Simplified Method for Evaluation of Control Strategies for Revision of State Implementation Plans

Michael R. Hoyles, Alaska Department of Environmental Conservation, Anchorage

A method is presented for demonstrating the effect of transportation control strategies and the degree of control needed to attain national ambient air quality standards by 1982, as mandated by the Clean Air Act Amendments of 1977. The emphasis is on the use of emission-concentration relationships to predict the average, rather than the maximum, concentration expected at a particular location in a given year. The statistical relation between the average concentration and the maximum is used to predict the corresponding maximum. This relation is calibrated by using all of the monitored data in the region of interest. The limitations associated with simulation modeling are minimized and the method is applicable to transportation planning methods. Carbon monoxide data collected for 8 h in the Seattle-Puget Sound region of Washington is used for illustration. The method may be applicable to oxidant control strategies.

The Clean Air Act Amendments of 1977 mandated revisions to the state implementation plans (SIP) and placed the burden for preparation of these at the local level. Although desirable on political and operational grounds, it has placed a heavy burden on the municipal planners because air quality expertise is not readily available to them. Nowhere is this more apparent than in the area of conversion of estimated emissions to concentrations. This is usually done by simulation modeling. Assuming that the models consider enough variables to predict air pollution concentrations accurately and are understood well enough to use them properly, there is still the problem of knowing what assumptions should be made in using them. This problem is often circumvented by the use of what has been termed worst-case methodology. This essentially means maximization of the model—all variables in the numerator are maximized and those in the denominator are minimized. The result is a model that overpredicts; however, this is resolved by comparison of modeled values to the monitored data to determine a calibration factor. The worst-case approach to modeling is not suited for use in developing control strategies because it tends to portray conditions to be much worse than they really are; however, selecting the proper conditions that will provide the best estimates is a challenge.

Models that estimate emissions rather than concentrations avoid this problem; however, the results still have to be converted to concentrations. A method is presented that allows one to make this comparison with a high degree of accuracy. The predictions are area specific and amenable to use in the development of SIP strategies. This model can be used to supplement existing methodology or can be used exclusively. It is inexpensive to use and is well suited to planning methods. The use of the model is demonstrated by using carbon monoxide (CO) data collected in the Puget Sound region of Washington in 1977.

MODEL DESCRIPTION

The model uses statistics to determine the maximum concentration once an average concentration has been predicted. Statistically the maximum concentration that occurs during a sampling period is equal to the mean of the data plus some number of standard deviations from the mean. Larsen (1) suggested that a log-normal distribution fits air pollution data better than does a normal one and, assuming this, one can express this relation as follows:

$$C_{max} = Mg \times Sg^{n} \quad (1)$$

where

C_{max} = the log of the maximum,
Mg = the geometric mean,
Sg = the geometric standard deviation, and
n = the number of standard deviations from the mean associated with the probability of the maximum occurring.

Larsen suggested that 3.81 be used for n in Equation 1 by assuming the maximum to be the highest value out of 8760 (the number of hours in a year) or a probability of 1/8760. Equation 1 can be rewritten as

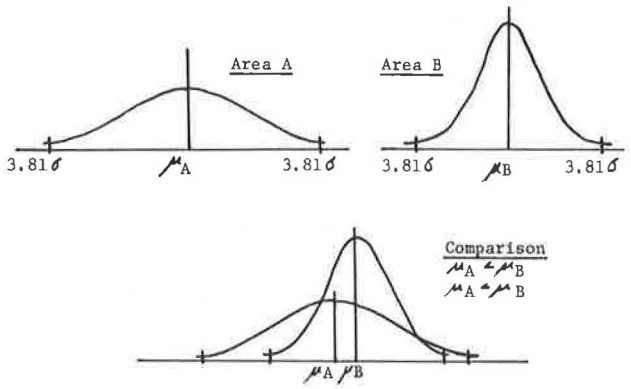
$$C_{max} = Mg \times Sg^{3.81} \quad (2)$$

