NIST BUILDING SCIENCE SERIES 174

Extreme Wind Distribution Tails: A 'Peaks Over Threshold' Approach

E. Simiu[1] and N. A. Heckert[2]

[1]Building and Fire Research Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-0001

and

Department of Civil Engineering
The Johns Hopkins University
Baltimore, MD 21218

[2]Computing and Applied Mathematics Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-0001

Sponsored in part by:

National Science Foundation
Arlington, VA 22230

March 1995

U.S. Department of Commerce
Ronald H. Brown, *Secretary*

Technology Administration
Mary L. Good, *Under Secretary for Technology*

National Institute of Standards and Technology
Arati Prabhakar, *Director*

National Institute of Standards and Technology Building Science Series 174
Natl. Inst. Stand. Technol. Bldg. Sci. Ser. 174, 72 pages (Mar. 1995)
CODEN: NBSSES

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1995

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402–9325

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iv
ABSTRACT	vi
1. INTRODUCTION	1
2. 'PEAKS OVER THRESHOLD' APPROACH	3
2.1 Generalized Pareto Distribution	3
2.2 Gumbel and Reverse Weibull Distributions	3
2.3 Mean Recurrence Intervals of Variate X as Functions of GPD Parameters and Exceedance Rate	4
3. DESCRIPTION OF DE HAAN ESTIMATION METHOD	5
4. WIND SPEED DATA	7
4.1 Uncorrelated Samples Obtained from Largest Daily Data Records	7
4.2 Largest Yearly Data Samples	8
5. ANALYSES AND RESULTS	9
5.1 Analyses of Uncorrelated Data Sets by the Probability Plot Correlation Coefficient Method	9
5.2 Estimation of Tail Length Parameter by 'Peaks over Threshold' Analyses of Uncorrelated Data Samples	11
5.3 Estimation of Tail Length Parameter by 'Peaks over Threshold' Analyses of Largest Yearly Data Samples	16
5.4 Estimation of Speeds with Specified Mean Recurrence Intervals by 'Peaks over Threshold' Analyses of Uncorrelated Data Samples	19
5.5 Influence of Data Errors on Analysis Results	19
6. LOAD FACTORS FOR WIND-SENSITIVE STRUCTURES	23
7. CONCLUSIONS	25
REFERENCES	27
APPENDIX 1 — Histograms of daily largest wind speeds and histograms of four-day interval uncorrelated wind speeds	29
APPENDIX 2 — Plots of parameter $\hat{c}$ and 95% confidence bounds versus threshold and number of threshold exceedances (based on samples of largest yearly wind speeds)	53
APPENDIX 3 — Plots of point estimates for 100-, 1000- and 100 000-year speeds versus threshold and number of threshold exceedances (based on four-day interval samples)	65

Page

APPENDIX 4 -- Instructions for accessing data sets and attendant programs . 71

LIST OF FIGURES[1]

Figure 1. Probability plots for four-day interval sample, Albany, New York.......... 10

Figure 2. Plots of parameter $\hat{c}$ and 95% confidence bounds versus threshold and number of threshold exceedances (based on four-day interval samples).......... 12

Figure 3. Point estimates of tail length parameters and 95% confidence bounds, versus number of exceedances (de Haan, 1990).......... 16

Figure 4. Mean of tail length parameter estimates weighted over 44 four-day interval samples versus order of threshold.......... 17

Figure 5. Mean of tail length parameter estimates weighted over 115 largest yearly data samples versus order of threshold.......... 18

Figure 6. Quantiles: point estimates and 95% confidence bounds versus number of exceedances (de Haan, 1990).......... 19

Figure 7. Plots of parameter $\hat{c}$ and 95% confidence bounds versus threshold and number of threshold exceedances (based on four-day interval samples, uncorrected data).......... 21

Figure 8. Point estimates for 100-, 1000- and 100 000-year speeds versus threshold and number of threshold exceedances (based on four-day interval samples, uncorrected data).......... 22

[1]Figures in Appendices are not included in this list.

ACKNOWLEDGMENTS

E. Simiu acknowledges with thanks partial support by the National Science Foundation (Grant No. CMS-9411642 to Department of Civil Engineering, The Johns Hopkins University), and the helpful interaction on this project with R.B. Corotis of the University of Colorado at Boulder. The writers wish to thank S.D. Leigh of the Center for Computing and Applied Mathematics, National Institute of Standards and Technology, who suggested the collection of large samples of daily data and the plotting format adopted in this report. Thanks are also due to Dr. S. Coles of Lancaster University for helpful criticism; Dr. J.J. Filliben of the National Institute of Standards and Technology for advice on testing the hypothesis that a universal reverse Weibull tail length parameter exists; and Drs. J.L. Gross and J.L. Lechner of the National Institute of Standards and Technology for valuable collaboration on earlier phases of this project, some of which are reviewed in this report.

COVER PICTURE:

Honoré Daumier (1808-1879)
"UN LÉGER COUP DE VENT" (A light gust)
Lithograph (newsprint)
In the Collection of the Corcoran Gallery of Art,
Gift of Dr. Armand Hammer

ABSTRACT

We seek to ascertain whether the reverse Weibull distribution is an appropriate extreme wind speed model by performing statistical analyses based on the 'peaks over threshold' approach. We use the de Haan method, which was found in previous studies to perform about as well or better than the Pickands and Cumulative Mean Exceedance methods, and has the advantage of providing estimates of confidence bounds. The data are taken principally from records of the largest daily wind speeds obtained over periods of 15 to 26 years at 44 U.S. weather stations in areas not subjected to mature hurricane winds. From these records we create samples with reduced mutual correlation among the data. In our opinion, the analyses provide persuasive evidence that extreme wind speeds are described predominantly by reverse Weibull distributions, which unlike the Gumbel distribution have finite upper tail and lead to reasonable estimates of wind load factors. Instructions are provided for accessing the data and attendant programs.

Key words: Building technology; building (codes); climatology; extreme value theory; load factors; structural engineering; structural reliability; threshold methods; wind (meteorology).

1. INTRODUCTION

A fundamental theorem in extreme value theory states that sufficiently large values of independent and identically distributed variates are described by one of three extreme value distributions: the Fréchet distribution (with infinite upper tail), the Gumbel distribution (whose upper tail is also infinite, but shorter than the Fréchet distribution's), and the reverse (negative) Weibull distribution, whose upper tail is finite (Castillo, 1988).

The wind loading provisions of the American National Standard ANSI A58.1-1972 were based on the assumption that a Fréchet distribution best fits extreme wind speeds blowing from any direction in regions not subjected to mature hurricane winds. An extensive study concluded, however, that the Gumbel distribution is a more appropriate model (Simiu, Filliben and Biétry, 1976). It is a physical fact that extreme winds are bounded. Their probabilistic model should reflect this fact. To the extent that an extreme value distribution would be a reasonable model of extreme wind behavior, one would intuitively expect the best fitting distribution to have finite tail, that is, to be a reverse Weibull distribution.

In addition to the certainty that wind speeds are bounded, there is at least one other indication — albeit indirect — that the Gumbel model might be an inappropriate model of extreme wind behavior. Estimated safety indices for wind-sensitive structures based on the Gumbel model imply unrealistically high failure probabilities (Ellingwood et al., 1980). This may be due, at least in part, to the use in those estimates of a distribution with an unrealistically long — infinite — upper tail.

In this report we seek to ascertain whether the reverse Weibull distribution is an appropriate extreme wind speed model by performing statistical analyses based on the 'peaks over threshold' approach. This approach enables the analyst to use all the data exceeding a sufficiently high threshold, and is more effective than the classical approach, which uses only the largest value in each of a number of basic comparable sets called epochs (typically, for extreme wind analysis an epoch consists of one year). To illustrate this point, consider, for example, two successive years in which the respective largest wind speeds are 30 m/s and 45 m/s. Assume that in the second year winds with speeds of 31 m/s, 37 m/s, 41 m/s and 44 m/s were also recorded (at dates separated by sufficiently long intervals to view the data as independent). For the purposes of threshold theory, for a 30 m/s threshold the two years would supply six data points. The classical theory would make use of only two data points. It may be argued that, by choosing a somewhat lower threshold, the number of data points used in the analysis could be considerably larger than six in our example. However, in view of the theory's basic assumption that the threshold is high, excessive lowering of the threshold would introduce a strong bias. Simulations reported by Gross et al. (1994) suggest that, in samples taken from normal or extreme value populations, optimal results are obtained if the threshold is chosen so that the number of exceedances is of the order of ten per year.

Given a sample of data exceeding a sufficiently high threshold, the analyst using the 'peaks over threshold' approach must choose an appropriate estimation method. In this report we use the estimation method proposed by de Haan (1994). Our choice is based on two reasons. First, Monte Carlo simulations suggest that the de Haan method performs about as well or better than two available alternative methods, the Pickands method and the Cumulative Mean Exceedance method (Gross et al, 1994). Second, the de Haan method has the advantage of providing estimates of confidence bounds.

Data used in this report are taken principally from records of the largest daily wind speeds obtained over periods of 15 to 26 years at 44 U.S. weather stations in areas not subjected to mature hurricane winds. A storm system usually affects a given location for longer than one day, so that wind speed data recorded on two or even more consecutive days are not necessarily independent. We describe the data samples and our procedure for creating, from the samples of largest daily speeds, samples with reduced mutual correlation among the data. In addition to samples of daily data we describe and analyze 115 samples consisting only of largest yearly speeds recorded over periods of 18 to 54 years at locations not subjected to mature hurricane winds. To our knowledge no tornado winds have affected any of our data. All the data samples used in our analyses are available in an anonymous data file. Instructions for accessing that file and attendant programs are available in Appendix 4.

In our opinion, the results presented in this report provide persuasive evidence that extreme wind speeds of extratropical origin and excluding tornadoes are described predominantly by reverse Weibull distributions. This result is in itself useful from a structural engineering viewpoint, and we discuss its potential implications for the estimation of load factors for wind-sensitive structures. Our analyses also suggest that estimates of extreme wind speeds based on the reverse Weibull model may not be obtainable with sufficient confidence from the information and by the method used in this report. This suggests the need for (i) data samples based on longer recording periods and (ii) more efficient estimation methods.

The report is organized as follows. Basic theoretical results pertaining to the 'peaks over threshold' approach are briefly reviewed in Section 2. The de Haan method is reviewed briefly in Section 3. The data used in the analyses are described in Section 4. Section 5 includes the main results of our investigation. Section 6 discusses the load factors issue. Section 7 presents our conclusions.

2. 'PEAKS OVER THRESHOLD' APPROACH

2.1 Generalized Pareto Distribution. The Generalized Pareto Distribution (GPD) is an asymptotic distribution whose use in extreme value theory rests on the fact that exceedances of a sufficiently high threshold are rare events to which the Poisson distribution applies. The expression for the GPD is

$$G(y) = \mathrm{Prob}[Y \le y] = 1-\{[1+(cy/a)]^{-1/c}\} \qquad a>0,\ (1+(cy/a))>0 \tag{1}$$

Equation (1) can be used to represent the conditional cumulative distribution of the excess $Y = X - u$ of the variate X over the threshold u, given $X > u$ for u sufficiently large (Pickands, 1975). The cases $c>0$, $c=0$ and $c<0$ correspond respectively to Fréchet (Type II Extreme Value), Gumbel (Type I Extreme Value), and reverse Weibull (Type III Extreme Largest Values) domains of attraction. For $c=0$ the expression between braces is understood in a limiting sense as the exponential $\exp(-y/a)$ (Castillo, 1988, p. 215).

Given the mean E(Y) and standard deviation s(Y) of the variate Y,

$$a = \tfrac{1}{2}\, E(Y)\{1 + [E(Y)/s(Y)]^2\} \tag{2}$$

$$c = \tfrac{1}{2}\{1-[E(Y)/s(Y)]^2\} \tag{3}$$

(Hosking and Wallis, 1987).

2.2 Gumbel and Reverse Weibull Distributions. We recall that the expressions for the Gumbel and reverse Weibull distributions for maxima are, respectively,

$$F_G(x) = \exp\{-\exp[-(x-\mu_G)/\sigma_G]\} \qquad -\infty<x<\infty \tag{4}$$

$$F_W(x) = \exp\{-[(\mu_W - x)/\sigma_W]^\gamma\}, \qquad x \le \mu_W \tag{5}$$

For the Gumbel distribution, the relations between distribution parameters and the mean E(X) and standard deviation s(X) are

$$\sigma_G = (6^{1/2}/\pi)s(X) \tag{6}$$

$$\mu_G = E(X) - 0.57722(6^{1/2}/\pi)s(X) \tag{7}$$

For the Weibull distribution,

$$\sigma_W = s(X)/\{\Gamma(1+2/\gamma)-[\Gamma(1+1/\gamma)]^2\}^{1/2} \tag{8}$$

$$\mu_W = E(X) + \sigma_w\, \Gamma(1 + 1/\gamma) \tag{9}$$

where Γ is the gamma function (Johnson and Kotz, 1972). For example, for $E(X) = 50$, $s(X) = 6.25$, and $\gamma = 2$, $\sigma_W = 13.49$ and $\mu_W = 61.96$. The tail length parameter γ is related to the parameter c in the GPD distributions as follows:

$$\gamma = -1/c \tag{10}$$

(Smith, 1989).

2.3 Mean Recurrence Intervals of Variate X as Functions of GPD Parameters and Exceedance Rate. The mean recurrence interval R of a given wind speed, in years, is defined as the inverse of the probability that that wind speed will be exceeded in any one year (see, e.g., Simiu and Scanlan, 1986, p. 525). In this section we give expressions that allow the estimation from the GPD of the value of the variate corresponding to any percentage point $1 - 1/(\lambda R)$, where λ is the mean crossing rate of the threshold u per year (i.e., the average number of data points above the threshold u per year). Set

$$\mathrm{Prob}(Y < y) = 1 - 1/(\lambda R) \qquad (11)$$

Using eq (1)

$$1 - [1 + cy/a]^{-1/c} = 1 - 1/(\lambda R) \qquad (12)$$

Therefore

$$y = -a[1 - (\lambda R)^c]/c \qquad (13)$$

(Davison and Smith, 1990). The value being sought is

$$x_R = y + u \qquad (14)$$

where u is the threshold used in the estimation of c and a.

3. DESCRIPTION OF DE HAAN ESTIMATION METHOD

Let the number of data above the threshold be denoted by k, so that the threshold u represents the (k+1)-th highest data point(s). We have $\lambda = k/n_{yrs}$, where n_{yrs} denotes the length of the record in years. The highest, second,..., k-th, (k+1)-th highest variates are denoted by $X_{n,n}$, $X_{n-1,n}$, $X_{n-(k+1),n}$, $X_{n-k,n}$=u, respectively. Compute the quantitites

$$M_n^{(r)} = \frac{1}{k} \sum_{i=0}^{k-1} \{\log(X_{n-1,n}) - \log(X_{n-k,n})\}^r \quad r=1,2 \tag{15}$$

The estimators of c and a are then

$$\hat{c} = M_n^{(1)} + 1 - \frac{1}{2\{1 - (M_n^{(1)})^2/(M_n^{(2)})\}} \tag{16}$$

$$\hat{a} = uM_n^{(1)}/\rho_1 \tag{17}$$

$$\rho_1 = \begin{cases} 1 & \hat{c}\geq 0 \\ 1(1-\hat{c}) & \hat{c}\leq 0 \end{cases} \tag{18}$$

The standard deviation of the asymptotically normal estimator of c is

$$\text{s.d.}(\hat{c})=[(1+\hat{c}^2)/k]^{1/2} \qquad \hat{c}\geq 0 \tag{19a}$$

$$\text{s.d.}(\hat{c}) = \{[(1-\hat{c})^2(1-2\hat{c})[4 - \frac{8(1-2\hat{c})}{(1-3\hat{c})} + \frac{(5-11\hat{c})(1-2\hat{c})}{(1-3\hat{c})(1-4\hat{c})}]/k\}^{1/2} \qquad \hat{c}<0 \tag{19b}$$

(de Haan, 1994).

4. WIND SPEED DATA

4.1 Uncorrelated Samples Obtained from Largest Daily Data Records. Sets of daily fastest mile wind speeds for winds blowing from any direction were obtained from the National Climatic Data Center, National Oceanic and Atmospheric Administration. In most samples a number of daily fastest miles were missing. The speeds on days with missing fastest mile data were estimated from speeds recorded on the respective days at 3-hour intervals, using observations of the approximate relation between these speeds and daily fastest mile speeds. Wind speeds so estimated exceeded 15.6 m/s (35 mph) only at the following stations and dates: Boise (4/26/87, 16.1 m/s); Portland, OR (11/14/81, 19.7 m/s), Salt Lake City (2/1/87, 16.1 m/s) and Toledo (2/6/86, 18.8 m/s). Forty-four samples were used in the analyses. For fourteen of these samples corrections based on the largest yearly records were effected.[1] The influence of these corrections on the results of the analyses is discussed in Section 5.5. The length of the records ranged from 15 to 26 years, the average length being about 18.5 years.

The anemometer elevations were changed during the period of record at the following stations: Duluth (16.2 m to 10/15/75; 6.4 m thereafter), Dayton (6.1 m to 2/4/64; 6.7 m/s thereafter), Missoula (6.1 m to 6/24/82; 9.8 m thereafter), Oklahoma City (16.8 m to 10/21/65; 6.1 m thereafter), Portland, OR (7.6 m to 3/1/73; 6.1 m thereafter), San Diego (6.4 m to 8/13/69, 6.1 m thereafter), Toledo, OH (6.1 m to 11/1/68; 10 m thereafter) and Winnemucca (10.4 m to 4/22/66; 6.1 m thereafter). For these stations the daily data were corrected to correspond to a common 10 m elevation using the logarithmic law for open terrain. For all other stations the anemometer elevations did not change during the period of record and (except for Denver, where the data were also corrected to correspond to a 10 m elevation), the original recorded data were used, that is, no elevation correction was effected.

From samples of largest daily wind speeds we obtained as follows samples that have reduced mutual dependence among the data. Partition the sample of daily maxima into small periods of size equal to or larger than the duration of typical storms in days. (A reasonable choice of the length of the period is four to eight days.) Pick the largest value in each period. If the maxima of two adjacent

---

[1] The following discrepancies between records of yearly maximum fastest miles and daily maximum fastest miles were found: Cheyenne (6.1 m elevation), 69 vs. 75 (72/3/6), 61 vs. 66 (1/18/75), 65 vs. 70 (6/14/76); Dayton (10 m), 44 vs. 49 (6/27/66), 61 vs. 86 (2/16/67), 52 vs. 67 (5/14/70); Fort Wayne (6.1 m), 51 vs. 55 (8/27/65); Greenville (6.1 m), 57 vs. 52 (7/15/66); Lander (6.1 m), 49 vs. 61 (4/12/68); Louisville (6.1 m), 57 vs. 61 (2/15/67); Milwaukee (6.1 m), 52 vs. 56 (6/16/73); Minneapolis (6.4 m), 52 2vs. 56 (7/10/66); Missoula (6.1 m), 67 vs. 61 (7/31/83); Portland, ME (6.1 m), 38 vs. 57 (12/24/70); Portland, OR (10 m), 48 vs. 42 (12/11/69); Pueblo (6.1 m), 65 vs. 70 (5/12/75); Richmond (6.1 m), 47 vs. 42 (7/11/73); Sheridan (6.1 m), 52 vs. 57 (3/2/74). For these stations the corrections were effected by replacing the recorded daily maximum by the yearly maximum recorded on the same day. The yearly maxima were checked for correctness against the original charts — see Section 4.2.

periods are less than half a period apart, replace the smaller of the two maxima by the next smaller value in the respective period which is at least half a period apart from the larger maximum. A data sample is thus obtained in which adjacent data are one period apart on the average and never less than half a period apart. We show below the daily maximum fastest miles at Boise, Idaho in the first six eight-day periods of the year 1965. The periods are separated by vertical bars. The data selected by the procedure just described are in bold type. In the sixth period we underlined the period maximum (26), discarded and replaced by the next largest value (18) because of the proximity to the larger maximum (31) of the adjacent period.

23,32,**35**,20,26,24,24,14 | 13,**16**, 5,11, 5,12,12, 7 | 6, 6, 9, 9,11,12,25,**26** |

15,12,12, 7,15,12,**29**,10 | 7,10,15,20,20,17,24,**31** | <u>26</u>,9,16,14,**18**,16,14,12|

In spite of our selection procedure, small correlations among data might subsist. Nevertheless, we refer to a sample obtained by the selection procedure just described as an **uncorrelated data sample** based on eight-day (four-day) intervals or, for short, an eight-day (four-day) interval sample. An assessment was made of differences between results of analyses based on four-day and eight-day interval samples at the same station. Since in all cases the differences were insignificant, we present in this report only results based on four-day interval samples.

Appendix 1 contains histograms of the full samples of daily data and of the four-day interval samples obtained from them for each of the 44 stations. Owing to the small scale of the graphs, in some cases high wind data are not perceptible on the daily data histograms; however, they can be seen clearly on the four-day interval data histograms. A comparison between the histograms of the full daily data samples and the histograms of the four-day interval samples shows that our selection procedure considerably reduces the number of lowest wind speed data. The selection procedure also results in a shifting of the highest ordinate of the histogram toward higher wind speeds.

4.2 Largest Yearly Data Samples. Also available were samples of largest yearly fastest miles for winds blowing from any direction, recorded over periods of 18 to 54 years at 115 U.S. stations not subjected to mature hurricane winds. The data in those samples were obtained and checked against original charts by M.J. Changery, Chief, Applied Climatology Branch, National Climatic Center, National Oceanic and Atmospheric Administration (letter to E. Simiu of December 20, 1988) and are an update of the information included in Simiu, Changery and Filliben (1979).

As noted earlier, all the data samples used in our analyses are available in an anonymous data file. Instructions for accessing that file and attendant programs for creating sets of uncorrelated data are available in Appendix 4.

5. ANALYSES AND RESULTS

5.1 Analysis of Uncorrelated Data Sets by the Probability Plot Correlation Coefficient Method. Before applying the 'peaks over threshold' approach, we estimated the best-fitting distributions for the four-day samples from among a set of seven distributions or families of distributions (normal, double exponential, lognormal, Gumbel, Fréchet, Weibull, and reverse Weibull). This analysis was viewed as a tentative step toward understanding the probabilistic structure of the populations from which the threshold exceedances were taken. The estimation of the best fitting distribution was based on the probability plot correlation coefficient (PPCC) (Filliben, 1975). As an example, Figure 1 shows the PPCC plots for the Albany, New York four-day interval samples. For this sample the mean and standard deviation were $E(X)=10.5$ m/s (23.5 mph) and $s.d.(X)=3.14$ m/s (7.03 mph); for the full sample of daily data $E(X)=7.8$ m/s (17.5 mph) and $s.d.(X)=3.1$ m/s (6.98 mph). The analyses were repeated for the eight-day samples, and the results were found to differ insignificantly from those based on the four-day samples.

The reverse Weibull distribution was found to best fit the data in the majority of the cases. Even in the cases where other distributions fitted the data better, the reverse Weibull was typically very close to being the best fitting distribution, that is, its PPCC differed only in the fourth or even fifth significant figure from the PPCC of the best fitting distribution. We therefore re-analyzed the eight-day interval samples by assuming that the populations for all stations have reverse Weibull distributions with a single, site-independent value of the tail length parameter, and site-dependent location and scale parameters. For each station we calculated the PPCC's by assuming that the shape parameter γ was $1,2,3,\ldots 50$. For samples of data based on eight-day intervals the mean and median of the PPCC's, taken over all the stations, were largest for $\gamma=11$ and $\gamma=13$, respectively. If it were true that a reverse Weibull distribution with a single tail length parameter characterized the extreme winds at all sites, then our analyses would indicate that the value of that parameter is $\gamma\approx12$.

The assumption that there exists a universal tail length parameter for extreme wind distributions is implicit in current practice, except that it is applied to the Gumbel distribution (for which $\gamma=\infty$). To see whether that assumption is tenable if applied to the reverse Weibull distribution with $\gamma\approx12$, 44 samples corresponding to 18-year record lengths based on 8-day intervals were generated from reverse Weibull populations with (1) $\gamma=8$, (2) $\gamma=12$, and (3) $\gamma=16$. The number of simulated samples for which the best fitting reverse Weibull distribution had shape parameters with $\gamma\leq12$, $13\leq\gamma\leq20$, and $\gamma\geq21$ are shown in Table 1. Also shown in Table 1 are the numbers of observed samples based on 8-day intervals for which the analysis yielded $\gamma\leq12$, $13\leq\gamma\leq20$, and $\gamma\geq21$. The results of Table 1 would suggest that a reverse Weibull distribution with $\gamma\approx12$ is an appropriate model for the populations of extreme winds representing data based on 8-day intervals, except for the larger number of samples with $\gamma\geq20$ among the observed samples than among the simulated samples. We interpret this larger number as reflecting the relatively frequent presence of outliers among the observed samples. This may suggest that, because wind speed populations are mixed (in addition to extremes they include ordinary winds whose meteorological structure may differ from that of the extremes), a sample taken from such a population is likely not to be a sound basis for inferences on extremes. It is therefore desirable to "let the tails speak for themselves." The application of the GPD-based 'peaks over threshold' approach is an attempt to do just this.

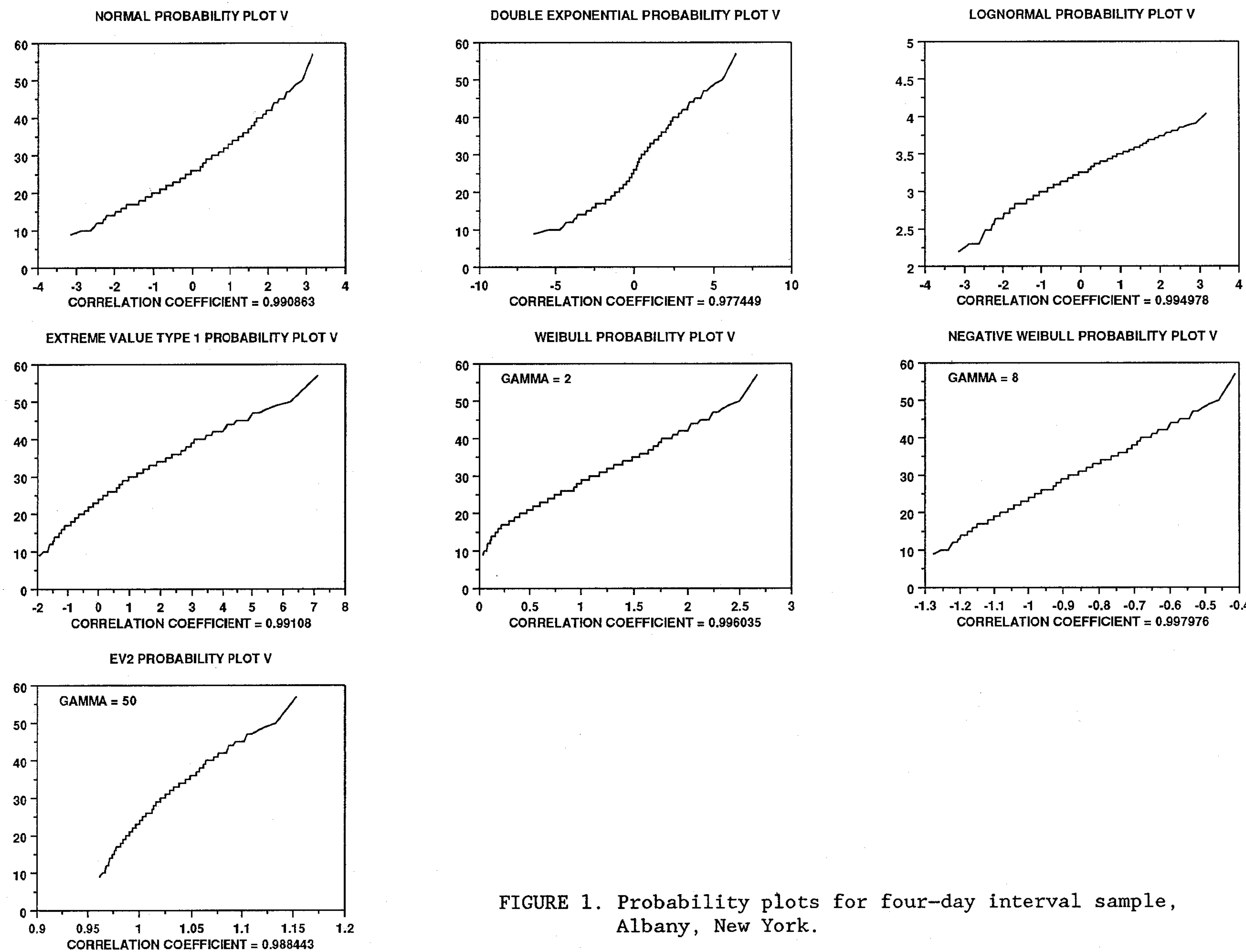


FIGURE 1. Probability plots for four-day interval sample, Albany, New York.

Table 1. Comparison of Results for Simulated and Observed Samples

	$\gamma \le 12$	$13 \le \gamma \le 20$	$\gamma \ge 21$
Simulated samples, $\gamma = 8$	44	0	0
Simulated samples, $\gamma = 12$	26	15	3
Simulated samples, $\gamma = 16$	7	22	15
Observed samples*	25	10	9

*Stations for which $\gamma \ge 21$ were: Green Bay, Greensboro, Huron, Lansing, Louisville, Macon, Moline, Portland, OR, San Diego. Those for which $13 \le \gamma \le 20$ were: Binghamton, Fort Smith, Fort Wayne, Grenville, Milwaukee, Minneapolis, Springfield, Topeka, Tucson, Yuma.

5.2 Estimation of Tail Length Parameter by 'Peaks Over Threshold' Analyses of Uncorrelated Data Samples. We applied the de Haan estimation method (Section 3) to the four-day interval samples using, for each sample, a highest threshold such that the number of its exceedances be equal to, or larger than and as close as possible to, 16; any higher threshold was deemed to result in data sets too small to yield useful statistics. Denoting a sample's maximum threshold by u_{max}, the next higher thresholds we considered were $u_{max}-1$, $u_{max}-2$, $\ldots, u_{max}-24$. The estimated values (point estimates) of c are shown for each station in the plots of Fig. 2. Also shown on the plots are 95 percent confidence bounds (i.e., lines corresponding to $\hat{c} \pm 2\text{s.d.}(\hat{c})$). On the horizontal coordinate axis of each plot we indicate the thresholds, in miles per hour, and the size of the data samples (i.e., the number of exceedances) for each threshold.

Figure 3 is an example of a similar plot ($\hat{c}$ versus number of threshold exceedances) presented for a different type of extreme value problem by de Haan (1990). Like the plots of Fig. 2, this plot exhibits fairly strong fluctuations in the region of the highest thresholds where the sample size is relatively small. In the region of the smaller thresholds the 95 percent confidence bounds become narrower — a result of the increasing sample size — but a bias sets in, which is due to the inclusion in the data samples of data not properly belonging to the tails. In de Haan's judgment, "it looks from the graph as if the value c=0 is not a bad choice in this case."

We propose to apply this type of qualitative judgment to the plots of Fig. 2. For example, it would appear that, for Abilene, $c<0$, perhaps $c \approx -0.25$. The plots of Fig. 2 indicate that $c<0$ for most, though not all, stations. This in itself is an interesting result, insofar as it would indicate that in most cases extreme wind distribution tails are indeed finite.

Let us again assume for a moment that extreme wind speeds in regions not subjected to mature hurricanes are described by a reverse Weibull distribution with site-dependent location and scale parameters and site-independent tail length parameter c. The weighted mean of c may be written as a function of threshold order q as (Gross et al., 1995):

$$\hat{c}_{wq} = \{ \sum_{i=1}^{44} \hat{c}_{iq}/s_{iq}^2 \} / \sum_{i=1}^{44} 1/s_{iq}^2 \qquad (20)$$

where the index q = 1,2,..,25 is the order of the highest, second highest,..,25-the highest threshold for the 44 samples being analyzed, and $\hat{c}_{iq}$, s_{iq} are the estimated value of c and the estimated standard deviation of c for station i, based on the threshold of order q. (Recall that the threshold corresponding to

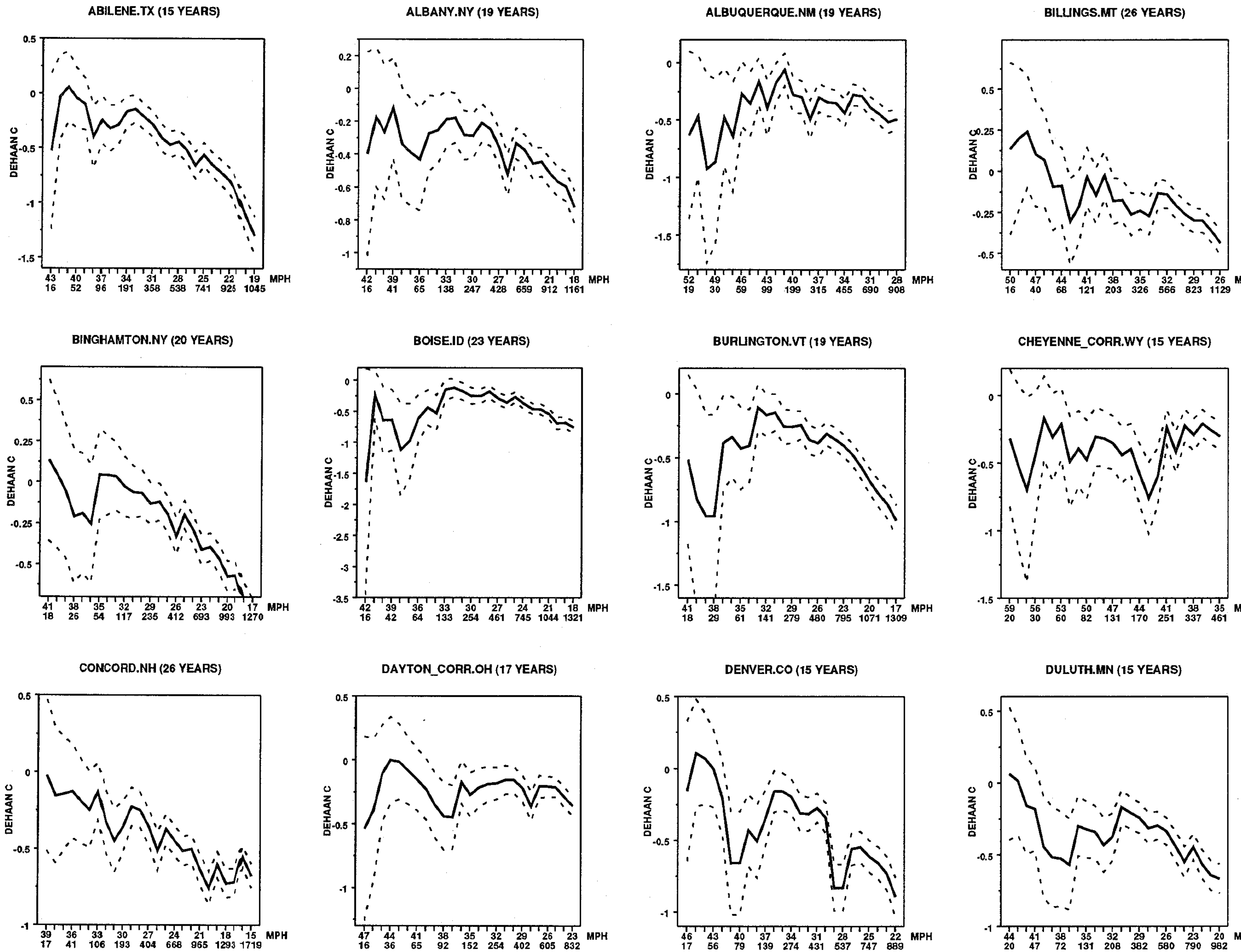


FIGURE 2. Plots of parameter $\hat{c}$ and 95% confidence bounds versus threshold and number of threshold exceedances (based on four-day interval samples).

