Final Report

A METHOD FOR CALCULATING DISPERSION MODELING UNCERTAINTY APPLIED TO THE REGULATION OF AN EMISSION SOURCE

SYSAPP-85/007

8 January 1985

Prepared for

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Source Receptor Analysis Branch
Research Triangle Park, NC 27711

Contract 68-02-3870

Prepared by

Anthony D. Thrall
Till E. Stoeckenius
C. Shepherd Burton

Systems Applications, Inc.
101 Lucas Valley Road
San Rafael, CA 94903

ACKNOWLEDGMENTS

We wish to acknowledge the valuable contributions of Mr. Joseph Tikvart and Mr. Bill Cox of the Source Receptor Analysis Branch, Monitoring and Data Analysis Division, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency who were very helpful in formulating the hypothetical example and the contents of our report. We also wish to thank Marianne Dudik, Mithra Moezzi, Paulus Irpan, and Belle Hudischewskyj of Systems Applications for their help in creating the figures.

DISCLAIMER

Although this report has been funded by the U.S. Environmental Protection Agency through Contract No. 68-02-3870, it has not been subject to the agency's peer and administrative review. Therefore, it does not necessarily reflect the views of the agency, and no official endorsement should be inferred.

ABSTRACT

A method for quantifying the uncertainty in dispersion model predictions is used to address three issues concerning model-based demonstrations of the attainment of National Ambient Air Quality Standards (NAAQS): (1) operational model performance, (2) the probability of NAAQS attainment, and (3) setting emission limits for an emission source. To illustrate the method, the use of a dispersion model to demonstrate attainment of the 24-hour SO_2 NAAQS near a 1300 MW coal-fired power plant is considered.

CONTENTS

ILLUSTRATIONS

TABLES

1 INTRODUCTION

Under contract to the U.S. Environmental Protection Agency, Systems Applications, Inc. has been developing methods for quantifying uncertainty in dispersion model results. During the past year, a method for the Calculation of Uncertainty Estimates (CUE) was developed. CUE utilizes comparisons of monitored (i.e., observed) concentrations with corresponding model predictions, together with a new statistical technique known as the bootstrap method (Efron, 1982), to calculate uncertainty in model predictions.

In particular, CUE is used to estimate the degree of uncertainty in model-predicted design-value concentrations.* In applications demonstrating attainment of National Ambient Air Quality Standards (NAAQS), ground-level concentrations are calculated at selected points (called receptors) in the vicinity of emitting sources. The design-value concentration is the maximum over all receptors of the second-highest concentrations occurring during a year. This quantity is often referred to as the "highest second high" concentration. When the highest second high concentration does not exceed the level prescribed by the NAAQS, the area is considered to be in attainment of the standard. When design values are available for more than one year, the largest of the available values is compared to the NAAQS.†

---

* The term "design-value concentrations" refers to concentrations used in the design of control strategies or derivation of emission limits.

† A different procedure is used for establishing attainment of the ozone standard, or when background concentrations are significant.

In this report we illustrate the application of the CUE method to a typical NAAQS attainment demonstration for a conventional steam power plant. The effects of uncertainty in model predictions on the attainment/nonattainment decision is discussed quantitatively. Techniques are presented for incorporating model uncertainty into model-based attainment decision making. In the course of illustrating the CUE method we explore three issues frequently raised by parties interested in the use of dispersion models to demonstrate NAAQS attainment:

(1) Operational model performance: What is the likelihood that the design value predicted by the model is higher or lower than the unknown monitoring design value, i.e., the design value we would otherwise obtain from measurements?

(2) Probability of NAAQS attainment: What is the likelihood that the unknown monitoring design value is higher or lower than the NAAQS?

(3) Setting emission limits: What level of emissions should be permitted from a source to manage the risk of violating the NAAQS while avoiding overly stringent emission limitations?*

* Here we only provide a procedure for calculating an emission limit for a prescribed acceptable risk of nonattainment. We do not discuss or recommend what such an acceptable risk might be.

2 CALCULATION OF UNCERTAINTY ESTIMATES (CUE)

The term "modeling uncertainty" can be defined in many ways. We define modeling uncertainty as the discrepancy between design-value concentrations predicted by a dispersion model and the corresponding design-value concentrations obtained from ground-level ambient monitoring instruments.* Under this definition, the estimation of modeling uncertainty is equivalent to the determination of the distribution or likely range for the unknown monitoring design value, i.e., the design-value concentration that would have been observed had monitoring data been available. Our goal then is to calculate this distribution.

An overview of the CUE method is presented in Figure 1 (see Appendix A for details). Meteorological and emissions data are fed to a dispersion model, which may be of any type, including a Gaussian model such as EPA's CRSTER model (EPA, 1977). Model predictions are usually available over a dense network of receptors (180 in the CRSTER or MPTER models) for one or more years. In contrast, the required monitoring data is usually collected at only a handful of locations for a year or two at most. These monitoring data, together with corresponding model predictions at the same few locations, form the model evaluation data base required by CUE. CUE uses the results of the model evaluation to calculate a range of adjustments to the design values predicted by the dispersion model. The resulting adjusted design values are interpreted as representing the likely range for the unknown monitoring design values corresponding to the

---

* This definition is intended to encompass all sources of uncertainty associated with the practice of regulatory air quality modeling including measurement errors and model formulation.

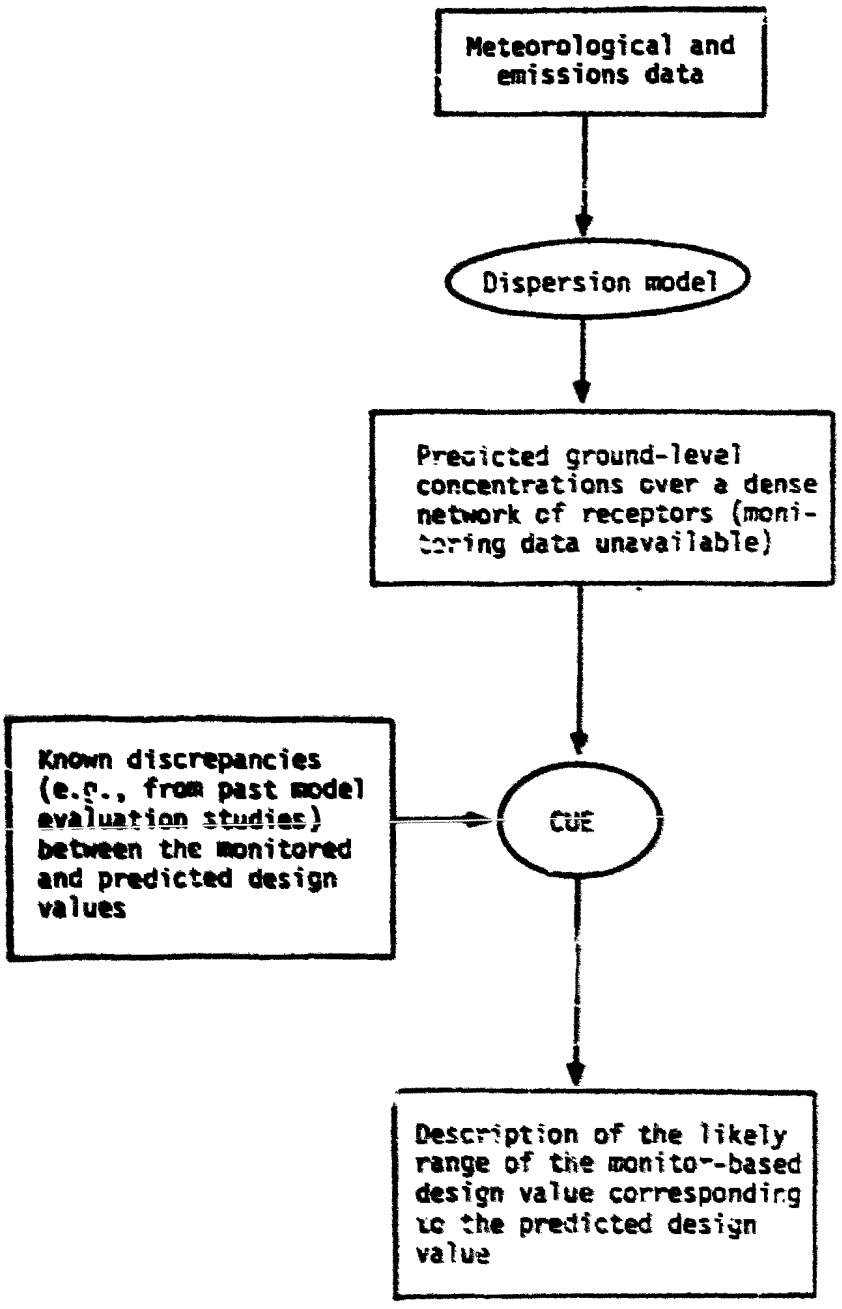


FIGURE 1. Illustration of the Calculation of Uncertainty Estimates (CUE).

design values predicted by the dispersion model. An example application of the CUE method is described in Section 3. We use this example to address the three issues raised in the introduction.

3 AN EXAMPLE

MODEL EVALUATION RESULTS

To illustrate the application of CUE we consider the use of a dispersion model to demonstrate attainment of the 24-hour SO_2 NAAQS at the Clifty Creek power plant near Madison, Indiana (see Stoeckenius et al., 1983, 1984). The Clifty Creek plant operates at an annual average emission rate of about 8400 g/s (see Appendix A for further details), but for the purposes of this report we assume a much higher hypothetical rate (15,625 g/s) to provide an example of a power plant that produces peak 24-hour average impact levels close to the SO_2 primary NAAQS (365 ug/m^3).

Meteorological data collected at Cincinnati (surface) and Dayton (upper air) for 1973 through 1977 were used in exercizing the MPTER dispersion model (Pierce and Turner, 1980). A polar grid of 180 receptors located at 10° intervals along five concentric rings centered on the source at radial distances of 1, 3, 6, 12, and 20 km was specified. The predicted 24-hour average ground-level concentrations were used to obtain the highest second-high concentration (design value) in each year. A comparison of these values with the 24-hour SO_2 NAAQS shows that the plant with this emission rate would be in violation of the standard in one of these five years (1976). Based on the design value for each individual year, or the five-year maximum design value, an emission limit could be calculated by current procedures (see Table 1).

Since the model predictions are not perfectly accurate, i.e., they contain uncertainties, a decision maker might wonder whether it is really necessary to designate the area as nonattainment and revise the plant's emission limit to eliminate possible violations of the ambient standard.

TABLE 1. Emission limits set by current practice. (See note)

Year	Given Emission Rate (g/s)	Highest Second-High Concentration ($\mu g/m^3$)	24-Hour SO_2 NAAQS ($\mu g/m^3$)	Calculated Emission Limit
1973	15,625	357	365	15,975
1974	15,625	313	365	18,221
1975	15,625	337	365	16,923
1976	15,625	367	365	15,498
1977	15,625	286	365	19,941
1973-77	15,625	357*	365	15,498

Note

$$Q_{lim} = \frac{X_{NAAQS}}{DV} \cdot Q_o, \text{ where}$$

Q_{lim} = calculated emission limit

X_{NAAQS} = 24-hour SO_2 NAAQS (365 $\mu g/m^3$)

Q_o = given emission rate, i.e., the hypothetical emission rate of 15,625 g/s used in the example analysis.

DV = highest second-high SO_2 concentration ($\mu g/m^3$) predicted by the dispersion model (known as the "design value"), i.e., the maximum of the second-highest SO_2 concentrations predicted at each receptor. (For purposes of illustration, background SO_2 concentrations are assumed to be negligible.)

* For the 1973-77 period, the design value concentration is the maximum of the five annual design value concentrations.

He might consider dismissing the 1976 model prediction as an aberration and consider the plant to be in "attainment." To resolve this dilemma the decision maker may wish to have an estimate of what the observed highest second-high concentration would have been in each year if monitoring data had been available at the 180 receptor sites. This is the issue of operational model performance, i.e., the agreement between predicted and monitored design values.

Were monitoring data available at each of the 180 receptor sites and were the monitoring design value less than the NAAQS reference concentration of 365 $\mu g/m^3$ in each of the five years, then the decision maker would have a firm basis for determining the NAAQS to have been attained and emission reductions to be unnecessary. In the absence of this dense network of monitors, CUE provides estimates of the monitoring design value. We can also use CUE to estimate the probability that the unknown design value would have exceeded the NAAQS reference concentration of 365 $\mu g/m^3$ in each of the five years.

In applying CUE to this situation, we use a model evaluation study carried out at Clifty Creek during 1975 (Mills et al., 1980; cf. Londergan et al., 1983). In this study, monitoring data and MPTER model results at six locations near the plant were used to obtain monitored and predicted highest second-high concentrations for the year. These values are much lower than those calculated from the MPTER results for the 180-receptor network (Table 1) because they were obtained under the plant's actual average 1975 emission rate of 8420 g/s. The actual average emission rate is used so that a useful comparison with monitored concentrations can be made.* Since ground-level concentrations can be reliably assumed to be linearly proportional to emission rates, model performance as measured by the ratio of monitored-to-predicted concentrations is unaffected. Thus, model evaluation results for the 24-hour average concentrations obtained

* In this example we are intentionally excluding treatment of the stochastic (variable) nature of the sulfur content of coal. This exclusion is a temporary simplification and is not a limitation of CUE.

in 1975 under the lower, actual average emission rate can be reliably applied to model predictions obtained for the larger emission rate since we are considering only the ratio of monitored-to-predicted 24-hour average concentrations.*

Taking the ratio of the monitored highest second-high concentration and corresponding predicted highest second-high concentration, we see from Figure 2 that the model prediction was 28 percent lower than the monitored design value. However, if the model evaluation experiment had been repeated for a different year, we may well have obtained a different ratio since meteorological conditions would have changed and since monitoring equipment is not perfectly precise. What would be most useful are data for many additional years, from which we could form a distribution of ratios that would represent the range of likely values. Although additional data are not available, CUE can provide us with such a distribution using the bootstrap technique. In the following paragraphs, we briefly describe the calculations carried out by the CUE method. (See Appendix A for additional details.)

CUE's USE OF MODEL EVALUATION RESULTS

The CUE method uses information obtained from a model evaluation study to produce a range of adjustments to the predicted design value, thereby producing a distribution for the unknown monitoring design value, i.e., the highest second-high concentration that might be monitored over a field of receptor sites. The steps involved in the CUE method can be summarized as follows:

---

* Model performance, as characterized by these ratios, would be improved were the model to use the actual, varying rates of emission rather than a constant, annual average rate.

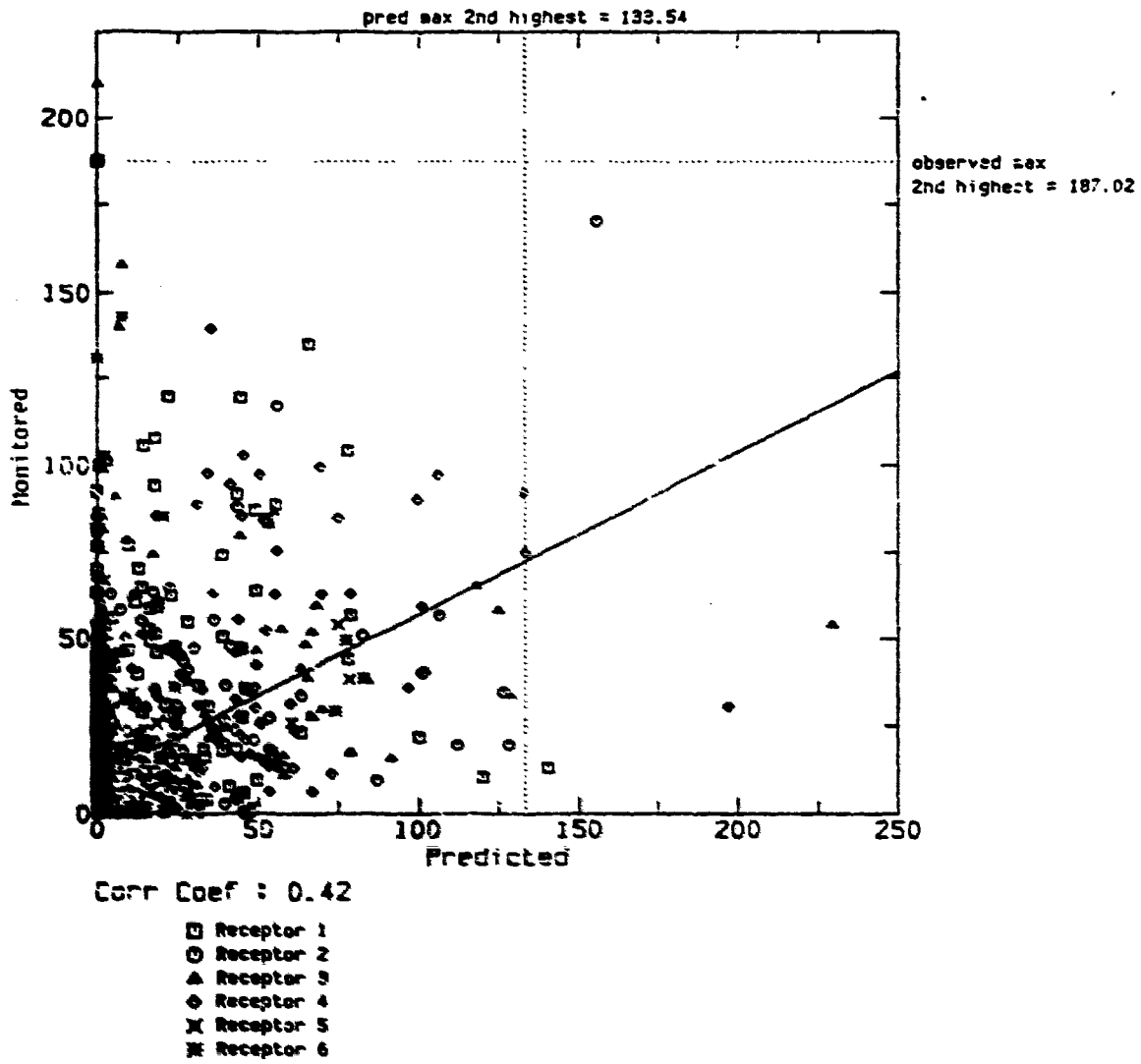


FIGURE 2. Clifty Creek 24-hour average model evaluation results.

