RC 20499 (90858) 7/12/96
Mathematics 16 pages

Research Report

The U.S. National Electronic Drought Atlas: Statistical Data Analysis with GIS-Based Presentation of Results

J. R. M. Hosking and J. R. Wallis

IBM Research Division
T. J. Watson Research Center
Yorktown Heights, NY 10598

Research Division
Almaden • T.J. Watson • Tokyo • Zurich

The U.S. National Electronic Drought Atlas: Statistical Data Analysis with GIS-Based Presentation of Results

J. R. M. Hosking and J. R. Wallis

IBM Research Division
P. O. Box 218
Yorktown Heights, NY 10598

Abstract. In 1989 the U.S. Army Corps of Engineers was charged with the responsibility of conducting a national study of water management during periods of drought. One of the results of the study is the National Drought Atlas, which contains analyses of data on monthly precipitation, streamflow, reservoir levels and the Palmer Drought Index for over 1000 measuring sites in the continental U.S.A. In this report we describe the precipitation analysis, which used the technique known as regional L-moments. We show how the results of the analysis are presented in user-friendly form in a GIS-based version of the National Drought Atlas.

Introduction

In 1989 the U.S. Army Corps of Engineers was charged with the responsibility of conducting a national study of water management during periods of drought. One of the results of the study is the National Drought Atlas (Willeke et al., 1995), which contains analyses of data on monthly precipitation, streamflow, reservoir levels and the Palmer Drought Index for over 1000 measuring sites in the continental U.S.A. In this report we describe only the precipitation analysis, which used the technique known as regional L-moments (Hosking and Wallis, 1996). We show how the results of the analysis are presented in user-friendly form in a GIS-based version of the National Drought Atlas.

Data

The data set used for the analysis was taken from the Historical Climatology Network (Karl et al., 1990). The Historical Climatology Network is a data base containing monthly temperature and precipitation data through 1989 for 1219 sites in the U.S. It was prepared for the U.S. Department of Energy by the National Climatic Data Center (NCDC). The stations in the network are considered to be the best long term records available. The analysis used data from sites that satisfied the following criteria: record length at least 60 years, note more than 10% of the monthly values missing, no more than 12 consecutive months of data missing. There were 1123 such sites, with an average record length of 85 years.

L-moments

L-moments are a system for describing the shapes of probability distributions. Historically they arose as modification of the “probability weighted moments” of Greenwood et al. (1979).

Probability weighted moments of a random variable X with cumulative distribution function $F(\cdot)$ were defined by Greenwood et al. (1979) to be the quantities

$$M_{p,r,s} = \mathrm{E}[X^p \{F(X)\}^r \{1 - F(X)\}^s] \; . \tag{1}$$

Particularly useful special cases are the probability weighted moments $\alpha_r = M_{1,0,r}$ and $\beta_r = M_{1,r,0}$. For a distribution that has a quantile function $x(u)$, we have

$$\alpha_r = \int_0^1 x(u)(1-u)^r \, du, \qquad \beta_r = \int_0^1 x(u) u^r \, du \,. \tag{2}$$

These equations may be contrasted with the definition of the ordinary moments, which may be written as

$$\mathrm{E}(X^r) = \int_0^1 \{x(u)\}^r du \ . \tag{3}$$

Conventional moments involve successively higher powers of the quantile function $x(u)$, while probability weighted moments involve successively higher powers of u or $1-u$, and may be regarded as integrals of $x(u)$ weighted by the polynomials u^r or $(1-u)^r$.

The probability weighted moments α_r and β_r have been used as the basis of methods for estimating parameters of probability distributions by Landwehr et al. (1979a,b), Greis and Wood (1981), Wallis (1981, 1982), Hosking et al. (1985b) and Hosking and Wallis (1987). However, they are difficult to interpret directly as measures of the scale and shape of a probability distribution. This information is carried in certain linear combinations of the probability weighted moments. For example, estimates of scale parameters of distributions are multiples of $\alpha_0 - 2\alpha_1$ or $2\beta_1 - \beta_0$ (Landwehr et al., 1979a; Hosking et al., 1985b). The skewness of a distribution can be measured by $6\beta_2 - 6\beta_1 + \beta_0$ (Stedinger, 1983). These linear combinations arise naturally from integrals of $x(u)$ weighted not by the polynomials u^r or $(1-u)^r$ but by a set of orthogonal polynomials.