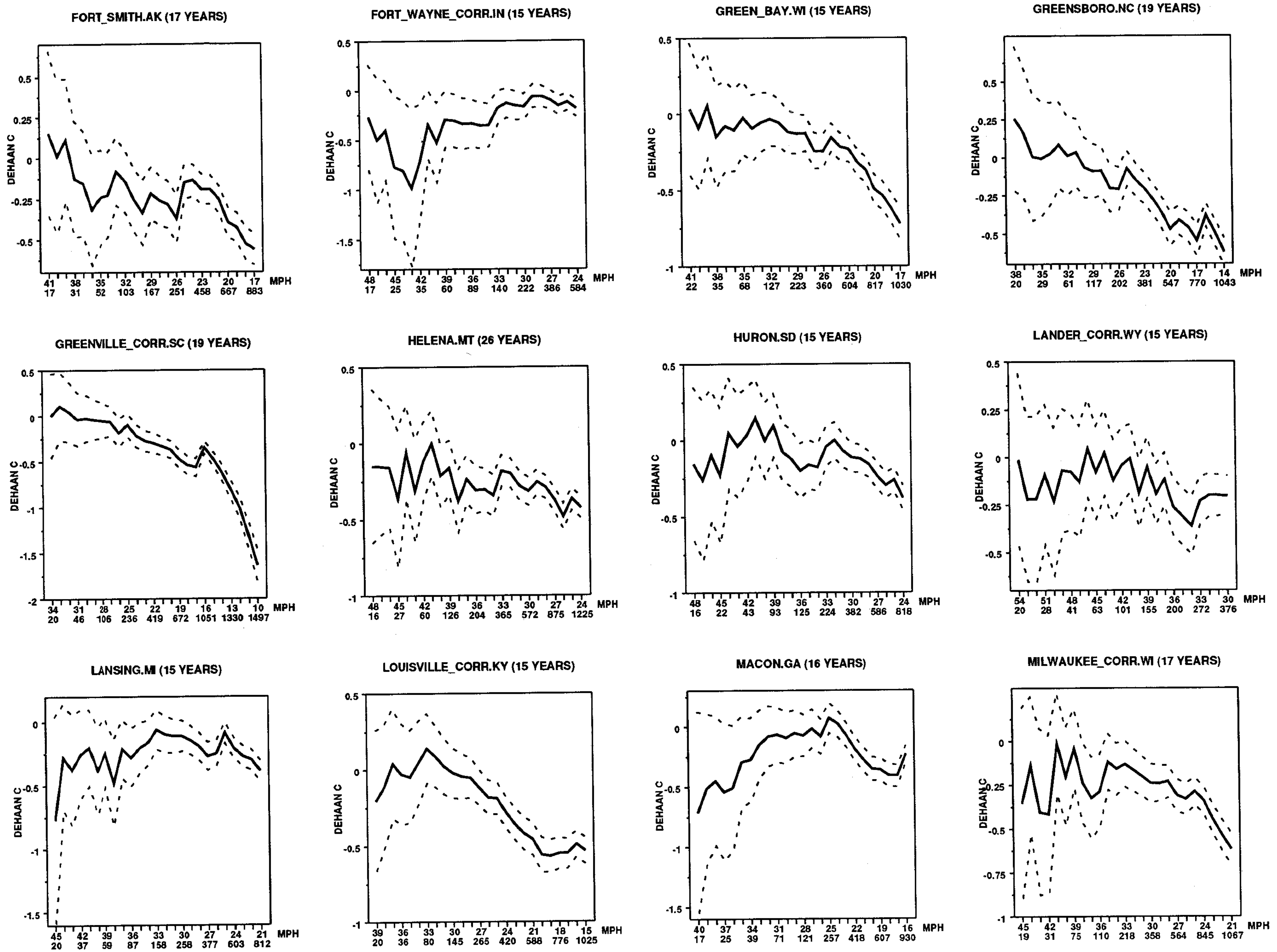


FIGURE 2 (continued)

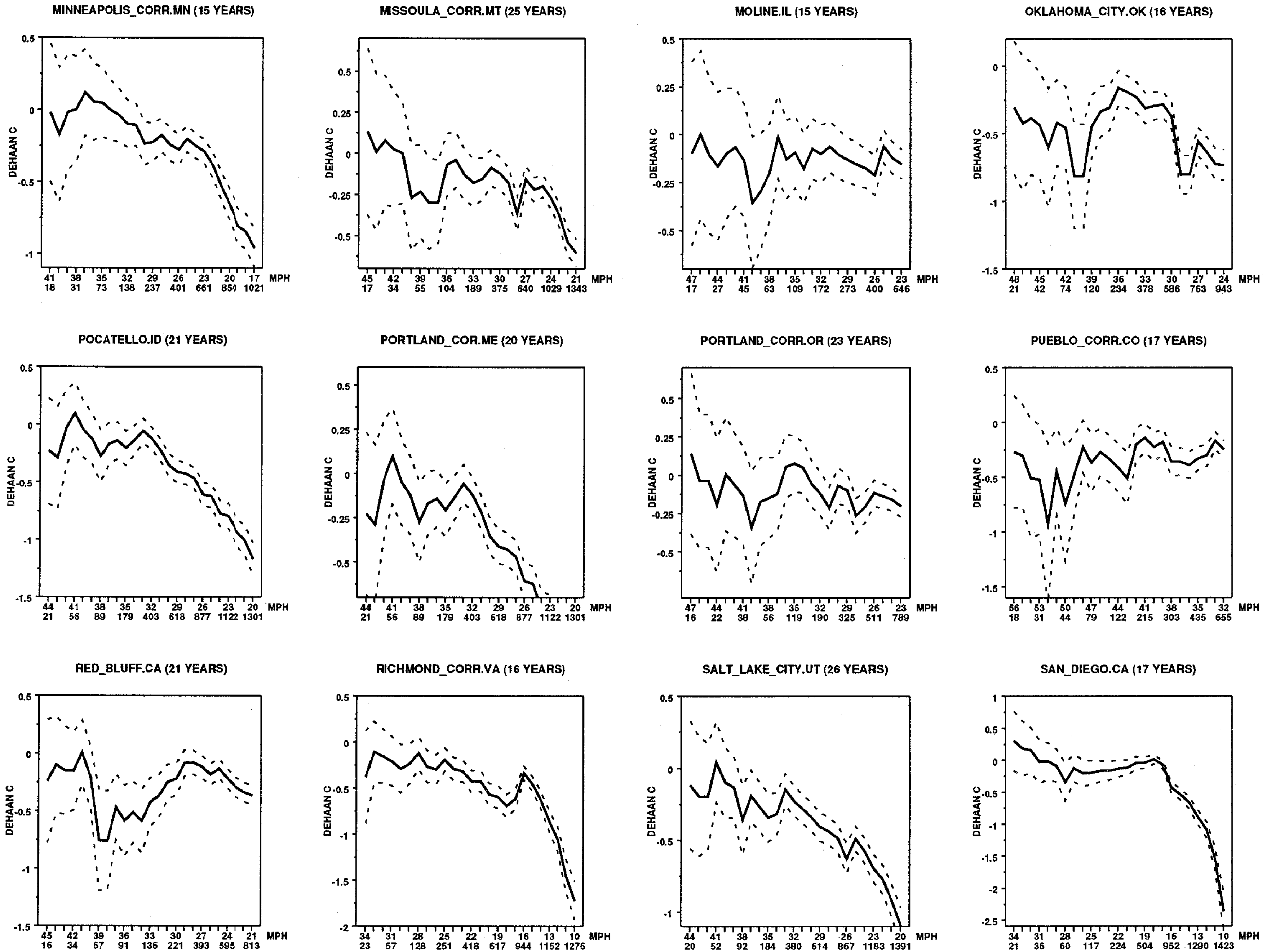


FIGURE 2 (continued)

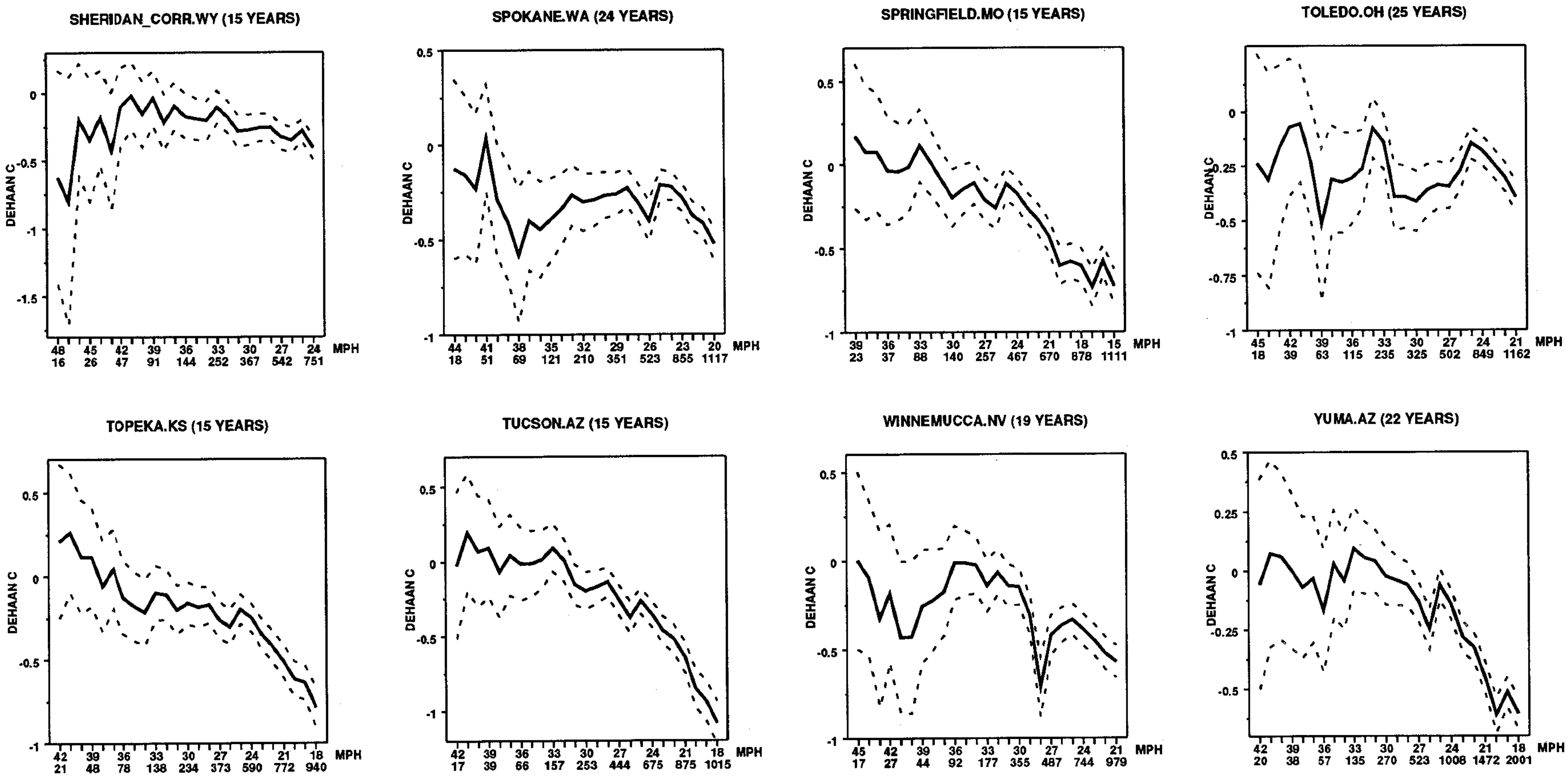


FIGURE 2 (continued)

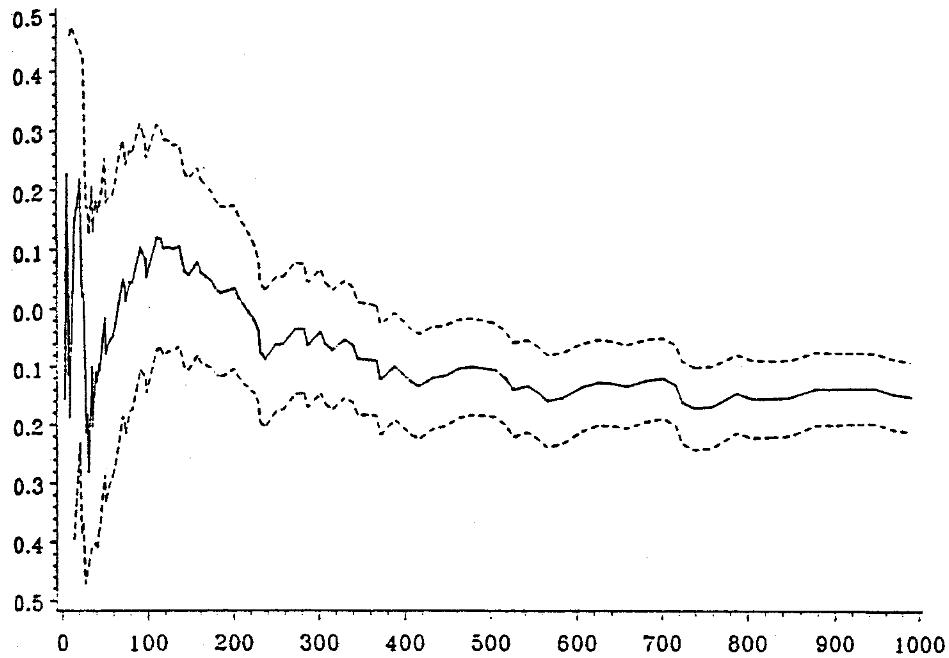


FIGURE 3. Point estimates of tail length parameters and 95% confidence bounds versus number of exceedances (de Haan, 1990).

$q = 1$ for each station was so chosen that at least 16 data points exceed that threshold.) The plot of c_{wq} is shown in Fig. 4 and, in our opinion, tends to confirm the view that, at most if not all stations, the estimated value of c is negative, perhaps $c \approx -0.2$ or $c \approx -0.25$. We note that, as suggested by Monte Carlo simulations (Gross et al., 1994), for sample sizes not exceeding about 10 percent of the total number of data, the bias in the estimation of c is about −0.05, that is, sufficiently small not to invalidate our judgement that, predominantly, $c<0$.[2]

5.3 Estimation of Tail Length Parameter by 'Peaks Over Threshold' Analyses of Largest Yearly Data Samples. Appendix 2 includes point estimates of c for each of the 115 stations for which largest yearly speeds were available. Also shown on the plots are 95 percent confidence bounds (i.e., lines corresponding to $\hat{c} \pm 2\text{s.d.}(\hat{c})$). The estimates are plotted against the threshold speed and the number of exceedances of the threshold, as in Fig. 2. For these plots the larger samples are not likely to be affected by bias, since the lowest wind speed in those samples is itself a largest yearly wind, and hence it will be within or close to the distribution tail. Though the plots are not always easy to interpret, in our opinion they confirm the view that at most stations c is negative. Figure 5, which shows the weighted average of the estimated tail length parameter for the 115 data samples (eq (20)), lends further credence to this view.

---

[2]We note a typographical error in Table 5, p. 147 of Gross et al. (1994). In the last line of the Table the population value of the parameter c should be −0.275 (as in line 7 of p. 142), rather than −0.50.

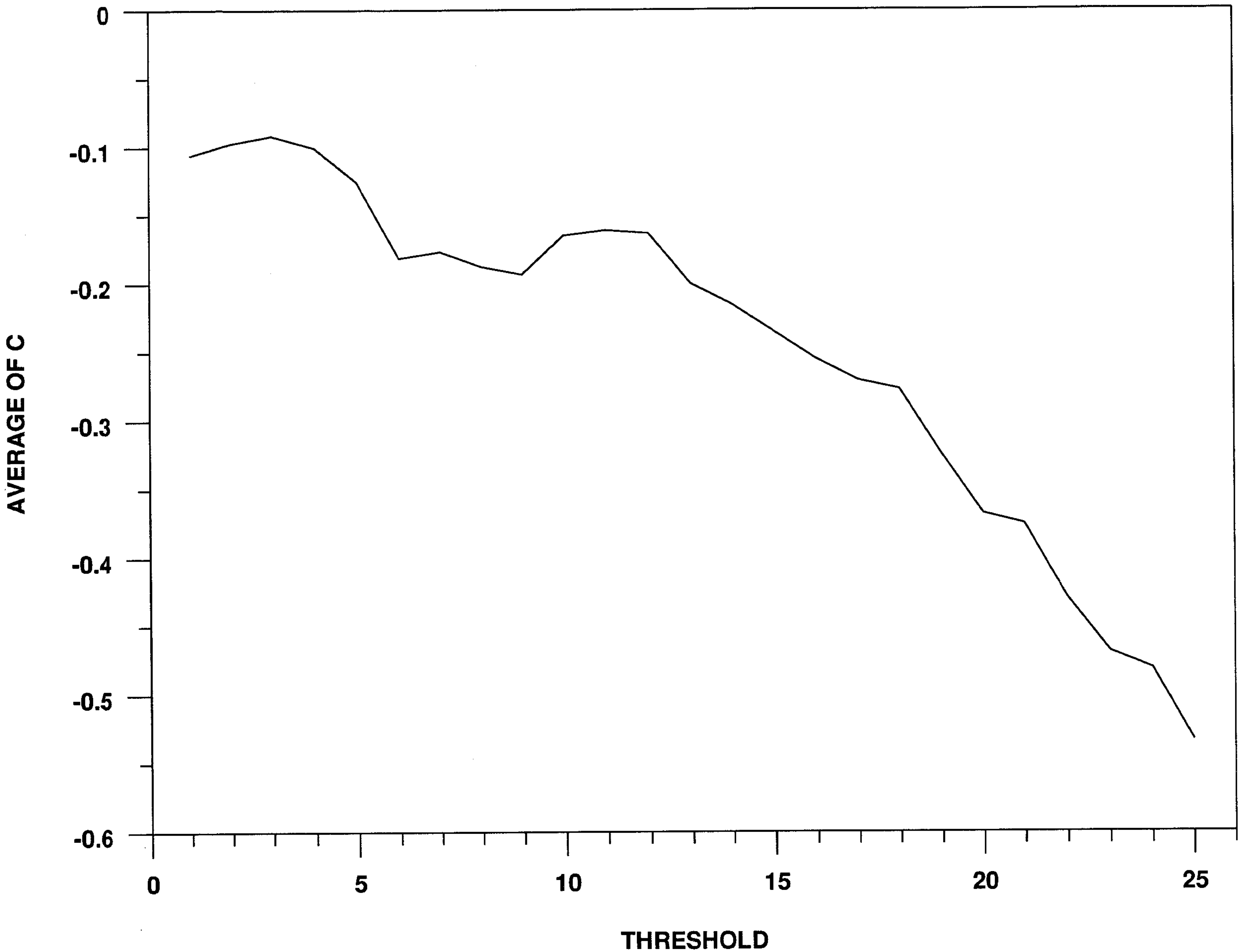


FIGURE 4. Mean of tail length parameter estimates weighted over 44 four-day interval samples versus order of threshold.

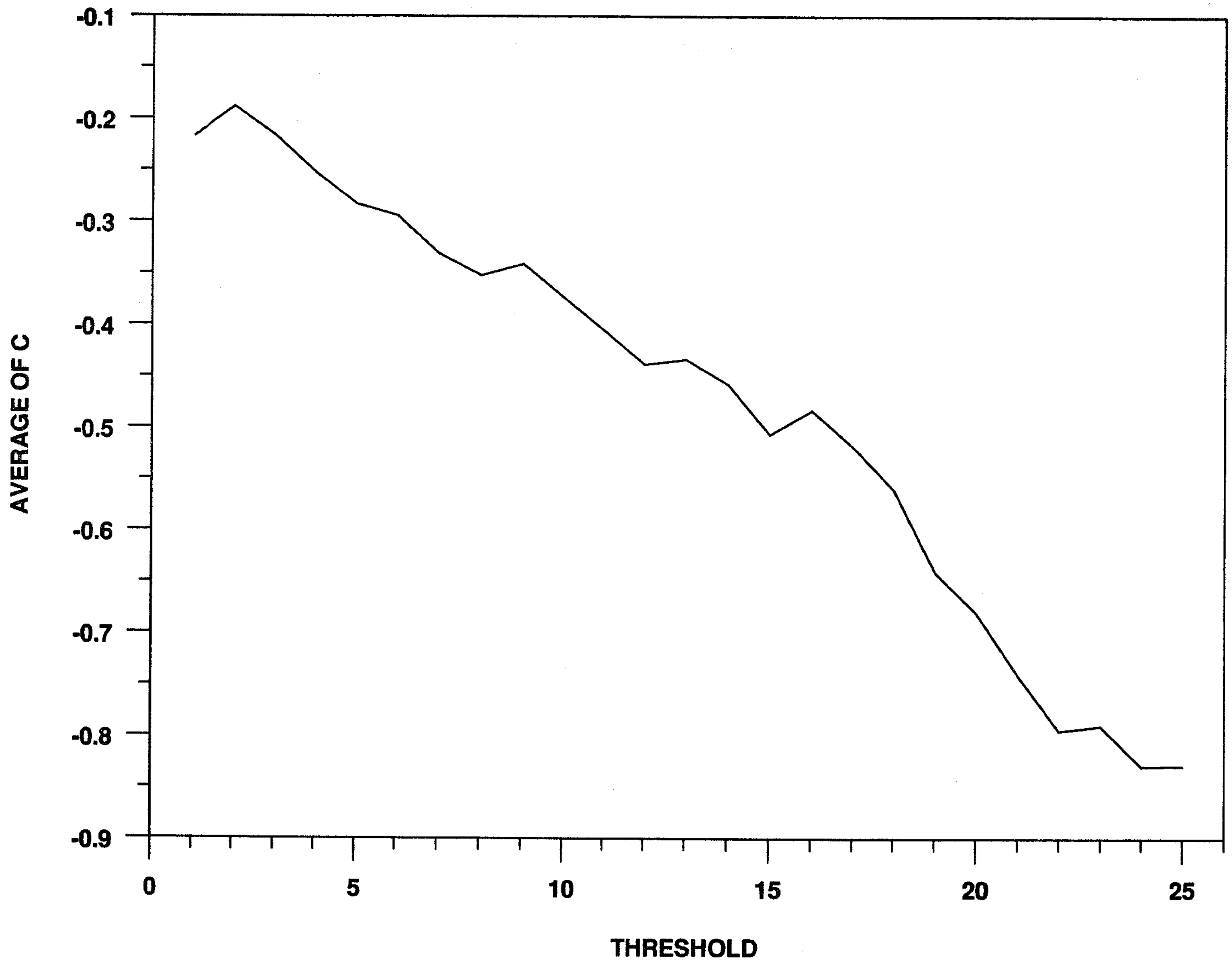


FIGURE 5. Mean of tail length parameter estimates weighted over 115 largest yearly data samples versus order of threshold.

5.4 Estimation of Speeds with Specified Mean Recurrence Intervals by 'Peaks over Threshold' Analyses of Uncorrelated Data Samples. Appendix 3 contains plots of point estimates of the extreme winds with mean recurrence intervals of 100, 1,000 and 100,000 years. The estimates were based on four-day interval samples at each of the 44 stations. They are plotted against the threshold speed and the number of exceedances of the threshold, as in Fig. 2.

We reproduce in Fig. 6 an example of a plot where the quantile fluctuates strongly as a function of threshold (de Haan, 1990). De Haan comments: "if one would be forced to give a point estimate a value of 5.1 m... would not be unreasonable." The comment is indicative of the spirit in which results based on the 'peaks over thresholds' method must be interpreted in cases where fairly large fluctuations are present, as is the case for Fig. 6 and many of the plots in Appendix 3. We do not attempt in this report to estimate extreme wind speeds for various mean recurrence intervals. Rather, having found that the tail length parameter c of the GPD is negative for the majority of the stations, we assess in the next section the potential implications of this finding for the estimation of load factors.

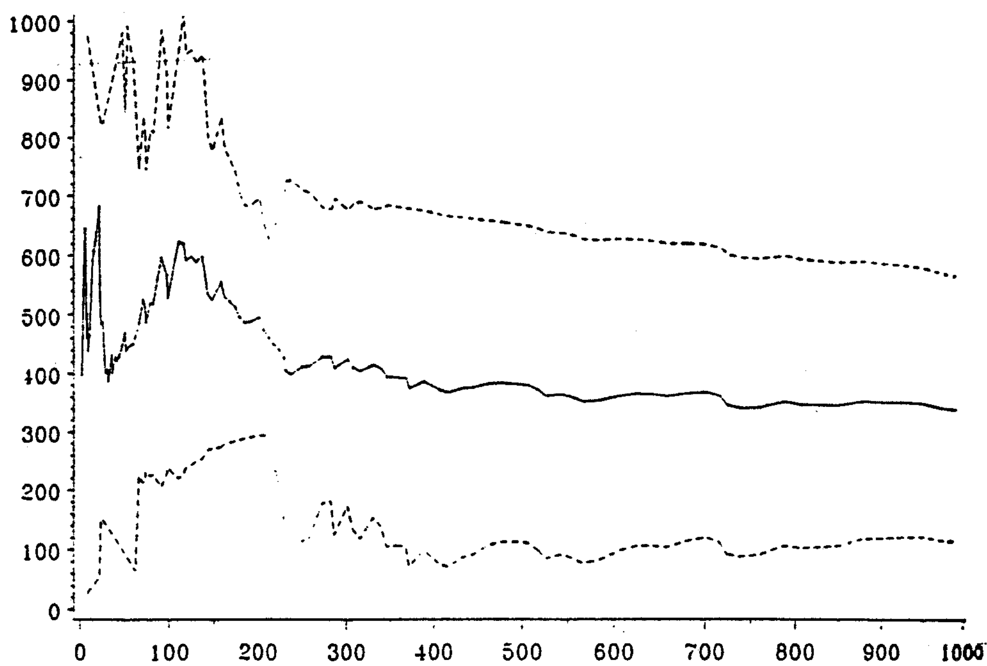


FIGURE 6. Quantiles: point estimates and 95% confidence bounds versus number of exceedances (de Haan, 1990).

5.5 Influence of Data Errors on Analysis Results. The footnote to Section 4.1 lists errors in the recorded daily data found at fourteen stations, and the respective corrected values. Figures 2 and 4 and Appendix 3 are based on the corrected data sets (these are marked in Fig. 2 and Appendix 3 by the suffix "corr"). Figures 7 and 8 include, respectively, plots of estimated tail length parameters and speeds with various mean recurrence intervals for twelve of these stations and are based on the uncorrected data. Comparisons between these plots and their counterparts in Fig. 2 and Appendix 3 show that estimates of speeds corresponding to various mean recurrence intervals are in some cases affected fairly significantly by the errors in the data. This is true to a much lesser extent for the tail length parameters. The plot of the weighted mean over all 44 stations, computed from results obtained by using the uncorrected data, was indistinguishable from Fig. 4.

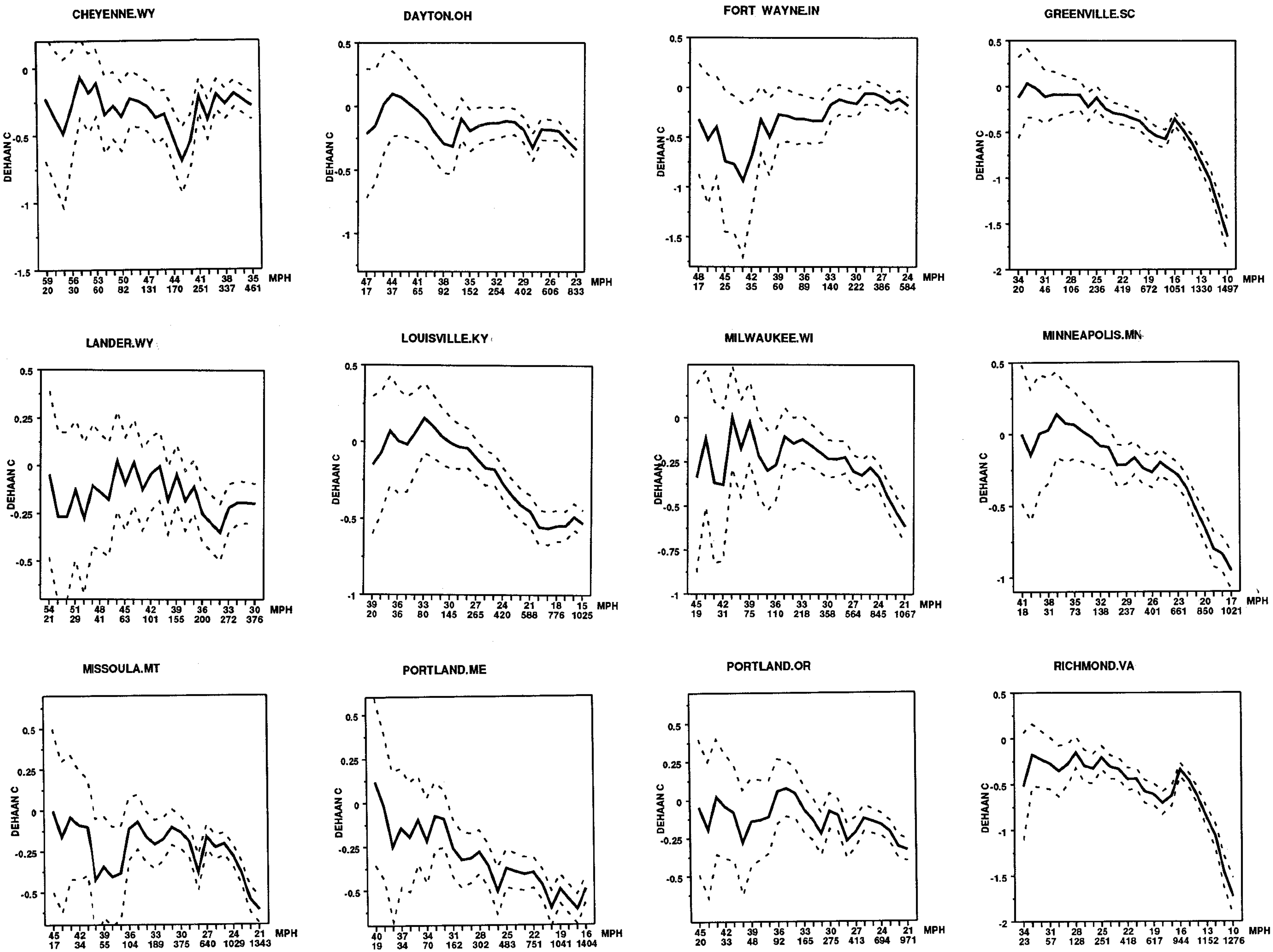


FIGURE 7. Plots of parameter $\hat{c}$ and 95% confidence bounds versus threshold and number of threshold exceedances (based on four-day interval samples, uncorrected data).