(1) Obtain model evaluation results at actual monitor locations (six in our example) consisting of pairs of monitored and predicted concentrations for each time interval (24-hour period in our example) at each location.

(2) Using the bootstrap technique, simulate 1000 sample years of model evaluation results. For each sample year determine the six-station monitored and predicted design values, as well as their ratio, R. Using the distribution of R, determine the probability with which the monitored design value exceeds the predicted design value (i.e., the probability with which R exceeds unity).

(3) Determine the design value predicted by the model for a large array of receptors (the number of receptors is typically 180, as in our example) based on a year of meteorological data.

(4) Multiply this 180-receptor predicted design value by the 1000 ratios, R, to produce a distribution for the 180-receptor monitoring design value. Using this distribution, find the probability that the monitoring* design value does not exceed the ambient standard.

(5) Perform steps 3 and 4 for a total of five years of meteorological data to produce an uncertainty distribution for the 180-receptor monitoring design value in each year. From these five uncertainty distributions, calculate the distribution of the five-year maximum of the monitoring design values.

---

* Throughout this report "monitored" design value refers to that obtained for the six stations, whereas "monitoring" design value refers to the value that would be obtained from monitoring at 180 receptors.

CHARACTERIZING MODEL PERFORMANCE

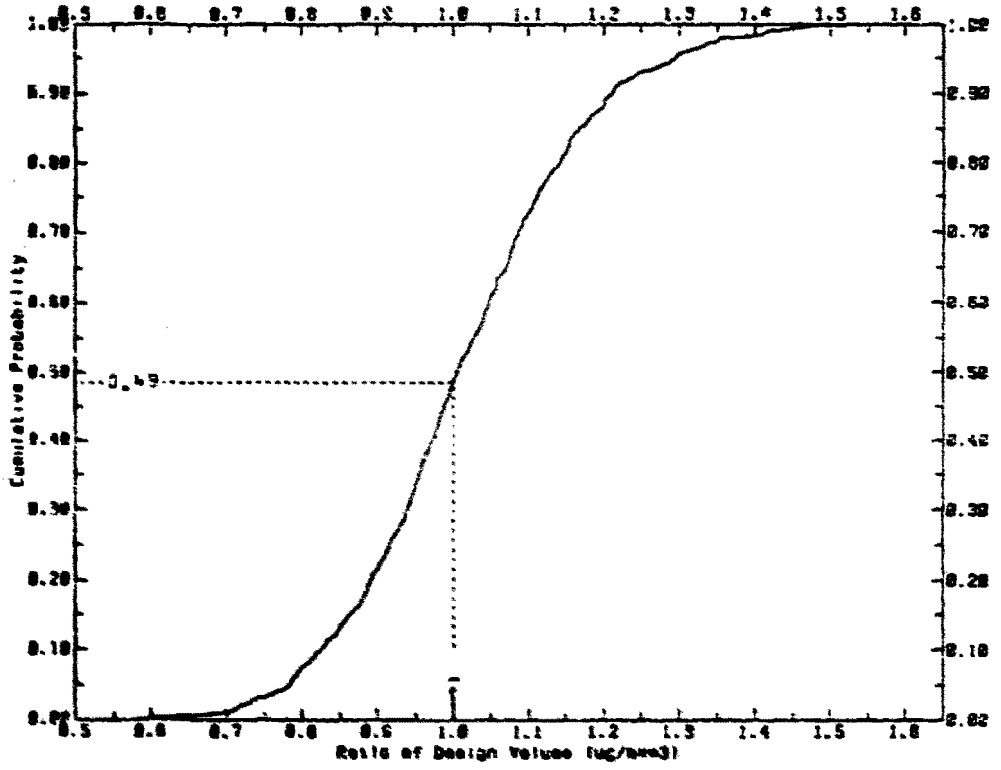
The 1000 ratios (R) of monitored-to-predicted design values generated by CUE from the model evaluation results characterize the operational performance of the model (the first issue raised in the introduction). That is, good operational performance by the model would be characterized by R's tightly distributed about unity. From Figure 3 we see that the model is unbiased in our example, since the 50th percentile of R is close to unity. The distribution is also fairly tight since R is between 0.8 and 1.2 (i.e., the predicted design value is within 20 percent of the monitored design value) about 80 percent of the time.

ESTIMATING THE PROBABILITY OF NAAQS ATTAINMENT ASSOCIATED WITH A GIVEN EMISSION RATE

Upon multiplying the 180-receptor predicted design value by the 1000 values of R, we obtain the likely range for the unknown monitoring design value (see Appendix A for details). These probability (or uncertainty) distributions are plotted in cumulative form for each year in Figure 4. Figure 4 also shows the probability distribution of the five-year maximum of the monitoring design values.

For each year the figure indicates the predicted design value, the NAAQS concentration of 365 $\mu g/m^3$, and the probability that the monitoring design value does not exceed the NAAQS concentration. Thus, for an average emission rate of 15,625 g/s in each year, there is an estimated 55 percent chance that the 24-hour NAAQS would have been attained in 1973, whereas there is a 94 percent chance that it would have been attained in 1977. The chance that the standard would have been attained in all five years during the 1973-77 period is estimated to be 15 percent.

An alternative, more compact, way of displaying the information contained in Figure 4 is to use a box plot (Figure 5). Box plots allow sketches of



Cliffy Creek 1675 (24-hr avg)

FIGURE 3. Cumulative distribution of R, the ratio of monitored to predicted design values.

84144

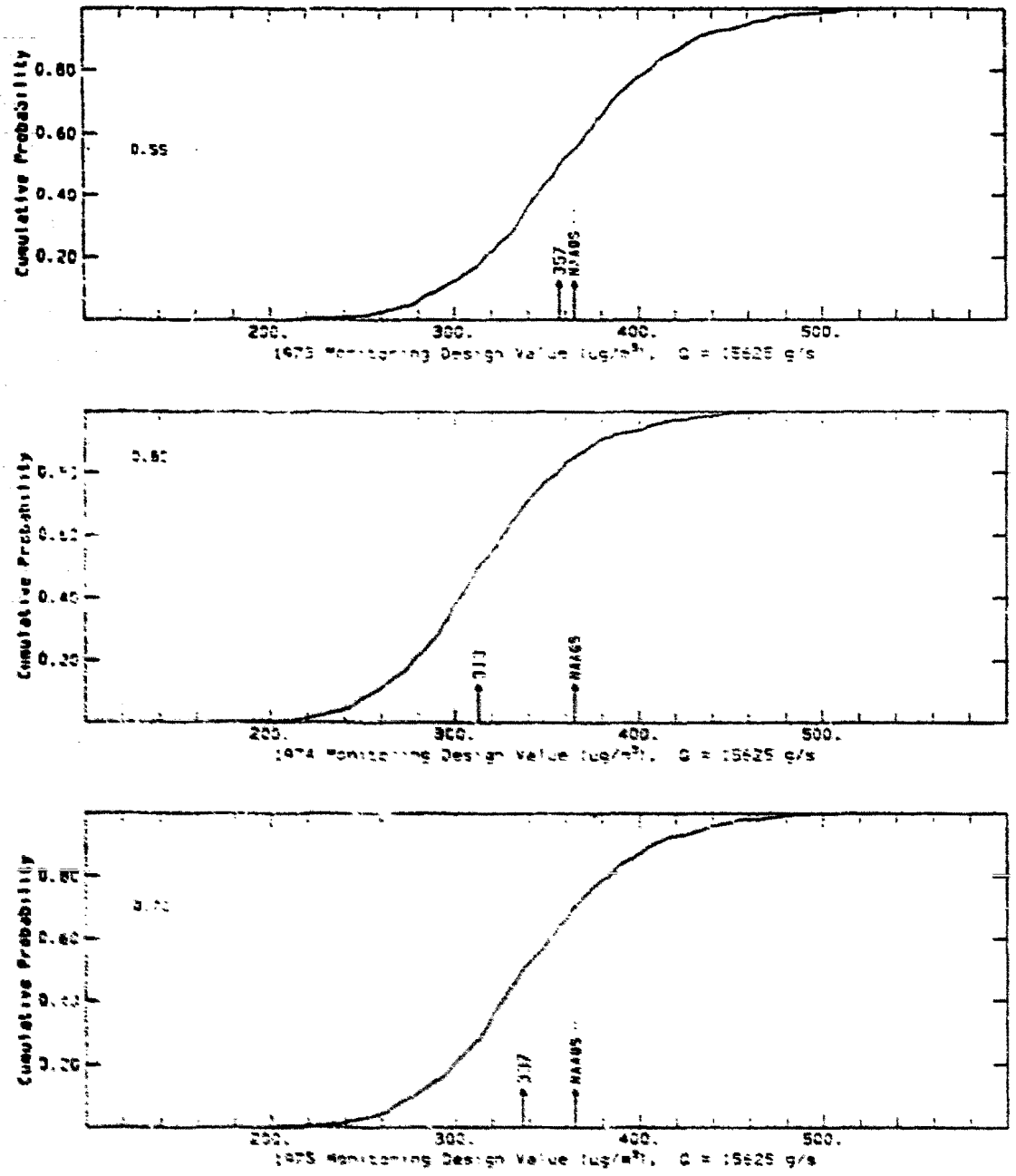
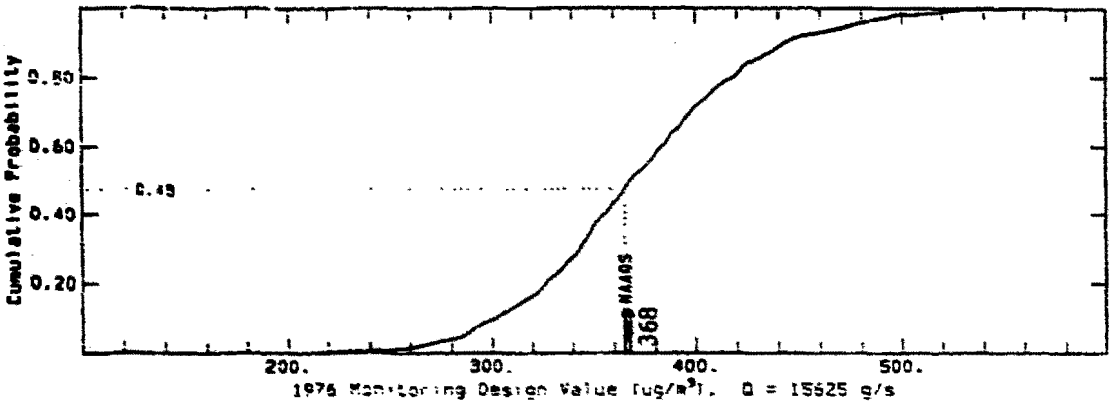
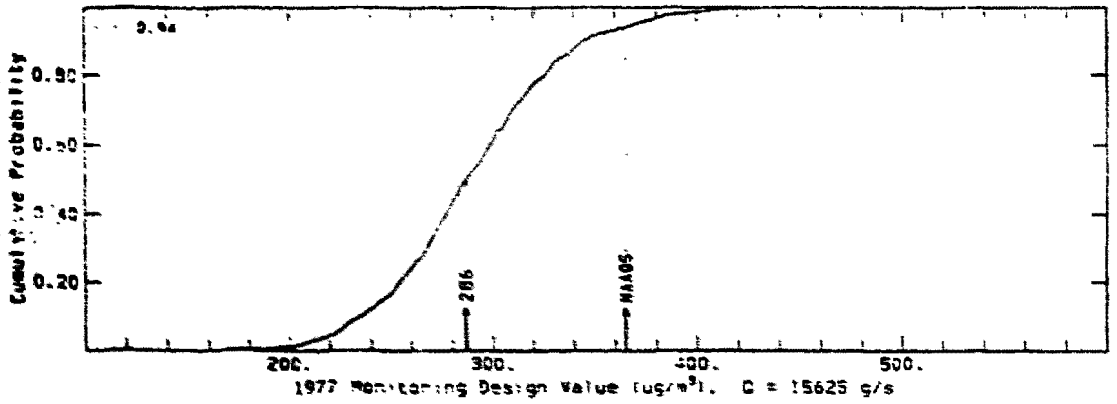


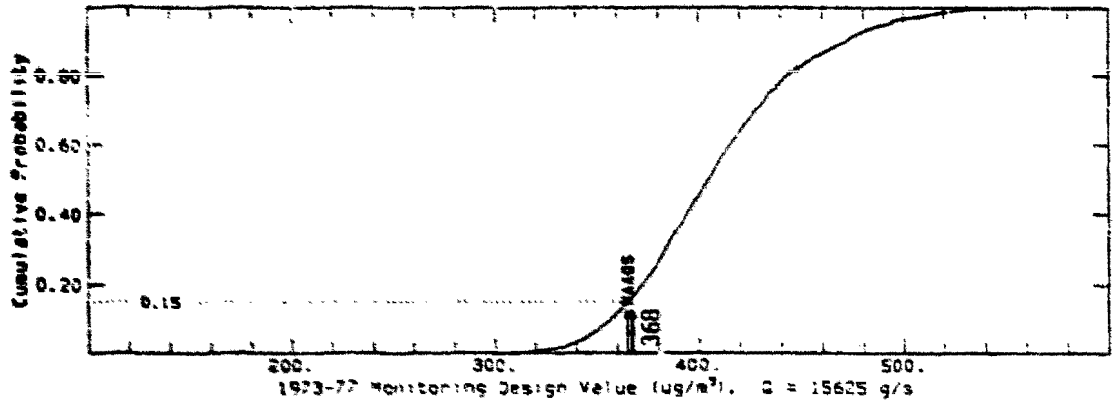
Figure 4. Cumulative probability distributions for 1973 - 1977 monitoring design values, assuming a 15,625 g/s annual average rate of emission. (Both the model-predicted concentration and the NAAQS concentration of 365 ug/m³ are indicated. The probability of NAAQS attainment is also shown.)



1976 Monitoring Design Value [ug/m³]. Q = 15625 g/s



1977 Monitoring Design Value [ug/m³]. Q = 15625 g/s



1973-77 Monitoring Design Value [ug/m³]. Q = 15625 g/s

Fig 4. concluded.

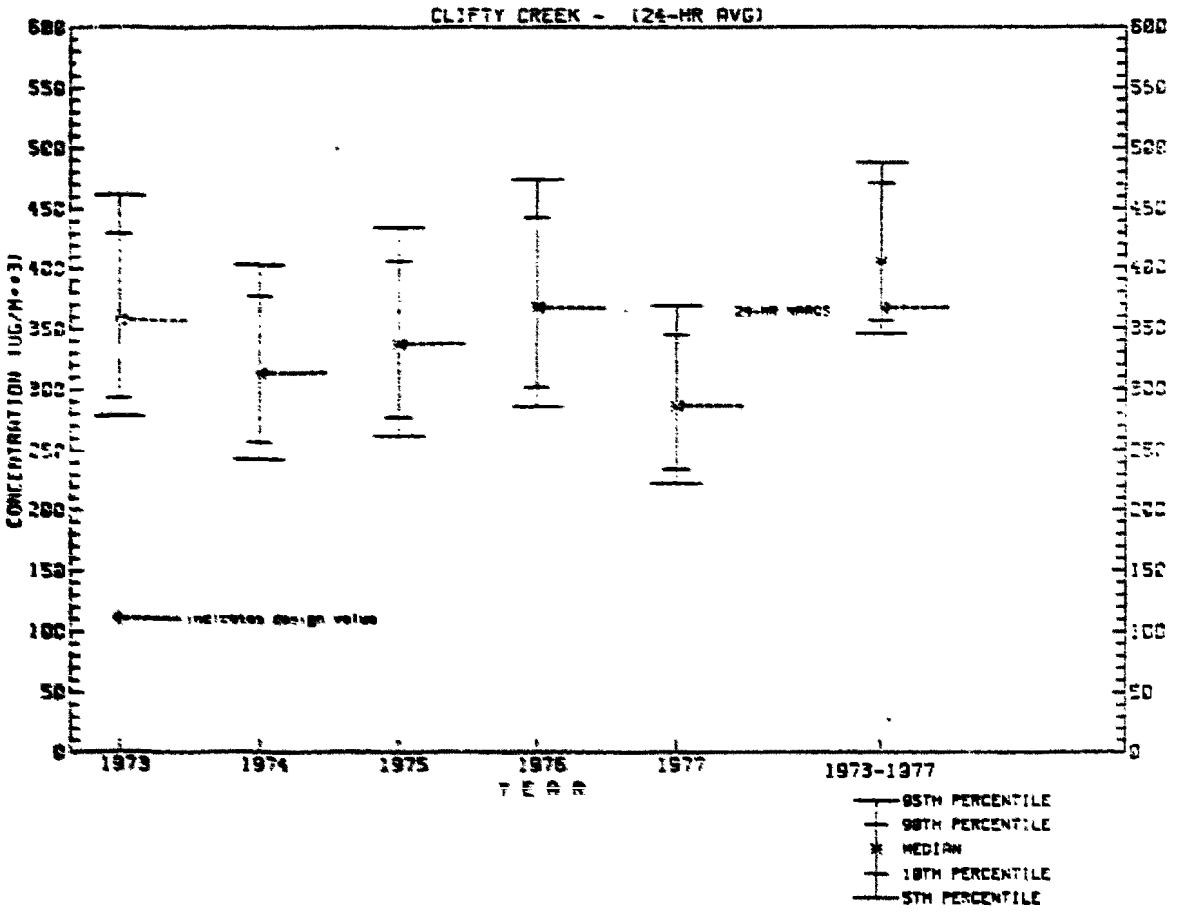


FIGURE 5. Box plots depicting probability distributions for monitoring design values, assuming a 15,625 g/s annual average rate of emission. The distribution for each year from 1973 to 1977 is depicted, along with the distribution for the maximum of the five annual design values.

design value distributions from more than one year to be displayed conveniently on a single page. As in the cumulative distribution plots of Figure 4, the arrows in Figure 5 indicate the design values predicted in each year.