We define polynomials $P_r^*(u)$, $r = 0, 1, 2, \ldots$, as follows.

(i) $P_r^*(u)$ is a polynomial of degree r in u.

(ii) $P_r^*(1) = 1$.

(iii) $\int_0^1 P_r^*(u) P_s^*(u)\, du = 0$ if $r \neq s$.

Condition (iii) is the orthogonality condition. These conditions define the *shifted Legendre polynomials* ("shifted", because the ordinary Legendre polynomials $P_r(u)$ are defined to be orthogonal on the interval $-1 \le u \le +1$, not $0 \le u \le 1$). The polynomials have the explicit form

$$P_r^*(u) = \sum_{k=0}^{r} p_{r,k}^* u^k \tag{4}$$

where

$$p^*_{r,k} = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} = \frac{(-1)^{r-k}(r+k)!}{(k!)^2\,(r-k)!}\,. \tag{5}$$

For a random variable X with quantile function $x(u)$, we now define the *L-moments* of X to be the quantities

$$\lambda_r = \int_0^1 x(u)\, P^*_{r-1}(u)\, du\,. \tag{6}$$

In terms of probability weighted moments, L-moments are given by

$$\lambda_1 = \alpha_0 \qquad\qquad = \beta_0\,, \tag{7}$$

$$\lambda_2 = \alpha_0 - 2\alpha_1 \qquad = 2\beta_1 - \beta_0\,, \tag{8}$$

$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 \qquad = 6\beta_2 - 6\beta_1 + \beta_0\,, \tag{9}$$

$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0\,, \tag{10}$$

and in general

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} p^*_{r,k}\alpha_k = \sum_{k=0}^{r} p^*_{r,k}\beta_k\,. \tag{11}$$

It is convenient to define dimensionless versions of L-moments; this is achieved by dividing the higher-order L-moments by the scale measure λ_2. We define the *L-moment ratios*

$$\tau_r = \lambda_r/\lambda_2, \qquad r = 3, 4, \ldots. \tag{12}$$

L-moment ratios measure the shape of a distribution independently of its scale of measurement.

We also define the *L-CV*

$$\tau = \lambda_2/\lambda_1. \tag{13}$$

This quantity is analogous to the ordinary coefficient of variation, C_v. (*L-CV* is not an abbreviation of "L-coefficient of variation": in words it would be more properly described as a "coefficient of L-variation".)

L-moments have been defined for a probability distribution, but in practice must often be estimated from a finite sample. Estimation is based on a sample of size n, arranged in ascending order. Let $x_{1:n} \le x_{2:n} \le \cdots \le x_{n:n}$ be the ordered sample. It is convenient to begin with an estimator of the probability weighted moment β_r. An unbiased estimator of β_r is

$$b_r = n^{-1} \binom{n-1}{r}^{-1} \sum_{j=r+1}^{n} \binom{j-1}{r} x_{j:n} \tag{14}$$

(Landwehr et al., 1979a). This may alternatively be written as

$$b_0 = n^{-1} \sum_{j=1}^{n} x_{j:n} \,, \tag{15}$$

$$b_1 = n^{-1} \sum_{j=2}^{n} \frac{(j-1)}{(n-1)} x_{j:n} \,, \tag{16}$$

$$b_2 = n^{-1} \sum_{j=3}^{n} \frac{(j-1)(j-2)}{(n-1)(n-2)} x_{j:n} \,, \tag{17}$$

and in general

$$b_r = n^{-1} \sum_{j=r+1}^{n} \frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)} x_{j:n} \,. \tag{18}$$

The *sample L-moments* are defined by

$$\ell_1 = b_0 \,, \tag{19}$$

$$\ell_2 = 2b_1 - b_0 \,, \tag{20}$$

$$\ell_3 = 6b_2 - 6b_1 + b_0 \,, \tag{21}$$

$$\ell_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \,, \tag{22}$$

and in general

$$\ell_{r+1} = \sum_{k=0}^{r} p^*_{r,k} b_k \,, \qquad r = 0, 1, \ldots, n-1, \tag{23}$$

the coefficients $p^*_{r,k}$ being as defined in (5). The sample L-moment ℓ_r is an unbiased estimator of λ_r.