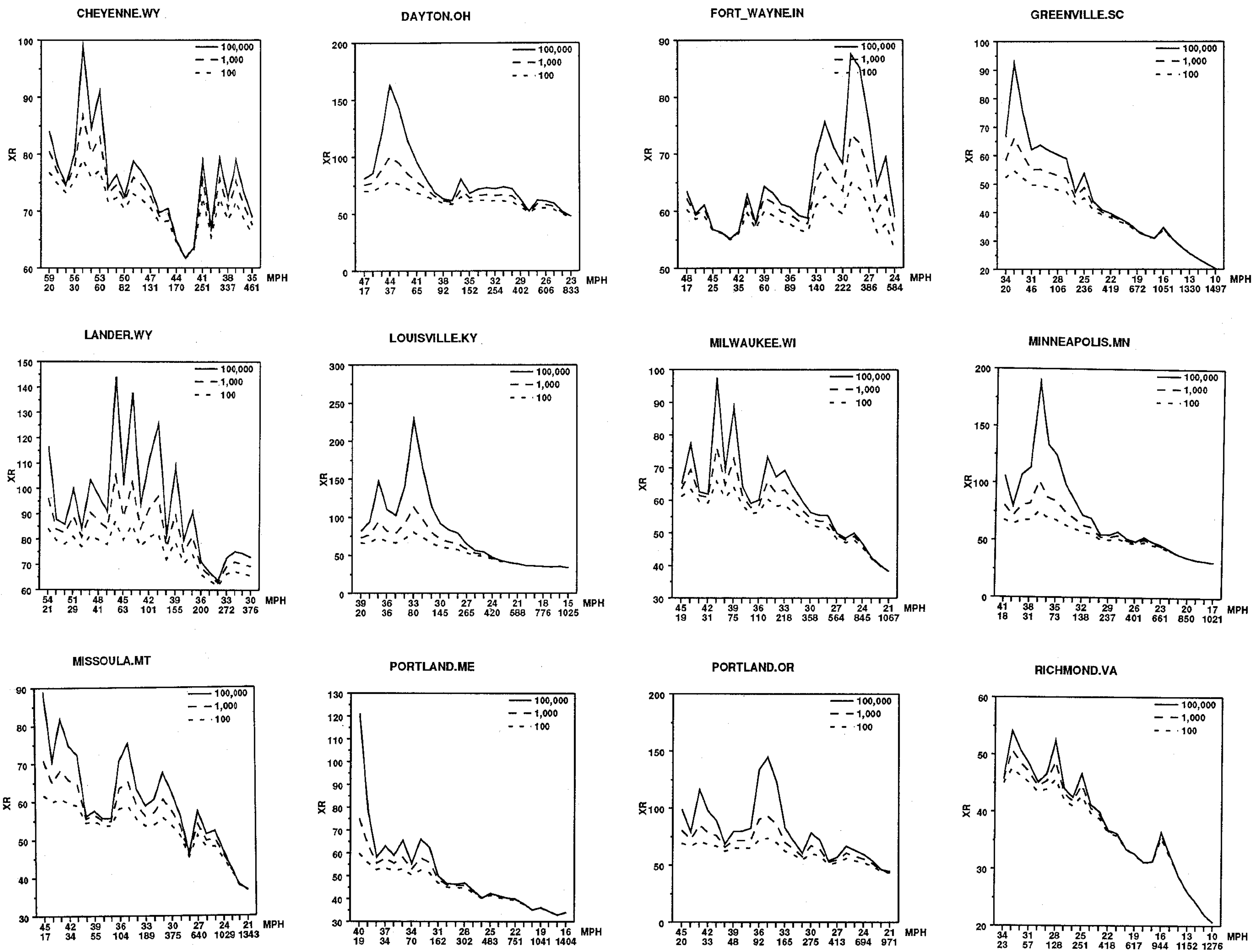


FIGURE 8. Point estimates of 100-, 1000- and 100 000-yr speeds versus threshold and number of threshold exceedances (based on 4-day interval samples, uncorrected data).

6. LOAD FACTORS FOR WIND-SENSITIVE STRUCTURES

Extreme wind loads used in design include nominal **basic design** wind loads (e.g., the 50-yr wind load) and **nominal ultimate** wind loads. A **basic design** wind load is an extreme load with specified probability of being exceeded during a basic time interval. In the United States that interval is usually 50 years. A basic design load with a 50-year mean recurrence interval has a probability of almost two thirds of being exceeded during a 50-year period.

A structure or element thereof is expected to withstand loads substantially in excess of a 50- or 100-year wind load without loss of integrity. The wind load beyond which loss of integrity can be expected is referred to as the **nominal ultimate** wind load. The nominal ultimate strength provided for by the designer is based on a nominal ultimate wind load equal to the design wind load times a wind load factor. This statement is valid for the simple case where wind is the dominant load. It needs to be modified if load combinations are considered, but for clarity we refer here only to this case.

The load factor should be selected so that the probability of occurrence of the nominal ultimate load is acceptably small. This probabilistic concept is important from an economic or insurance point of view. To the extent that evacuation or similar measures cannot be counted on to prevent loss of life, it is also important from a safety point of view.

A probabilistic approach has proven helpful in a number of cases, particularly for relative assessments of alternative design provisions, for example for mobile homes. However, in most cases the difficulties of obtaining wind load factors by probabilistic methods have proven to be substantial if not prohibitive. For this reason code writers have largely relied on wind load factors implicit in traditional codes and standards. For example, the American Society of Civil Engineers Standard A7-93 (1993) specifies a wind load factor of 1.3. In a very large number of applications the wind load is proportional to the square of the wind speed, so that a basic design wind speed and a nominal ultimate wind speed may be defined that are proportional to the square root of the basic design wind load and the square root of the nominal ultimate wind load, respectively. For example, for Lander, Wyoming, the American Society of Civil Engineers Standard A7-93 (1993) specifies a basic design 50-year design speed of 35.8 m/s (80 mph fastest-mile) at 10 m (33 ft) elevation. The corresponding nominal ultimate wind speed would then be $1.3^{1/2}35.8=40.8$ m/s (91.2 mph).

Reliance on traditional code values is part of the process sometimes referred to as "calibration against existing practice." Traditional codes were generally adequate for many types of structures, but questions remain on whether safety margins implicit in those codes may be applied to modern structures, which can differ substantially from their predecessors in their materials and design/construction techniques. For this reason an assessment of wind load factors used in codes and standards would be desirable. For example, one would wish to answer the question: what is the approximate mean recurrence interval of the nominal ultimate wind speed?

The answer to this question depends strongly upon the probability distribution assumed to best fit the extreme wind speeds. For example, a PPCC analysis of largest yearly fastest-mile speeds recorded at Denver between 1951-1977, based on the assumption that the best fitting distribution is Gumbel, yielded a 27.9 m/s (62 mph) estimate of the 50-year wind speed at 10 m above ground. The corresponding nominal ultimate wind speed would be $1.3^{1/2}27.9=31.8$ m/s (71.0 mph), to which there would correspond, under the Gumbel assumption, a mean recurrence interval of about 500 years (Simiu, Changery and Filliben, 1979). If taken at face value this would be an alarmingly short recurrence interval, since it would entail an unacceptably large probability of exceedance of the nominal ultimate wind load during the life of the structure.

However, the 500 years mean recurrence interval is based on the Gumbel model. Since our results support the assumption that, predominantly, the appropriate model is a reverse Weibull distribution, rather than a Gumbel distribution, we wish to answer the question: what is the mean recurrence interval corresponding to $1.3^{1/2}$ times the wind speed with a 50-year mean recurrence interval? For Denver, if one had to estimate the tail length parameter $\hat{c}$ from the plots of Fig. 2 and Appendix 2, and the 50-year speed from the plot of Appendix 3, one might choose, say, $\hat{c}=-0.2$ (a conservative choice: according to the plots $\hat{c}$ is likely to be somewhat lower, that is, the distribution tail is likely to be somewhat shorter than that corresponding to $c=-0.2$), and $x_{50}=26.8$ m/s (60 mph). For a threshold of 16.5 m/s (37 mph) — a value that is roughly consistent with these choices, see Denver plot, Fig. 2 — we have $\lambda=139/15=9.27$/year. Assuming $x_{50}=26.8$ m/s (60 mph), it would follow from eqs (13) and (14) $\hat{a}=2.5$ m/s (5.6 mph). The estimated maximum possible wind speed corresponding to the parameters $\hat{c}=-0.2$ and $\hat{a}=2.5$ m/s (5.6 mph) is obtained by letting $R \rightarrow \infty$ in eqs (13, 14). Its value is $x_{max}=u-\hat{a}/\hat{c} = 16.5+2.5/0.2 = 29.0$ m/s (65 mph). The estimated mean recurrence interval of the nominal ultimate wind speed $1.3^{1/2}x_{50}=1.3^{1/2}x26.8=30.6$ m/s (68 mph) is therefore infinity (i.e., such a wind speed is estimated to never occur). This estimate is of course subject to sampling errors: the actual maximum possible wind speed may be higher than 29.0 m/s (65 mph), and the mean recurrence interval of the 30.6 m/s (68 mph) speed may in fact be finite, though likely much longer than 500 years. In spite of the uncertainty inherent in our estimates, our result suggests that a load factor of 1.3 — specified in the ASCE A7-93 Standard on the basis of practical experience — is in fact adequate from a probabilistic point of view. This is contrary to what would be concluded if the analysis were based on the assumption that the Gumbel distribution holds.[3]

---

[3]In this instance code writers attended to the apparent insufficiency of the nominal ultimate design speed by substantially inflating the value of the 50-year wind speed: the ASCE A7-93 Standard specifies for Denver a 35.8 m/s (80 mph) 50-year speed at 10 m above ground. Since the estimated standard deviation of the sampling error for the estimated 50-year wind is $s_{s.e.} \approx 1.3$ m/s (3.0 mph) (Simiu, Changery and Filliben, 1979), the specified 35.8 m/s (80 mph) value differs from the estimated 27.9 m/s (62 mph) value by almost six standard deviations. One may view the actual load factor implicit in the ASCE A7-93 Standard as being equal to $(1.3)(80/62)^2=2.14$, rather than just 1.3. The mean recurrence interval for the 40.7 m/s (91 mph) (i.e., $1.3^{1/2}80 = 2.14^{1/2}x62$) nominal ultimate load inherent in the ASCE A7-93 Standard provisions for Denver, based on the Gumbel model, would be about 80,000 years (Simiu, Changery and Filliben, 1979), which is much more acceptable than 500 years.

7. CONCLUSIONS

In this report we presented estimates of the tail length parameters of extreme wind distributions for non-tornadic winds blowing from any direction in regions unaffected by mature hurricanes. In our opinion, these estimates support the view that the reverse Weibull distribution is an appropriate probabilistic model in most if not all cases. They also suggest that load factors for wind sensitive structures specified by current standards provide for reasonable safety margins against wind loads, and that the adoption of the Gumbel model likely results in an unrealistic assessment of structural reliability under wind loads.

However, owing to fluctuations of our estimates with the threshold value, it is difficult to provide reliable quantitative estimates of the tail length parameters. This difficulty is even more pronounced for quantile estimates.
We tentatively ascribe these difficulties to the relatively small size of our samples (15 to 26 years). It would therefore be desirable to assemble data for longer records than those used in this report. In addition, more efficient estimation methods should be developed. Efforts to develop such methods are currently in progress (Coles, 1994).

REFERENCES

American National Standard A58.1-1972 (1972), American National Standards Institute, New York.

ASCE Standard A7-93 (1993), American Society of Civil Engineers, New York.

Bingham, N.H. (1990), "Discussion of the Paper by Davison and Smith," Journal of the Royal Statistical Society, Series B, 52, p. 431.

Castillo, E. (1988), Extreme Value Theory in Engineering, Academic Press, New York.

Coles, S. (1994), private communication.

Davison, A.C., and Smith, R.L. (1990), "Models of Exceedances Over High Thresholds," Journal of the Royal Statistical Society, Series B, 52, pp. 339-442.

de Haan, L. (1990), "Fighting the Arch-enemy with Mathematics," Statistica Neerlandica 44, pp. 45-68.

de Haan, L. (1994), "Extreme Value Statistics," in Extreme Value Theory and Applications, Vol. 1 (J.Galambos, J. Lechner and E. Simiu, eds.), Kluwer Academic Publishers, Dordrecht and Boston, 1994.

Ellingwood, B. et al. (1980), Development of a Probability-Based Load Criterion for American National Standard A58, NBS Special Publication 577, National Bureau of Standards (U.S.), Washington, DC.

Filliben, J.J., "The Probability Plot Correlation Plot Test for Normality," Technometrics, 17, pp. 111-117.

Gross, J., Heckert, A. Lechner, J., and Simiu, E., (1994), "Novel Extreme Value Procedures: Application to Extreme Wind Data," in Extreme Value Theory and Applications, Vol. 1 (J.Galambos, J. Lechner and E. Simiu, eds.), Kluwer Academic Publishers, Dordrecht and Boston, 1994.

Gross, J.L., Heckert, N.A., Lechner, J.A. and Simiu, E. (1995), "A Study of Optimal Extreme Wind Estimation Procedures," Proceedings, Ninth International Conference on Wind Engineering, New Delhi.

Hosking, J.R.M. and Wallis, J.R. (1987), "Parameter and Quantile Estimation for the Generalized Pareto Distribution," Technometrics, 29, pp. 339-349.

Johnson, N.L. and Kotz, S. (1972), Distributions in Statistics: Continuous Multivariate Distributions, Wiley, New York.

Pickands, J. (1975), "Statistical Inference Using Order Statistics," Annals of Statistics, 3, 119-131.

Simiu, E., Changery, M.J. and Filliben, J.J. (1979), Extreme Wind Speeds at 129 Stations in the Contiguous United States, NBS Building Science Series 118, National Bureau of Standards (U.S.), Washington, DC.

Simiu, E., Filliben, J.J. and Biétry, J., (1978) "Sampling Errors in the Estimation of Extreme Wind Speeds," Journal of the Structural Division, ASCE, **104**, pp. 491-501.

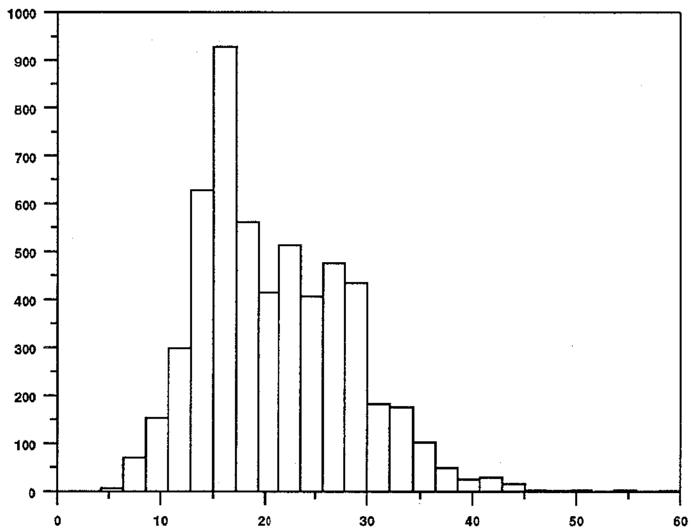
Simiu, E., and Scanlan, R.H. (1986), Wind Effects on Structures, Second Edition, Wiley-Interscience, New York, 1986.

Smith, R. L. (1989), Extreme Value Theory, in Handbook of Applicable Mathematics, Supplement edited by W. Ledermann, E. Lloyd, S. Vajda, and C. Alexander, pp. 437-472, John Wiley and Sons, New York.

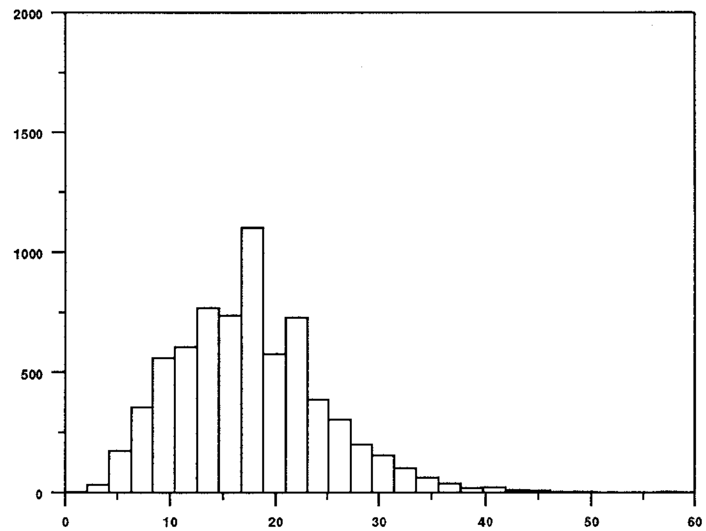
APPENDIX 1

Histograms of daily largest wind speeds (top) and histograms of four-day interval uncorrelated wind speeds (bottom)

