United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
Research Triangle Park NC 27711

EPA-450/4-85-010
August 1985

Air

EPA

An Examination Of 1982 - 1983 Particulate Matter Ratios And Their Use In The Estimation Of PM_{10} NAAQS Attainment Status

EPA-450/4-85-010
August 1985

AN EXAMINATION OF 1982 - 1983 PARTICULATE MATTER RATIOS AND THEIR USE IN THE ESTIMATION OF PM_{10} NAAQS ATTAINMENT STATUS

By

A. K. Pollack
A. B. Hudischewskyj
A. D. Thrall

Systems Applications, Inc.
San Rafael, CA 94903

EPA Contract No 68-02-4306

EPA Project Officer: Edwin L. Meyer, Jr.

U.S. Environmental Protection Agency
Region V, Library
230 South Dearborn Street
Chicago, Illinois 60604

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office Of Air And Radiation
Office Of Air Quality Planning And Standards
Research Triangle Park, North Carolina 27711

This report has been reviewed by the Office Of Air Quality Planning And Standards, U.S. Environmental Protection Agency, and approved for publication as received from the contractor. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, neither does mention of trade names or commercial products constitute endorsement or recommendation for use.

EPA-450/4-85-010

U.S. Environmental Protection Agency

CONTENTS

1 INTRODUCTION..... 1

2 THE DATA BASE..... 3

Constructing the Data Base..... 3

Examination of High Ratios..... 5

Definition of Valid Annual Averages..... 6

The Number of Ratios Analyzed..... 9

3 THE USE OF PARTICULATE MATTER RATIOS TO PREDICT THE LIKELY PM10 STATUS..... 11

The Pace-Frank Methodology..... 11

National Particulate Matter Distributions..... 13

4 DIFFERENCES IN CONCENTRATION RATIOS WHEN GROUPED BY TIME, LOCATION, OR TOTAL CONCENTRATION..... 25

24-Hour Ratios Undifferentiated by TSP Concentration..... 26

24-Hour Ratios on Days with High TSP Concentrations..... 29

Annual Average Ratios Undifferentiated by TSP Concentration..... 31

Annual Average Ratios Differentiated by High TSP Concentration..... 33

5 APPLICATION OF CROSS-VALIDATION TECHNIQUES TO ESTIMATE MISCLASSIFICATION RATES..... 56

Cross-Validation Applied to 24-Hour Ratio Distributions..... 57

Cross-Validation Applied to Annual Average Ratio Distributions..... 67

6 ESTIMATING DESIGN VALUES USING PM10/TSP RATIOS..... 73

Design Values Estimated from TSP Data and Ratio Distributions..... 73

Comparison with Design Values Estimated from PM10 Data.... 74

7 DISCUSSION AND RECOMMENDATIONS..... 79

Sites Selected to Monitor PM10 and PM15..... 79

National Versus Local Distributions of Ratios..... 80

Ratios Differentiated by TSP Concentration..... 82

PM10/TSP Versus PM15/TSP Distributions..... 83

Attainment Probabilities When Only PM15 Data Are Available..... 84

Conclusion..... 84

References..... 85

Appendix A: Monitoring Sites Used in the Analysis..... 88

Appendix B: Listing of Site-Days on Which at Least One 24-Hour Ratio Exceeds 1.0..... 93

Appendix C: Determination of Criteria for a Valid Annual Mean..... 101

Appendix D: Recommendation of a Constant to Use for Converting PM15 to PM10 for Use with PM15/TSP Ratios..... 113

Appendix E: Determination of Cutoff Values for High TSP..... 116

1 INTRODUCTION

The U.S. Environmental Protection Agency is proposing new short- and long-term National Ambient Air Quality Standards (NAAQS) for particulate matter having an aerodynamic diameter of less than 10 micrometers (PM10). The current NAAQS for particulate matter refers to total suspended particulate matter (TSP) concentrations without a size specification. Until PM10 data are more widely available, TSP monitoring data must be used to estimate the likelihood of attainment of the PM10 NAAQS. Pace and Frank (1984) have developed a method for estimating that likelihood. Their approach relies on the distributions of particulate matter ratios. This report addresses two key questions about the implementation of the Pace and Frank methodology:

> Are there differences in ratio distributions across time, among geographical units, from site to site? Or is it reasonable to use a national distribution?
>
> Should the ratio distribution used to determine attainment status be PM10/TSP or some other particulate matter ratio?

In this report three ratio distributions are examined: PM10/TSP, PM10/PM15, and PM15/TSP, where PM15 denotes particulate matter less than 15 micrometers in aerodynamic diameter. In Section 2 we discuss the construction of the data base used in the analysis. In Section 3 we describe the Pace and Frank methodology and the national distributions of particulate matter ratios. In Sections 4 and 5 we examine differences in ratio distributions among different locations and time periods, first using analysis of variance techniques and then using statistical cross-validation procedures to see the practical differences among the distribu-

tions. The use of the Pace and Frank methodology to estimate design values is discussed in Section 6. The results of our analysis and our recommendations are discussed in Section 7.

2 THE DATA BASE

The data base consists of 24-hour average concentrations of particulate matter collected during 1982-1983 at monitoring sites across the nation. Three types of particulate matter are monitored: total suspended particulates (TSP), particulate matter up to 15 μm in aerodynamic diameter (PM15), and particulate matter up to 10 μm in aerodynamic diameter (PM10). The data were collected in two monitoring networks. EPA has established the Inhalable Particulates (IP) network of approximately 200 monitors across the country, at which at least two of the three types of particulate matter are measured. TSP data from the national network of federal, state, and local agency monitors were also used; these data are maintained by EPA in a central data base (Storage and Retrieval of Air Quality Data, SAROAD).

CONSTRUCTING THE DATA BASE

Although TSP data were available from the IP network, TSP data from collocated monitors in the SAROAD network were used instead because of possible TSP data inconsistencies in the IP network. During the period under consideration (1982-83), the glass fiber filters used by the IP network to collect TSP were manufactured by different companies. Although important filter specifications were maintained, some of the physical characteristics of the filters changed, which appear to have had significant impact on the ambient TSP measurements (Frank, 1984). The filters used by the IP network to collect PM15 and PM10 samples did not change, however, and therefore the IP network data for these aerosols were used for our

study. The collocated TSP monitors in the SAROAD network did not undergo such filter changes and so were considered more consistent. More importantly, the filter medium used in monitoring TSP at the SAROAD sites is identical to that of sites whose TSP data will ultimately be used to draw inferences about PM10 attainment status. Collocated TSP monitoring data are available for all IP network sites except three in Durham, North Carolina.

A data file from the IP network for the period 1979 to 1983 was received on magnetic tape. Each record contained station identification information, the date, and up to 23 types of particulate measurements. PM10 and PM15 data for 1982 and 1983, along with corresponding dates and station identification codes, were extracted from this file. Any missing value, which was recorded as a blank, was recoded to -999.9999. The new file was then merged with the 1982-1983 TSP data obtained from the SAROAD network.

The IP network and SAROAD data were matched and merged by site identification codes. The site identification code is a 12-character alphanumeric code indicating state (characters 1 and 2), area (county or city, characters 3 through 6), site number within the area (characters 7 through 9), controlling agency (character 10), and project classification (characters 11 and 12); the first nine characters uniquely identify a monitoring location. Because of changes in controlling agencies and project classification, a given site may have two or more 12-character identification codes. Data corresponding to only one site identification code from the SAROAD network and one from the IP network were used for each unique monitoring location. In cases of multiple site identification codes at a given site, the particular identification code chosen was the first in an alphanumeric sort; this assured that data from a federal agency were chosen instead of data from, say, a city agency.

This matching and omitting of stations based on station identification codes and dates resulted in 166 station codes with unique nine-digit prefixes. A simpler three-digit site code varying from 1 to 166 was assigned

to each of these stations. The stations and their codes are listed in Appendix A along with state and city/county names, EPA region, and climatic classification.

To reduce the size of the data file, the data were screened for completeness and cases for which missing values appeared for all three variables (PM10, PM15, and TSP) for a given station at a given date were omitted.

A data file containing 1982 and 1983 annual averages for PM10, PM15, and TSP was created from this 24-hour data file. Since the three ratios PM10/PM15, PM10/TSP, and PM15/TSP of these pollutants were to be analyzed, separate annual averages for the numerator and demoninator of each ratio were computed in pairs. For the ratio PM10/PM15, e.g., an annual average PM10 value for 1983 was computed from only those 1983 cases for which a corresponding PM15 measurement was available. Likewise, an annual average PM15 value for 1983 was calculated from only those 1983 cases for which a corresponding PM10 measurement was available. The same method was followed in the calculation of annual average PM10 and TSP values for the second ratio and annual average PM15 and TSP for the third ratio.

From the data base of 24-hour particulate matter concentrations, three 24-hour ratios were calculated: PM10/PM15, PM10/TSP, and PM15/TSP. All three particulate matter concentrations and all three ratios were examined for possible data errors. Only one value was deleted from the data base after careful checking. At site 147 on 21 October 1983 the PM10, PM15, and TSP concentrations reported were 35.05, 1.36, and 56.0 $\mu g/m^3$, respectively; the PM15 concentration of 1.36 $\mu g/m^3$ was deleted.

EXAMINATION OF HIGH RATIOS

Theoretically, 24-hour PM10 concentrations should be less than 24-hour average PM15 concentrations, which in turn should be less than 24-hour TSP concentrations; that is, theoretically all three ratios should be less than 1.0. In actuality, however, some ratios exceed 1.0. These high

ratios may be caused, for example, by sampling variability, machine malfunction, and filter problems. In general, high ratios are seen at low particulate matter concentrations, where sampling variability produces larger percentage differences. In Table 2-1 the total number of 24-hour ratios, the percentages of ratios above 1.0 and 1.5, and the maximum ratio are given for each of the three particulate matter ratios examined. A full listing of all site-days on which at least one ratio is greater than 1.0 is given in Appendix B.

The percentage of annual average ratios exceeding unity should be less than the percentge of 24-hour ratios exceeding unity, because averages are much less variable than individual measurements. There are, in fact, only four annual averages with ratios greater than 1.0; these are listed in Table 2-2. The highest annual average (1.61 for PM15/TSP) is actually only a single 24-hour ratio. However, the second-highest annual average (1.17, for PM10/PM15) is based on 32 pairs of 24-hour concentrations.

DEFINITION OF VALID ANNUAL AVERAGES

An annual average based on only a few pairs of 24-hour concentrations is certainly less reliable than an annual average based on a full complement of 61 pairs of observations (from a once-in-six-day sampling schedule). Two approaches for dealing with different numbers of observations contributing to an annual mean are possible. The first approach is to specify that an annual average is considered valid (and thus included in the analysis) only if the number of observations exceeds a specified minimum. The second approach is to use all annual averages, but to preferentially weight, by some appropriate method, the annual averages composed of the greatest number of 24-hour observations. In the interests of simplicity and project resources the first approach was adopted. An annual average was included in analyses in this study only if there are at least 10 pairs of 24-hour concentrations contributing to the annual average; the choice of 10 as the cutoff value is justified in Appendix C.

TABLE 2-1. Percentage of 24-hour ratios exceeding 1.0.

	Total No.	Percent > 1	Percent > 1.5	Maximum Ratio
PM-10/PM-15	567	13.8	0.7	2.23
PM-10/TSP	1432	2.2	0.5	2.37
PM-15/TSP	4397	4.6	1.2	5.79

TABLE 2-2. Annual ratios exceeding 1.0.

Site Code	Year	Ratio	Number of 24-hour Observations	Numerator	Denominator
97	1982	PM15/TSP = 1.05	32	36.6	34.7
125	1982	PM15/TSP = 1.01	3	25.3	25.0
96	1983	PM10/PM15 = 1.17	32	26.5	22.7
102	1983	PM15/TSP = 1.61	1	56.4	35.0

THE NUMBER OF RATIOS ANALYZED

The number of observations available for each of the three 24-hour and (valid) annual ratios is presented in Table 2-3 by site, city, and state. The number of observations per site, city, and state varies widely. This is because the sites in the IP network were chosen as a stratified sample. One must be cautious, therefore, in generalizing from this set of selected sites.

TABLE 2-3. Numbers of sites, cities, states, and EPA regions with valid particulate matter ratios.

Geographic Unit	Number	No. of Ratios Per Geographic Unit
(a) 24-Hour Ratios		
PM10/PM15 (567 Ratios)		
Site	11	1 to 86
City	9	30 to 86
State	8	30 to 156
EPA Region	6	30 to 156
PM10/TSP (1432 Ratios)		
Site	34	1 to 89
City	34	1 to 89
State	22	1 to 172
EPA Region	10	33 to 231
PM15/TSP (4397 Ratios)		
Site	112	1 to 89
City	99	2 to 162
State	43	8 to 434
EPA Region	10	133 to 986
(b) Annual Ratios		
PM10/PM15 (18 Ratios)		
Site	10	1 to 2
City	9	2
State	8	2 to 4
EPA Region	6	2 to 4
PM10/TSP (39 Ratios)		
Site	31	1 to 2
City	31	1 to 2
State	20	1 to 4
EPA Region	10	1 to 7
PM15/TSP (133 Ratios)		
Site	98	1 to 2
City	89	1 to 4
State	41	1 to 14
EPA Region	10	3 to 27

3 THE USE OF PARTICULATE MATTER RATIOS TO PREDICT THE LIKELY PM10 STATUS

THE PACE-FRANK METHODOLOGY

Pace and Frank (1984) recommend various procedures for assessing the likelihood that a given monitoring site is in attainment of an annual or 24-hour PM10 standard. The recommendations depend on the type of measurements recorded at the site. In this report we focus on those cases in which TSP or PM15 measurements are available but PM10 measurements are not. For such cases, Pace and Frank recommend a procedure based on the sample distributions of particulate matter ratios. The procedure is based on the assumption that some proportion of the measured TSP or PM15 concentration is composed of PM10. The likely proportion, and thus the likely PM10 concentration, is estimated from the sample distributions of particulate matter ratios. If, for example, only TSP measurements are available, then the likely proportion (and hence the concentration of PM10) is based on the distribution of PM10/TSP ratios. The estimated probability of exceeding the PM10 standard, given the TSP concentration is expressed mathematically as follows:

$$\hat{Pr}\ (PM10 > C_{NAAQS} \mid TSP) = \hat{Pr}\left(\frac{PM10}{TSP} > \frac{C_{NAAQS}}{TSP} \;\middle|\; TSP\right) \quad ,$$

$$= \hat{Pr}\left(R > \frac{C_{NAAQS}}{TSP} \;\middle|\; TSP\right)$$

i.e., the estimated probability of exceeding the PM10 standard (denoted C_{NAAQS}), given a TSP concentration, is equal to the estimated probability that the ratio R of PM10 to TSP is greater than the standard divided by the given TSP concentration. Thus, the equation simply expresses the identity between the probability that the standard is exceeded given a TSP value and the probability that the ratio R (PM10/TSP) is greater than the

quantity C_{NAAQS}/TSP, given the same TSP concentration. The distribution of PM10/TSP ratios is used to estimate this probability. In fact, for the annual NAAQS and for sites with high TSP values for the 24-hour NAAQS, it is the lower portion of the distribution that is most important for the determination of the likely PM10 attainment status. For example, the fifth percentile of the distribution of PM10/TSP together with C_{NAAQS} determine the TSP value corresponding to a 95 percent probability of exceeding the level of the standard.

An analogous procedure is followed for estimating PM10 attainment likelihood from PM15 concentrations; PM15 is substituted for TSP in the above equations and the distribution of PM10/PM15 ratios is used to estimate the exceedance probabilities. In this report we also look at using PM15/TSP ratios to assess likely PM10 attainment status. The same procedure outlined above is followed except that PM15/TSP ratios are converted to PM10/TSP ratios by multiplying by the estimated fraction of PM15 composed of PM10. The fraction we use is 0.85; this choice is justified in Appendix D.

The procedures discussed above are applied differently to the annual and the 24-hour standards. To estimate the likelihood of attainment of the annual average standard, distributions of annual average particulate ratios are used in the method described. However, estimation of likely attainment status for the 24-hour standard is more complex. Distributions of 24-hour ratios are used in the procedure described above to estimate the probability of exceedance of the 24-hour standard for each day for which monitoring data are available. These probabilities are then mathematically combined to estimate the probability that the 24-hour standard has been attained. Complete details and examples are given in the report by Pace and Frank (1984).

Exceedance probabilities are estimated using the Pace and Frank methodology with sample distributions of particulate matter ratios. Pace and Frank used data from the IP network from January 1980 to December 1982 to

derive sample distributions. Because PM10 data were collected in the network only in 1982, their 24-hour sample distribution is based on a small set of PM10/TSP ratios observed in the network in 1982. Because of the paucity of PM10 annual averages, they recommended that the sample distribution of PM15/TSP ratios be used in estimating the likelihood of attainment of the annual PM10 standard; they multiplied the PM15/TSP ratios by 0.8 to estimate PM10/TSP ratios.

Annual and 24-hour distributions of PM15/TSP are discussed by Thrall and Burton (1983). They also examined the IP network database, but only 1980 and 1981 data were available at the time of their study. Their analyses have been updated (Thrall and Hudischewskyj, 1984) to include 1982 IP network data, and in fact the sample distributions they derive are the same as those in the Pace and Frank report. One of the aims of the study reported here is to update the previously derived sample particulate matter ratio distributions by including 1983 IP network data.

Thrall and Hudischewskyj (1984) deleted all particulate matter ratios exceeding 1.05 from the sample distributions, on the recommendation of EPA technical staff. These high ratios occurred in the IP Network data and are just as likely to occur in other monitoring networks. Omission of high ratios from the sample biases the estimation of the ratio distribution and decreases the probability of non-attainment of the PM10 standards. We have therefore chosen not to delete any of the ratios exceeding unity.

NATIONAL PARTICULATE MATTER DISTRIBUTIONS

The percentage points of the cumulative distribution functions (cdf) and summary statistics for the three ratios of 24-hour particulate matter concentrations are given in Table 3-1; the cdf for each of the three ratios is plotted in Figure 3-1. (In this figure and succeeding plots of 24-hour cdf's, the most extreme ratios are not shown. The entire cdf's are not depicted because of the loss of resolution that would occur if all

TABLE 3-1. Percentage points from the cumulative distribution functions of 24-hour ratios.

PM10/PM15		PM10/TSP		PM15/TSP	
Percentage	Percentile	Percentage	Percentile	Percentage	Percentile
Minimum	0.335	Minimum	0.029	Minimum	0.006
1	0.425	1	0.168	1	0.223
5	0.606	5	0.262	5	0.341
10	0.675	10	0.312	10	0.406
20	0.754	20	0.374	20	0.475
30	0.800	30	0.421	30	0.520
40	0.834	40	0.458	40	0.559
50	0.865	50	0.494	50	0.599
60	0.893	60	0.532	60	0.640
70	0.923	70	0.583	70	0.687
80	0.964	80	0.643	80	0.748
90	1.043	90	0.737	90	0.849
95	1.204	95	0.836	95	0.984
99	1.410	99	1.180	99	1.580
Maximum	2.234	Maximum	2.368	Maximum	5.790
Average	0.872	Average	0.521	Average	0.632
Standard deviation	0.184	Standard deviation	0.205	Standard deviation	0.270
Number of cases	567	Number of cases	1432	Number of cases	4397

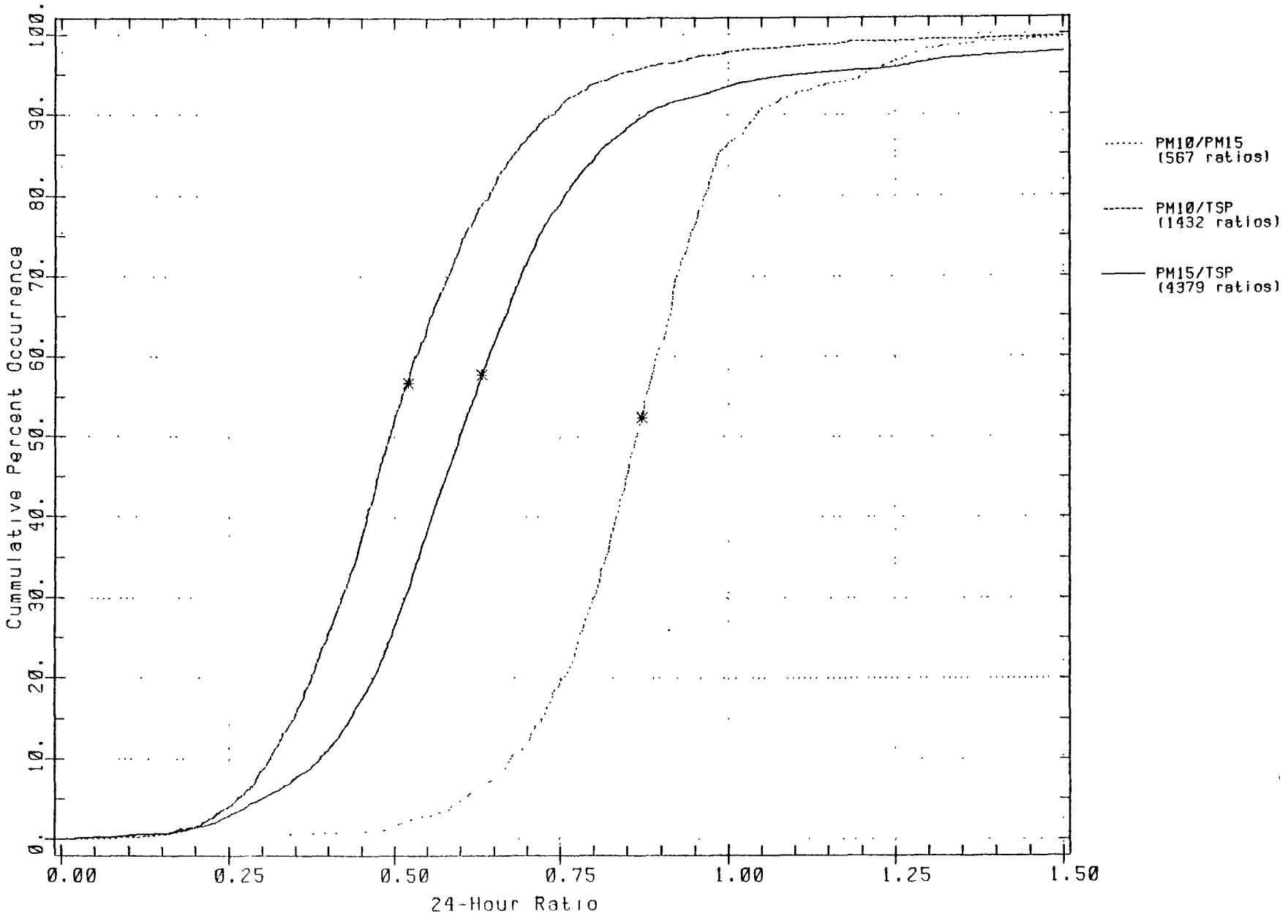

FIGURE 3-1. Cumulative distributions of 24-hour particulate matter ratios derived from the IP network, 1982-1983.

ratios were included.) An asterisk in the cdf's in Figure 3-1 and in all succeeding cdf plots indicates the mean of the distribution.

Percentage points and summary statistics of the cdf's of annual averages ratios are listed in Table 3-2, and the cdf's for all three ratios are plotted in Figure 3-2. There is much variability in these ratio distributions, particularly in the 24-hour distributions. It is exactly because of the large variability in the ratios, in fact, that Pace and Frank recommend their probabilistic approach rather than just using constants to convert PM15 or TSP concentrations to PM10 concentrations.

Although the particulate matter ratios are more variable at lower concentrations, it is the ratios corresponding to high levels of TSP that are of interest because PM10 concentrations are not likely to be high unless TSP concentrations are also high. Pace and Frank as well as Thrall and Hudischewskyj suggest examining ratio distributions at high levels of TSP because the distributions at low levels are less stable. A scatterplot of PM15/TSP ratios against TSP shows that this is indeed the case--the highest ratios occur at low levels of TSP, and there is much less variance in the ratios at high levels of TSP (Figure 3-3). The ratio distributions for 24-hour TSP concentrations greater than 100 $\mu g/m^3$ and annual average TSP concentrations greater than 55 $\mu g/m^3$ are more stable than the full distributions; further statistical analyses leading to the choice of these two cutoff values can be found in Appendix E.

Percentage points and summary statistics for these so-called high-TSP distributions of 24-hour and annual average ratios are listed in Tables 3-3 and 3-4, respectively.* The cdf's are plotted in Figures 3-4 and 3-5. The standard deviations of these distributions are, on the average, about 20 percent less than those of the full range distributions. Also,

---

* For the 24-hour PM10/PM15 distribution, ratios were included only if TSP was measured concurrently and was above 100 $\mu g/m^3$; similarly, for the annual PM10/PM15 distribution, ratios were included only if the annual average TSP concentration was above 55 $\mu g/m^3$.

TABLE 3-2. Percentage points from the cumulative distribution functions of annual ratios.

PM10/PM15		PM10/TSP		PM15/TSP	
Percentage	Percentile	Percentage	Percentile	Percentage	Percentile
Minimum	0.74	Minimum	0.28	Minimum	0.32
1	-	1	-	1	0.32
5	0.74	5	0.32	5	0.44
10	0.75	10	0.34	10	0.47
20	0.77	20	0.41	20	0.51
30	0.80	30	0.44	30	0.54
40	0.83	40	0.46	40	0.57
50	0.86	50	0.49	50	0.60
60	0.88	60	0.51	60	0.63
70	0.89	70	0.55	70	0.66
80	0.91	80	0.58	80	0.71
90	0.96	90	0.63	90	0.74
95	0.97	95	0.66	95	0.78
99	-	99	-	99	0.99
Maximum	1.17	Maximum	0.69	Maximum	1.05
Average	0.87	Average	0.50	Average	0.61
Standard deviation	0.10	Standard deviation	0.10	Standard deviation	0.12
Number of cases	18	Number of cases	39	Number of cases	133

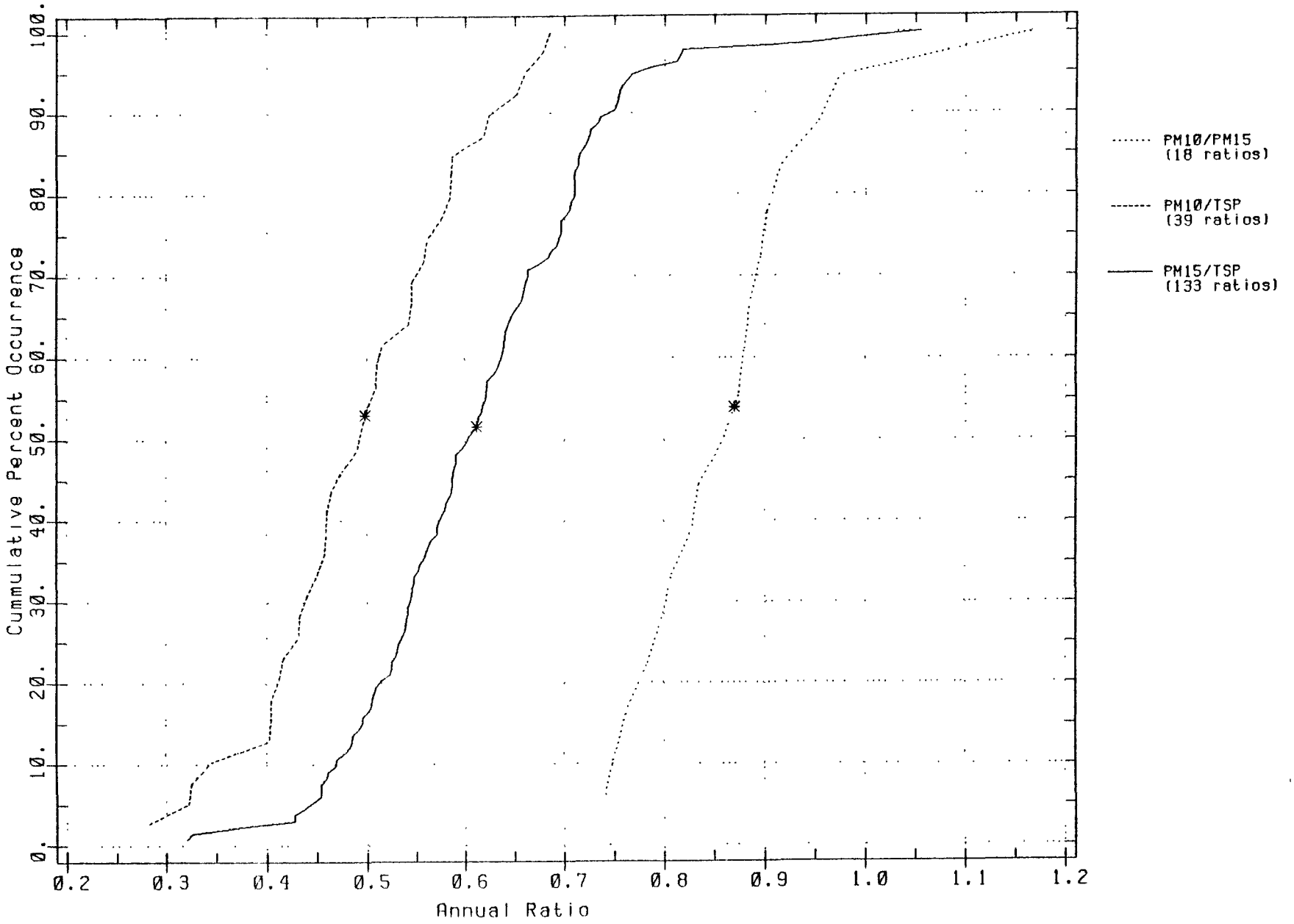

FIGURE 3-2. Cumulative distributions of annual particulate matter ratios derived from the IP network, 1982-1983.

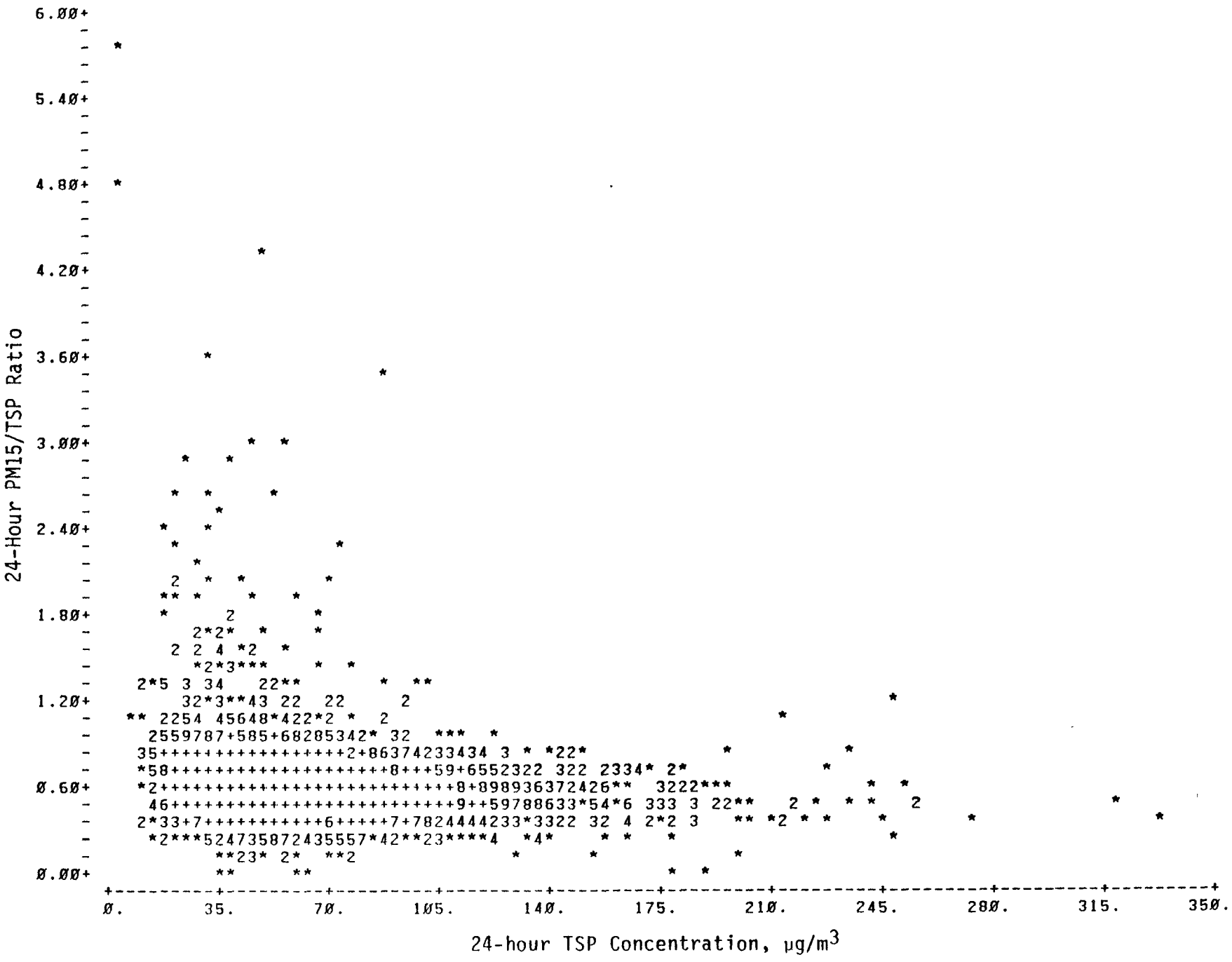

FIGURE 3-3. 24-hour PM15/TSP versus TSP.

TABLE 3-3. Percentage points from the cumulative distribution functions of ratios of high TSP 24-hour averages.

PM10/PM15		PM10/TSP		PM15/TSP	
Percentage	Percentile	Percentage	Percentile	Percentage	Percentile
Minimum	0.341	Minimum	0.029	Minimum	0.006
1	0.357	1	0.140	1	0.218
5	0.621	5	0.223	5	0.303
10	0.687	10	0.275	10	0.361
20	0.747	20	0.334	20	0.431
30	0.783	30	0.396	30	0.479
40	0.815	40	0.433	40	0.515
50	0.843	50	0.472	50	0.541
60	0.867	60	0.507	60	0.574
70	0.906	70	0.547	70	0.613
80	0.936	80	0.597	80	0.671
90	0.970	90	0.687	90	0.734
95	1.181	95	0.754	95	0.803
99	1.393	99	0.950	99	0.932
Maximum	1.664	Maximum	1.181	Maximum	1.293
Average	0.852	Average	0.478	Average	0.549
Standard deviation	0.167	Standard deviation	0.165	Standard deviation	0.152
Number of cases	122	Number of cases	351	Number of cases	579

TABLE 3-4. Percentage points from the cumulative distribution functions of high TSP annual averages.

PM10/PM15		PM10/TSP		PM15/TSP	
Percentage	Percentile	Percentage	Percentile	Percentage	Percentile
Minimum	0.75	Minimum	0.28	Minimum	0.32
1	-	1	-	1	0.32
5	0.75	5	0.32	5	0.43
10	0.76	10	0.34	10	0.45
20	0.78	20	0.40	20	0.50
30	0.80	30	0.43	30	0.53
40	0.83	40	0.46	40	0.55
50	0.85	50	0.47	50	0.58
60	0.87	60	0.51	60	0.59
70	0.88	70	0.54	70	0.62
80	0.90	80	0.56	80	0.66
90	0.93	90	0.59	90	0.71
95	0.96	95	0.63	95	0.73
99	-	99	-	99	0.76
Maximum	0.97	Maximum	0.66	Maximum	0.94
Average	0.85	Average	0.48	Average	0.58
Standard deviation	0.07	Standard deviation	0.09	Standard deviation	0.10
Number of cases	15	Number of cases	35	Number of cases	91

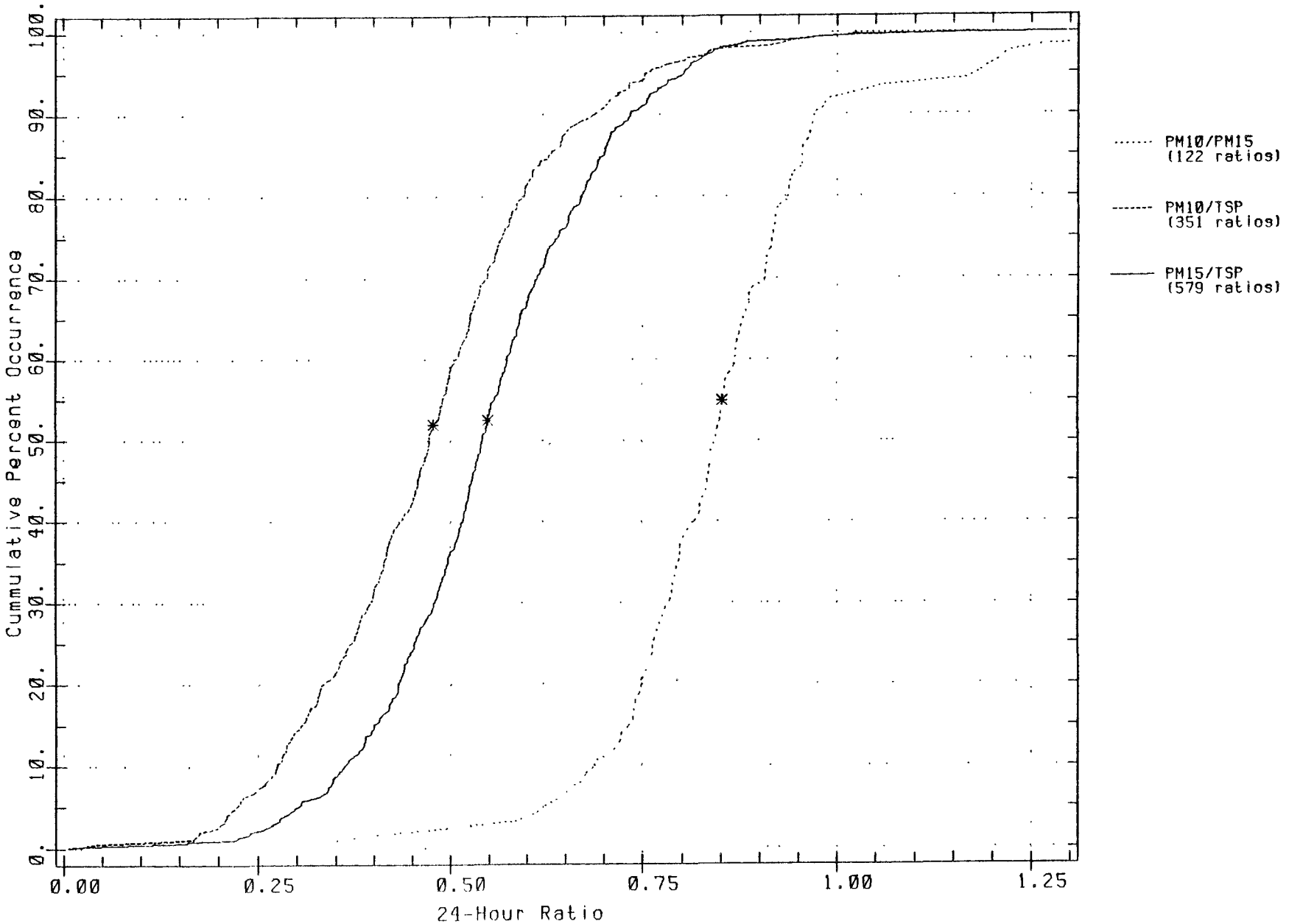

FIGURE 3-4. Cumulative distributions of 24-hour particulate matter ratios derived from the IP network, 1982-1983, for TSP >= 100 ug/m3.