From (14) and (23), ℓ_r is a linear combination of the ordered sample values from $x_{1:n}, \ldots, x_{n:n}$, and we can write

$$\ell_r = n^{-1} \sum_{j=1}^{n} w^{(r)}_{j:n} x_{j:n} \,. \tag{24}$$

Parameter estimation using L-moments

A common problem in statistics is the estimation, from a random sample of size n, of a probability distribution whose specification involves a finite number, p, of unknown parameters. Analogously to the usual method of moments, the *method of L-moments* obtains parameter estimates by equating the first p sample L-moments to the corresponding population quantities. This requires expressions for the parameters in terms of the L-moments. Such expressions have been obtained for many standard distributions: examples are given in Hosking (1990), and Hosking and Wallis (1996).

Hosking et al. (1985b) and Hosking and Wallis (1987) found that with small and moderate samples the method of L-moments is often more efficient than maximum-likelihood. These results are for estimators based on a single sample of data, and are not directly relevant to regional frequency analysis. However, they demonstrate that the method of L-moments yields efficient and computationally convenient estimates of parameters and quantiles, and we may reasonably expect these properties to continue to hold in index-flood procedures for regional frequency analysis.

Regional analysis using an index-flood procedure

Index-flood procedures are a convenient way of pooling summary statistics from different data samples. The term "index flood" arose because early applications of the procedure were to flood data in hydrology (e.g., Dalrymple, 1960), but the method can be used with any kind of data.

Suppose that data are available at N sites, with site i having sample size n_i and observed data Q_{ij}, $j=1,\ldots,n_i$. Let $Q_i(F)$, $0<F<1$, be the quantile function of the frequency distribution at site i. The key assumption of an index-flood procedure is that the sites form a *homogeneous region*, i.e. that the frequency distributions of the N sites are identical apart from a site-specific scaling factor, the *index flood*. We may then write

$$Q_i(F) = \mu_i q(F), \qquad i = 1, \ldots, N. \tag{25}$$

Here μ_i is the index flood. We shall take it to be the mean of the at-site frequency distribution, though any location parameter of the distribution may be used instead. Smith (1989), for example, used the 90% quantile $Q_i(0.9)$. The remaining factor in (25), $q(F)$, is the *regional growth curve*, a dimensionless quantile function common to every site. It is the quantile function of the *regional frequency distribution*, the common distribution of the Q_{ij}/μ_i.

The index flood is naturally estimated by $\hat{\mu}_i = \bar{Q}_i$, the sample mean of the data at site i. Other location estimators such as the median or a trimmed mean could be used instead.

The dimensionless rescaled data $q_{ij} = Q_{ij}/\hat{\mu}_i$, $j = 1, \ldots, n_i$, $i = 1, \ldots, N$, are the basis for estimating the regional growth curve $q(F)$, $0 < F < 1$. It is usually assumed that the form of $q(F)$ is known apart from p undetermined parameters $\theta_1, \ldots, \theta_p$, so we write $q(F)$ as $q(F; \theta_1, \ldots, \theta_p)$. In the regional L-moment algorithm, these parameters are the L-moment ratios $\tau, \tau_3, \ldots$, defined in (12)–(13). The mean of the regional frequency distribution is not an unknown parameter: taking μ_i in (25) to be the mean of the frequency distribution at site i ensures that the regional frequency distribution has mean 1. In our approach the parameters are estimated separately at each site, the site-i estimate of θ_k being denoted by $\hat{\theta}_k^{(i)}$. These at-site estimates are combined to give regional estimates:

$$\hat{\theta}_k^{\mathrm{R}} = \sum_{i=1}^{N} n_i \hat{\theta}_k^{(i)} \Big/ \sum_{i=1}^{N} n_i \,. \tag{26}$$

This is a weighted average, with the site-i estimate given weight proportional to n_i because for regular statistical models the variance of $\hat{\theta}_k^{(i)}$ is inversely proportional to n_i. Substituting these estimates into $q(F)$ gives the estimated regional growth curve $\hat{q}(F) = q(F; \hat{\theta}_1^R, \ldots, \hat{\theta}_p^R)$. This method of obtaining regional estimates is essentially that of Wallis (1981), except that the weighting proportional to n_i is a later addition, suggested by Wallis (1982). Somewhat different methods were used by Dalrymple (1960) and the Natural Environment Research Council (1975).