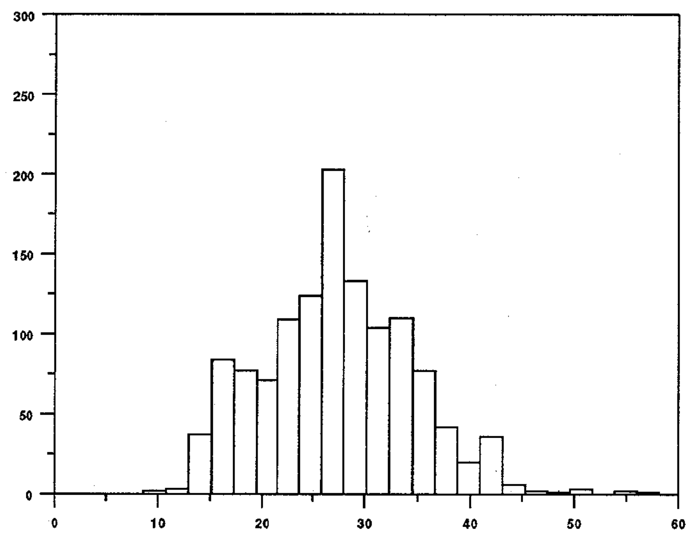
ABILENE.TX - DAILY MAXIMA



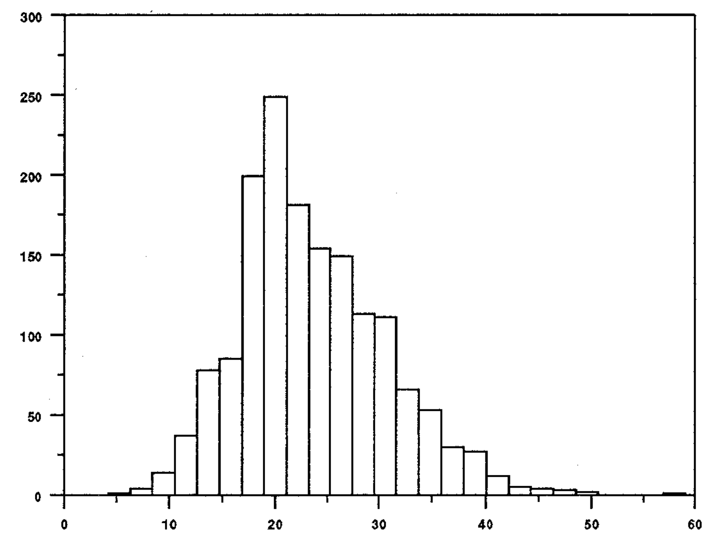
ALBANY.NY - DAILY MAXIMA



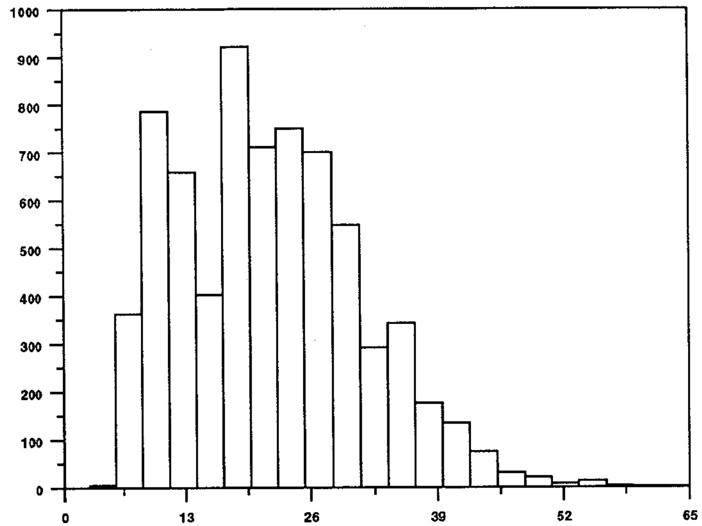
ABILENE.TX - 4 DAY INTERVAL



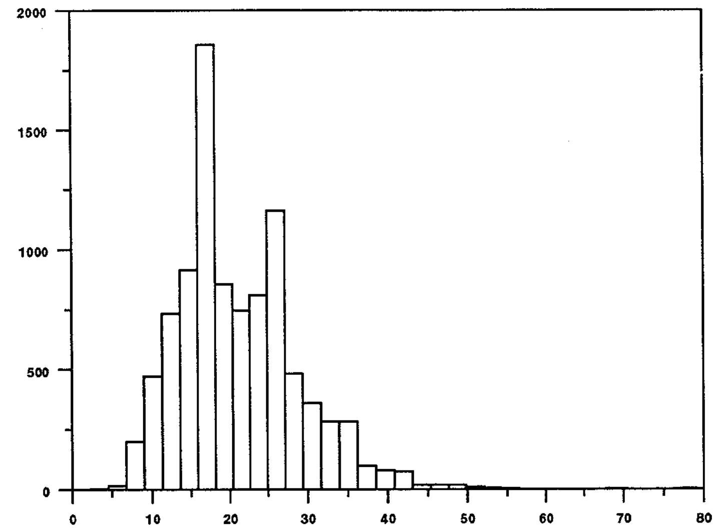
ALBANY.NY - 4 DAY INTERVAL



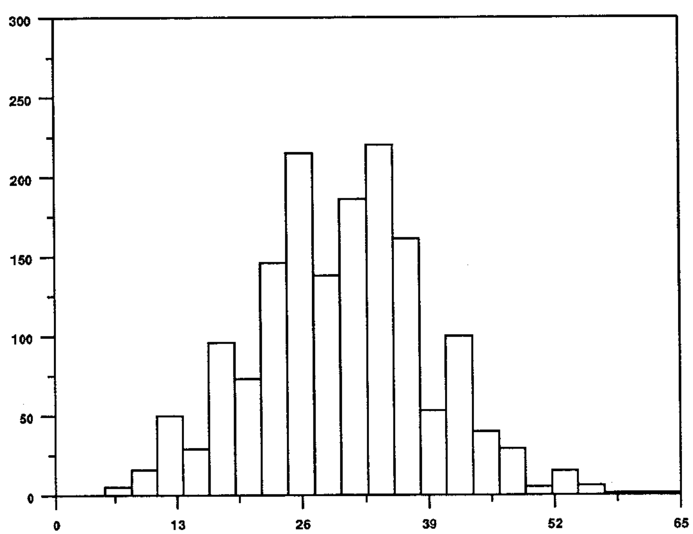
ALBUQUERQUE.NM - DAILY MAXIMA



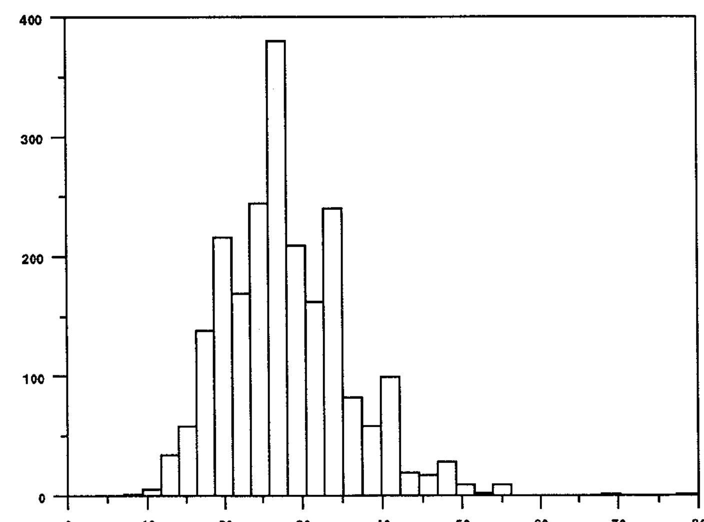
BILLINGS.MT - DAILY MAXIMA



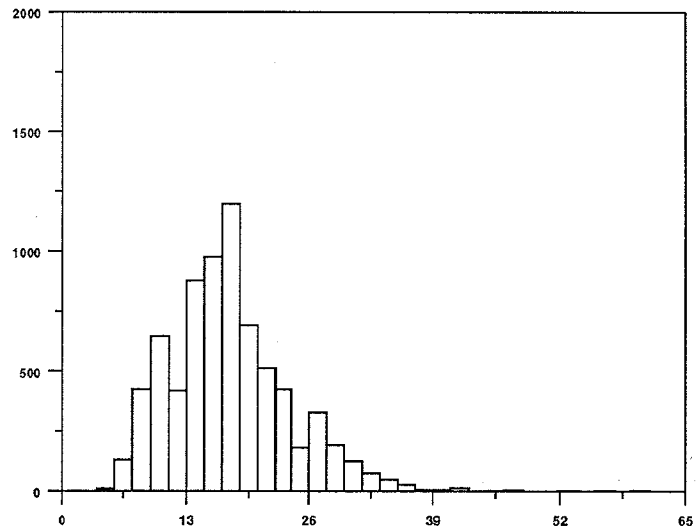
ALBUQUERQUE.NM - 4 DAY INTERVAL



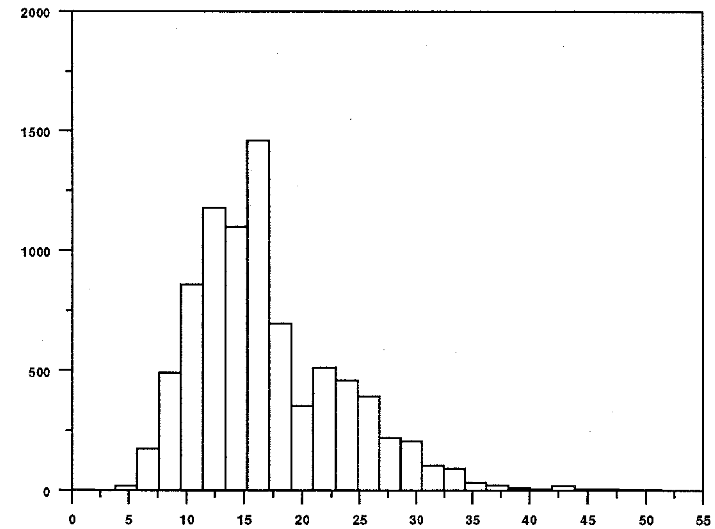
BILLINGS.MT - 4 DAY INTERVAL



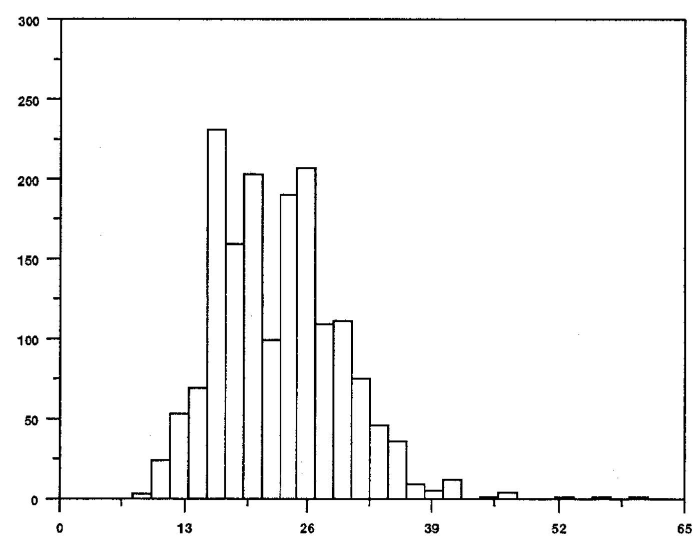
BINGHAMTON,NY - DAILY MAXIMA



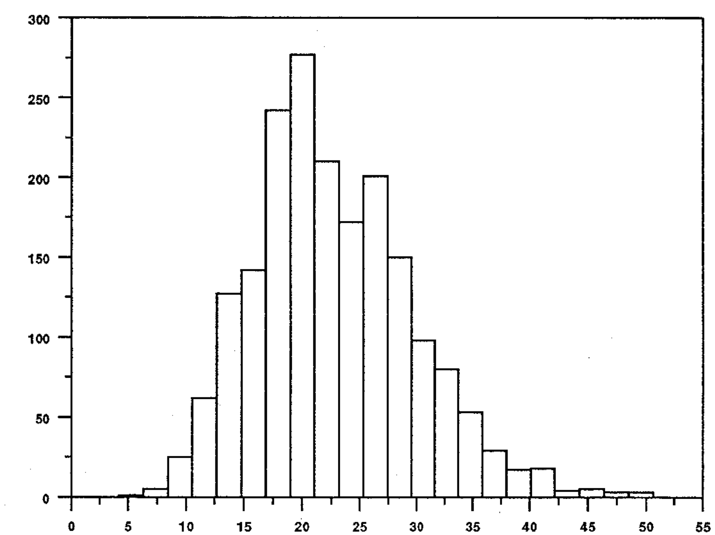
BOISE,ID - DAILY MAXIMA



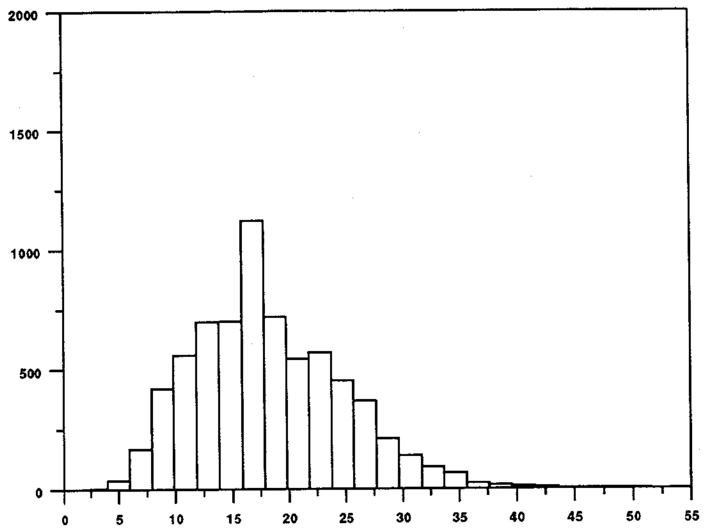
BINGHAMTON,NY - 4 DAY INTERVAL



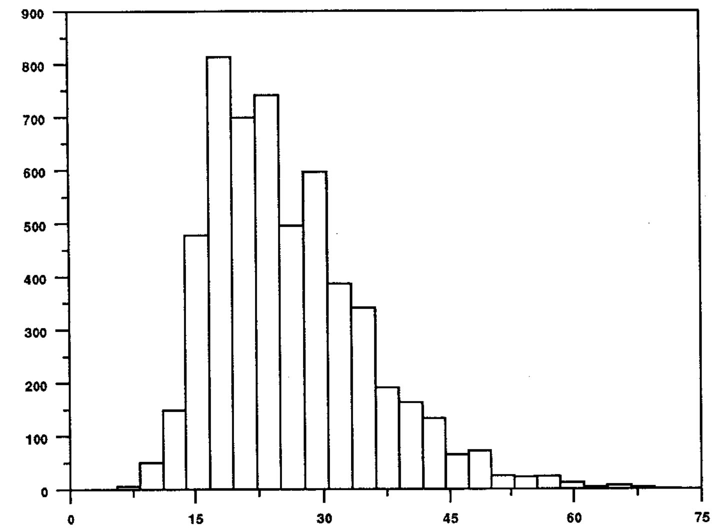
BOISE,ID - 4 DAY INTERVAL



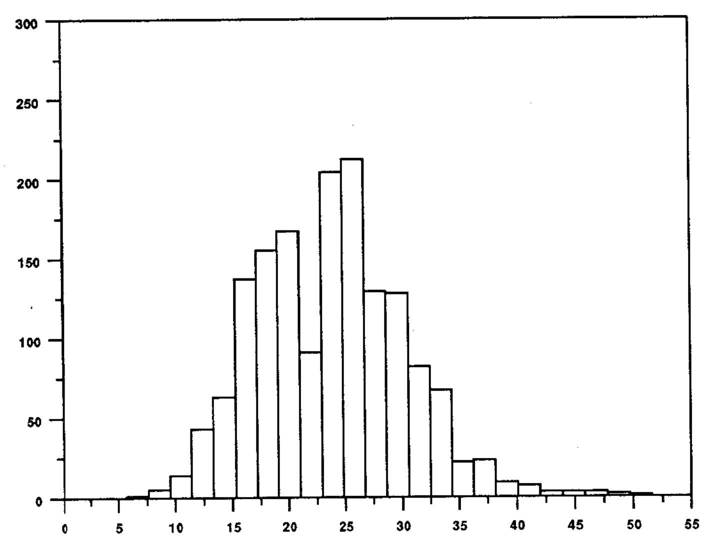
BURLINGTON.VT - DAILY MAXIMA



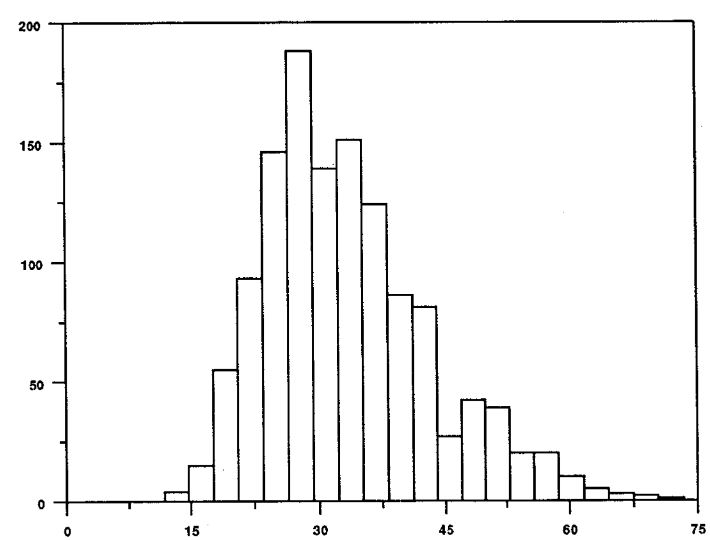
CHEYENNE_CORR.WY - DAILY MAXIMA



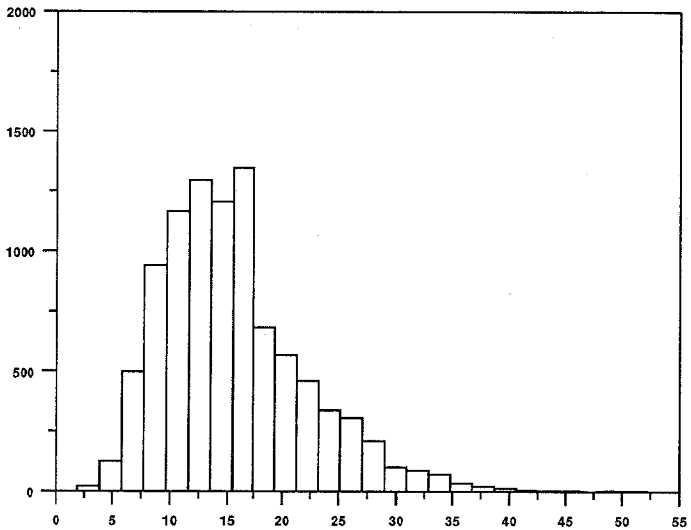
BURLINGTON.VT - 4 DAY INTERVAL



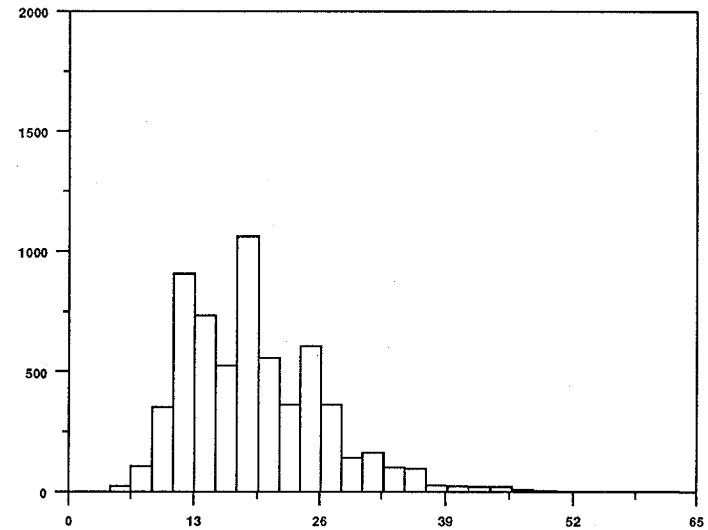
CHEYENNE_CORR.WY - 4 DAY INTERVAL



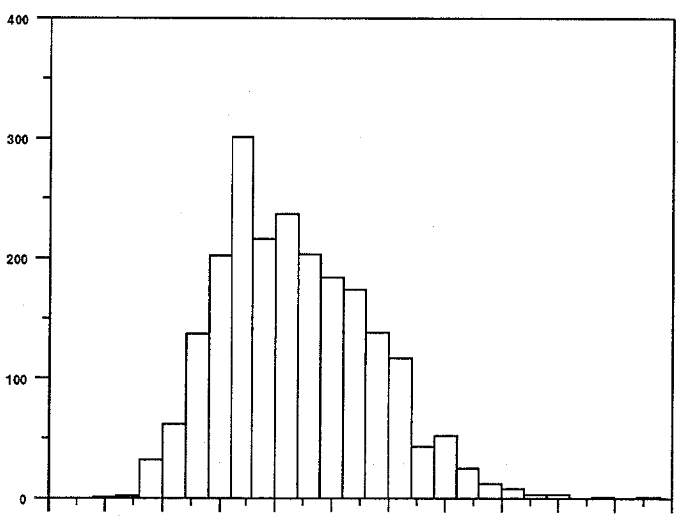
CONCORD.NH - DAILY MAXIMA



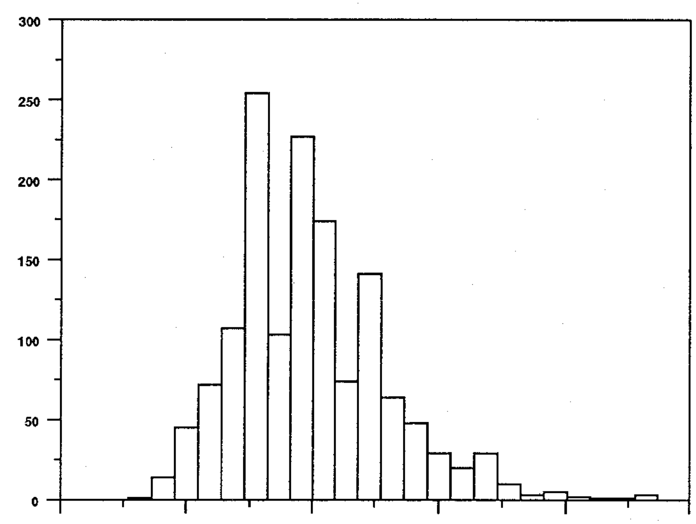
DAYTON_CORR.OH - DAILY MAXIMA



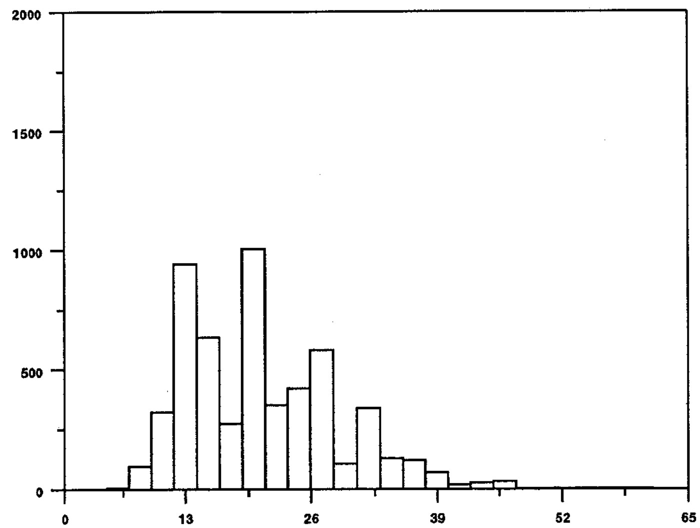
CONCORD.NH - 4 DAY INTERVAL



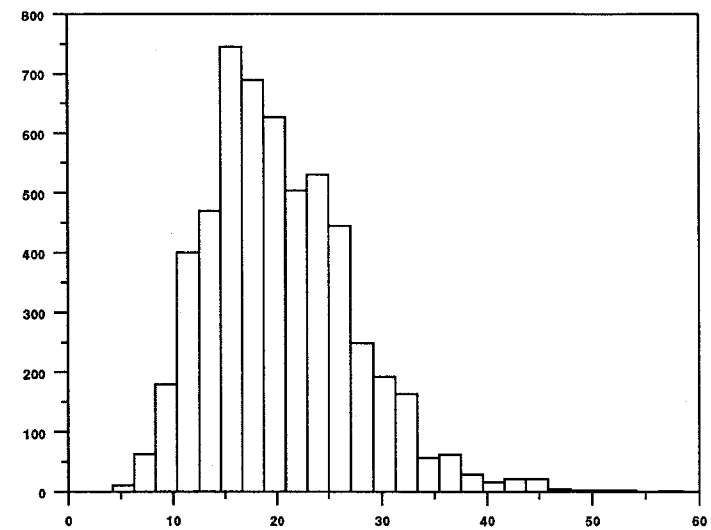
DAYTON_CORR.OH - 4 DAY INTERVAL



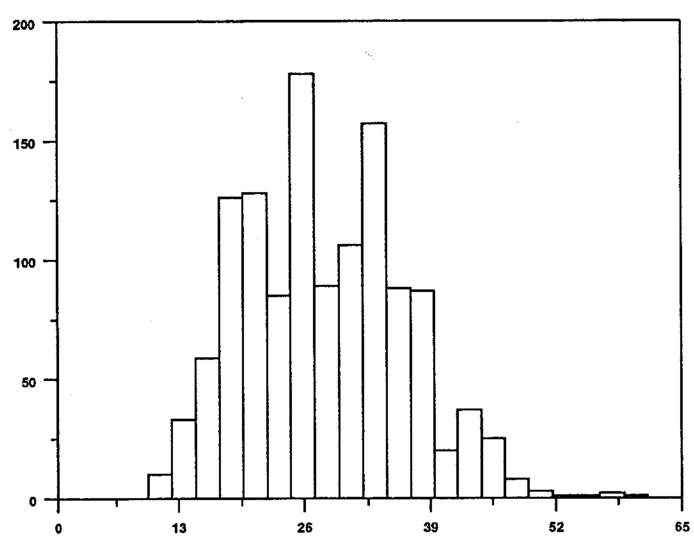
DENVER.CO - DAILY MAXIMA



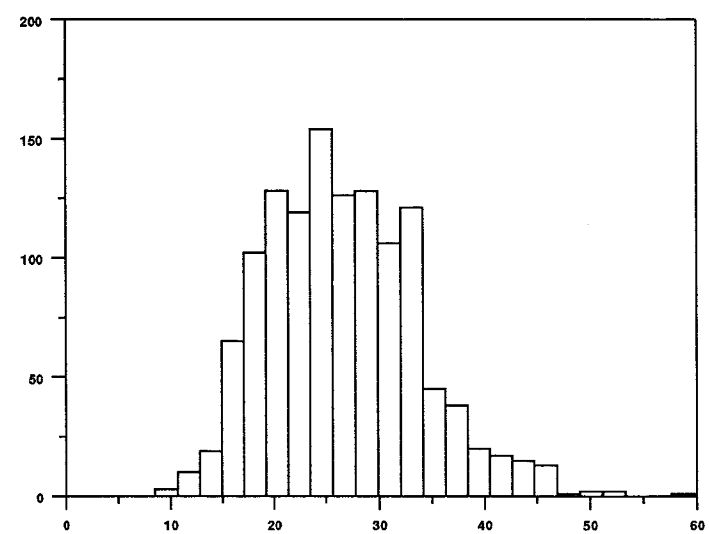
DULUTH.MN - DAILY MAXIMA



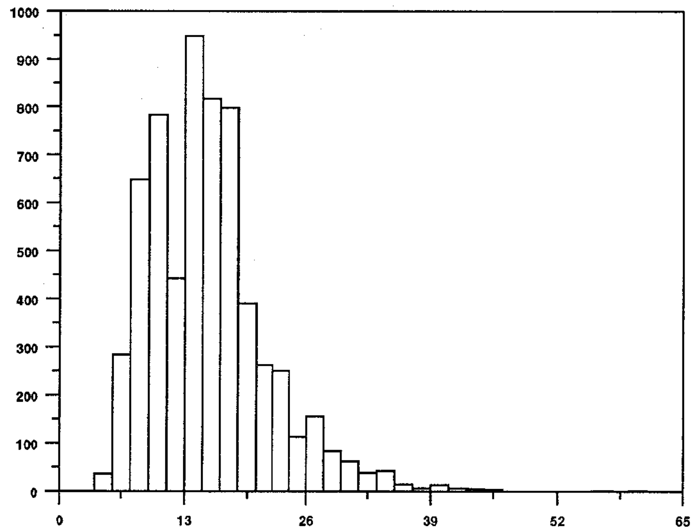
DENVER.CO - 4 DAY INTERVAL



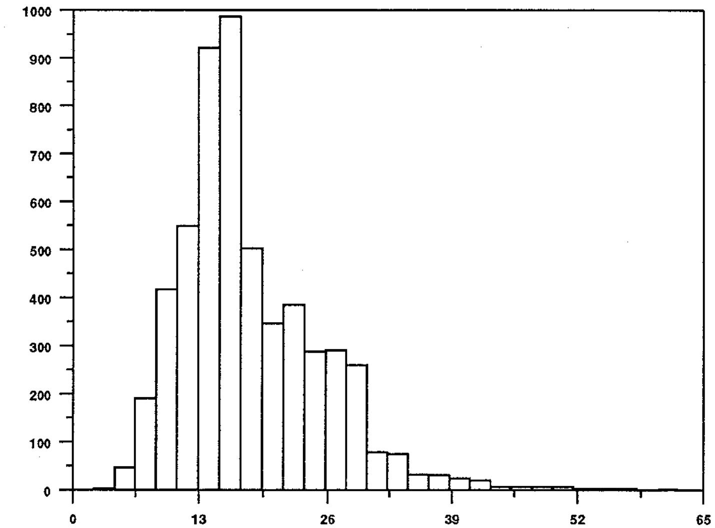
DULUTH.MN - 4 DAY INTERVAL



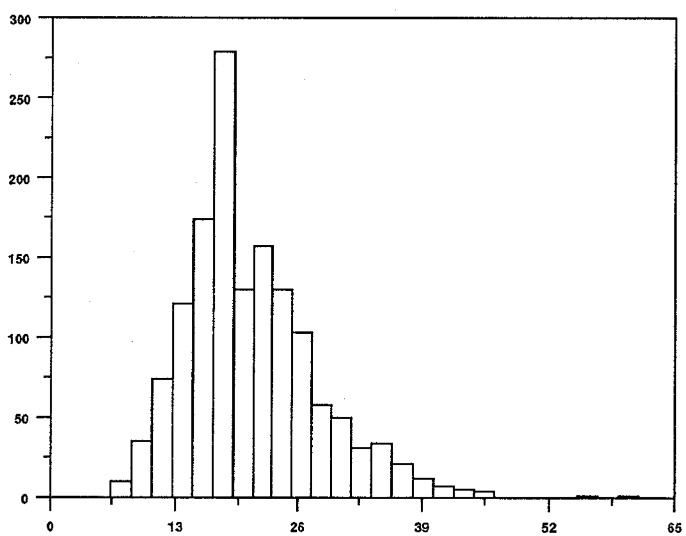
FORT_SMITH.AK - DAILY MAXIMA



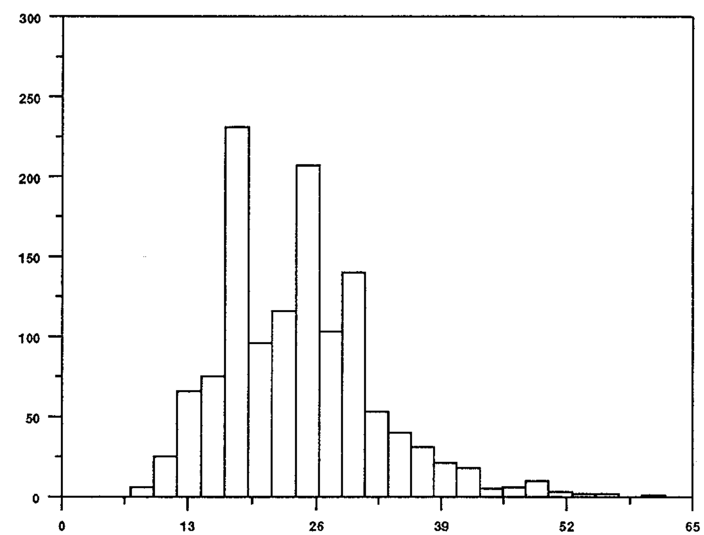
FORT_WAYNE_CORR.IN - DAILY MAXIMA



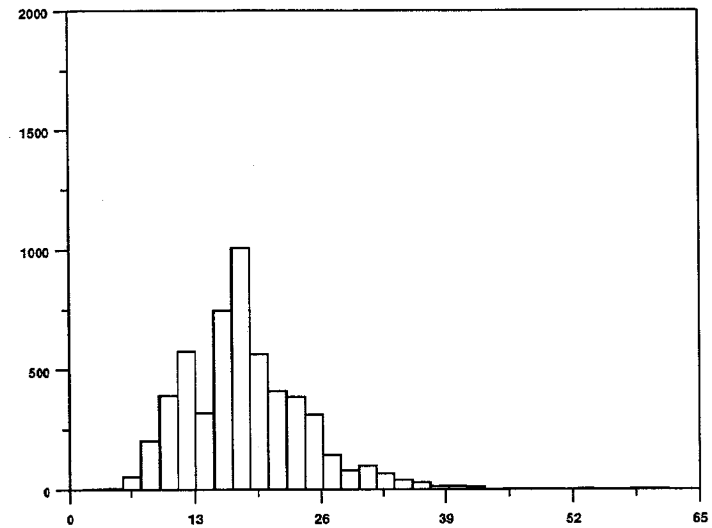
FORT_SMITH.AK - 4 DAY INTERVAL



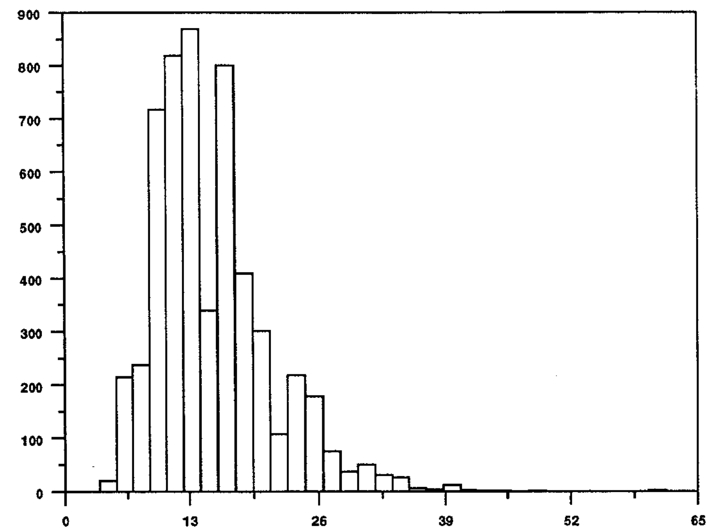
FORT_WAYNE_CORR.IN - 4 DAY INTERVAL



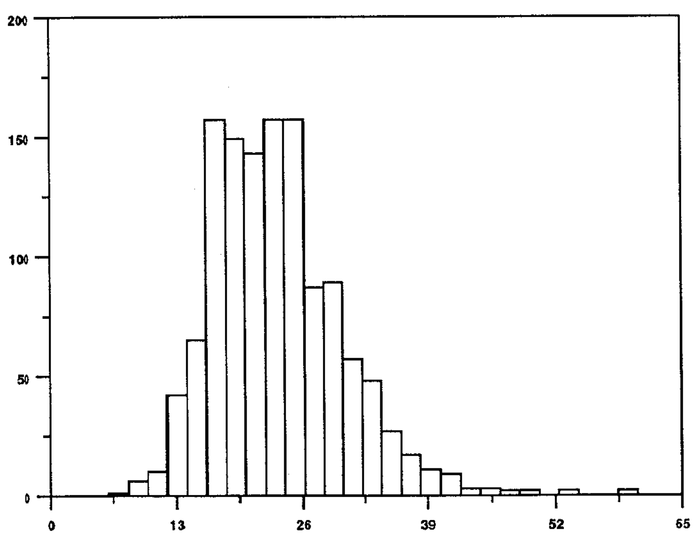
GREEN_BAY.WI - DAILY MAXIMA



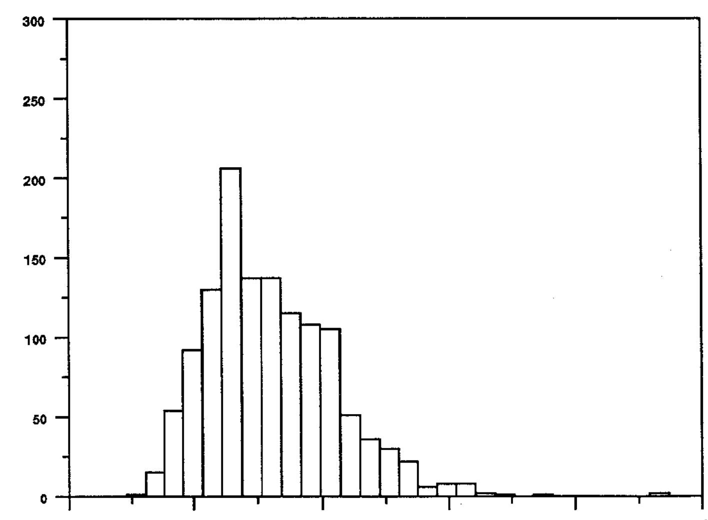
GREENSBORO.NC - DAILY MAXIMA



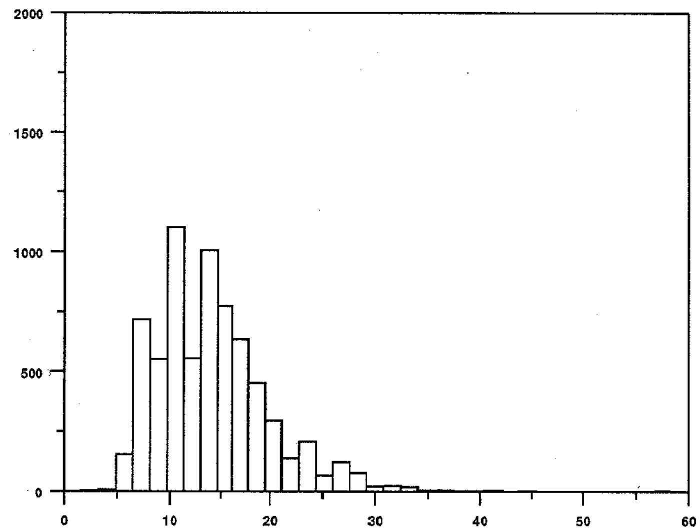
GREEN_BAY.WI - 4 DAY INTERVAL



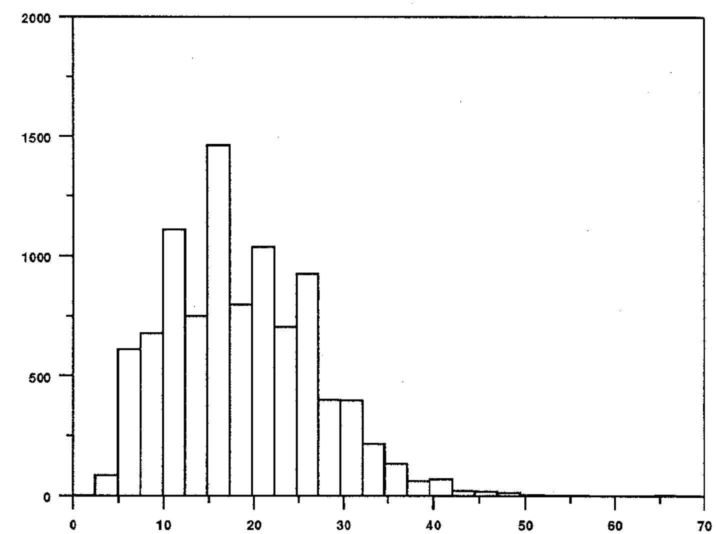
GREENSBORO.NC - 4 DAY INTERVAL



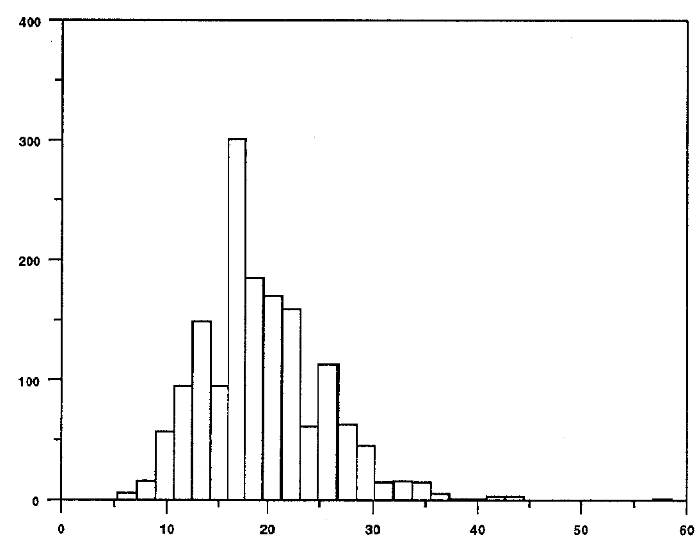
GREENVILLE_CORR.SC - DAILY MAXIMA



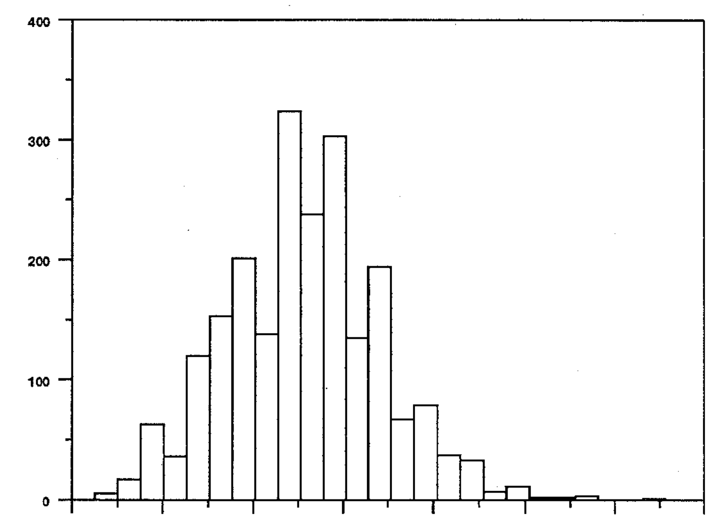
HELENA.MT - DAILY MAXIMA



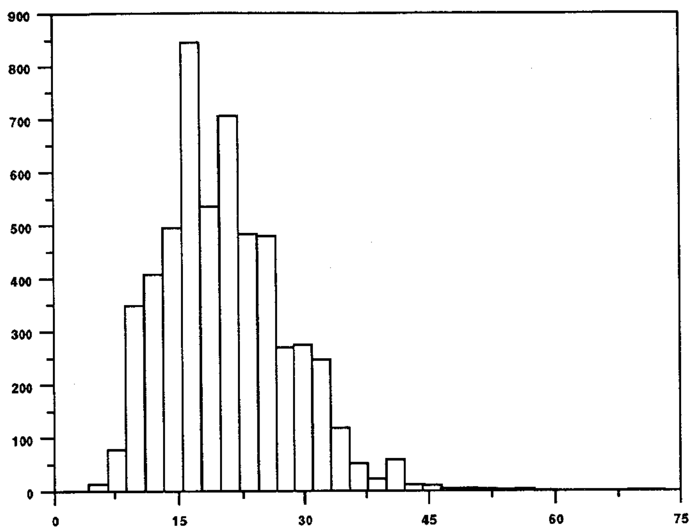
