes 5yr, etc.). Then repeat for each duration (e.g., after the first mean annual maximum iteration, the duration starts at 60-minute, then 2-hour, etc.)

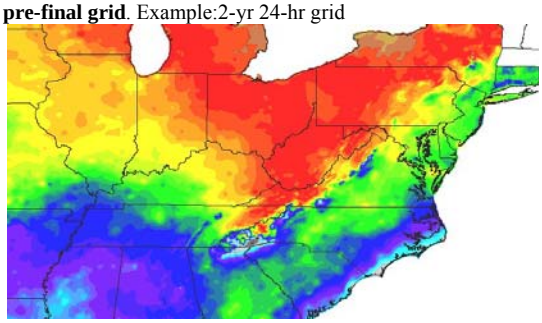
Continued on next page.

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 advanced spatial smoothing algorithm that smoothes estimates in areas with flat terrain and similar climates, but retains patterns in complex terrain where more variability is expected and appropriate. Then apply small (5x5 grid cell) center-weighted, block filter to the entire pre-final filtered grid to blend the smoothing threshold boundaries, lightly filter the coastlines and complex terrain, and 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*. To achieve a smoothed result, the spatial resolution was reduced from 30-seconds to 1-minute before spatially interpolating the normalized residuals with the IDW algorithm. Sensitivity tests were conducted to determine the optimum resolution to avoid "over-smoothing" the normalized residuals which would cause the maps to deviate from the quantile estimates achieved through the L-moment analysis. The results were then re-sampled back to a 30-second resolution for the remainder of the process. 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:

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

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).

The *normalized residual grid* was 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.

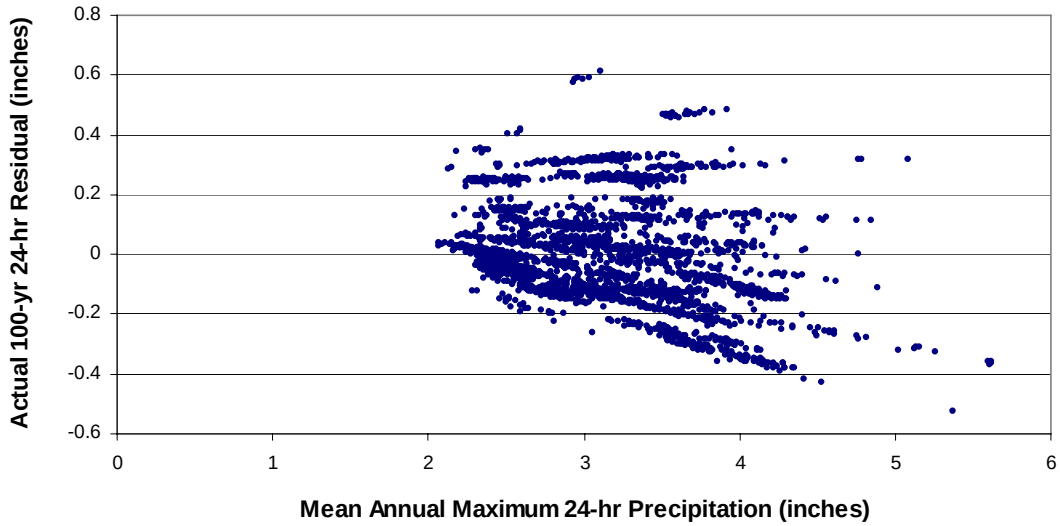


Figure 4.8.3. The relationship between the 100-year 24-hour *actual residuals* and the mean annual maximum precipitation from NOAA Atlas 14 Volume 2.

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 remove unnatural variability in the spatially distributed precipitation frequency estimates the *pre-final grid* was smoothed using an advanced spatial smoothing algorithm different than that used in NOAA Atlas 14 Volume 1. The algorithm smoothes estimates in areas with flat terrain, but retains patterns in complex terrain where more variability is expected and appropriate. The degree of spatial smoothing applied to a grid cell is dictated by its surrounding terrain and proximity to a coastline. In areas where terrain or the proximity of the coastline is important in defining patterns of precipitation, less spatial smoothing was applied.