The quantile estimates at site i are obtained by combining the estimates of μ_i and $q(F)$:

$$\hat{Q}_i(F) = \hat{\mu}_i \hat{q}(F). \tag{27}$$

This index-flood procedure makes the following assumptions.

(i) Observations at any given site are identically distributed.

(ii) Observations at any given site are serially independent.

(iii) Observations at different sites are independent.

(iv) Frequency distributions at different sites are identical apart from a scale factor.

(v) The mathematical form of the regional growth curve is correctly specified.

The first two assumptions are plausible for many kinds of data, particularly for annual totals or extremes, which are free from seasonal variations. The last three assumptions are unlikely to be satisfied by environmental data. At best they may be approximately attained, by careful selection of the sites that are to be regarded as forming a region and by careful choice of a frequency distribution that is consistent with the data. Therefore an index-flood procedure can be appropriate only if it is robust to physically plausible departures from these three assumptions. Recent research (Hosking et al., 1985a; Lettenmaier and Potter, 1985; Wallis and Wood, 1985; Lettenmaier et al., 1987; Hosking and Wallis, 1988) has shown that it is possible to construct index-flood procedures that yield suitably robust and accurate quantile estimates.

Initial formation of precipitation regions

Regions were formed by identifying clusters in a space of site characteristics. The analysis is described by Guttman (1993). Clustering was based on site characteristics only and did not involve at-site statistics measuring the shape of the frequency distribution of precipitation. At-site statistics are used to assess the homogeneity of the regions (Hosking and Wallis, 1993) that are formed in the clustering procedure, and the validity of this assessment is compromised if the same data are used both to form regions and to test their homogeneity.

The site characteristics used were judged to be of importance in defining a site's precipitation climate, and included indicators of precipitation amounts and of the distribution of the amounts through the year, and the site's geographic location. Seven variables were chosen to describe a precipitation climate: site latitude; site longitude; site elevation; mean annual precipitation; the ratio of the mean precipitation for the two consecutive months with the lowest mean amount in the year to that for the two months with the highest mean amount; the beginning month of the two consecutive months with the highest mean amount in the year; the beginning month of the two consecutive months with the lowest mean amount in the year.

The observed scales of the variables are very different, and standard methods of cluster analysis are very sensitive to such scale differences. The variables were therefore transformed so that their ranges were comparable. The location, precipitation amount and precipitation ratio variables were rescaled so that their values lay between 0 and 1. The other two variables represent a point along an annual cycle and were transformed by representing the months by a sine curve with a period of one year; the range of the transformed variables if from -1 to $+1$.

Cluster analysis was performed using SAS average linkage and Ward's minimum variance hierarchical clustering software (SAS, 1988). In Ward's method, the distance between two clusters is the sum of squares between the two clusters summed over all the variables. The method tends to join clusters that contain a small number of sites and is strongly biased towards producing clusters containing approximately equal numbers of sites.

Clusterings in which the 1123 sites were divided into between 8 and 60 clusters were initially obtained. The clusters were reviewed to assess whether they were spatially continuous and physically reasonable. A clustering containing about 40 clusters was selected as the basis for further progress. Its clusters were subjectively judged to be reasonable in that the area that they covered could easily be justified on the basis of the physical processes that control precipitation.

The discordancy and heterogeneity measures D_i and H defined in Hosking and Wallis (1993) were computed from the annual precipitation totals for each region identified by the clustering procedure. When the computed heterogeneity measure H exceeded 2, indicating that a region was "definitely heterogeneous", the sites in the region were separated by the clustering algorithms into smaller groups. The discordancy measure occasionally indicated that several neighboring sites in a regions were discordant with the rest of the region. In these cases a new region was formed containing just the discordant sites.

These procedures for subdivision of heterogeneous regions continued until no further progress could be made. At this point there were 109 regions, 82 of which were "acceptably homogeneous" or only "possibly heterogeneous", with $H < 2$. The 27 "definitely heterogeneous" regions were mostly in the western states.