GREENVILLE_CORR.SC - 4 DAY INTERVAL



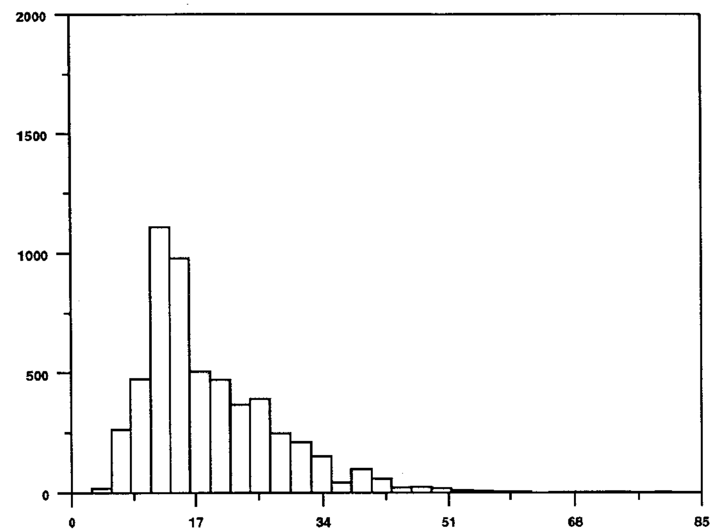
HELENA.MT - 4 DAY INTERVAL



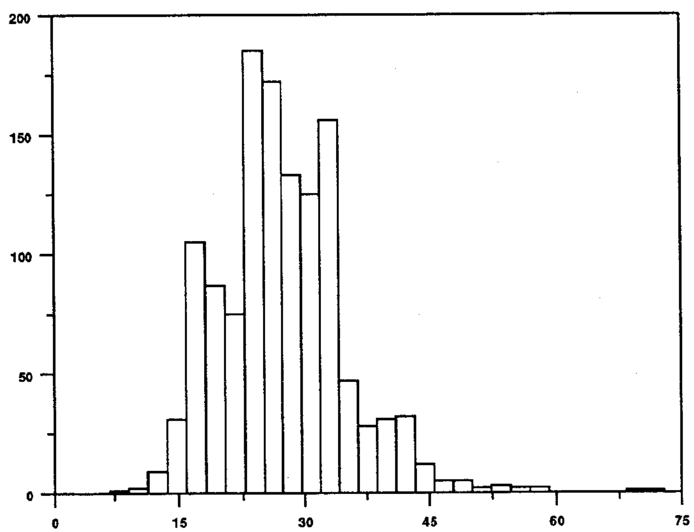
HURON.SD - DAILY MAXIMA



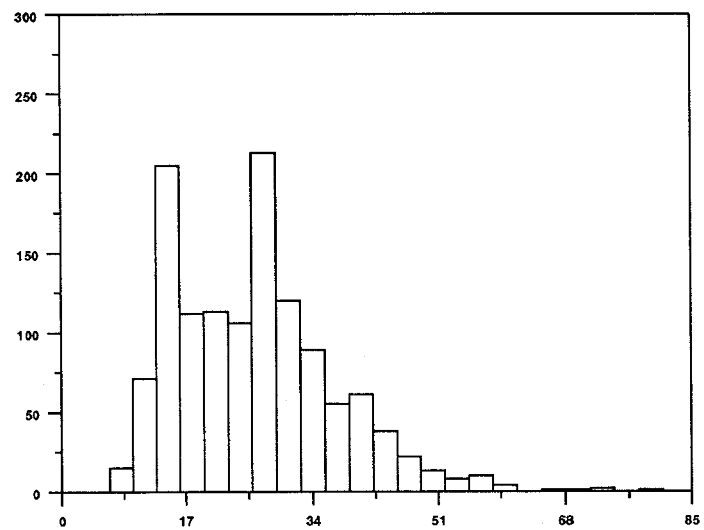
LANDER_CORR.WY - DAILY MAXIMA



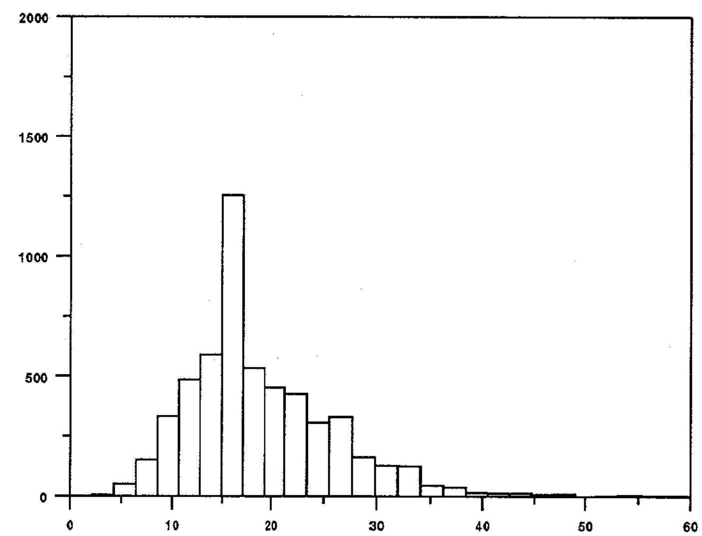
HURON.SD - 4 DAY INTERVAL



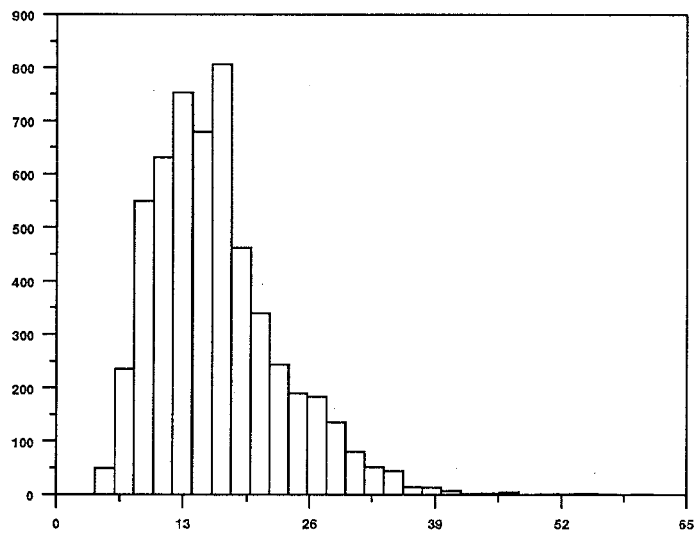
LANDER_CORR.WY - 4 DAY INTERVAL



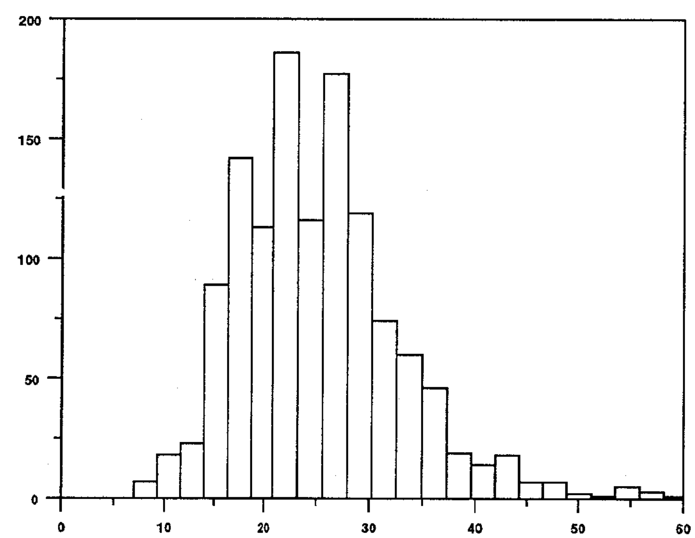
LANSING.MI - DAILY MAXIMA



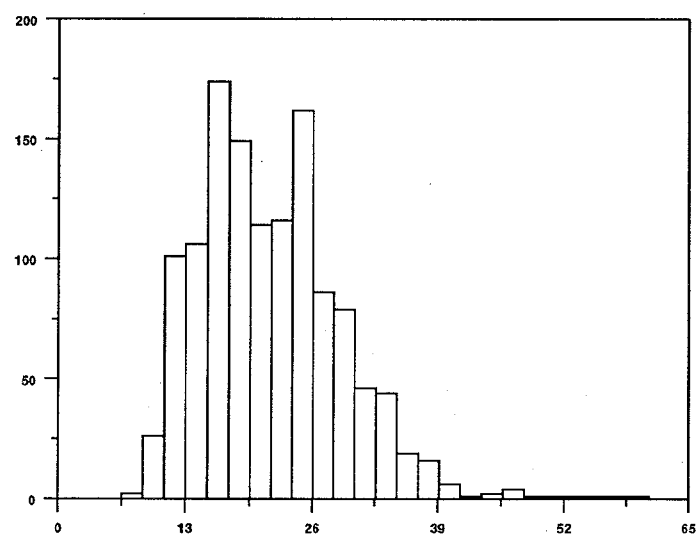
LOUISVILLE_CORR.KY - DAILY MAXIMA



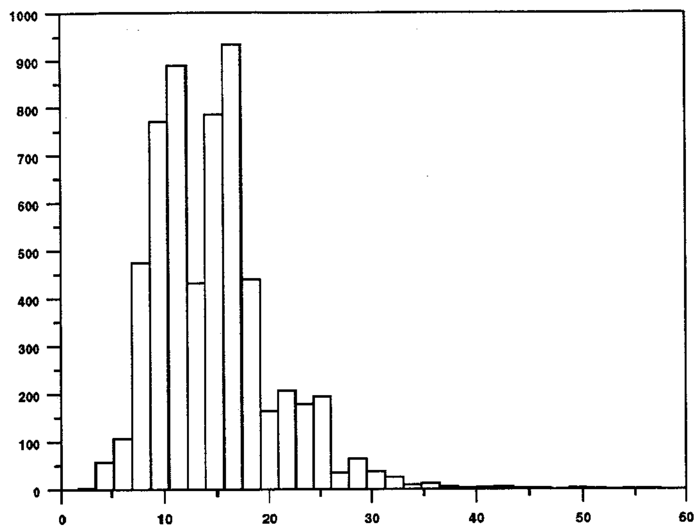
LANSING.MI - 4 DAY INTERVAL



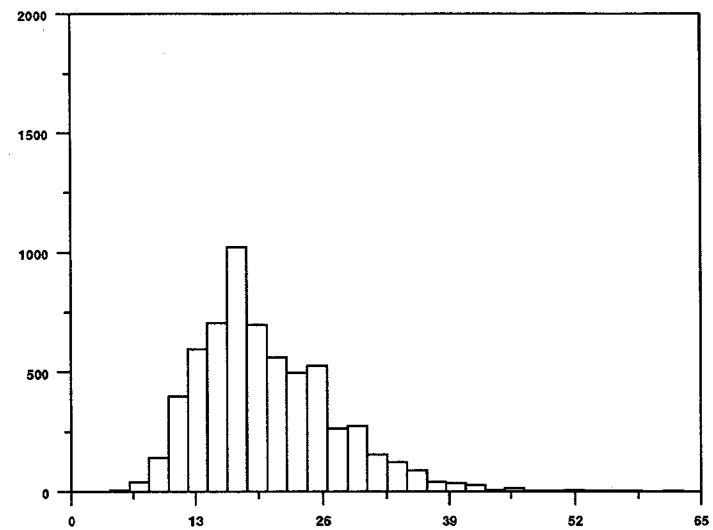
LOUISVILLE_CORR.KY - 4 DAY INTERVAL



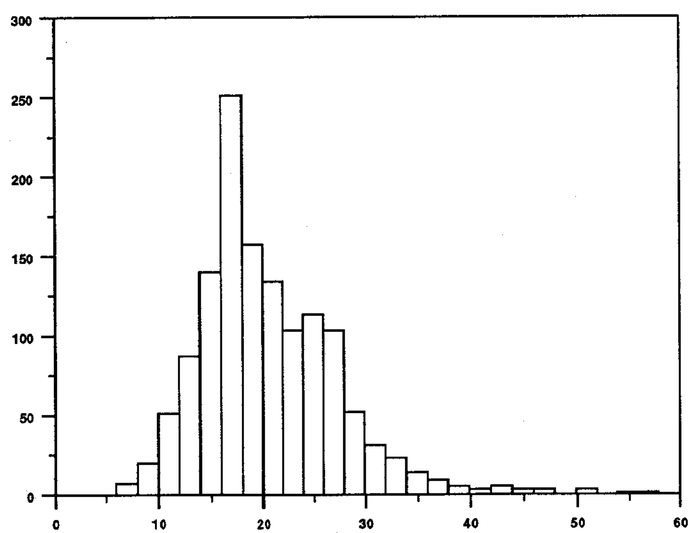
MACON.GA - DAILY MAXIMA



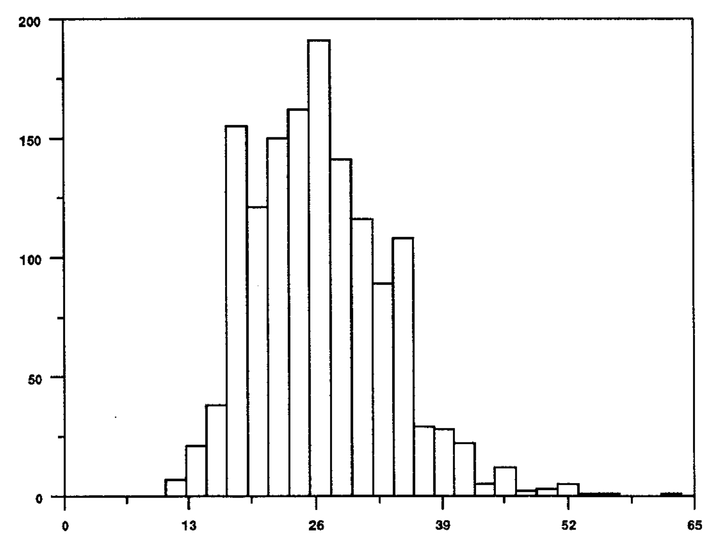
MILWAUKEE_CORR.WI - DAILY MAXIMA



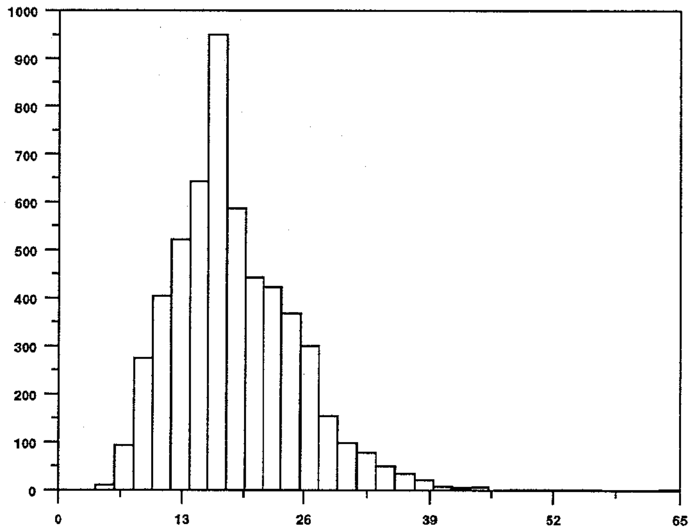
MACON.GA - 4 DAY INTERVAL



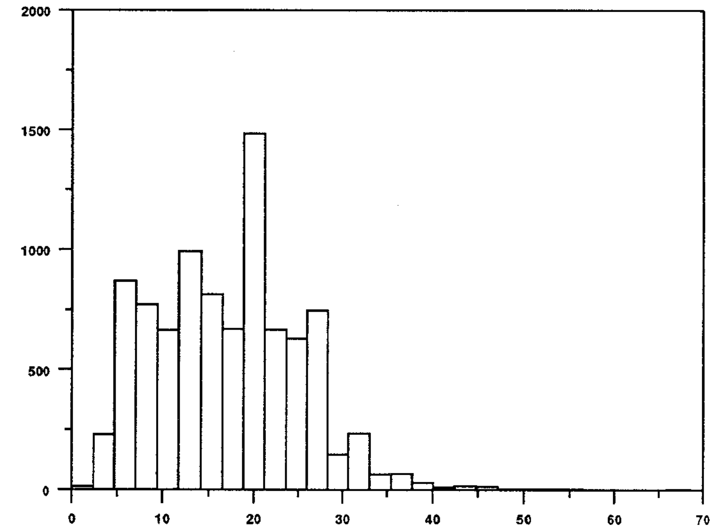
MILWAUKEE_CORR.WI - 4 DAY INTERVAL



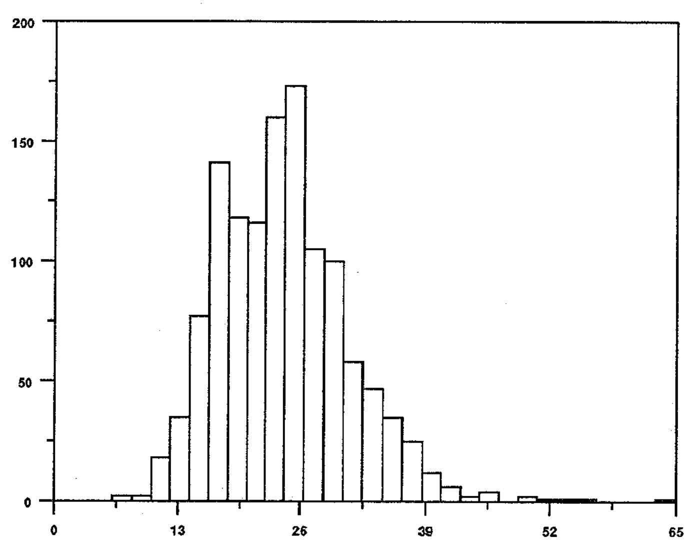
MINNEAPOLIS_CORR.MN - DAILY MAXIMA



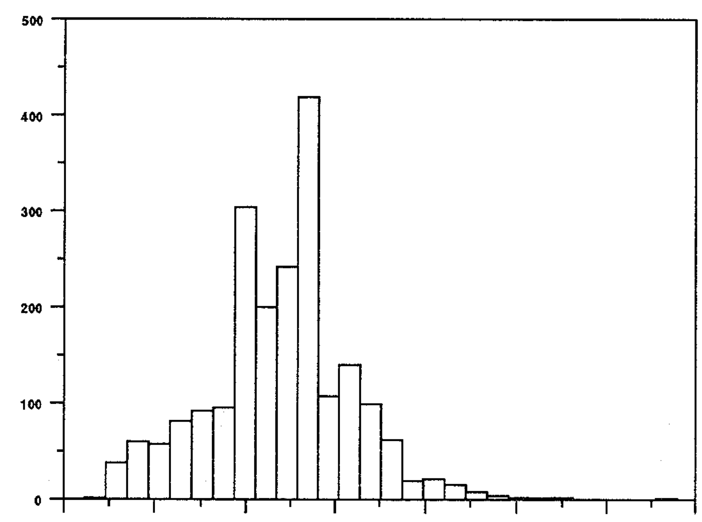
MISSOULA_CORR.MT - DAILY MAXIMA



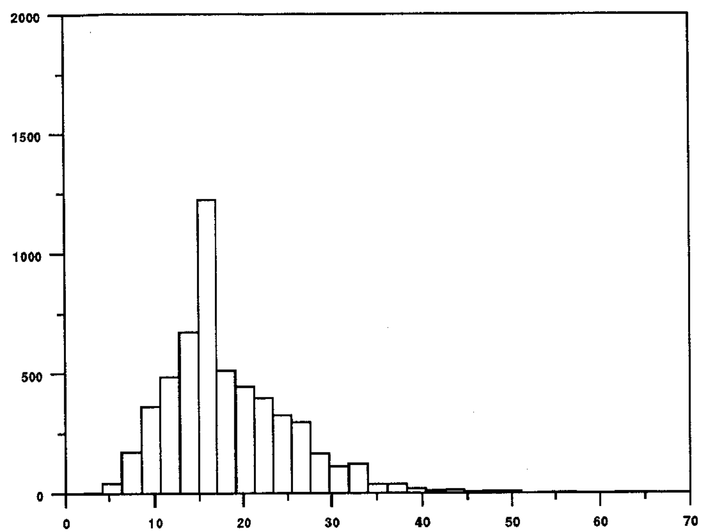
MINNEAPOLIS_CORR.MN - 4 DAY INTERVAL



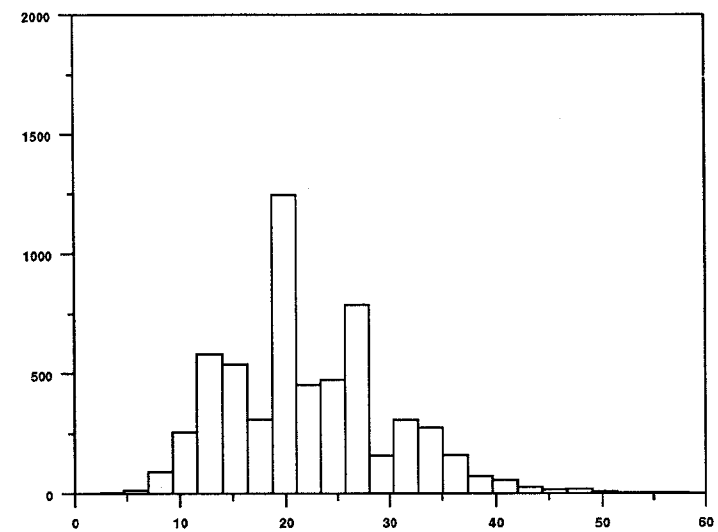
MISSOULA_CORR.MT - 4 DAY INTERVAL



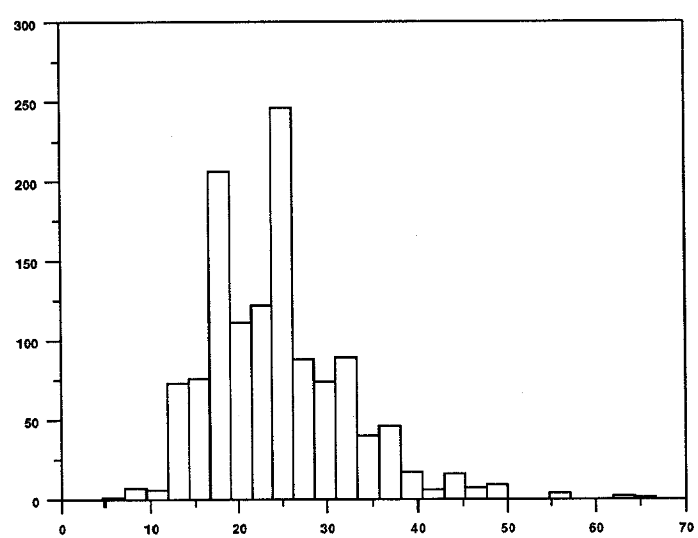
MOLINE.IL - DAILY MAXIMA



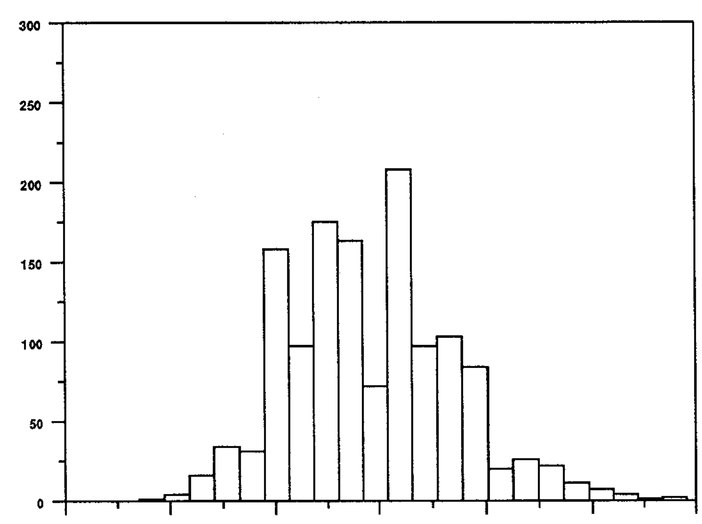
OKLAHOMA_CITY.OK - DAILY MAXIMA



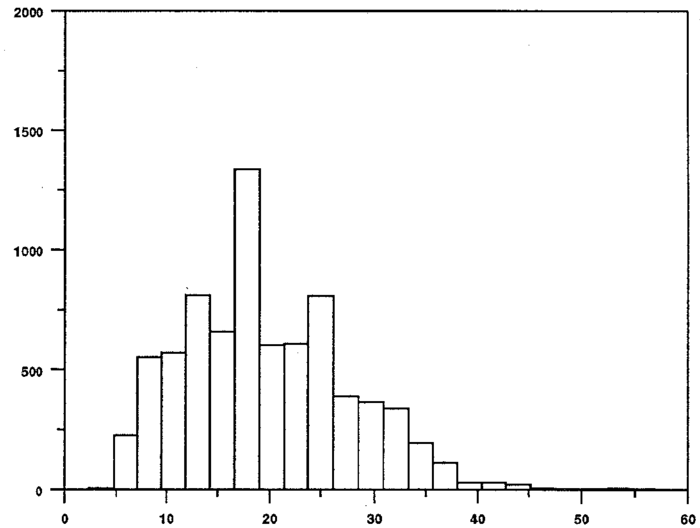
MOLINE.IL - 4 DAY INTERVAL



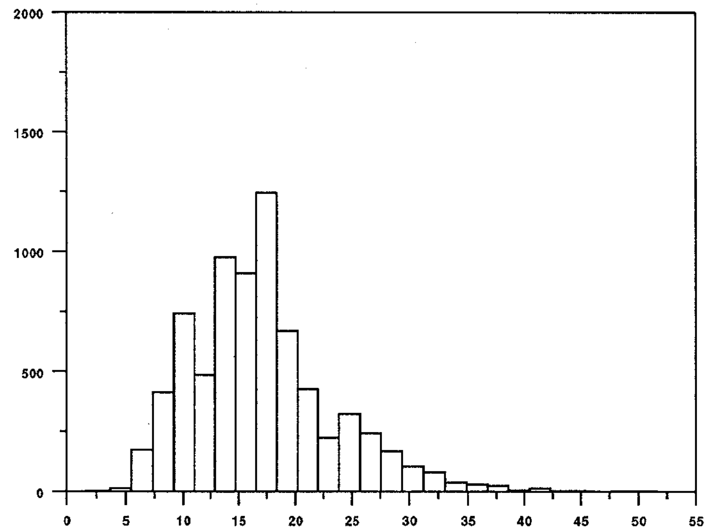
OKLAHOMA_CITY.OK - 4 DAY INTERVAL



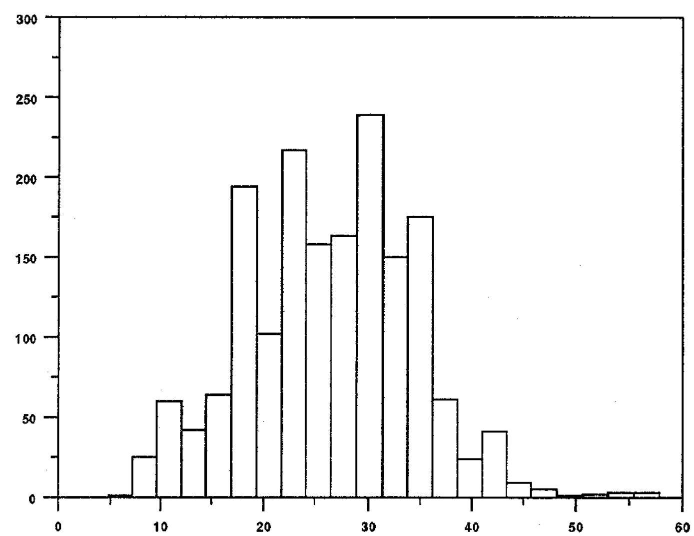
POCATELLO.ID - DAILY MAXIMA



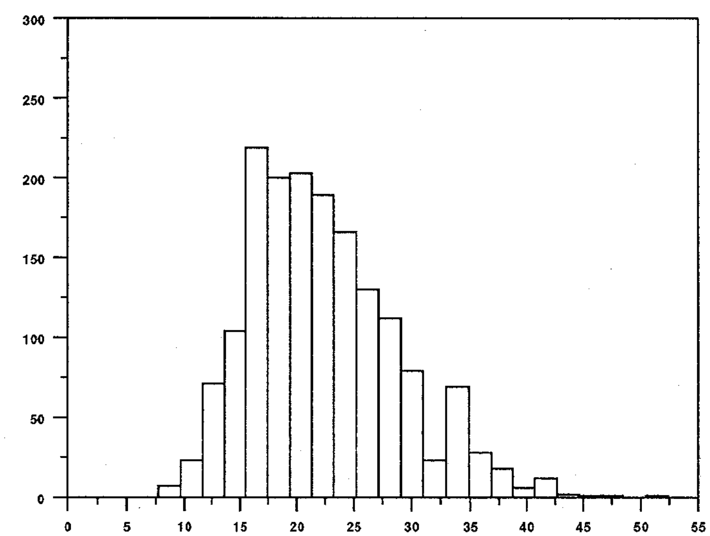
PORTLAND_CORR.ME - DAILY MAXIMA



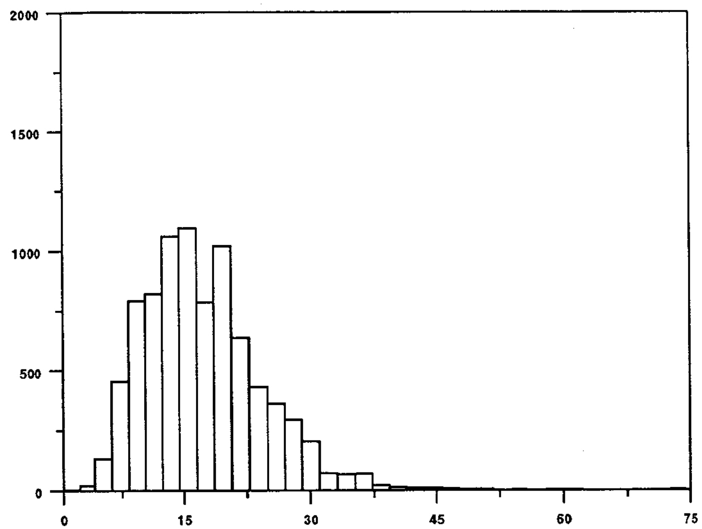
POCATELLO.ID - 4 DAY INTERVAL



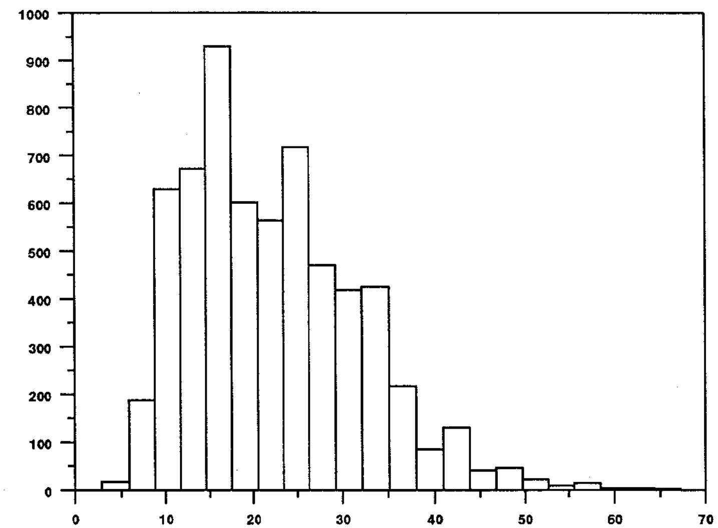
PORTLAND_CORR.ME - 4 DAY INTERVAL



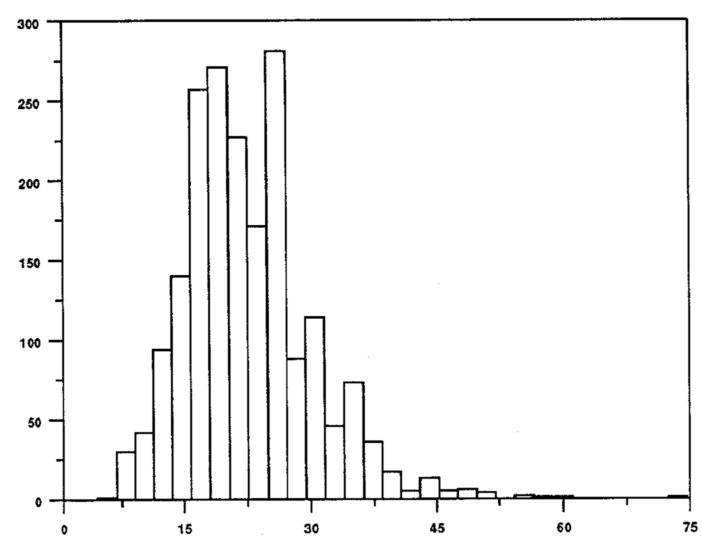
PORTLAND_CORR.OR - DAILY MAXIMA



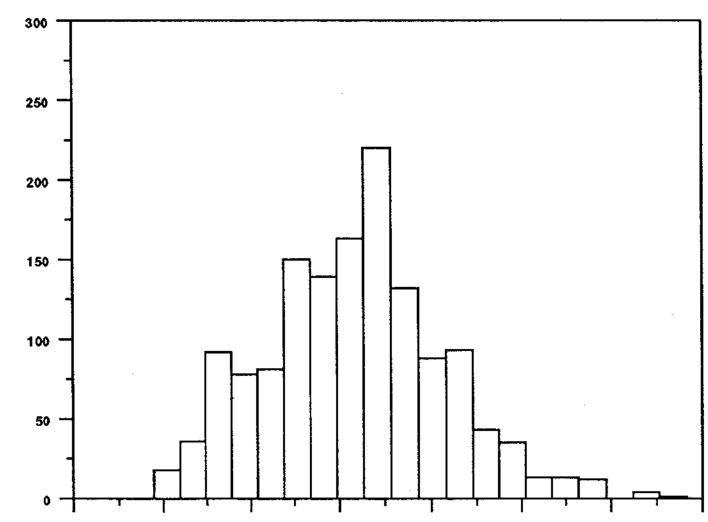
PUEBLO_CORR.CO - DAILY MAXIMA



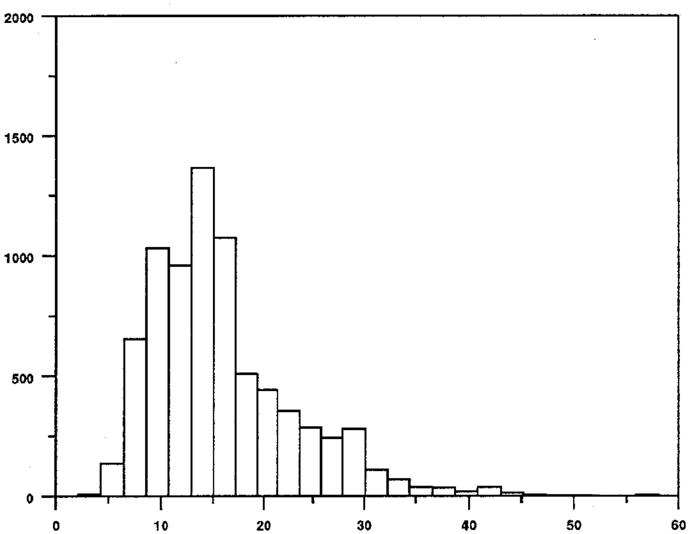
PORTLAND_CORR.OR - 4 DAY INTERVAL



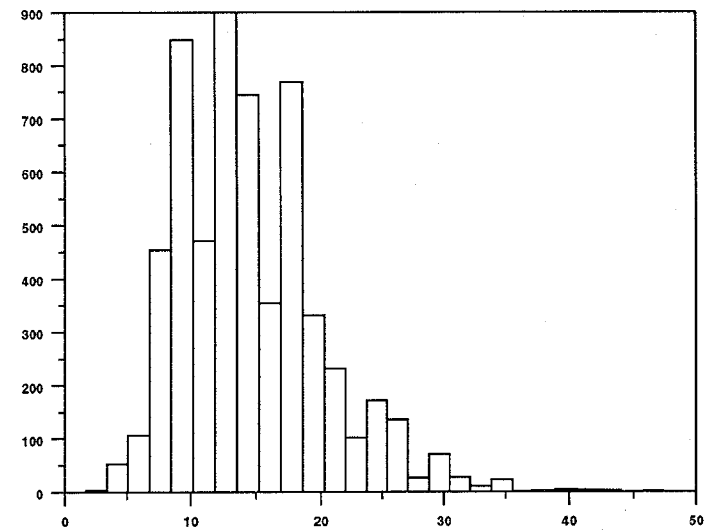
PUEBLO_CORR.CO - 4 DAY INTERVAL



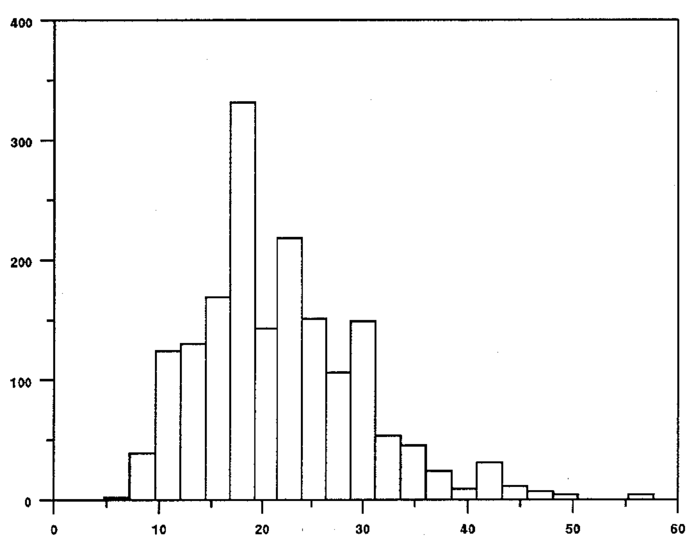
RED_BLUFF.CA - DAILY MAXIMA



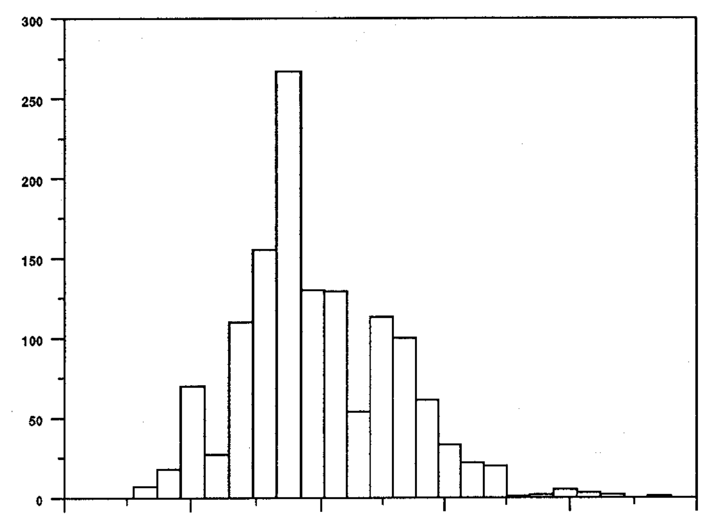
RICHMOND_CORR.VA - DAILY MAXIMA



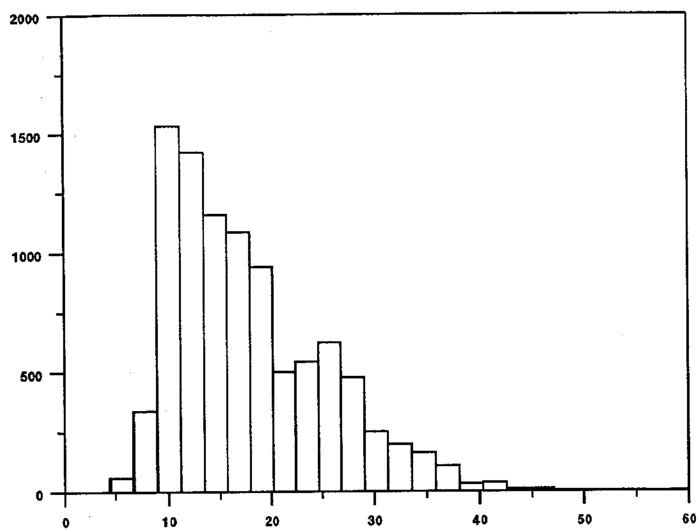
RED_BLUFF.CA - 4 DAY INTERVAL



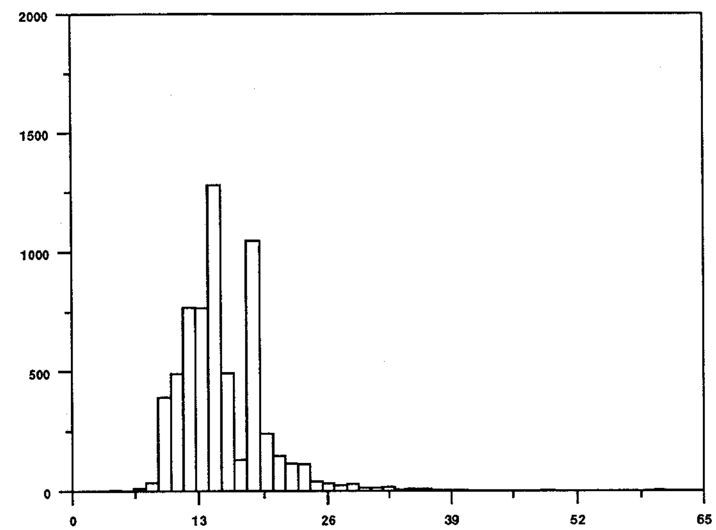
RICHMOND_CORR.VA - 4 DAY INTERVAL