As shown, the maximum concentration that occurs in an area is related to both the mean of the data and its standard deviation. The importance of considering Sg is due to the fact that it is a predictor of the probability that an extreme value will occur. This can be illustrated by comparing two hypothetical distributions, as shown in Figure 1. Area A has a flatter curve and, hence, a larger standard deviation, which indicates higher variability in the data and a greater chance of a larger extreme occurring. The two areas are superimposed, assuming area A to be less polluted than area B, as indicated by the bulk (and the mean) of the data being to the left of that in area B. A portion of area A can be seen that extends to the right of area B, which indicates that the probability is higher that area A will experience a larger maximum than will area B. An additional reason for considering the standard deviation is seen in Equation 1, where the mean is directly proportional to the maximum, whereas the effect of the standard deviation is nearly to the fourth power.

MODEL CALIBRATION

The statistical relationship depicted in Equation 1 is calibrated by using Mg and Sg of the data to calculate a C_{max}, which is then compared to the actual monitored maximum. One could use the second highest maximum, but here the extreme value is used for conservative reasons. This comparison is made for a large number of data sets, and the resultant data pairs (monitored and calculated maximums for each data set) can be compared by regression to produce the final empirically calibrated equation. In the study reported here, the data were collected for a six-week period in 12 separate metropolitan areas in the Puget Sound region. These 12 areas were chosen because they are potential hot spots, as indicated by high motor vehicle traffic density.

Figure 1. Comparison of CO distributions of two hypothetical areas.



A more complete description of the study is found elsewhere (2). The data are 8-h averages and the resultant equation will be used to calculate similar averaging time maximums.

Mg can be used directly or it can be calibrated. The best approach is dictated by which provides the best estimate in the final equation. Calibration has the advantage of accounting for the skewness of the data. To calibrate, the data are grouped and ranked in ascending order and the percentage of the time a data value group was exceeded is calculated. The data values and percentiles are made linear and a least-squares line is determined. The logarithm of the concentrations provides the necessary adjustment for these data. The percentile transformation is more involved but easily done. The position of the percentiles are symmetrical about the center of the scale and equal to their corresponding number of standard deviations from the mean. Any set of tables that gives the solution to the normal probability function can be used to provide this transformation. A final adjustment is made by converting everything to one linear scale rather than two that emanate from the center. This process can be illustrated by referring to the bell-shaped curves in Figure 1. The 3.81 standard deviations are 3.81 linear units from the center. An adjustment would produce a scale that is 7.62 linear units long and the mean or 50th percentile would be 3.81 linear units from the left origin. Similarly, the probability associated with the 3.81 units on the left would now be assigned the value of 0 and that associated with the 3.81 units on the right would be assigned the value of 7.62. This transformation is made for each percentile and the regression is done. Mg is the 50th percentile concentration and Sg is determined by using the concentration associated with the 15.87th percentile (16P) or 1.00 units to the left of the center (this is 2.81 units from the origin) and dividing it by Mg.

For each data set, a calculated and graphically determined Mg is found. A least-squares regression of these pairs provides the calibration of Mg. For the Seattle data, this was found to be

$$Mg = 0.15 + 1.01(50P) \quad (3)$$

Similarly, the calibration of Sg was accomplished and the resultant equation was found to be

$$Sg = -0.27 + 1.25(16P/50P) \quad (4)$$

The correlation coefficieints (r) for the two were 0.99 and 0.96, respectively. Substitution of these into Equation 1 provides the following equation:

$$C_{max} = [0.15 + 1.01(50P)]\ [-0.27 + 1.25(16P/50P)]^n \quad (5)$$

The best value of n is determined by arbitrarily assigning it a value and using Equation 5 to calculate a maximum for each data set, which is compared to the actual monitored maximum. Linear regression of the data pairs provides a correlation coefficient. The process is repeated with another value of n until the best fit is found. Figure 2 illustrates this process. For the Seattle data, the best value of n was 2.00.

Once the best n has been found, the corresponding regression slope and intercept become part of the final equation. For these data, the final equation was

$$C_{max} = 1.686 + 0.922\{[0.15 + 1.01(50P)] \ldots [-0.27 + 1.25(16P/50P)]^{2.00}\} \quad (6)$$

The correlation coefficient of this relationship was found to be 0.93 and the comparison between maximums predicted by this equation and those actually occurring is depicted in Figure 3. (These models were designed for U.S. customary units only; therefore, values in figures are not given in SI units.)