To gauge the effectiveness of terrain to influence the precipitation frequency estimates, PRISM's effective terrain height grid was used. The effective terrain height grid, developed by the Spatial Climate Analysis Service (Daly and Neilson, 1992; Daly et al., 1994; Daly et al., 1997; Daly et al., 2002), is based on a 2.5-minute digital elevation model (DEM) in meters. The effective terrain height grid was prepared by first finding the minimum elevation within a 40-km radius of each grid cell. The minimum elevations were spatially averaged to produce a smooth, base elevation grid. Then, to obtain the effective terrain height grid, the base grid was subtracted from the original DEM grid and filtered to produce a smooth grid, which has units of meters. For more details, please refer to http://www.ocs.orst.edu/pub/prism/docs/effectiveterrain-daly.pdf.

For the CRAB process, the effective terrain height grid was smoothed further to prevent discontinuities at the boundaries of different degrees of spatial smoothing. It was spatially averaged

over a 40-km radius of each grid cell. To account for the impact of precipitation frequency estimate patterns as a result of coastal influences, PRISM's coastal proximity grid was used. The coastal proximity grid, also produced by Spatial Climate Analysis Service, is composed of grid cell values denoting a measure of the shortest distance from a cell to the water (Daly and Neilson, 1992; Daly et al., 1994; Daly et al., 1997; Daly et al., 2002). The measure is the distance to the nearest ocean pixel, divided into 10 distance classes. For more details, visit: http://www.ocs.orst.edu/pub/prism/docs/prisguid.pdf.

Based on the effective terrain height and coastal proximity grids, HDSC developed three degrees of initial spatial smoothing: heavy, moderate and none. The map in Figure 4.8.4 indicates the areas receiving the different degrees of smoothing.

1. HEAVY: Flat areas were determined if effective terrain height is less than 100 m (328 ft), and then a 17x17 grid cell (approximately 15 miles by 15 miles), center-weighted filter was used at the longer durations and a 25x25 grid cell (approximately 25 miles by 25 miles) filter at the shorter (<24-hour) durations. The shorter durations were subjected to greater smoothing because the lower station density was prone to cause unnatural variability.

2. MODERATE: Moderately complex terrain areas were determined if effective terrain height was greater than 100 meters (328 feet) and less than 200 meters (656 feet), and then a 11x11 grid cell (approximately 5.5 miles x 5.5 miles), center weighted filter was used for all durations.

3. LIGHT: Complex terrain areas and coastlines were determined if effective terrain height was greater than 200 meters (656 feet) or if the coastal proximity grid (a grid of values indicating distance from coast) was <=5, and then no filter was used at this stage. However, light smoothing was conducted during the next stage.

Once the above filtering was complete, a final 5x5 grid cell (approximately 2.5 mile by 2.5 mile), center-weighted filter was applied to the entire grid to blend the smoothing threshold boundaries, lightly filter the coastlines and complex terrain, remove extraneous "noise" in the spatial interpolation and promote smooth contour lines when interpolated, thus creating the *smoothed pre-final grid*.