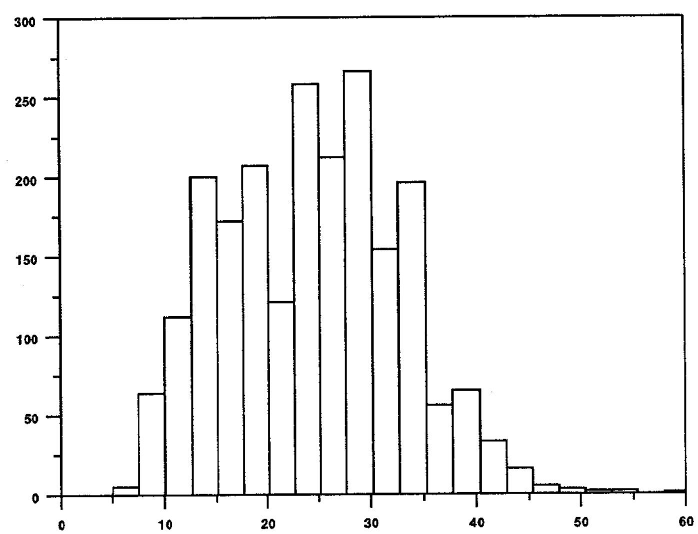
SALT_LAKE_CITY.UT - DAILY MAXIMA



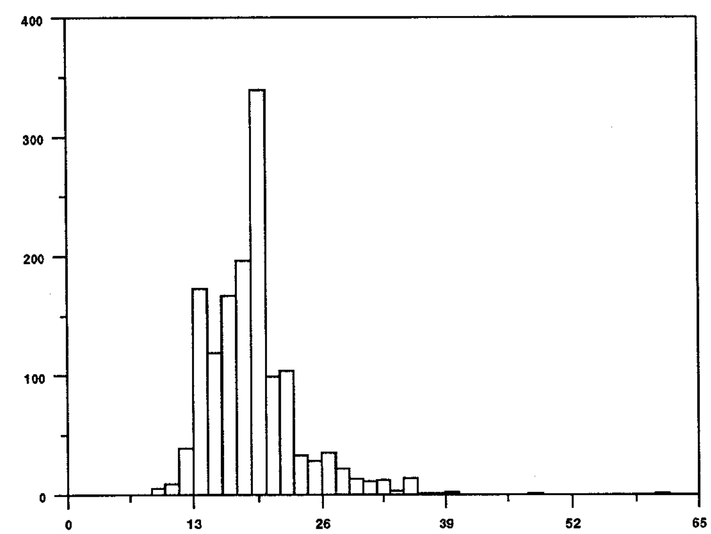
SAN_DIEGO.CA - DAILY MAXIMA



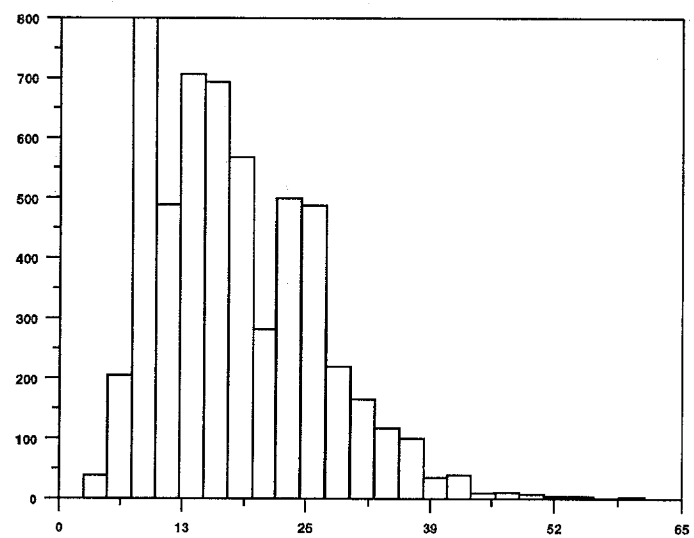
SALT_LAKE_CITY.UT - 4 DAY INTERVAL



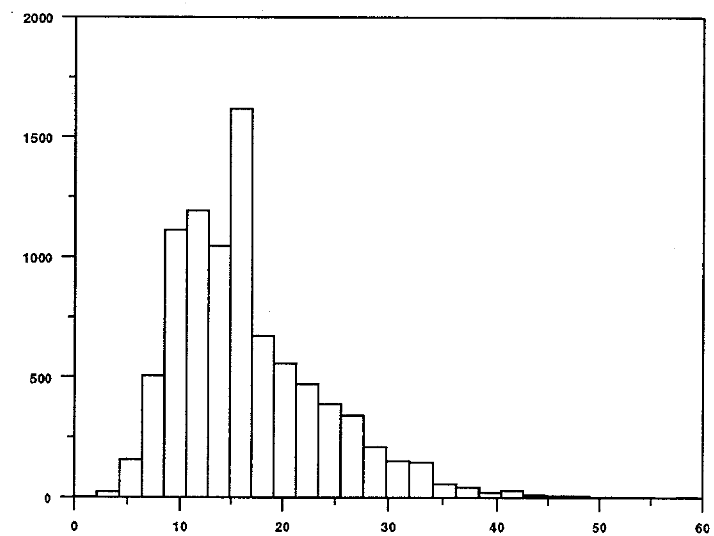
SAN_DIEGO.CA - 4 DAY INTERVAL



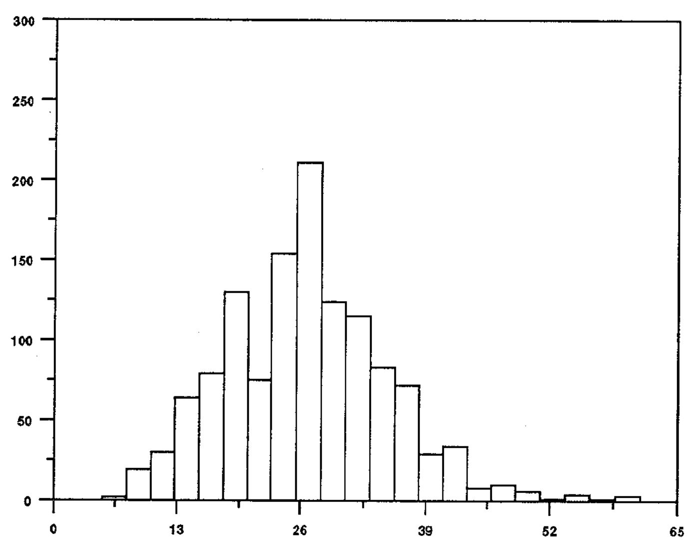
SHERIDAN_CORR.WY - DAILY MAXIMA



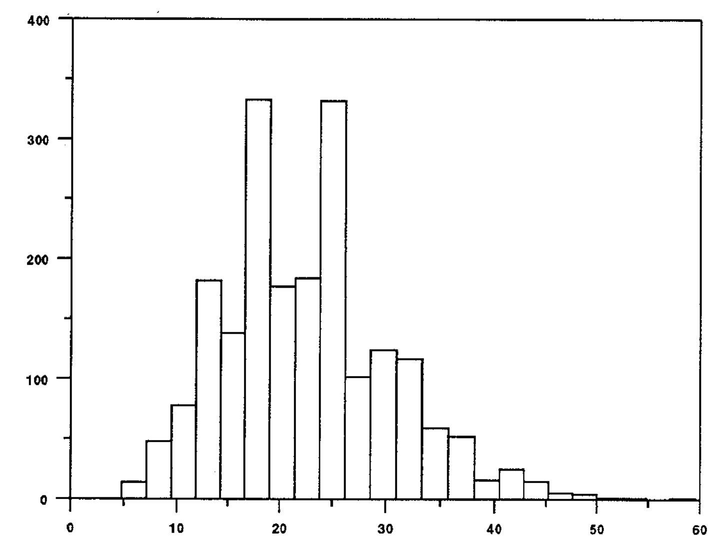
SPOKANE.WA - DAILY MAXIMA



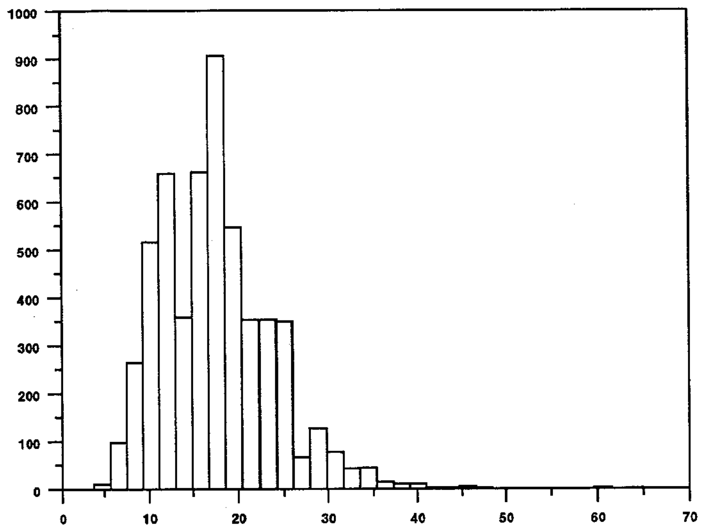
SHERIDAN_CORR.WY - 4 DAY INTERVAL



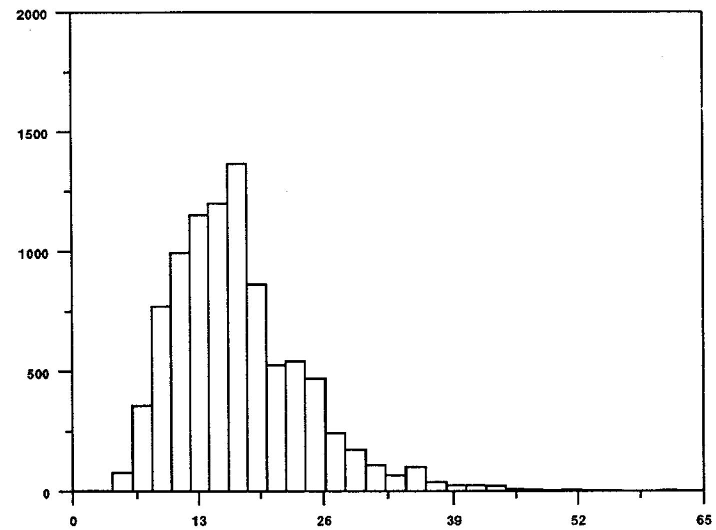
SPOKANE.WA - 4 DAY INTERVAL



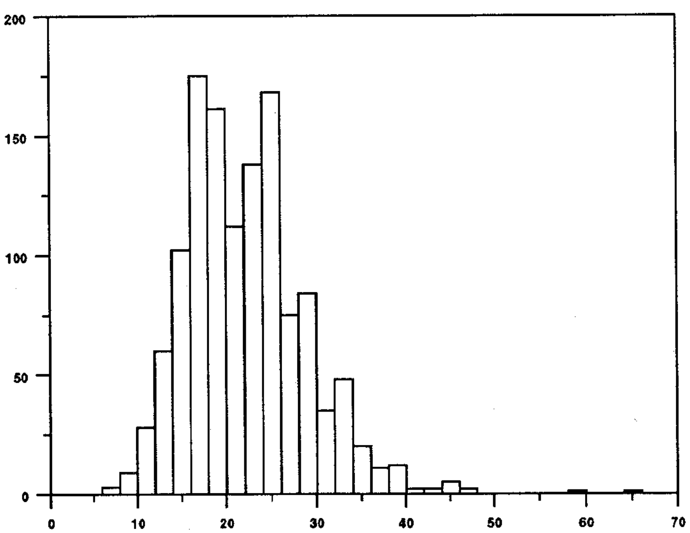
SPRINGFIELD.MO - DAILY MAXIMA



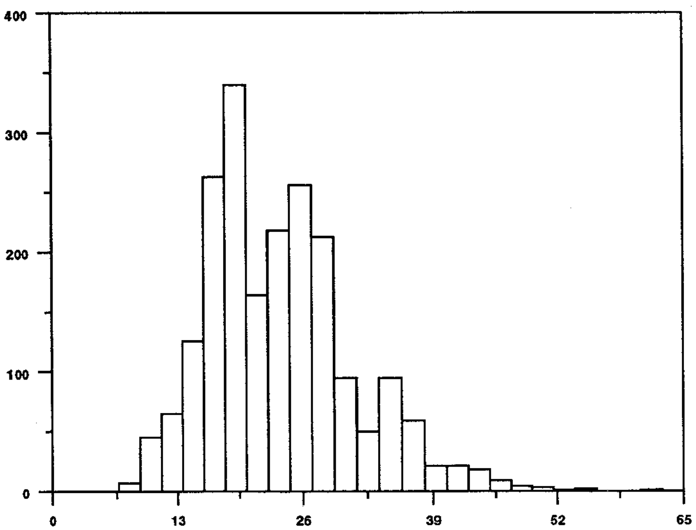
TOLEDO.OH - DAILY MAXIMA



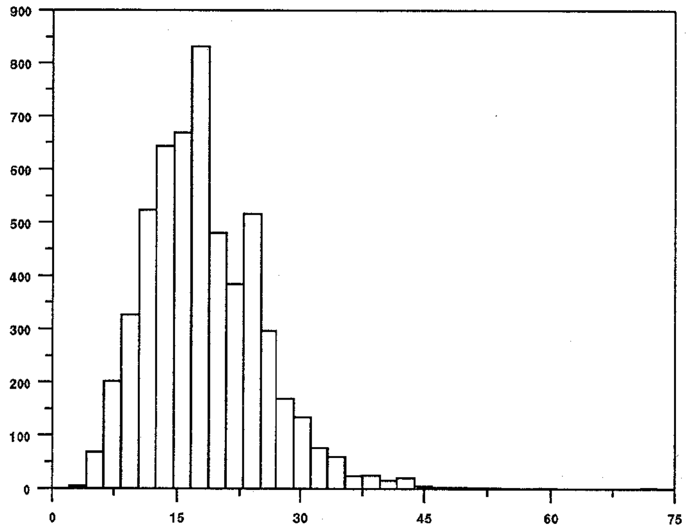
SPRINGFIELD.MO - 4 DAY INTERVAL



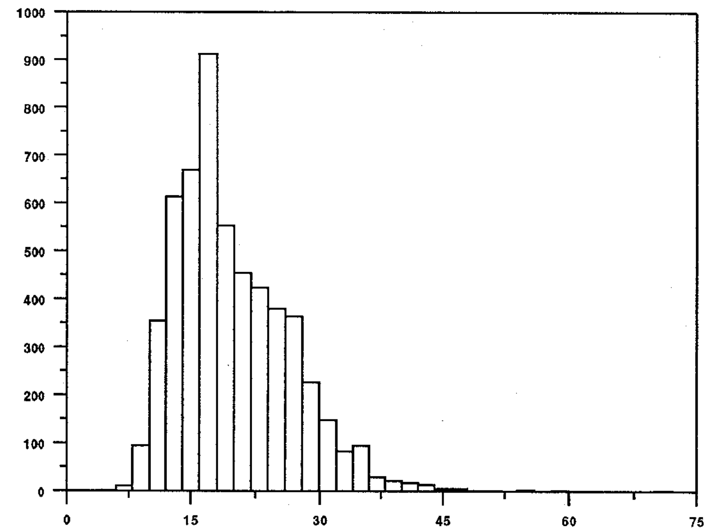
TOLEDO.OH - 4 DAY INTERVAL



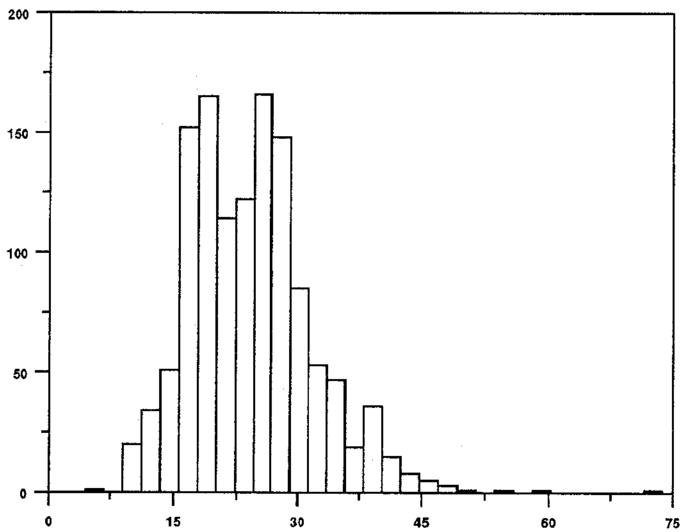
TOPEKA.KS - DAILY MAXIMA



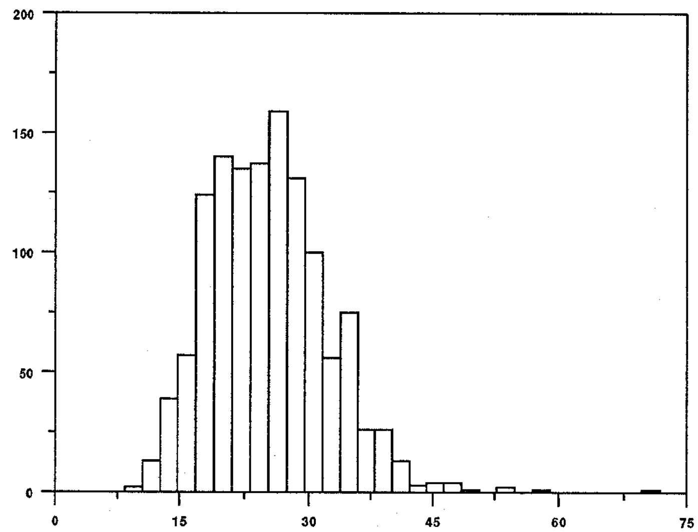
TUCSON.AZ - DAILY MAXIMA



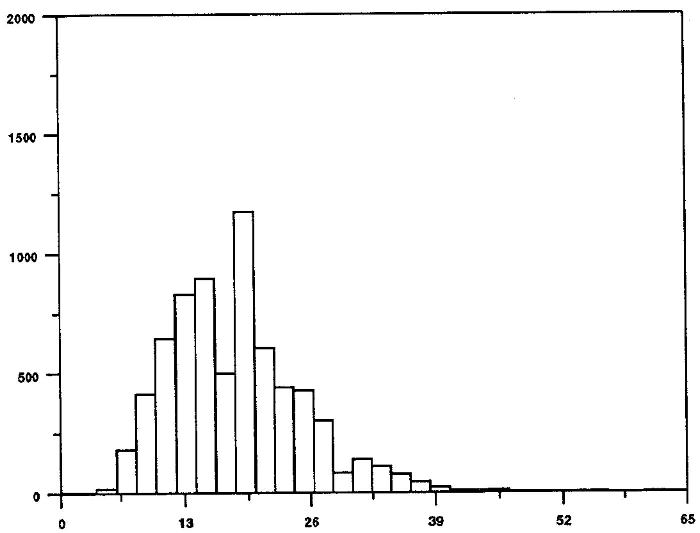
TOPEKA.KS - 4 DAY INTERVAL



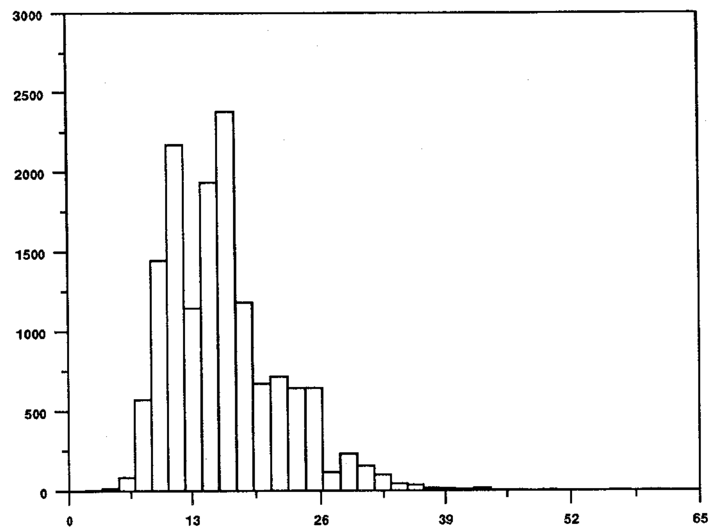
TUCSON.AZ - 4 DAY INTERVAL



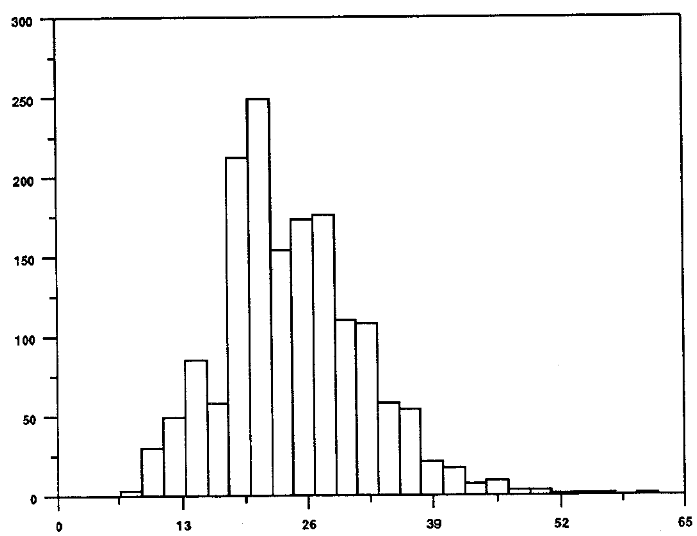
WINNEMUCCA,NV - DAILY MAXIMA



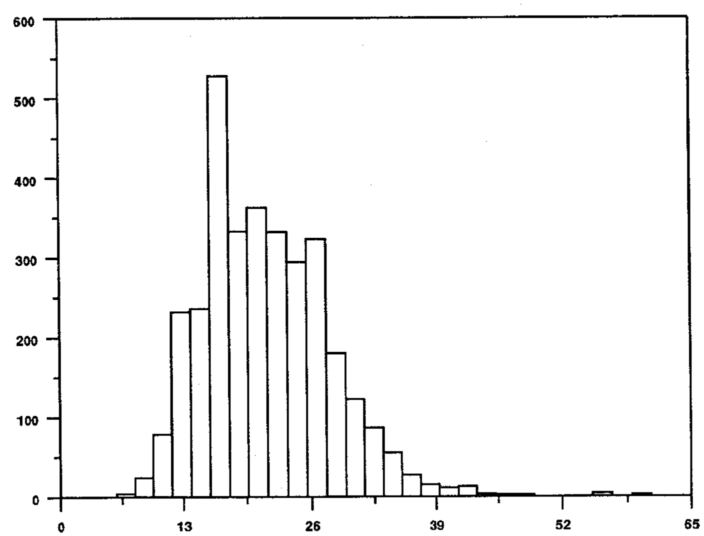
YUMA,AZ - DAILY MAXIMA



WINNEMUCCA,NV - 4 DAY INTERVAL

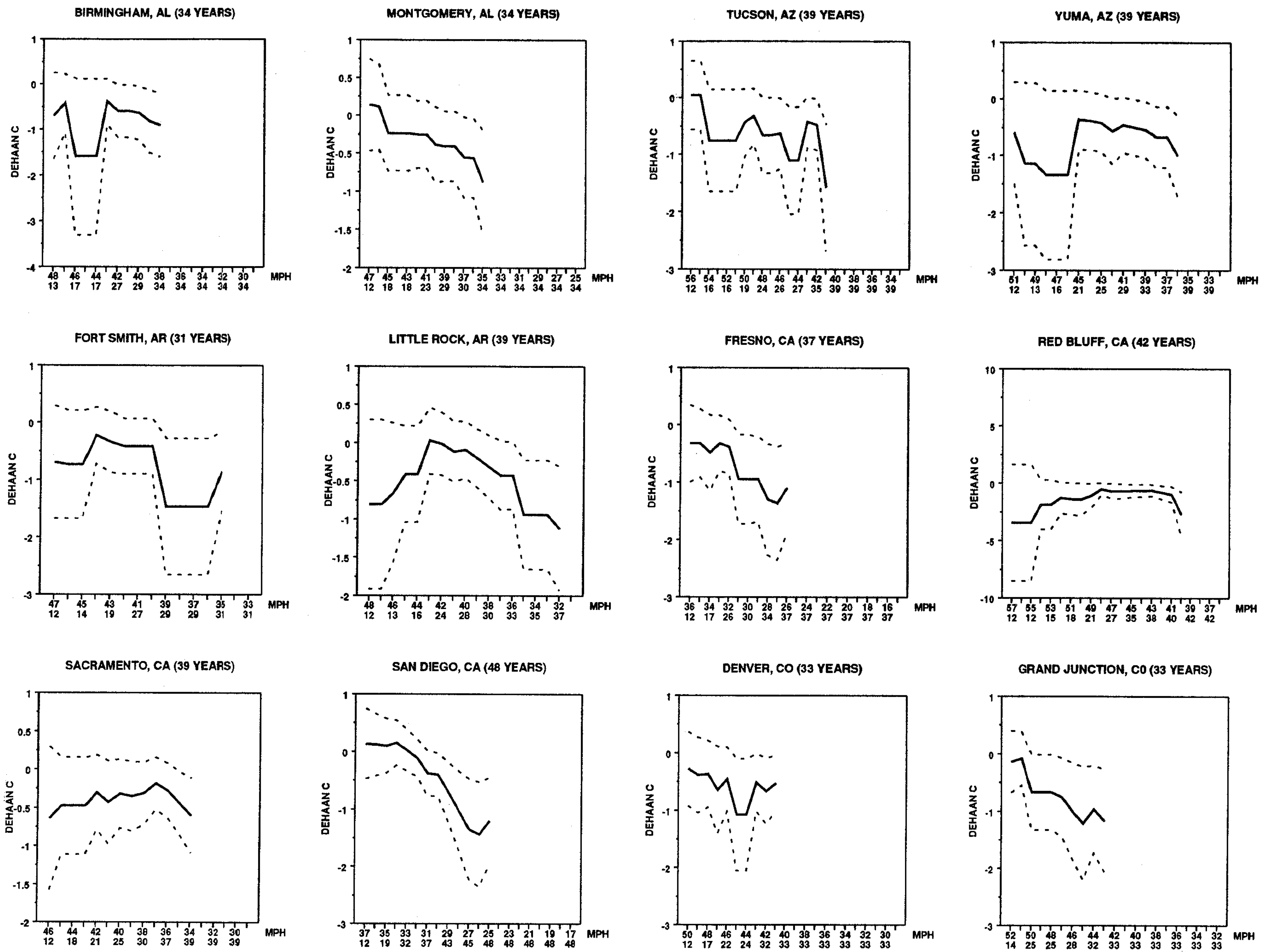


YUMA,AZ - 4 DAY INTERVAL

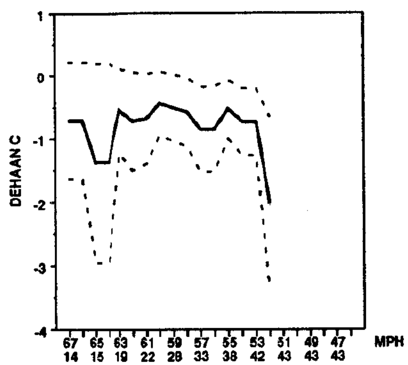


APPENDIX 2

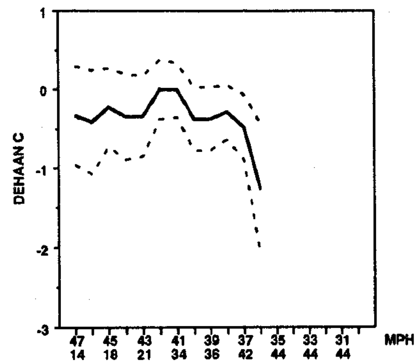
Plots of parameter $\hat{c}$ and 95% confidence bounds versus threshold and number of threshold exceedances (based on samples of largest yearly wind speeds)


BIRMINGHAM, AL (34 YEARS)
MONTGOMERY, AL (34 YEARS)
TUCSON, AZ (39 YEARS)
YUMA, AZ (39 YEARS)
FORT SMITH, AR (31 YEARS)
LITTLE ROCK, AR (39 YEARS)
FRESNO, CA (37 YEARS)
RED BLUFF, CA (42 YEARS)
SACRAMENTO, CA (39 YEARS)
SAN DIEGO, CA (48 YEARS)
DENVER, CO (33 YEARS)
GRAND JUNCTION, C0 (33 YEARS)
DEHAAN C
MPH

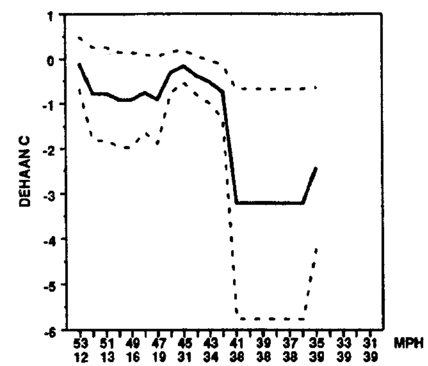
PUEBLO, CO (43 YEARS)



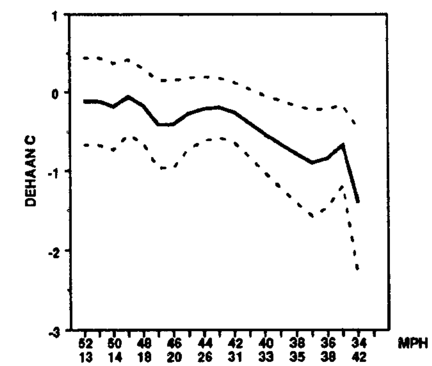
HARTFORD, CT (44 YEARS)



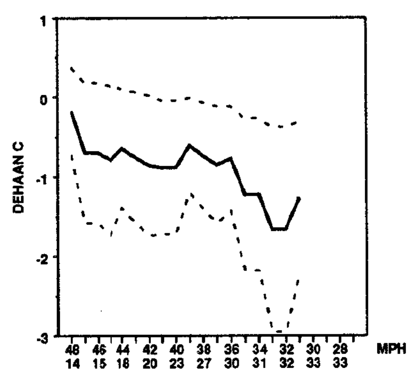
WASHINGTON, DC (39 YEARS)



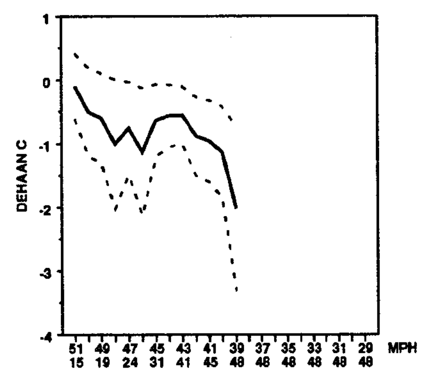
ATLANTA, GA (42 YEARS)



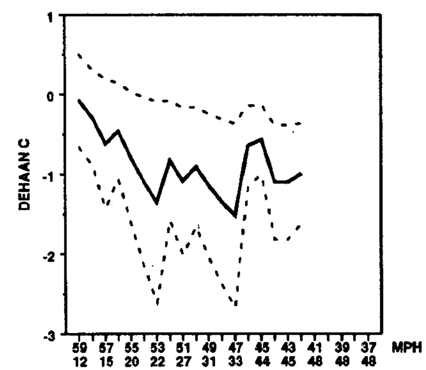
MACON, GA (33 YEARS)



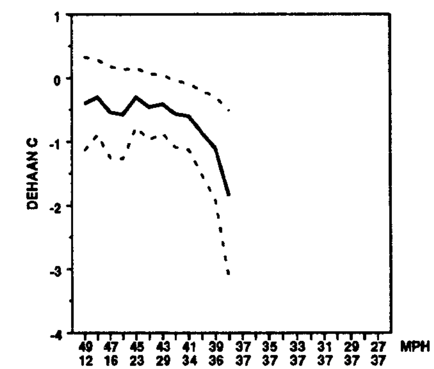
BOISE, ID (48 YEARS)



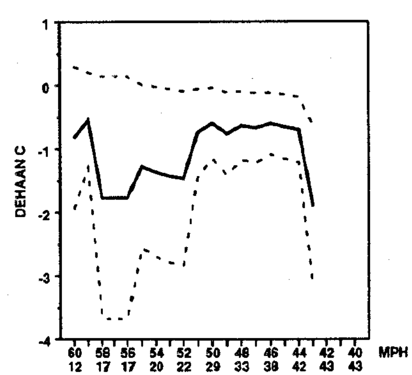
POCATELLO, ID (48 YEARS)



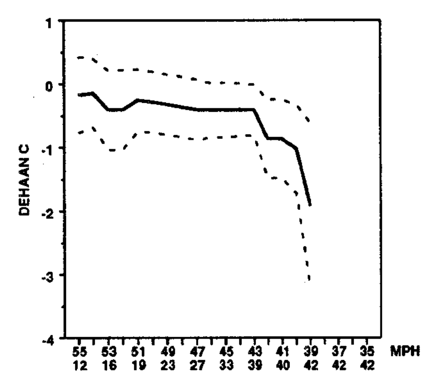
CHICAGO MIDWAY, IL (37 YEARS)