The large difference between the exponent of 2.00 and that proposed by Larsen may seem striking; however, there are two possible explanations of this. First, the data were collected for only six weeks (approximately 900 h) rather than a full year (8760 h). This probability would dictate an exponent of 3.0, which is midway between the two extremes. There is also a stronger reason for this discrepancy, which is in the area of the independence of the data values from each other in the same data set. This is especially true in regard to 8-h averages. If one assumes that 2.00 is correct, it would correspond to 44 samples/year. This compares favorably to the 50 samples/year indicated by Neustadter (3) and is certainly closer to the one-month study recommended by Meisel (4). A larger sample size would only increase the precision of the statistical parameters without appreciably changing their values.

MODEL UTILIZATION

Once the statistical relationship between major parameters of monitored data has been calibrated, it is used for CO by assuming

1. The average concentration that occurs for a pollutant in a given area is directly related to the average emissions of the pollutant in the same area, and
2. The standard deviation of the occurrence of a pollutant at a given location will not change from year to year.

Comparison of average rather than extreme values emphasizes the central tendency of the data, which is easily determined and is the strongest indicator of the data. Also, it is statistically more valid to compare the means of data sets rather than other parameters. Once a new mean has been determined, the statistical distribution of those data (such as Equation 6) is used to calculate the extreme value.

The second assumption is based on the hypothesis that the shape of the distribution of a population is due to factors independent of source strength (such as location of monitoring site with respect to pollutant sources, demographic and meteorological characteristics of the area, and the nature of the pollutant). A change in source strength should not affect the shape of the distribution but only shift it up or down scale. This

Figure 2. Comparison of regression equations and correlation coefficients for different exponents.

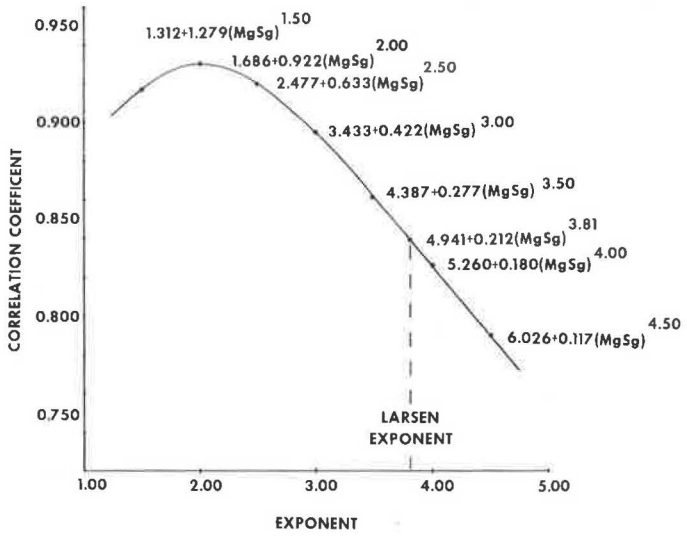


Figure 3. Comparison of calculated and monitored maximum CO 8-h averages.