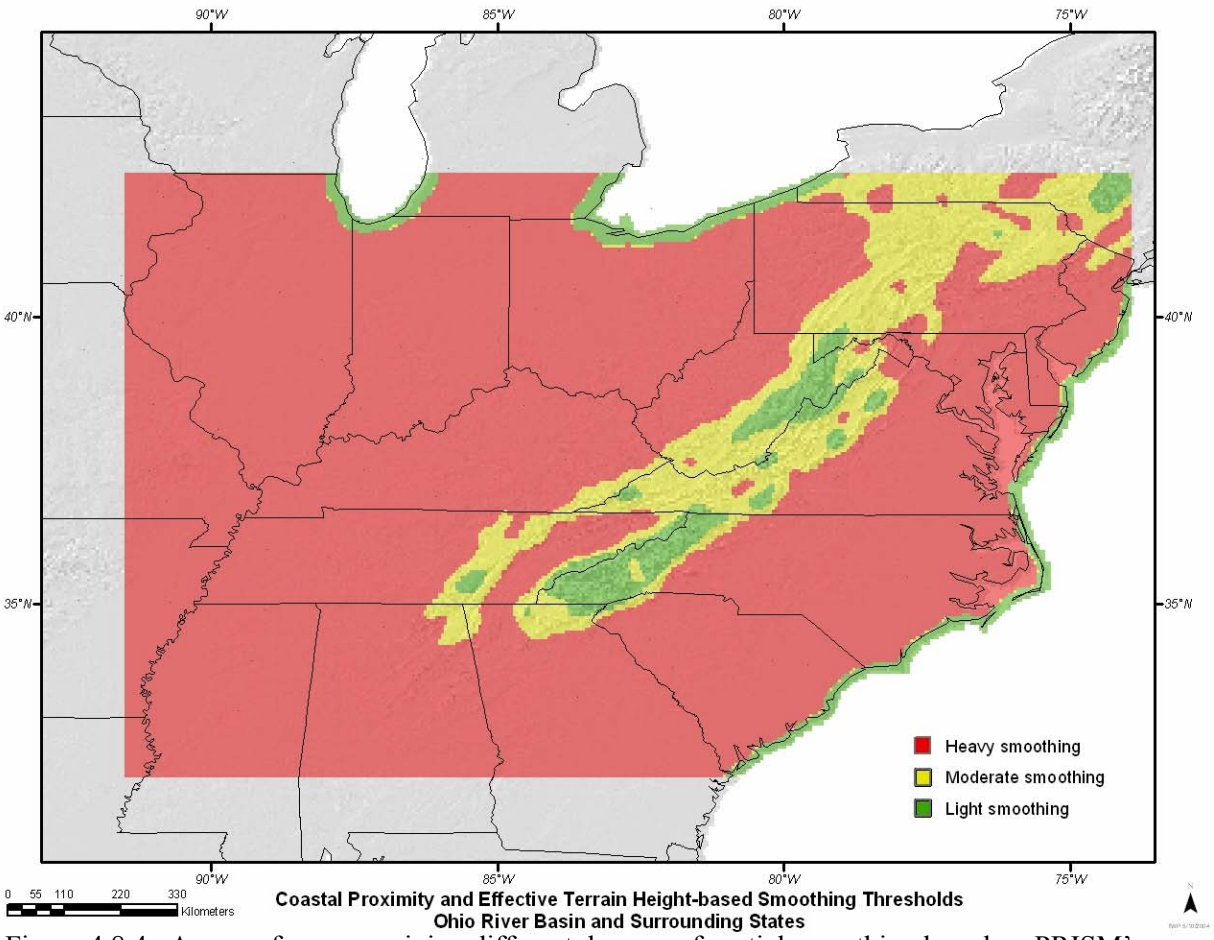


Figure 4.8.4. A map of areas receiving different degrees of spatial smoothing based on PRISM's effective terrain height and coastal proximity grids.

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. 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. Because there were so few n-minute

stations in the project area, larger regional (northern and southern) ratios of n-minute to 60-minute estimates were used. Unlike Volume 1, each duration, frequency and larger region (north and south) had its own unique n-minute ratio. Tables 4.8.2 and 4.8.3 again show the n-minute ratios for the northern and southern regions, respectively, for all durations and annual exceedance probabilities. Grids for each ratio were generated using a grid that delineates the larger northern and southern regions. At the boundary between the larger regions, the grid cells were subjected to spatial averaging for a distance of approximately 90 miles, thereby providing a wide band of gradually changing ratios from north to south. Then, these grids of ratios were multiplied by the appropriate 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 the northern region of NOAA Atlas Volume 2: 5-, 10-, 15- and 30-minute to 60-minute. *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).

Annual Exceedance Probability	**5-min**	**10-min**	**15-min**	**30-min**
1 in 1.58 (1-year ARI)	0.325	0.505	0.619	0.819
1 in 2	0.319	0.498	0.609	0.815
1 in 5	0.305	0.474	0.582	0.797
1 in 10	0.298	0.460	0.566	0.786
1 in 25	0.289	0.442	0.546	0.771
1 in 50	0.283	0.429	0.531	0.759
1 in 100	0.277	0.417	0.518	0.748
1 in 200	0.272	0.406	0.505	0.737
1 in 500	0.266	0.391	0.488	0.723
1 in 1,000	0.261	0.380	0.475	0.712

Table 4.8.3. N-minute ratios for the southern region of NOAA Atlas Volume 2: 5-, 10-, 15- and 30-minute to 60-minute. *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).