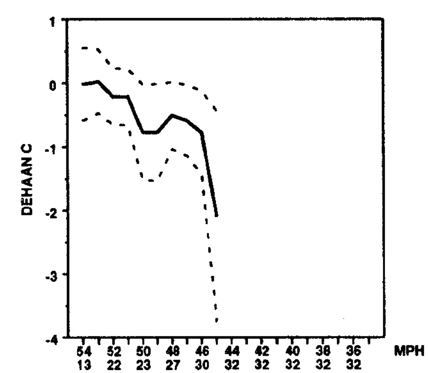
MOLINE, IL (44 YEARS)



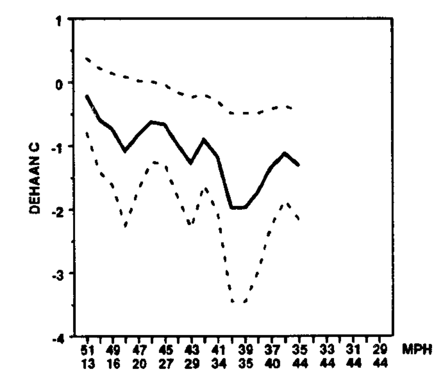
PEORIA, IL (42 YEARS)



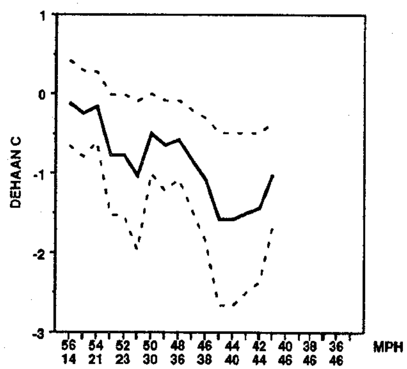
SPRINGFIELD, IL (32 YEARS)



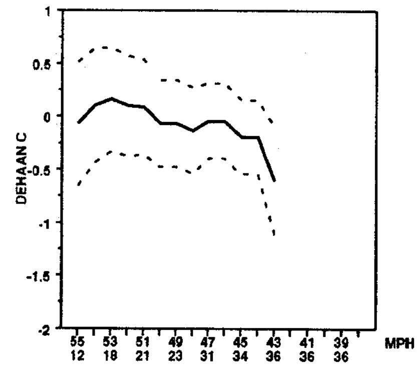
EVANSVILLE, IN (44 YEARS)



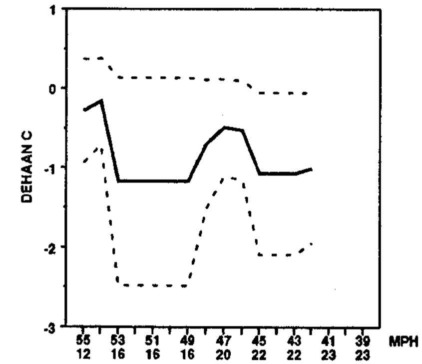
FORT WAYNE, IN (46 YEARS)



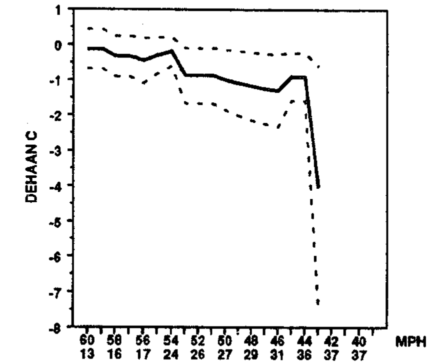
INDIANAPOLIS, IN (36 YEARS)



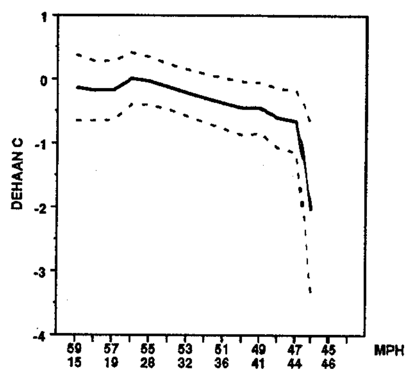
BURLINGTON, IA (23 YEARS)



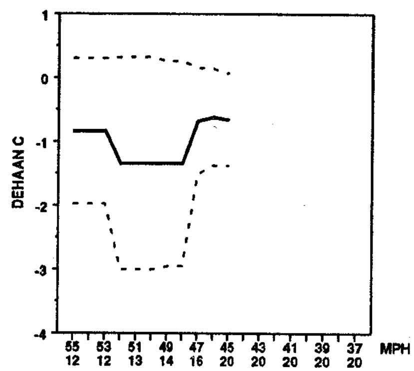
DES MOINES, IA (37 YEARS)



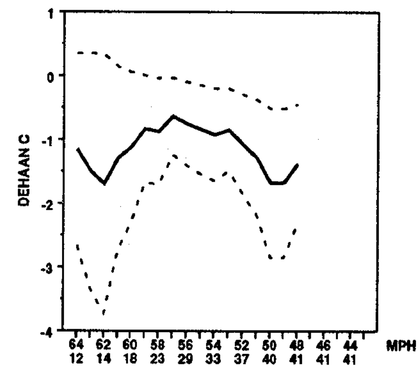
SIOUX CITY, IA (46 YEARS)



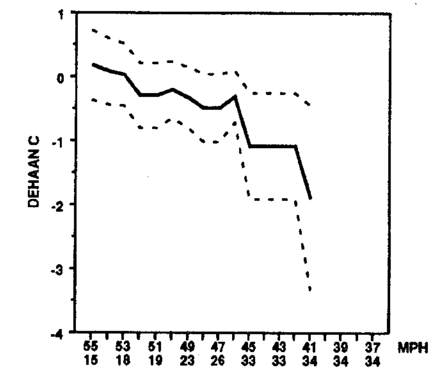
CONCORDIA, KS (20 YEARS)



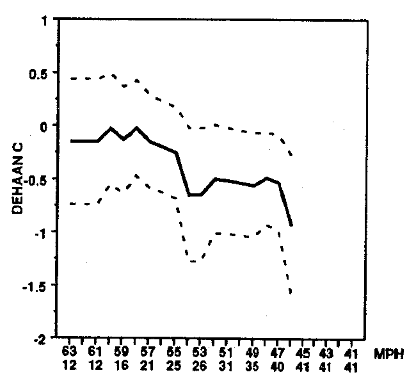
DODGE CITY, KS (41 YEARS)



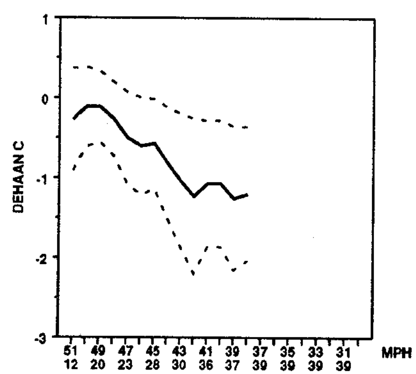
TOPEKA, KS (34 YEARS)



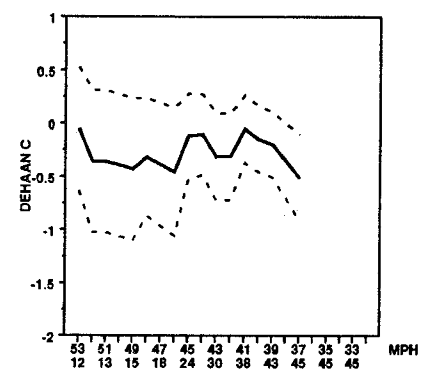
WICHITA, KS (41 YEARS)



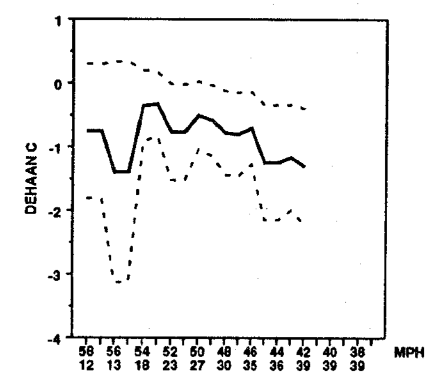
LOUISVILLE, KY (39 YEARS)



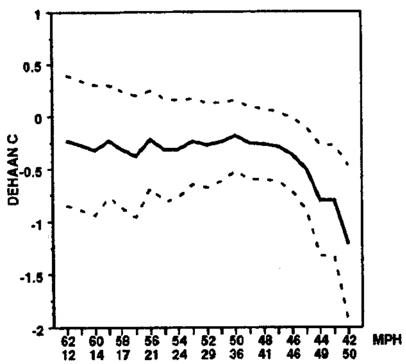
PORTLAND, ME (45 YEARS)



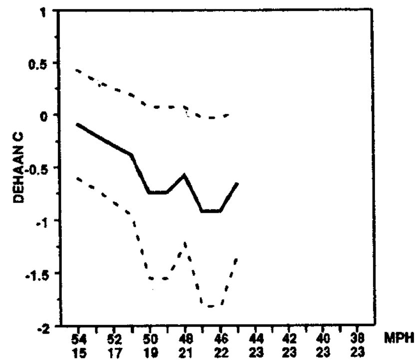
BALTIMORE, MD (39 YEARS)



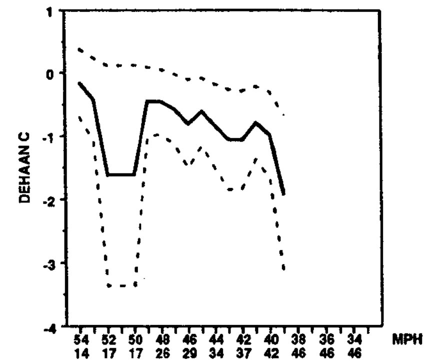
BOSTON, MA (50 YEARS)



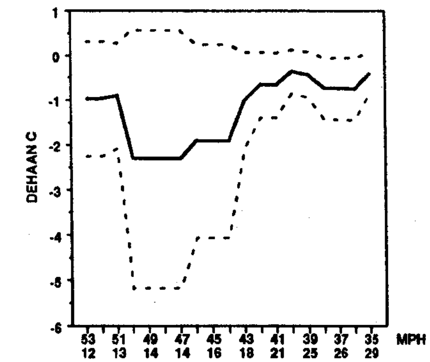
NANTUCKET, MA (23 YEARS)



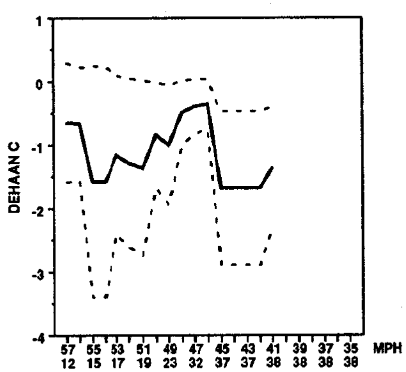
DETROIT, MI (46 YEARS)



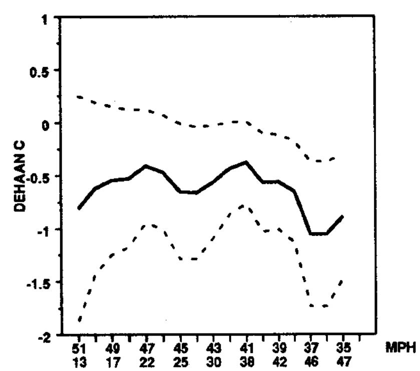
GRAND RAPIDS, MI (29 YEARS)



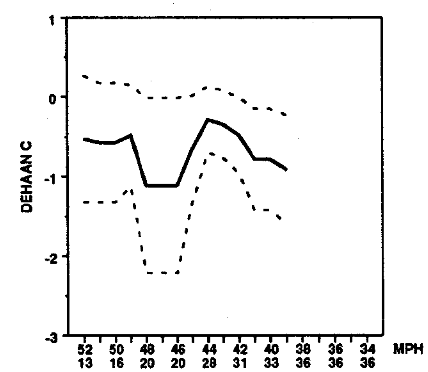
LANSING, MI (38 YEARS)



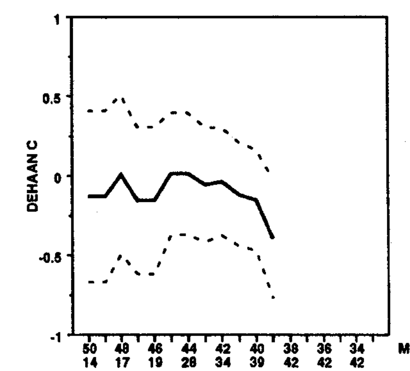
SAULT STE MARIE, MI (47 YEARS)



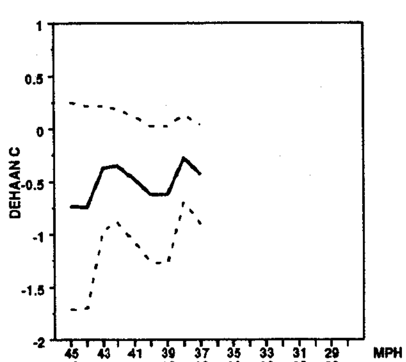
DULUTH, MN (36 YEARS)



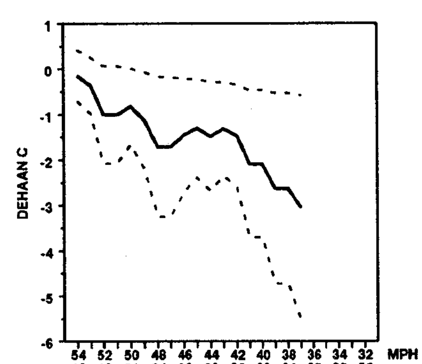
MINNEAPOLIS, MN (42 YEARS)



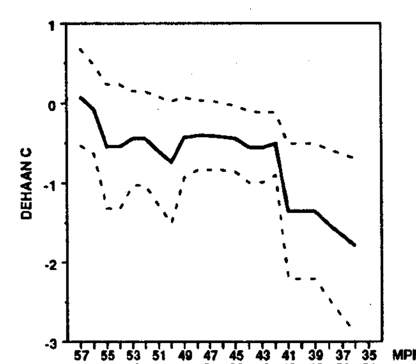
JACKSON, MS (29 YEARS)



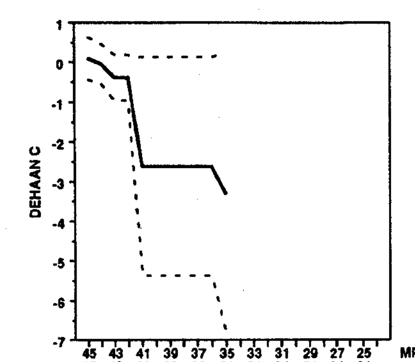
COLUMBIA, MO (35 YEARS)



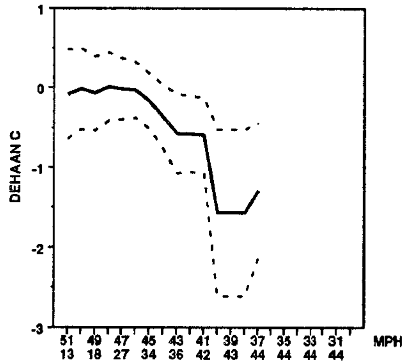
KANSAS CITY, MO (51 YEARS)



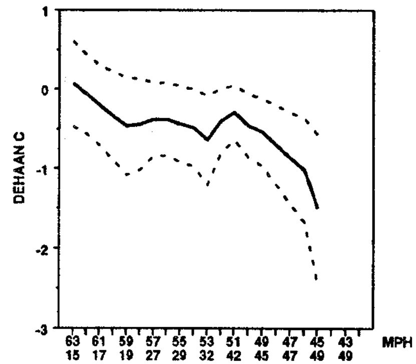
ST.LOUIS, MO (21 YEARS)



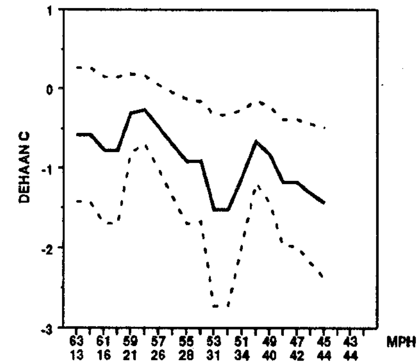
SPRINGFIELD, MO (44 YEARS)



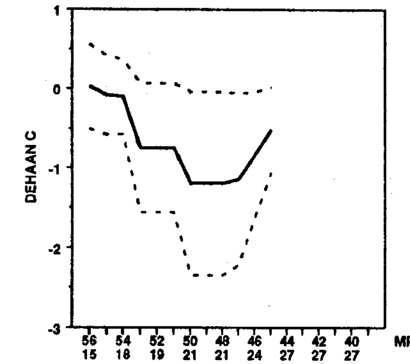
BILLINGS, MT (49 YEARS)



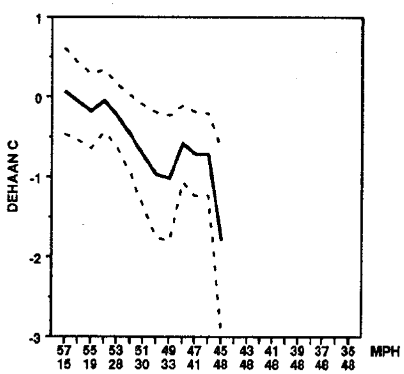
GREAT FALLS, MT (44 YEARS)



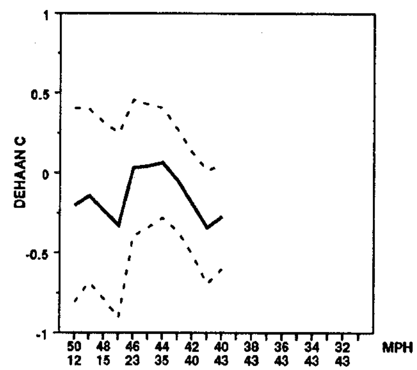
HAVRE, MT (27 YEARS)



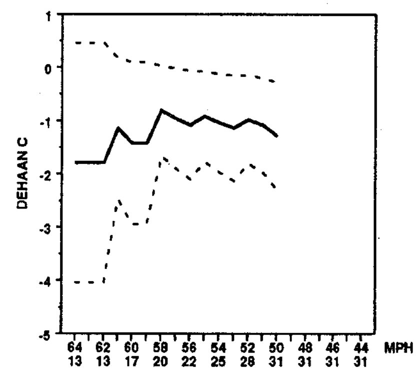
HELENA, MT (48 YEARS)



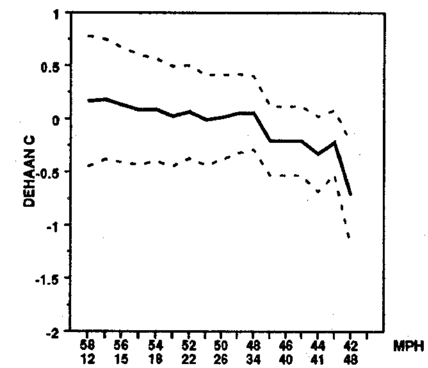
MISSOULA, MT (43 YEARS)



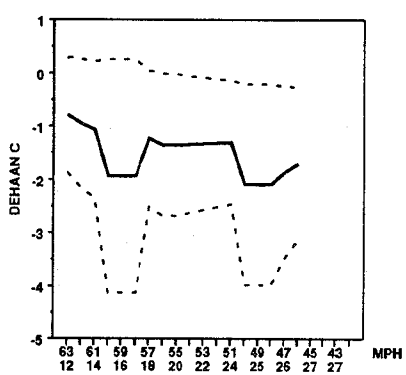
NORTH PLATTE, NE (31 YEARS)



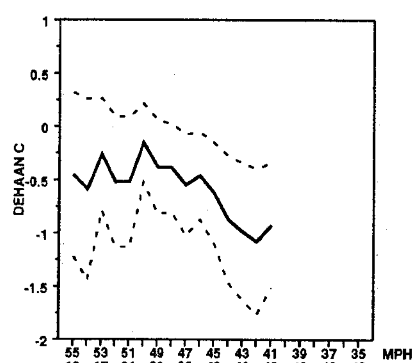
OMAHA, NE (51 YEARS)



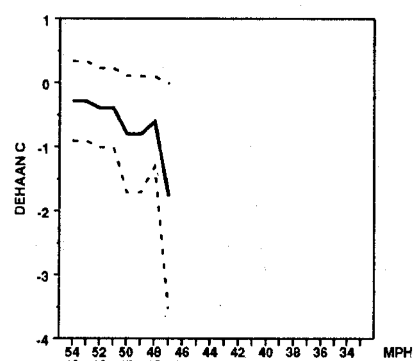
VALENTINE, NE (27 YEARS)



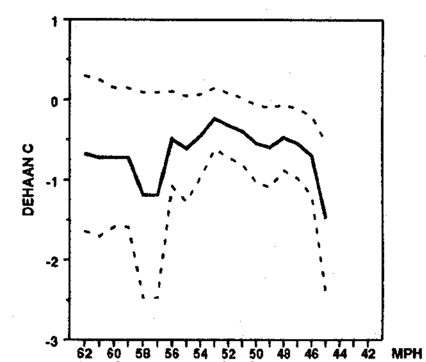
ELY, NV (49 YEARS)

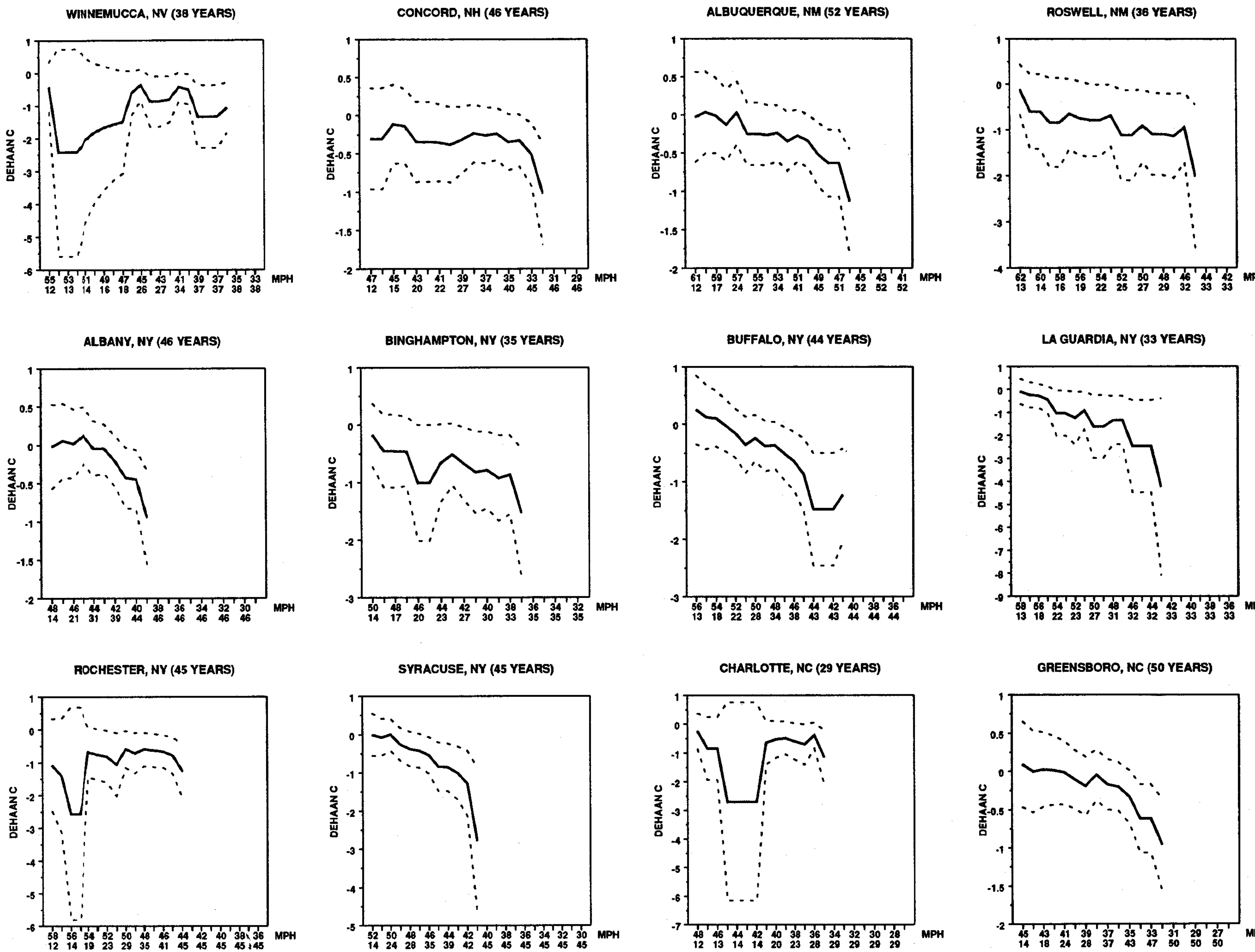


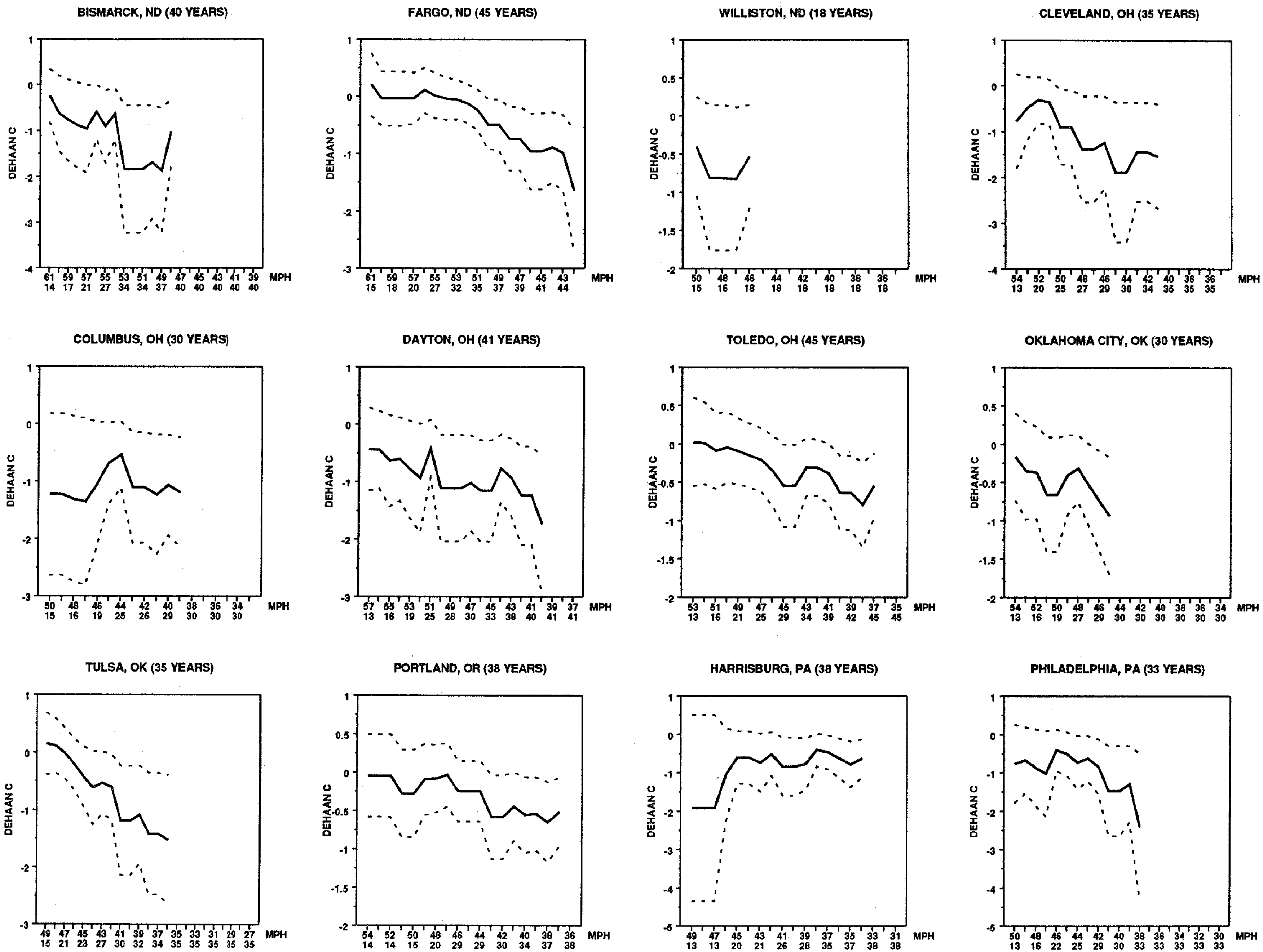
LAS VEGAS, NV (20 YEARS)



RENO, NV (45 YEARS)

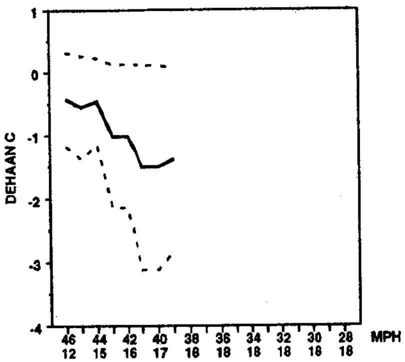



WINNEMUCCA, NV (38 YEARS)
CONCORD, NH (46 YEARS)
ALBUQUERQUE, NM (52 YEARS)
ROSWELL, NM (36 YEARS)
ALBANY, NY (46 YEARS)
BINGHAMPTON, NY (35 YEARS)
BUFFALO, NY (44 YEARS)
LA GUARDIA, NY (33 YEARS)
ROCHESTER, NY (45 YEARS)
SYRACUSE, NY (45 YEARS)
CHARLOTTE, NC (29 YEARS)
GREENSBORO, NC (50 YEARS)
DEHAAN C
MPH

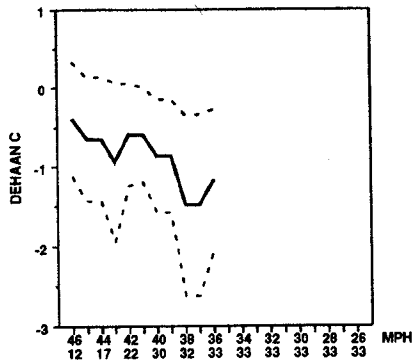


BISMARCK, ND (40 YEARS)
FARGO, ND (45 YEARS)
WILLISTON, ND (18 YEARS)
CLEVELAND, OH (35 YEARS)
COLUMBUS, OH (30 YEARS)
DAYTON, OH (41 YEARS)
TOLEDO, OH (45 YEARS)
OKLAHOMA CITY, OK (30 YEARS)
TULSA, OK (35 YEARS)
PORTLAND, OR (38 YEARS)
HARRISBURG, PA (38 YEARS)
PHILADELPHIA, PA (33 YEARS)
DEHAAN C
MPH

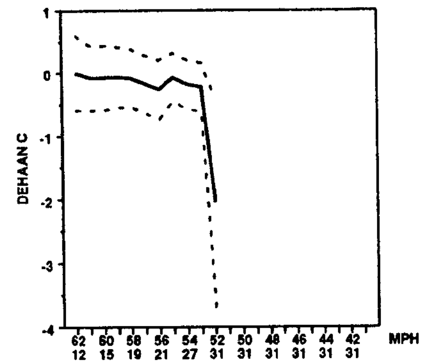
PITTSBURGH, PA (18 YEARS)



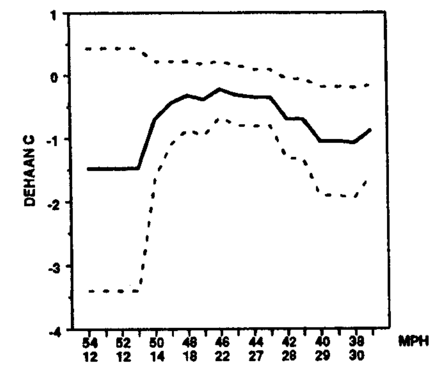
SCRANTON, PA (33 YEARS)



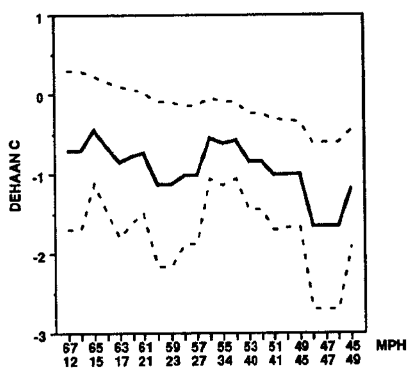
BLOCK ISLAND, RI (31 YEARS)



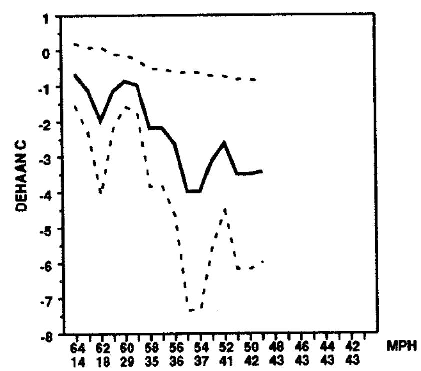
GREENVILLE, SC (43 YEARS)



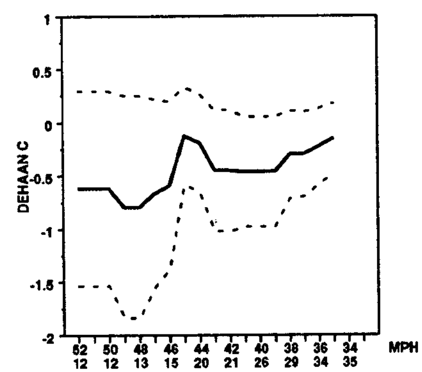
HURON, SD (49 YEARS)