Refinement of regions

The regions obtained from the automatic clustering procedure were adjusted manually. Inspection of the clusters, taking into account the topography and spatial patterns of mean annual precipitation in the areas covered by the clusters, suggested several natural and physically reasonable modifications to the clusters which resulted in more nearly homogeneous clusters; four sites were deleted from the data set during this process.

The final set of regions is illustrated in Figure 1. The remaining 1119 sites were grouped into 111 regions. 108 regions were loosely categorized as homogeneous, with heterogeneity measures $H < 2$; of these 108 regions, 73 were "acceptably homogeneous", with $H < 1$, and 35 were "possibly heterogeneous", with $1 \leq H < 2$. Only 3 regions remained "definitely heterogeneous", with $H \geq 2$. The regions contain between 1 and 48 sites; the median region size is 8 sites and 645 station-years of data; half of the sites are in regions containing at least 13 sites.

Choice of distribution

The goodness of fit statistic defined by Hosking and Wallis (1993) was used to assess each of five candidate probability distributions: generalized logistic, generalized extreme-value, lognormal, Pearson type III and generalized Pareto. The analysis is described by Guttman et al. (1993). Counts were made by duration, region, and starting month of the number of times a distribution was acceptable. The Pearson type III was found to be acceptable most often for precipitation totals over all durations: it was acceptable for about 80% of all combinations of regions, duration and starting month. The lognormal and generalized extreme-value distributions were acceptable almost as often as the Pearson type III for durations longer than 6 months. Based on the counts of acceptable fits, the Pearson type III distribution was chosen for use in homogeneous regions for which it gave an acceptable fit, and the Wakeby distribution was used in the other homogeneous regions and in all heterogeneous regions.

The PC version of the National Drought Atlas

The National Drought Atlas is a bulky document that many users may find cumbersome and relatively difficult to use. A PC version, the National Electronic Drought Atlas (NEDA), that runs under Windows 3.1, Windows 95 and the DOS box of OS/2 Warp has been prepared by Teqnical Services Inc. based upon the MAP-

INFO GIS product (information on the availability of NEDA can be obtained at http://www.teqservices.com).

The illustrations and discussion which follow are based upon the β test version of the NEDA program. In NEDA, by just "pointing and clicking" one can enlarge any portion of the precipitation starburst map, locate a site of interest, and determine the amount of precipitation at any one of 9 probabilities for 1, 2, 3, or 6 months starting in any month, as well as for periods of 12, 24, 36, and 60 months starting in January. Tabular values of precipitation at the identified site for all 8 durations and 9 cumulative probabilities are displayed in a single table. The default starting month is taken from the PC clock but other months can be selected from a drop down window (see Figure 2).

The cumulative distribution function for a region expressed in terms of recurrence intervals (see Figure 3) can be displayed for any site or region by merely clicking on the line graph box of Figure 2.

If precise interpolation or limited extrapolation to other probabilities is required, this is accomplished by means of the *cloud icon* from the NEDA power bar (see Figure 4).

References

Dalrymple, T. (1960). Flood frequency analyses. *Water Supply Paper 1543-A*, U.S. Geological Survey, Reston, Va.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C. & Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, **15**, 1049–54.

Greis, N. P. & Wood, E. F. (1981). Regional flood frequency estimation and network design. *Water Resources Research*, **17**, 1167–77. Correction: *Water Resources Research*, **19** (1983), 589–90.

Guttman, N. B. (1993). The use of L-moments in the determination of regional precipitation climates. *Journal of Climate*, **6**, 2309–25.

Guttman, N. B., Hosking, J. R. M. & Wallis, J. R. (1993). Regional precipitation quantile values for the continental U.S. computed from *L*-moments. *Journal of Climate*, **6**, 2326–40.

Hosking, J. R. M. (1990). *L*-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B*, **52**, 105–24.

Hosking, J. R. M. & Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, **29**, 339–49.

Hosking, J. R. M. & Wallis, J. R. (1988). The effect of inter-site dependence on regional flood frequency analysis. *Water Resources Research*, **24**, 588–600.

Hosking, J. R. M. & Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. *Water Resources Research*, **29**, 271–81. Correction: *Water Resources Research*, **31** (1995), 251.

Hosking, J. R. M. & Wallis, J. R. (1996). *Regional Frequency Analysis: An Approach Based Upon L-moments*. To be published by Cambridge University Press.