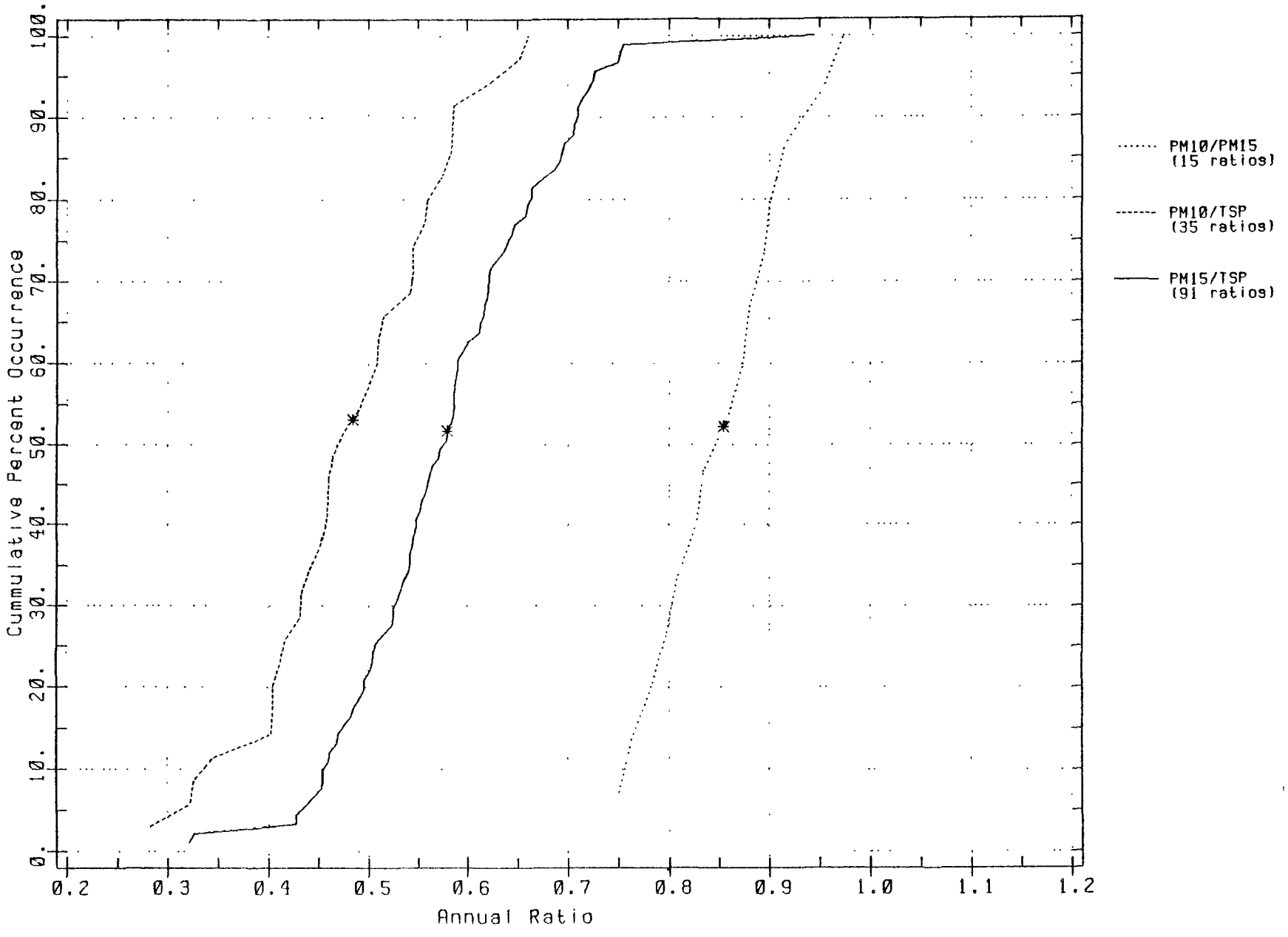

FIGURE 3-5. Cumulative distributions of annual particulate matter ratios derived from the IP network, 1982-1983, for TSP >= 55 ug/m3.

because of the removal of most of the high ratios, the ratio distributions are shifted downward, and thus the ratios at corresponding percentage points are lower; this is especially true for the PM15/TSP ratio.

4 DIFFERENCES IN CONCENTRATION RATIOS WHEN GROUPED BY TIME, LOCATION, OR TOTAL CONCENTRATION

The ratio distributions described in the previous section have large ranges. The variability of each distribution, however, may be due to differences in ratios by time or location. In this section we examine distributions of ratios grouped by time, location, and total concentration to see if these factors account for the observed variability. If large differences in these distributions are observed, it is worthwhile to investigate the use of the Pace-Frank methodology with distinct ratio distributions for each geographic region instead of using one national distribution.

Differences in ratio distributions are examined with one-way analysis of variance (ANOVA) models (Winer, 1971) using the computer package BMDP7D (Dixon et al., 1983). ANOVA tests the hypothesis that the underlying mean ratio is the same from one group to another. This is a standard technique for testing differences among groups. Passing the test does not imply that entire distributions are equivalent; if group means are similar but the variances differ, then the test will be passed. Failing the equality of means test indicates that the distributions differ beyond that expected for groups with equal means, given the variability in the distributions. ANOVA provides a relatively simple way to document statistically significant group differences when testing the hypothesis of equal group means. Large group differences produce large F-statistics and small p-values (the probability of observing an F-statistic as large if the group means are in fact equal, given the variability in the distributions).

The results of an ANOVA test on differences in distributions across groups should be interpreted cautiously, for several reasons. First, it is possible for the distributions of particulate matter ratios to be markedly

different from one group to the next even if the average ratio is nearly the same in each group. Also, the finding that differences in average ratios are "statistically significant" does not address the magnitude of the differences, i.e., the practical significance, since small differences are statistically detectable if the sample size is sufficiently large; small differences can be statistically significant but not practically significant. The only way to evaluate the practical differences among distributions of different groups of ratios is to compare the likelihood of attainment calculated with the Pace and Frank methodology, using different distributions; this is done in Section 5. Finally, the ANOVA results should be interpreted with the number and representativeness of sites in each group in mind. This point is discussed further in later sections.

Another technique for testing differences in distributions is the nonparametric Kolmogorov-Smirnov test, which looks not at differences in means but differences in percentage points of distributions. This test is computationally intensive for testing differences between two distributions; the difficulty increases significantly for testing differences among three or more distributions. We did not use these tests extensively because of funding limits. In any case, these tests are much less powerful (i.e., less able to detect significant differences) when screening many groups of small samples. Thrall and Hudischewskyj did conduct some Kolomogorov-Smirnov tests, and the findings were not different from their ANOVA tests.

24-HOUR RATIOS UNDIFFERENTIATED BY TSP CONCENTRATION

Table 4-1 lists the results of ANOVA tests on differences between distributions of different groupings of ratios for PM10/PM15, PM10/TSP, and PM15/TSP. The distributions are compared by year, quarter of the year, high or low TSP concentration (i.e., above or below 100 $\mu g/m^3$), climatic region, EPA region, state, city, and site. If the p value from the test of equality of group means is low (below 0.05, for example), then differences between group mean ratios are said to be statistically significant.

TABLE 4-1. Differences among 24-hour ratios by groups.

Grouping Factor	No. of Groups	Average No. Obs. Per Group	Group Average		ANOVA
			Minimum	Maximum	Tail Prob.
Year					
PM10/PM15	2	284	.853	.890	.006
PM10/TSP	2	716	.511	.557	.001
PM15/TSP	2	2199	.627	.643	.031
Quarter					
PM10/PM15	4	142	.859	.905	.114
PM10/TSP	4	358	.480	.546	.003
PM15/TSP	4	1099	.609	.660	.000
High/Low TSP (100 $\mu g/m^3$ cutoff)					
PM10/PM15	2	235	.852	.889	.060
PM10/TSP	2	716	.478	.535	.000
PM15/TSP	2	2199	.549	.644	.000
Climatic Region					
PM10/PM15	3	189	.748	.900	.000
PM10/TSP	3	477	.449	.572	.000
PM15/TSP	3	1466	.517	.652	.000
EPA Region					
PM10/PM15	6	95	.777	1.019	.000
PM10/TSP	10	143	.408	.639	.000
PM15/TSP	10	440	.515	.708	.000
State					
PM10/PM15	8	71	.748	1.019	.000
PM10/TSP	22	65	.323	.691	.000
PM15/TSP	43	102	.422	.833	.000
City					
PM10/PM15	9	63	.748	1.019	.000
PM10/TSP	34	42	.307	.695	.000
PM15/TSP	99	44	.350	1.024	.000
Site					
PM10/PM15	11	52	.748	1.019	.000
PM10/TSP	34	42	.307	.695	.000
PM15/TSP	112	39	.333	1.024	.000

Differences in the distributions of 24-hour PM10/PM15, PM10/TSP, and PM15/TSP ratios by year are shown in Figure 4-1 (Figure 4-1 and all other figures for this section appear at the end of the section). PM10/PM15 and PM15/TSP ratios are higher in 1983 than in 1982; the converse is true for PM10/TSP ratios. Although the ANOVA test found the difference between years statistically significant (Table 4-1), Figure 4-1 shows that the differences in the percentiles of the distribution are small. There are statistically significant differences in the distributions of PM10/TSP and PM15/TSP ratios by quarter, but not in the distribution of PM10/PM15 ratios. However, the patterns of the quarterly means are different for all three ratio distributions, i.e., there are no consistent changes in the ratio from one quarter to the next.

Differences in the distributions of 24-hour ratios differentiated by TSP concentration are shown in Figure 4-2 for PM10/PM15, PM10/TSP, and PM15/TSP. The ANOVA test revealed statistically significant difference in group means for all three ratio distributions (Table 4-1). In all three plots, percentiles of distributions corresponding to TSP concentrations above 100 $\mu g/m^3$ are lower than the corresponding percentiles of the low TSP distributions. These differences tend to be larger in the upper percentage points of the ratio distributions.

Differences in ratio distributions by climatic region (West Coast, arid, and "other") are depicted in Figure 4-3. In all three cases ratios in arid areas are the lowest and ratios in non-arid areas outside the West Coast are the highest. The ANOVA tests show climate effects for all three particulate matter ratios. However, for the PM10/PM15 ratio, both the West Coast and the arid climate distributions are based on one site only, so inferences cannot be made about actual climatic differences.

Figure 4-4 shows differences in ratio distributions by EPA region. The ANOVA tests show statistically significant differences among the regional means (Table 4-1). Most of the regional differences observed in PM10/PM15 distributions are attributable to site differences because three of the

regions are represented by only one site and only one region is represented by more than two sites. If any generalization can be made about regional differences in PM10/TSP and PM15/TSP differences, it is that ratios in regions VI, VII, VIII, IX, and X tend to be higher than in other regions. It is difficult to assess regional patterns, however, because large differences are based on comparisons among only a few sites; where there are groups with many sites there are relatively small differences.

The ANOVA results also show significant differences among states, cities, and sites for all three kinds of ratios. Ratio distributions among states and cities show significant differences, but because of the small number of sites (often only one) per geographic unit, the observed differences may only be site-to-site differences. Only three states have a large number of sites: California, Ohio, and Texas. Only a few cities are represented in the PM10/PM15 and PM15/TSP ratios by more than one site; no cities are represented in the PM10/TSP ratios by more than one site (in Table 4-1, there are 34 sites and 34 cities of PM10/TSP ratios).

24-HOUR RATIOS ON DAYS WITH HIGH TSP CONCENTRATIONS

The results of ANOVA applied only to 24-hour ratios on days with high TSP ($> 100\ \mu g/m^3$) are presented in Table 4-2. Comparing Tables 4-2 and 4-1 (results for all ratios undifferentiated by TSP concentration) we see that the only change in ANOVA results is that PM10/TSP ratios are no longer significantly different between quarters. Significant differences from year to year are still detected, though the 1983 means are no longer consistently higher than the 1982 means. For groupings by quarter, only PM15/TSP ratios show significant differences: means of the first two quarters are significantly lower than those of the last two quarters. This quarterly difference may represent site-to-site differences since most sites did not record measurements in all four quarters.

For all kinds of geographical groupings, ANOVA detects significant differences among the means for all three three kinds of ratios. As with the

TABLE 4-2. Differences among 24-hour high TSP ratios by groups.

Grouping Factor	No. of Groups	Average No. Obs. Per Group	Group Average		ANOVA Tail Prob.
			Minimum	Maximum	
Year					
PM10/PM15	2	61	.820	.896	.012
PM10/TSP	2	176	.469	.508	.026
PM15/TSP	2	290	.534	.587	.000
Quarter					
PM10/PM15	4	31	.831	.872	.460
PM10/TSP	4	88	.459	.489	.505
PM15/TSP	4	145	.522	.585	.000
Climatic Region					
PM10/PM15	3	41	.768	.905	.000
PM10/TSP	3	117	.419	.537	.000
PM15/TSP	3	193	.496	.590	.000
EPA Region					
PM10/PM15	6	20	.738	.966	.000
PM10/TSP	10	35	.324	.661	.000
PM15/TSP	10	58	.472	.621	.000
State					
PM10/PM15	7	17	.738	.966	.000
PM10/TSP	18	20	.274	.661	.000
PM15/TSP	39	15	.306	.675	.000
City					
PM10/PM15	8	15	.738	.981	.000
PM10/TSP	30	12	.264	.684	.000
PM15/TSP	77	8	.306	1.063	.000
Site					
PM10/PM15	8	15	.738	.981	.000
PM10/TSP	30	12	.264	.684	.000
PM15/TSP	83	7	.306	1.063	.000

undifferentiated 24-hour ratios, however, many of the geographical groupings are represented by only one or two sites. Therefore, one cannot tell if the ANOVA tests are significant because of geographical differences or simply site-to-site differences. We do not have a large enough sample of sites, nor a random distribution of sites, to conclusively detect geographical differences.

ANNUAL AVERAGE RATIOS UNDIFFERENTIATED BY TSP CONCENTRATION

Results of the ANOVA tests for differences between distributions of different groupings of annual average ratios are presented in Table 4-3. There are fewer significant differences among groups of annual ratios than among 24-hour ratios, but there are also far fewer observations in each distribution. In addition, fewer site-years of data are available for the analysis of annual ratios, because annual ratios composed of fewer than 10 sample-days have been omitted.

The differences in annual average ratio distributions by year are not significant for any of the three kinds of ratios, unlike the results for the 24-hour ratios. Also, while 24-hour ratios for 1983 are consistently higher than those for 1982, there is no such pattern in the annual average ratios.

Differences in the distributions of annual average PM10/TSP and PM15/TSP ratios differentiated by TSP concentration are shown in Figure 4-5. The test for differences in PM10/PM15 distributions could not be performed because the low-TSP distribution is represented by only one site for one site-year. The ANOVA results for the other two ratios indicate statistically significant differences between ratio distributions differentiated by TSP concentration. In Figure 4-5, we see that for a given percentage point, the ratio corresponding to high TSP is lower than that for low TSP.

Figure 4-6 shows differences in ratio distributions by climatic region. No significant differences in annual average ratios for PM10/PM15 are

TABLE 4-3. Differences among annual ratios by groups.

Grouping Factor	No. of Groups	Average No. Obs. Per Group	Group Average		ANOVA
			Minimum	Maximum	Tail Prob.
Year					
PM10/PM15	2	9	.854	.877	.653
PM10/TSP	2	20	.491	.522	.443
PM15/TSP	2	67	.608	.618	.647
High/Low TSP (55 $\mu g/m^3$ cutoff)					
PM10/PM15	2	9	.854	1.169	*
PM10/TSP	2	20	.484	.619	.009
PM15/TSP	2	67	.580	.679	.000
Climatic Region					
PM10/PM15	3	6	.756	.890	.164
PM10/TSP	3	13	.431	.540	.006
PM15/TSP	3	44	.512	.629	.002
EPA Region					
PM10/PM15	6	3	.786	1.014	.043
PM10/TSP	10	4	.404	.614	.044
PM15/TSP	10	13	.482	.701	.000
State					
PM10/PM15	8	2	.756	1.014	.124
PM10/TSP	20	2	.301	.686	.044
PM15/TSP	41	3	.428	.815	.020
City					
PM10/PM15	9	2	.756	1.014	.203
PM10/TSP	31	1	.280	.686	.005
PM15/TSP	89	1	.406	1.054	.016
Site					
PM10/PM15	10	2	.740	1.014	.203
PM10/TSP	31	1	.280	.686	.005
PM15/TSP	98	1	.374	1.054	.000

* ANOVA not performed because, of the two groups, one has only one ratio.

evident, but the West Coast and arid climatic regions are represented by only two site-years (one site) each. Significant differences are observed for the other two kinds of ratios, with the same pattern as observed in the 24-hour ratio distributions: ratios of arid regions are the lowest, and ratios of non-arid regions outside the West Coast are the highest.

Differences by EPA region are shown in Figure 4-7. The ANOVA test for these differences is barely significant for annual average PM10/PM15 and PM10/TSP ratios; there are very few annual average ratios for most of the regions. For PM15/TSP, however, at least seven site-years are represented in each group (representing at least six sites) and significant differences among the distributions by region are evident (Table 4-3).

Differences between states, cities, or sites are not significant for PM10/PM15 ratios, but are significant for PM10/TSP and PM15/TSP ratios. As with the 24-hour ratios, because of the paucity of sites per state or city, all three tests are more a comparison among sites than among geographical units.

ANNUAL AVERAGE RATIOS DIFFERENTIATED BY HIGH TSP CONCENTRATION

Table 4-4 contains the ANOVA results for annual average ratios for site-years with high TSP concentrations (annual average TSP $\geq$ 55 $\mu g/m^3$). As for the undifferentiated set of annual ratios, there are no significant yearly differences. There are differences in the ANOVA comparisons of Tables 4-3 and 4-4. In Table 4-4 significant climatic differences appear for all ratios, whereas in Table 4-3 no significant climatic differences are evident for the PM10/PM15 ratio. There are other differences in the ANOVA results for the geographical groupings of EPA region, state, and city, but these groupings are composed of so few site-years that the results are difficult to interpret. Finally, site-to-site differences are still significant for all three kinds of ratios.

TABLE 4-4. Differences among high TSP annual ratios by groups.

Grouping Factor	No. of Groups	Average No. Obs. Per Group	Group Average		ANOVA Tail Prob.
			Minimum	Maximum	
Year					
PM10/PM15	2	8	.854	.854	.985
PM10/TSP	2	18	.472	.522	.186
PM15/TSP	2	46	.567	.611	.053
Climatic Region					
PM10/PM15	3	5	.756	.879	.027
PM10/TSP	3	12	.425	.520	.019
PM15/TSP	3	30	.494	.592	.014
EPA Region					
PM10/PM15	6	3	.786	.921	.006
PM10/TSP	10	4	.404	.588	.338
PM15/TSP	10	9	.486	.665	.017
State					
PM10/PM15	7	2	.756	.921	.006
PM10/TSP	18	2	.301	.588	.209
PM15/TSP	34	3	.428	.715	.051
City					
PM10/PM15	8	2	.756	.937	.012
PM10/TSP	29	1	.280	.652	.011
PM15/TSP	64	1	.427	.750	.020
Site					
PM10/PM15	8	2	.756	.937	.012
PM10/TSP	28	1	.280	.652	.011
PM15/TSP	69	1	.326	.835	.026

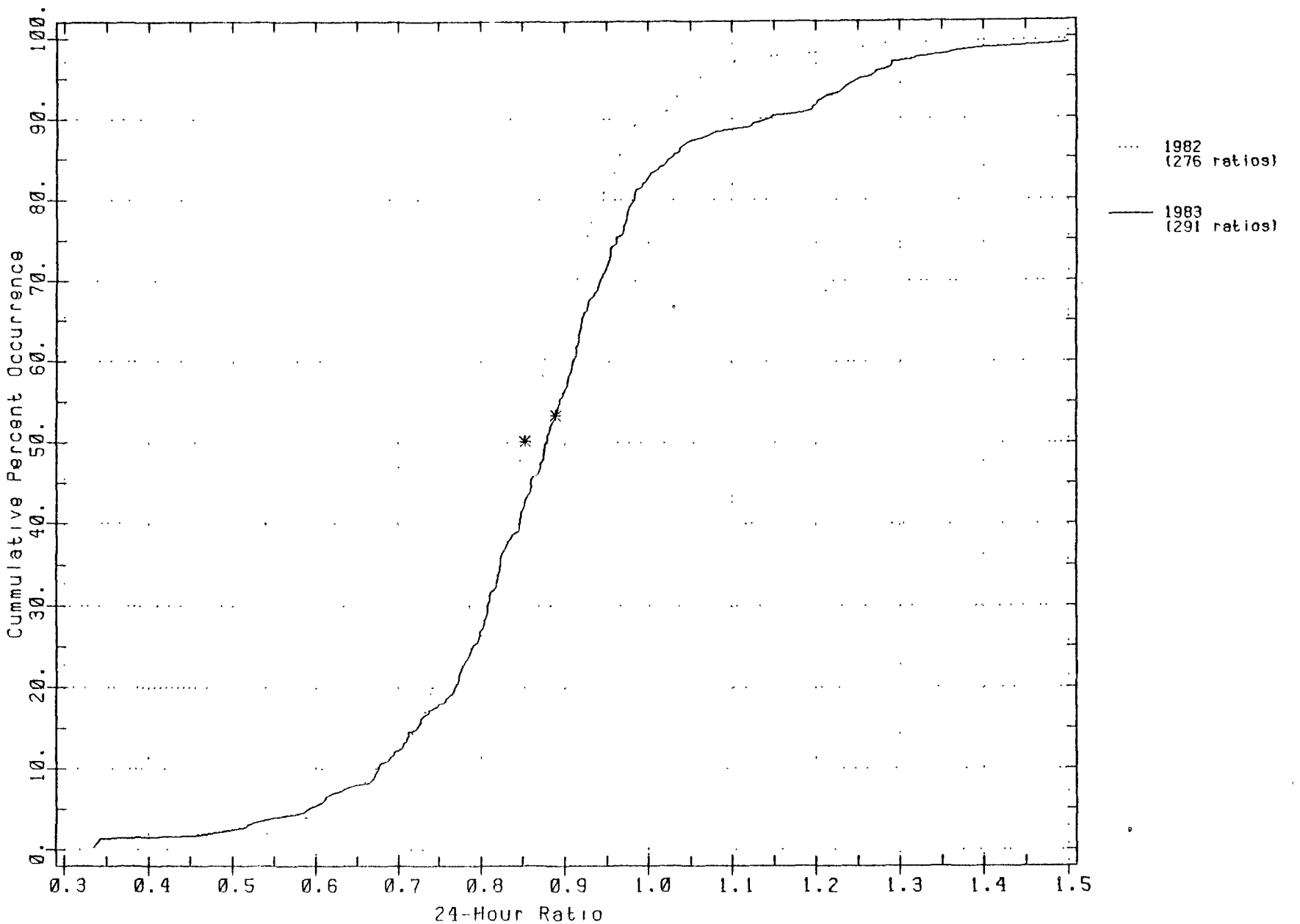

FIGURE 4-1a. Cumulative distributions of 24-hour PM10/PM15 ratios derived from the IP network, 1982-1983, by year.

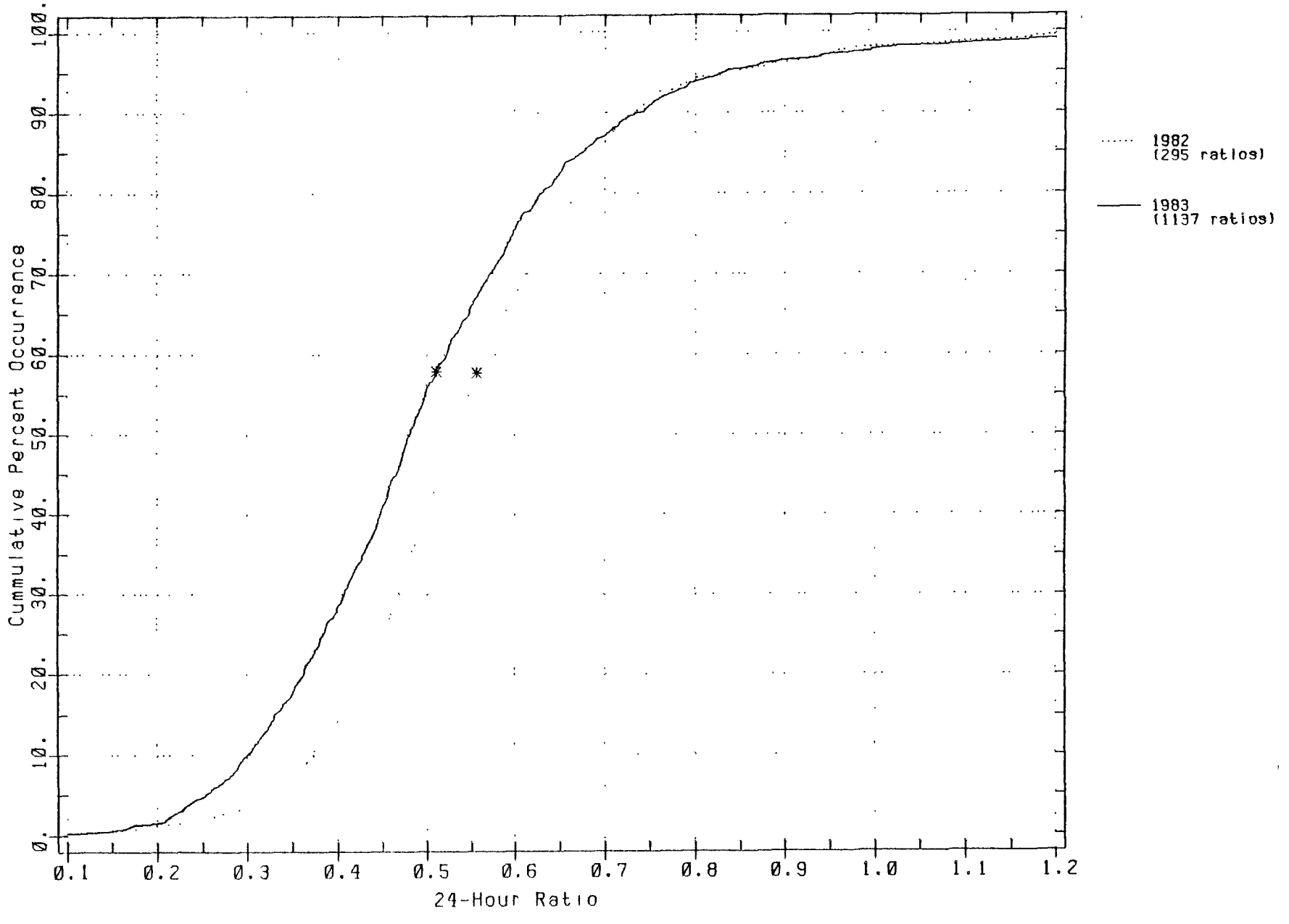

FIGURE 4-1b. Cumulative distributions of 24-hour PM10/TSP ratios derived from the IP network, 1982-1983, by year.

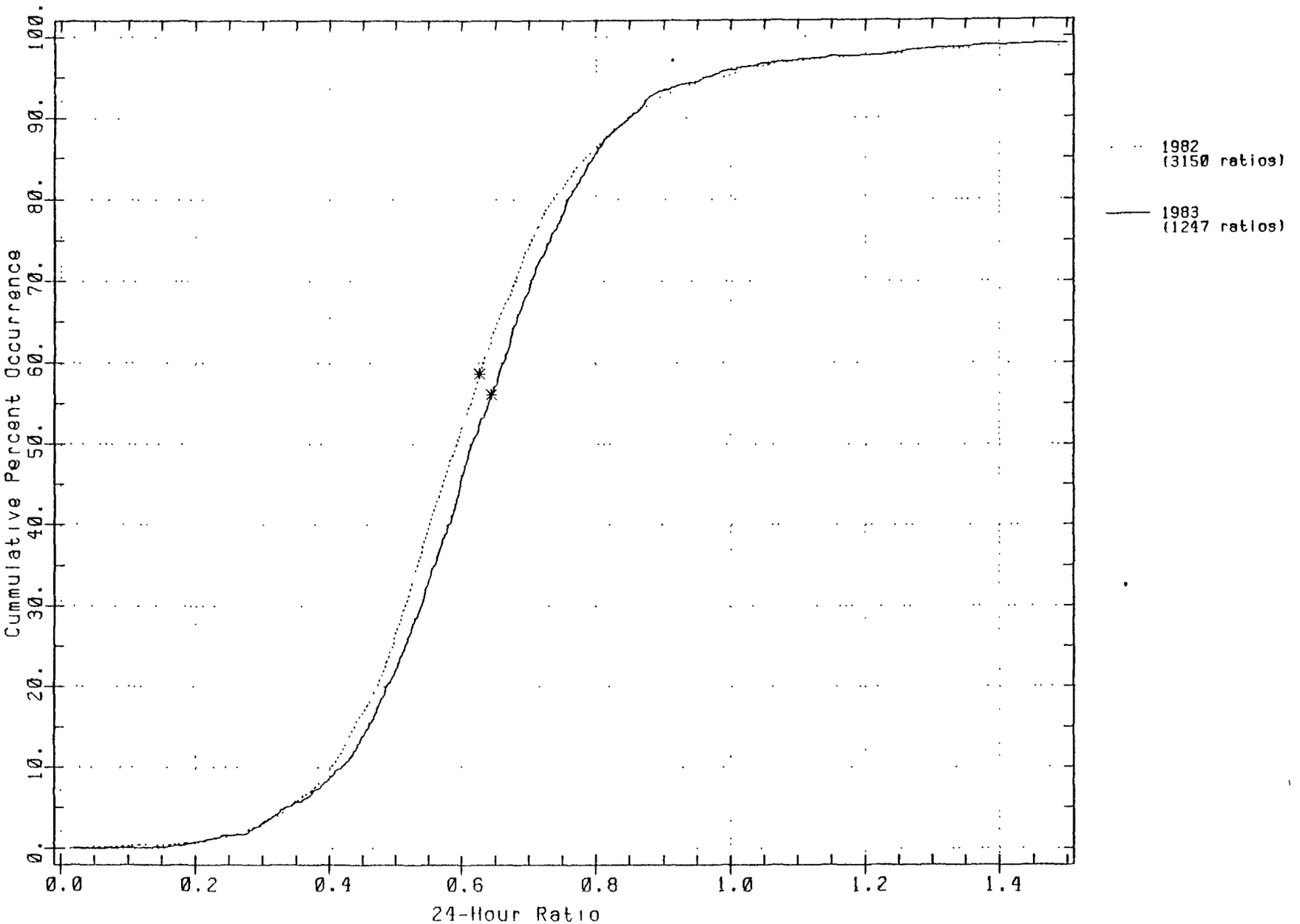

FIGURE 4-1c. Cumulative distributions of 24-hour PM15/TSP ratios derived from the IP network, 1982-1983, by year.

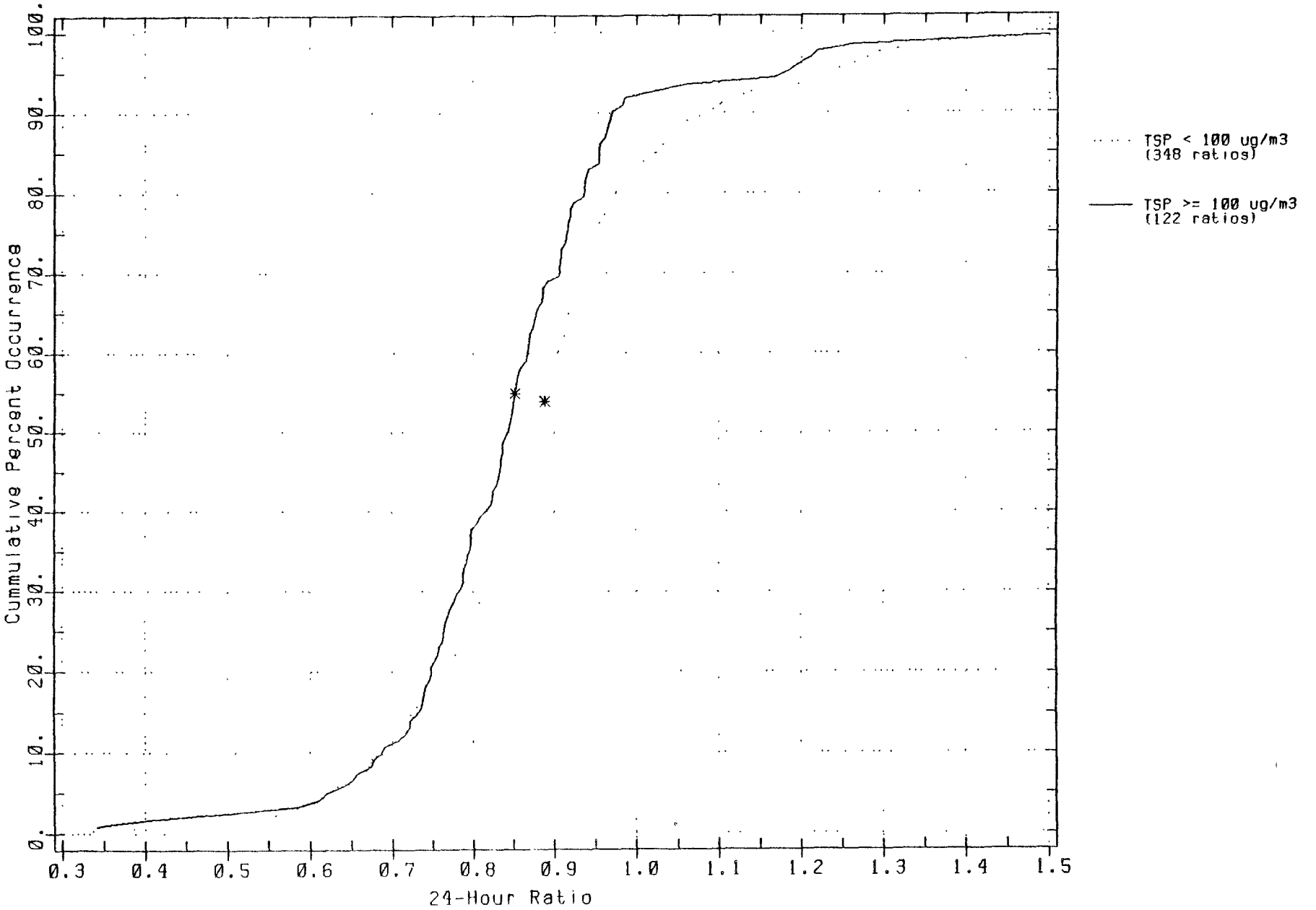

FIGURE 4-2a. Cumulative distributions of 24-hour PM10/PM15 ratios derived from the IP network, 1982-1983, by TSP value.

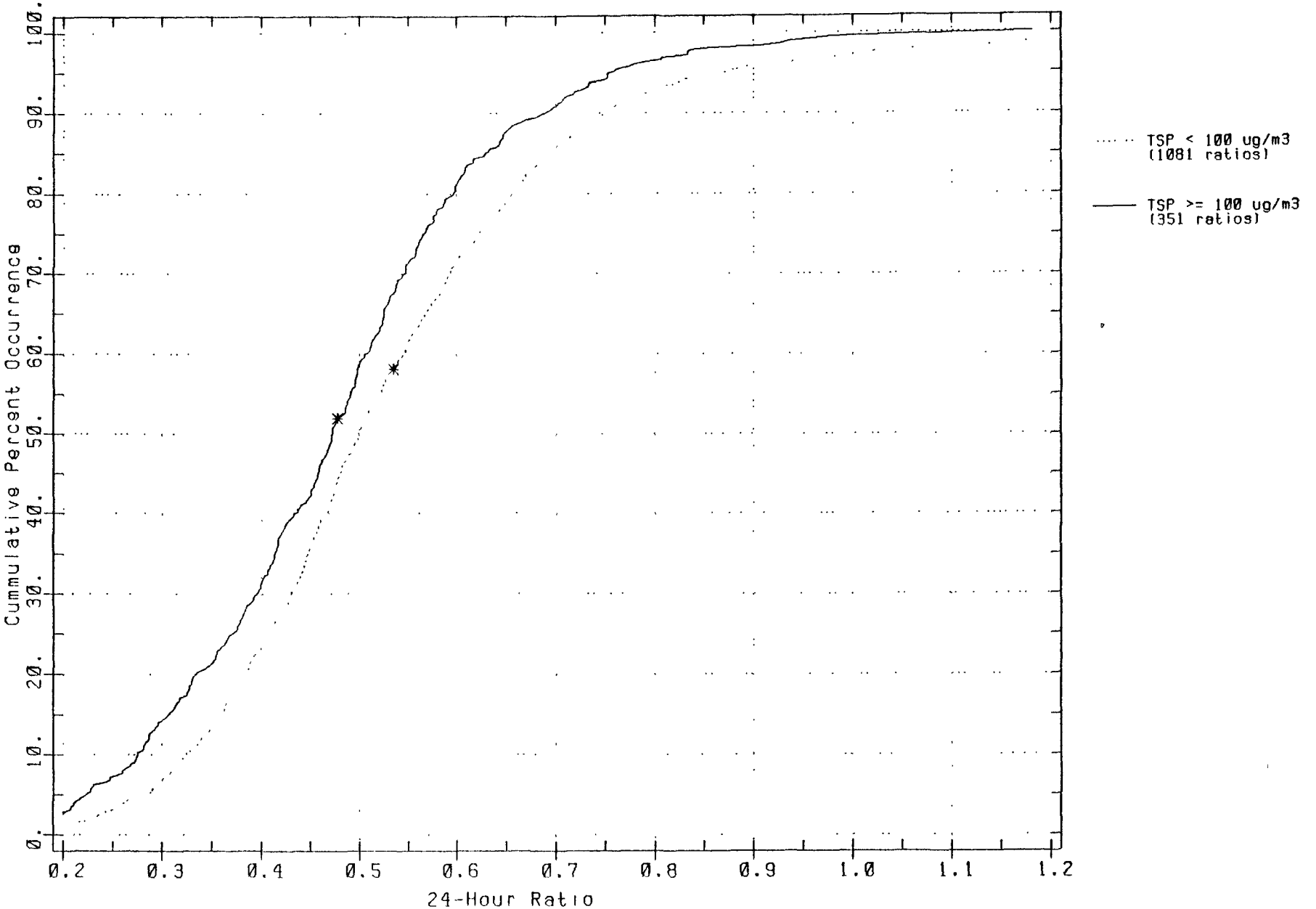

FIGURE 4-2b. Cumulative distributions of 24-hour PM10/TSP ratios derived from the IP network, 1982-1983, by TSP value.