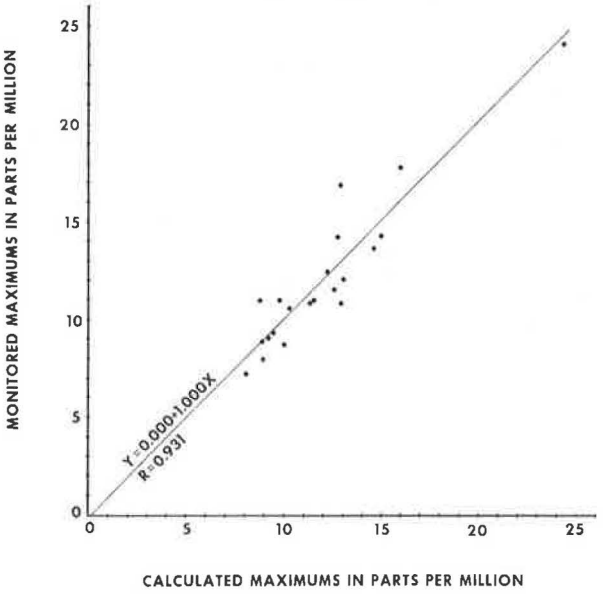
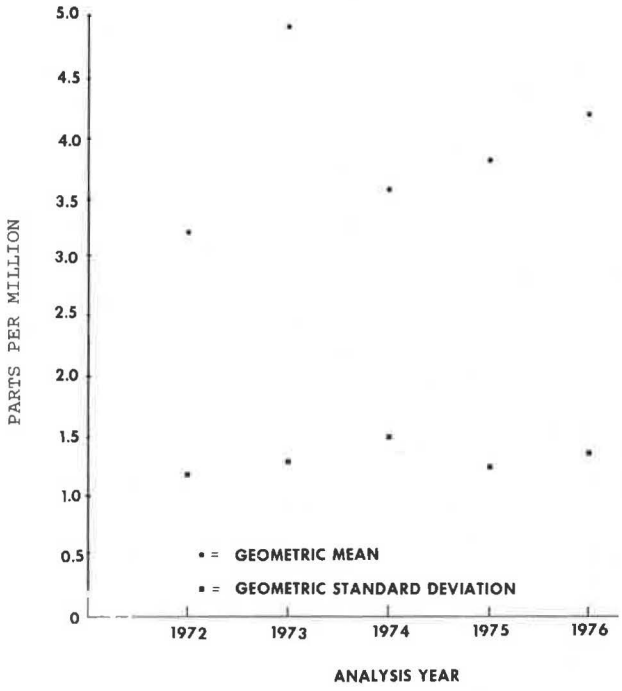


Figure 4. Trends in geometric mean and standard deviation for CO 8-h average concentrations.



hypothesis was tested for CO by analyzing data collected continuously for five years by the Washington State Department of Ecology at their James Street monitoring site in Seattle, Washington. Fifty data values were selected at random from each year of data and the mean and standard deviations were calculated and plotted (Figure 4) as a function of the year they represent. As can be seen, the standard deviation remains reasonably constant, but the means vary widely.

One should use monthly data to calibrate the model. All the data in the air shed should be used together for calibration. The only exception is that if different instrument methods of analysis are involved, it may be necessary to group the data accordingly; however, the data should at least be tested all together since the greater data size is more desirable even at the cost of some loss in precision. A month of data is used for each data set in order to consider seasonal variations and to allow more dynamic contingencies in the SIPs.

The calibrated relationship can be used to demonstrate the emission reduction needed to attain the standards and the influence of the strategies in the scenarios. A generalized form of Equation 6 will help in the discussion and is as follows:

$$C = ko + kl\ [Mg(Sg^n)] \quad (7)$$

Once the constants have been determined, the equation can be solved for Mg:

$$Mg = C - ko/[(kl)\ (Sg^n)] \quad (8)$$

By using the 8-h national ambient air quality standard (NAAQS) of 9 ppm and the standard deviation that was monitored in the location of interest, one can calculate a Mg that would correspond to attainment. Then by comparing this new mean to that monitored, one can determine the extent of the problem. In other words, rather

Table 1. Total CO.

District	1975 (kg/day)	1982 (kg/day)	Reduction from 1975 (%)	1990 (kg/day)	Reduction from 1975 (%)
Bellevue	23 406	13 956	40.4	5 346	77.2
Tacoma mall	15 382	9 114	40.8	3 695	76.0
Tacoma CBD	36 128	21 322	41.0	8 310	77.0
Everett	28 830	17 192	40.4	6 057	79.0
Renton	28 536	17 000	40.4	8 088	71.7
Tukwila	24 395	14 431	40.8	5 598	77.1
Duwamish	16 940	10 123	40.2	3 595	78.8
Montlake	6 995	4 165	40.5	1 411	79.8
University	26 760	15 990	40.3	6 060	77.4
Northgate	11 347	6 741	40.6	2 303	79.7
Seattle CBD	68 138	40 559	40.8	14 997	78.0
Bremerton CBD	24 210	-	-	4 853	80.0

Note: 1 kg = 2.2 lb.