Hosking, J. R. M., Wallis, J. R. & Wood, E. F. (1985a). An appraisal of the regional flood frequency procedure in the UK Flood Studies Report. *Hydrological Sciences Journal*, **30**, 85–109.

Hosking, J. R. M., Wallis, J. R. & Wood, E. F. (1985b). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, **27**, 251–61.

Karl, T. R., Williams, C. N., Quinlan, F. T. & Boden, T. A. (1990). United States historical climatology network (HCN) serial temperature and precipitation data.

ORNL/CDIAC-30, NDP-019/R1, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn.

Landwehr, J. M., Matalas, N. C. & Wallis, J. R. (1978). Some comparisons of flood statistics in real and log space. *Water Resources Research*, **14**, 902–20. Correction: *Water Resources Research*, **15** (1979), 983–4.

Landwehr, J. M., Matalas, N. C. & Wallis, J. R. (1979a). Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, **15**, 1055–64.

Landwehr, J. M., Matalas, N. C. & Wallis, J. R. (1979b). Estimation of parameters and quantiles of Wakeby distributions. *Water Resources Research*, **15**, 1361–79. Correction: *Water Resources Research*, **15** (1979), 1672.

Lettenmaier, D. P. & Potter, K. W. (1985). Testing flood frequency estimation methods using a regional flood generation model. *Water Resources Research*, **21**, 1903–14.

Lettenmaier, D. P., Wallis, J. R. & Wood, E. F. (1987). Effect of regional heterogeneity on flood frequency estimation. *Water Resources Research*, **23**, 313–23.

Natural Environment Research Council (1975). *Flood Studies Report, vol. 1.* London: Natural Environment Research Council.

SAS (1988). *SAS/STAT User's Guide*, Release 6.03. Cary, N.C.: SAS Institute.

Stedinger, J. R. (1983). Estimating a regional flood frequency distribution. *Water Resources Research*, **19**, 503–10.

Wallis, J. R. (1981). Risk and uncertainties in the evaluation of flood events for the design of hydraulic structures. In *Piene e Siccità*, edited by E. Guggino, G. Rossi and E. Todini, pp. 3–36. Catania, Italy: Fondazione Politecnica del Mediterraneo.

Wallis, J. R. (1982). Hydrologic problems associated with oilshale development. In *Environmental Systems and Management*, edited by S. Rinaldi, pp. 85–102. Amsterdam: North-Holland.

Wallis, J. R. & Wood, E. F. (1985). Relative accuracy of log Pearson III procedures. *Journal of Hydraulic Engineering*, **111**, 1043–56.

Willeke, G. E., Hosking, J. R. M., Wallis, J. R. & Guttman, N. B. (1995). The National Drought Atlas (draft). *IWR Report 94-NDS-4*, U.S. Army Corps of Engineers, Fort Belvoir, Va.