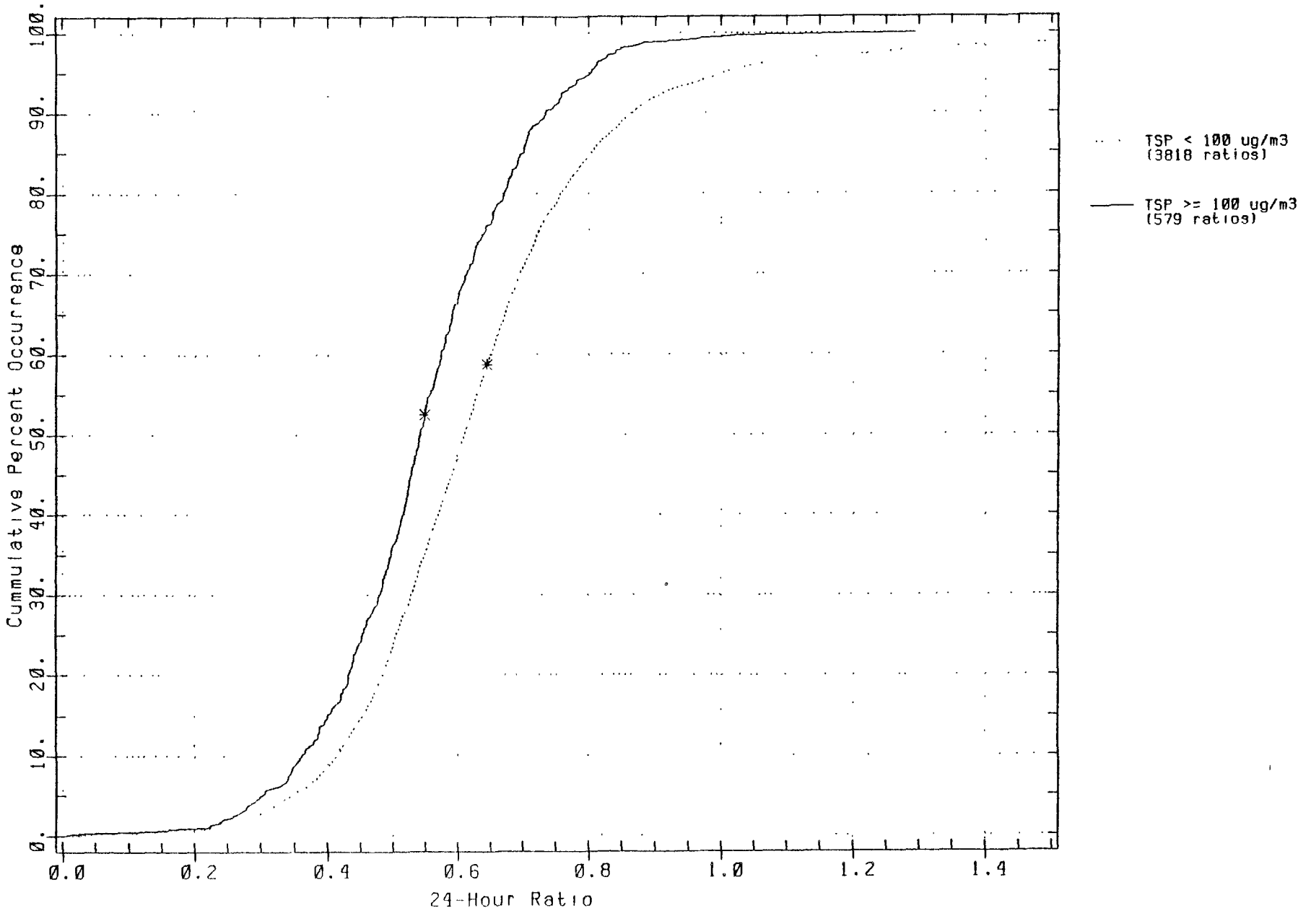

FIGURE 4-2c. Cumulative distributions of 24-hour PM15/TSP ratios derived from the IP network, 1982-1983, by TSP value.

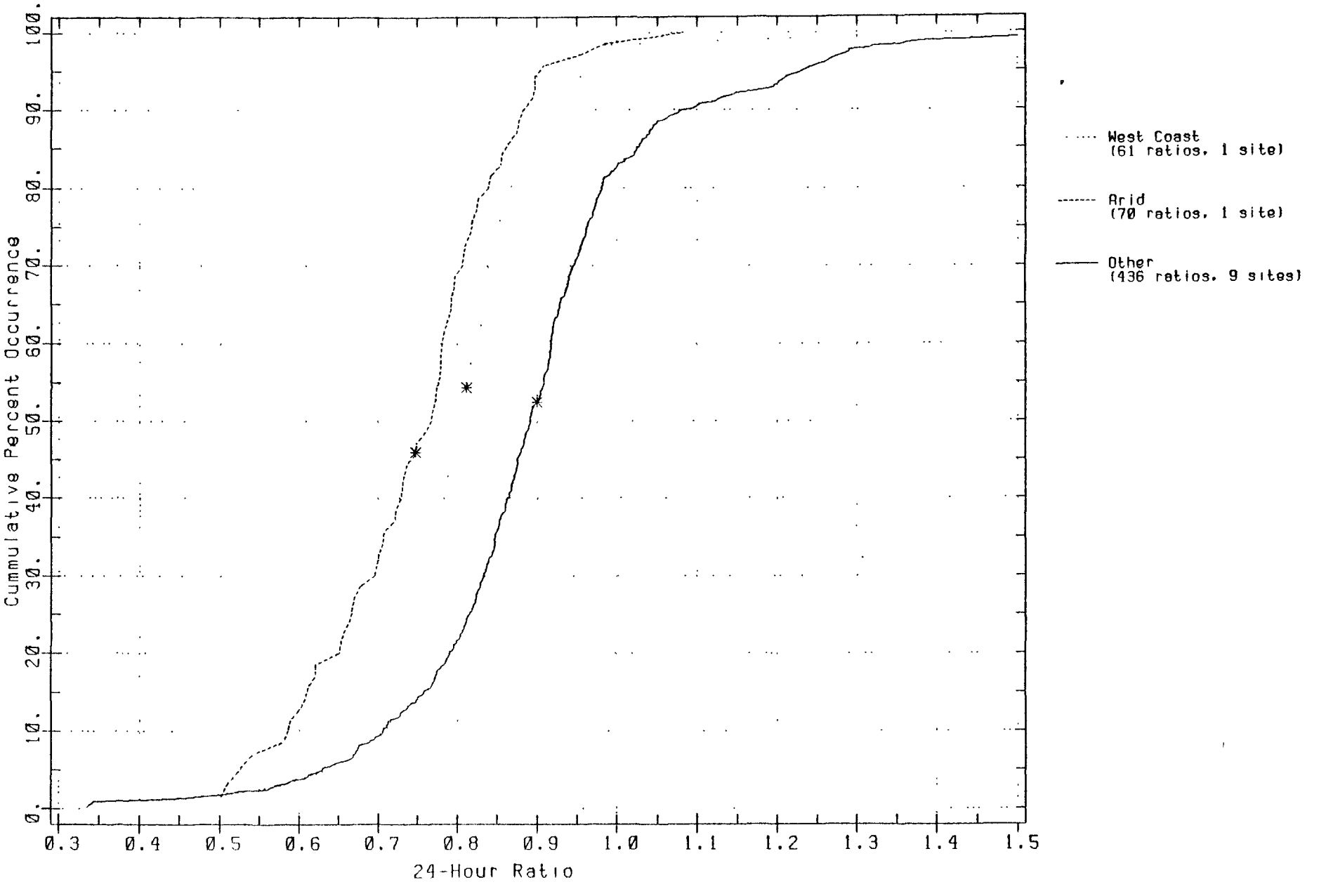

FIGURE 4-3a. Cumulative distributions of 24-hour PM10/PM15 ratios derived from the IP network, 1982-1983, by climatic region.

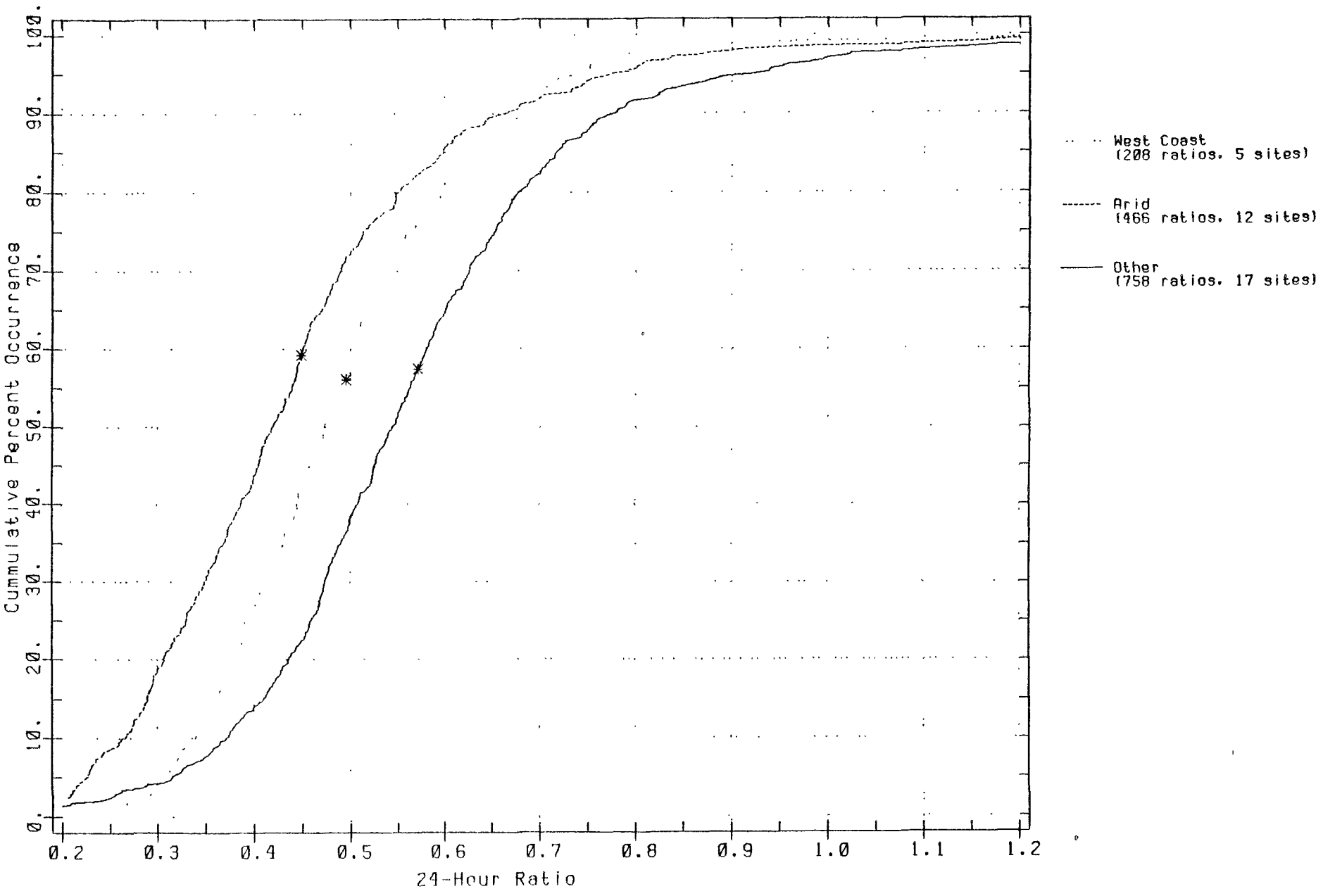

FIGURE 4-3b. Cumulative distributions of 24-hour PM10/TSP ratios derived from the IP network, 1982-1983, by climatic region.

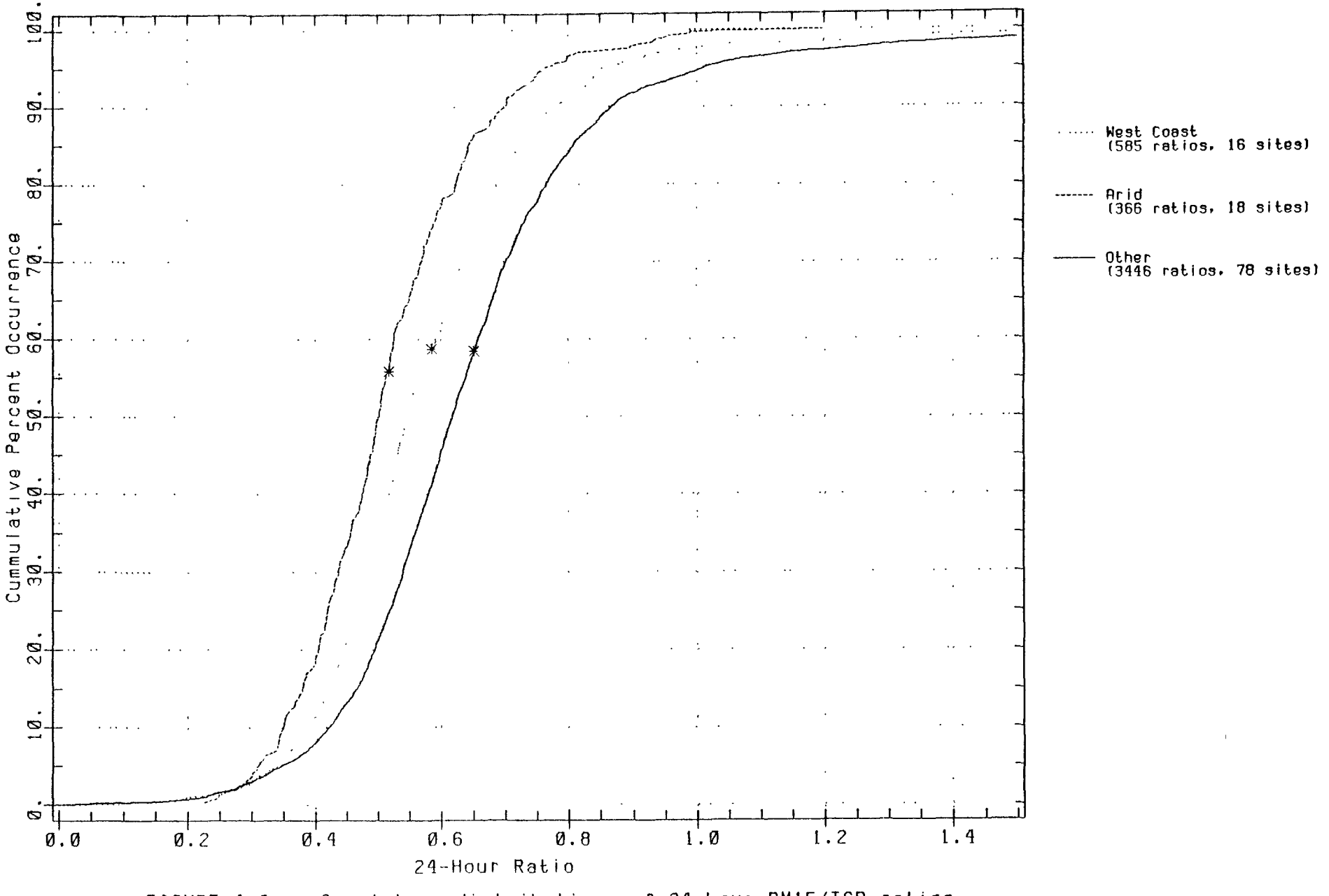

FIGURE 4-3c. Cumulative distributions of 24-hour PM15/TSP ratios derived from the IP network, 1982-1983, by climatic region.

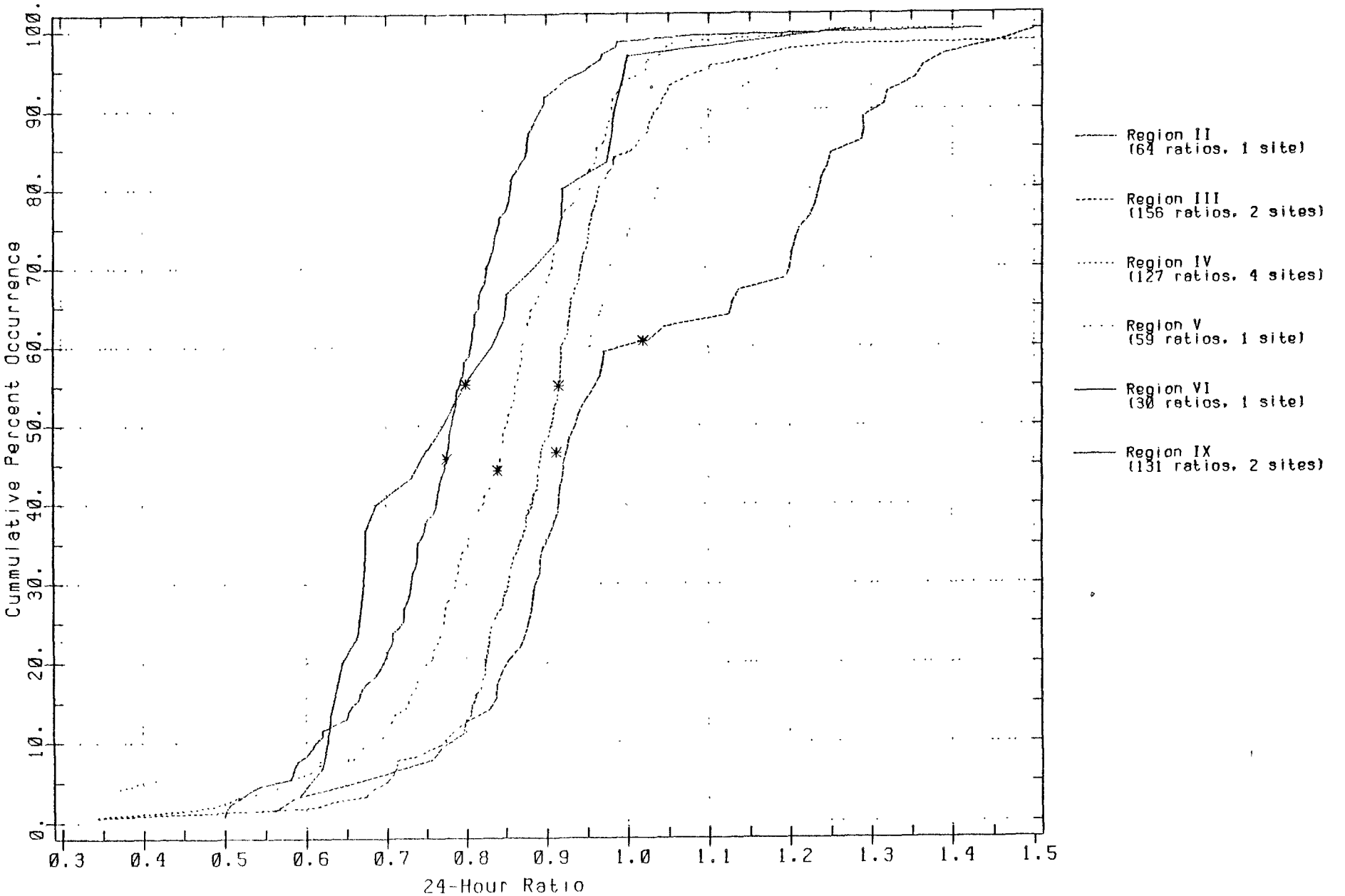

FIGURE 4-4a. Cumulative distributions of 24-hour PM10/PM15 ratios derived from the IP network, 1982-1983, by EPA region.

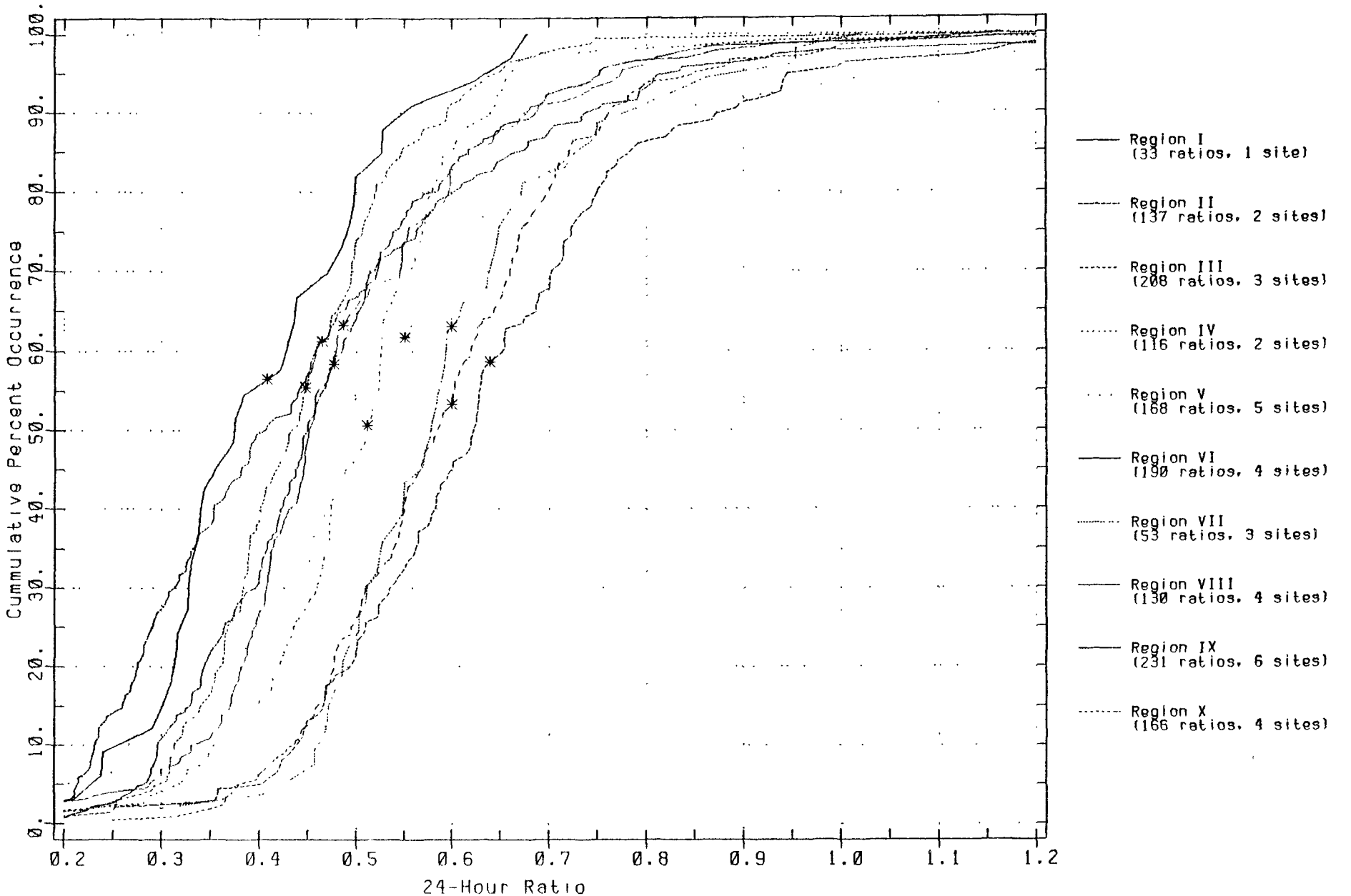

FIGURE 4-4b. Cumulative distributions of 24-hour PM10/TSP ratios derived from the IP network, 1982-1983, by EPA region.

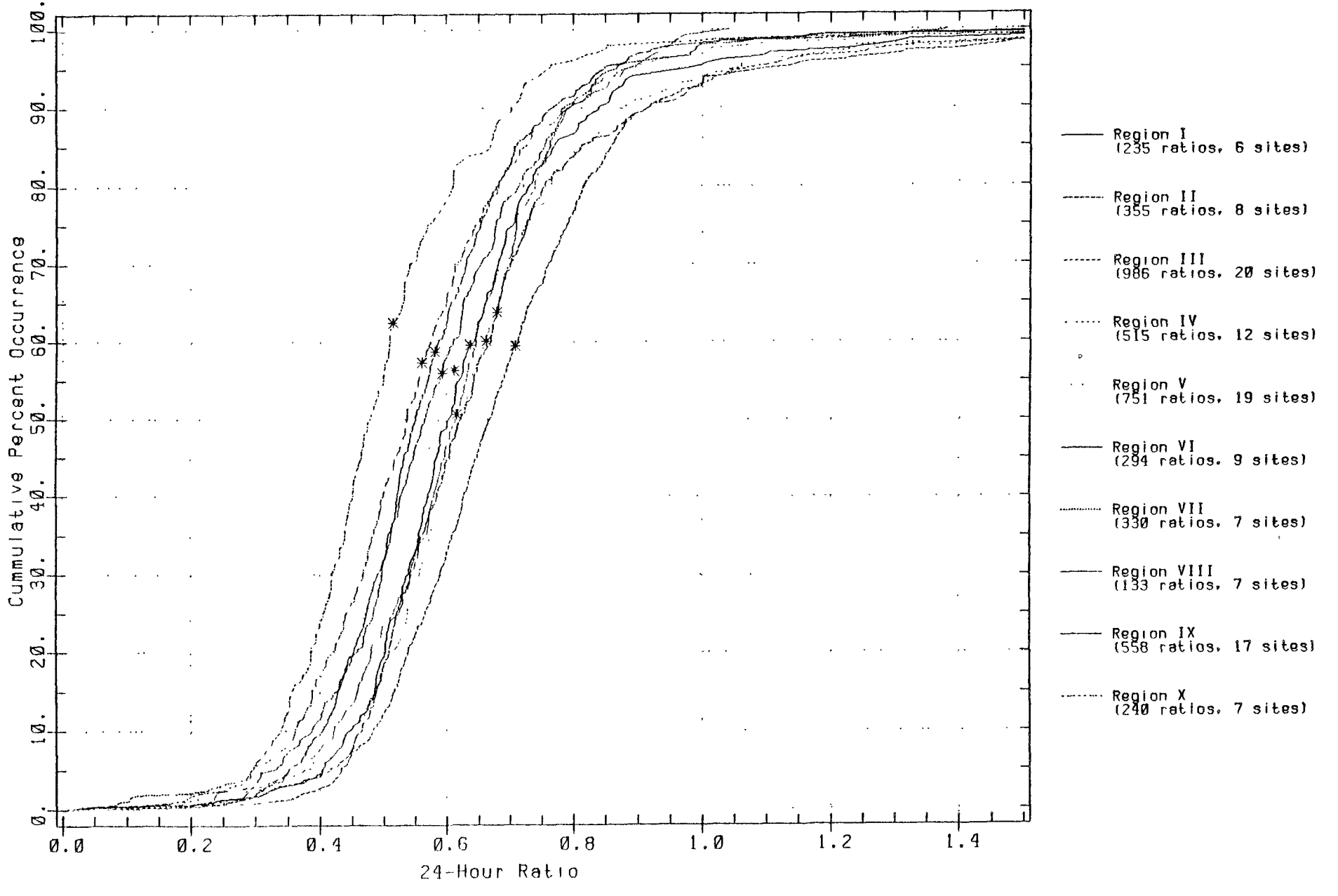

FIGURE 4-4c. Cumulative distributions of 24-hour PM15/TSP ratios derived from the IP network, 1982-1983, by EPA region.

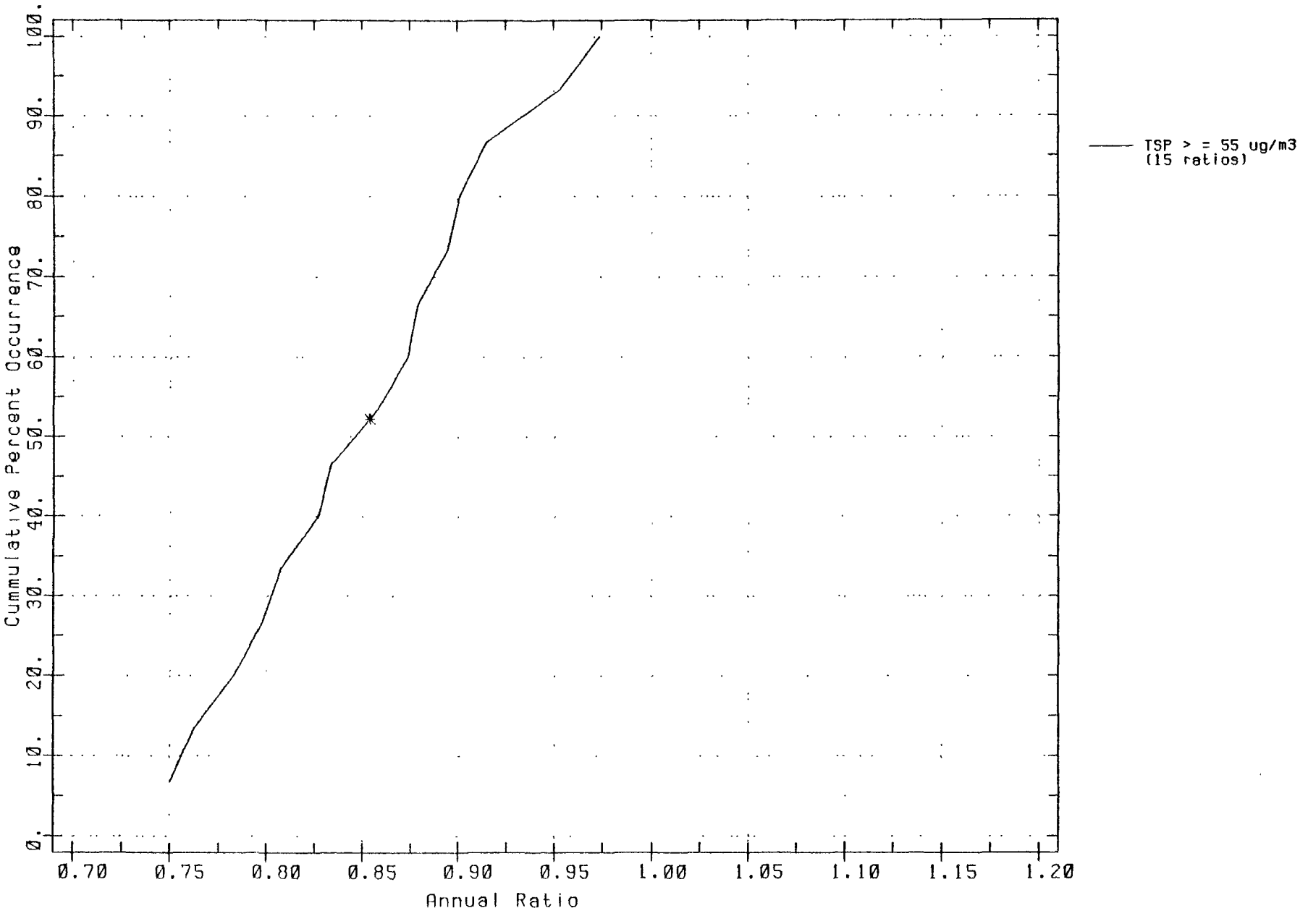

FIGURE 4-5a. Cumulative distributions of annual PM10/PM15 ratios derived from the IP network, 1982-1983, by TSP value.

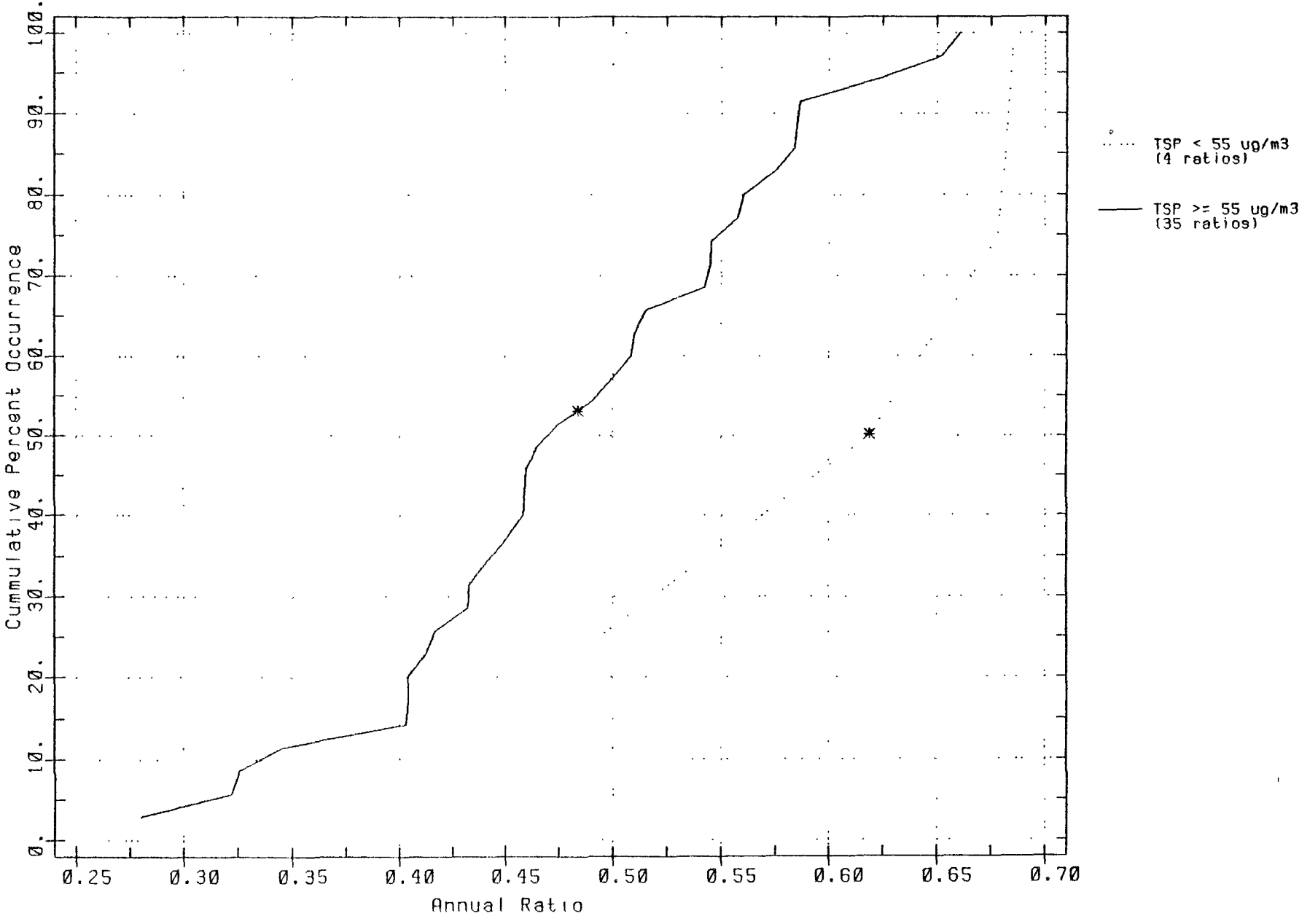

FIGURE 4-5b. Cumulative distributions of annual PM10/TSP ratios derived from the IP network, 1982-1983, by TSP value.

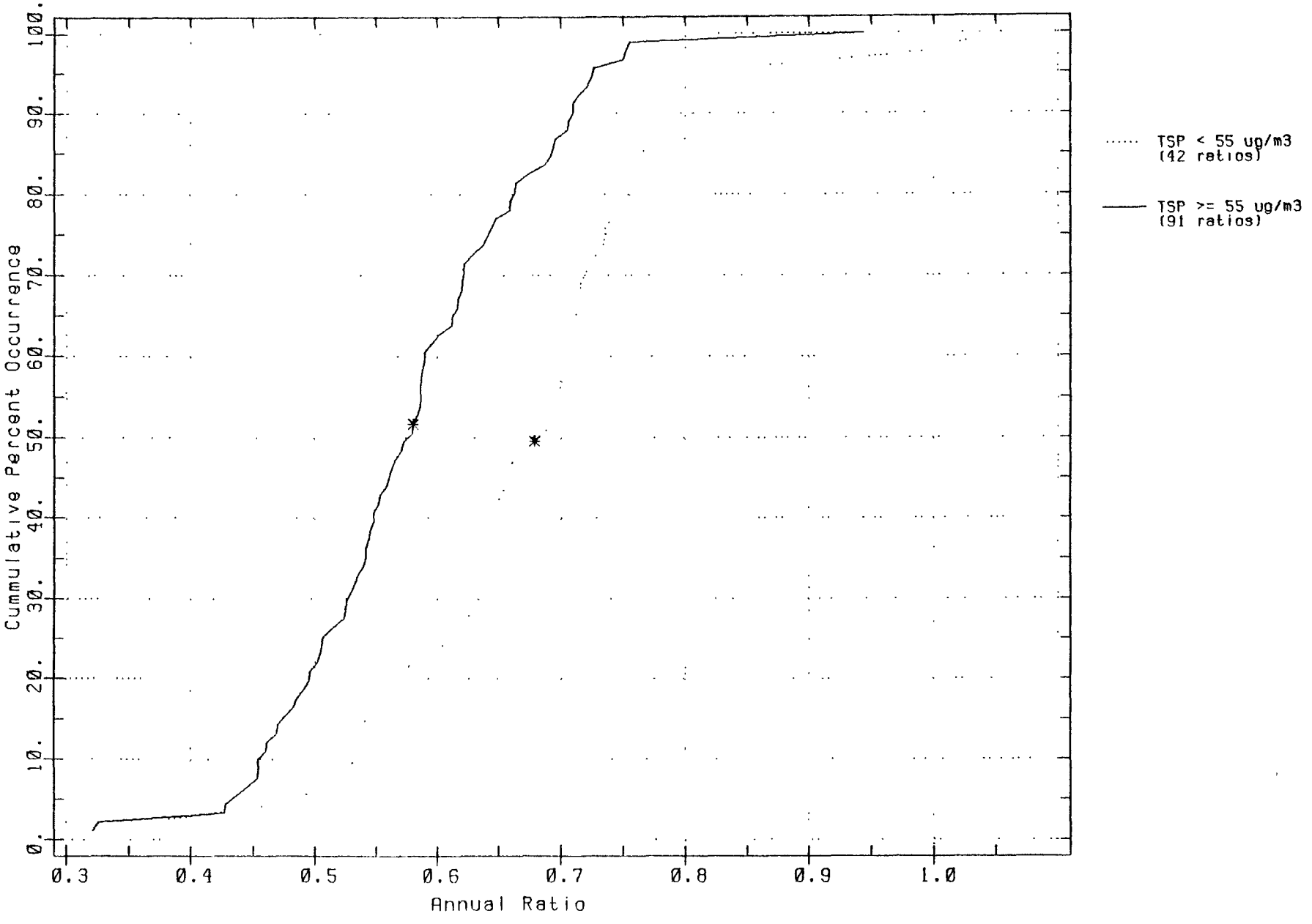

FIGURE 4-5c. Cumulative distributions of annual PM15/TSP ratios derived from the IP network, 1982-1983, by TSP value.