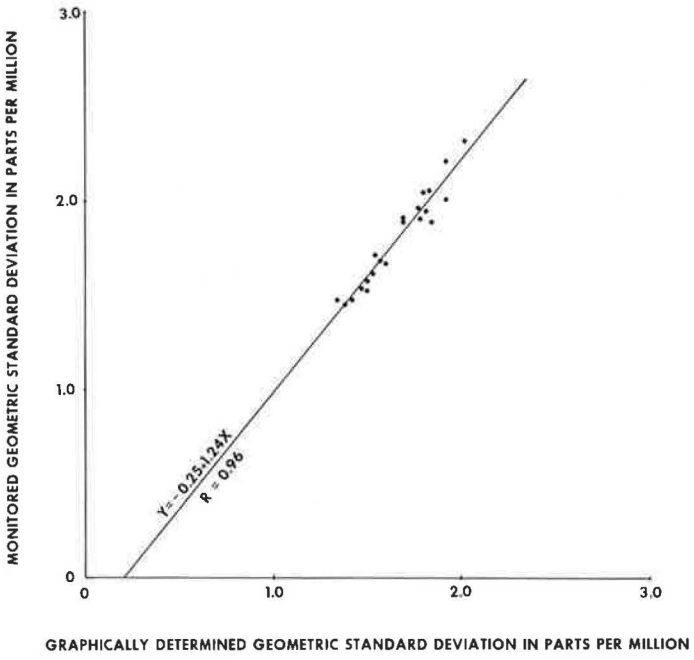


Figure 5. Comparison of graphically determined and monitored geometric standard deviation for CO 8-h averages.

than judging control strategies against the standards (maximum values), the criterion should be the mean, which is a much more behaved parameter and is statistically related to the maximum. Using the first assumption, the percentage decrease between the existing mean and that determined by Equation 8 is the percentage decrease needed in average emissions to attain the standards.

An analysis of the long-range element (LRE) transportation system for the base year and 1982 (by using the respective emission factors to determine the average emissions for each year) allows one to see the effect of the planned system without additional strategies. This was done for the Seattle data and is given in Table 1. The LRE projections for Bellevue, for example, show a decrease in average emissions of 40 percent between 1975 and 1982. (This would be the effect of no additional control strategies.) By using Equation 8 with the Bellevue data and the parameters in Equation 6, the decrease needed to attain the standards was determined to be 48 percent. This demonstrates that additional air pollution strategies are needed to reduce average emissions by 8 percent. Further, if a suggested strategy can be shown to reduce the average emissions by 4 percent, then one has demonstrated that more strategies are needed to achieve the remaining 4 percent reduction. One should remember that Sg from each data location can be used independently in Equation 8 to provide the corresponding Mg needed for that location. This allows one to establish strategies unique to each location. If analysis shows that a location cannot attain the standards by 1987, then this type of information would be a demonstration that the section 176 sanctions of the Clean Air Act Amendments of 1977 should be applied only to that part of the attainment area. Along these lines, the shrinking of the attainment area can be demonstrated on a yearly basis and, by using data that represent a month-long sampling period, the violation season can also be demonstrated to be shrinking. For example, in a given location, the violation season may be a six-month period. If attainment is expected in 1982, then, perhaps, by 1980 the violation season may be reduced to two months. If this can be demonstrated, then the more adverse strategies can be eased for the other four months. The advantages of this are readily apparent.

Another application of this method is in determination of the attainment date of the standards. Once a Mg has been determined by considering changes in average emissions, Equation 6 can be used to calculate the corresponding maximum. After this is done for several years, a plot can be made to depict the trend. This can be done for each location. Such a plot was prepared for the Seattle data and is shown in Figure 5.

Once strategies are considered, their effect on the attainment date can be illustrated by this method. This serves as a demonstration that the strategies will allow the standards to be attained.