The 1973-77 design value is represented in Figures 4 and 5 because in many cases decision makers are concerned with the largest design value taken over a five-year period. As noted earlier (Table 1), the largest predicted design value just exceeds the NAAQS concentration. Decision makers might wish to know the probability with which the largest of the five monitoring design values would have been greater than (or less than) the NAAQS concentration. Figure 4 shows that there is a 15 percent chance that the largest design value taken over the five years would be less than the NAAQS concentration. This probability is much lower than the probability in individual years because we are focusing attention on the maximum of five design values drawn from the first five distributions in Figure 4, any one of which might be larger than the standard.

Solely on the basis of the original, conventionally interpreted, model results, a decision maker might conclude that the Clifty Creek plant, operating at an average emission rate of 15,625 g/s, was in violation of the standard in one of the five years, namely, 1976. The decision maker might then designate the area "nonattainment" and in so doing seek to have the emission limit of the plant reduced to bring the area into "attainment." However, Figure 4 shows that there was a 48 percent chance of attaining the NAAQS in 1976, i.e., a 44 percent chance that a decision to designate the area "nonattainment" would be incorrect. When all five years of model results are considered, however, the decision maker would find that there is only a 15 percent probability that the NAAQS was attained in all five years. From this broader perspective, the decision to designate the area "nonattainment" has an 85 percent chance of being correct.

Alternatively, the probability of attainment in one year or during a five-year period can be interpreted as the probability that no reduction in the given emission rate would be required on the basis of monitoring across a network of 180 receptors.

ESTIMATING THE PROBABILITY OF ATTAINMENT ASSOCIATED WITH CURRENT PRACTICE EMISSION LIMITS

The above emission rate of 15,625 g/s was chosen to illustrate the use of CUE when attainment status is difficult to determine reliably by conventional methods. The CUE method quantifies the risk of making an incorrect decision. In a similar vein, we can use CUE to obtain the probability of NAAQS attainment for an average emission rate equal to the emission limit as currently derived from the dispersion model (see Table 1).

For each individual year from 1973 through 1977 and for the five-year period 1973-77, Figure 6 shows the probability distribution of the monitoring design value for an average emission rate equal to the emission limit prescribed by the dispersion model. Recall that the emission limit is the emission rate at which the model predicts a design value concentration equal to the NAAQS concentration of 365 ug/m^3. The corresponding monitoring design value concentration, i.e., that likely to result from an emission rate equal to the emission limit, can therefore be approximated by multiplying a monitored-to-predicted ratio of design values by the NAAQS concentration.

Performing this multiplication for the 1000 ratios represented in Figure 3, we obtain the probability distribution for individual years shown in Figure 6. Note that the emission limit varies from year to year but the probability distribution remains the same, with the NAAQS concentration occurring at the 49th percentile corresponding to the percentile at which unity occurs in the distribution of ratios.

Figure 6 also shows the distribution of the 1973-77 maximum monitoring design value for an emission rate equal to the minimum of the five annual emission limits. Use of the most restrictive emission limit conforms to current regulatory practice (see Table 1).

Table 2 summarizes the attainment probabilities associated with the hypothetical emission rate of 15,625 g/s and with the emission limits set by

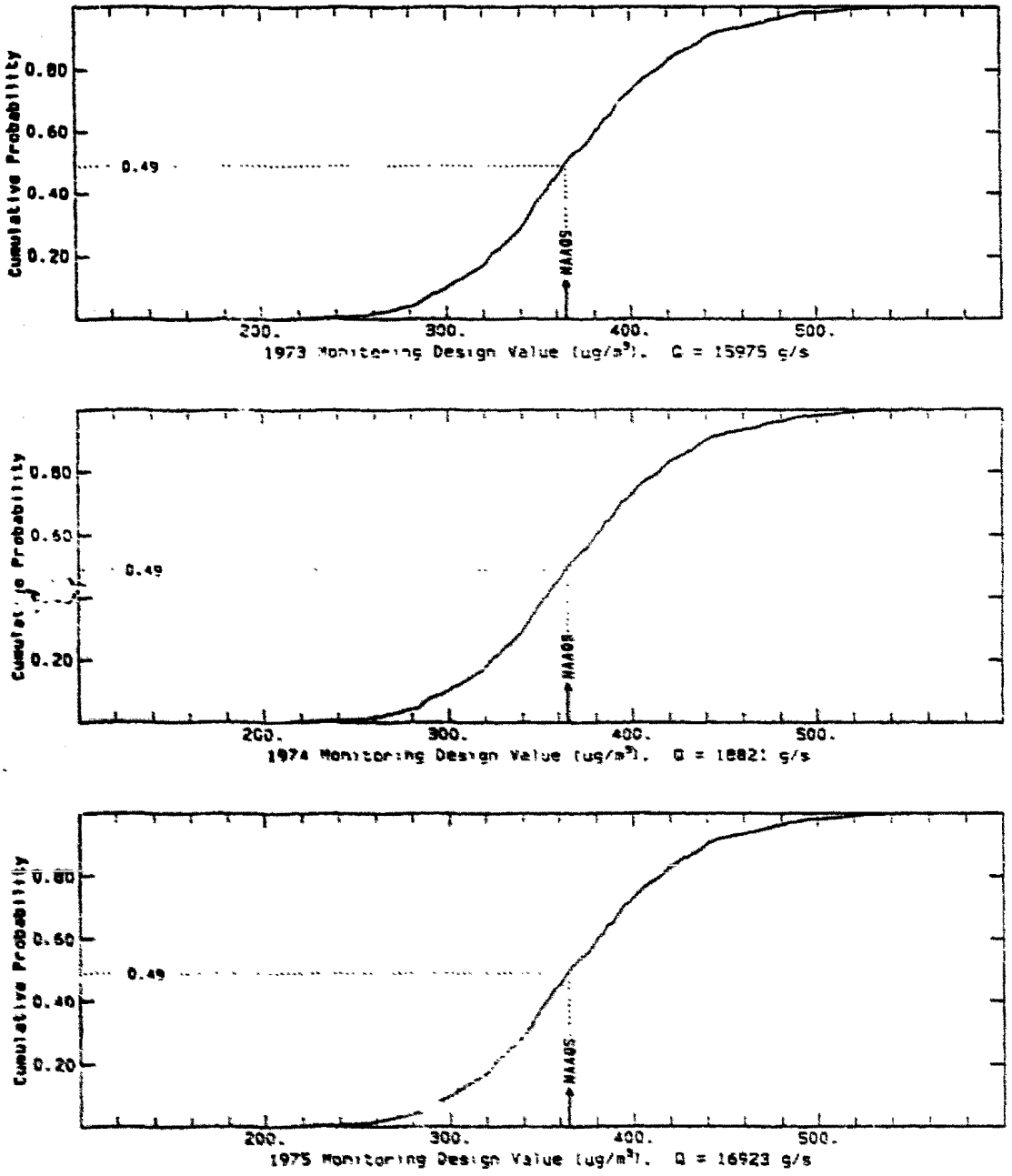
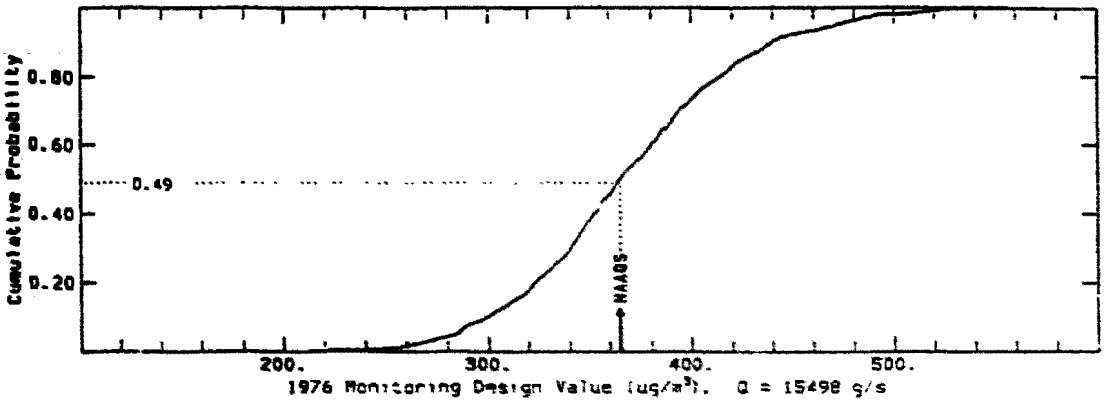
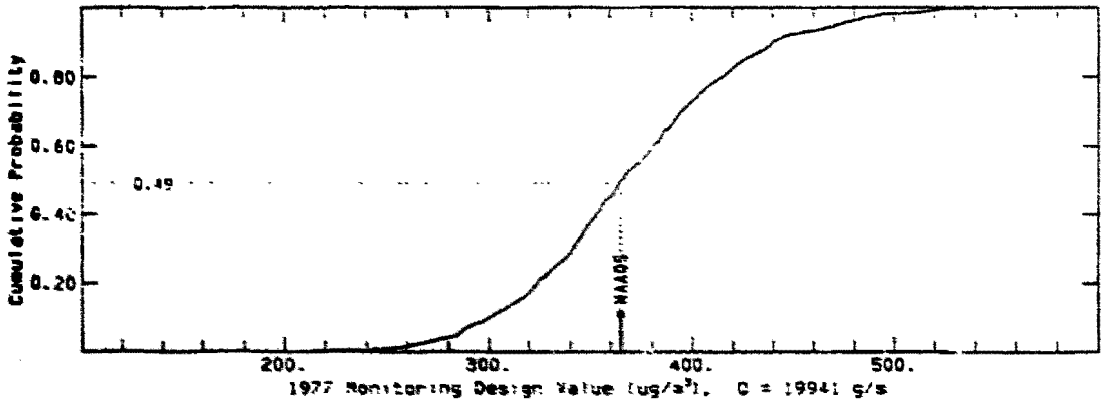


Figure 6. Cumulative probability distributions for 1973 - 1977 monitoring design values, assuming the average rate of emission in each year is equal to the emission limit set by current practice. (The model-predicted concentration coincides with the NAAQS concentration of 365 ug/m³. The resulting probability of NAAQS attainment is 0.49 for each year.)





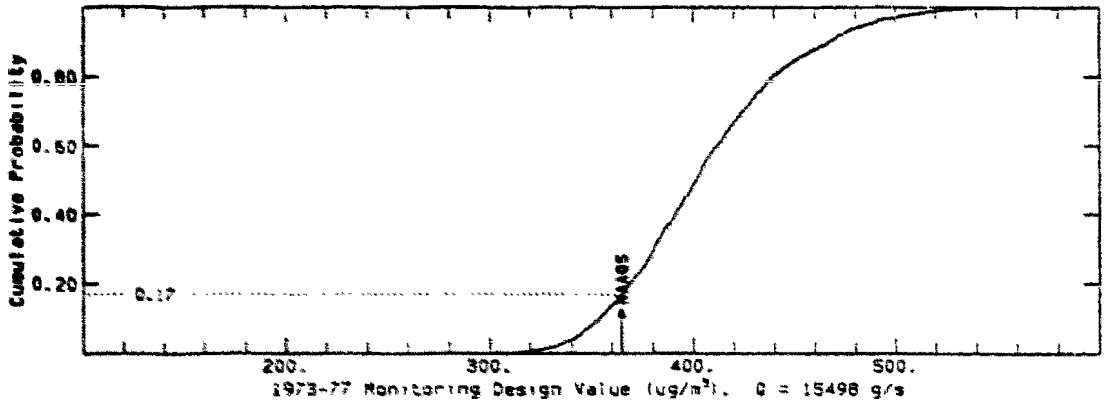


Fig 6. concluded.

TABLE 2. Probabilities of NAAQS attainment associated with the given emission rate and with emission limits set by current practice.

	Given Emission Rate		Calculated Emission Limit	
Year	Emission Rate (g/s)	Associated Probability (%)	Emission Rate (g/s)	Associated Probability (%)
1973	15,625	55	15,975	49
1974	15,625	85	18,221	49
1975	15,625	70	16,923	49
1976	15,625	48	15,498	49
1977	15,625	94	19,941	49
1973-77	15,625	15*	15,498	17*

* This is the probability that the 1973-77 maximum monitoring design value concentration resulting from the indicated 1973-77 emission rate would not exceed the NAAQS concentration of 365 $\mu g/m^3$.

current practice. As a consequence of the operational unbiasedness of the model, i.e., the fact that the monitoring design value falls above or below the predicted design value with nearly equal probability, the attainment probability associated with the current practice emission limit is 49 percent for each year. Using the five-year minimum emission limit results in a 17 percent probability that the standard would be attained in each of the five years, i.e., an 83 percent probability that the standard would not be attained in at least one of the five years.

It is tempting to conclude from the 83 percent probability of non-attainment that the emission limits set by current practice are too lenient. Such a conclusion is not warranted at this stage, however, for several reasons. First, the magnitude of the probability is overstated because it was derived from ratios in Figure 3 that characterize a model that assumes a constant rate of emission over the one- or five-year period. This assumption exaggerates the uncertainty or imprecision of model predictions. Second, as we discuss in subsequent sections, the adequacy of the emission limits derived from a model by current practice should be judged not by the corresponding probability of attainment, but by the probability that can be achieved with only a slight alteration of the emission limit. Thus, an unbiased model would produce an annual emission limit having an associated 50 percent probability of attainment, but if the model is sufficiently precise the emission limit corresponding to a 90 percent (or perhaps even a 99 percent) probability of attainment would be only slightly more stringent. Finally, additional mode. evaluation studies and meteorological records should be examined before generalizing the results presented here.

Although we have cautioned against generalizations about model perfo mance or the apparent leniency of current practice, the foregoing discussion demonstrates how CUE can be used in a given setting to assess the effects of modeling uncertainty on attainment/nonattainment decisions. We thus see how the CUE method can be used to answer the second question raised in the introduction. We now turn to the third question: what level of emissions should be permitted from a source in order to bring the risk of

nonattainment to an acceptably low level while avoiding unnecessarily stringent emission limits?

ESTIMATING THE EMISSION RATE ASSOCIATED WITH A DESIRED PROABILITY OF NAAQS ATTAINMENT

Since we can assume that concentrations are linearly proportional to emission rates, we can choose an emission rate for each year from 1973 through 1977, or for the 1973-77 period, such that the resulting probability of NAAQS attainment is some desired value. Table 3 shows emission limits corresponding to attainment probabilities of 50, 90, 95, and 99 percent.

We see that the 1973-77 emission limit based on current practice would be 15,498 g/s, slightly less (< 1 percent) than the hypothetical value of 15,625 g/s. We also see that, to achieve a 50 percent probability of attaining the standard in all five years, the emission limit would be set at 14,092 g/s, or reduced by about 9 percent from current practice. Alternatively, if a 90 percent probability were desired, the required emission rate would be 12,098 g/s, or 22 percent lower than current practice.

The emission limit determined by current practice for 1976, the year in which the maximum predicted design value occurs, is by definition the emission limit obtained for the 1973-77 period. This is not the case, however, for emission limits determined by CUE since the maximum design value may occur in any of the five years (albeit with different probabilities). If, for example, the desired probability of attainment is 50 percent, then the emission limit obtained by CUE is 15,442 g/s for 1976, whereas it is 14,092 g/s for the 1973-77 period.

The results in Table 3 are summarized in Figure 7, which illustrates emission limits corresponding to current practice and two probabilities of attainment: 50 and 90 percent.

TABLE 3. Comparison of emission limits set by current practice and those calculated by CUE to yield a desired probability of NAAQS attainment.

Year	Highest Second-High Concentration ($\mu g/m^3$)	Emission Limit (g/s)	Associated Probability of NAAQS Attainment* (%)	Emission Limit (g/s) for Desired Probability of NAAQS Attainment			
				50%	90%	95%	99%
1973	357	15,975	49	15,901	13,207	12,331	11,294
1974	313	18,221	49	18,150	15,075	14,075	12,891
1975	337	16,923	49	16,838	13,985	13,057	11,959
1976	368	15,498	49	15,442	12,825	11,974	10,967
1977	286	19,941	49	19,824	16,466	15,373	14,080
1973-77**	368	15,498	17	14,092	12,098	11,084	10,886

* Computed using CUE as in Table 2.

** This row refers to the probability that the 1973-77 maximum design value concentration resulting from the indicated 1973-77 emission rate would not exceed the NAAQS concentration of 365 $\mu g/m^3$.

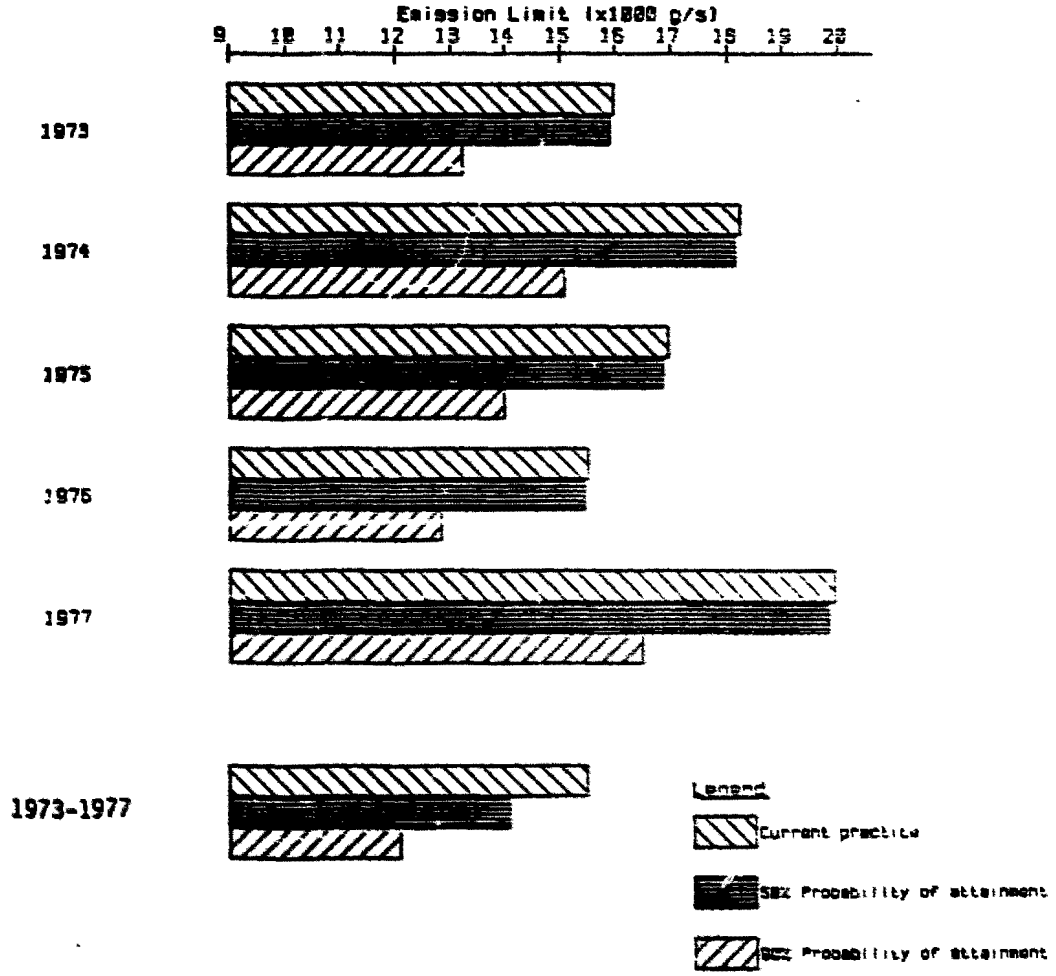


FIGURE 7. Comparison of emission limits obtained via current practice with emission limits required to achieve 50 and 90 percent probabilities of NAAQS attainment.