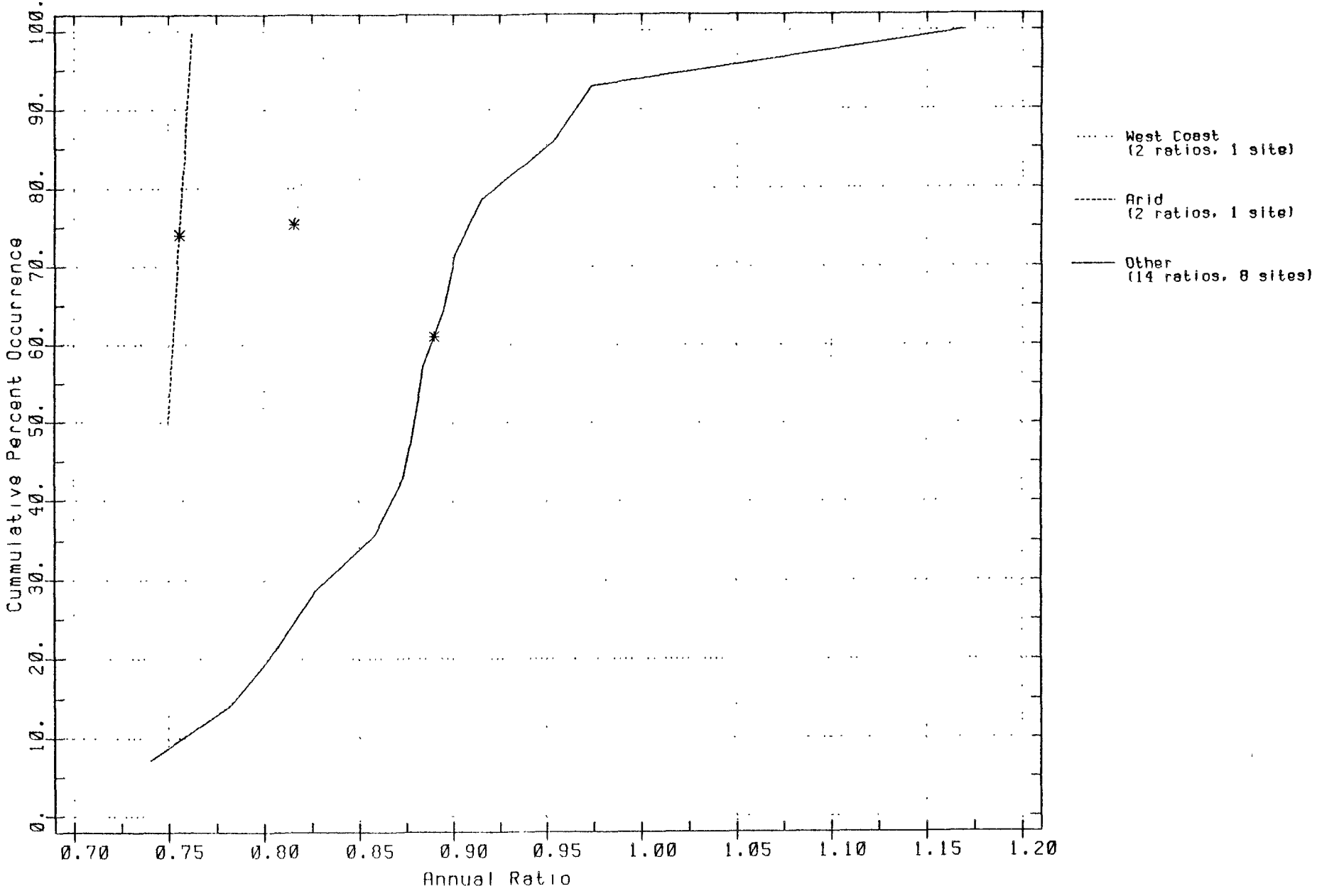

FIGURE 4-6a. Cumulative distributions of annual PM10/PM15 ratios derived from the IP network, 1982-1983, by climatic region.

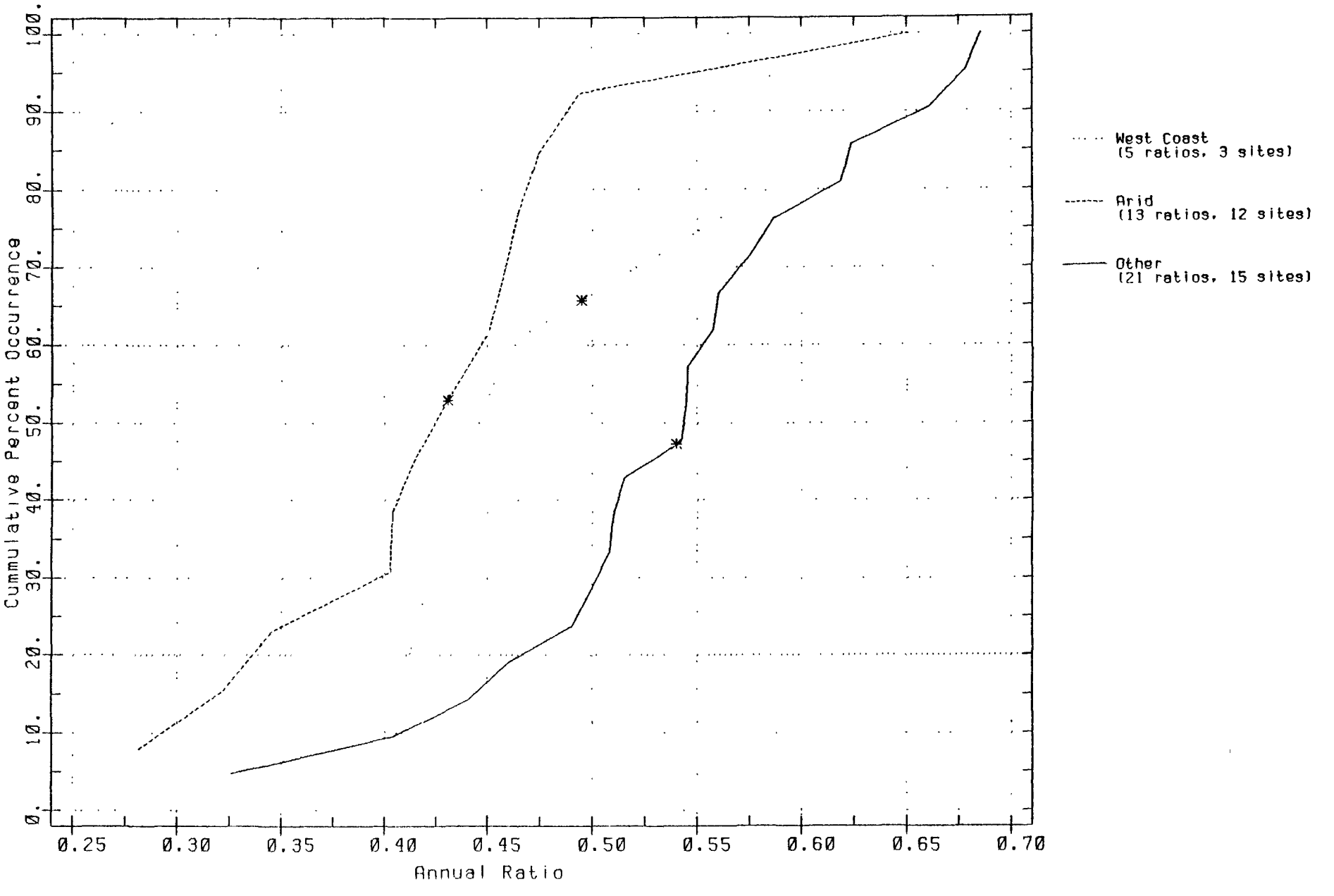

FIGURE 4-6b. Cumulative distributions of annual PM10/TSP ratios derived from the IP network, 1982-1983, by climatic region.

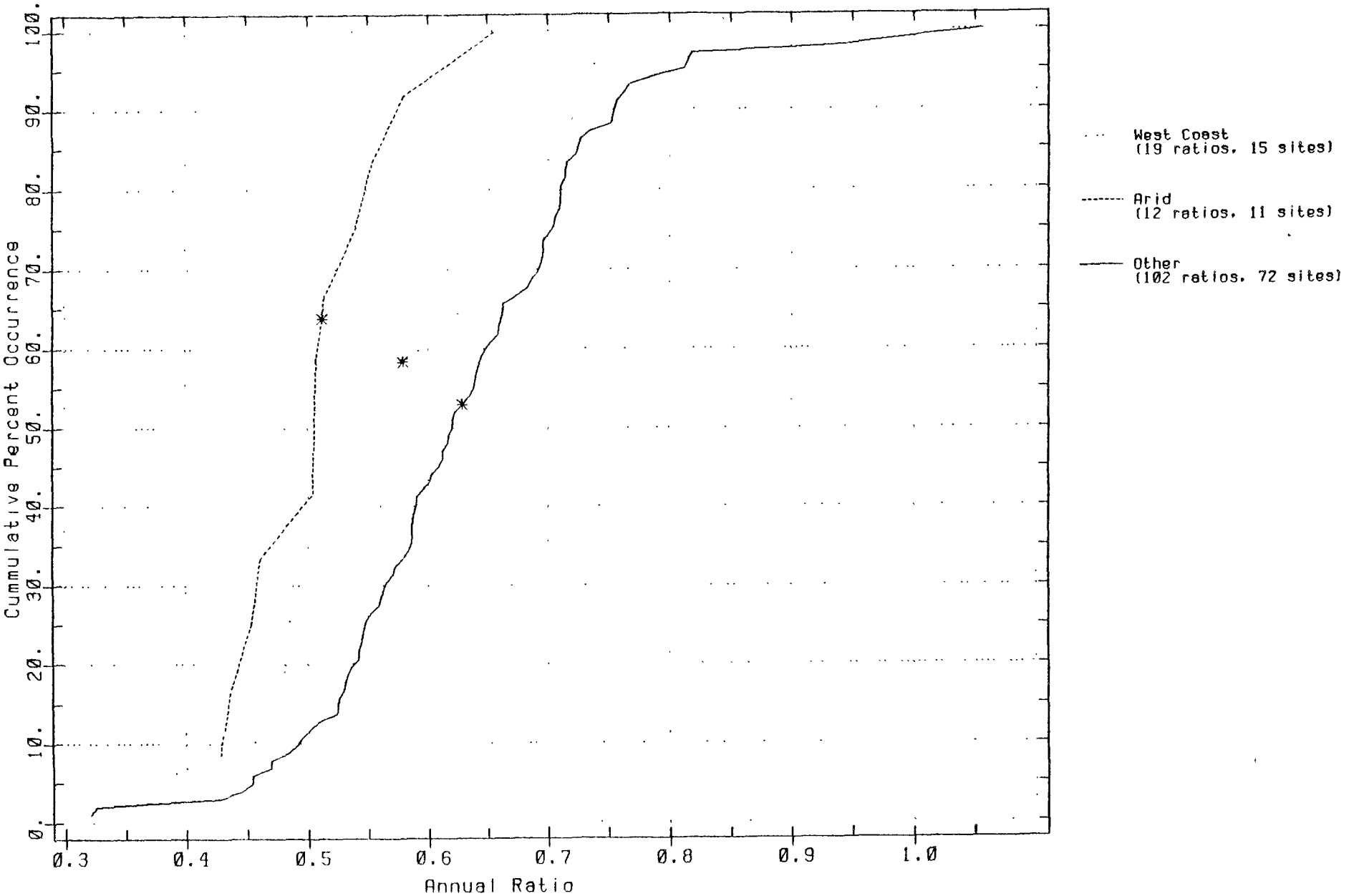

FIGURE 4-6c. Cumulative distributions of annual PM15/TSP ratios derived from the IP network, 1982-1983, by climatic region.

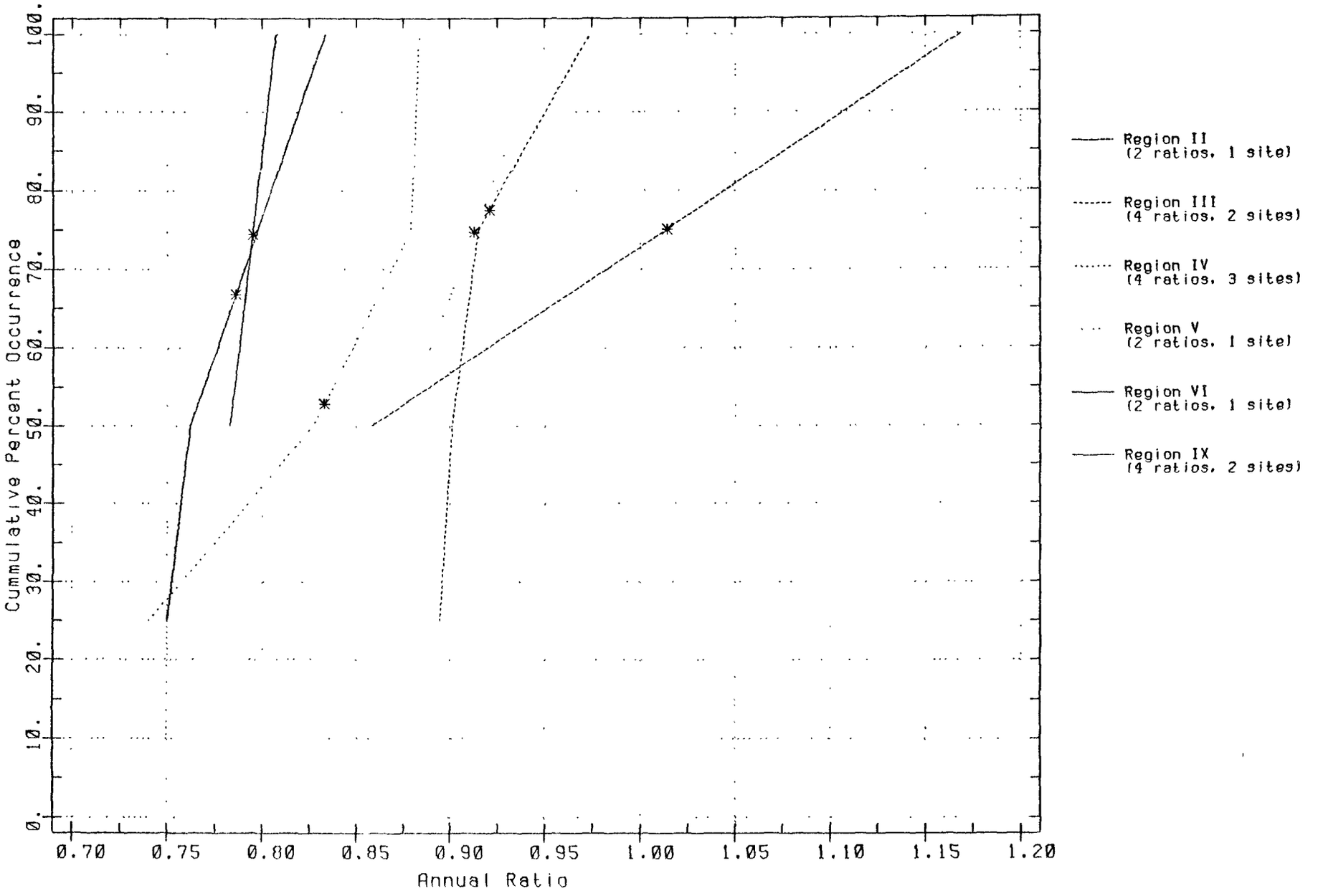

FIGURE 4-7a. Cumulative distributions of annual PM10/PM15 ratios derived from the IP network, 1982-1983, by EPA region.

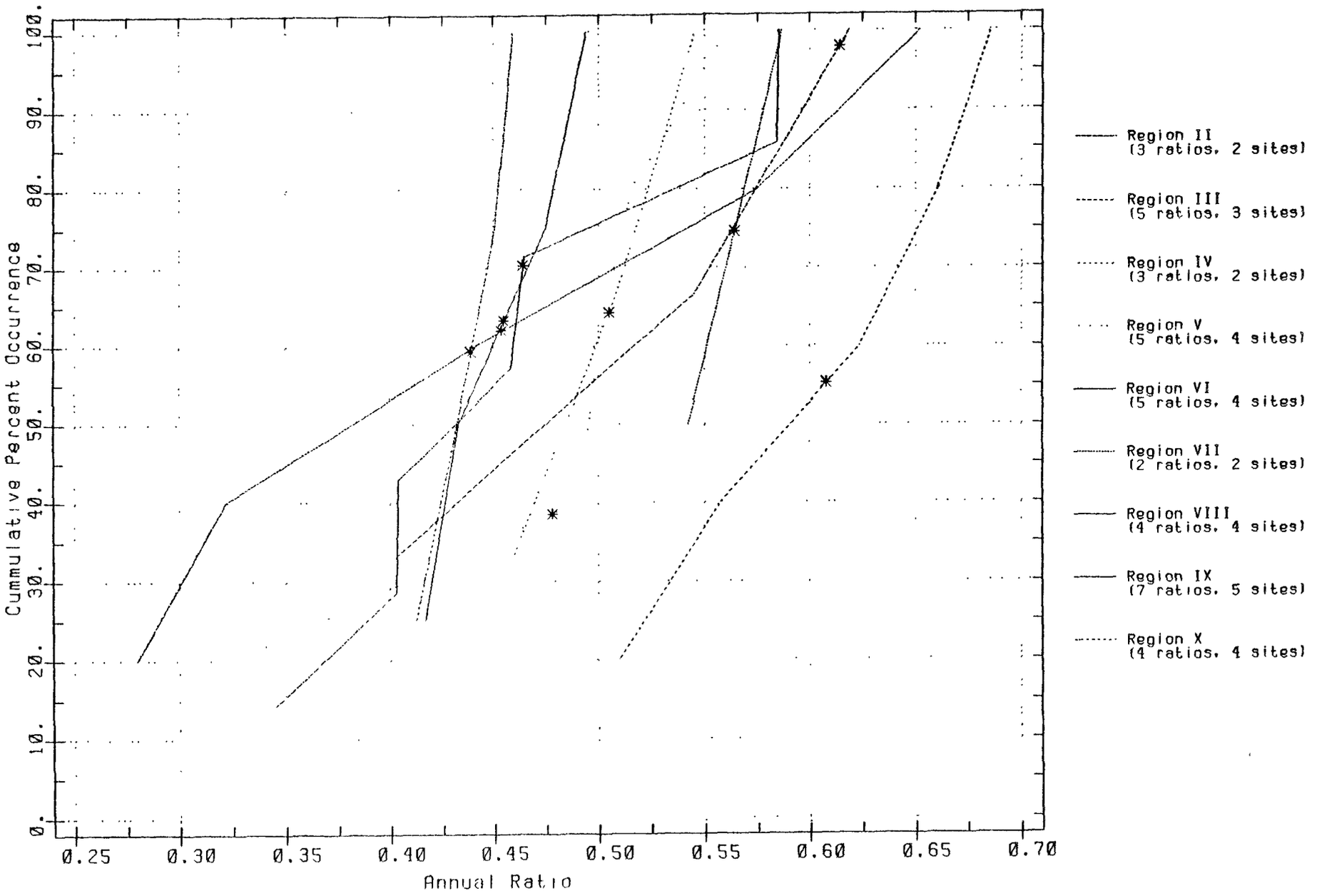

FIGURE 4-7b. Cumulative distributions of annual PM10/TSP ratios derived from the IP network, 1982-1983, by EPA region.

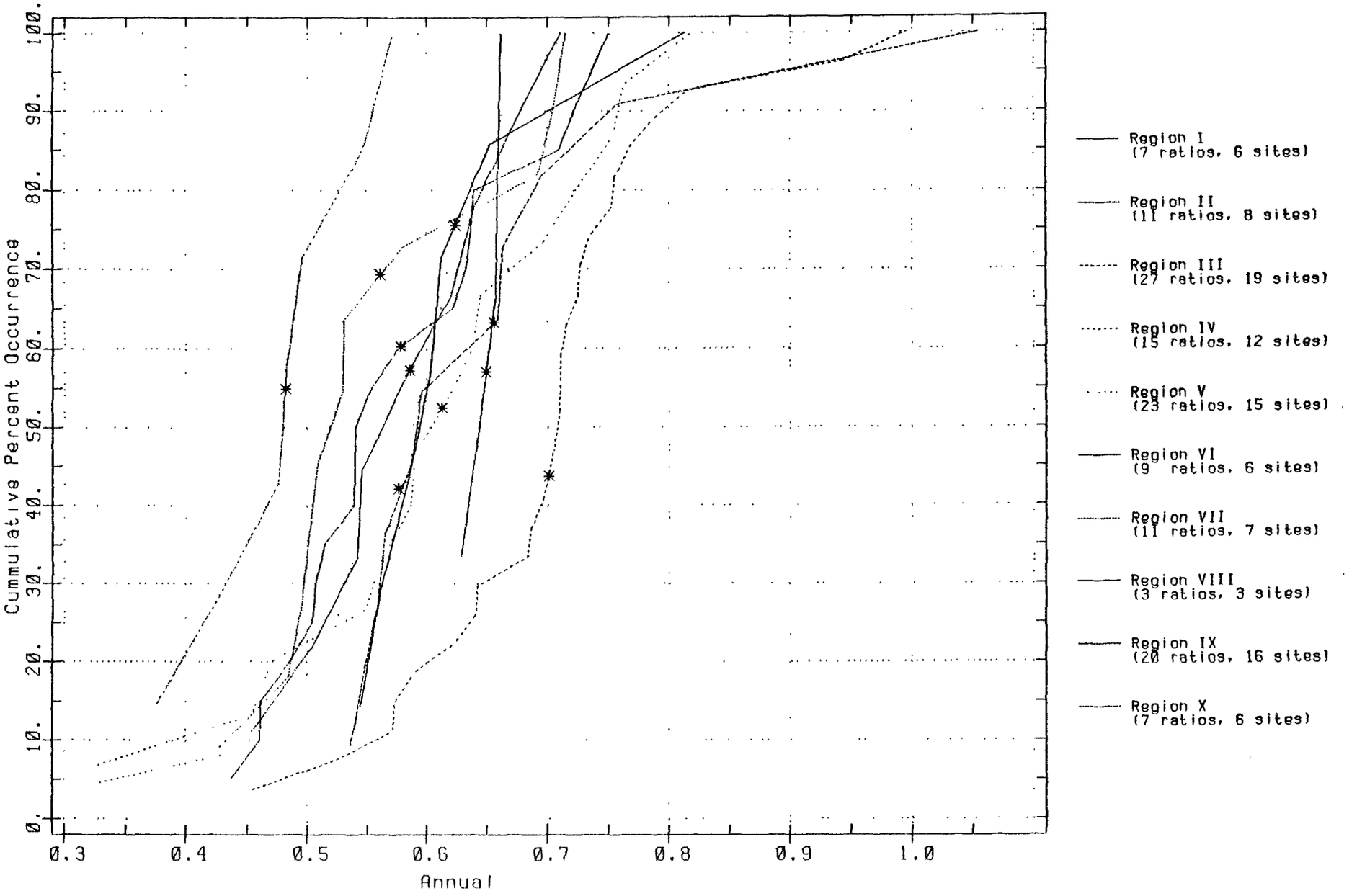

FIGURE 4-7c. Cumulative distributions of annual PM15/TSP ratios derived from the IP network, 1982-1983, by EPA region.

5 APPLICATION OF CROSS-VALIDATION TECHNIQUES TO ESTIMATE MISCLASSIFICATION RATES

In the previous section we examined distributions for different groups of ratios. There were some statistically significant differences among distributions by geographical groupings, e.g., there appeared to be significant differences in both 24-hour and annual average ratio distributions among the three climatic regions. In this section we address the practical differences between using a national ratio distribution and separate ratio distributions for different climatic regions, under the assumption that separate climatic distributions are indeed valid.

Different ratio distributions can be assessed by using them to predict the likelihood of attainment and comparing the predicted attainment probability with actual attainment status. There are several limitations, however, in following such a procedure with the available IP network data. First of all, the 24-hour standard is an expected annual exceedance standard, with attainment normally demonstrated with three years of monitoring data. For the IP network there are only two years of PM10 data available, and in fact only one year for many of the sites. We therefore consider a one-year attainment test for the 24-hour standard.

The second problem is that, because valid particulate matter samples are not available for every day of the year, we cannot know the actual attainment status. A simple approach for determining attainment status for all site years was chosen. For the annual standard the annual mean is estimated from available sampling days; for the 24-hour standard, since no site year has more than 61 valid sampling days, no exceedances are allowed, and

so the 24-hour standard is considered failed if the maximum PM10 concentration in the year exceeds the standard. Otherwise the standard is considered passed.

The third problem is the limited number of sites in each geographical region. No cross-validation analyses could be performed on PM10/PM15 ratios; for climate types, for example, there was only one arid site and one West Coast site with any PM10/PM15 ratios. For PM10/TSP and PM15/TSP ratios, only climatic and EPA regional distributions could be compared for the 24-hour standard and only climatic distributions could be examined for the annual standard.

The only particulate matter data available to estimate misclassification rates from different ratio distributions are the actual data used to develop the ratio distributions. This presents a statistical problem: a statistical model should not be validated with the data used to develop the model because the results will be biased; the model will most likely perform better than it would on a different data set. To avoid this problem, we use a simple cross-validation procedure, the "leave-out-one" technique, first described by Mosteller and Tukey (1968). Conceptually, the procedure is to develop a distinct ratio distribution for each site; the distribution contains data from all sites except the one under consideration. This distribution is then used to estimate the probability of nonattainment at the site, and the nonattainment probability is compared to the attainment status of the site, as determined from monitoring data at the site. The comparison is made for each site, and the results are used to evaluate the predictive ability of one type of ratio distribution. Efficient computer algorithms were developed to substantially reduce the number of required calculations.

CROSS-VALIDATION APPLIED TO 24-HOUR RATIO DISTRIBUTIONS

We have chosen three levels of a 24-hour PM10 standard to examine. Although the level of the 24-hour standard is proposed to be between

150 and 250 $\mu g/m^3$ there were so few 24-hour PM10 concentrations above 150 $\mu g/m^3$ that the probability of nonattainment would be nearly zero for most site-years, no matter what ratio distribution was used. The three chosen levels of the standard are 75, 100, and 150 $\mu g/m^3$; this range is wide enough to see if the results of the cross-validation are dependent on the level of the standard.

The probability of nonattainment of a 24-hour 100 $\mu g/m^3$ PM10 standard as calculated from a national PM10/TSP distribution is plotted against the probability calculated from climatic PM10/TSP distributions in Figure 5-1. Each point in the plot is a separate site-year. The plotting symbol indicates the climate type; where plotted points coincide, the number of points in the same plotting position is indicated. On each axis, 20 percent and 95 percent probabilities of nonattainment are indicated. These probability levels are important because they will be used to determine the required frequency of PM10 monitoring for each site. Those sites with low probability of nonattainment, i.e., less than 20 percent, will be required to maintain a PM10 sampling schedule of once every six days. Those sites with a medium probability of nonattainment, i.e., between 20 and 95 percent, will be required to sample every other day; and those sites with high probability of nonattainment, i.e., greater than 95 percent, will be required to sample every day (49 Fed. Reg. 10,437, 1984).

A similar plot comparing nonattainment probabilities derived from a national PM10/TSP distribution with those derived from 10 regional PM10/TSP distributions is contained in Figure 5-2. (Hereafter we will refer to such a plot as a P-P plot.) In both Figures 5-1 and 5-2 the largest differences in nonattainment probabilities appear to be in the medium probability range, i.e., between 20 and 95 percent. There are no site-years with a low probability (less than 20 percent) on one axis and a high probability (greater than 95 percent) on the other.

In Table 5-1 site-years are classified as having a low, medium, or high probability of nonattainment of three 24-hour PM10 standards (75, 100, and 150 $\mu g/m^3$) based on a single national PM10/TSP distribution, on climatic

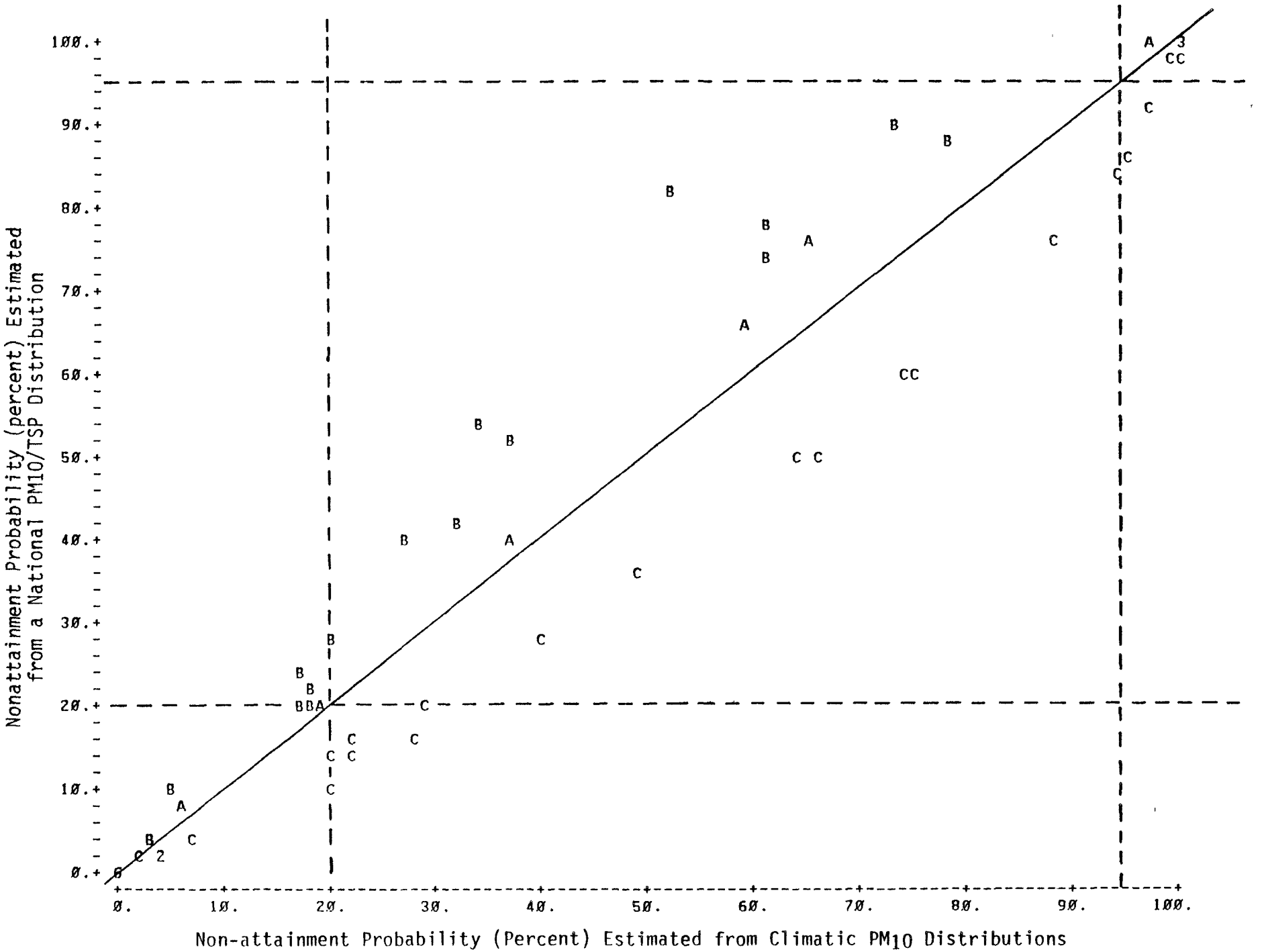

FIGURE 5-1. Probability of non-attainment of a 24-hour 100 $\mu g/m^3$ PM_{10} standard estimated from national and climatic PM_{10}/TSP distributions. A = West Coast, B = arid, C = other.

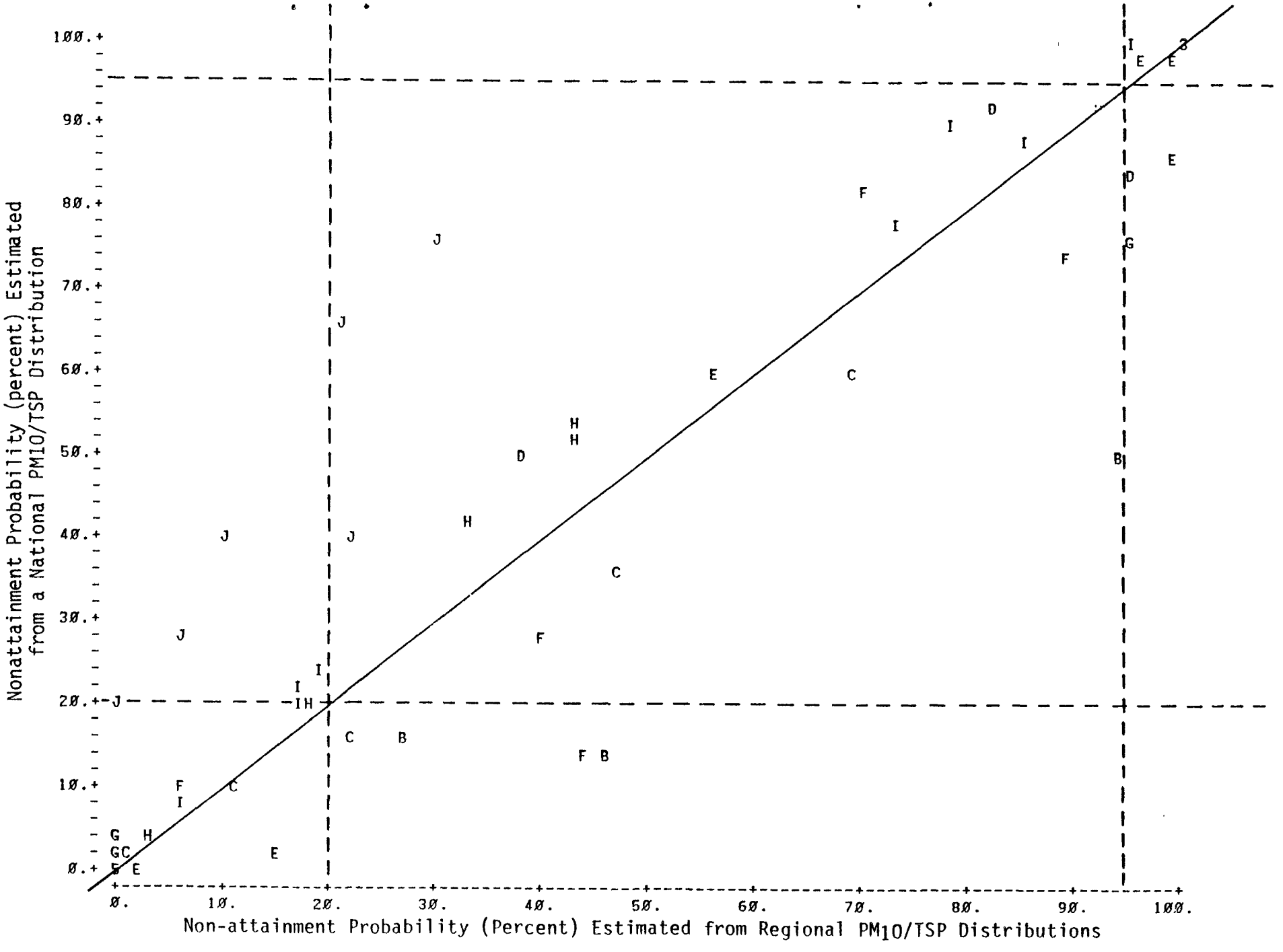

FIGURE 5-2. Probability of non-attainment of a 24-hour 100 $\mu g/m^3$ PM_{10} standard estimated from national and regional PM10/TSP distributions. A = Region I, B = Region II, etc.

TABLE 5-1. Number of sites classified as low, medium, or high probability of nonattainment of three one-year 24-hour PM10 standards compared with actual attainment status: comparison among national, climate, and region PM10/TSP cumulative distribution functions.

Standard		National cdf				Climate cdf's				EPA Region cdf's			
		Low	Med	High		Low	Med	High		Low	Med	High	
75 µg/m^3	Pass	10	9	2	21	8	11	2	21	9	10	1	20
	Fail	1	14	17	32	2	13	17	32	3	15	14	32
		11	23	19	53	10	24	19	53	12	25	15	52
100 µg/m^3	Pass	21	17	0	38	18	20	0	38	20	17	0	37
	Fail	1	8	6	15	1	6	8	15	1	7	7	15
		22	25	6	53	19	26	8	53	21	24	7	52
150 µg/m^3	Pass	42	5	2	49	40	8	1	49	38	8	2	48
	Fail	2	2	0	4	2	2	0	4	2	2	0	4
		44	7	2	53	42	10	1	53	40	10	2	52

distributions, and on regional distributions. Additionally, the site-years are classified into actual attainment and nonattainment of each standard on the basis of the maximum PM10 concentration observed in the site-year. Ideally, one would like to see all nonattainment site-years classified as having a high probability of nonattainment, and all attainment site-years as having a low probability of nonattainment. The worst misclassifications result when the converse occurs, i.e., when attainment sites are classified as having a high probability of nonattainment and nonattainment sites are classified as having a low probability of nonattainment. In the former case the penalty is that unnecessary additional monitoring as well as a State Implementation Plan within nine months will be required, and in the latter case not enough monitoring will be required and exceedance days may not be detected.