Annual Exceedance Probability	5-min	10-min	15-min	30-min
1 in 1.58 (1-year ARI)	0.293	0.468	0.585	0.802
1 in 2	0.287	0.459	0.577	0.797
1 in 5	0.271	0.434	0.549	0.780
1 in 10	0.262	0.419	0.530	0.768
1 in 25	0.251	0.400	0.507	0.751
1 in 50	0.243	0.387	0.490	0.738
1 in 100	0.236	0.375	0.474	0.726
1 in 200	0.229	0.363	0.458	0.713
1 in 500	0.220	0.348	0.438	0.697
1 in 1,000	0.214	0.337	0.423	0.685

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. For the test, 261 stations, which is half of the hourly stations in the core project area, were selected to get a representative sample of terrain and climate evenly distributed around the core area. 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.

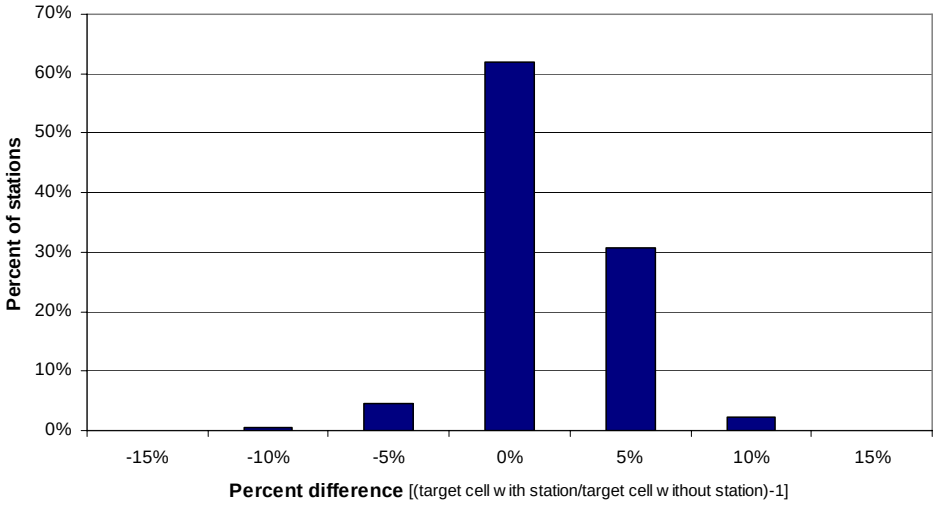


Figure 4.8.5. NOAA Atlas 14 Volume 2 100-year 60-minute jackknife cross-validation results.

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.