All the emission limits reported here conform to current practice in one respect. They are based on model evaluation results and attainment demonstrations in which the dispersion model uses a constant rate of emission (i.e., they do not allow for the vari ility of actual rates of emission). Had the model used fluctuating rather than constant, peak emissions then the current-practice emission limits in Table 3 would be somewhat higher (or less stringent). The magnitude of this change depends on the magnitude of the fluctuation in emission rates relative to the average emission rate, 10-30 percent being typical for 24-hour average rates of emission from large coal-fired boilers.

Additionally, the model's use of actual fluctuating emissions would most likely improve the performance of the model, as characterized by design value ratios (Figure 3). Moreover, if the CUE procedure were modified to include the randomly varying fuel-sulfur content (i.e., incorporating an ExEx-like methodology into CUE), then the resulting emission limits from the modified CUE procedure would also be altered. In this case, however, it is not possible to estimate the change in stringency because the shift is strongly influenced by policy choices about statistical interpretations of the NAAQS and the setting and enforcement of emission limits derived form such interpretations. Suffice it to say here that there is enough latitude in these policy options to allow for any such shift.

Thus, by recognizing the variability of emission rates, as recorded in the model evaluation study and as projected due to natural, random variations in coal-sulfur content, the estimated risk of nonattainment may be substantially altered without substantial relaxation of current practice emission limits.

THE EFFECTS OF MODEL BIAS AND IMPRECISION

The foregoing results represent a hypothetical situation in which there is little direct evidence of model bias (i.e., the tendency of the model to overestimate or underestimate monitored design values). In this example, therefore, the principal source of the uncertainty in providing design value estimates is model imprecision. In Appendix B we examine the effects of model bias and imprecision on NAAQS attainment probabilities and emission limits. The results can be summarized by discussing emission limits set by CUE under several cases of model bias and imprecision.

We illustrate the separate and combined effects of different levels of bias and imprecision by making hypothetical but reasonable adjustments to the ratios characterizing operational model performance given in Figure 3. We consider three levels of bias (one of which is no bias, cf. Figure 3) and three levels of precision (one of which is unaltered from the results in Figure 3), resulting in nine cases, eight of which are new.

The levels of bias chosen were ±30 percent. Bias was introduced by shifting the overall distribution of ratios of monitored to predicted design values (cf. Figure 3) by a factor of 1.3 (underprediction) and by a factor of 1/1.3 (overprediction), thus altering the geometric mean ratio by these factors. The operational precision of the model was altered by, respectively, squaring or taking the square root of the geometric standard deviation of the ratios. In a certain sense, stated in Appendix B, the operational imprecision of the model was doubled or halved, respectively, by these alterations.

The effects of changes in overall model accuracy on emission limits obtained from CUE are summarized in Table 4 for the year 1973. The entries in Table 4 represent allowable emissions for a 90 percent probability of NAAQS attainment. The entry in the center of the table, for the case of unaltered precision and no bias, is identical to the allowable 1973 emissions given in Table 3. Considering first changes in precision only (no bias), we see that when model precision is enhanced, the emission

TABLE 4. 1973 emission limits set by CUE and designed to achieve a 90 percent probability of NAAQS attainment for different scenarios of model bias and imprecision.

	Underprediction ($\beta < 0$)	Unbiasedness ($\beta = 0$)	Overprediction ($\beta > 0$)
Enhanced Precision ($\phi = 1/2$)*	14,528	14,528	14,528
Unaltered Precision ($\phi = 1$)*	13,207	13,207	13,207
Diminished Precision ($\phi = 2$)*	10,915	10,915	10,915

* See Appendix B.

limit increases approximately 10 percent, from 13,207 to 14,528 g/s; higher emissions can be permitted because of the greater reliability of the dispersion model. Similarly, when precision is diminished, the emission limit decreases 17 percent, from 13,207 to 10,915 g/s. The lower allowable emissions result from the need to compensate for greater uncertainty.

Turning to the results for cases of model bias, we see that the CUE method provides values for emission limits that are identical to those obtained for the unbiased results; that is, in the 30 percent underprediction case CUE provides the same results (10,915, 13,207, and 14,528 g/s) as obtained for an unbiased model. This, of course, is as it should be. These results (cases of model bias) may be compared with what would be obtained from current practice for, say, the case of 30 percent underprediction. In this situation current practice would give an emission limit 30 percent higher (viz., 20,768 = 1.3 x 15,975, cf. Table 3), which CUE would adjust downward to account for the bias.

The effect of model imprecision on emission limits will be particularly pronounced for emission limits based on the five-year maximum design value (see Table 3). Here we again note that gains may be made both in model precision and in the realism of the regulatory use of models by incorporating the actual, varying rates of emission in model evaluation results and then incorporating these varying rates in the linkage provided by CUE between an annual average emission rate and the corresponding probability of NAAQS attainment.

Thus, by incorporating variable rates of emission, CUE will provide more realistic emission limits while maintaining an objective means of incorporating modeling uncertainty into regulatory decisions.

4 SUMMARY AND CONCLUSIONS

SUMMARY OF RESULTS

This report has described a hypothetical regulatory application in which a dispersion model was used to estimate the air quality impact of an electric steam power plant. Model results, stated in terms of design values, showed the plant to be in violation of the 24-hour NAAQS for SO_2 in one of the five modeled years, while predicted design values in the other four years were less than the standard. Given only this information, a decision maker might conclude that the plant is in violation of the NAAQS and that emissions reductions are needed (Table 1). However, model results are known to contain uncertainties and therefore a decision maker might want to know what the chance is that the plant would cause a violation. By this we mean that a decision maker might wish to have some information regarding the likely monitoring design values, i.e., the values that would have been obtained from monitoring at each of the 180 model receptors for each of the five years.

The CUE method uses information obtained from a given model evaluation study to "correct" the predicted design values and thereby to estimate the highest second-high concentration that might be measured over a field of receptor sites. The steps involved in the CUE method are summarized in Section 3 and are further discussed in Appendix A.

In our hypothetical example, the probability of NAAQS attainment in each year from 1973 to 1977 ranged from 48 to 94 percent. However, the probability of attaining the standard in all five years was only 15 percent (Table 2). Thus, a regulatory policy that focuses on the maximum design value taken over a five-year period is much more likely to result

in a finding that the plant is in violation than one that looks only at the design value for an individual year. It should be noted that this increased stringency results primarily because the plant (operating at its hypothetical emission rate) is producing peak 24-hour average concentrations very close to the NAAQS in each of the five years. Larger year-to-year differences in the design values, coupled with decreases in modeling uncertainty (greater agreement between monitored and modeled design values), would result in less dramatic differences in the effects of these two alternative policies than is the case in the present example.

Are emissions reductions required at this plant to avoid an excessively high risk of NAAQS nonattainment? If we suppose that a 99 percent probability of attainment, i.e., a 1 percent probability of violation, represents an acceptable level of risk for decision makers, then a reduced emission limit is required, since Table 2 shows attainment probabilities below 99 percent in all cases. Alternatively, if a 50 percent level of risk were acceptable, then the need for a reduced emission limit cannot be determined without considering which attainment period policy is preferred, viz., the five-year or individual year attainment period. If the former were chosen, then a reduced emission limit would be required; if the latter were chosen, then a reduced emission limit would not be required. The amount that emission limits would have to be reduced for any alternative probability-of-NAAQS-attainment policy can be calculated as shown in Table 3.

At the end of Section 3, we briefly considered the effects of model bias and imprecision on the emission limits, derived by CUE, required to achieve a desired probability of NAAQS attainment. We illustrated how CUE, unlike current practice, compensates for model bias. We also illustrated how CUE-derived emission limits would increase with improvements in model precision and decrease with degradation in model precision.

We believe the example presented here is encouraging because it represents a framework and a method for incorporating modeling uncertainty into attainment-demonstration decision making. However, we strongly caution

against generalization of the results and policy implications from this one hypothetical example. Our understanding of how results depend on source characteristics, source location, results from different model evaluation studies, explicit treatment of fuel-sulfur variability and policy options is, as yet, incomplete.

CONCLUSIONS

CUE can be used to address the three issues raised in the introduction, namely,

> Characterizing operational model performance in the form of a distribution of ratios of monitored-to-predicted design values (Figure 3), a distribution centered about unity indicating "operational unbiasedness," and the steepness of the slope in the cumulative distribution indicating "operational precision";
>
> Estimating the probability of NAAQS attainment for an annual average emission rate, either given or prescribed by the dipsersion model (see Table 2);
>
> Estimating the annual average emission rate needec to achieve a desired probability of NAAQS attainment (see Table 3).

With respect to the emission limits derived by CUE, we have seen that

> Unlike current practice, CUE compensates for model bias
>
> CUE provides higher (less stringent) emission limits as the model becomes more precise and provides lower (more stringent) emission limits for imprecise models to compensate for the greater uncertainty.

Finally, model precision is likely to be increased when we provide the model with the actual, varying rates of emission. Moreover, once model evaluation results of this kind are available, it is possible to modify CUE so that this greater model precision is incorporated in the calculation of more realistic emission limits. CUE thus offers an objective method for deriving emission limits that compensates for a model's operational bias, uses a model's precision, and incorporates alternative policies for managing the risk attendant to these uncertainties.

REFERENCES

Efron, B. 1982. The Jackknife, the Bootstrap and Other Resampling Plans. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

EPA. 1977. User's Manual for Single-Source (CRSTER) Model. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Londergan, R. J., D. H. Minott, D. J. Wackter, T. M. Kincaid, and D. M. Bonitata. 1982. "Evaluation of Rural Air Quality Simulation Models." U.S. Environmental Protection Agency, Research Triangle Park, North Carolina (EPA-450/4-83-003).

Mills, M. T., R. Caiazza, D. D. Hergert, and A. Lynn. 1980. "Evaluation of Point Source Dispersion Models." Teknekron Research, Inc., Waltham, Massachusetts.

Pierce, Thomas E., and D. B. Turner. 1980. User's Guide for MPTER: A Multiple Point Gaussian Dispersion Algorithm with Optional Terrain Adjustment. U.S. Environmental Protection Agency, Research Triangle Park, NC (EPA-600/8-80-016).

Rao, K., and L. Satterfield. 1982. MPTER-DS: The MPTER Model Including Deposition and Sedimentation. User's Guide. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Stoeckenius, T. E., D. Thrall, and C. S. Burton. 1983. "Evaluating the Effect of Modeling Uncertainty on the Determination of Emission Limits Using the "Bootstrap" Technique." Systems Applications, Inc., San Rafael, California.

Stoeckenius, T. E., D. Thrall, and C. S. Burton. 1984. "Incorporating Estimates of Modeling Uncertainty into Dispersion Model Results." Systems Applications, Inc., San Rafael, California.

Appendix A

DETAILED DESCRIPTION OF CUE

Appendix A

DETAILED DESCRIPTION OF CUE

MODEL EVALUATION RESULTS

In our example of the application of CUE we consider an attainment demonstration of a typical regulatory situation involving a large point source of SO_2 emissions. As in our previous reports (Stoeckenius et al., 1983, 1984), we use the Clifty Creek power plant near Madison, Indiana as the basis of our example. Relevant plant parameters are listed in Table A-1. Although the Clifty Creek plant operates at an annual average emission rate of about 8400 g/s, for the purposes of this report, we assume a much higher hypothetical rate (15,625 g/s) to provide an example of a power plant that produces peak 24-hour average impact levels close to the SO_2 primary NAAQS (365 $\mu g/m^3$).

Meteorological data collected at Cincinnati (surface) and Dayton (upper air) for 1973 through 1977 were used in exercising the MPTER dispersion model (Pierce and Turner, 1980). A polar grid of 180 receptors located at 10° intervals along five concentric rings, centered on the source at radial distances of 1, 3, 6, 12, and 20 km, was specified; the resulting 24-hour average ground-level concentrations were used to calculate the maximum second highest concentration ($M2H_{180}$) in each year. Table A-2 shows $M2H_{180}$ for each year along with a modified design value which we call $MDVl_{180}$ ($MDVl_{180}$ is used in simulations, which we shall describe shortly).

A comparison of the maximum second highest concentration with the 24-hour SO_2 NAAQS shows that the plant with this emission rate would be in

TABLE A-1. Clifty Creek stack parameters. (Source: Mills et al. 1980).*

Stack	Height (m)	Diameter (m)	Stack Gas Temperature (°k)	Stack Gas Exit Velocity
1	207.87	4.63	445.37	49.9
2	207.87	4.63	445.37	49.9
3	207.87	4.63	445.37	49.9

* Total plant capacity = 1300 MW; total plant emissions (hypothetical) = 15,625 g/s.

TABLE A-2. 24-hour average design values predicted by the MPTER dispersion model for a network of 180 receptors using a hypothetical emissions rate of 15,625 g/s.

Year	Maximum Second Highest Concentration $\mu g/m^3$ (ppm)*	Maximum DVI (MDVI)* ($\mu g/m^3$)
1973	357	360
1974	313	365
1975	337	368
1976	367	348
1977	286	313

* DVI is the concentration value exceeded exactly once per year as determined from a tail exponential fit to the distribution of daily concentrations.

violation of the standard in only one of these five years (1976). By current procedures, we can derive an emission limit based on each year's design value by multiplying the 15,625 g/s emission rate by the ratio of χ_{NAAQS} (the 365 $\mu g/m^3$ level of the current 24-hour SO_2 NAAQS) by the calculated maximum second highest concentration, $M2H_{180}$ (see Table A-3).

Since the model predictions are not perfectly accurate (i.e., they contain uncertainties), a decision maker might wonder whether it is really necessary to designate the area as nonattainment and begin the process of modifying the plants emission limit to ensure that the plant does not cause violations of the ambient standard, or whether the 1976 model prediction can be dismissed as an aberration. What the decision maker may wish to know is what the observed maximum second highest concentration would have been in each year if monitoring data had been available at the 180 receptor site. We can denote these values by $M2H_{180}$. If $M2H_{180}$ is less than the NAAQS in all 5 years, then the decision maker could assume that the plant is in attainment of the NAAQS and that no reduction in emission rate is required. CUE allows us to estimate what $M2H_{180}$ would have been for each year.

In applying CUE to this situation, we use a model evaluation study carried out at Clifty Creek during 1975 (Mills et al., 1980). In this study, monitoring data and MPTER model results at six locations near the plant were used to obtain observed and predicted maximum second highest concentrations ($M2H_o$ and $\hat{M2H}_p$, respectively) for the year, as illustrated in Figure A-1. These values are much lower than those calculated from the MPTER results for the 180-receptor network (Table A-2) because they were obtained under the plant's actual 1975 emission rate of 8420 g/s. However, since ground-level concentrations can be reliably assumed to be linearly proportional to emission rates, model performance as measured by the ratio of observed-to-predicted concentrations is unaffected. Thus, model validation results for the 24-hour average concentrations obtained in 1975 under the actual lower emission rate can be reliably applied to

TABLE A-3. Emission limits set by current practice. (See note)

Year	Given Emission Rate (g/s)	Maximum Second Highest Concentration ($\mu g/m^3$)	24-Hour SO_2 NAAQS ($\mu g/m^3$)	Calculated Emission Limit
1973	15,625	357	365	15,975
1974	15,625	313	365	18,221
1975	15,625	337	365	16,923
1976	15,625	367	365	15,498
1977	15,625	285	365	19,941
1973-77	15,625	367	365	15,498

Note:

$$Q_{lim} = \frac{X_{NAAQS}}{DV} \cdot Q_o, \text{ where}$$

Q_{lim} = calculated emission limit

X_{NAAQS} = 24-hour SO_2 NAAQS (365 $\mu g/m^3$)

Q_o = given emission rate, i.e., the hypothetical emission rate of 15,625 g/s used in the example analysis

DV = maximum second highest concentration ($\mu g/m^3$) calculated by the dispersion model (known as the "design value" due to its role in determining emission limits)

model predictions obtained for the hypothetical emission rate as long as we are considering only the ratio of observed-to-predicted 24-hour average concentrations.

HOW CUE USES MODEL EVALUATION RESULTS

Taking the ratio of $M2H_6/\hat{M}2H_6$ from the values given in Figure A-1, we see that the model underpredicted the design value by a factor of 1.4. However, if the model validation experiment had been repeated for a different year, we may well have obtained a different ratio since meteorological conditions would have changed and since monitoring equipment is not perfectly precise. What would be most useful are data for many additional years, from which we could form a distribution of ratios that would represent the range of likely $M2H_6/\hat{M}2H_6$ values. Although additional data are not available, CUE can provide us with such a distribution using the bootstrap technique. In the following paragraphs, we provide a brief description of the calculations carried out in the CUE method.