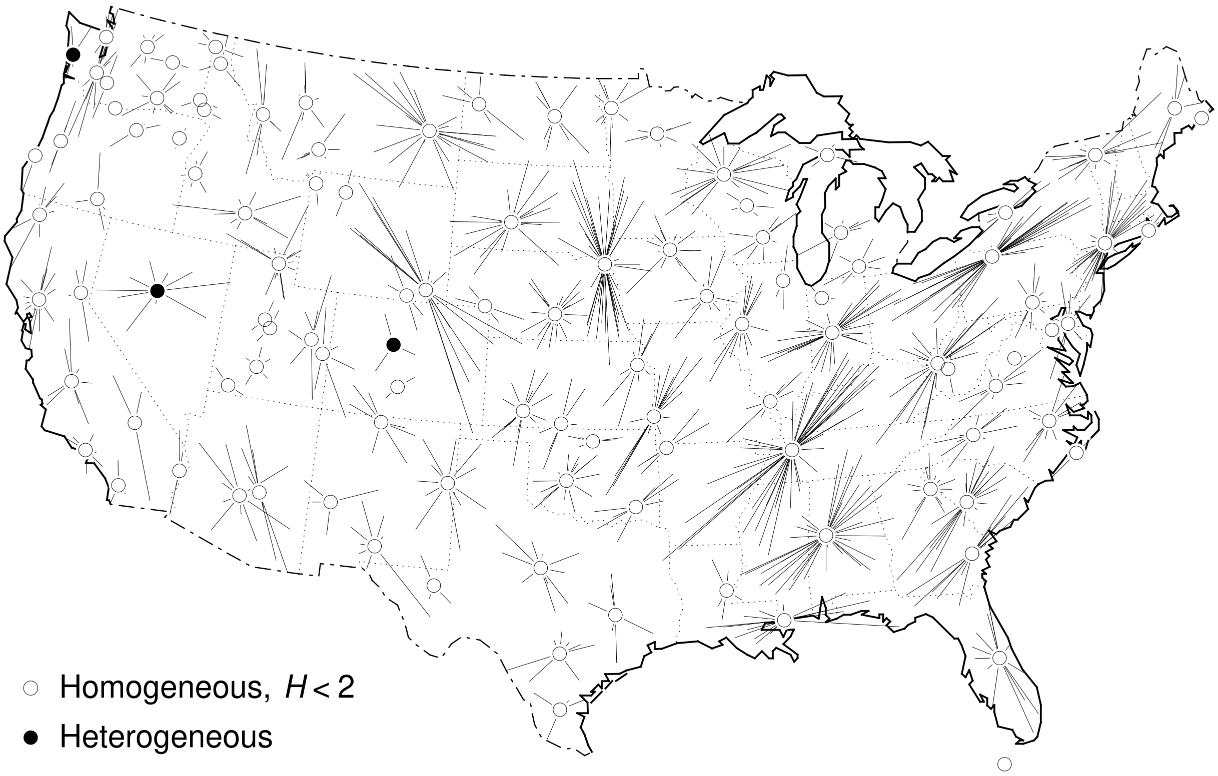
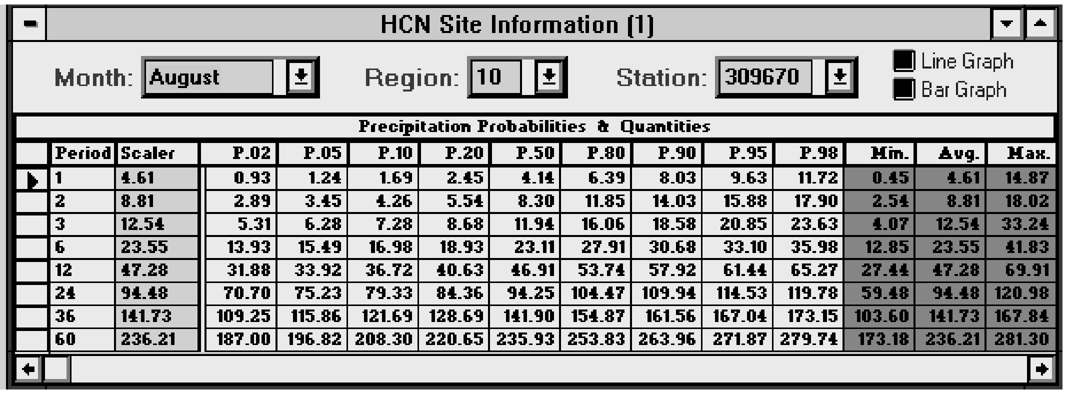


Figure 1. Precipitation regions, final version. The symbol at the center of a region indicates whether the region is homogeneous. Lines radiating from the center end at the locations of stations in the region.



HCN Site Information (1)

Month: August Region: 10 Station: 309670

Line Graph

Bar Graph

Precipitation Probabilities & Quantities

Period	Scaler	P.02	P.05	P.10	P.20	P.50	P.80	P.90	P.95	P.98	Min.	Avg.	Max.
1	4.61	0.93	1.24	1.69	2.45	4.14	6.39	8.03	9.63	11.72	0.45	4.61	14.87
2	8.81	2.89	3.45	4.26	5.54	8.30	11.85	14.03	15.88	17.90	2.54	8.81	18.02
3	12.54	5.31	6.28	7.28	8.68	11.94	16.06	18.58	20.85	23.63	4.07	12.54	33.24
6	23.55	13.93	15.49	16.98	18.93	23.11	27.91	30.68	33.10	35.98	12.85	23.55	41.83
12	47.28	31.88	33.92	36.72	40.63	46.91	53.74	57.92	61.44	65.27	27.44	47.28	69.91
24	94.48	70.70	75.23	79.33	84.36	94.25	104.47	109.94	114.53	119.78	59.48	94.48	120.98
36	141.73	109.25	115.86	121.69	128.69	141.90	154.87	161.56	167.04	173.15	103.60	141.73	167.84
60	236.21	187.00	196.82	208.30	220.65	235.93	253.83	263.96	271.87	279.74	173.18	236.21	281.30