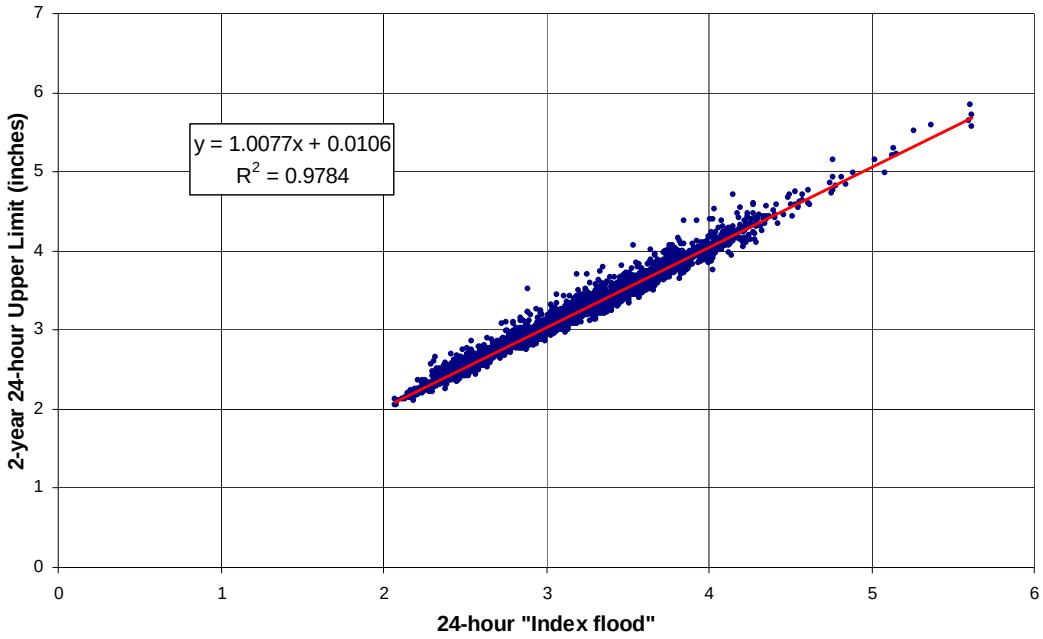


Figure 4.8.6. Scatter plot of the 24-hour mean precipitation frequency estimates vs. the 2-year 24-hour upper limit showing a coefficient of determination of 0.9784 in NOAA Atlas 14 Volume 2.

4.8.3. Pseudo data

Since each duration was computed independently, it was possible for inconsistencies from duration to duration at a given location to occur. In the spatial interpolation, this was a particular concern at hourly-only and daily-only station locations. However, such inconsistencies were rare.

At hourly-only station locations, inconsistencies could occur because calculated 60-minute through 48-hour estimates anchored the interpolation while 4-day through 60-day estimates at those locations were computed during the spatial interpolation process that was based on estimates at nearby daily stations. During the evaluation phase of the grids, HDSC evaluated the results for inconsistencies in the precipitation frequency estimates from 48-hour to 4-day, but none were found. Such inconsistencies occurred in NOAA Atlas 14 Volume 1 due to unreliable 48-hour data derived from accumulated hourly observations. The practical adjustments applied in this project compensated for any such inconsistencies.

Likewise, there were 21 cases where inconsistencies arose at daily-only station locations because calculated 24-hour through 60-day estimates anchored the interpolation while 60-minute through 12-hour estimates at those locations were computed during the spatial interpolation process that was based on estimates at nearby hourly stations. In these 21 cases, the ≤ 12-hour interpolated precipitation frequency estimates were considerably lower and inconsistent with the surrounding calculated ≥ 24-hour precipitation frequency estimates. This caused unreasonable changes in the precipitation frequency estimates from 12-hours to 24-hours at those locations.

These cases were objectively identified using grids that indicated the difference between the 100-year 12-hour and 100-year 24-hour precipitation frequency estimates. By using these grids, spatial artifacts were differentiated from climatologically-driven patterns. In general, if the difference between the 100-year 12-hour and 100-year 24-hour grid cell value was ≥ 1.40”, the daily-only stations in that area were scrutinized. The 21 locations with such inconsistencies were identified and verified for data accuracy.

Table 4.8.4. Hourly pseudo stations used in the preparation of NOAA Atlas 14 Volume 2.

Station ID	Station Name	State
11-0338	AURORA COLLEGE	IL
11-2223	DE KALB	IL
11-4530	JOLIET BRANDON RD DAM	IL
11-4535	JOLIET	IL
11-7354	ROCHELLE	IL
11-9221	WHEATON 3 SE	IL
12-4662	KOKOMO POST OFFICE	IN
12-5174	LOWELL	IN
18-6620	OAKLAND 1 SE	MD
28-0690	BELLEPLAIN	NJ
31-0184	ANDREWS	NC
31-0241	ARCOLA	NC
31-6031	NANTAHALA	NC
31-6044	NASHVILLE	NC
31-6135	NEW HOLLAND	NC
38-0972	BRANCHVILLE 6 S	SC
38-5628	MCCLELLANVILLE	SC
38-7313	RIMINI	SC
44-0385	BACK BAY WILDLIFE REFUGE	VA
44-0993	BREMO BLUFF PWR	VA
44-6456	OYSTER 1 W	VA

So-called pseudo data were used to mitigate the inconsistencies at these 21 locations. Table 4.8.4 lists the hourly pseudo stations generated for this Atlas. 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 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. Locations where an inconsistency between 12-hour and 24-hour estimates could not be expressly proved were assumed accurate based on climate and not mitigated. Pseudo data were used only where deemed absolutely necessary to produce consistent results.

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 and resolution 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 the partial duration series-based precipitation frequency grids. The maps were created using Environmental Systems Research Institute, ArcGIS© 8.3 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 1:2,000,000 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 40, 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 user interface for accessing estimates.