The first step involves the construction of a large number (say, 1000) of sample years. Each sample year is made up of 365 days drawn at random from the 365 days in calendar year 1975. Any given day may be selected more than once for inclusion in the same sample year. For each day drawn at random for inclusion in a sample year, the corresponding observed and predicted concentrations at each of the six receptors used in the model validation study are recorded. Sets of observed and predicted concentrations obtained in this way for a complete sample year are then used to calculate a pair of observed and predicted design values for that year. Figure A-2 shows how the 1000 pairs of design values are generated.

Because of the way the bootstrap technique works, the maximum second highest concentration is not a design value suitable for this procedure (Stoeckenius et al., 1983). We use instead a related design value known as MDV1, which is calculated as follows: For a given sample year at each receptor, the value exceeded exactly once per year is determined from a

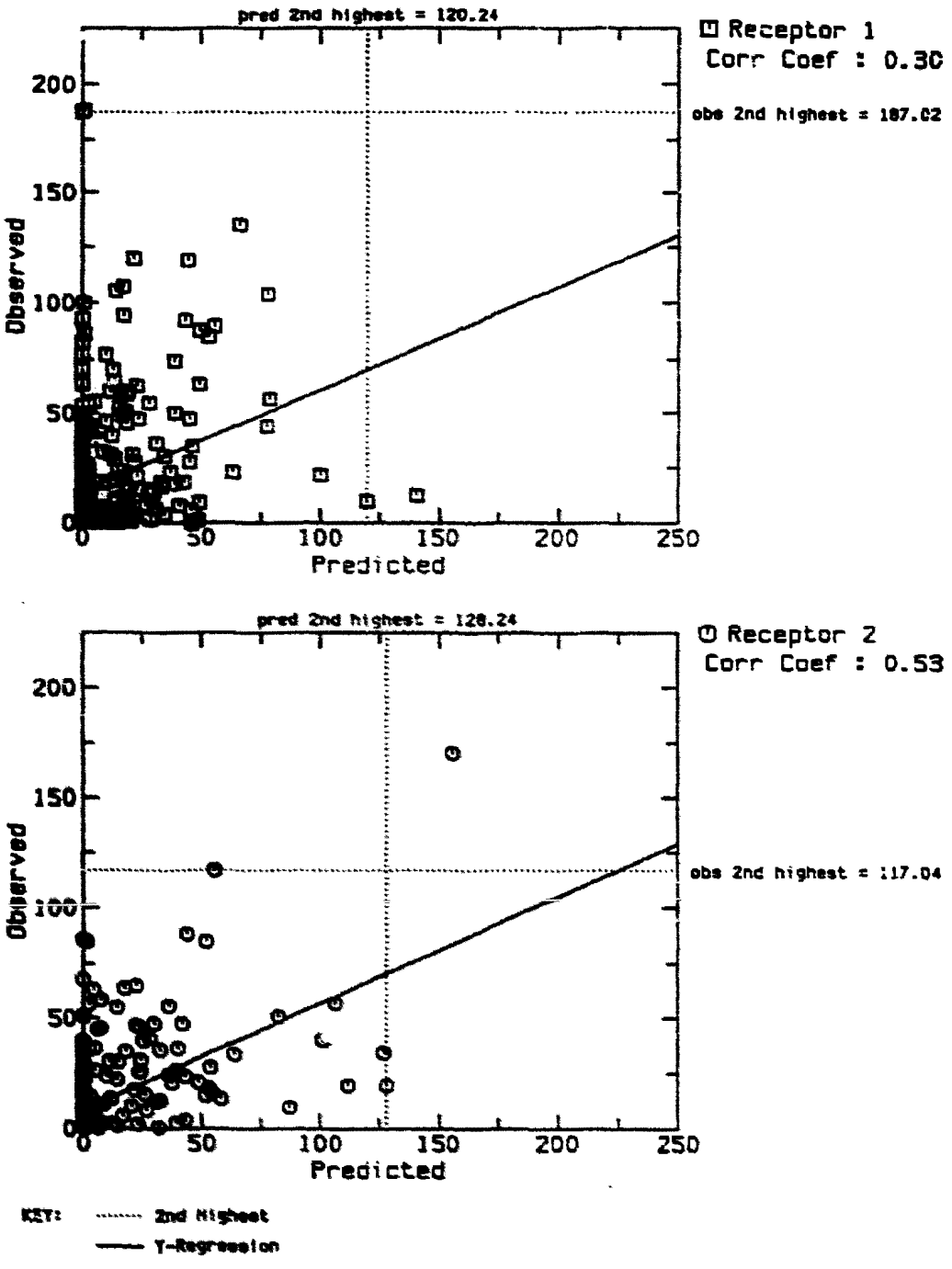


FIGURE A-1. Clifty Creek 24-hour average model validation results.

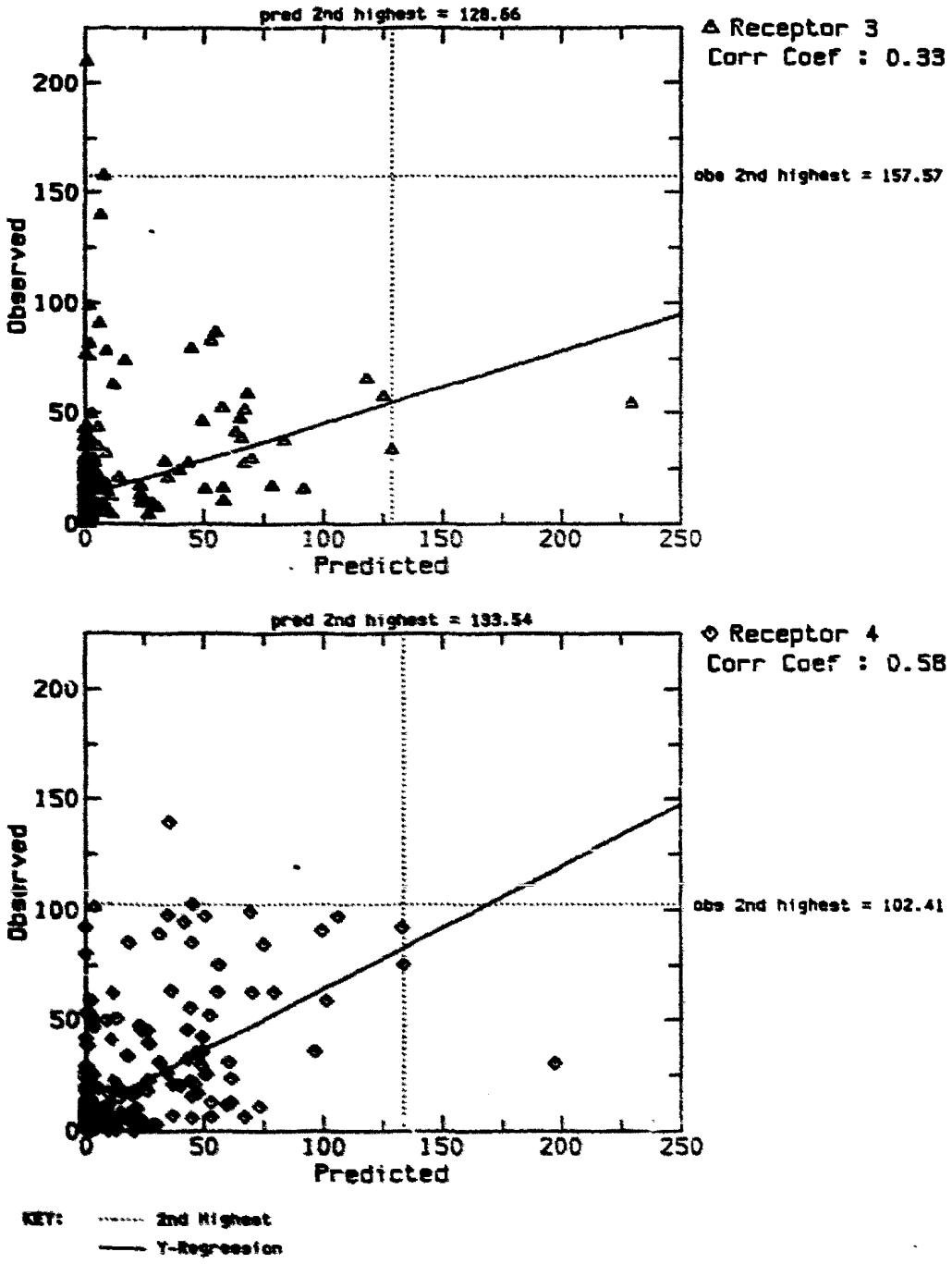


FIGURE A-1 continued

84144

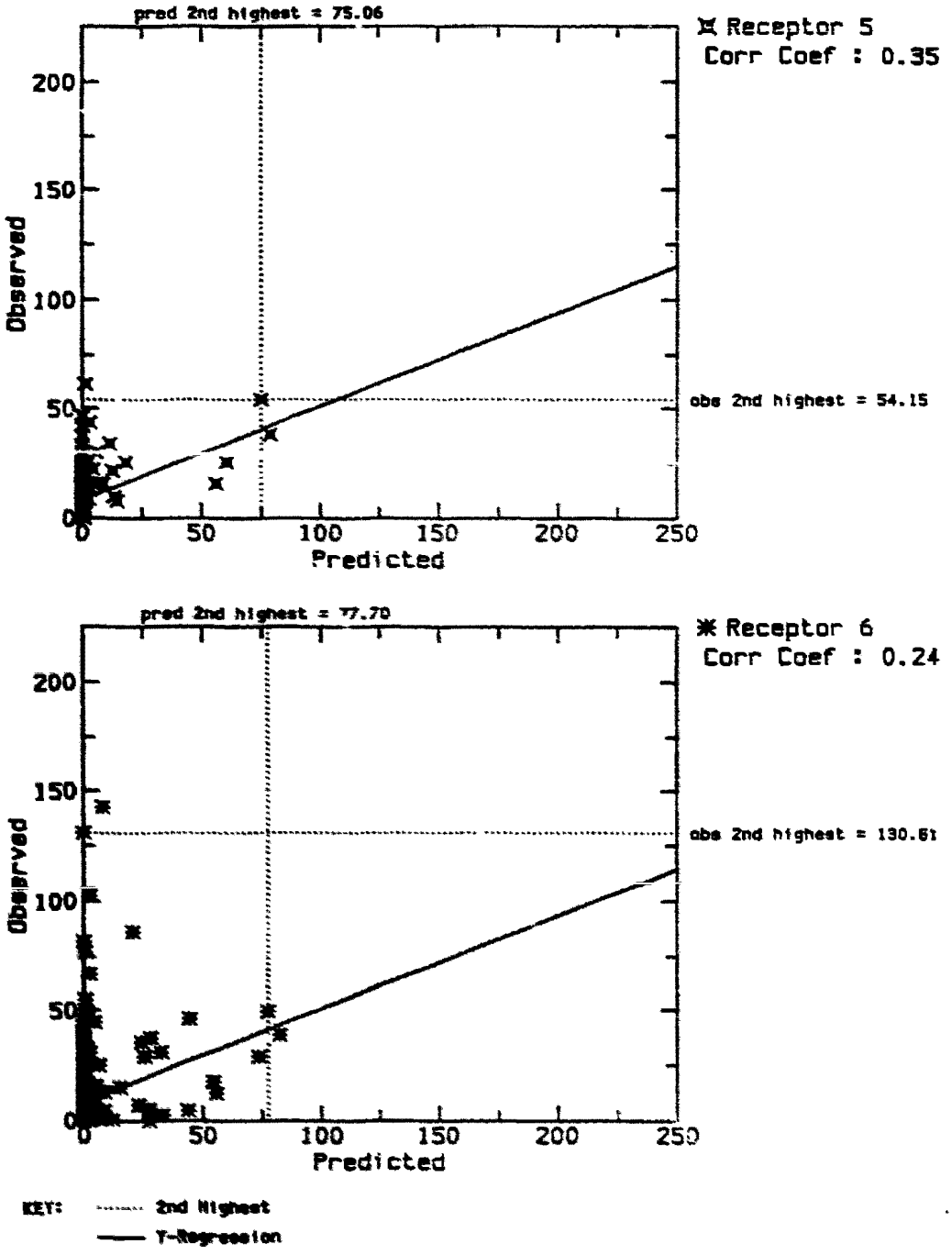


FIGURE A-1 continued

84144

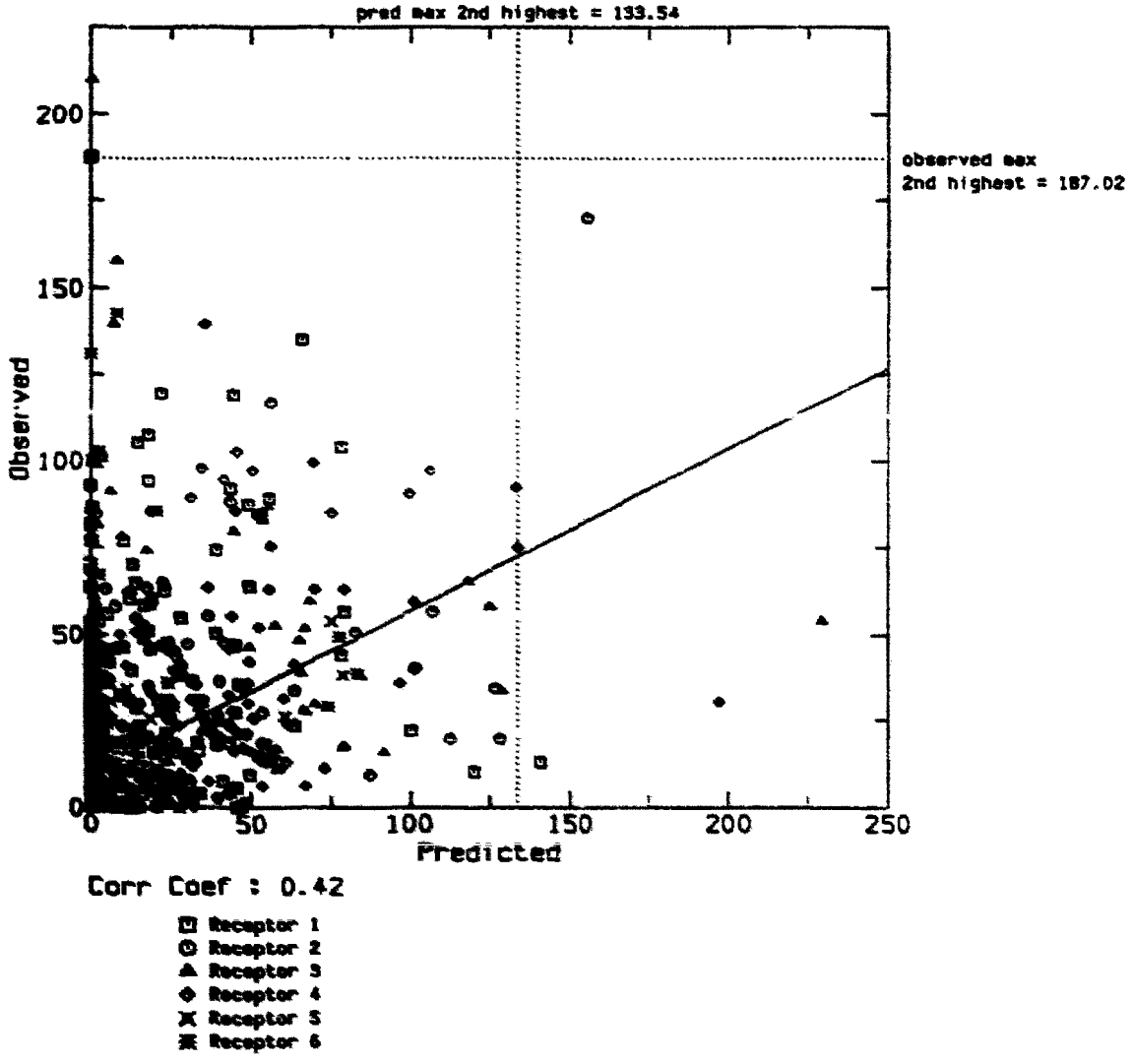


FIGURE A-1 concluded

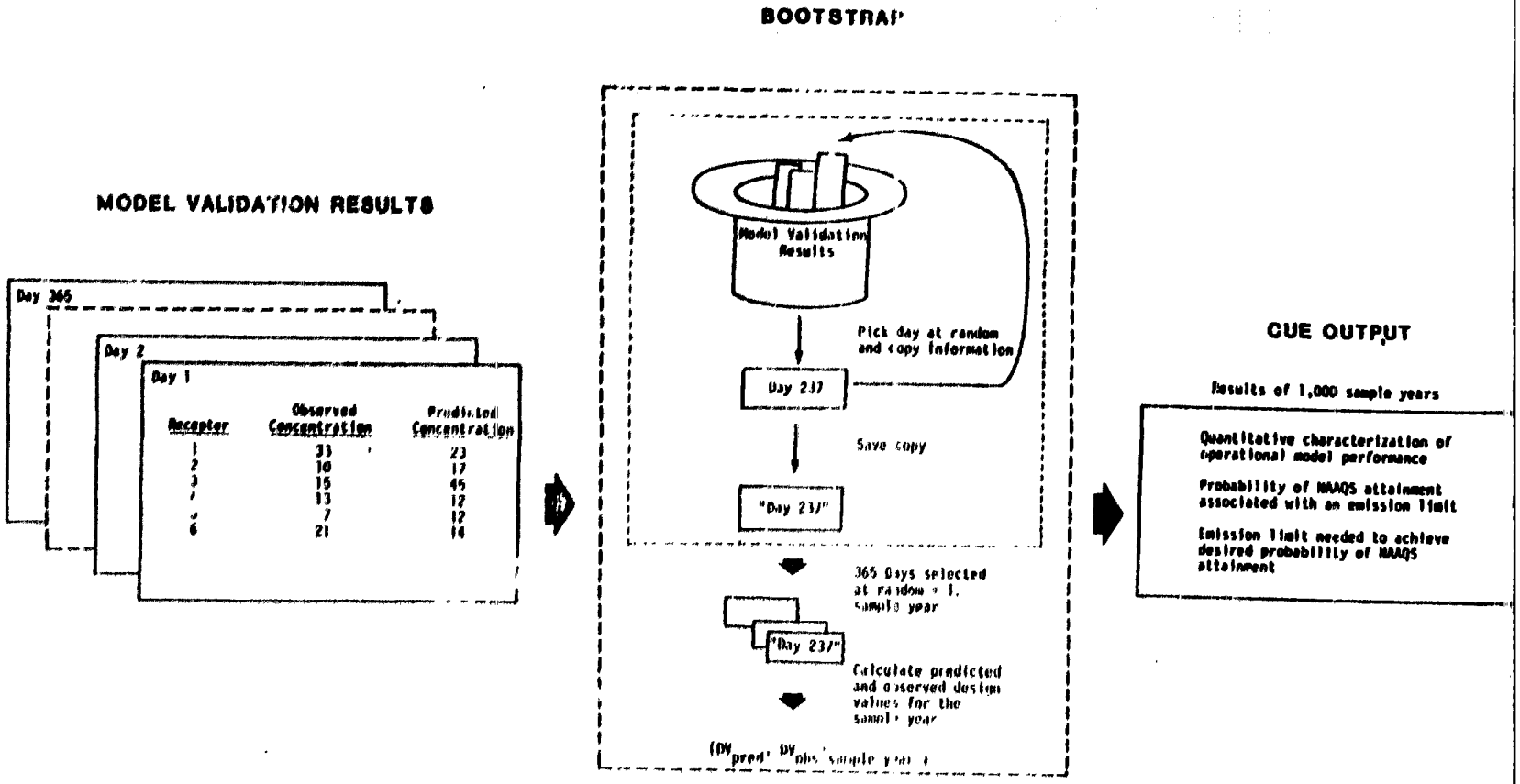


Figure A-2. How CUE works.

tail exponential fit (Breiman et al., 1978) to the upper 10 percent of the 365 observed concentrations. The maximum of these values taken across all 6 receptors is the design value for the sample year. We refer to this value as $MDVI_6$. The predicted design value ($\hat{MDVI}_6$) is calculated in a similar manner. MDVI values are generally comparable to $M2H_6$ values and the smoothing imparted by the tail exponential fit makes them suitable for use in the bootstrap method. A set of 1000 ($MDVI_6$, $\hat{MDVI}_6$) thus generated by the bootstrap technique is shown in Figure A-3.