The number of nonattainment sites incorrectly classified as having a low probability of nonattainment, for each standard, is no higher with a national distribution than with climatic or regional distributions. In fact, under a 75 $\mu g/m^3$ standard the national distribution performs better than climatic or regional distributions. The national distribution also performs well if one considers the number of nonattainment or attainment site-years classified as having a high or medium probability of nonattainment.

Figures 5-3 and 5-4 show P-P plots of nonattainment probabilities for a 24-hour 100 $\mu g/m^3$ standard calculated from a national PM15/TSP distribution versus climatic and regional distributions, respectively. The patterns are similar to those observed in Figures 5-1 and 5-2, with far more site-years of data available for the PM15/TSP distribution. These nonattainment probabilities are classified as low, medium, or high in Table 5-2, where the classifications derived from the national distribution are compared to those derived from the climatic and regional distributions for the three 24-hour standards. Actual misclassifications were not calculated, because there were not enough site-days on which both a PM15/TSP

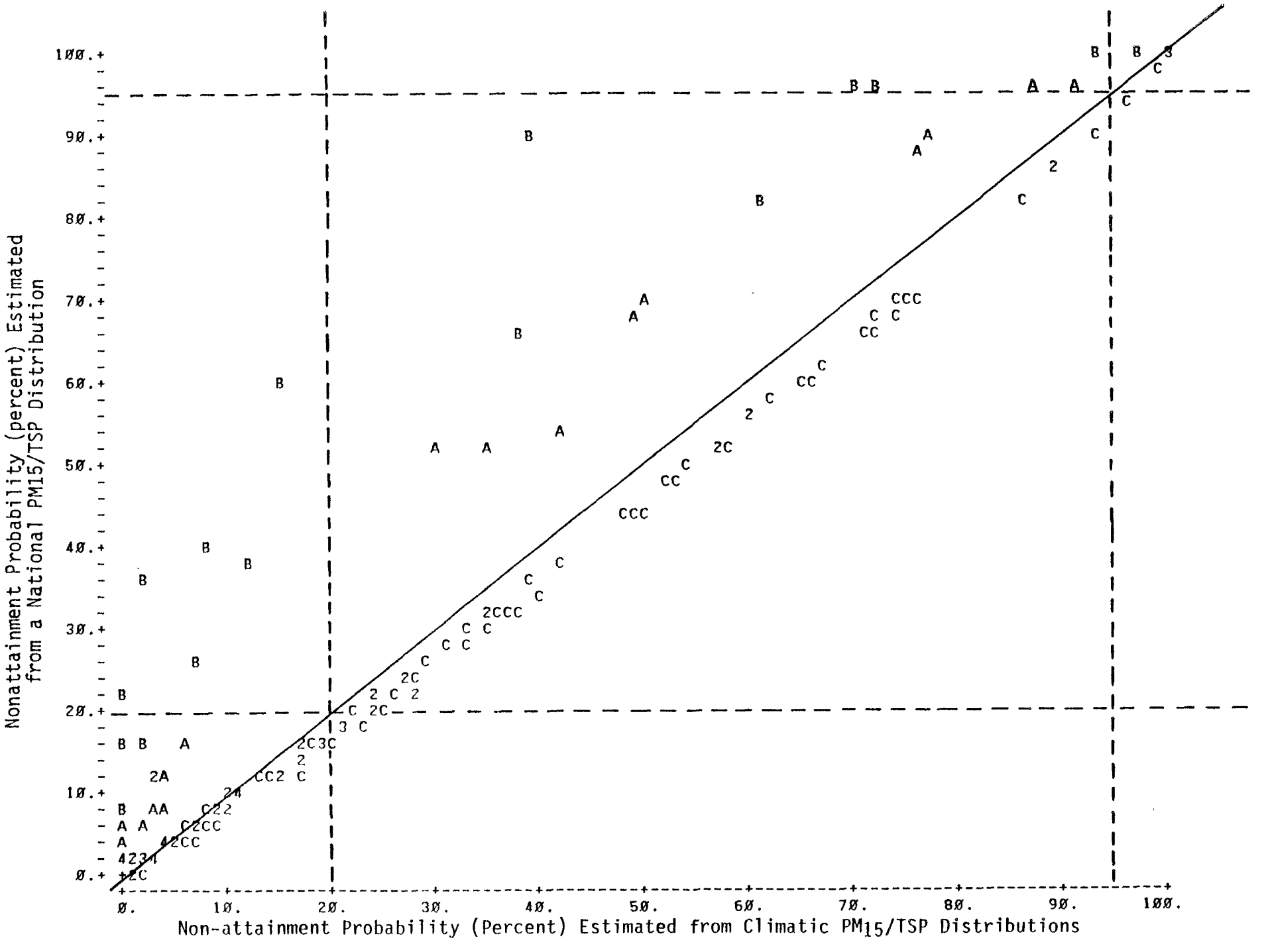

FIGURE 5-3. Probability of non-attainment of a 24-hour 100 $\mu g/m^3$ PM_{10} standard estimated from national and climatic PM_{15}/TSP distributions. A = West Coast, B = arid, C = other.

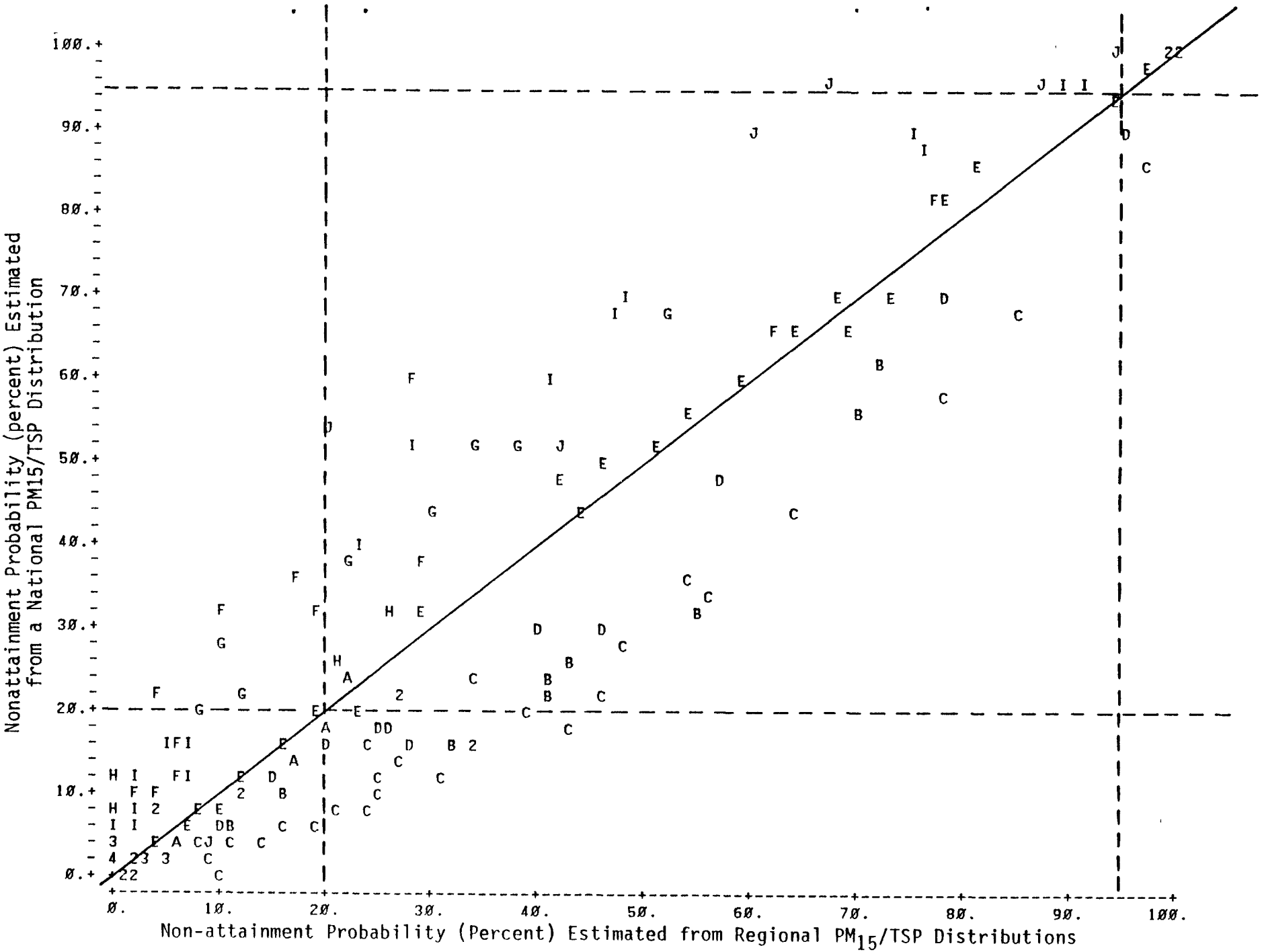

FIGURE 5-4. Probability of non-attainment of a 24-hour 100 $\mu g/m^3$ PM_{10} standard estimated from national and regional PM_{15}/TSP distributions. A = Region I, B = Region II, etc.

TABLE 5-2. Number of sites classified as low, medium, or high probability of nonattainment of three one-year 24-hour PM10 standards using a national cumulative distribution function of PM15/TSP ratios compared with probability classifications from climate and region distributions.

Standard		Climate cdfs				EPA Region cdf's			
		Low	Med	High		Low	Med	High	
75 µg/m^3	Low	46	1	0	47	41	6	0	47
	Med	9	74	6	89	8	77	4	89
	High	0	3	31	34	0	9	25	34
		55	78	37	170	49	92	29	170
100 µg/m^3	Low	86	6	0	94	79	15	0	94
	Med	6	59	1	66	7	58	1	66
	High	0	5	5	10	0	5	5	10
		94	70	6	170	86	78	6	170
150 µg/m^3	Low	150	1	0	151	145	6	0	151
	Med	6	13	0	19	3	16	0	19
	High	0	0	0	0	0	0	0	0
		156	14	0	170	86	78	6	170

ratio and a PM10 concentration were available.* At all three standards, there are no sites for which the use of the national distribution results in a high probability of nonattainment and use of one of the geographical distributions results in a low probability of nonattainment, or vice-vera. For the 75 and 100 $\mu g/m^3$ standards, use of the national and geographic distributions results in different probability classifications for about 11 to 16 percent of the site-years; in these cases more monitoring is dictated by the use of one distribution than the other. For the 150 $\mu g/m^3$ standard, which is in the range of the standards being considered, classification differences occur in only about 5 percent of the site-years. At even higher standards it seems likely that the percentage of sites with different classifications would be even smaller because the vast majority of the sites will be classified as having a low probability of nonattainment.

One notable difference between Tables 5-1 and 5-2 is that, for any standard, the percentage of sites classified as having a high probability of nonattainment is greater when using the PM10/TSP distribution (Table 5-1) than when using the PM15/TSP distribution (Table 5-2). In fact, in Table 5-2 no sites are classified as having a high probability of nonattainment of the 150 $\mu g/m^3$ 24-hour standard, while in Table 5-1 about 4 percent of the sites are so classified. This discrepancy is due to the difference in the sites used to construct the two ratio distributions. Those sites with both PM10 and TSP data have significantly higher TSP concentrations than those sites with both PM15 and TSP data. In addition, ratios corresponding to the lower percentage points (i.e., below the median) of the PM10/TSP distribution are lower than the corresponding ratios from the PM15/TSP distribution multiplied by 0.85, so that a given TSP concentration results in a higher probability of exceedance using the PM10/TSP

---

* We considered multiplying the PM15 concentrations by 0.85 to estimate PM10 concentrations and then determine attainment status. However, because of the variability in PM10/PM15 ratios, we considered this procedure invalid in this context.

distribution than with the PM15/TSP distribution. This topic will be discussed further in Section 7.

CROSS-VALIDATION APPLIED TO ANNUAL AVERAGE RATIO DISTRIBUTIONS

The range of concentrations proposed for the annual average primary PM10 standard is 50 to 65 $\mu g/m^3$. Because so few of the site-years in the IP network data base exceed even 50 $\mu g/m^3$, we have chosen to perform the cross-validation analysis assuming annual average PM10 standards of 30, 35, and 50 $\mu g/m^3$, a range that is wide enough to determine whether the results depend upon the level of the standard.

Figures 5-5 and 5-6 show P-P plots of nonattainment probabilities for an annual average 35 $\mu g/m^3$ standard derived from a national distribution and from climatic distributions of PM10/TSP and PM15/TSP, respectively. As in the 24-hour P-P plots, each plotted point represents one site-year. In both plots we see that sites in the West Coast and arid climates have a lower probability of nonattainment under the climate-specific distributions than under the national distribution, while the converse is true for sites in the third climatic region; this is the same pattern as in Figures 5-1 and 5-2 for the 24-hour standard of 100 $\mu g/m^3$. This is because the percentiles of the West Coast and arid distributions are always less than those of the national and the third climatic distribution; this can be seen in the cdf comparison plots of Figures 4-3 and 4-6 for the 24-hour and annual average ratios, respectively.

In Table 5-3 the probability of nonattainment (low, medium, and high, as described above) resulting from the use of a national PM10/TSP distribution and climatic distributions are compared to the actual attainment status of a site-year for each of the three standards considered. Actual attainment status was determined by using the sample annual average PM10 concentration for the site-year as an estimate for the true annual average; only site-years with at least 10 days of sampling both PM10 and TSP are considered. For the most severe misclassifications--nonattainment

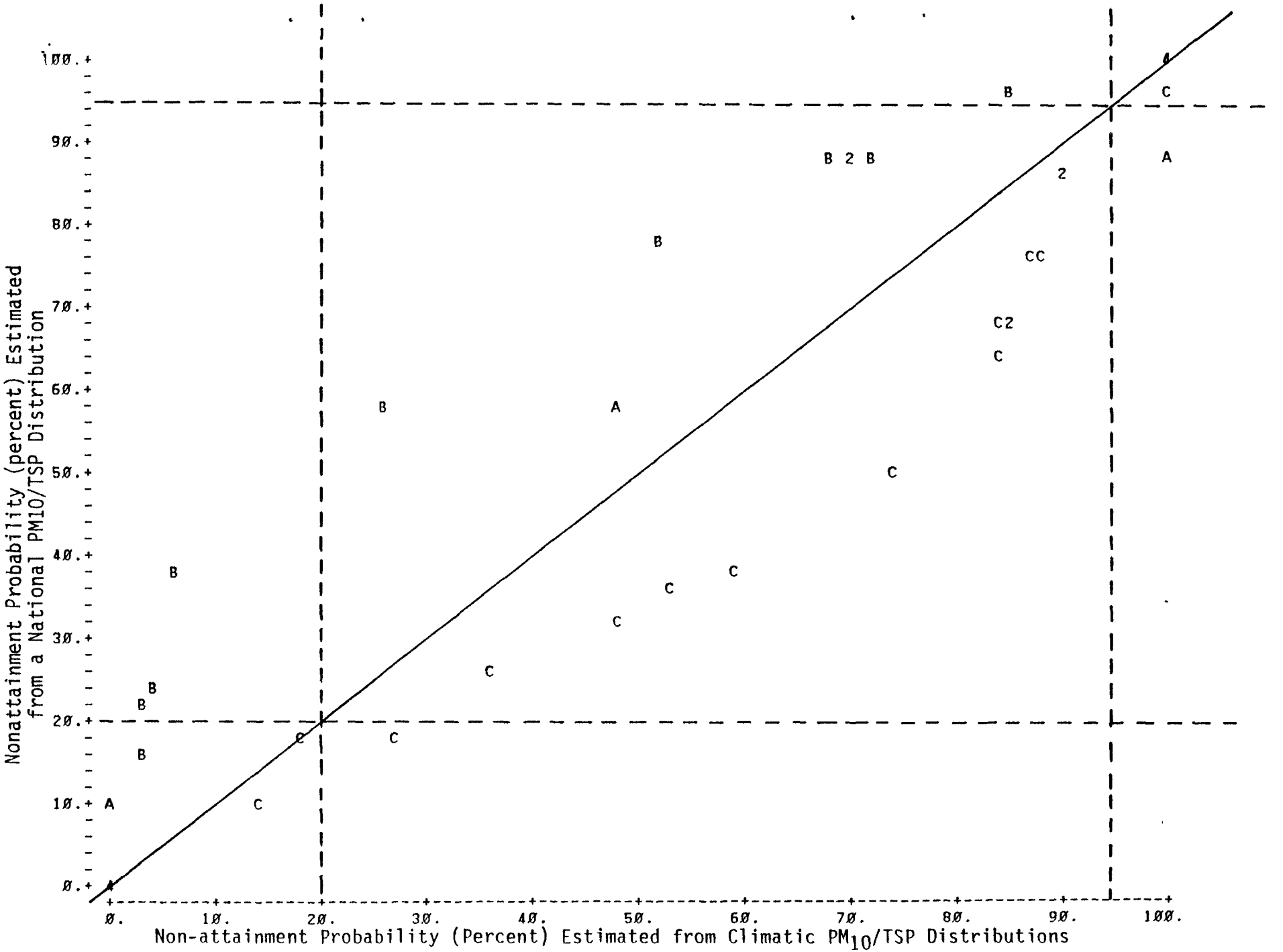

FIGURE 5-5. Probability of non-attainment of a 35 $\mu g/m^3$ annual average PM_{10} standard estimated from national and climatic PM10/TSP distributions. (A = West Coast, B = arid, C = others)

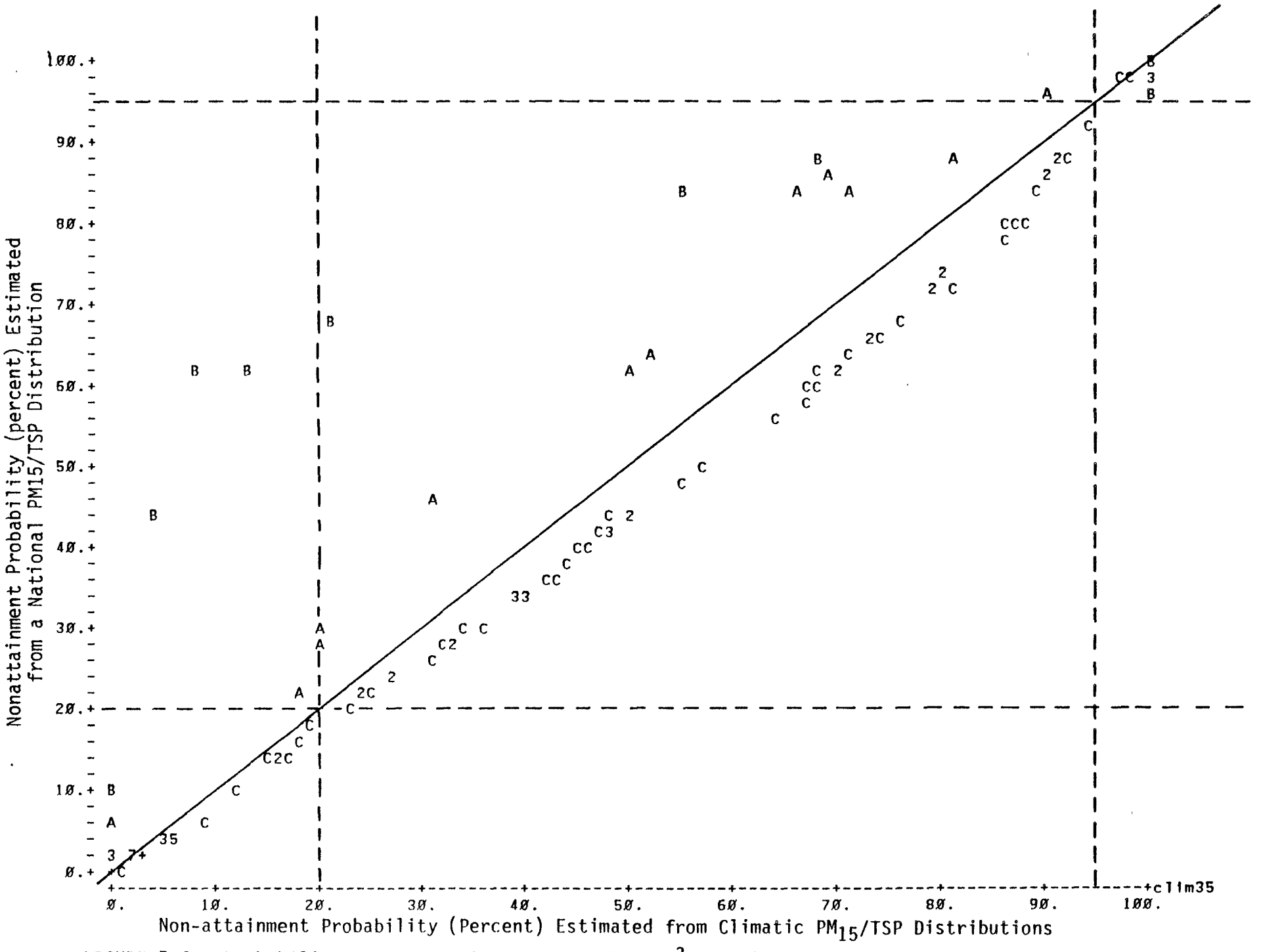

FIGURE 5-6. Probability of non-attainment of a 35 $\mu g/m^3$ annual average PM_{15} standard estimated from national and climatic PM_{15}/TSP distributions. (A = West Coast, B = arid, C = other)

TABLE 5-3. Number of sites classified as low, medium, or high probability of nonattainment of three annual PM10 standards compared with actual attainment status (pass = attainment, fail = nonattainment): comparison between national and climate PM10/TSP cumulative distribution functions.

Standard		National cdf				Climate cdf's			
		Low	Med	High		Low	Med	High	
30 µg/m^3	Pass	3	8	0	11	6	5	0	11
	Fail	1	18	9	28	1	19	8	28
		4	26	9	39	7	24	8	39
35 µg/m^3	Pass	8	8	1	17	10	6	1	17
	Fail	1	16	5	22	1	16	5	22
		9	24	6	39	11	22	6	39
50 µg/m^3	Pass	26	7	0	33	27	6	0	33
	Fail	0	5	1	6	1	2	3	6
		26	12	1	39	28	8	3	39

sites classified as having a low probability of nonattainment or attainment sites classified as having a high probability of nonattainment--there is almost no difference between using the national and the climatic PM10/TSP distributions. The only difference is at the 50 $\mu g/m^3$ standard, where the national distribution performs slightly better. In comparing site-years in which the standard was not attained but the probability of nonattainment was declared not to be high (so that not enough monitoring would be required), there is almost no difference between the misclassication rates based on the national and the climatic distributions. The climatic distributions perform slightly better, however, in avoiding the misclassification of attainment sites deemed medium or high probability of nonattainment.

The probabilities of nonattainment based on a national and separate climatic PM15/TSP distributions are compared in Table 5-4. There is no evidence that the use of climatic distributions improves the classification of sites. As with the 24-hour standard, there is no site-year for which the national distribution predicts a high nonattainment probability while the climatic distribution predicts a low probability of nonattainment, or vice versa. The percentage of site-years in which the probability classifications are different is only about 5 percent for all three standards. In most of these cases the nonattainment probability classification predicted by the national distribution is higher than that predicted by the climatic distributions, i.e., the national distribution dictates more monitoring for these sites than does the climatic distribution.

TABLE 5-4. Number of sites classified as low, medium, or high probability of nonattainment of three annual PM10 standards using a national cumulative distribution function of PM15/TSP ratios versus probability classifications from climate distributions.

Standard		Climate cdfs			
		Low	Med	High	
30 $\mu g/m^3$	Low	29	0	0	29
	Med	1	78	2	81
	High	0	4	19	23
		30	82	21	133
35 $\mu g/m^3$	Low	49	0	0	49
	Med	6	70	0	76
	High	0	1	7	8
		55	71	7	133
50 $\mu g/m^3$	Low	115	0	0	155
	Med	5	13	0	18
	High	0	0	0	0
		120	13	0	133

6 ESTIMATING DESIGN VALUES USING PM10/TSP RATIOS

Annual and 24-hour PM10 design values can be predicted by using the observed PM10/TSP ratio distributions. Annual average design values are simple to predict, but 24-hour design value estimates require repeated calculations. The methods for predicting design values will be summarized below, but for a more complete explanation and examples the reader is referred to Appendix B of the PM10 SIP Development Guideline (EPA, 1984).

DESIGN VALUES ESTIMATED FROM TSP DATA AND RATIO DISTRIBUTIONS

Annual average PM10 design values can be estimated by multiplying the observed annual average TSP concentrations by a factor which represents the distribution of the PM10/TSP ratios of annual averages. The factor can be either the mean of the distribution or the median; we prefer the latter because it is less influenced by outliers. As can be seen in Tables 3-2 and 3-4, the medians of the annual average PM10/TSP distributions are 0.49 and 0.47 for all annual ratios and for those based on high annual average TSP concentrations, respectively.

The 24-hour design value is that PM10 concentration which has an expected exceedance rate of once per year. This value can be estimated from the distribution of 24-hour PM10/TSP values by following these steps:

(1) Assume a design value, e.g., the level of the standard

(2) Using the 24-hour PM10/TSP ratio distribution, calculate the probability of exceedance for each day for which there is a valid TSP concentration.

(3) Calculate the expected number of exceedances by summing the exceedance probabilities calculated in Step 2. Adjust for missing sampling days by multiplying by the ratio of the number of days in the year divided by the number of sampling days with valid TSP concentrations (this assumes that the exceedance rate in the unsampled days is the same as the exceedance rate in the sampled days).

(4) If the expected number of exceedances calculated in Step 3 is equal to one, then the design value is found. If it is less than one, then go back to Step 2 with a lower design value. If it is greater than one, then go back to Step 2 with a higher design value.

A computer program which performs the required iterative calculations is available from EPA (Freas, 1984); this program was modified for the present study to consider a one-year attainment period for the 24-hour standard rather than a three-year period.

COMPARISON WITH DESIGN VALUES ESTIMATED FROM PM10 DATA

Annual and 24-hour PM10 design values were calculated for 24 site-years for which 35 or more PM10/TSP ratios were available. The PM10 design values estimated from the TSP data for these site-years are listed in Table 6-1 along with design values calculated from the PM10 data for these site years. Annual design values were calculated from the PM10 data simply as the observed annual average.*

---

* We also looked at linear and loglinear regressions of annual PM10 average on annual TSP average. The intercepts in the regression were insignificant, and the slopes were approximately the same as the 0.47 and 0.49 factors from the median of the distribution of annual average PM10/TSP.

TABLE 6-1. Estimation of PM10 design values from PM10 data and from PM10/TSP ratio distributions.

SITECODE	YEAR	NSAMP	DV24TAIL	DV24ALL	DV24HIGH	PM10AVG	TSPAVG	ANNDVALL
2	83	50	140	157	143	47.5	87.2	42.7
11	82	43	138	151	137	38.6	95.8	46.9
27	82	41	214	212	189	73.6	126.1	61.8
27	83	38	257	247	228	77.9	133.2	65.3
46	83	38	102	129	120	27.1	60.1	29.4
60	83	50	148	159	146	35.5	77.3	37.9
88	83	45	107	111	100	27.5	59.1	29.0
93	83	45	72	144	132	26.3	81.8	40.1
94	83	52	139	259	238	33.8	120.6	59.1
96	82	35	123	136	125	35.3	64.9	31.8
96	83	49	104	97	91	29.1	47.0	23.0
98	83	49	117	100	93	34.1	50.3	24.6
117	83	38	213	198	181	46.6	83.2	40.8
118	83	53	128	154	143	40.9	83.4	40.9
123	83	51	113	102	92	25.7	55.9	27.4
128	82	43	134	121	113	37.6	60.2	29.5
128	83	46	217	131	119	47.5	71.8	35.2
129	82	36	98	88	84	34.8	62.3	30.5
129	83	47	112	96	90	30.5	59.8	29.3
144	83	46	160	145	133	58.1	89.0	43.6
149	83	46	122	100	90	25.0	50.6	24.8
150	83	47	124	116	109	29.1	61.4	30.1
156	83	36	117	75	68	28.1	41.0	20.1
158	83	49	102	137	124	31.0	75.1	36.8

Twenty-four-hour PM10 design values were estimated by fitting a tail exponential model to the highest 20 percent of the PM10 concentrations. This model assumes that, no matter what the underlying distribution, the upper 20 percent of the values drop off exponentially; for a variety of underlying distributions such a model fits well. Complete details of the model, its derivation, and its application to estimating design values can be found in Breiman et al. (1978). After obtaining the parameters of a tail exponential distribution for a given site-year, the design value is found by calculating the 364/365 percentile of the distribution, i.e., that value which is exceeded only once per year.

Figure 6-1 compares 24-hour PM10 design values estimated from the national 24-hour PM10/TSP distribution and from the PM10 tail exponential model. For many site-years the differences between the TSP-based estimate and the PM10-based estimate is small. Nevertheless, there are some very large differences, as high as 119 $\mu g/m^3$; the average absolute difference is 25.5 $\mu g/m^3$. A similar plot for design values estimated from high TSP data (i.e., those ratios with TSP greater than or equal to 100 $\mu g/m^3$) is not provided because it appears almost the same as Figure 6-1, since these design values are all between 5 and 10 percent less than the design values estimated from the ratio distribution undifferentiated by TSP concentration.

A comparison of PM10 annual design values estimated from TSP data (by multiplying by 0.49, the median of the complete annual average PM10/TSP distribution) and from the annual PM10 average is shown in Figure 6-2. Here too there is much scatter in the plot, and some large discrepancies between the two estimates can be seen. The largest difference between the PM10-based and the TSP-based annual design value estimates is 25 $\mu g/m^3$; the mean absolute difference is approximately 7 $\mu g/m^3$. The observed differences are approximately the same for TSP-based design values estimated from the median of the annual average PM10/TSP distribution for high TSP concentration (i.e., ratios for which the annual average TSP is at least 55 $\mu g/m^3$); since the multiplier is 0.47 instead of 0.49, there is very little difference between the resulting annual design value estimates.

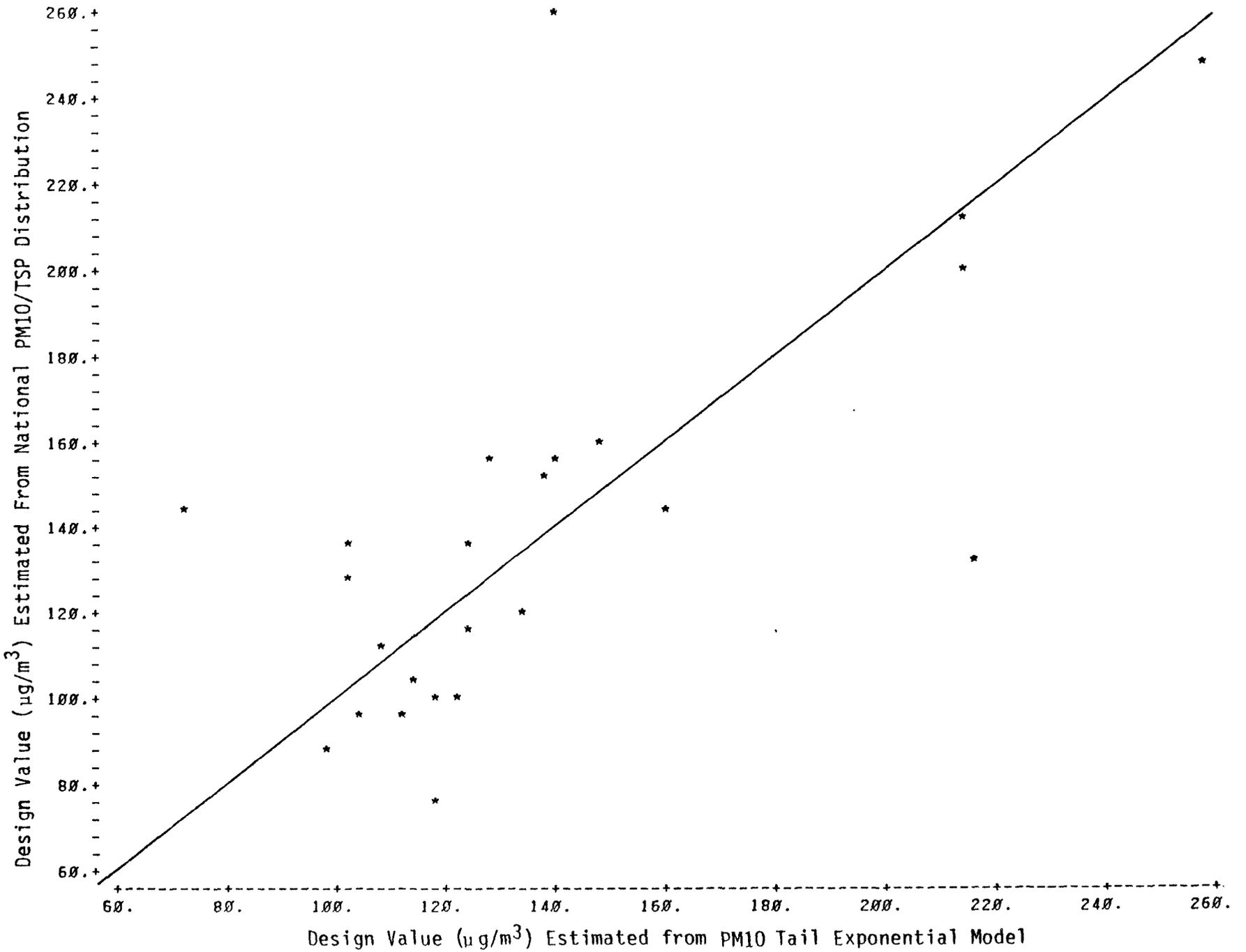

FIGURE 6-1. 24-hour PM10 design values estimated from PM10 data (using tail exponentail model) and from TSP data (using PM10/TSP distribution).

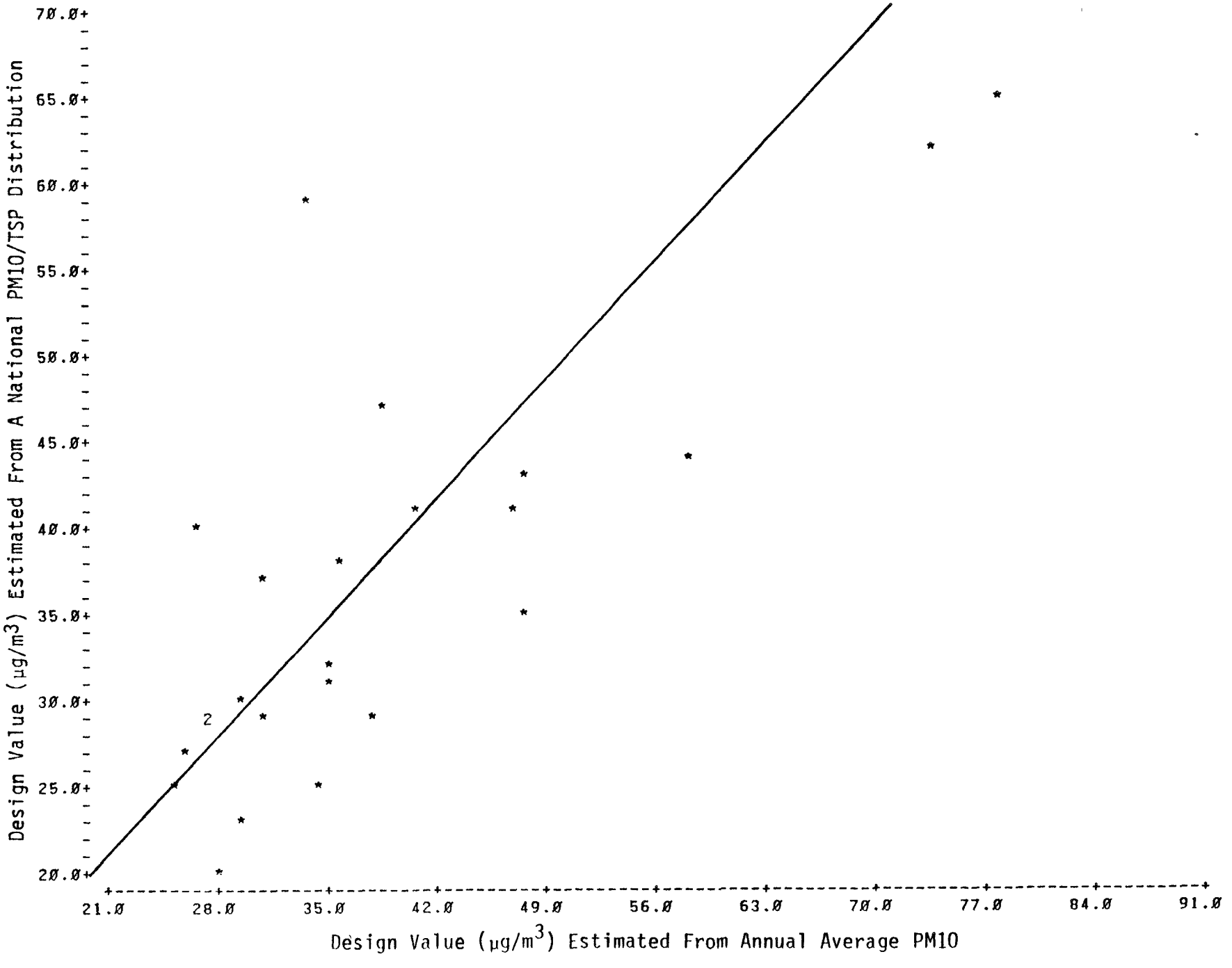

FIGURE 6-2. Anuual PM10 design values estimated from PM10 data (using annual average) and from TSP data (using annual average PM10/TSP distribution).

7 DISCUSSION AND RECOMMENDATIONS

In this section we summarize the results of our analyses and recommend which ratio distributions to use to determine the probability of attainment of 24-hour and annual PM10 standards. Our results are based on the PM10 and PM15 data from the IP netowrk for 1982-1983, and on TSP data from collocated monitors, as described in Section 1.