OXIDANTS

This method may have an application in the area of photochemical oxidant control, where the average oxidant value is predicted by the usual methods. Once the statistical relation has been calibrated, then the predicted mean concentration is used to determine the maximum. There is a limitation in that the model allows a comparison between two variables rather than the large number needed to consider the complex interaction between pollulants and other conditions. A sophisticated method of grouping the data may resolve this problem.

This would not be a problem at all; one may be able to show a reasonable correlation between the change in average traffic emissions and the change in average oxidant concentrations to produce a new average oxidant concentration for 1982. Then one would proceed as above to determine the associated maximum. Certainly, this method could be used when simple one-to-one relationships have been found or a simple approach is being considered. For example, the well-known relationship betwen the 6:00-9:00 a.m. nonmethane hydrocarbon (NMHC) concentrations and the afternoon oxidant maximum may be even stronger if one correlates the average 6:00-9:00 a.m. NMHC value that occurs during a month with the average oxidant value from that same month. The correlation results from a large number of data sets, where each set represents one month of data. This relationship is used to determine the average oxidant expected due to a reduction in NMHC emissions.

CONCLUSION

This method requires a regional approach to analysis. This is due to the need for numerous data sets for calibration; however, this is not undesirable due to the nature of the transportation planning process, the nature of the air pollution problem, and the need for comprehensive strategies. Specifically, the transportation system is regionwide in approach and considers regional growth and development because air pollution characteristics of one area are not completely independent of those in another area in the same city or air shed. It is recommended that one use all of the data in the region for the calibration of the statistical relationship. By analyzing for each area, one can differentiate between them. For example, resources can be properly allocated, strategies can be formulated that do not inhibit adjacent areas, and existing problem-area boundaries can be redefined where and as needed to ensure the attainment of NAAQS without undue restriction on neighboring community growth and development.

REFERENCES

1. R. I. Larsen. A New Mathematical Model of Air Pollutant Concentration, Averaging Time, and Frequency. Presented at the 61st Annual Meeting of the Air Pollution Control Association, St. Paul, MN, APCA 68-147, June 23-27, 1968.
2. M. R. Hoyles. A Method for Determining the Attainment Date of the National Ambient Air Quality Standards. Presented at the 71st Annual Meeting of the Air Pollution Control Association, Houston, APCA 78-16.3, June 25-30, 1978.
3. H. E. Neustadter and S. M. Sidik. On Evaluating Compliance with Air Pollution Levels "Not To Be Exceeded More Than Once a Year." Journal of the Air Pollution Control Association, Vol. 24, No. 6, 1974, pp. 559-563.
4. W. S. Meisel. Monitoring Carbon Monoxide Concentrations in Urban Areas. NCHRP, Rept. 200, HR 20-14, March 1978.

Analysis of Air Quality Sensitivity to Development Pattern Changes and Growth Levels

George J. Scheuernstuhl and Jeffrey May, Denver Regional Council of Governments

To determine the impact of population and employment distribution changes as well as additional population and employment growth on air quality, regional vehicle kilometers of travel and emissions were obtained for four land-development-pattern scenarios for the year 2000. The scenarios include two activity-center scenarios, a dispersion scenario, and a centralization scenario. A fifth scenario was developed on the basis of additional growth beyond the forecast level. The distribution pattern of population and employment had little or no effect on ambient air quality; the alternative patterns showed little variation. The predicted differences in ambient air quality were less than the potential margin of modeling error. Carbon monoxide levels varied by only 6 percent. The centralization scenario produced the highest concentration, but one that is still 43 percent below the federal standard. Ozone predictions showed even less variation; the range was only 2 percent. Given a regional total, the distribution pattern appears to have little effect on regional air quality. A second result is that, although the absolute level of population has an effect on air pollution levels, these two variables are not directly proportional. A 27 percent increase in population resulted in a 16 percent increase in carbon monoxide emissions but only in an 8 percent increase in predicted maximum ambient concentrations. Large changes in population and employment produced smaller changes in ambient air quality.

Improved air quality is a goal in most metropolitan areas of the country. The Federal Highway Administration (FHWA) and Urban Mass Transportation Administration (UMTA) joint regulations and, more recently,