CHARACTERIZING MODEL PERFORMANCE

The next step in CUE is to use the 1000 ($\hat{MDVI}_6$, $MDVI_6$) pairs shown in Figure A-3 to calculate 1000 ratios R, where $R = MDVI_6/\hat{MDVI}_6$. A cumulative distribution of these ratios is shown in Figure A-4. In the final step, the 1000 values of R are used to "correct" or adjust the M2H design values predicted by the MPTER model for each of 5 years at the 180-receptor network ($\hat{M2H}_{180}$), thus providing an estimate of the monitoring design value, $M2H_{180}$:

$$M2H_{180} = R * \hat{M2H}_{180} \quad . \qquad (1)$$

ESTIMATING THE PROBABILITY OF NAAQS ATTAINMENT ASSOCIATED WITH AN EMISSION RATE

When each of the 1000 values of R is substitued into Equation 1, 1000 $M2H_{180}$ values are produced, forming a distribution describing the likely range of the monitoring design value, i.e., the design value that would have been observed had monitoring data been available over the 180-receptor network. The modeled design value, $\hat{M2H}_{180}$, for each year thus rescales the distribution of R to form a probability, or uncertainty, distribution for $M2H_{180}$. These distributions are plotted in cumulative form for each year in Figure A-5. For each year the figure indicates the

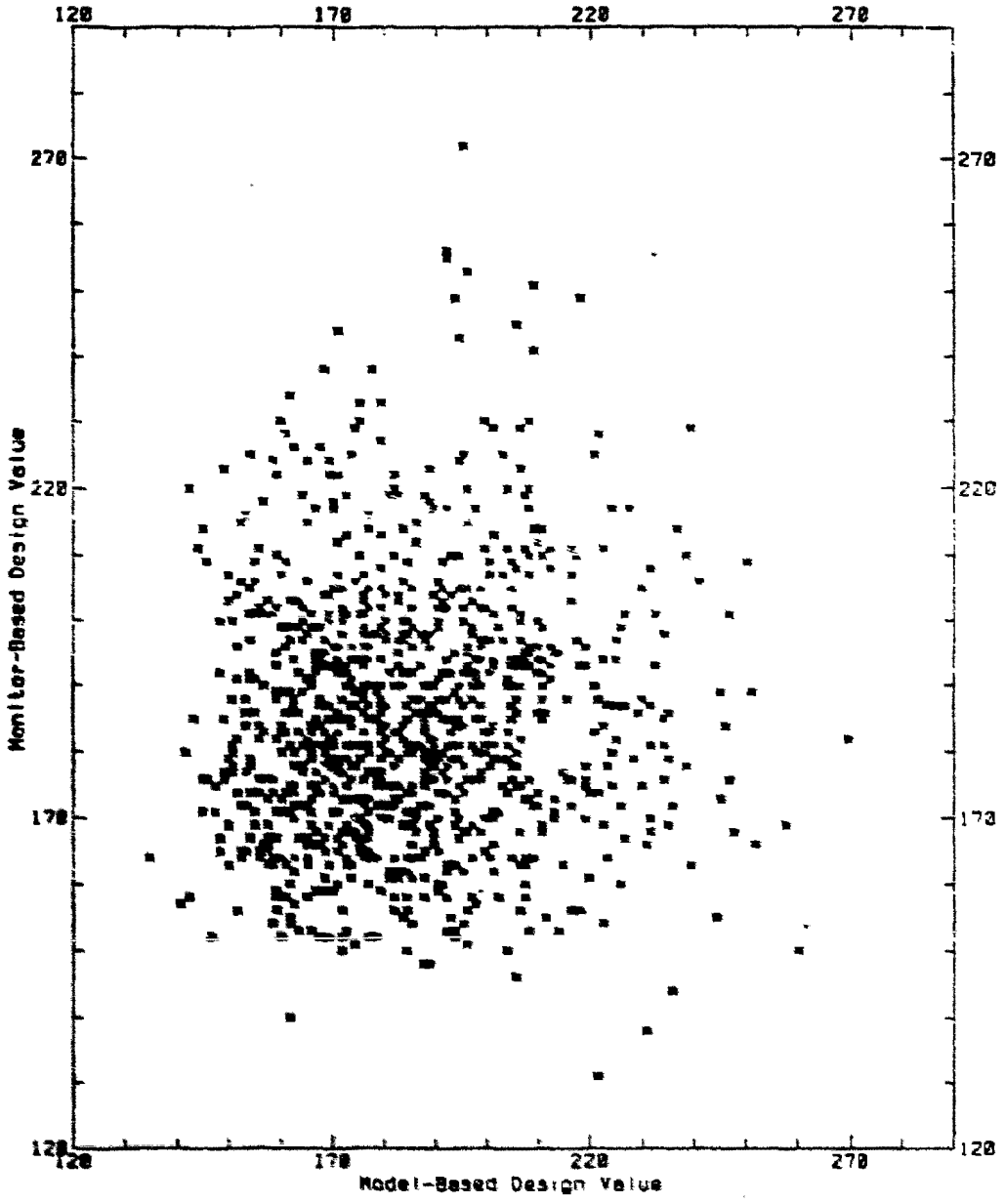
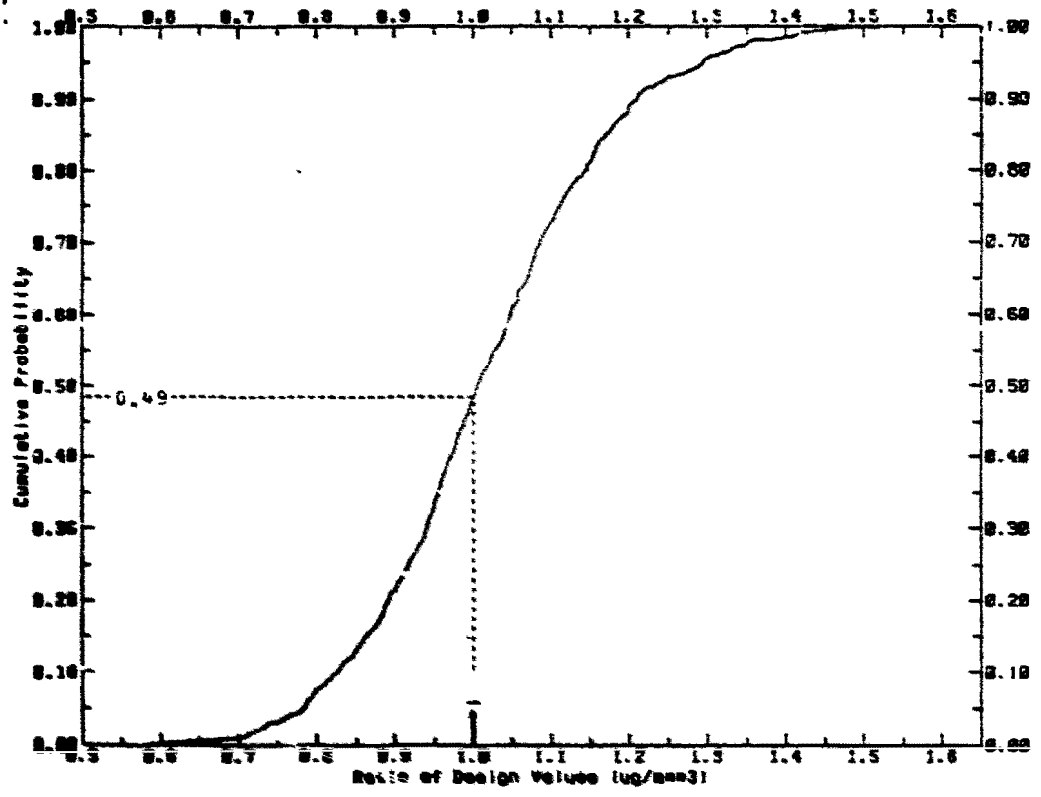


FIGURE A-3. Scatter plot of observed (monitor-based) and predicted (model-based) MDVI design values generated by the bootstrap method.



Cliffy Creek 1975 (24-hr avg)

FIGURE A-4. Cumulative distribution of $R = MDV1_6/\hat{MDV1}_6$ obtained from bootstrap results in Figure A-3.

84144

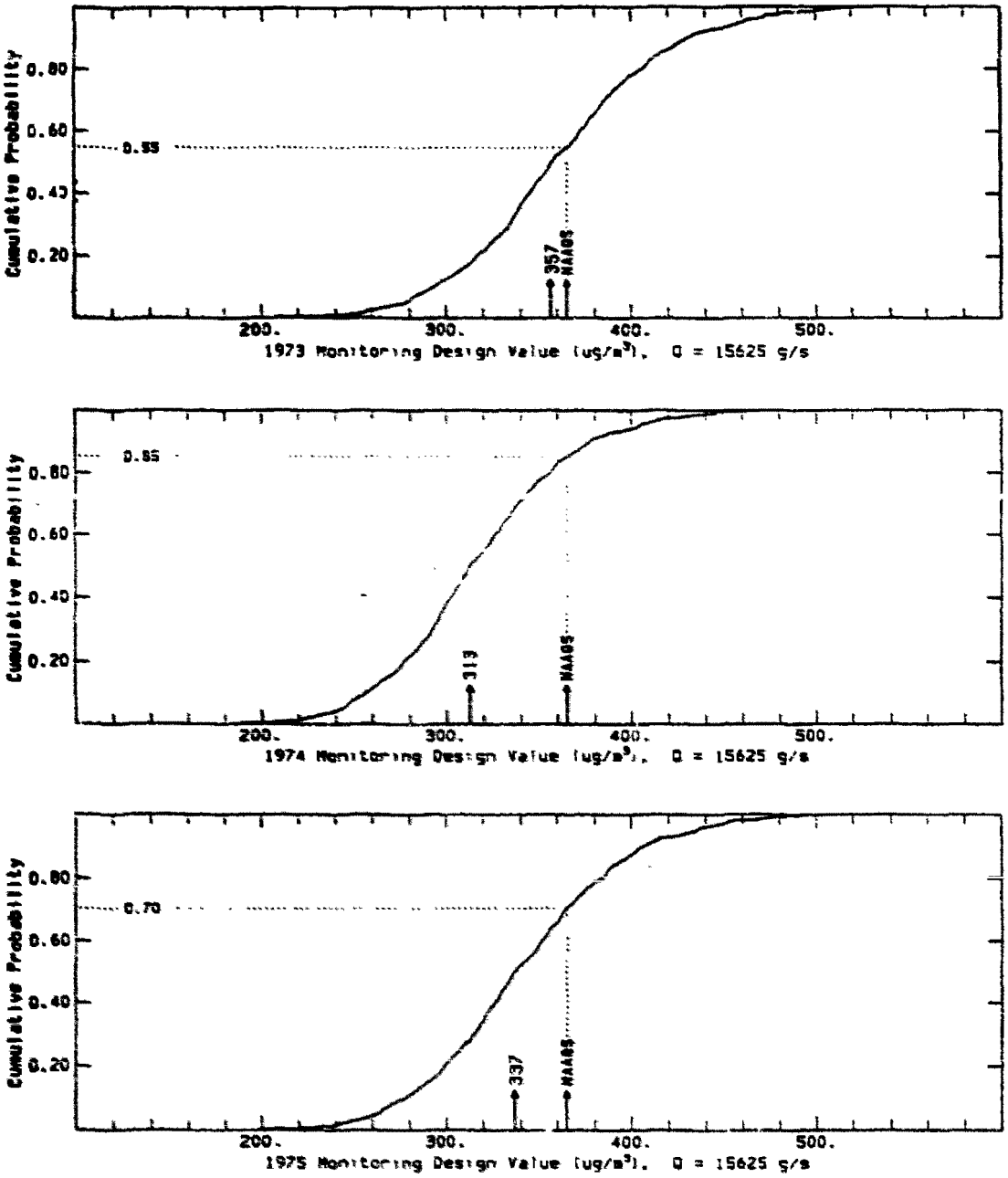
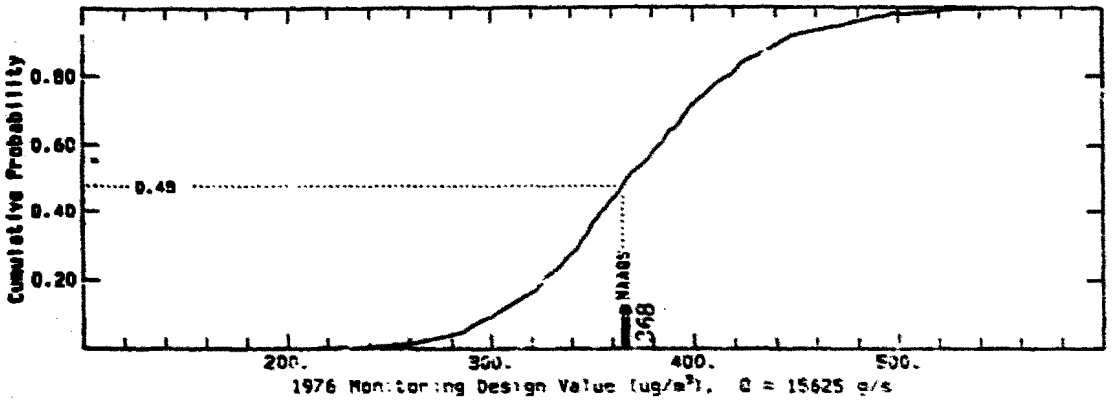
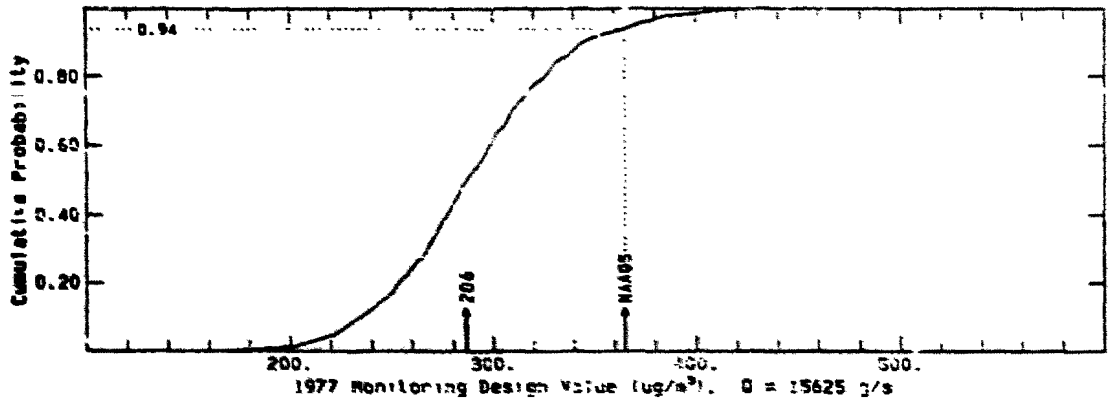


FIGURE A-5. Cumulative probability distributions for 1973 - 1977 monitoring design values, assuming a 15,625 g/s annual average rate of emission. (Both the model-predicted concentration and the NAAQS concentration of 365 ug/m³ are indicated. The probability of NAAQS attainment is also shown.)





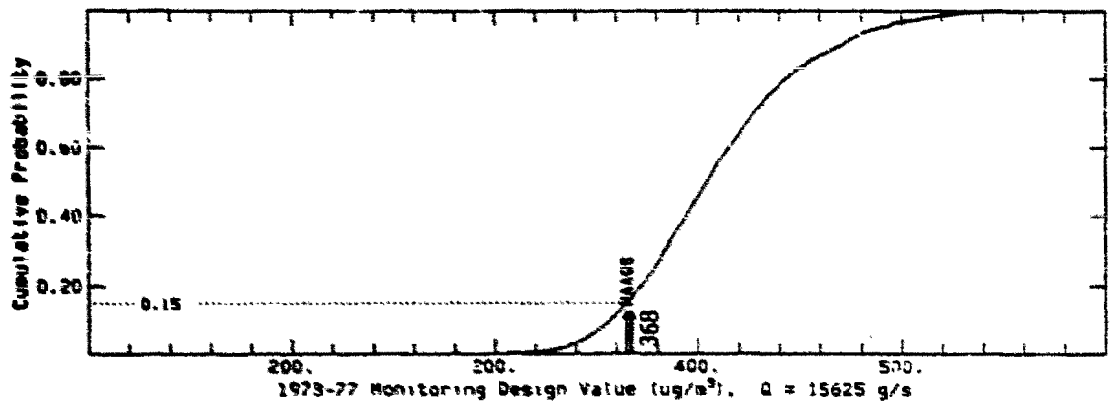


FIGURE A-5. concluded.