Figure 2. The values shown are for August at HCN site 309670#, (Yorktown Heights, N.Y.). The little down arrows allow one to select other starting months and sites. The black pointer on the left identifies the active period (duration) for the precipitation estimate, 1 month in this example, with longer durations on succeeding lines. The scaler column has as its default value the long term average precipitation for the corresponding periods. If this field is altered then the estimated precipitation values for each of the 9 cumulative probability columns to its right automatically reflect the change; an option used when estimating precipitation probabilities for locations other than those specifically identified on the Starburst map.

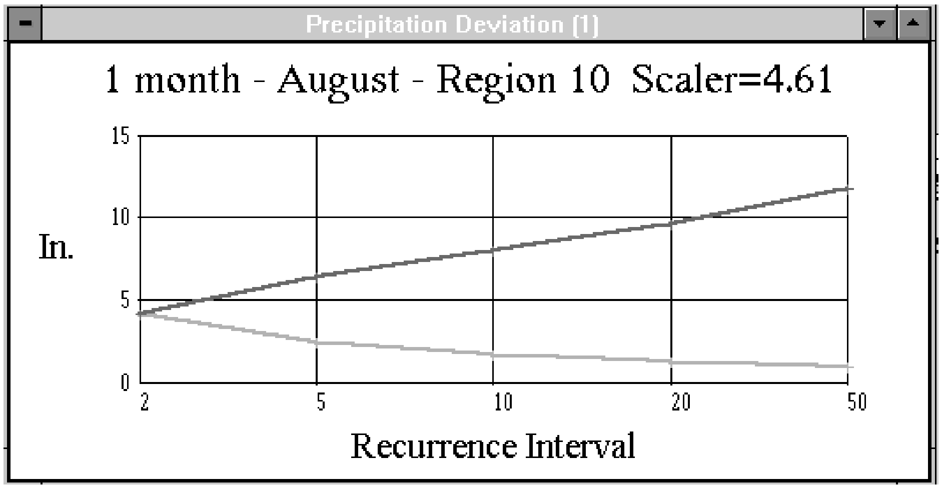


Figure 3. This shows the cumulative probability distribution expressed in terms of recurrence intervals. The lower line shows the chances of observing precipitation amount below the indicated amounts, and the upper line gives the chances of observing precipitation amounts in excess of the indicated amounts.

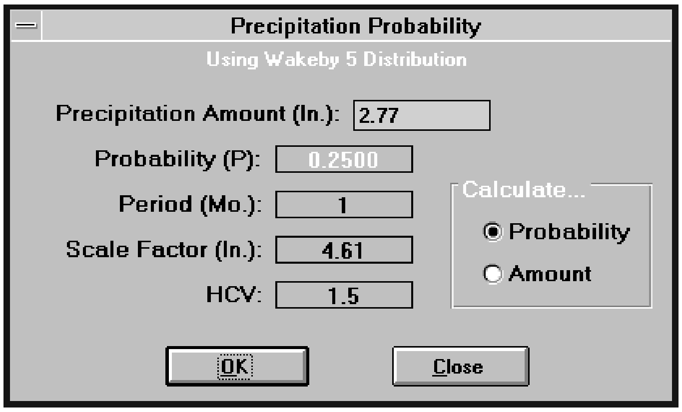


Figure 4. If precise interpolation or limited extrapolation to other probabilities are required this is accomplished by means of the cloud icon (from the NEDA power bar). For instance, Yorktown Heights in August 1995 experiences hot dry weather with only 2.77 inches of rainfall. Entering 2.77 in the amount field of the "cloud table" results in a probability estimate of 0.25, i.e. 1 out of 4 Augusts at Yorktown Heights can be expected to have less rainfall than was measured in 1995—a dry month but not an extremely dry month.