SITES SELECTED TO MONITOR PM10 AND PM15

Analyses of the IP network data are complicated by the way in which sites were selected to measure PM10 and PM15. The IP network sites were selected to represent larger urban areas. After the initial selection of sites, a few more sites were added in smaller rural locations in the West. State and regional EPA offices then requested additional sites. Since the sites were not selected randomly and are not geographically representative, inferences to the larger population of monitors across the country must be made carefully.

Initially, all IP network sites monitored PM15; however, they did not all use the same kind of sampler. Those sites that did not have dichotomous samplers had problems and did not record valid data. Those sites selected to monitor PM10 were sites where the highest TSP concentrations were observed, i.e., the sites where the highest PM10 concentrations were expected. Initially it was thought that the proposed particulate matter standard would be based on PM15 concentrations. However, when it was decided that the standard would be based on PM10 instead, the inlet on the

PM15 dichotomous sampler was changed to sample PM10 (Pace, personal communication). Thus there are very few monitors which sample both PM10 and PM15, and the monitors that sample PM10 are at sites where high concentrations occur.

NATIONAL VERSUS LOCAL DISTRIBUTIONS OF RATIOS

One of the main issues that this report addresses is whether the likelihood of attainment for a site should be based on a single national distribution of particulate matter ratios or on a more local distribution. Analyses of variance were performed to see if there are differences in the means of ratios grouped geographically (see Section 4). For both 24-hour and annual average ratios, statistically significant differences in ratio distributions were observed among climatic regions, EPA regions, states, cities, and sites, for all three kinds of ratios examined (PM10/TSP, PM15/TSP, and PM10/PM15). The results did not change when the ratios were limited to a subset with high TSP concentrations. Statistical issues in the interpretation of ANOVA results were provided in that section, and the cumulative distribution functions for geographical groupings permit one to see the practical differences.

The analyses show that ratio distributions by site are significantly different. It is our judgment that geographic differences, e.g., among EPA regions, may be statistically significant not because of actual regional differences but because each geographical region comprises only a few sites--in many cases, only one or two sites, especially in smaller geographical groupings, such as state and city--and there are major differences between sites. A larger data base is required to be able to detect actual regional differences, if they exist.

Observed differences by year (and quarter) may also be attributed to site-to-site differences: Not all of the monitors were placed in service at

the same time and many monitors did not operate for long periods. Thus the ratios are rarely derived from the same set of sites over a given period.

The practical implications of using a single national ratio distribution instead of climate or EPA regional distributions were also examined (see Section 5). The results must be viewed in the light of the large differences in ratio distributions across sites. Using the "leave-out-one" cross-validation approach, nonattainment probabilities were calculated for each site using a ratio distribution that did not include ratios from that site. The observed differences in nonattainment probability for a site between one calculated from a national distribution and from a climatic or regional distribution may only be an artifact of the large site-to-site differences.

We recommend the use of a single national distribution except when sufficient data at a site permit the site-specific distribution to be reliably estimated. The two major reasons for this recommendation are (1) site-to-site differences are so much larger than geographic differences, and (2) attainment classifications based on a one-year test do not appear to be improved by the adoption of climatic or regional distributions.

There are other statistical considerations favoring this recommendation. First, because sites have not been chosen randomly within geographical regions, the observed regional distributions cannot necessarily be considered representative of regions in which they are located. Second, geographical distributions are based on much smaller sample sizes than the national distribution and are thus less reliable.

At the same time the nonrandom location of sites could also be used as an argument against the use of a single national distribution. However, since the national distribution is based primarily on ratios observed at

urban sites with high concentrations, and since the probability of nonattainment is zero unless the TSP concentration is above the PM10 standard, the Pace and Frank methodology will only be used by similar, high-concentration sites. In addition, given the large site-to-site differences, the national distribution represents a greater range of sites than any one geographic distribution.

RATIOS DIFFERENTIATED BY TSP CONCENTRATION

A second major issue is whether the probability of nonattainment should be estimated from distributions differentiated by TSP concentration. We saw significant differences in particulate matter ratios corresponding to high or low TSP concentration (see Sections 2 and 3). Ratios at lower TSP concentrations are more variable than ratios at higher TSP concentrations. We recommend a national ratio distribution based on high TSP concentrations (above 100 $\mu g/m^3$ for 24-hour concentrations and 55 $\mu g/m^3$ for annual average concentrations) for two reasons. First, the distribution based on high TSP concentrations is less prone to measurement error than the undifferentiated distribution. Second, the ratio distributions will only be applied to site-years and site-days with high TSP concentrations, i.e., higher than the level of the standard, because otherwise the exceedance probability is nil (since PM10 is assumed to be less than TSP). It is logical, then, to have a ratio distribution based on high TSP concentrations only.

It should be noted here that in cross-validation analyses (see Section 6), complete national distributions were used rather than the high-TSP distributions. This was done because the sample sizes for the high-TSP distributions for specific geographical units were too small to perform the cross-validation analyses. We assume that the results using the high-TSP distribution, if enough data were available, would be similar. In fact, it is likely that differences in nonattainment probabilities observed

between national and geographic distributions (see Section 5) would be even smaller if high-TSP distributions were used, because of the reduced variability in the ratio distributions.

Estimation of design values using the Pace and Frank methodology was performed for a small subset of sites (see Section 6). In the analysis both the complete and the high-TSP national distributions were used, and there was little difference in the resulting design value estimates. These estimates did not agree well in some cases with design values estimated from actual PM10 data, but this may be because very little PM10 data was available for each site-year (at most 53 samples out of a possible 365). Further analyses are recommended when more PM10 data are available.

PM10/TSP VERSUS PM15/TSP DISTRIBUTIONS

The third and final issue is which ratio distribution to use for calculating probabilities of attainment. With only TSP monitoring data available, there are two choices. One may use either the PM10/TSP ratio distribution or the PM15/TSP ratio distribution multiplied by 0.85. In previous studies the PM10/TSP distribution was recommended for the 24-hour standard and the PM15/TSP distribution (multiplied by an appropriate factor) was recommended for the annual standard (Pace and Frank, 1982; Thrall and Hudischewskyj, 1984). In these studies only 1982 Network PM10 data were available, and annual average PM10/TSP distributions were based on too few samples to be considered reliable.

With the 1983 IP Network PM10 data, however, there is a much larger sample size for the annual average PM10/TSP distribution, and we recommend its use as well as the 24-hour distribution. Although the annual average PM15/TSP distribution is based on a substantially larger number of site years, multiplying by 0.85 or any other factor assumes that PM10/PM15 is relatively constant, which has been shown not to be the case.

ATTAINMENT PROBABILITIES WHEN ONLY PM15 DATA ARE AVAILABLE

There may be some sites for which the only particulate matter monitored is PM15. Until actual PM10 data become available for these sites, attainment/nonattainment classification may be made either by using a PM15 to PM10 conversion factor or by using a PM10/PM15 distribution. Because of the variation in the distribution of PM10/PM15 ratios, we recommend against the use of a single factor. We recommend that a single national distribution based on days with high TSP be used for the reasons listed above. The percentiles of these 24-hour and annual distributions may be found in the first column of Tables 3-3 and 3-4, respectively.

CONCLUSION

In summary, we recommend the use of PM10/TSP distributions based on high TSP concentrations. Unless enough data are available to develop a site-specific PM10/TSP distribution, we recommend that a national distribution of annual average and 24-hour ratios be used to estimate nonattainment probabilities for the respective standards. The percentage points and summary statistics of the recommended distributions may be found in the middle column of Tables 3-3 and 3-4.

REFERENCES

Breiman, L., J. Gins, and C. Stone. 1978. "Statistical Analysis and Interpretation of Peak Air Pollution Measurements." Technology Service Corporation, Santa Monica, CA (TSC-PD-A190-10).

Dixon, W. J., ed. 1983. BMDP Statistical Software. University of California Press.

EPA. 1984. PM10 SIP Development Guideline. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Federal Register. 1984. Proposed Revisions to the National Ambient Air Quality Standards for Particulate Matter, 49(55):10408.

Frank, N. H. 1984. "Nationwide Trends in Total Suspended Particulate Matter and Associated Changes in the Measurement Process." Air Pollution Control Association/American Society for Quality Control Specialty Conference on Quality Assurance in Air Pollution Measurements. Boulder, CO (14-18 October).

Freas, W. P. 1984. User's Guide for PM10 Probability Guideline Software. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Mosteller, F., and J. W. Tukey. 1968. "Data Analysis Including Statistics." In Handbook of Social Psychology, Vol. 1, G. Lindsey and E. Aronson, eds., Addeson-Wesley, Reading, MA.

Pace, T. G., and N. H. Frank. 1984. Procedures for Estimating Probability of Nonattainment of a PM10 NAAQS Using Total Suspended Particulate or Inhalable Particulate Data. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Thrall, A. D., and C. S. Burton. 1983. "Characterizing Ratios of Particulate Concentrations: A Preliminary Step in Assessing Likely Attainment Status Under a PM10 National Ambient Air Quality Standard." Systems Applications, Inc., San Rafael, CA (SYSAPP-83/078).

Thrall, A. D., and A. B. Hudischewskyj. 1984. "An Update on the Use of Particulate Ratios to Assess Likely PM10 Attainment Status." Systems Applications, Inc., San Rafael, CA (SYSAPP-84/100).

Weiner, B. J. 1971. Statistical Principles in Experimental Design, 2nd edition. McGraw-Hill, New York.

Appendix A

MONITORING SITES USED IN THE ANALYSIS

Appendix A

MONITORING SITES USED IN THE ANALYSIS

On the following pages is a complete list of the monitoring stations used in our analyses. PM10 and PM15 data were taken from IP network monitoring records; TSP data were taken from monitoring records of collocated SAROAD TSP monitors. For this project we assigned to each collocated pair of IP network and SAROAD TSP monitors a unique three-digit site number. Additionally, the EPA assigns to each monitoring site a unique 12-digit ID for entry in the SAROAD system. The first nine characters uniquely identify the physical location of a monitor; these appear in the column labeled SAROAD. The remaining three characters indicate controlling agency (one letter) and project classification (two digits). The column labeled IP contains the agency and project codes for the IP Network monitor; the corresponding codes for the colocated TSP monitor are in the column labeled TSP. State and city names were extracted from the AEROS Manual of codes (EPA, 1983) for the SAROAD system, which also contains a map of EPA regions. Climate classifications were provided by the project officer.

Reference

EPA. 1983. AEROS Manual Series Volume V: AEROS Manual of Codes. U.S. Environmental Protection Agency, Research Triangle Park, NC (EPA-450/2-76-005a).

SITECODE	SAROAD	IP	TSP	STATE	CITY	EPA REG	CLIMATE
2	010380023	A07	G02	Alabama	Birmingham	IV	Other
6	012540001	A07	G03	Alabama	Mountain Brook	IV	Other
8	020040003	A07	I01	Alaska	Anchorage	X	West Coast
11	030600002	A07	G01	Arizona	Phoenix	IX	Arid
12	030600004	A07	G01	Arizona	Phoenix	IX	Arid
13	041440001	A07	F01	Arkansas	Little Rock	VI	Other
14	050500002	A07	I01	California	Azusa	IX	West Coast
16	050520004	A07	F01	California	Bakersfield	IX	Arid
17	051260002	A07	F01	California	Chico	IX	West Coast
18	052220003	A07	I01	California	El Cajon	IX	West Coast
19	052800005	A07	F01	California	Fresno	IX	Arid
20	052820002	A07	I01	California	Fresno County	IX	Arid
22	054020003	A07	I01	California	Livermore	IX	West Coast
23	054080002	A07	I01	California	Lompoc	IX	West Coast
24	054180103	A07	I01	California	Los Angeles	IX	West Coast
26	056300003	A07	I01	California	Richmond	IX	West Coast
27	056535001	A07	I01	California	Rubidoux	IX	West Coast
29	056980004	A07	I01	California	San Jose	IX	West Coast
31	060580001	A07	F01	Colorado	Denver	VIII	Arid
32	061260001	A07	F01	Colorado	Lakewood	VIII	Arid
33	061820001	A07	F01	Colorado	Pueblo	VIII	Arid
35	070420003	A07	F01	Connecticut	Hartford	I	Other
36	070478001	A07	F03	Connecticut	Litchfield County	I	Other
38	080180001	A07	F01	Delaware	New Castle County	III	Other
39	090020017	A07	I01	District of Columbia	Washington	III	Other
41	104360035	A07	G02	Florida	Tampa	IV	Other
42	110200001	A07	F01	Georgia	Atlanta	IV	Other
43	110200039	A07	F01	Georgia	Atlanta	IV	Other
44	114500017	A07	F02	Georgia	Savannah	IV	Other
45	120370004	A07	F01	Hawaii	Pearl City	IX	West Coast
46	130220003	A07	F01	Idaho	Boise City	X	Arid
47	141220014	A07	H01	Illinois	Chicago	V	Other
48	141220022	A07	H01	Illinois	Chicago	V	Other
50	148320007	A07	F03	Illinois	Will County	V	Other
52	152040021	A07	F01	Indiana	Indianapolis	V	Other
54	162500003	A07	F01	Iowa	Marshalltown	VII	Other
56	171800011	A07	F02	Kansas	Kansas City	VII	Other
57	173560007	A07	F02	Kansas	Topeka	VII	Other
58	173740012	A07	F02	Kansas	Wichita	VII	Other
60	183090001	A07	G01	Kentucky	Okolona	IV	Other
63	210120001	A07	F01	Maryland	Baltimore	III	Other
65	210120009	A07	F01	Maryland	Baltimore	III	Other
67	211380007	A07	G01	Maryland	Rockville	III	Other
68	220240012	A07	F01	Massachusetts	Boston	I	Other
70	222160011	A07	F01	Massachusetts	Springfield	I	Other

SITECODE	SAROAD	IP	TSP	STATE	CITY	EPA REG	CLIMATE
71	222640016	A07	F01	Massachusetts	Worcester	I	Other
72	231180015	A07	G01	Michigan	Detroit	V	Other
76	242260049	A07	F01	Minnesota	Minneapolis	V	Other
77	242260051	A07	F01	Minnesota	Minneapolis	V	Other
78	243300003	A07	H01	Minnesota	St. Paul	V	Other
79	251260003	A07	F01	Mississippi	Jackson	IV	Other
80	260030001	A07	G01	Missouri	Affton	VII	Other
81	262380002	A07	H01	Missouri	Kansas City	VII	Other
83	264280007	A07	H01	Missouri	St. Louis	VII	Other
84	270160005	A07	F01	Montana	Butte - Silver Bow	VIII	Other
86	271100020	A07	G01	Montana	Missoula County	VIII	Other
87	281880028	A07	G01	Nebraska	Omaha	VII	Other
88	290480001	A07	I01	Nevada	Reno	IX	Arid
89	290580001	A07	F01	Nevada	Winnemucca	IX	Arid
90	310720005	A07	F01	New Jersey	Camden	II	Other
91	311380001	A07	F01	New Jersey	Essex County	II	Other
92	312320005	A07	F01	New Jersey	Jersey City	II	Other
93	320040001	A07	H01	New Mexico	Albuquerque	VI	Arid
94	320090001	A07	F01	New Mexico	Bayard	VI	Arid
96	330660010	A07	F01	New York	Buffalo	II	Other
97	332000003	A07	F01	New York	Erie County	II	Other
98	333520001	A07	F01	New York	Lackawanna	II	Other
99	334680005	A07	F01	New York	New York City	II	Other
101	334680079	A07	F01	New York	New York City	II	Other
102	340700010	A07	G01	North Carolina	Charlotte	IV	Other
104	341160006	A07		North Carolina	Durham	IV	Other
105	341160101	A07		North Carolina	Durham	IV	Other
106	341160102	A07		North Carolina	Durham	IV	Other
107	360060014	A07	H01	Ohio	Akron	V	Other
108	361220020	A07	G01	Ohio	Cincinnati	V	Other
109	361300013	A07	H01	Ohio	Cleveland	V	Other
110	361300021	A07	H01	Ohio	Cleveland	V	Other
111	361300041	A07	H01	Ohio	Cleveland	V	Other
113	361660014	A07	G01	Ohio	Dayton	V	Other
114	363080010	A07	H01	Ohio	Ironton	V	Other
116	364340005	A07	G01	Ohio	Middletown	V	Other
117	366420012	A07	I01	Ohio	Steubenville	V	Other
118	367760002	A07	I02	Ohio	Youngstown	V	Other
119	372200035	A07	F01	Oklahoma	Oklahoma City	VI	Arid
122	380560013	A07	F01	Oregon	Eugene	X	West Coast
123	381460015	A07	F01	Oregon	Portland	X	West Coast
124	390100064	A07	G01	Pennsylvania	Allegheny County	III	Other
125	390100068	A07	G01	Pennsylvania	Allegheny County	III	Other
126	390400002	A07	G01	Pennsylvania	Avalon	III	Other
127	390780725	A07	F02	Pennsylvania	Bethlehem	III	Other

SITECODE	SAROAD	IP	TSP	STATE	CITY	EPA REG	CLIMATE
128	396620001	A07	G01	Pennsylvania	North Braddock	III	Other
129	397140003	A07	H01	Pennsylvania	Philadelphia	III	Other
133	397140024	A07	H01	Pennsylvania	Philadelphia	III	Other
135	397140036	A07	H01	Pennsylvania	Philadelphia	III	Other
137	397140038	A07	H01	Pennsylvania	Philadelphia	III	Other
139	410300012	A07	F01	Rhode Island	Providence	I	Other
140	420560003	A07	F02	North Carolina	Charleston County	IV	Other
141	440380006	A07	G02	Tennessee	Chattanooga	IV	Other
142	442540006	A07	G01	Tennessee	Murfreesboro	IV	Other
143	451310050	A07	H01	Texas	Dallas	VI	Arid
144	451700002	A07	F01	Texas	El Paso	VI	Arid
146	452330024	A07	F01	Texas	Harris County	VI	Other
147	452560034	A07	F01	Texas	Houston	VI	Other
148	454715001	A07	F01	Texas	Seabrook	VI	Other
149	460520001	A07	F02	Utah	Magna	VIII	Arid
150	460920001	A07	F01	Utah	Salt Lake City	VIII	Arid
154	481560002	A07	F02	Virginia	Hopewell	III	Other
155	482140007	A07	F01	Virginia	Norfolk	III	Other
156	482630001	A07	G01	Virginia	Reston	III	Other
158	491840057	A07	I02	Washington	Seattle	X	West Coast
159	491840073	A07	I01	Washington	Seattle	X	West Coast
160	492040013	A07	F01	Washington	Spokane	X	West Coast
161	500280004	A07	F01	West Virginia	Charleston	III	Other
162	502000002	A07	F02	West Virginia	Weirton	III	Other
163	502120002	A07	F01	West Virginia	Wheeling	III	Other
164	510240002	A07	F01	Wiconsin	Beloit	V	Other
166	511180009	A07	F01	Wisconsin	Green Bay	V	Other

Appendix B

LISTING OF SITE-DAYS ON WHICH AT LEAST
ONE 24-HOUR RATIO EXCEEDS 1.0

SITECODE	SAROAD: ID	YEAR	MONTH	DAY	PM10	PM15	TSP	PM10/PM15	PM10/TSP	PM15/TSP
2	010380023	82	1	17	*	63.620	60.0	*	*	1.06033
2	010380023	82	5	29	110.300	91.290	*	1.2082	*	*
2	010380023	82	6	4	66.230	64.820	*	1.0218	*	*
2	010380023	82	8	3	120.500	94.180	*	1.2795	*	*
2	010380023	82	11	19	25.990	44.490	44.0	0.5842	0.59068	1.01114
2	010380023	83	5	30	27.100	26.040	45.0	1.0407	0.60222	0.57867
2	010380023	83	7	5	86.460	*	48.0	*	1.80125	*
6	012540001	82	2	4	*	51.710	34.0	*	*	1.52088
6	012540001	82	2	22	*	17.950	17.0	*	*	1.05588
8	020040003	82	12	31	*	7.156	7.0	*	*	1.02229
8	020040003	83	3	13	*	16.900	13.0	*	*	1.30000
11	030600002	82	3	12	13.410	12.390	26.3	1.0823	0.50989	0.47110
13	041440001	82	6	4	*	37.940	36.0	*	*	1.05389
13	041440001	83	4	6	*	27.060	25.0	*	*	1.08240
13	041440001	83	6	29	*	39.090	34.0	*	*	1.14971
13	041440001	83	9	21	*	49.530	30.0	*	*	1.65100
14	050500002	82	10	8	*	303.600	86.0	*	*	3.53023
14	050500002	82	11	7	*	129.300	100.0	*	*	1.29300
16	050520004	83	2	23	71.670	*	45.0	*	1.59267	*
17	051260002	83	2	17	*	45.330	44.0	*	*	1.03023
17	051260002	83	12	26	*	22.900	18.0	*	*	1.27222
22	054020003	82	1	11	*	82.230	70.0	*	*	1.17471
22	054020003	82	4	11	*	20.750	16.0	*	*	1.29687
22	054020003	82	8	9	*	61.080	49.0	*	*	1.24653
26	056300003	82	2	22	*	67.070	43.0	*	*	1.55977
27	056535001	82	2	22	199.100	217.000	212.0	0.9175	0.93915	1.02358
27	056535001	82	4	11	139.300	96.980	118.0	1.4364	1.18051	0.82186
29	056980004	82	1	29	*	122.600	66.0	*	*	1.85758
31	060580001	83	4	18	152.300	*	142.0	*	1.07254	*
36	070478001	82	5	29	*	13.000	10.0	*	*	1.30000
36	070478001	82	6	22	*	29.860	27.0	*	*	1.10593
36	070478001	82	6	28	*	67.520	51.0	*	*	1.32392
36	070478001	82	7	22	*	18.560	18.0	*	*	1.03111
36	070478001	82	8	9	*	21.620	17.0	*	*	1.27176
36	070478001	82	8	15	*	28.160	23.0	*	*	1.22435
36	070478001	82	11	19	*	74.990	47.0	*	*	1.59553
38	080180001	82	8	21	*	35.270	29.0	*	*	1.21621
38	080180001	82	9	26	*	36.170	35.0	*	*	1.03343
38	080180001	82	11	19	*	64.950	64.0	*	*	1.01484
38	080180001	82	12	1	*	47.040	42.0	*	*	1.12000
38	080180001	83	10	3	*	67.520	64.0	*	*	1.05500
38	080180001	83	10	9	*	63.190	50.0	*	*	1.26380
38	080180001	83	11	8	*	93.700	93.0	*	*	1.00753
38	080180001	83	11	14	*	65.200	57.0	*	*	1.14386
39	090020017	82	1	5	*	48.130	47.0	*	*	1.02404

SITECODE	SAROAD ID	YEAR	MONTH	DAY	PM10	PM15	TSP	PM10/PM15	PM10/TSP	PM15/TSP
39	090020017	82	1	11	*	46.250	36.0	*	*	1.28472
39	090020017	82	1	23	*	62.290	48.0	*	*	1.29771
39	090020017	82	2	4	*	51.790	50.0	*	*	1.03580
39	090020017	82	2	10	*	51.980	48.0	*	*	1.08292
39	090020017	82	2	16	*	60.340	54.0	*	*	1.11741
39	090020017	82	9	26	*	40.880	36.0	*	*	1.13556
41	104360035	82	5	23	*	59.380	56.0	*	*	1.06036
44	114500017	82	4	5	*	85.720	43.0	*	*	1.99349
44	114500017	82	4	11	*	40.250	40.0	*	*	1.00625
44	114500017	82	4	17	*	40.850	29.0	*	*	1.40862
44	114500017	82	4	29	*	47.600	47.0	*	*	1.01277
44	114500017	82	5	17	*	70.670	60.0	*	*	1.17783
44	114500017	82	5	23	*	49.870	29.0	*	*	1.71966
44	114500017	82	5	29	*	67.400	37.0	*	*	1.82162
44	114500017	82	7	4	*	75.010	60.0	*	*	1.25017
44	114500017	82	8	9	*	71.280	56.0	*	*	1.27286
44	114500017	82	9	14	*	73.800	69.0	*	*	1.06957
44	114500017	82	9	20	*	82.580	49.0	*	*	1.68531
44	114500017	82	11	19	*	61.490	57.0	*	*	1.07877
48	141220022	82	6	16	*	55.240	35.0	*	*	1.57829
50	148320007	82	5	5	*	74.250	26.0	*	*	2.85577
52	152040021	82	9	8	*	53.970	48.0	*	*	1.12437
52	152040021	82	9	26	*	60.270	35.0	*	*	1.72200
52	152040021	82	11	19	*	55.120	34.0	*	*	1.62118
56	171800011	83	12	14	47.000	*	46.0	*	1.02174	*
58	173740012	82	9	14	*	84.190	83.0	*	*	1.01434
58	173740012	83	1	30	*	16.380	12.0	*	*	1.36500
60	183090001	82	6	22	*	53.020	50.0	*	*	1.06040
63	210120001	82	5	11	*	57.470	56.0	*	*	1.02625
63	210120001	82	6	4	*	21.420	21.0	*	*	1.02000
63	210120001	82	8	3	*	64.270	57.0	*	*	1.12754
63	210120001	82	8	9	*	83.700	72.0	*	*	1.16250
63	210120001	82	9	20	*	70.440	53.0	*	*	1.32906
63	210120001	82	9	26	*	63.610	39.0	*	*	1.63103
65	210120009	83	3	1	*	74.430	46.0	*	*	1.61804
67	211380007	82	9	20	*	41.050	32.0	*	*	1.28281
67	211380007	82	9	26	*	31.290	23.0	*	*	1.36043
67	211380007	82	10	2	*	40.670	39.0	*	*	1.04282
67	211380007	82	11	19	*	55.810	55.0	*	*	1.01473
67	211380007	82	12	25	*	47.340	21.0	*	*	2.25429
67	211380007	83	4	24	*	10.240	9.0	*	*	1.13778
67	211380007	83	5	30	*	57.590	40.0	*	*	1.43975
67	211380007	83	6	5	*	46.120	42.0	*	*	1.09810
71	222640016	82	4	5	*	47.450	28.0	*	*	1.69464
71	222640016	82	8	3	*	44.510	44.0	*	*	1.01159

SITECODE	SAROAD ID	YEAR	MONTH	DAY	PM10	PM15	TSP	PM10/PM15	PM10/TSP	PM15/TSP
77	242260051	82	3	24	*	143.900	69.0	*	*	2.08551
79	251260003	82	6	16	*	74.330	31.0	*	*	2.39774
79	251260003	82	7	10	*	69.360	48.0	*	*	1.44500
79	251260003	82	7	28	*	78.450	70.0	*	*	1.12071
79	251260003	82	11	1	*	48.280	35.0	*	*	1.37943
81	262380002	83	7	29	*	107.100	106.0	*	*	1.01038
81	262380002	83	8	4	*	93.070	87.0	*	*	1.06977
81	262380002	83	8	10	*	113.300	93.0	*	*	1.21828
81	262380002	83	8	22	*	107.800	78.0	*	*	1.38205
81	262380002	83	8	28	*	127.000	97.0	*	*	1.30928
83	264280007	82	8	9	*	45.740	41.3	*	*	1.10751
84	270160005	83	11	14	*	43.700	42.0	*	*	1.04048
87	281880028	82	6	4	*	91.730	73.0	*	*	1.25658
87	281880028	82	8	15	*	40.390	40.0	*	*	1.00975
87	281880028	82	10	20	*	44.790	34.0	*	*	1.31735
88	290480001	83	11	8	74.910	*	68.0	*	1.10162	*
90	310720005	82	8	3	*	87.370	56.0	*	*	1.56018
90	310720005	83	2	11	*	54.160	37.0	*	*	1.46378
91	311380001	82	2	28	*	64.670	46.0	*	*	1.40587
91	311380001	82	4	5	*	31.260	16.0	*	*	1.95375
91	311380001	82	4	11	*	38.260	35.0	*	*	1.09314
92	312320005	83	2	11	*	56.720	46.0	*	*	1.23304
93	320040001	83	3	19	28.200	*	24.0	*	1.17500	*
96	330660010	82	1	11	*	45.630	32.0	*	*	1.42594
96	330660010	82	4	11	33.070	31.650	48.0	1.0449	0.68896	0.65937
96	330660010	82	4	17	48.370	82.440	77.0	0.5867	0.62818	1.07065
96	330660010	82	11	13	*	9.520	2.0	*	*	4.76000
96	330660010	82	12	31	21.450	19.000	38.0	1.1289	0.56447	0.50000
96	330660010	83	1	6	47.060	34.770	61.0	1.3535	0.77148	0.57000
96	330660010	83	1	24	17.050	*	17.0	*	1.00294	*
96	330660010	83	1	30	41.290	32.010	44.0	1.2899	0.93841	0.72750
96	330660010	83	2	17	31.170	25.950	33.0	1.2012	0.94455	0.78636
96	330660010	83	2	23	24.700	16.940	33.0	1.4581	0.74848	0.51333
96	330660010	83	3	13	11.740	11.410	22.0	1.0289	0.53364	0.51864
96	330660010	83	3	19	8.074	7.180	9.0	1.1245	0.89711	0.79778
96	330660010	83	3	25	28.060	22.450	61.0	1.2499	0.46000	0.36803
96	330660010	83	3	31	18.990	16.700	45.0	1.1371	0.42200	0.37111
96	330660010	83	4	12	19.890	16.240	38.0	1.2248	0.52342	0.42737
96	330660010	83	4	18	18.570	15.530	26.0	1.1958	0.71423	0.59731
96	330660010	83	4	24	14.460	11.670	25.0	1.2391	0.57840	0.46680
96	330660010	83	4	30	24.400	16.190	28.0	1.5071	0.87143	0.57821
96	330660010	83	5	6	26.650	22.220	56.0	1.1994	0.47589	0.39679
96	330660010	83	5	12	23.640	17.960	55.0	1.3163	0.42982	0.32655
96	330660010	83	6	5	35.210	27.350	54.0	1.2874	0.65204	0.50648
96	330660010	83	6	11	48.370	37.500	83.0	1.2899	0.58277	0.45181
96	330660010	83	10	3	66.560	54.970	106.0	1.2108	0.62792	0.51858

SITECODE	SAROAD ID	YEAR	MONTH	DAY	PM10	PM15	TSP	PM10/PM15	PM10/TSP	PM15/TSP
96	330660010	83	10	9	8.052	5.794	49.0	1.3897	0.16433	0.11824
96	330660010	83	11	2	37.960	30.840	46.0	1.2309	0.82522	0.67043
96	330660010	83	11	8	38.710	32.100	55.0	1.2059	0.70382	0.58364
96	330660010	83	11	14	46.870	35.520	75.0	1.3195	0.62493	0.47360
96	330660010	83	11	20	18.830	15.110	20.0	1.2462	0.94150	0.75550
96	330660010	83	11	26	19.320	15.650	*	1.2345	*	*
96	330660010	83	12	14	19.370	14.200	26.0	1.3641	0.74500	0.54615
96	330660010	83	12	26	*	30.450	20.0	*	*	1.52250
97	332000003	82	2	4	*	30.210	23.0	*	*	1.31348
97	332000003	82	4	5	*	25.080	25.0	*	*	1.00320
97	332000003	82	4	11	*	25.000	19.0	*	*	1.31579
97	332000003	82	6	10	*	39.340	34.0	*	*	1.15706
97	332000003	82	6	16	*	25.120	24.0	*	*	1.04667
97	332000003	82	6	28	*	73.110	73.0	*	*	1.00151
97	332000003	82	7	10	*	69.930	39.0	*	*	1.79308
97	332000003	82	7	16	*	96.200	65.0	*	*	1.48000
97	332000003	82	8	3	*	24.020	24.0	*	*	1.00083
97	332000003	82	8	15	*	139.500	47.0	*	*	2.96809
97	332000003	82	9	14	*	47.790	46.0	*	*	1.03891
97	332000003	82	9	20	*	34.010	34.0	*	*	1.00029
98	333520001	82	5	17	*	98.900	87.0	*	*	1.13678
98	333520001	82	11	25	*	42.750	35.0	*	*	1.22143
98	333520001	82	12	25	16.250	*	14.0	*	1.16071	*
98	333520001	83	3	19	15.750	*	14.0	*	1.12500	*
98	333520001	83	5	18	54.300	*	46.0	*	1.18043	*
98	333520001	83	7	17	94.930	*	74.0	*	1.28284	*
98	333520001	83	11	2	31.410	*	24.0	*	1.30875	*
101	334680079	83	3	1	*	55.000	48.0	*	*	1.14583
101	334680079	83	5	12	*	74.230	58.0	*	*	1.27983
102	340700010	82	5	23	*	22.960	22.0	*	*	1.04364
102	340700010	82	6	10	*	52.730	44.0	*	*	1.19841
102	340700010	82	6	16	*	42.470	40.0	*	*	1.06175
102	340700010	82	7	22	*	56.220	49.0	*	*	1.14735
102	340700010	82	9	2	*	45.590	44.0	*	*	1.03614
102	340700010	82	11	7	*	82.750	69.0	*	*	1.19928
102	340700010	82	11	13	*	39.420	32.0	*	*	1.23187
102	340700010	82	12	7	*	58.400	40.0	*	*	1.46000
102	340700010	82	12	31	*	78.600	30.0	*	*	2.62000
102	340700010	83	2	5	*	56.370	35.0	*	*	1.61057
104	341160006	82	12	31	19.220	18.170	*	1.0578	*	*
104	341160006	83	1	30	35.910	35.090	*	1.0234	*	*
104	341160006	83	3	7	15.640	14.960	*	1.0455	*	*
105	341160101	83	5	18	26.180	25.620	*	1.0219	*	*
108	361220020	82	4	17	*	57.670	27.0	*	*	2.13593
108	361220020	82	8	15	*	76.110	76.0	*	*	1.00145

SITECODE	SAROAD ID	YEAR	MONTH	DAY	PM10	PM15	TSP	PM10/PM15	PM10/TSP	PM15/TSP
111	3613000041	82	7	22	*	47.500	32.0	*	*	1.48437
113	3616600014	82	3	30	*	53.590	53.0	*	*	1.01113
113	3616600014	82	11	7	*	51.290	51.0	*	*	1.00569
114	3630800010	82	2	4	*	52.390	35.0	*	*	1.49686
116	364340005	83	1	30	*	51.450	34.0	*	*	1.51324
116	364340005	83	4	30	*	54.310	29.0	*	*	1.87276
117	366420012	82	2	22	41.400	49.290	42.0	0.8399	0.98571	1.17357
117	366420012	82	3	12	43.770	51.640	50.0	0.8476	0.87540	1.03280
117	366420012	82	5	5	90.510	74.280	187.0	1.2185	0.48401	0.39722
117	366420012	82	5	11	84.380	93.260	92.0	0.9048	0.91717	1.01370
117	366420012	82	8	3	85.960	77.990	79.0	1.1022	1.08810	0.98722
117	366420012	82	12	7	27.050	34.770	22.0	0.7780	1.22955	1.58045
117	366420012	82	12	19	17.280	16.040	55.0	1.0773	0.31418	0.29164
117	366420012	83	1	12	*	24.040	18.0	*	*	1.33556
117	366420012	83	2	5	23.080	28.070	27.0	0.8222	0.85481	1.03963
117	366420012	83	4	12	*	27.730	26.0	*	*	1.06654
117	366420012	83	4	18	*	11.500	2.0	*	*	5.79000
117	366420012	83	5	30	37.110	51.040	49.0	0.7271	0.75735	1.04163
117	366420012	83	6	5	52.240	*	51.0	*	1.02431	*
117	366420012	83	7	11	46.300	56.460	56.0	0.8200	0.82679	1.00821
117	366420012	83	8	4	86.660	88.140	46.0	0.9832	1.88391	1.91609
117	366420012	83	8	10	50.030	52.460	45.0	0.9537	1.11178	1.16578
117	366420012	83	8	22	47.460	47.000	97.0	1.0098	0.48928	0.48454
117	366420012	83	8	28	39.840	38.630	85.0	1.0313	0.46871	0.45447
117	366420012	83	9	3	123.500	119.100	96.0	1.0369	1.28646	1.24062
117	366420012	83	9	27	95.820	90.150	115.0	1.0629	0.83322	0.78391
117	366420012	83	10	3	58.680	63.720	58.0	0.9209	1.01172	1.09862
117	366420012	83	10	15	42.260	40.750	30.0	1.0371	1.40867	1.35833
117	366420012	83	10	21	16.060	14.650	42.0	1.0962	0.38238	0.34881
117	366420012	83	11	2	85.040	71.870	215.0	1.1832	0.39553	0.33428
117	366420012	83	11	14	20.180	18.810	79.0	1.0728	0.25544	0.23810
117	366420012	83	11	20	54.010	42.730	110.0	1.2640	0.49100	0.38845
117	366420012	83	11	26	28.630	25.530	28.0	1.1214	1.02250	0.91179
117	366420012	83	12	8	23.020	18.100	31.0	1.2718	0.74258	0.58387
117	366420012	83	12	20	36.030	31.380	57.0	1.1482	0.63211	0.55053
117	366420012	83	12	26	17.880	17.870	23.0	1.0006	0.77739	0.77696
118	367760002	82	12	7	67.830	*	65.0	*	1.04354	*
118	367760002	82	12	25	45.000	*	19.0	*	2.36842	*
123	381460015	82	7	28	*	38.390	16.0	*	*	2.39937
124	390100064	82	7	28	*	44.930	43.0	*	*	1.04488
125	390100068	82	12	1	*	24.630	23.0	*	*	1.07087
125	390100068	82	12	19	*	28.860	16.0	*	*	1.80375
125	390100068	83	9	3	*	64.860	62.0	*	*	1.04613
126	390400002	82	7	4	*	112.000	65.0	*	*	1.72308
126	390400002	82	7	10	*	45.780	36.0	*	*	1.27167