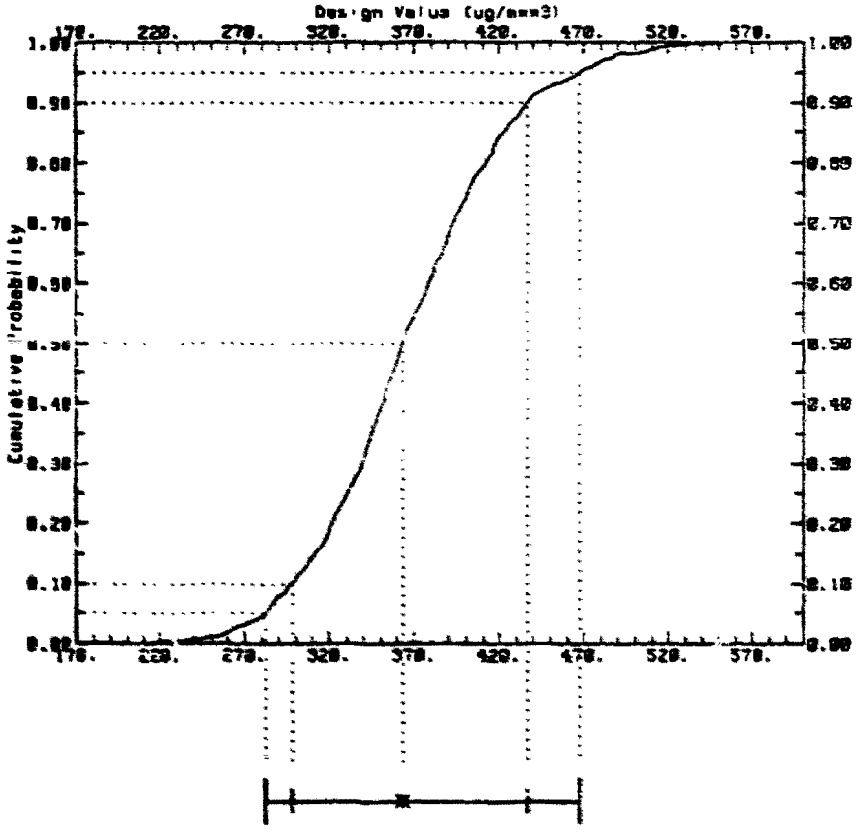
estimated probability that $M2H_{180}$ is less than or equal to the NAAQS concentration of 365 $\mu g/m^3$. Thus, the probability of NAAQS attainment is shown for each year.

In many cases, decision makers are concerned with the largest design value taken over a five-year period. As noted earlier (Table A-2), the largest $\hat{M2H}_{180}$ just exceeds the NAAQS. Decision makers might wish to know with what probability the largest of the monitoring design values would have been greater than (or less than) the NAAQS. Let's call this largest design value the 1973-77 monitoring design value, denoted as $X_{1:5}$ (the largest value out of a set of five). We can find the distribution of $X_{1:5}$ by assuming that the design value in each year is a random value, whose distribution is shown in Figure A-5, independent of the value in any other year. Denoting the design value in year i by X^i, we have $X_{1:5} = \max(X^1, X^2, X^3, X^4, X^5)$. Then for any concentration x_0

$$\Pr(X_{1:5} < x_0) = \prod_{i=1}^{5} \Pr(X^i < x_0) \quad . \tag{2}$$

Equation 2 defines the cumulative distribution of $X_{1:5}$, which is also shown in Figure A-5. Of particular interest is the result obtained from Equation 2, where $x_0 = 365$ $\mu g/m^3$, the 24-hr SO_2 NAAQS. As shown in Figure A-5, there is a 15 percent chance that the largest design value taken over the five years would be less than or equal to the standard. This probability is much lower than the probability in individual years because we are focusing attention on the maximum of five design values, any one of which might be larger than the standard, drawn from the five annual distributions in Figure A-5.

An alternative, more compact, way of displaying the information contained in Figure A-5 is to use a box plot. As shown schematically in Figure A-6, a box plot is constructed by simply noting the concentrations corresponding to various percentiles of the distribution with special symbols. Box plots allow information from more than one year to be



Cliffy Creek 1973 (24-hr avg)

FIGURE A-6. Example of the derivation of a box plot from a cumulative distribution.

displayed conveniently and clearly on a single page (Figure A-7). As in the cumulative distribution plots (Figure A-5), the arrows in Figure A-7 indicate the location of $\hat{M2H}_{180}$ in each year. Figure A-7 shows that there is a chance that the NAAQS would not have been exceeded by an observed design value in each year. These probabilities are listed in Table A-4, which illustrates NAAQS attainment probabilities ranging from 48 percent for 1976 to 94 percent for 1977. Using Equation 2, and thus pooling information across the 5 years, we see that the chance that 5-year maximum design value is below 365 $\mu g/m^3$ is only 15 percent. As we shall discuss shortly, Table A-4 also shows probabilities of NAAQS attainment corresponding to emission rates equal to the emission limits prescribed by current practice.

Solely on the basis of the original, conventionally interpreted model results ($\hat{M2H}_{180}$ values), a decision maker might conclude that the Clifty Creek plant was in violation of the standard in one (1976) of the five years. From our CUE calculations of modeling uncertainty, however, we estimate that there is a 15 percent chance that this would be an incorrect conclusion based on the worst-in-five year design value and that there is a 48 percent chance that this would be an incorrect conclusion based on just the 1976 design value. Alternatively, if a decision maker concluded, based on conventional use of model results (i.e., all five years), that the plant did not demonstrate attainment, then such a decision maker could be 85 percent confident of being correct. However, if a decision maker concluded that the plant did not demonstrate attainment based on one year (1976), an estimate of the probability that such a decision is correct is 52 percent. To reiterate, the 15 and 48 percent probability estimates represent the probabilities that no emission reduction is needed.

Since we can assume that concentrations are linearly proportional to emission rates, the results shown in Figure A-7 can be easily rescaled to show what would happen if the Clifty Creek plant were to operate at a different emission rate. In particular, we can rescale the results to show the probability of NAAQS attainment that would result if the plant were to operate at an emission limit, Q_{lim}, calculated strictly on the basis of the $\hat{M2H}_{180}$ values listed in Table A-2:

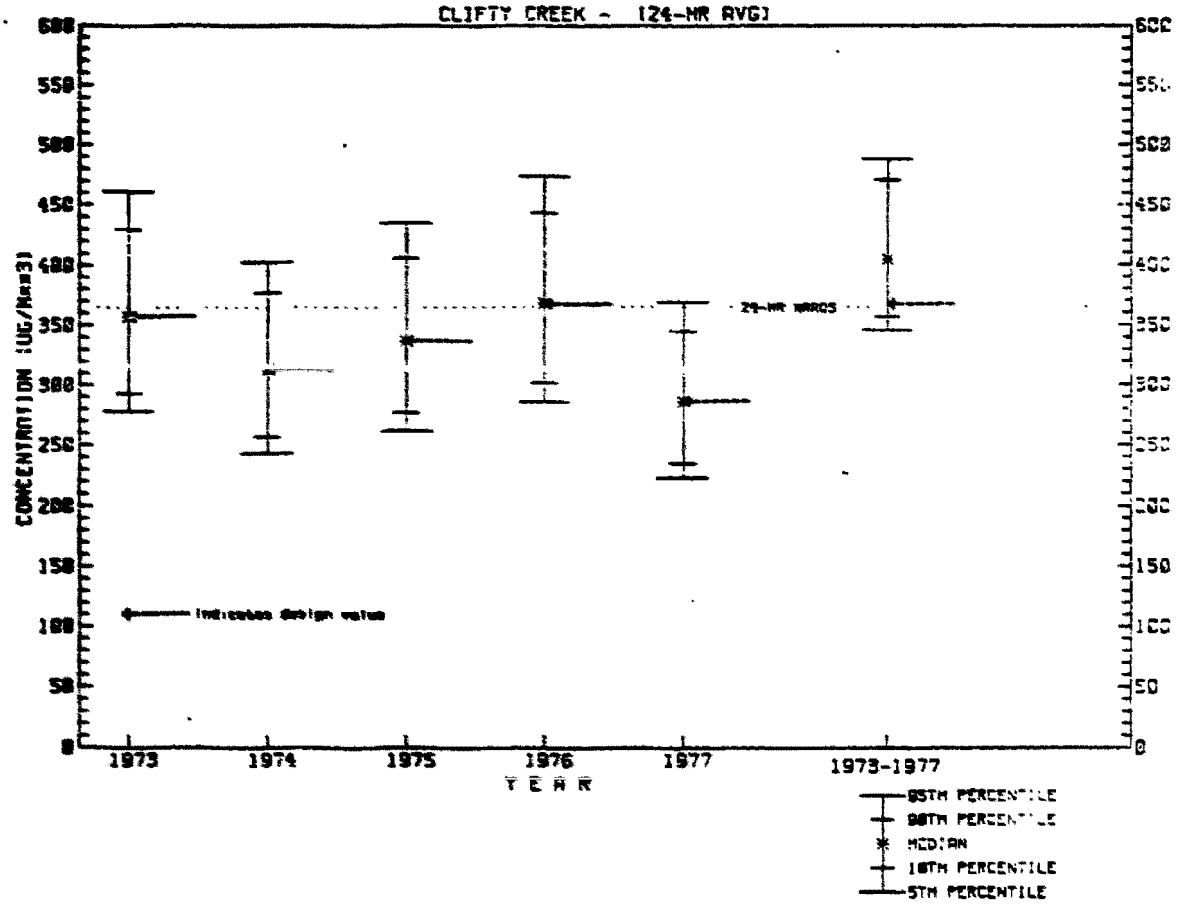


FIGURE A-7. Box plots depicting probability distributions for monitoring design values, assuming a 15,625 g/s annual average rate of emission. The distribution for each year from 1973 to 1977 is depicted, along with the distribution for the maximum of the five annual design values.

TABLE A-4. Probabilities of NAAQS attainment associated with the given emission rate and with emission limits set by current practice.

	Given Emission Rate		Calculated Emission Limit	
Year	Emission Rate (g/s)	Associated Probability (%)	Emission Rate (g/s)	Associated Probability (%)
1973	15,625	55	15,975	49
1974	15,625	85	18,2[illegible]	49
1975	15,625	70	16,923	49
1976	15,625	48	15,498	49
1977	15,625	94	19,941	49
1973-77	15,625	15*	15,498	17*

* The probability that the 1973-77 maximum monitoring design value concentration resulting from the indicated 1973-77 emission rate would not exceed the NAAQS concentration of 365 $\mu g/m^3$.

$$Q_{lim} = \frac{\chi_{NAAQS}}{\widetilde{M2H}_{180}} \cdot Q_o$$

where χ_{NAAQS} = the 24-hour NAAQS for SO_2 (365 $\mu g/m^3$), and Q_o = the original emission rate used in constructing Figure A-7 (15,625 g/s).

Values of Q_{lim} correspond to current EPA practice for setting emission limits. These values, along with the probabilities of attainment that would result from their implementation, are listed in Table A-4.

ESTIMATING THE EMISSION RATE ASSOCIATED WITH A DESIRED PROBABILITY OF NAAQS ATTAINMENT

It is also possible to calculate an emission rate, Q_α, such that any one of the six uncertainty distributions in Figure A-5 we choose to focus attention on will be altered to indicate an α percent probability of NAAQS attainment. We can calculate Q_α from

$$Q_\alpha = \frac{\chi_{NAAQS}}{DV_\alpha} \cdot Q_o \qquad (3)$$

where DV_α = the design value corresponding to an α percent probability of attainment as determined from Figure A-4 for the distribution of interest.

Table A-5 provides values of Q_{lim} and Q_α, the latter corresponding to α = 50, 90, 95, and 99 percent probabilities of attainment determined from Equation 3. The various emission limits are shown for each year and the worst-in-five years. From Table A-5 we see that the emission limit based on current practice would be 15,498 g/s. We also see that considering model uncertainty and accepting a 50 percent probability of NAAQS attainment based on a worst-in-five year policy, the emission limit would

TABLE A-5. Comparison of emission limits set by current practice and those calculated by CUE to yield a desired probability of NAAQS attainment.

Year	Highest Second-High Concentration ($\mu g/m^3$)	Emission Limit (g/s)	Associated Probability of NAAQS Attainment* (%)	Emission Limit (g/s) for Desired Probability of NAAQS Attainment			
				50%	90%	95%	99%
1973	357	15,975	49	15,901	13,207	12,331	11,294
1974	313	18,221	49	18,150	15,075	14,075	12,891
1975	337	16,923	49	16,838	13,985	13,057	11,959
1976	368	15,498	49	15,442	12,825	11,974	10,967
1977	286	19,941	49	19,824	16,466	15,373	14,080
1973-77**	368	15,498	17	14,092	12,098	11,684	10,886

* Computed using CUE as in Table 2.

** This row refers to the probability that the 1973-77 maximum design value concentration resulting from the indicated 1973-77 emission rate would not exceed the NAAQS concentration of 365 $\mu g/m^3$.

be set at 14,092 g/s, or reduced by about 9 percent from current practice. Alternatively, if an acceptable probability of attainment is established at 90 percent for a worst-in-five year policy, the resulting emission limit would be 12,098 g/s, or 22 percent lower than current practice. It is also of interest to note that if the estimate of NAAQS attainment probability is based on a single year, say 1976 because it produces the highest design value estimates, and an acceptable probability is established to be 50 percent, then the resulting emission limit would be 15,442 g/s, which is virtually identical to the emission limit that results from current practices (15,498 g/s).

Many other comparison can also be made between alternative emission limit setting and model uncertainty risk management policies. To facilitate such comparisons, Figure A-8 illustrates emission limits obtained from current practice and two probability of attainment levels: 50 and 90 percent.

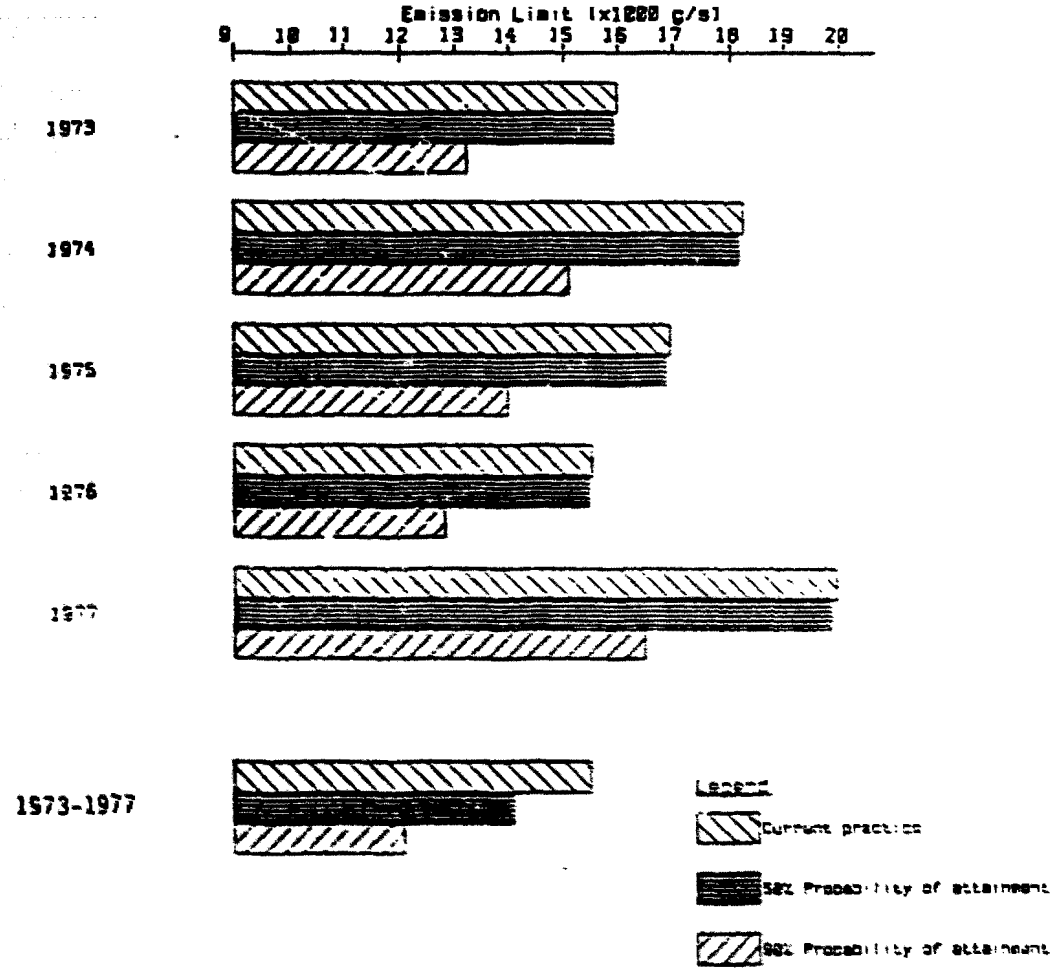


FIGURE A-8. Comparison of emission limits obtained via current practice with emission limits required to achieve 50 and 90 percent probabilities of NAAQS attainment.

Appendix B

THE APPLICATION OF CUE TO BIASED OR IMPRECISE MODELS

Appendix B

THE APPLICATION OF CUE TO BIASED OR IMPRECISE MODELS

When we used Clifty Creek 1975 data and CUE to exercise the Gaussian dispersion model MPTER, it exhibited almost no bias in design value and was also fairly precise. But it would be useful to see applications of CUE to a variety of settings in which model performance is markedly better or worse than that of MPTER for Clifty Creek. In consideration of available resources, we have had to simulate these other settings by artifically altering the design values derived from our study of MPTER with the Clifty Creek data.