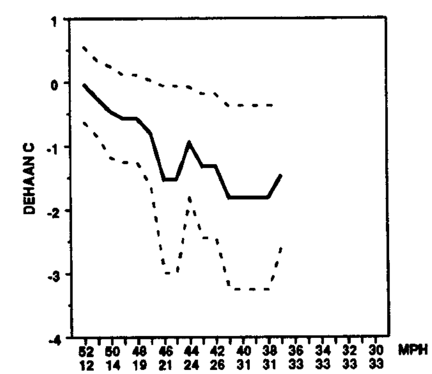
RAPID CITY, SD (43 YEARS)



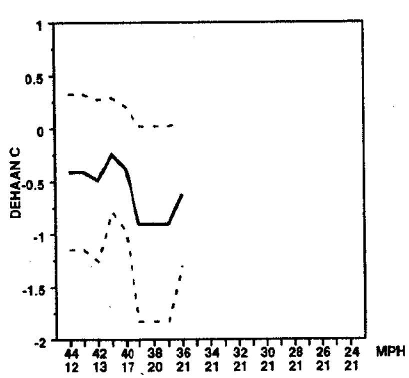
CHATTANOOGA, TN (35 YEARS)



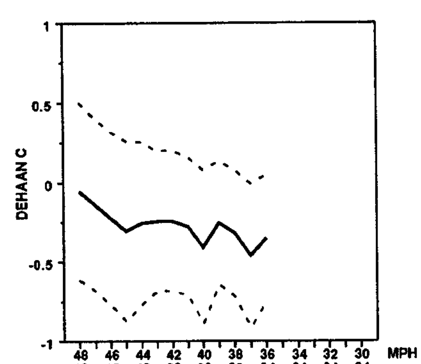
KNOXVILLE, TN (33 YEARS)



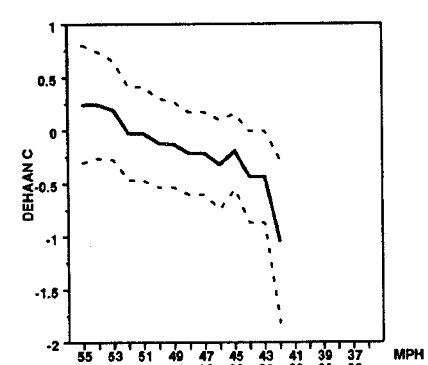
MEMPHIS, TN (21 YEARS)



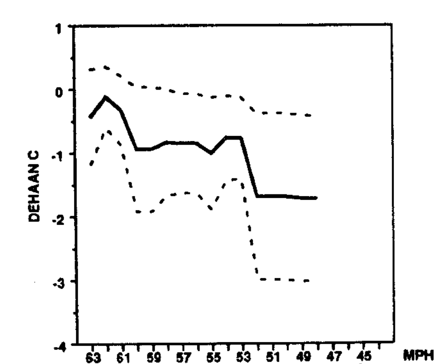
NASHVILLE, TN (34 YEARS)

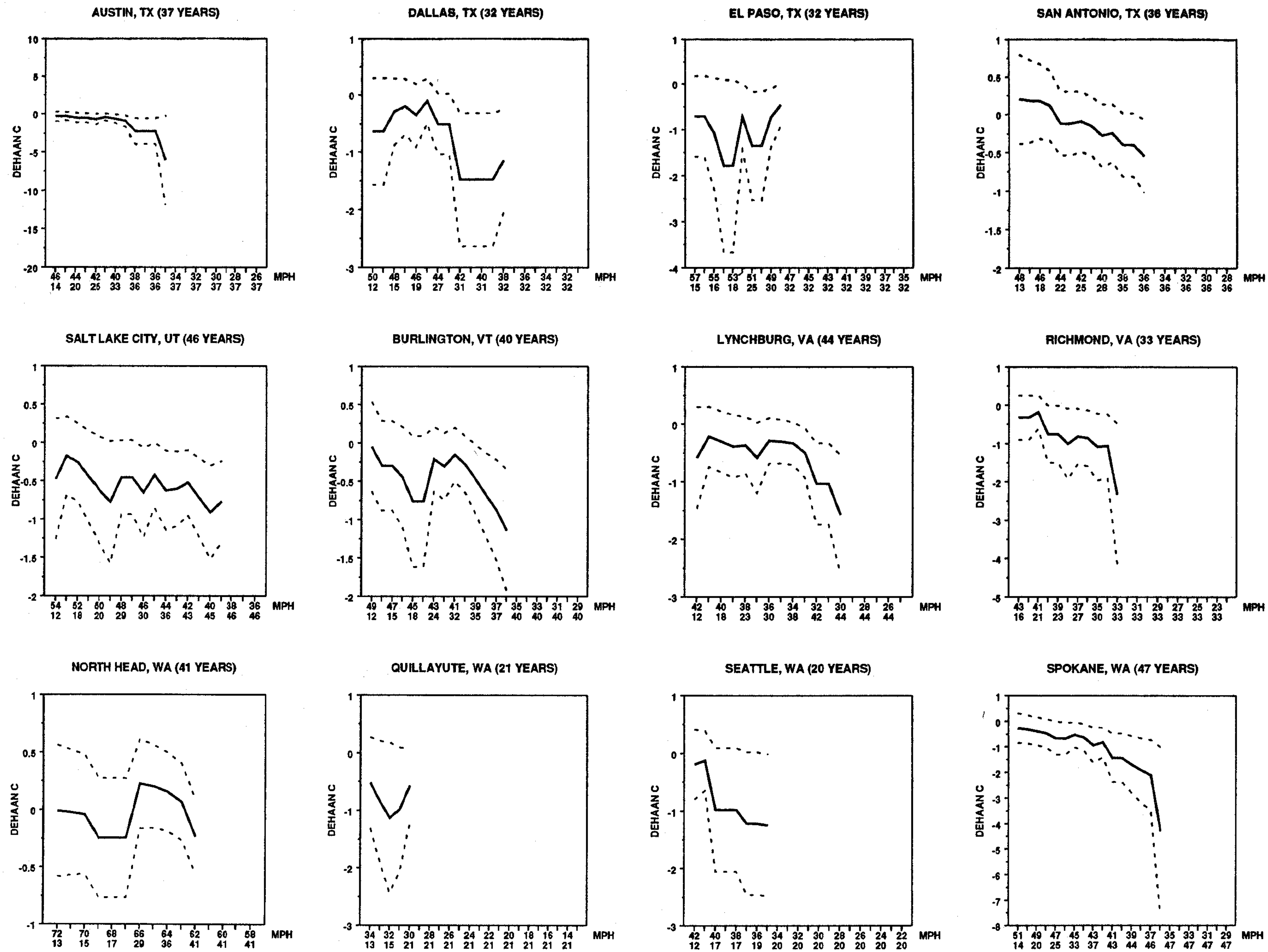


ABILENE, TX (36 YEARS)

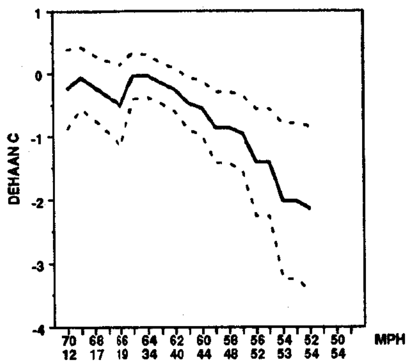


AMARILLO, TX (34 YEARS)

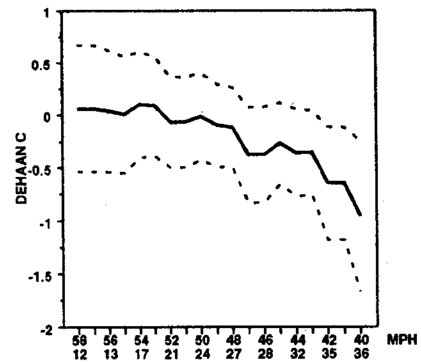



AUSTIN, TX (37 YEARS)
DALLAS, TX (32 YEARS)
EL PASO, TX (32 YEARS)
SAN ANTONIO, TX (36 YEARS)
SALT LAKE CITY, UT (46 YEARS)
BURLINGTON, VT (40 YEARS)
LYNCHBURG, VA (44 YEARS)
RICHMOND, VA (33 YEARS)
NORTH HEAD, WA (41 YEARS)
QUILLAYUTE, WA (21 YEARS)
SEATTLE, WA (20 YEARS)
SPOKANE, WA (47 YEARS)
DEHAAN C
MPH

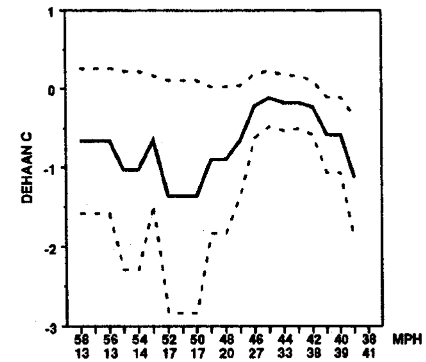
TATOOSH ISLAND, WA (54 YEARS)



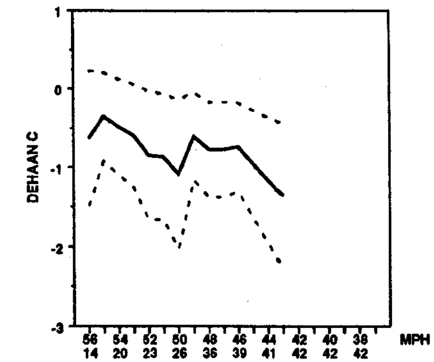
GREEN BAY, WI (36 YEARS)



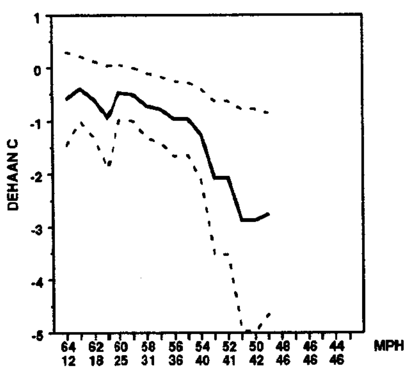
MADISON, WI (41 YEARS)



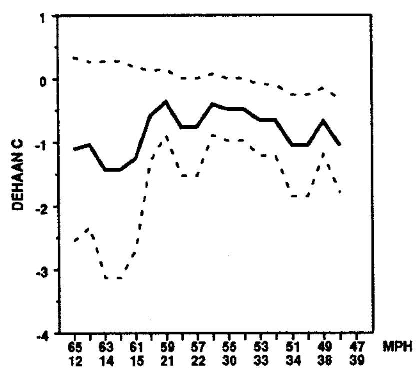
MILWAUKEE, WI (42 YEARS)



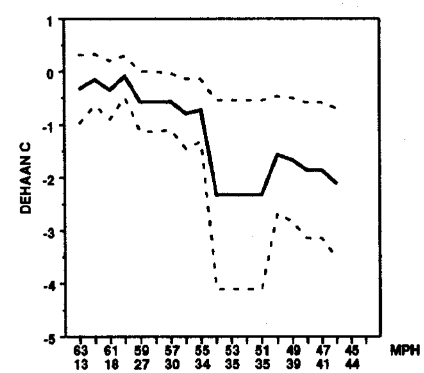
CHEYENNE, WY (46 YEARS)



LANDER, WY (42 YEARS)



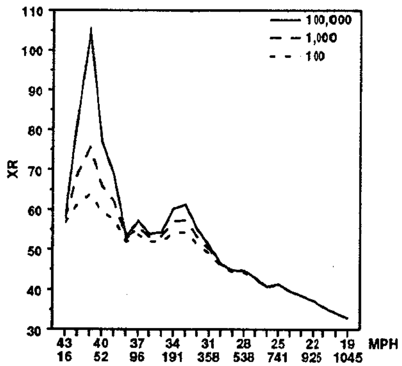
SHERIDAN, WY (44 YEARS)



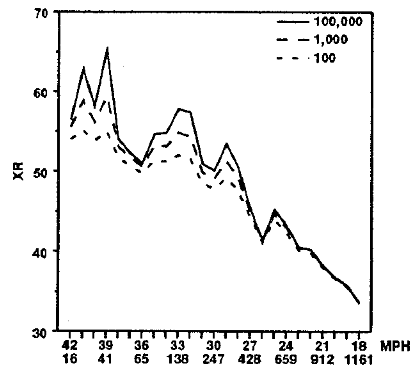
APPENDIX 3

Plots of point estimates for 100-year, 1000-year and 100 000-year speeds versus threshold and number of threshold exceedances (based on four-day interval data samples)

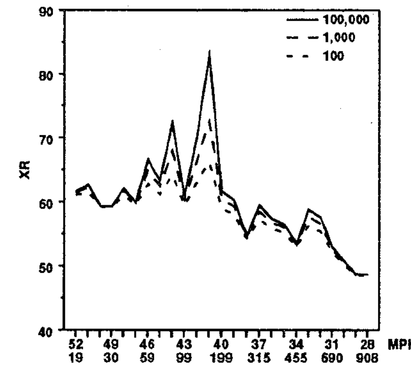
ABILENE.TX (15 YEARS)



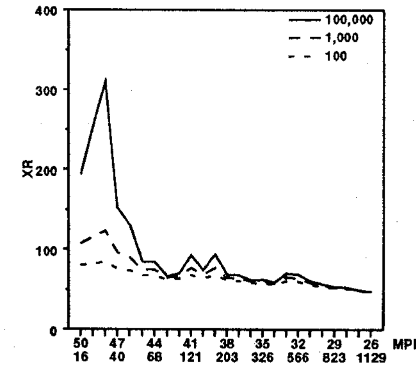
ALBANY.NY (19 YEARS)



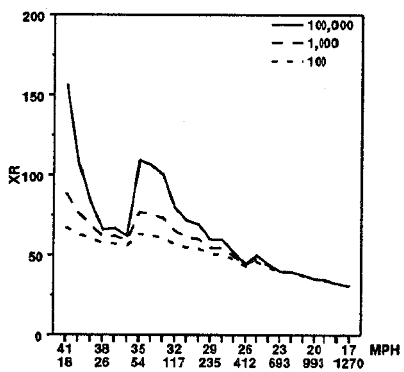
ALBUQUERQUE.NM (19 YEARS)



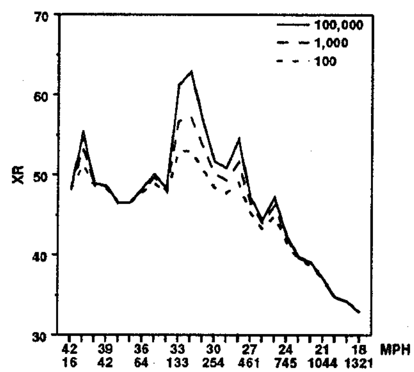
BILLINGS.MT (26 YEARS)



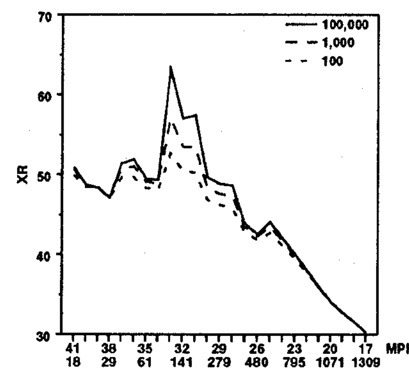
BINGHAMTON.NY (20 YEARS)



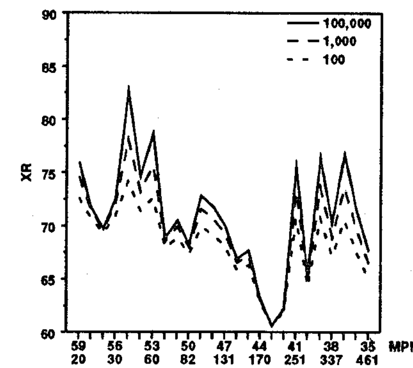
BOISE.ID (23 YEARS)



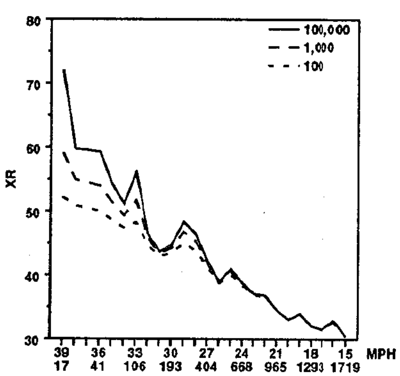
BURLINGTON.VT (19 YEARS)



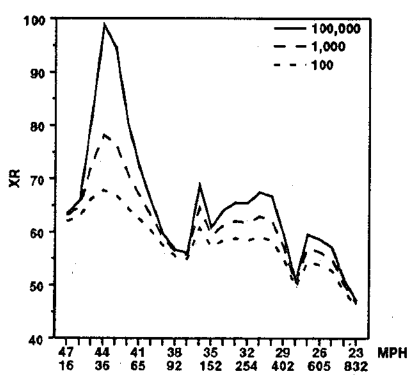
CHEYENNE_CORR.WY (15 YEARS)



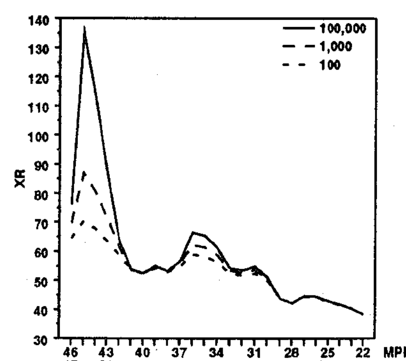
CONCORD.NH (26 YEARS)



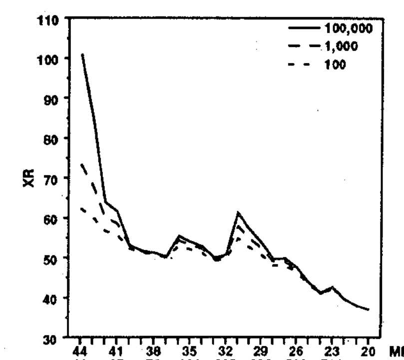
DAYTON_CORR.OH (17 YEARS)



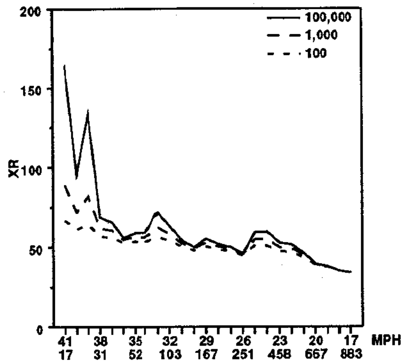
DENVER.CO (15 YEARS)



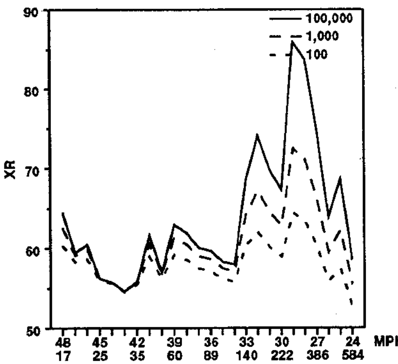
DULUTH.MN (15 YEARS)



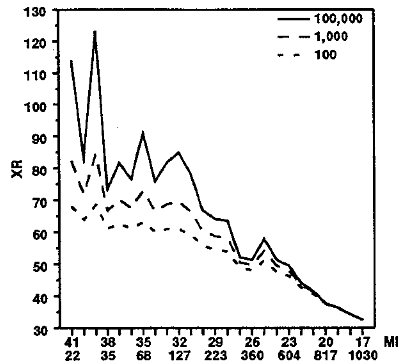
FORT_SMITH.AK (17 YEARS)



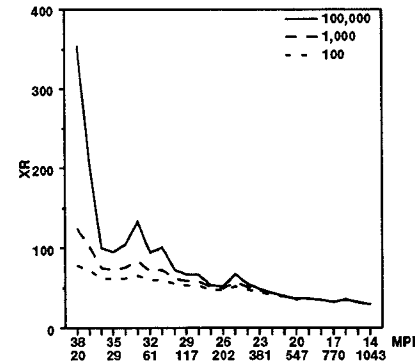
FORT_WAYNE_CORR.IN (15 YEARS)



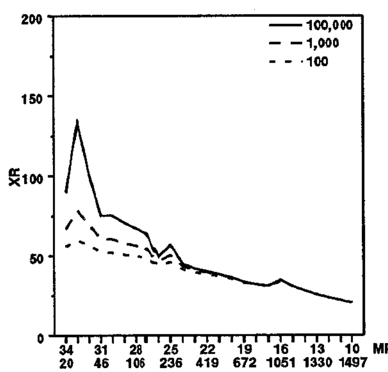
GREEN_BAY.WI (15 YEARS)



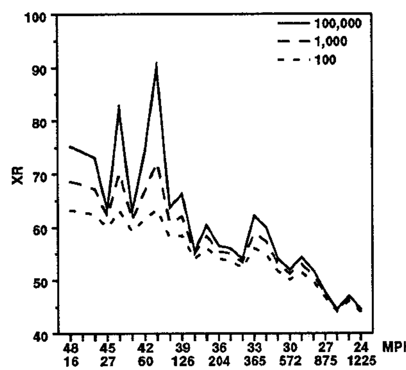
GREENSBORO.NC (19 YEARS)



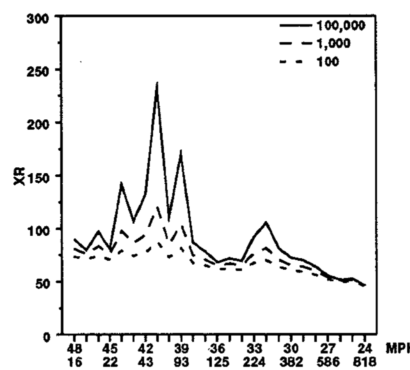
GREENVILLE_CORR.SC (19 YEARS)



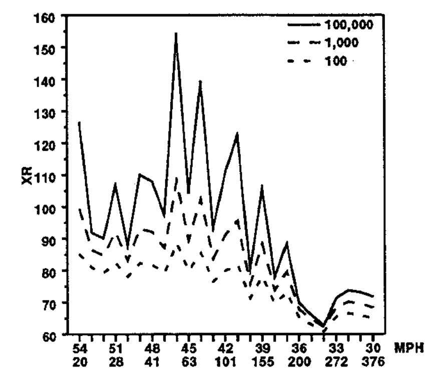
HELENA.MT (26 YEARS)



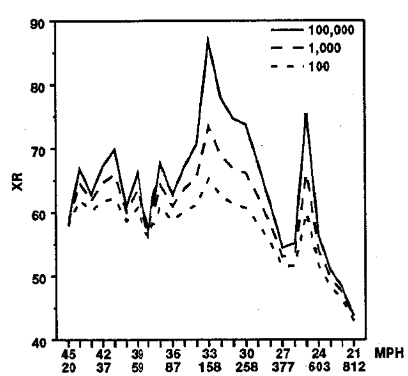
HURON.SD (15 YEARS)



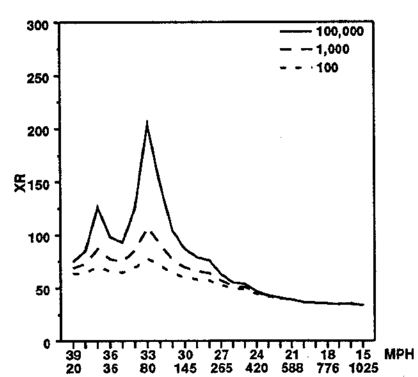
LANDER_CORR.WY (15 YEARS)



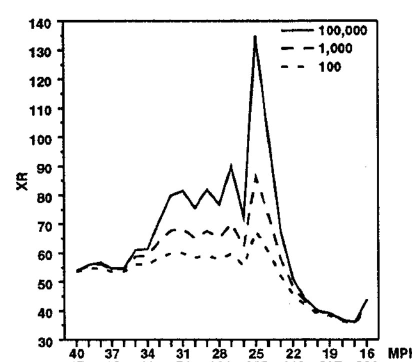
LANSING.MI (15 YEARS)



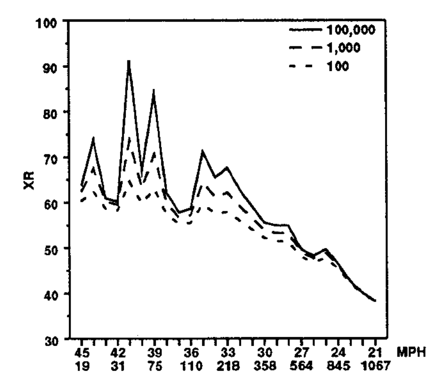
LOUISVILLE_CORR.KY (15 YEARS)

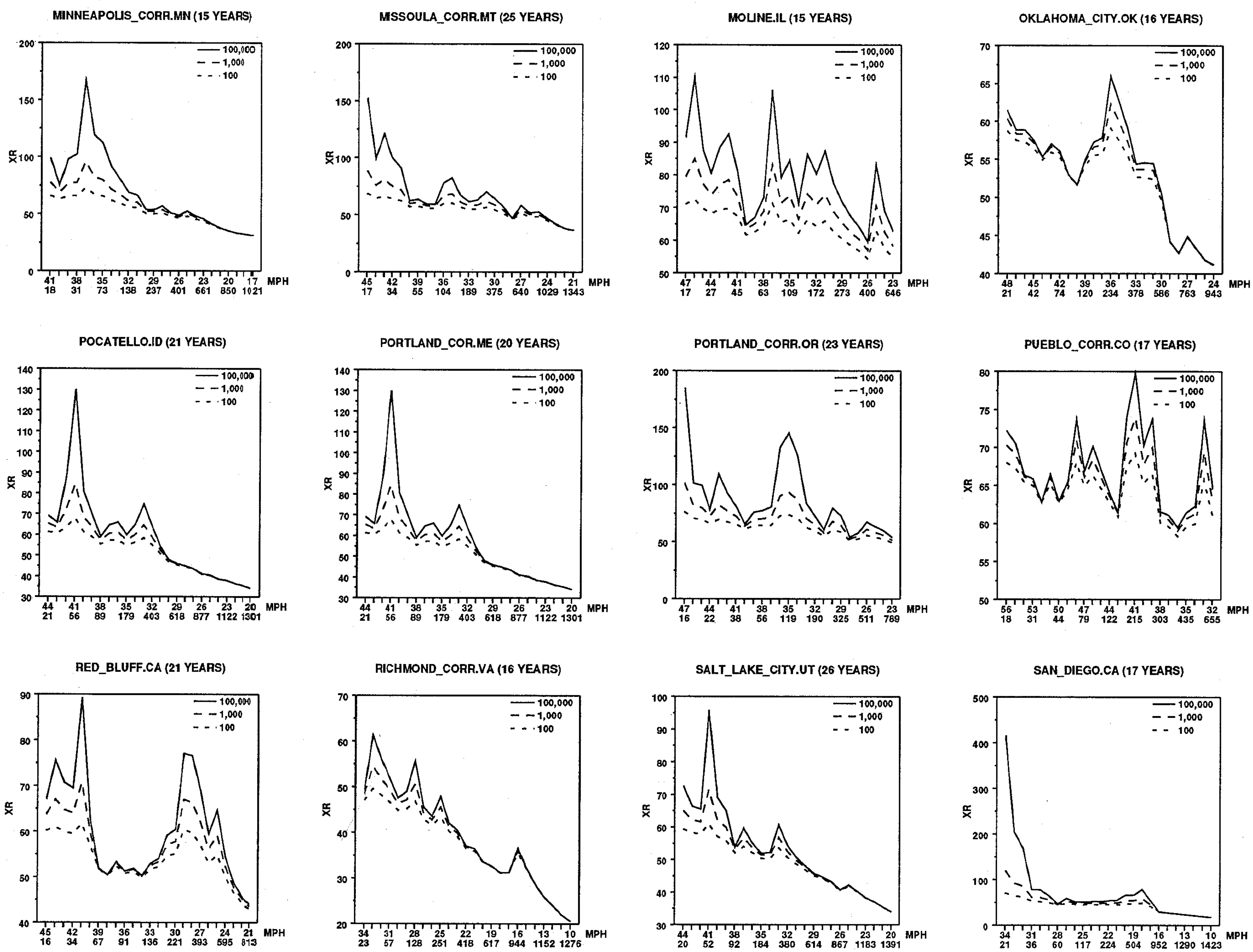


MACON.GA (16 YEARS)

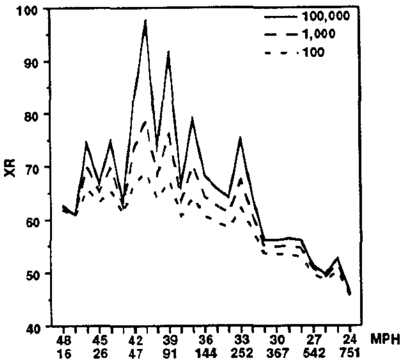


MILWAUKEE_CORR.WI (17 YEARS)

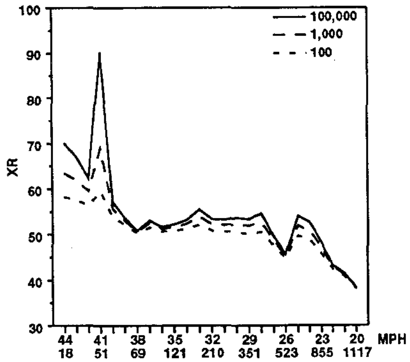



MINNEAPOLIS_CORR.MN (15 YEARS)
MISSOULA_CORR.MT (25 YEARS)
MOLINE.IL (15 YEARS)
OKLAHOMA_CITY.OK (16 YEARS)
POCATELLO.ID (21 YEARS)
PORTLAND_COR.ME (20 YEARS)
PORTLAND_CORR.OR (23 YEARS)
PUEBLO_CORR.CO (17 YEARS)
RED_BLUFF.CA (21 YEARS)
RICHMOND_CORR.VA (16 YEARS)
SALT_LAKE_CITY.UT (26 YEARS)
SAN_DIEGO.CA (17 YEARS)
100,000
1,000
100
XR
MPH

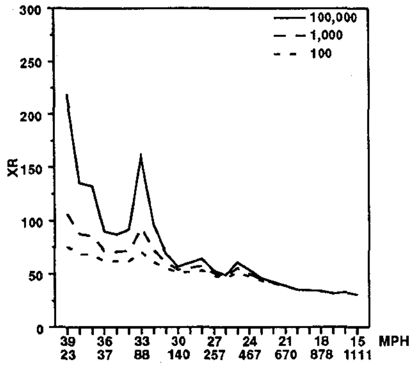
SHERIDAN_CORR.WY (17 YEARS)



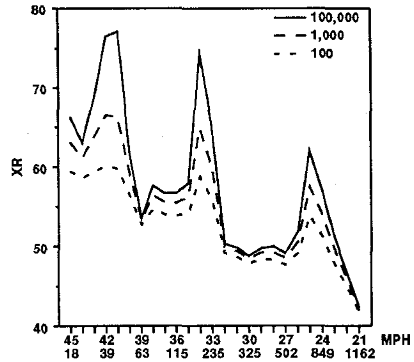
SPOKANE.WA (23 YEARS)



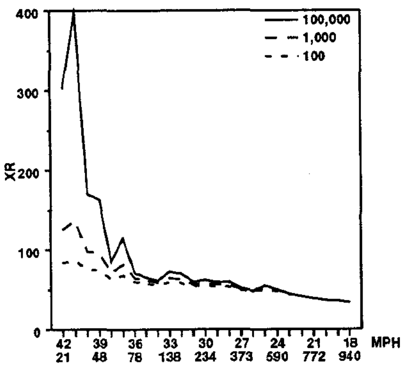
SPRINGFIELD.MO (24 YEARS)



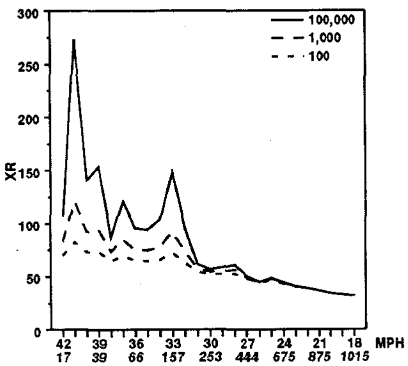
TOLEDO.OH (15 YEARS)



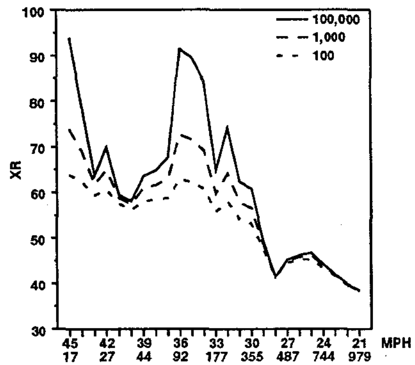
TOPEKA.KS (25 YEARS)



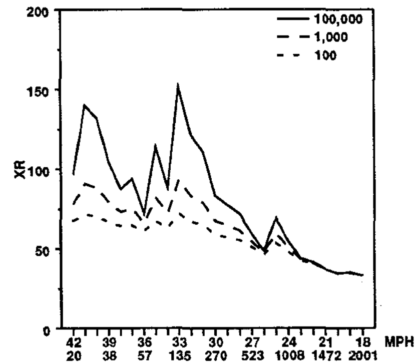
TUCSON.AZ (15 YEARS)



WINNEMUCCA.NV (15 YEARS)



YUMA.AZ (19 YEARS)



APPENDIX 4

Instructions for accessing data sets and attendant programs

NOTE. Only corrected data are included in the data files.

```
ftp enh.nist.gov (or: ftp 129.6.16.1)
>user anonymous
enter password >guest
>cd emil/datasets (to access data)
>cd emil/programs (to access programs)
>prompt off
>mget * (this copies all the data files)
>dir (this lists the available files)
>get <enh name> <local name> (this copies a specific file; example: get boise.id
 boise.id)
>quit
```


☆U.S. GOVERNMENT PRINTING OFFICE: 1995-386-627/37909