SITECODE	SAROAD ID	YEAR	MONTH	DAY	PM10	PM15	TSP	PM10/PM15	PM10/TSP	PM15/TSP
126	390400002	82	7	22	*	114.400	58.0	*	*	1.97241
126	390400002	82	7	28	*	213.200	50.0	*	*	4.26400
126	390400002	82	8	21	*	28.780	27.0	*	*	1.06593
126	390400002	82	10	20	*	55.460	55.0	*	*	1.00836
126	390400002	82	11	25	*	65.650	56.0	*	*	1.17232
126	390400002	83	2	17	*	71.680	65.0	*	*	1.10277
126	390400002	83	5	24	*	111.400	31.0	*	*	3.59355
127	390780725	82	2	10	*	57.140	41.0	*	*	1.39366
127	390780725	82	7	28	*	45.280	40.0	*	*	1.13200
127	390780725	82	8	9	*	30.960	28.0	*	*	1.10571
127	390780725	82	12	13	*	41.790	37.0	*	*	1.12946
127	390780725	83	2	11	*	50.220	40.0	*	*	1.25550
128	396620001	82	2	10	42.190	38.260	59.0	1.1027	0.71508	0.64847
128	396620001	82	3	6	35.510	32.420	51.0	1.0953	0.69627	0.63569
128	396620001	82	5	5	62.070	53.200	102.0	1.1667	0.60853	0.52157
128	396620001	82	5	17	43.620	41.550	72.0	1.0498	0.60583	0.57708
128	396620001	82	7	10	63.290	61.260	81.0	1.0331	0.78136	0.75630
128	396620001	82	7	16	127.200	124.100	169.0	1.0250	0.75266	0.73432
128	396620001	82	7	22	51.840	48.680	76.0	1.0649	0.68211	0.64053
128	396620001	82	8	9	46.040	44.080	52.0	1.0445	0.88538	0.84769
128	396620001	82	8	15	61.490	59.840	92.0	1.0276	0.66837	0.65043
128	396620001	82	9	8	30.470	29.550	46.0	1.0311	0.66239	0.64239
128	396620001	82	9	26	31.180	30.880	46.0	1.0097	0.67783	0.67130
128	396620001	82	10	2	26.670	26.040	37.0	1.0242	0.72081	0.70378
128	396620001	83	4	24	12.650	12.410	16.0	1.0193	0.79062	0.77562
128	396620001	83	6	11	153.700	92.360	157.0	1.6641	0.97898	0.58828
128	396620001	83	7	5	*	114.700	40.0	*	*	2.86750
128	396620001	83	7	17	76.860	67.120	77.0	1.1451	0.99818	0.87169
128	396620001	83	7	23	40.760	41.750	41.0	0.9763	0.99415	1.01829
128	396620001	83	10	3	97.040	81.230	101.0	1.1946	0.96079	0.80426
128	396620001	83	10	27	24.080	22.360	36.0	1.0769	0.66889	0.62111
129	397140003	82	3	12	43.160	41.480	61.0	1.0405	0.70754	0.68000
129	397140003	82	6	28	66.310	79.290	78.0	0.8363	0.85013	1.01654
129	397140003	82	7	16	56.080	53.820	85.0	1.0420	0.65976	0.63318
129	397140003	82	7	22	25.150	11.260	48.0	2.2336	0.52396	0.23458
129	397140003	82	11	25	31.630	25.080	47.0	1.2612	0.67298	0.53362
129	397140003	83	4	6	70.330	32.930	61.0	2.1357	1.15295	0.53984
129	397140003	83	5	30	30.550	29.110	*	1.0495	*	*
129	397140003	83	9	3	38.530	38.030	69.0	1.0131	0.55841	0.55116
129	397140003	83	9	15	18.510	18.460	53.0	1.0027	0.34925	0.34830
133	397140024	82	1	29	*	163.900	55.0	*	*	2.98000
133	397140024	82	4	17	*	51.820	50.0	*	*	1.03640
133	397140024	82	4	23	*	45.070	43.0	*	*	1.04814
133	397140024	82	4	29	*	60.300	30.0	*	*	2.01000
133	397140024	82	5	5	*	133.400	51.0	*	*	2.61569

SITECODE	SAROAD: ID	YEAR	MONTH	DAY	PM10	PM15	TSP	PM10/PM15	PM10/TSP	PM15/TSP
133	397140024	82	5	23	*	33.550	25.0	*	*	1.34200
133	397140024	82	6	16	*	166.000	73.0	*	*	2.27397
133	397140024	82	10	8	*	114.900	88.0	*	*	1.30568
135	397140036	82	1	17	*	30.000	25.0	*	*	1.20000
139	410300012	82	10	26	*	25.160	23.0	*	*	1.09391
139	410300012	83	2	11	*	42.560	28.0	*	*	1.52000
144	451700002	82	11	7	*	297.600	249.0	*	*	1.19518
144	451700002	83	4	30	94.370	*	67.0	*	1.40851	*
144	451700002	83	10	27	98.300	*	49.0	*	2.00612	*
147	452560034	83	9	3	83.230	91.120	36.0	0.9134	2.31194	2.53111
147	452560034	83	10	15	52.890	41.670	76.0	1.2693	0.69592	0.54829
147	452560034	83	10	21	35.050	1.363	56.0	25.7153	0.62589	0.02434
150	460920001	83	6	5	41.910	*	36.0	*	1.16417	*
155	482140007	82	4	11	*	45.130	22.0	*	*	2.05136
155	482140007	82	8	9	*	53.430	49.0	*	*	1.09041
156	482630001	82	12	19	*	37.910	20.0	*	*	1.89550
156	482630001	83	4	24	6.202	*	5.7	*	1.08807	*
156	482630001	83	7	29	65.090	*	43.2	*	1.50671	*
158	491840057	83	11	2	61.000	*	45.0	*	1.35556	*
160	492040013	82	7	4	*	55.170	21.0	*	*	2.62714
161	500280004	82	7	4	*	46.450	29.0	*	*	1.60172
161	500280004	82	8	27	*	38.080	38.0	*	*	1.00210
161	500280004	83	3	25	*	33.850	27.0	*	*	1.25370
161	500280004	83	4	12	*	32.690	26.0	*	*	1.25731
162	502000002	83	5	30	*	54.260	44.0	*	*	1.23318
163	502120002	82	12	13	*	45.480	22.0	*	*	2.06727
163	502120002	83	5	18	*	70.640	70.0	*	*	1.00914
164	510240002	82	3	6	*	40.080	37.0	*	*	1.08324
164	510240002	82	12	1	*	41.630	33.0	*	*	1.26152

Appendix C

DETERMINATION OF CRITERIA FOR A VALID ANNUAL MEAN

Appendix C

DETERMINATION OF CRITERIA FOR A VALID ANNUAL MEAN

The variability of an annual ratio depends on the number of sample days used to calculate the ratio. An annual ratio calculated from only a few sample days is much more variable than an annual ratio calculated from 20 or 30 sample days. We decided to include annual ratios, i.e., consider them valid, only if at least 10 sample days contributed to the annual average ratio. Justification of 10 days as the cutoff point is described here.

The annual average is calculated by dividing an annual average numerator by the annual average demoninator (not by calculating the average across the individual ratios). Suppose the annual averages are composed of n 24-hour measurements and denote the annual average numerator by $\bar{x}$ and the annual average demoninator by $\bar{y}$. Then the variance of the annual average ratio can be approximated as

$$\hat{\mathrm{Var}}\,\frac{\bar{x}}{\bar{y}} \simeq \frac{1}{n}\left(\frac{\bar{x}}{\bar{y}}\right)^2\left(\frac{S_x^2}{\bar{x}^2}+\frac{S_y^2}{\bar{y}^2}-2\frac{S_{x\,y}}{\bar{x}\,\bar{y}}\right),$$

where S_x^2, S_y^2, and $S_{x\,y}$ are, respectively, the sample variance of x and y and their sample covariance. The approximation is valid provided n is large enough for $\bar{x}$ and $\bar{y}$ to have an approximate joint normal distribution (Ku, 1979). The approximation ignores any autocorrelation in the 24-hour measurements, but this is probably not large, particularly for a once-per-six-day monitoring scheduling, which is common.

If we have K annual average ratios R_1, ..., R_K, from K site-years, with n_1, ..., n_K sample days in the respective numerators and denominators, then the variance of the average of these annual ratios is

$$\text{Var}(\bar{R}) = \text{Var}\left[\frac{1}{K}(R_1 + \ldots + R_K)\right]$$

$$= \frac{1}{K^2}\,\text{Var}(R_1 + \ldots + R_K)$$

$$= \frac{1}{K^2}[\text{Var}(R_1) + \ldots + \text{Var}(R_K)]$$

if we make the reasonable assumption that the K ratios are uncorrelated (any actual correlation is probably quite small). If we make the further simplifying assumption that the variance and covariance of 24-hour collocated measurements are approximately the same for all sites, then the variance of the average across the annual average ratios is

$$\text{Var}(\bar{R}) \approx \frac{C}{K^2}\left(\frac{1}{n_1} + \ldots + \frac{1}{n_K}\right) .$$

We estimated the mean and variance of the average ratio for each of the ratios PM10/PM15, PM10/TSP, and PM15/TSP, for all possible cutoff values of the numbers of sample days required, and then examined the results to find the most reasonable cutoff value. There are trade-offs to be made in the choice of a cutoff value, for as we require more 24-hour measurements per site, the site-specific ratios included in our analysis become more reliable, but the analysis rests on those few sites having a large number of measurements.

The complete calculations for the PM10/TSP annual average ratio are given in Table C-1. There are 53 annual average PM10/TSP ratios from the

1982-1983 IP Network and SAROAD data; they are sorted in the table by the number of sample days contributing to each ratio. The first column, K, is the number of ratios included in each successive calculation. The second column is the annual average PM10/TSP ratio, and the third column is the number of sampling days for the ratio. The fourth column is the inverse of the number of samples; these are successively summed in the fifth column. The sixth column is the inverse of the square of K, the number of ratios. The seventh column is the approximate variance of the mean ratio (calculated from K ratios) without the common multiplier C, i.e.,

$$C \cdot \mathrm{Var}(\bar{R}) \approx \frac{1}{K^2}\left(\frac{1}{n_1} + \cdots + \frac{1}{n_K}\right),$$

and the seventh column is its square root, i.e., the standard deviation without the multiplier $\sqrt{C}$. The final column is the successive means of the ratios as each additional ratio is introduced.

The successive means and variances about the means are plotted against K, the number of annual ratios included, in Figures C-1 and C-2, respectively. The vertical line on both plots corresponds to a cutoff value of 10 sampling days required. At this cutoff the mean seems relatively stable. The variance decreases continuously until ratios with about 10 samples are included; then the variance increases as annual ratios based on only a few samples are included. Of the 53 annual average PM10/TSP ratios available, 14 (26 percent) are based on fewer than 10 sample days and were excluded from the analyses in this report.

Similar plots for annual average PM10/PM15 ratios are in Figures C-3 and C-4 for the mean ratio and its variance, respectively. Again the vertical line indicates the value of K that corresponds to at least 10 sample days contributing to the annual average. At 10 samples the mean is relatively stable and the variance begins its upward trend. Of the 20 annual ratios available, 2 (10 percent) have fewer than 10 sample days and are therefore excluded. Similar patterns can be seen in the mean PM15/TSP and variance

of the mean plots in Figures C-5 and C-6, respectively. Of the 170 available PM15/TSP, 37 (22 percent) are based on fewer than 10 samples and are therefore considered invalid.

Reference

Ku, H. H. 1966. Notes on the use of Propagation Error Formulas. Precision Measurement and Calibration: Selected NBS Papers on Statistical Concepts and Procedures, Harry H. Ku, ed. National Bureau of Standards, United States Department of Commerce (NBS Special Publication 300-Volume 1).

TABLE C-1. Calculation of the mean and approximate variance of the average of the annual average PM10/TSP ratios as ratios with fewer observations are successively added.

K	ratio	nsamp	1/n	sum(1/n)	1/K**2	var(m)	stdev(m)	mean
1	0.490300	53	0.01887	0.01887	1.000000	0.0188679	0.137361	0.490300
2	0.280121	52	0.01923	0.03810	0.250000	0.0095247	0.097594	0.385210
3	0.459348	51	0.01961	0.05771	0.11111	0.0064118	0.080074	0.409923
4	0.545236	50	0.02000	0.07771	0.06250	0.0048567	0.069690	0.443751
5	0.460069	50	0.02000	0.09771	0.04000	0.0039083	0.062516	0.447015
6	0.678036	49	0.02041	0.11811	0.02778	0.0032810	0.057280	0.485518
7	0.618555	49	0.02041	0.13852	0.02041	0.0028270	0.053170	0.504524
8	0.412549	49	0.02041	0.15893	0.01563	0.0024833	0.049833	0.493027
9	0.474548	47	0.02128	0.18021	0.01235	0.0022248	0.047168	0.490974
10	0.510146	47	0.02128	0.20148	0.01000	0.0020148	0.044887	0.492891
11	0.494091	46	0.02174	0.22322	0.00826	0.0018448	0.042951	0.493000
12	0.652375	46	0.02174	0.24496	0.00694	0.0017011	0.041245	0.506281
13	0.660701	46	0.02174	0.26670	0.00592	0.0015781	0.039726	0.518160
14	0.321994	45	0.02222	0.28892	0.00510	0.0014741	0.038394	0.504148
15	0.464458	45	0.02222	0.31115	0.00444	0.0013829	0.037187	0.501502
16	0.623726	43	0.02326	0.33440	0.00391	0.0013063	0.036142	0.509141
17	0.402997	43	0.02326	0.35766	0.00346	0.0012376	0.035179	0.502897
18	0.584067	41	0.02439	0.38205	0.00309	0.0011792	0.034339	0.507407
19	0.585117	38	0.02632	0.40836	0.00277	0.0011312	0.033633	0.511497
20	0.560211	38	0.02632	0.43468	0.00250	0.0010867	0.032965	0.513932
21	0.450649	38	0.02632	0.46100	0.00227	0.0010453	0.032332	0.510919
22	0.685719	36	0.02778	0.48877	0.00207	0.0010099	0.031778	0.518864
23	0.557745	36	0.02778	0.51655	0.00189	0.0009765	0.031248	0.520555
24	0.545000	35	0.02857	0.54512	0.00174	0.0009464	0.030764	0.521573
25	0.586320	33	0.03030	0.57543	0.00160	0.0009207	0.030343	0.524163
26	0.404159	33	0.03030	0.60573	0.00148	0.0008960	0.029934	0.519548
27	0.515386	29	0.03448	0.64021	0.00137	0.0008782	0.029635	0.519393
28	0.575146	29	0.03448	0.67469	0.00128	0.0008606	0.029336	0.521385
29	0.499870	25	0.04000	0.71469	0.00119	0.0008498	0.029152	0.520643
30	0.432125	21	0.04762	0.76231	0.00111	0.0008470	0.029104	0.517692
31	0.458420	21	0.04762	0.80993	0.00104	0.0008428	0.029031	0.515780
32	0.404100	18	0.05556	0.86549	0.00098	0.0008452	0.029072	0.512290
33	0.325678	16	0.06250	0.92799	0.00092	0.0008521	0.029192	0.506635
34	0.508616	16	0.06250	0.99049	0.00087	0.0008568	0.029272	0.506693
35	0.440820	14	0.07143	1.06192	0.00082	0.0008669	0.029443	0.504811
36	0.416779	14	0.07143	1.13334	0.00077	0.0008745	0.029572	0.502366
37	0.345503	14	0.07143	1.20477	0.00073	0.0008800	0.029665	0.498126
38	0.432967	14	0.07143	1.27620	0.00069	0.0008838	0.029729	0.496412
39	0.542340	11	0.09091	1.36711	0.00066	0.0008988	0.029980	0.497589
40	0.470117	6	0.16667	1.53378	0.00063	0.0009586	0.030961	0.496903
41	0.754780	6	0.16667	1.70044	0.00059	0.0010116	0.031805	0.503192
42	0.578608	6	0.16667	1.86711	0.00057	0.0010585	0.032534	0.504988
43	0.486979	5	0.20000	2.06711	0.00054	0.0011180	0.033436	0.504569
44	0.607438	5	0.20000	2.26711	0.00052	0.0011710	0.034220	0.506907
45	0.532089	5	0.20000	2.46711	0.00049	0.0012183	0.034905	0.507467
46	0.667021	4	0.25000	2.71711	0.00047	0.0012841	0.035834	0.510935
47	0.791877	4	0.25000	2.96711	0.00045	0.0013432	0.036650	0.516913
48	0.666686	2	0.50000	3.46711	0.00043	0.0015048	0.038792	0.520033
49	0.480677	2	0.50000	3.96711	0.00042	0.0016523	0.040648	0.519230
50	0.747551	2	0.50000	4.46711	0.00040	0.0017868	0.042271	0.523796
51	0.508341	2	0.50000	4.96711	0.00038	0.0019097	0.043700	0.523493
52	0.322857	1	1.00000	5.96711	0.00037	0.0022068	0.046976	0.519635
53	0.635886	1	1.00000	6.96711	0.00036	0.0024803	0.049802	0.521828

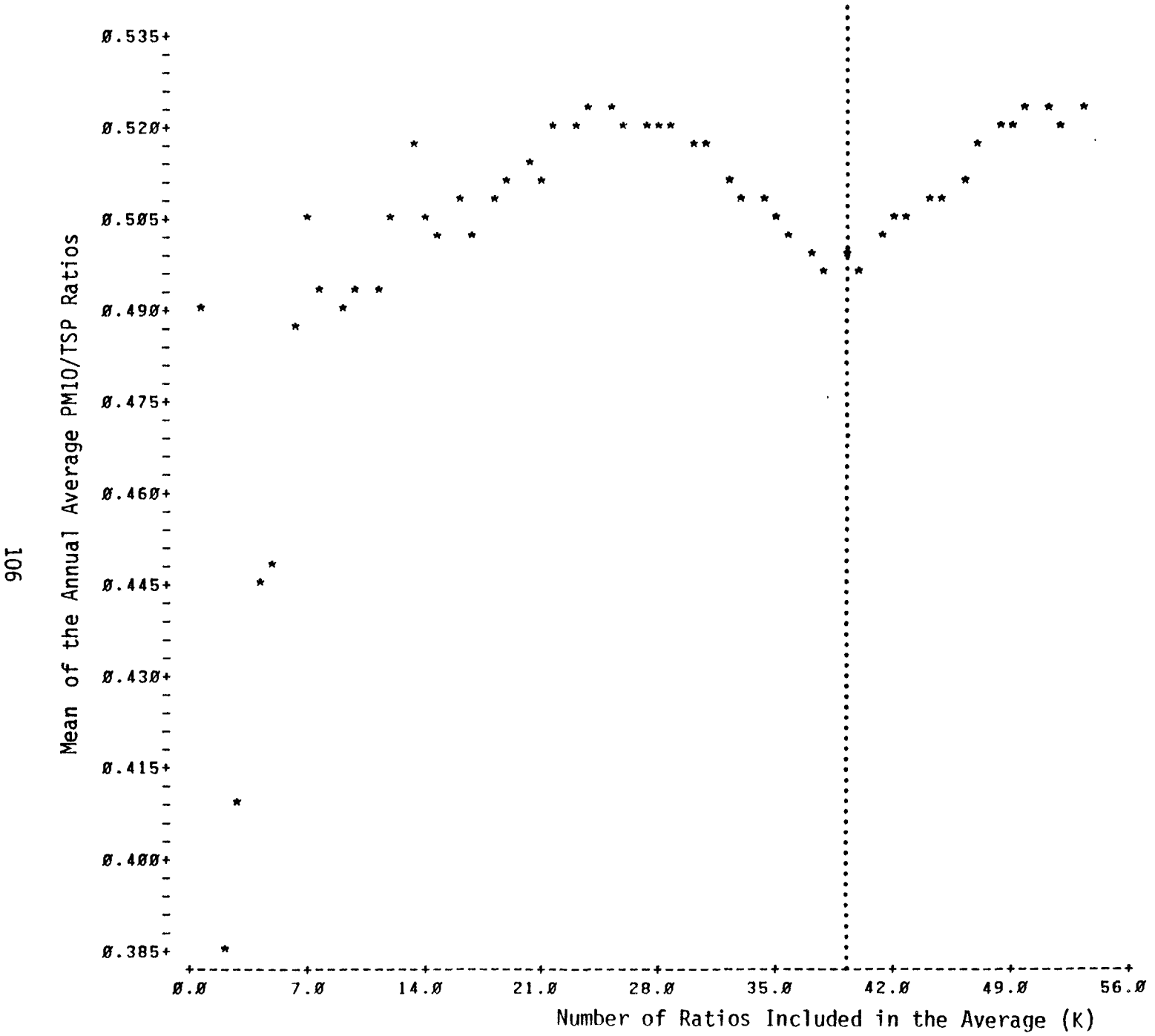

FIGURE C-1. Mean annual average PM10/TSP ratio versus K, the number of ratios included in the average. As K increases, ratios are composed of decreasing numbers of samples.

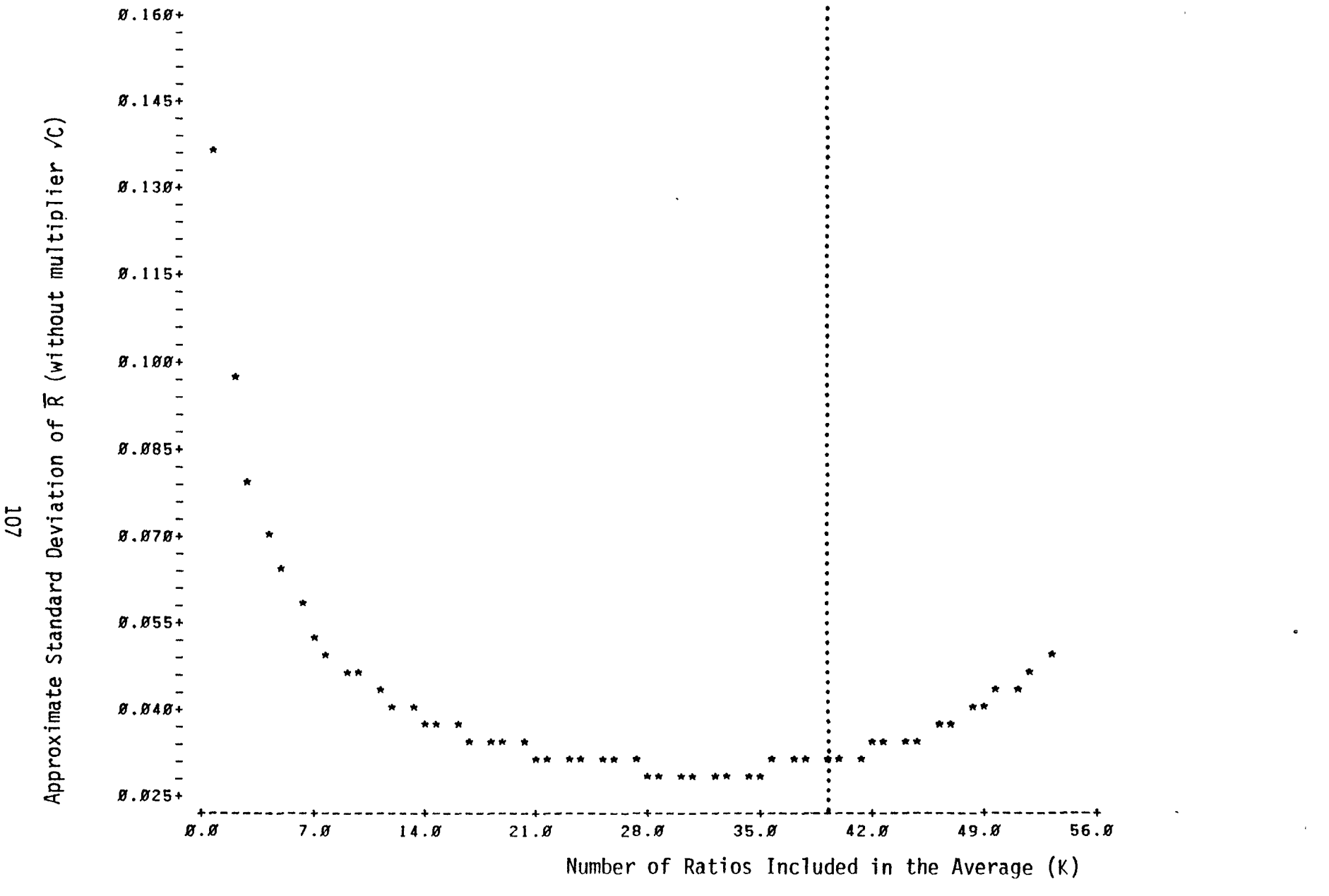

FIGURE C-2. Approximate standard deviation of $\bar{R}$, the mean of the annual average PM10/TSP ratios, versus K, the number of ratios included in the average. As K increases, ratios are composed of decreasing numbers of samples.

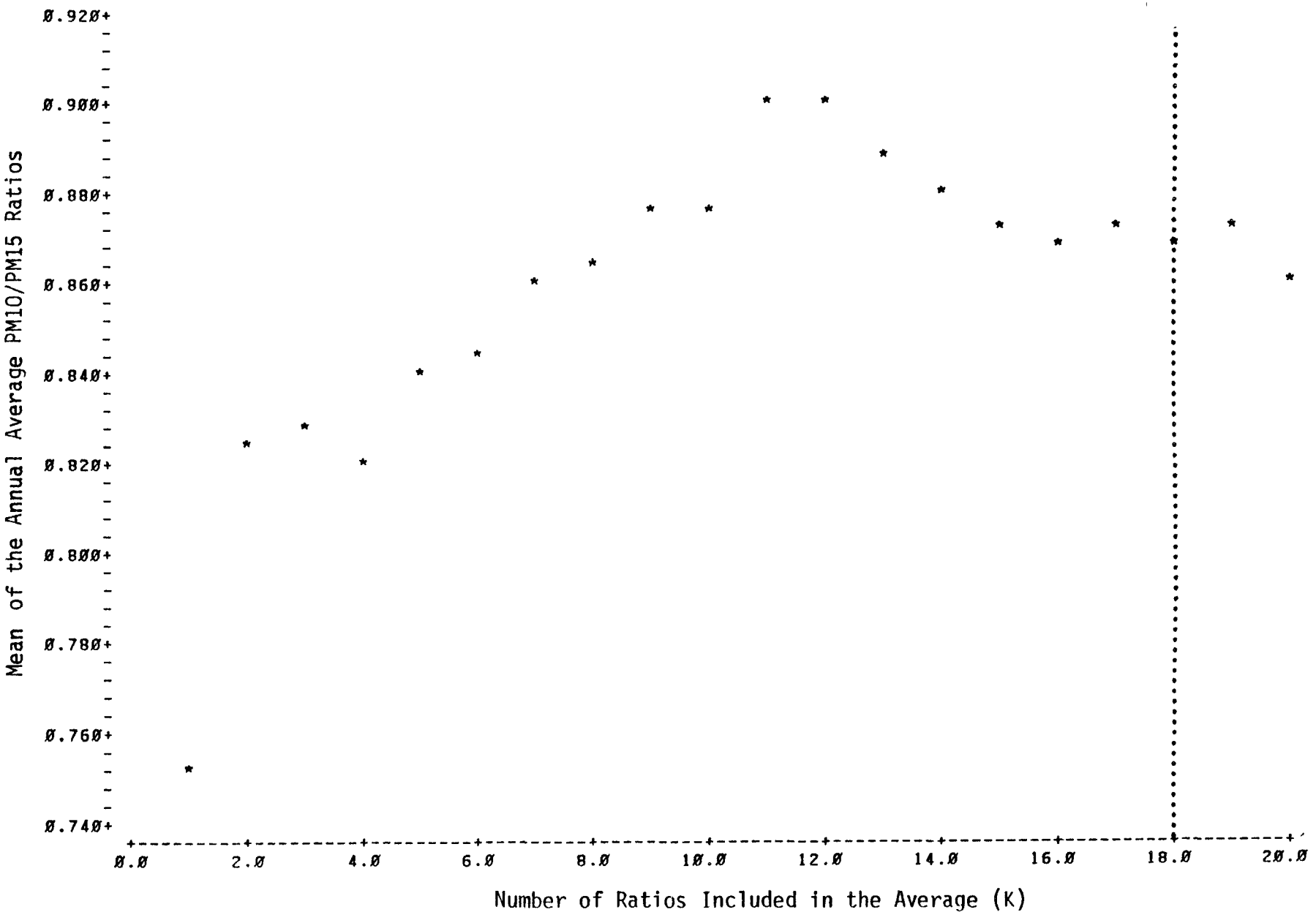

FIGURE C-3. Mean annual average PM10/PM15 ratio versus K, the number of ratios included in the average. As K increases, ratios are composed of decreasing numbers of samples.

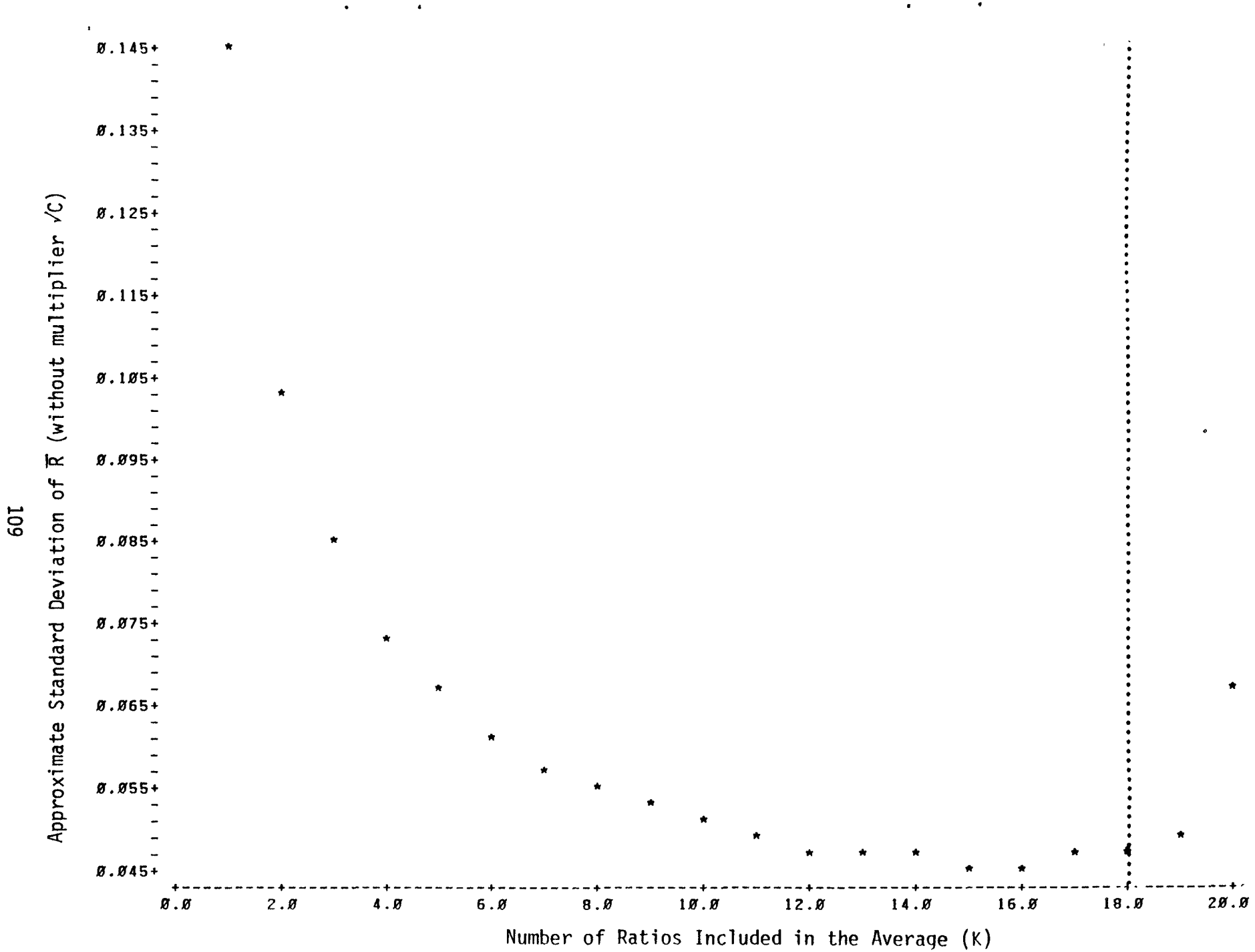

FIGURE C-4. Approximate standard deviation of $\bar{R}$, the mean of the annual average PM10/PM15 ratios, versus K, the number of ratios included in the average. As K increases, ratios are composed of decreasing numbers of samples.

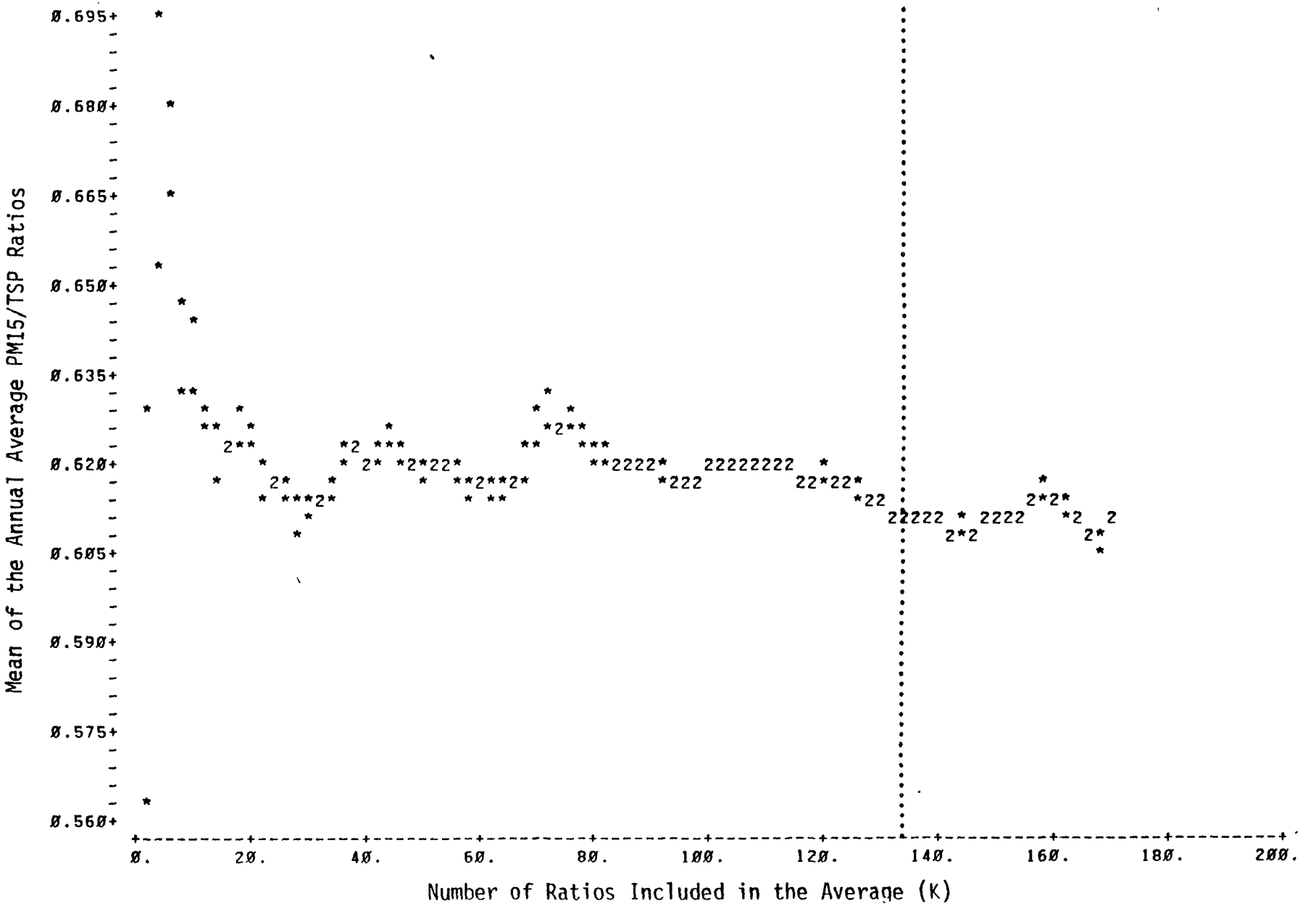

FIGURE C-5. Mean annual average PM15/TSP ratio versus K, the number of ratios included in the average. As K increases, ratios are composed of decreasing numbers of samples.

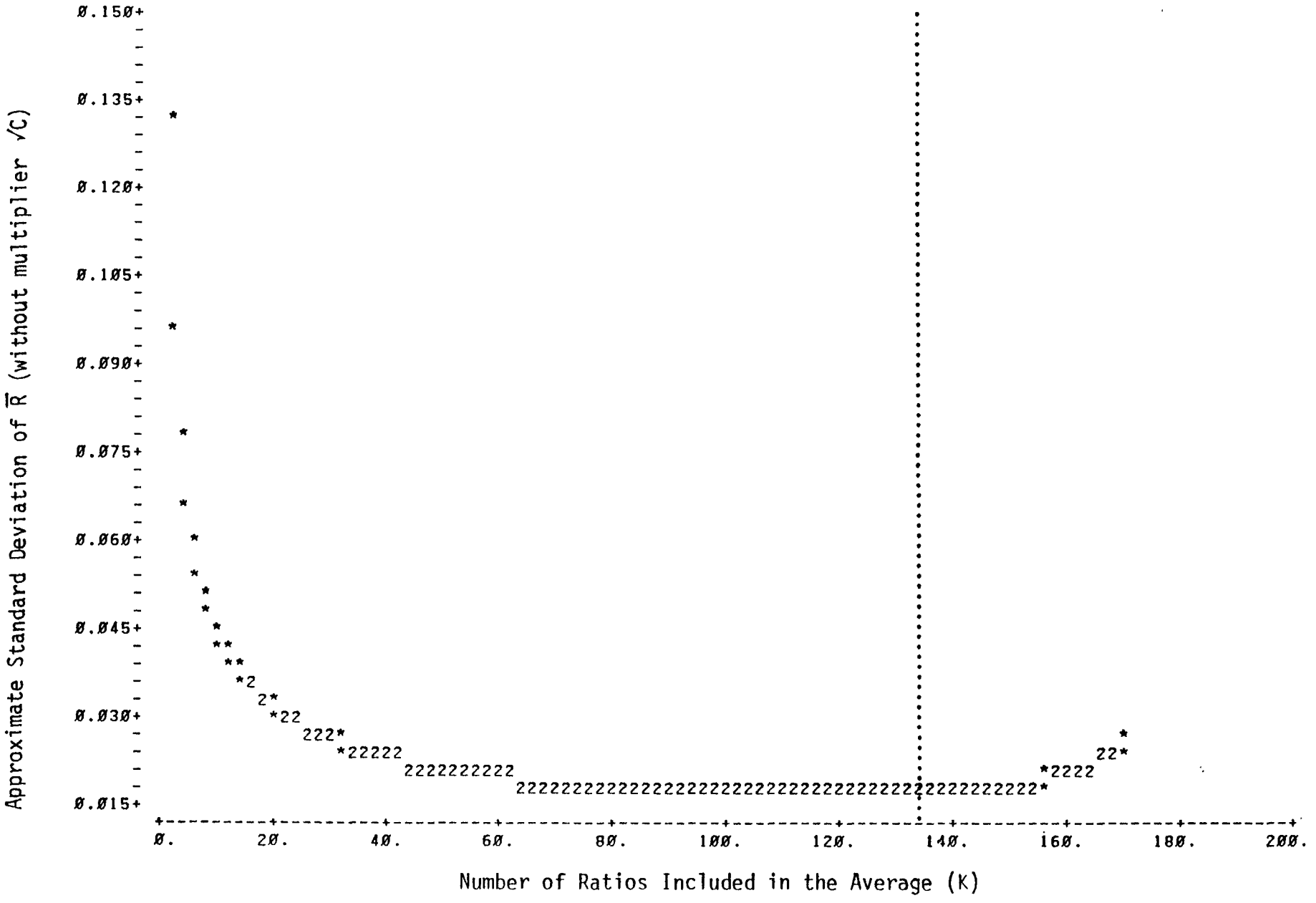

FIGURE C-6. Approximate standard deviation of $\bar{R}$, the mean of the annual average PM15/TSP ratios, versus K, the number of ratios included in the average. As K increases, ratios are composed of decreasing numbers of samples.