There are many ways in which design values could be altered for the purpose of altering a model's bias or imprecision. We decided to utilize observation-prediction ratios since (1) they play a key role in our work, and (2) their characterization of model performance is familiar to the modeling community. For example, suppose that a model is "within a factor of 2 (of monitored values) half the time." Specifically, suppose that the 25th percentile of observed/predicted ratios is equal to one-half, and that the 75th percentile is equal to 2; or, taking logarithms, that the 25th percentile of the logged ratio is equal to - ln 2, and the 75th percentile is equal to + ln 2. Then, in a sense, the model is unbiased because the distribution of logged ratios is centered about zero, but is also rather imprecise, since the distribution of logged ratios is fairly broad.

This point of view suggests that we alter our bootstrapped design values by altering the distribution of logged ratios. That is, we can bias the model by shifting the distribution away from zero, or we can make the model more or less precise by multiplying logged ratios by a constant that is less than or greater than unity.

What we have in mind, then, is to alter the ratio (R) of monitored design values (m) to predicted design values (p) to produce a new ratio ($\tilde{R}$) such that

$$(-\ln \tilde{R}) = \beta + \phi \cdot (-\ln R).$$

(The reason for the minus sign will be made clear in a moment.) If we agree to leave the monitored design value unaltered, i.e., to alter bias and impecision by altering the predicted design value (p) alone, we obtain:

$$\tilde{p} = e^{\beta} \cdot m^{1-\phi} \cdot p^{\phi} ,$$

where $\tilde{p}$ is the altered version of p.

Tables B-1 and B-2 show that the altered model will have a greater tendency to underpredict (or less of a tendency to overpredict) if β is negative, and will have a greater tendency to overpredict (or less of a tendency to underpredict) if β is positive. (We decided to work with - ln R rather than ln R so that the correspondence between the sign of β and model overprediction or underprediction would be more natural.) Similarly, the altered model will be more or less precise if ϕ is, respectively, less than or greater than unity, so that ϕ is an index of model imprecision.

Figures B-1 and B-2 show, respectively, the distribution of - ln $\tilde{R}$ and the relationship between $\tilde{p}$ and m as determined by values for β and ϕ. To introduce an approximate 30 percent bias, we selected three values for

TABLE B-1. The behavior of $\tilde{R}$ for certain combinations of β and φ.

	$\beta < 0$	$\beta = 0$	$\beta > 0$
$\varphi = 0$	$-\ln \tilde{R} = \beta < 0$	$-\ln \tilde{R} = 0$	$-\ln \tilde{R} = \beta > 0$
$0 < \varphi < 1$		$\lvert\ln \tilde{R}\rvert < \lvert \ln R\rvert$	
$\varphi = 1$	$-\ln \tilde{R} < -\ln R$	$-\ln \tilde{R} = -\ln R$	$-\ln \tilde{R} > -\ln R$
$\varphi > 1$		$\lvert\ln \tilde{R}\rvert > \lvert\ln R\rvert$	

TABLE B-2. The behavior of $\tilde{p}$ for certain combinations of β and φ.

	$\beta < 0$	$\beta = 0$	$\beta > 0$
$\varphi = 0$	$\tilde{p} = e^{\beta} \cdot m < m$	$\tilde{p} = m$	$\tilde{p} = e^{\beta} \cdot m > m$
$0 < \varphi < 1$		$\tilde{p}$ closer than p to m	
$\varphi = 1$	$\tilde{p} = e^{\beta} \cdot p < p$	$\tilde{p} = p$	$\tilde{p} = e^{\beta} \cdot p > p$
$\varphi > 1$		$\tilde{p}$ farther than p to m	

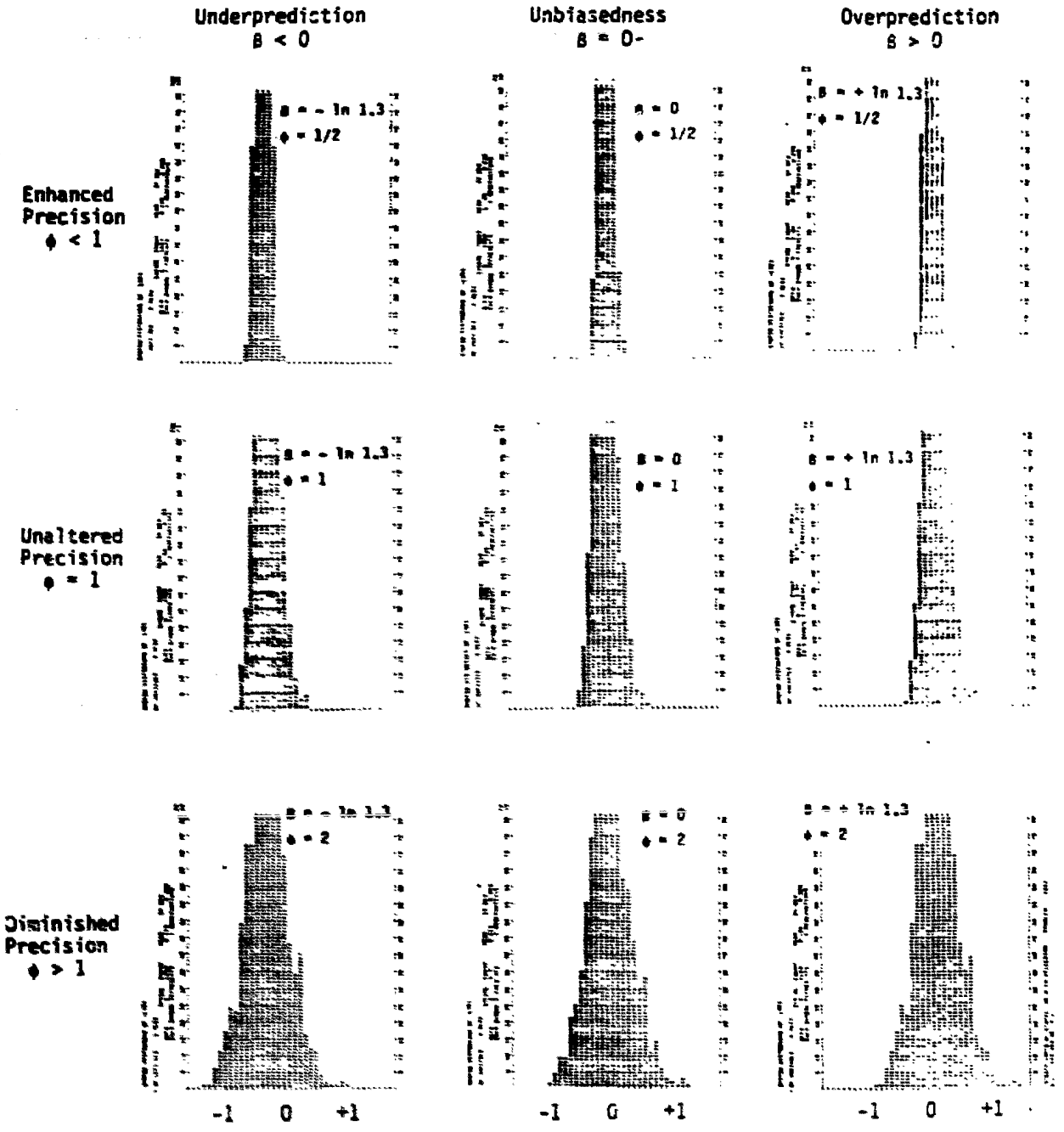


FIGURE B-1. Histograms of -ln R for certain combinations of β and φ.

84144

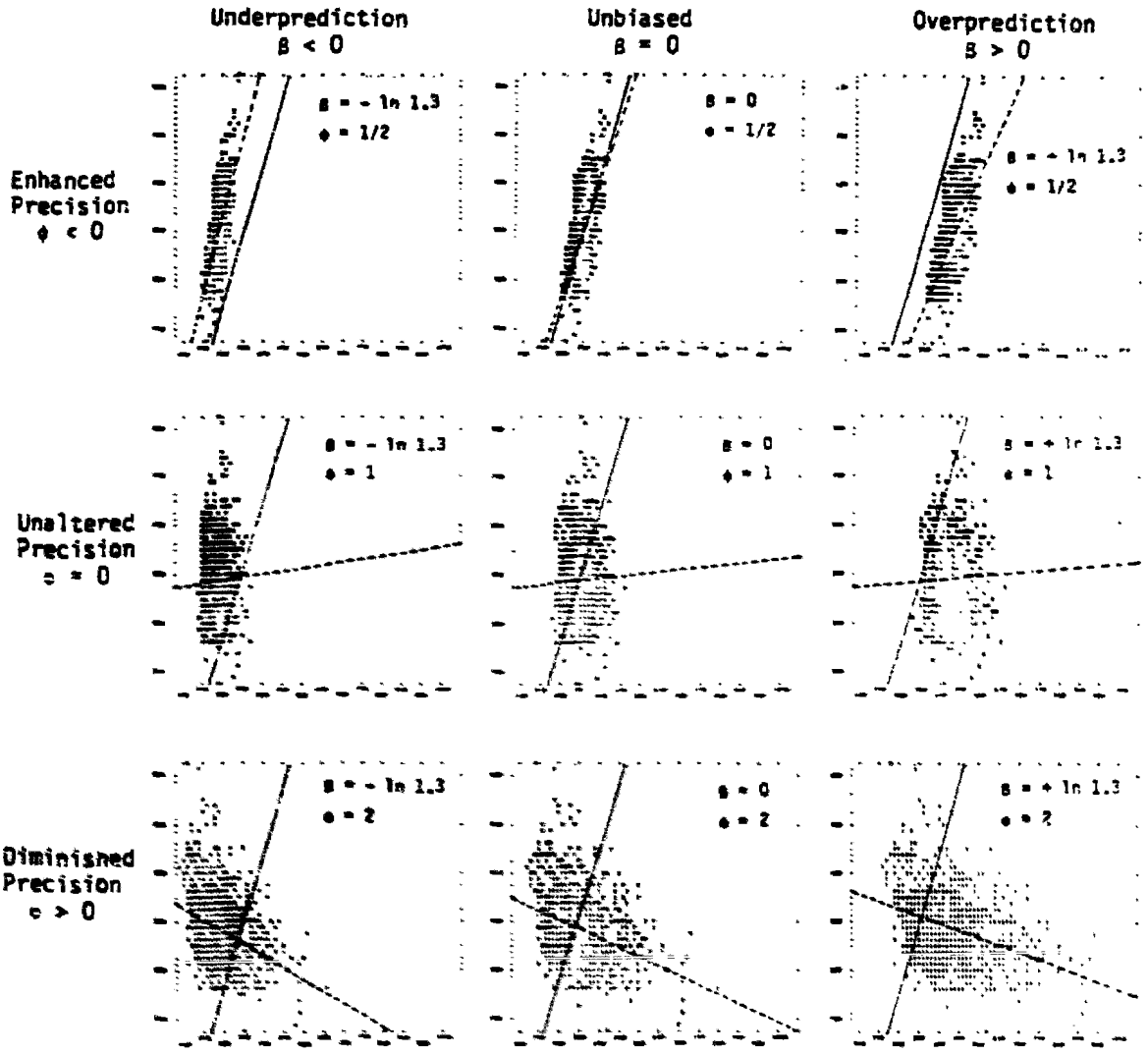


FIGURE B-2. Scatter diagrams of altered model-based design values (horizontal axis) versus monitor-based design values (vertical axis) for certain combinations of β and φ.

β, namely, zero and $\pm \ln 1.3$. A change in precision by a factor of 2 also seemed to be a natural choice so we have assigned ϕ the values of 1/2, 1, and 2. Summary statistics for the resulting distributions of predicted design values and design value differences are shown, respectively, in Tables B-3 and B-4.

To carry our discussions a step further, it is interesting to calculate the effect of altering the bias and precision of a model on ratios, design values, and emission limits. Table B-5 shows these effects using the notation of Section 3.

Since ratios are of observed divided by predicted design values, an increased tendency of the model to overestimate monitored design values leads to smaller ratios and, conversely, an increased tendency to underestimate leads to larger ratios. Thus, if model predictions are altered by a factor of e^{β}, these ratios are altered by a factor of $e^{-\beta}$. Also, the distribution of the altered ratios will be narrower than the distribution of original ratios if ϕ is less than unity, and the altered distribution will be wider if ϕ is greater than unity.

If we alter the bootstrapped predicted design values by a factor of e^{β}, it seems reasonable to alter $\hat{DV}$, the design value predicted from the original data (prior to bootstrapping), by the same factor. Since the original prediction tends to fall in the center of the bootstrap distribution, it seems reasonable to make no further adjustment to $\hat{DV}$.

To construct the distribution for the unknown monitoring design value, CUE multiplies $\hat{DV}$ by bootstrapped ratios, R. Specifically, quantiles for the unknown monitoring design value are obtained directly from quantiles of the distribution of ratios, as indicated in Table B-3. The same procedure would be followed for the altered model, as indicated.

Finally, emission limits are obtained as indicated in Table B-5. That is, the limit is calculated from either the original predicted design value, $\hat{DV}$ (current practice) or DV_{α} (CUE). Again, the same procedures would be followed for the altered model.

TABLE B-3. Summary statistics for predicted design values (MDV1PR) for certain combinations of β and ϕ.

	$\beta = -\ln 1.3$		$\beta = 0$		$\beta = +\ln 1.3$	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
$\phi = 1/2$	263.10	20.899	342.03	27.169	444.65	35.320
$\phi = 1$	263.37	31.305	342.38	40.701	445.10	52.911
$\phi = 2$	268.67	71.325	349.27	92.725	454.06	120.54

TABLE B-4. Summary statistics for design value differences (MDV1PR-MDV10B) for certain combinations of β and ϕ.

	$\beta = -\ln 1.3$		$\beta = 0$		$\beta = +\ln 1.3$	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
$\phi = 1/2$	- 80.59	26.38	- 1.66	26.39	+ 100.96	28.55
$\phi = 1$	- 80.32	46.49	- 1.31	53.04	+ 101.41	62.65
$\phi = 2$	- 75.02	90.22	+ 5.58	110.40	+ 110.37	137.18

TABLE B-5. The effect of altering the bias and precision of a model on ratios, design values, and emission limits.

	Description	Original Model	Altered Model
Ratios of observed-to-predicted design values	αth quantile of bootstrap distribution	R_α	$R'_\alpha = e^{-\theta} \cdot R_\alpha^{\phi}$
Design values	Original prediction	$\hat{DV}$	$\hat{DV}' = e^{\theta} \cdot \hat{DV}$
	αth quantile of bootstrap distribution	$DV_\alpha = R_\alpha \cdot \hat{DV}$	$DV'_\alpha = R'_\alpha \cdot \hat{DV}'$
(c) Emission limits	Current practice	$Q_{lim} = \frac{X_{NAAQS}}{\hat{DV}} \cdot Q_0$	$Q'_{lim} = \frac{X_{NAAQS}}{\hat{DV}'} \cdot Q_0$
	CUE	$Q_\alpha = \frac{X_{NAAQS}}{\hat{DV}} \cdot Q_0$	$Q'_\alpha = \frac{X_{NAAQS}}{DV'_\alpha} \cdot Q_0$

Putting these results together we obtain

$$DV'_\alpha = R'_\alpha \cdot DV'$$
$$= R_\alpha^\phi \cdot \hat{DV}$$

so that the distribution of the unknown monitoring design value generated by CUE is not affected by model bias. That is, CUE compensates for bias and produces the distribution that would be derived for an unbiased model. We see, however, that the distribution is affected by the alteration of model precision. As the model increases in precision (as ϕ approaches zero), the distribution of R narrows about unity, and the distribution of the unknown monitoring design value narrows about the original, unaltered, predicted design value, DV.

We may also write

$$DV'_\alpha = R_\alpha^\phi \cdot \hat{DV}$$
$$= R_\alpha^{\phi-1} \cdot R_\alpha \cdot \hat{DV}$$
$$= R_\alpha^{\phi-1} \cdot DV_\alpha$$

to obtain

$$Q'_\alpha = \frac{X_{NAAQS}}{DV'_\alpha} \cdot Q_0$$
$$= R_\alpha^{1-\phi} \frac{X_{NAAQS}}{DV_\alpha} \cdot Q_0$$
$$= R_\alpha^{1-\phi} \cdot Q_\alpha$$

Again, we see that the alteration of model bias does not alter the emission limits produced by CUE. However, as the model becomes more precise (as ϕ approaches zero), the emission limit, Q'_α, needed to achieve a desired probability, α, of NAAQS attainment converges to the emission limit obtained from current practice.

The emission limit derived by current practice is, of course, affected by the alteration of model bias, since

$$\begin{aligned} Q'_{lim} &= \frac{X_{NAAQS}}{DV'} \cdot Q_0 \\ &= \frac{1}{e^{\delta}} \cdot \frac{X_{NAAQS}}{\hat{DV}} \cdot Q_0 \\ &= \frac{1}{e^{\delta}} \cdot Q_{lim} \quad . \end{aligned}$$

Thus, the more the model tends to overpredict design values, the smaller (less lenient) the emission limit becomes; conversely, the more the model tends to underpredict, the greater (more lenient) the emission limit becomes.

Thus, an important feature of CUE is that its emission limits compensate for model bias, whereas those obtained by current practice do not.

Table B-6 illustrates these effects for emission limits obtained from 1973 Clifty Creek data under the various alterations of model bias and precision.

TABLE B-6. 1973 emission limits set by current policy and policies designed to achieve 50 and 90 percent probabilities of NAAQS attainment for different scenarios of model bias and imprecision.

	Underprediction ($\beta < 0$)	Unbiasedness ($\beta = 0$)	Overprediction ($\beta > 0$)
Enhanced	20,768[a]	15,975	12,289
Precision	15,941[b]	15,941	15,941
($\phi = 1/2$)	14,528[c]	14,528	14,528
Unaltered	20,768	15,975	12,289
Precision	15,901	15,901	15,901
($\phi = 1$)	13,207	13,207	13,207
Diminished	20,768	15,975	12,289
Precision	15,821	15,821	15,821
($\phi = 2$)	10,915	10,915	10,915

[a] Current policy
[b] 50 percent policy
[c] 90 percent policy