Appendix D

RECOMMENDATION OF A CONSTANT TO USE FOR CONVERTING PM15 TO PM10 FOR USE WITH PM15/TSP RATIOS

Appendix D

RECOMMENDATION OF A CONSTANT TO USE FOR CONVERTING PM15 TO PM10 FOR USE WITH PM15/TSP RATIOS

In some situations PM15/TSP ratios may be used to determine attainment status for the PM10 standards rather than PM10/TSP ratios. In this case, a multiplier is needed to convert the PM15 data to PM10 data. The distributions of PM10/PM15 ratios are variable and should not be considered constant. However, a single conversion factor has been used in the past and in this study because of the limited availability of PM10 data on days having both PM15 and TSP data. This appendix contains justification for our recommendation of 0.85 as the multiplier.

Four data sets of PM10/PM15 ratios were examined: all 24-hour ratios, 24-hour ratios corresponding to high TSP values ($> 100\ \mu g/m^3$), all valid annual ratios, and valid annual ratios corresponding to high annual average TSP concentrations ($> 55\ \mu g/m^3$). For each of these four data sets, the number of available ratios, mean and median ratio, and TSP to PM10 regression slope are given in Table D-1. The regression slope is the linear regression estimate of the best multiplier for predicting PM10 from PM15; i.e., it is the constant a that minimizes the sum of squared differences

$$\sum_{i=1}^{n} (PM10 - a \cdot PM15)^2 \quad .$$

The multipliers in the table cover a very small range from a minimum of 0.84 to a maximum of 0.87. Recognizing that most of the sample sizes in Table D-1 are small and therefore that the variances about these estimates are not small, we chose 0.85 as a rounded number in the middle of the range of multipliers.

TABLE D-1. Summary statistics for 24-hour and annual PM10/TSP ratios.

Date Set	Number of Ratios	Mean Ratio	Median Ratio	Regression Slope
All 24-hour ratios	567	0.87	0.87	0.85
24-hour ratios corresponding to TSP > 100 $\mu g/m^3$	122	0.85	0.84	0.84
All valid annual ratios	18	0.87	0.87	0.84
Valid annual ratios corresponding to annual average TSP > 55 $\mu g/m^3$	17	0.85	0.86	0.84

Appendix E

DETERMINATION OF CUTOFF VALUES FOR HIGH TSP

Appendix E

DETERMINATION OF CUTOFF VALUES FOR HIGH TSP

PM10/TSP and PM15/TSP ratios, both 24-hour and annual, are more variable at low TSP concentrations. Most of the usually large ratios occur at low TSP concentrations, where sampling errors have a greater relative effect than at high TSP concentrations. Because of these outlying ratios, the variances of the ratio distributions are large. Rather than arbitrarily omit all ratios greater than a specified level, we chose to consider ratios corresponding to a TSP concentration above a fixed level. In this appendix, justification for the choice of the high TSP cutoff values of 100 $\mu g/m^3$ for 24-hour concentrations and 55 $\mu g/m^3$ for annual average concentrations is given.

Figure E-1 contains a plot of 24-hour PM10/TSP ratios versus 24-hour TSP concentrations. The extremely high PM10/TSP ratios occur at less than 60 $\mu g/m^3$ TSP, but even beyond that cutoff value there are ratios greater than one. Figure E-2 is a plot of PM15/TSP ratios versus TSP, also 24-hour concentrations. Here, too, the largest ratios occur at low TSP concentrations, below approximately 90 $\mu g/m^3$ TSP.

The higher the TSP cutoff value, the more ratios are excluded. In general variances decrease with larger sample sizes, but in these cases the variance about the mean most likely decreases as more low TSP ratios are excluded, because they are the more variable ratios. A wide range of cutoff values for TSP concentrations, from 50 to 150 $\mu g/m^3$, was considered. The means and standard deviations of the 24-hour PM10/TSP and PM15/TSP ratio distributions corresponding to each of the 24-hour TSP cutoff values considered are shown in Figures E-3 and E-4, respectively.

Because of the relationship between high ratios and low TSP concentrations, the mean ratio decreases almost monitonically. The standard deviations of the ratio distributions, however, decrease when high ratios are omitted but then increase as there are fewer and fewer ratios. For both PM10/TSP and PM15/TSP ratios, the standard deviation is minimized at 100 $\mu g/m^3$, the chosen cutoff value.

Annual average PM10/TSP and PM15/TSP ratios are also dependent on TSP levels (i.e., annual average TSP), but the relationships shown in Figures E-5 and E-6 are not as striking as those of E-1 and E-2, because averaging over multiple sample days reduces the effects of outliers. There are no annual average PM10/TSP ratios above one, and only one annual average PM15/TSP ratio above one. There are, however, some unusually high ratios. Cutoff values for annual average TSP of 40 to 80 $\mu g/m^3$ were considered. The means and standard deviations of the resulting ratio distributions for each cutoff value considered are shown in Figures E-7 and E-8, respectively. As with the 24-hour ratios, the mean decreases almost monotonically with increasing cutoff values. Except for the very highest cutoff values above 75 $\mu g/m^3$, where most of the data are excluded, the standard deviations of the ratio distributions are minimized at 55 $\mu g/m^3$, the chosen cutoff value.

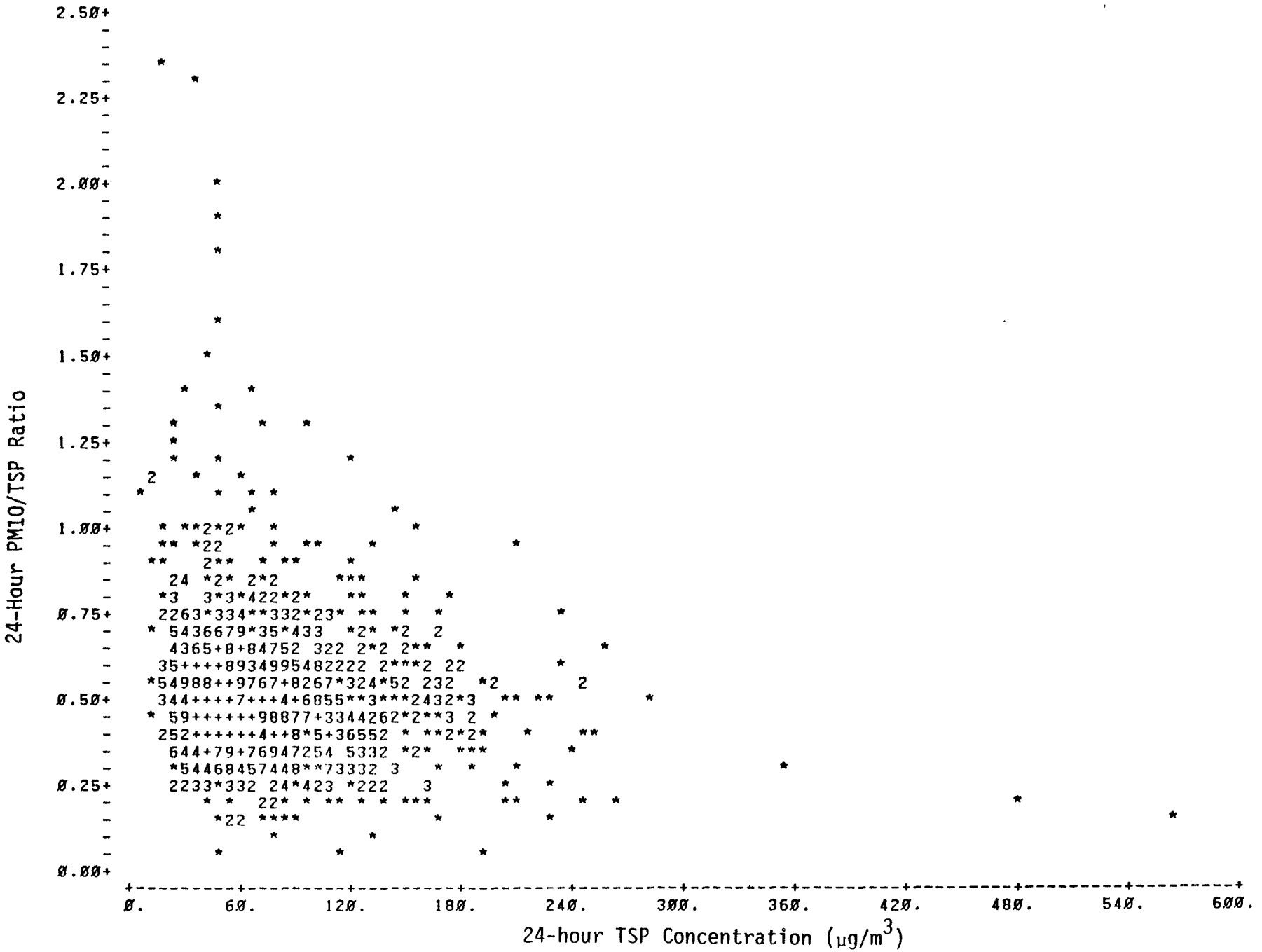

FIGURE E-1. 24-hour PM10/TSP ratio versus TSP concentration.

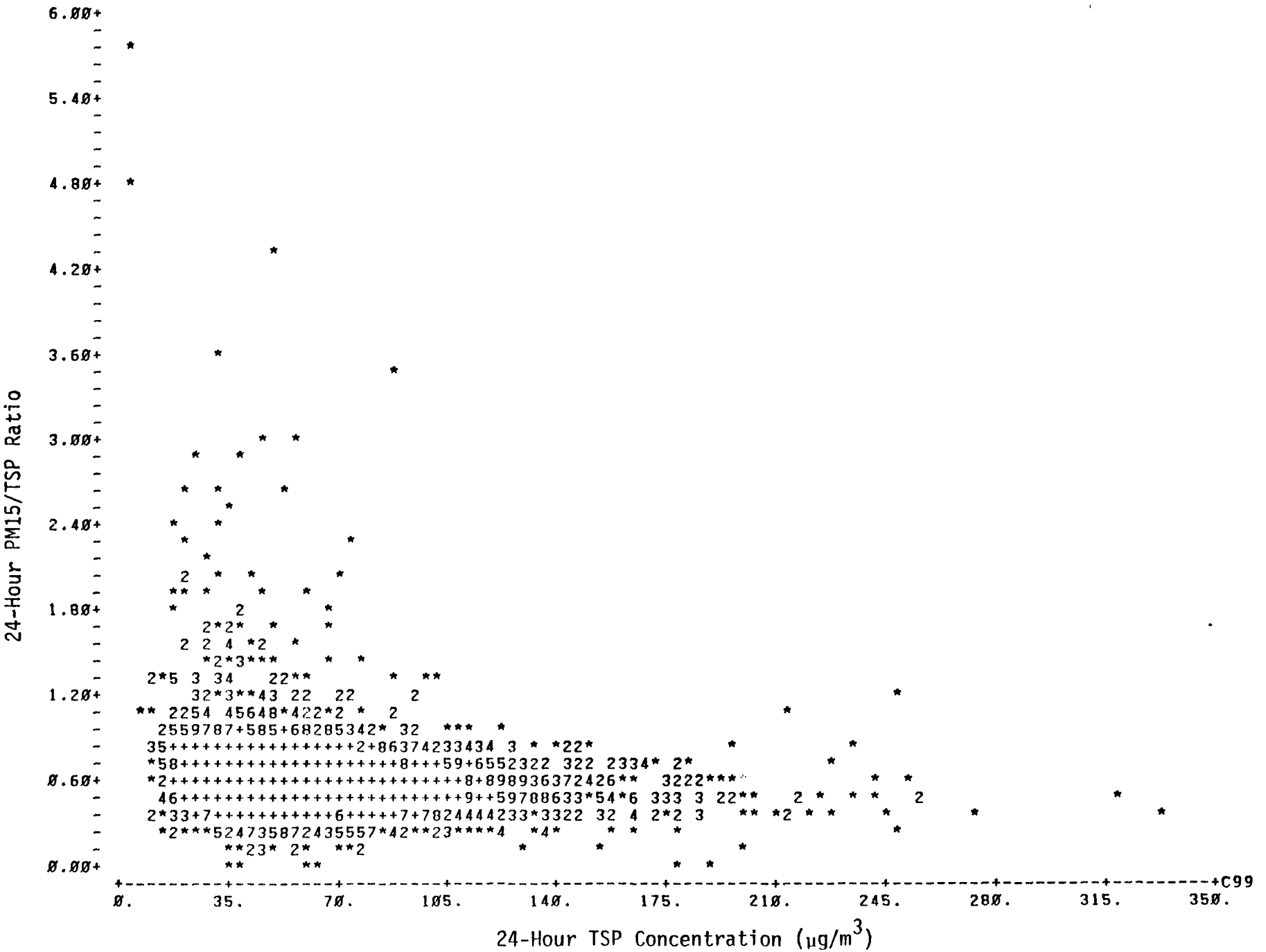

FIGURE E-2. 24-hour PM15/TSP ratio versus TSP concentration.

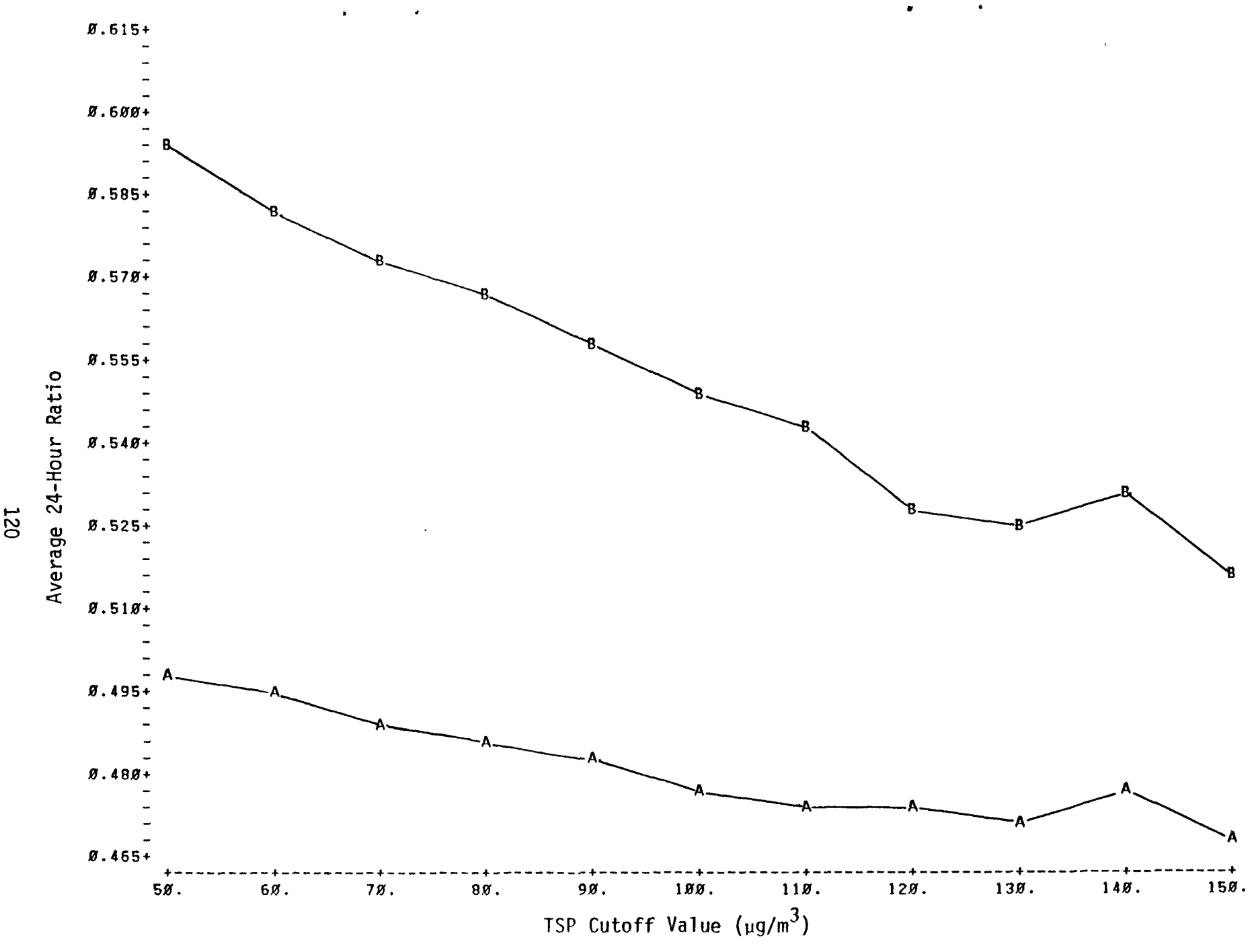

FIGURE E-3. Average 24-hour ratios calculated from days with TSP concentrations above different cutoff values. (A = PM10/TSP, B = PM15/TSP)

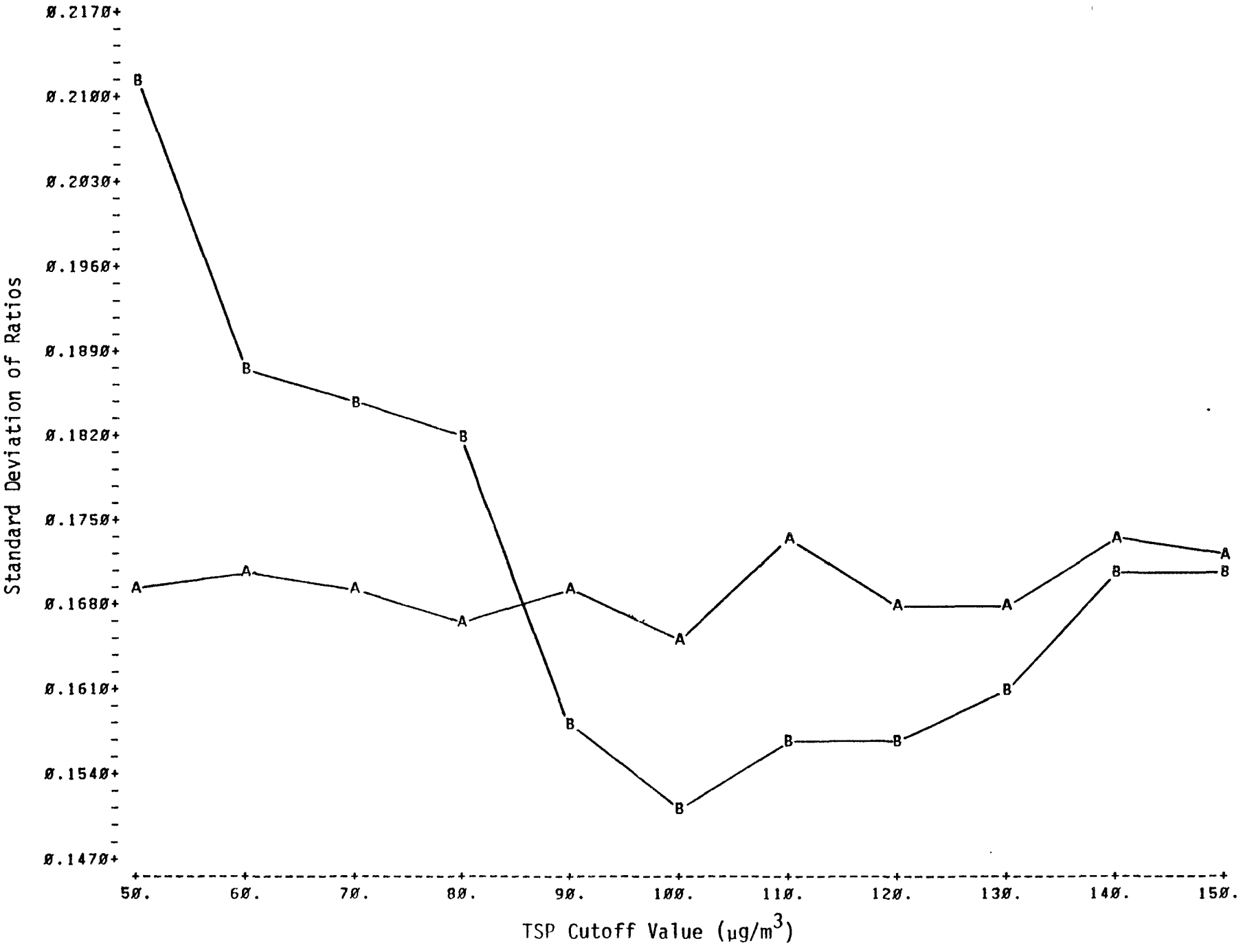

FIGURE E-4. Standard deviation of 24-hour ratios calculated from days with TSP concentrations above different cutoff values. (A = PM10/TSP, B = PM15/TSP)

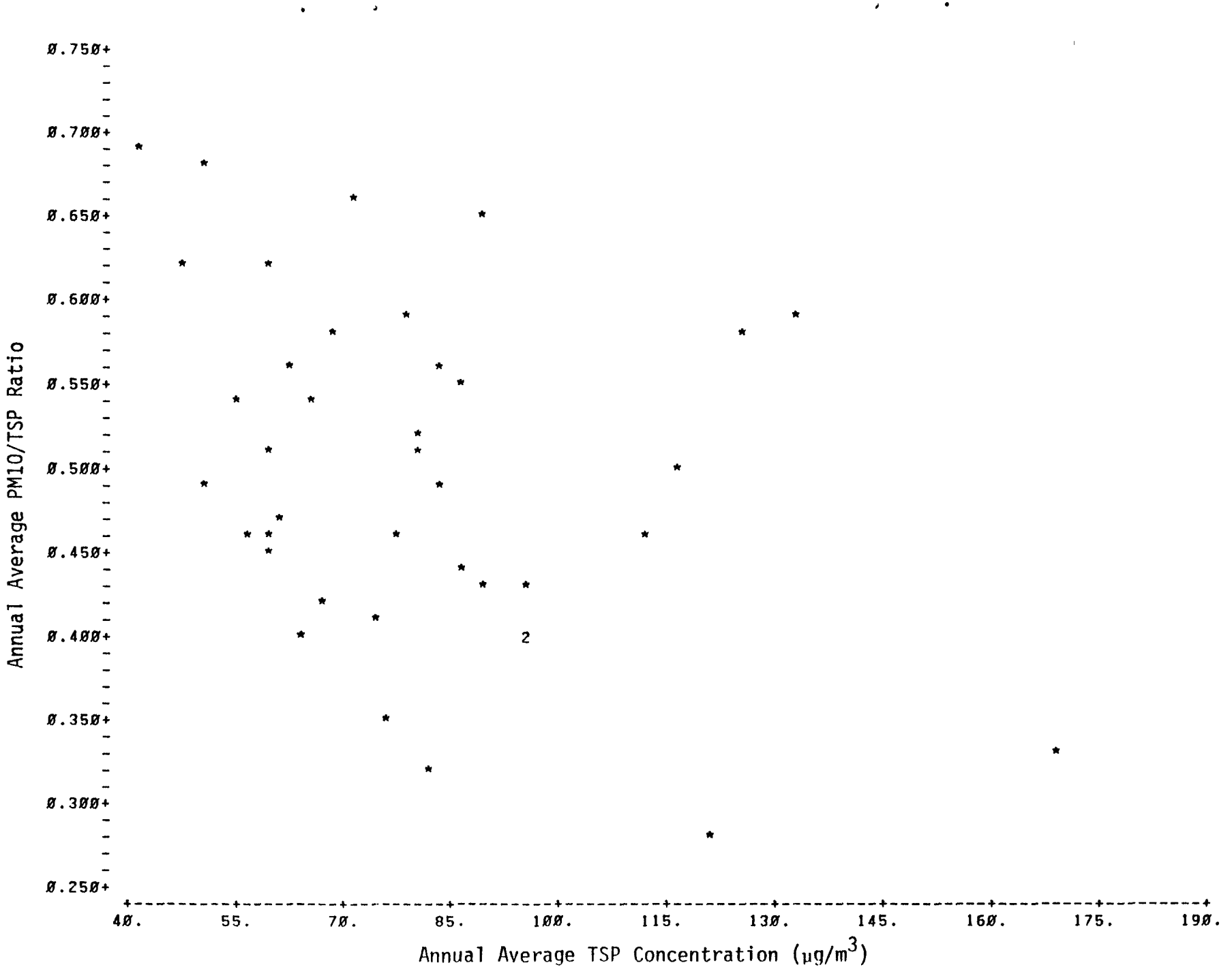

FIGURE E-5. Annual average PM10/TSP versus TSP concentration.

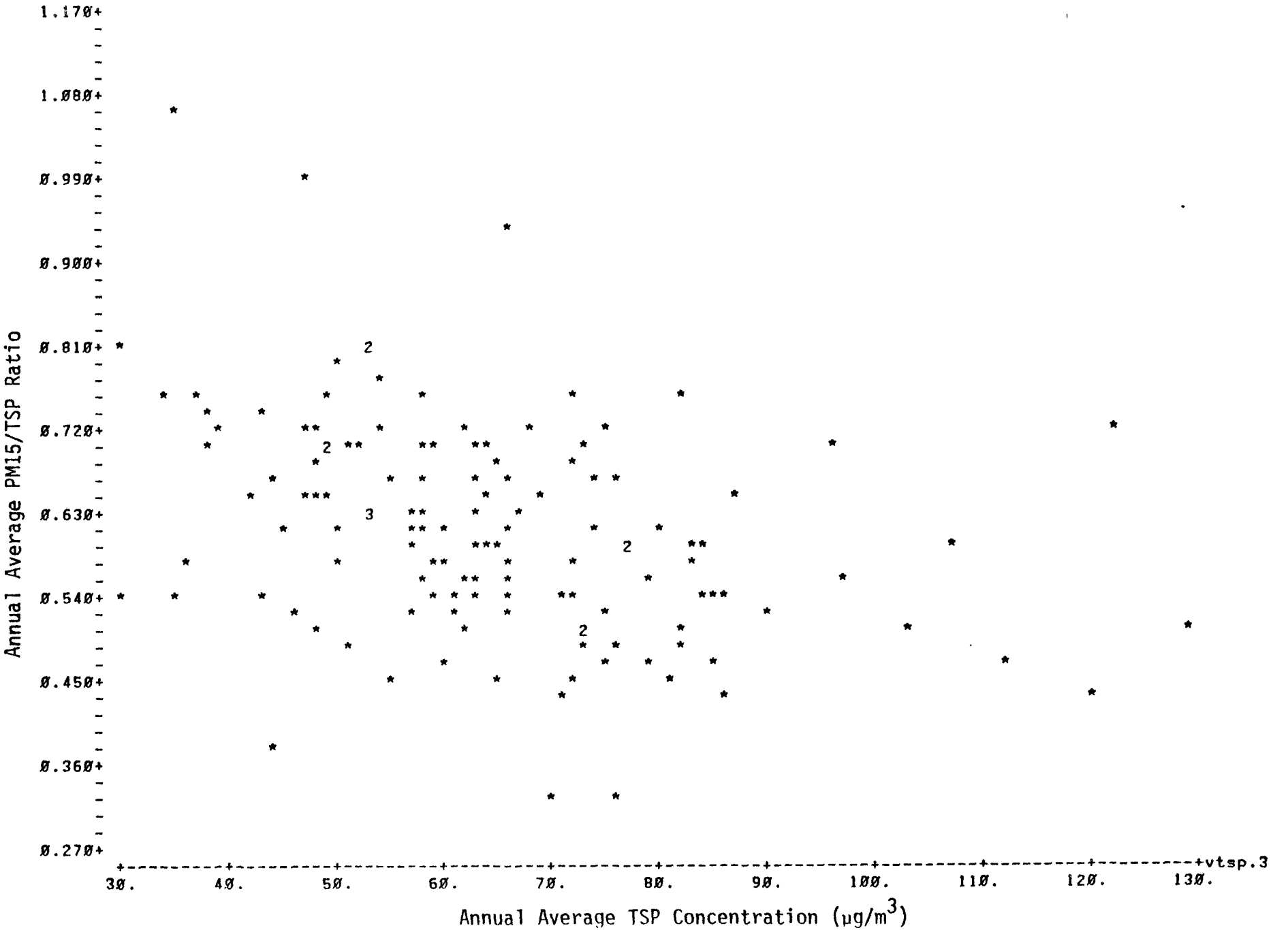

FIGURE E-6. Annual average PM15/TSP ratio versus TSP concentration.

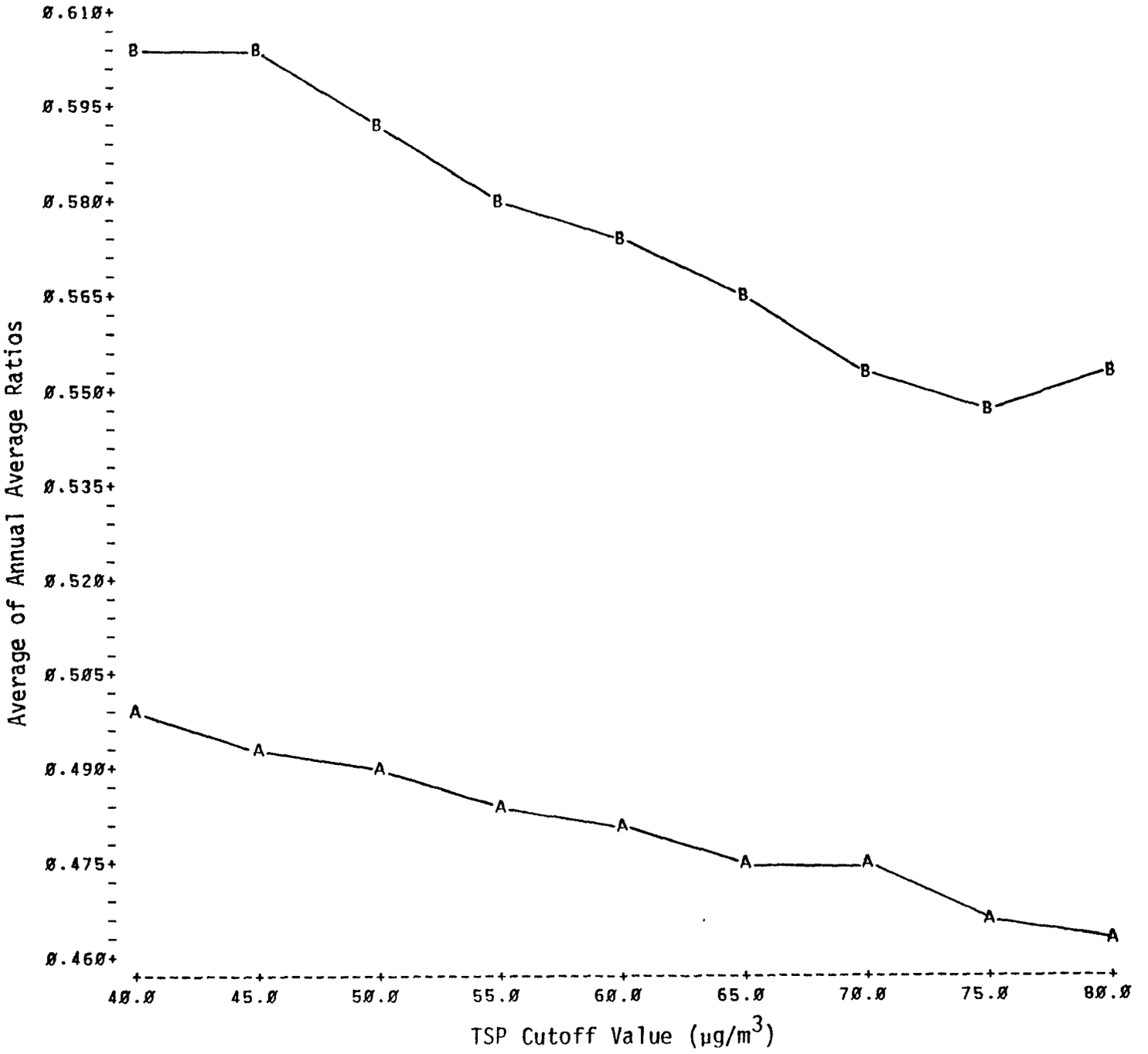

FIGURE E-7. Average of annual average ratios calculated from site-years with annual average TSP concentrations above different cutoff values. (A = PM10/TSP, B = PM15/TSP)

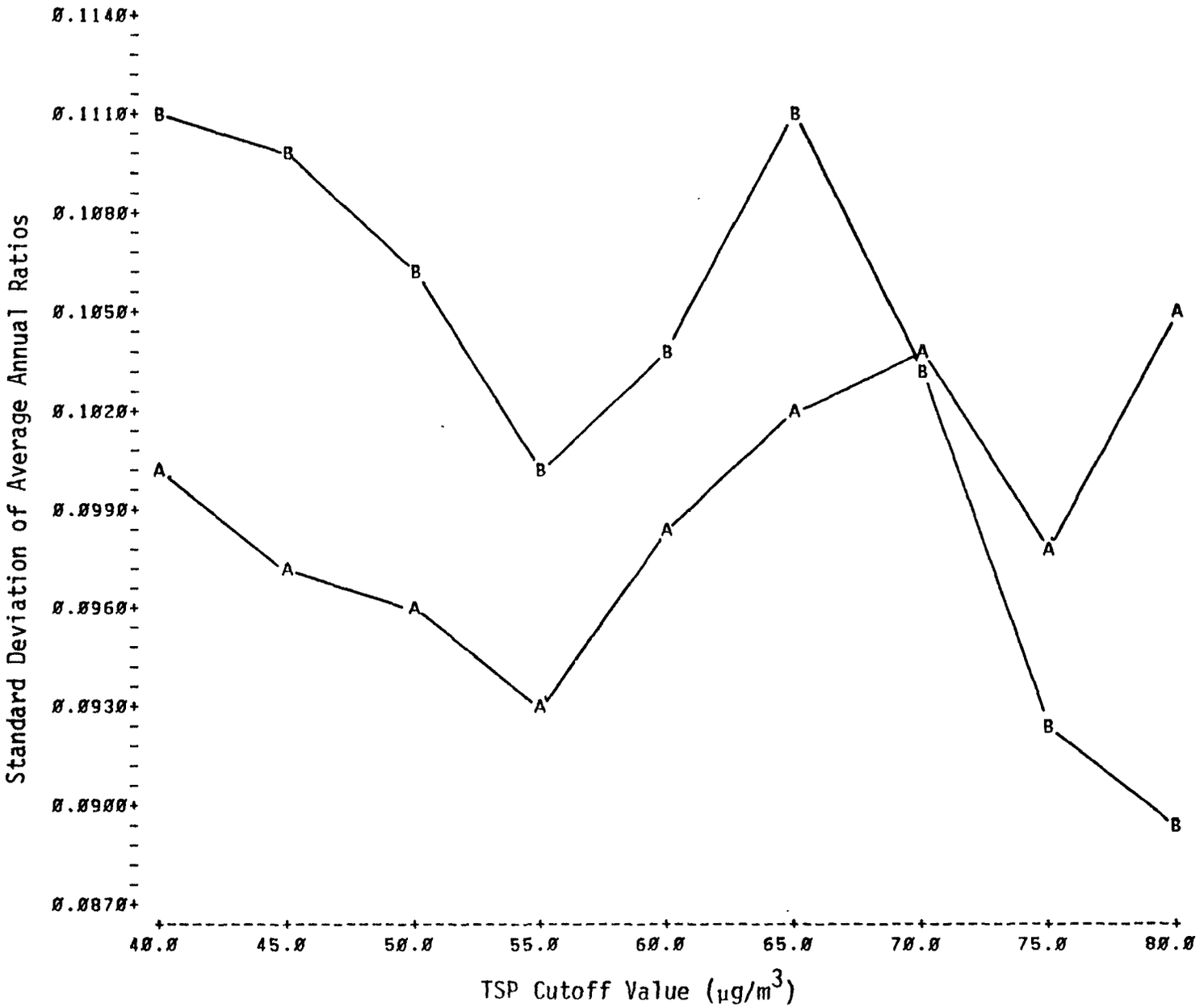

FIGURE E-8. Standard deviations of annual average ratios calculated from site-years with annual average TSP concentrations above different cutoff values. (A = PM10/TSP, B = PM15/TSP)

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2	3. RECIPIENT'S ACCESSION NO.
EPA-450/4-85-010		

4. TITLE AND SUBTITLE	5. REPORT DATE
An Examination of 1982-83 Particulate Matter Ratios and Their Use In The Estimation of PM_{10} NAAQS Attainment Status	August 1985
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.
A.K. Pollack et al.	
9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT NO.
Systems Applications, Inc. San Rafael, California 94903	11. CONTRACT/GRANT NO
	68-02-4306
12. SPONSORING AGENCY NAME AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
Air Management Technology Branch (MD-14) U. S. Environmental Protection Agency Research Triangle Park, N.C. 27711	14 SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

EPA Project Officer: Edwin L. Meyer, Jr.

16. ABSTRACT

The U. S. Environmental Protection Agency is proposing new short- and long-term National Ambient Air Quality Standards (NAAQS) for particulate matter having an aerodynamic diameter of less than 10 micrometers (PM_{10}). The current NAAQS for particulate matter refers to total suspended particulate matter (TSP) concentrations without a size specification. Until PM_{10} data are more widely available, TSP monitoring data must be used to estimate the likelihood of attainment of the PM_{10} NAAQS. Pace and Frank (1984) have developed a method for estimating that likelihood. Their approach relies on the distributions of particulate matter ratios.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
NAAQS PM_{10} Attainment Particulate Particulate Matter Ratios		

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS (This Report)	21. NO. OF PAGES
		129
	20. SECURITY CLASS (This page)	22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE