



# An Air Quality Data Analysis System for Interrelating Effects, Standards, and Needed Source Reductions: Part 6. Calculating Concentration Reductions Needed to Achieve the New National Ozone Standard

Ralph I. Larsen , Thomas C. Curran & William F. Hunt Jr.

To cite this article: Ralph I. Larsen , Thomas C. Curran & William F. Hunt Jr. (1980) An Air Quality Data Analysis System for Interrelating Effects, Standards, and Needed Source Reductions: Part 6. Calculating Concentration Reductions Needed to Achieve the New National Ozone Standard, Journal of the Air Pollution Control Association, 30:6, 662-669, DOI: [10.1080/00022470.1980.10465992](https://doi.org/10.1080/00022470.1980.10465992)

To link to this article: <https://doi.org/10.1080/00022470.1980.10465992>



Published online: 13 Mar 2012.



Submit your article to this journal [↗](#)



Article views: 48



View related articles [↗](#)

# An Air Quality Data Analysis System for Interrelating Effects, Standards, and Needed Source Reductions: Part 6. Calculating Concentration Reductions Needed to Achieve the New National Ozone Standard

Ralph I. Larsen, Thomas C. Curran, and William F. Hunt, Jr.  
U.S. Environmental Protection Agency

The new ozone National Ambient Air Quality Standard specifies that the expected number of days per calendar year that the 1 hour average ozone concentration can exceed 0.12 ppm must be equal to or less than 1. This paper describes a method to calculate design frequency, design concentration, and the percentage concentration reduction necessary to achieve this standard. The design frequency is once per year (1/365) if daily maximum hour ozone concentrations are available at a particular site for an equal number of days in summer (April through September) and winter (October through March). An equation is used to adjust design frequency as a function of the number of summer and winter samples available. The design concentration (the ambient concentration measured at the design frequency) needs to be reduced to 0.12 ppm. Graphical and digital methods for determining the design concentration are presented. Percentage concentration reductions needed to achieve the standard are calculated for each site which has ozone concentration data available in the National Aerometric Data Bank for at least half of the days in one summer of years 1975 through 1977. The degree of reduction calculated for the site with the highest concentrations in each county is indicated by shading on a map of the United States.

This paper describes a method which can be employed by hand or computer to calculate the percent concentration reduction needed to achieve the new ozone National Ambient Air Quality Standard (NAAQS).<sup>1</sup>

Design concentration is the ambient concentration that needs to be reduced to the standard. The percent concentration reduction is the reduction needed to accomplish this. This paper treats only concentration reductions, not emission reductions. States are to apply one of the four following techniques in determining the degree of hydrocarbon emission reduction required to attain the standard:<sup>2</sup> photochemical

dispersion models; empirical kinetics modeling (EKMA); empirical and statistical models; or proportional rollback (Ref. 3, Eq. 28).

In 1971, NAAQS were issued for suspended particulates and five gases<sup>4</sup> (carbon monoxide, hydrocarbons, nitrogen dioxide, photochemical oxidants, and sulfur dioxide) for short-term (1 to 24 hour) and long-term (1 year) averaging times. All of the short-term standards specify concentrations that are not to be exceeded more than once a year.

## The New Ozone Standard

The new ozone standard replaces the old photochemical oxidant standard. The new standard specifies that the "expected number of days per calendar year" that the 1 hr average ozone concentration can exceed 0.12 ppm (parts per million) must be "equal to or less than 1." This is a frequency of 1/365 (0.2740%), which will be called the design frequency. The design concentration used here is the ambient concentration that plots at this frequency on cumulative frequency distribution graph paper. The standard is achieved by reducing ozone precursor emissions until the design concentration equals the standard concentration.

The *Guideline for Interpretation of Ozone Air Quality Standards*<sup>5,6</sup> specifies that the last three calendar years of ozone air quality data will usually be analyzed to determine the design concentration. The *Guideline* presents four methods that can be used to determine a design concentration for a site:

1. Fitting one statistical distribution to several years of data;
2. Fitting a separate distribution for each year of data;
3. Using the empirical frequency distribution of several years of data;
4. Table look up (using the number of samples available to enter a table that shows which one of the top few measured concentrations to use as the design value).

The method presented here is of type 3. It is an easy to use method for determining a design concentration. Some limitations of this method will be discussed in a later section, "Variability in Calculated Design Concentrations."

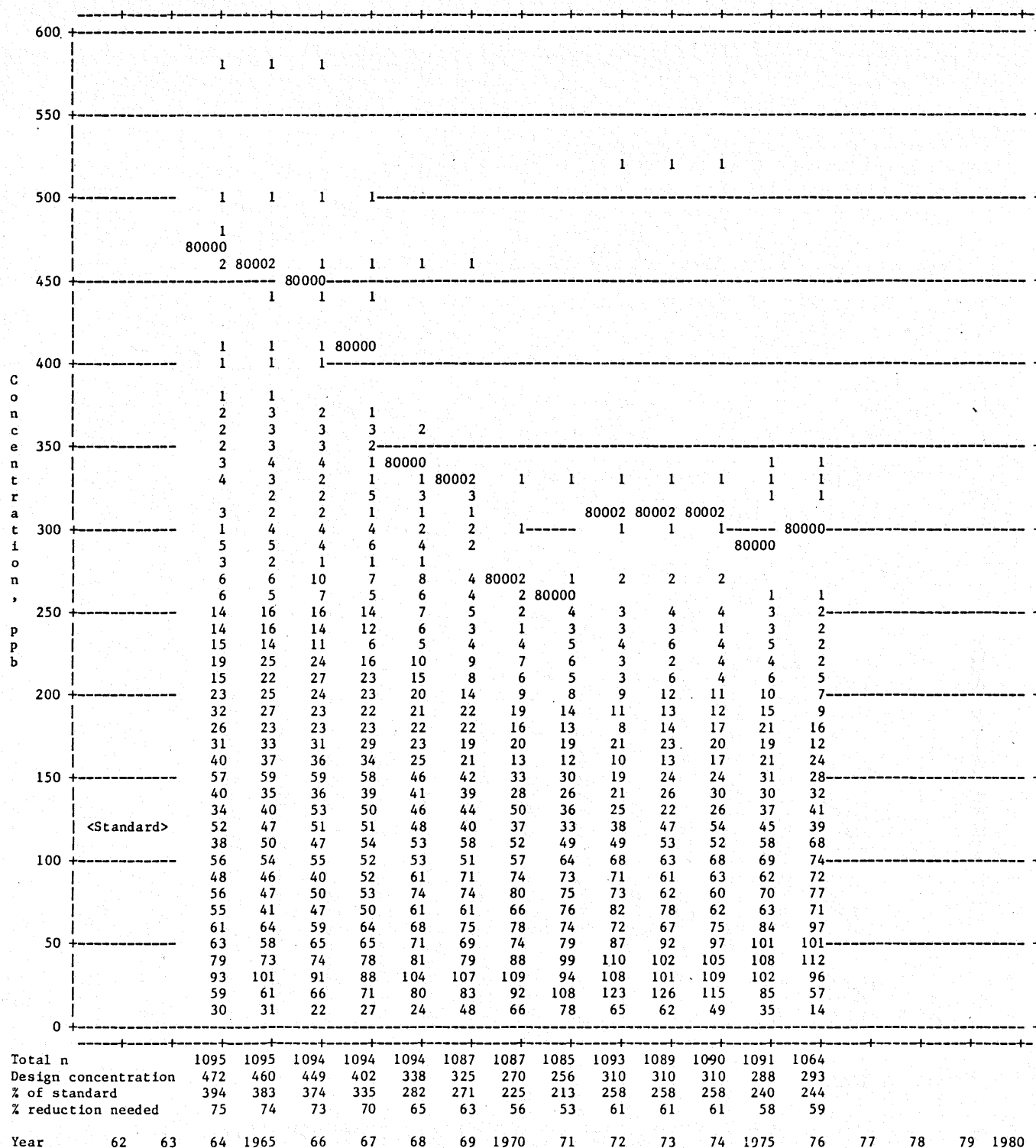


Figure 1. Number of daily maximum hours in various concentration classes for pollutant-method 4410114,<sup>7</sup> total oxidants, for each stated year plus preceding year and following year at site 054180001101,<sup>8</sup> 434 S. San Pedro Street, Los Angeles, CA. (The concentration class that contains the design concentration has been designated by adding 80000 to it.)

### Calculating Plotting Frequency

Total oxidant concentrations measured in downtown Los Angeles have been entered into the National Aerometric Data Bank (NADB) for almost every hour, starting with year 1963. Three-year running frequency distributions of the daily maximum hour concentrations for these data are shown in Figure 1. For years 1975 through 1977, 1064 (97%) of the possible 1096 values were available.

The plotting frequency for these data on cumulative frequency distribution graph paper can be calculated with the

following equation:<sup>9</sup>

$$f = (r - c) / (n - 2c + 1) \quad (1)$$

where

$f$  = the plotting frequency of a sample

$r$  = the rank of the sample, from highest to lowest concentration

$c$  = a constant

$n$  = the number of samples

The correct value of  $c$  depends on the type of distribution and  $n$ , but  $c = 0.5$  has been found generally acceptable for a wide variety of distributions<sup>9</sup> and is used here as a "first approximation," giving

$$f = (r - 0.5)/n \quad (2)$$

Eq. 2 has been used to calculate the plotting frequency of the median sample in each concentration class, as listed in Figure 2, where frequency is expressed as samples per million (one spm equals 0.0001%). For instance, the third highest con-

centration measured from 1975 through 1977, as indicated in Figure 1, is in the 320 ppb (parts per billion) concentration class (315–325 ppb). Its plotting frequency, at a plotting concentration of 320 ppb, calculates as

$$f = (3 - 0.5)/1064 \quad (3)$$

$$f = 0.002350$$

$$f = 0.2350\%$$

$$f = 2350 \text{ spm (as listed in Figure 2)}$$

The median sample in each concentration class in Figure

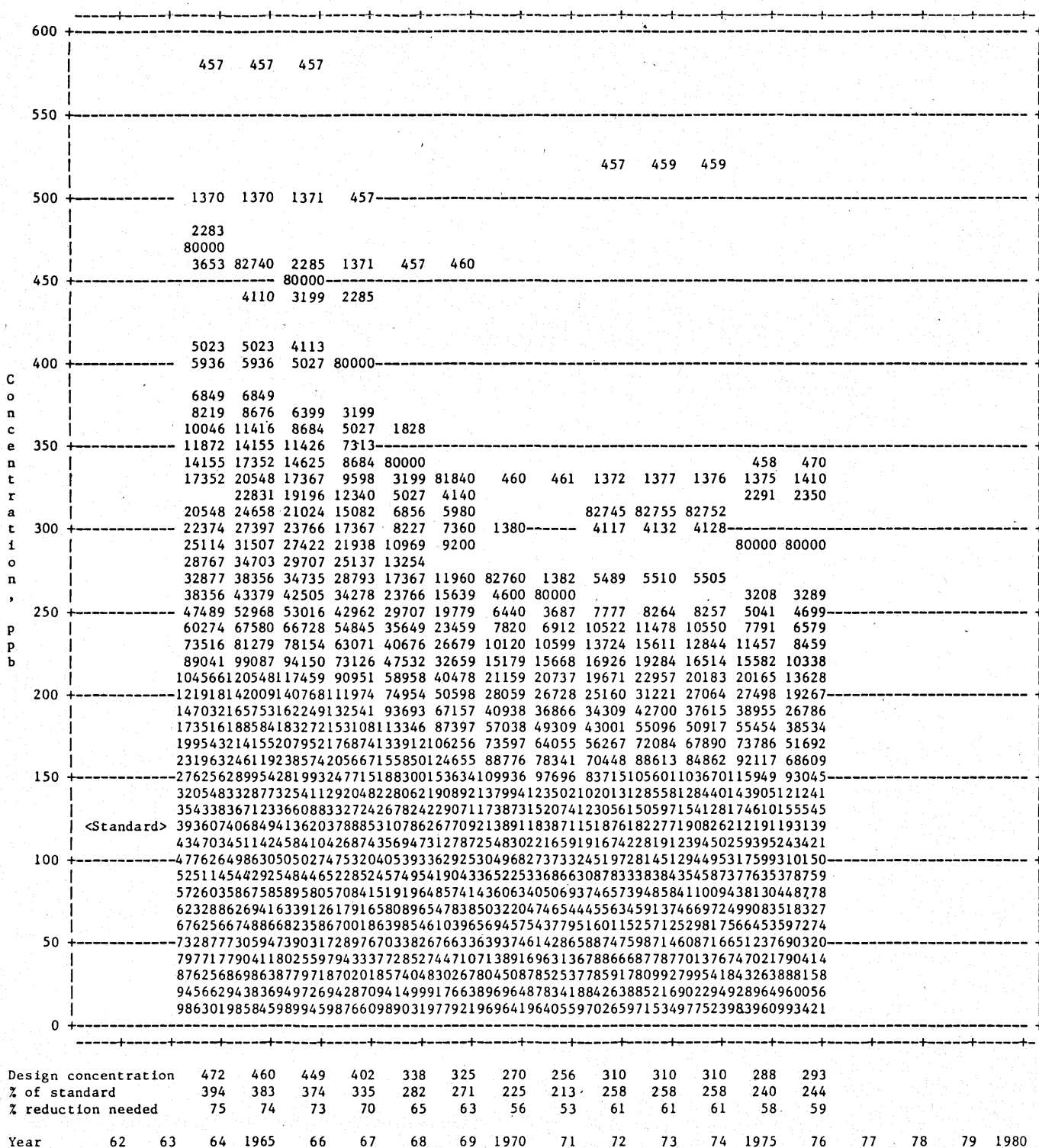


Figure 2. Cumulative plotting frequency (in samples per million) of median sample in each concentration class of total oxidants, pollutant-method 4410114, for each stated year plus preceding year and following year at site 05418000101, 434 S. San Pedro Street, Los Angeles, CA (The concentration class that contains the design concentration has been designated by adding 80000 to it.)

2 for years 1975 through 1977 has been plotted in Figure 3 on normal probability, log-probability, and semi-logarithmic graph papers. If an even number of samples is in a class, the point has been plotted at a rank midway between the two middle samples. Thus, the two samples in the 250 ppb concentration class are plotted as rank 5.5. The empirical frequency distribution for each plot has been described by joining the plotted points with straight lines.

If these data could be characterized as an independently identically distributed normal random variable, then the following equation (Ref. 10, Eq. 26), which closely approximates the mean position of ranked normal deviates (Ref. 11, p. 175), could be used to determine the plotting position of each sample in the top half of the distribution:

$$f = (r - 0.4)/n \quad (4)$$

When  $n$  is greater than 100 and  $c = 0.4$ , then Eq. 1 and 4 will both give almost the same value for  $f$ . Eq. 1 can be used for the entire distribution, not just the top half.

Neustadter and Sidik<sup>12</sup> used Monte Carlo simulation to calculate plotting positions for lognormally distributed data. They calculated the position for the second highest of 365 values. Expressing their results in terms of Eq. 1, they calculated  $c = 0.567$ . For normally distributed data, they obtained  $c = 0.440$ . The plot of the Los Angeles data is slightly concave upward on normal probability graph paper (Figure 3a) and slightly concave downward on log-probability graph paper (Figure 3b), suggesting that the true distribution might lie between the two. A value of  $c = 0.5$  is about midway between the two, and Eq. 2 is therefore used here. The "first approximation" of  $c = 0.5$  used above thus becomes the final value used throughout this paper. The accuracy obtained by using Eq. 2 depends on how far the data depart from being an independently identically distributed random variable. This point will be discussed later.

### Calculating Design Concentration

Figure 3c is plotted on semi-logarithmic graph paper<sup>13</sup> as are most of the figures in the *Guideline*. The line connecting the third and fourth samples intersects the design frequency of 1/365 (2740 spm). The design concentration can be calculated by taking the equation of this line and rearranging it to solve for the design concentration at the line's intersection with the design frequency:

$$c_d = c_1 + (\ln f_d - \ln f_1)(c_2 - c_1)/(\ln f_2 - \ln f_1) \quad (5)$$

where

$c_d$  = design concentration

$c_1$  = concentration of the highest of the two points

$c_2$  = concentration of the lowest of the two points

$\ln$  = natural logarithms

$f_d$  = design frequency (expressed below as samples per million)

$f_1$  = plotting frequency for point 1

$f_2$  = plotting frequency for point 2

For the example,

$$c_d = 320 + (\ln 2740 - \ln 2350)(260 - 320)/(\ln 3289 - \ln 2350)$$

$$c_d = 292.6 \text{ ppb}$$

The design concentration has been highlighted in Figures 1 and 2 by adding 80,000 to the concentration class that contains it.

Since the line used to determine the design concentration is so short, the type of distribution selected to describe this line is of minor importance. Eq. 5 was chosen because it is the easiest of the three to calculate. The line connecting points 3 and 4 on normal probability graph paper (Figure 3a) intersects the design frequency at a calculated concentration of 292.8 ppb. Calculations for the lognormal plot (Figure 3b) give

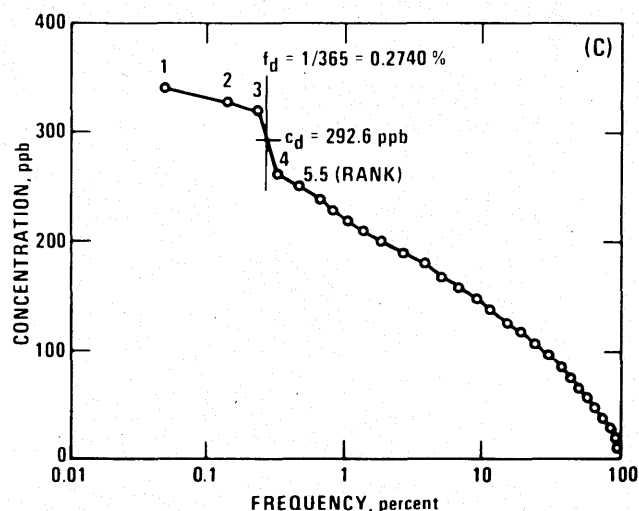
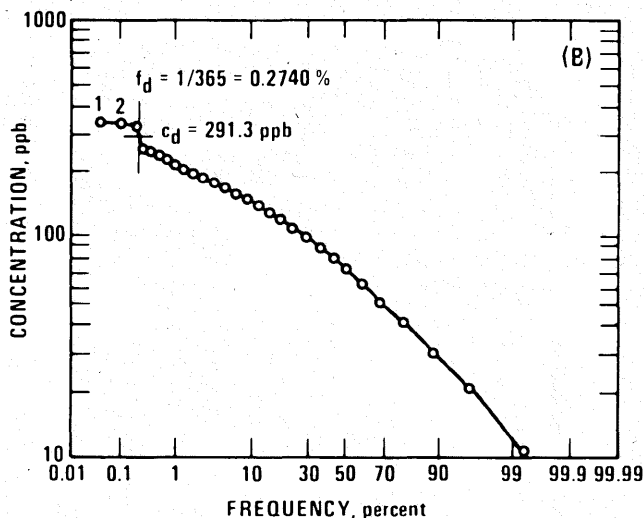
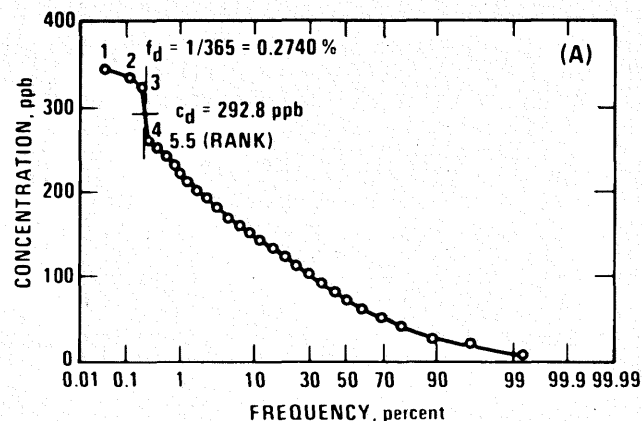


Figure 3. Frequency of daily maximum hour total oxidant concentrations equal to or greater than stated concentration at site 054180001101, 434 S. San Pedro Street, Los Angeles, CA, for years 1975 through 1977, plotted on (A) normal probability, (B) log probability, and (C) semi-logarithmic graph papers.

291.3 ppb. Differences in the values calculated by these three methods are insignificant, and this is for a very steep line. Lines with shallower slopes would show even smaller differences.

The 293 ppb design concentration is 244% of the 120 ppb standard. A 59% reduction in the design concentration would thus be needed to lower it to the standard. This information appears near the bottom of Figures 1 and 2.

Similar calculations have been made for all 3-yr periods, beginning with years 1963 through 1965. Note that the four top concentrations in each column of Figure 1 changed in pattern over the past 5 yr but that the design concentration remained fairly constant, varying only from 288 ppb to 310 ppb. The concentration reduction needed to achieve the standard varied only from 58 to 61%.

### Seasonal Variation in Ozone Concentrations

In the above example only the four highest daily maxima measured in a 3-yr period were used to determine the design concentration. At some sites, the four highest daily maxima in one of three years may be higher than all of the daily maxima for the other two years. This data analysis system focuses on concentrations measured during months when the four highest daily maxima for the year are most likely to occur.

The daily maximum ozone concentrations measured in 1976 at a warm winter site in Florida were distributed by concentration and month (Figure 4). For computer programming ease, the days per month were divided (only in these monthly distribution figures) into the following regular pattern: 30, 31, 30, 31, 30, 31, 30, 31, 30, 31, 30, and 31 for a December leap year (otherwise 30). Note that the four highest daily maxima occurred during the second and third quarters of the year (April through September). The highest hour concentration measured in the highest month was about twice the highest hour measured in the lowest month (December was excluded since it had so few samples).

The monthly distribution pattern at a Rochester, New York, cold winter sampling site (Figure 5) was similar to that for the Florida site but demonstrated greater differences between summer (defined here as April through September) and winter (October through March). The highest hour in the year was approximately four times higher than the highest measured in the lowest month. The important fact, though, is that the four highest daily maxima all occurred in summer. This does not imply that the standard level is not exceeded during the winter but that the higher concentrations that determine the design value usually occur in the summer. The applicability of this assumption may vary in some areas but can be verified on a site specific basis.

Samples were available throughout the year for the Los Angeles example. The design frequency was 1/365. If ozone is sampled only during the 183 days of summer and the four highest samples of each year are expected to occur only in summer, then the design frequency is 1/183. If samples are taken diligently during summer and either sparsely or diligently in winter, then the design frequency will vary between 1/365 and 1/183 and is approximated by the following equation:

$$f_d = 2n_s/365n_t \quad (6)$$

where

$f_d$  = design frequency  
 $n_s$  = number of summer samples  
 $n_t$  = total number of samples

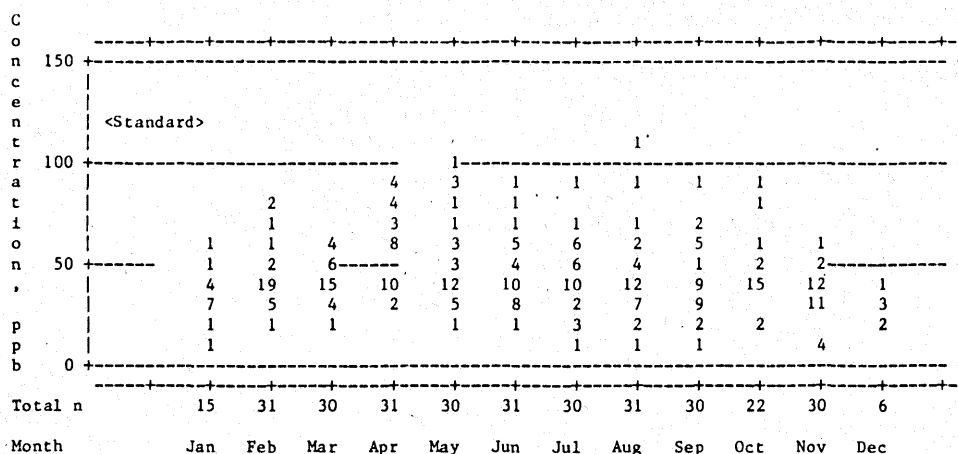


Figure 4. Number of daily maximum hours in various concentration classes for ozone, pollutant-method 4420111, by month of year 1976 at site 104360022G02, 111 N. Dale Harry Hwy, Tampa, FL.

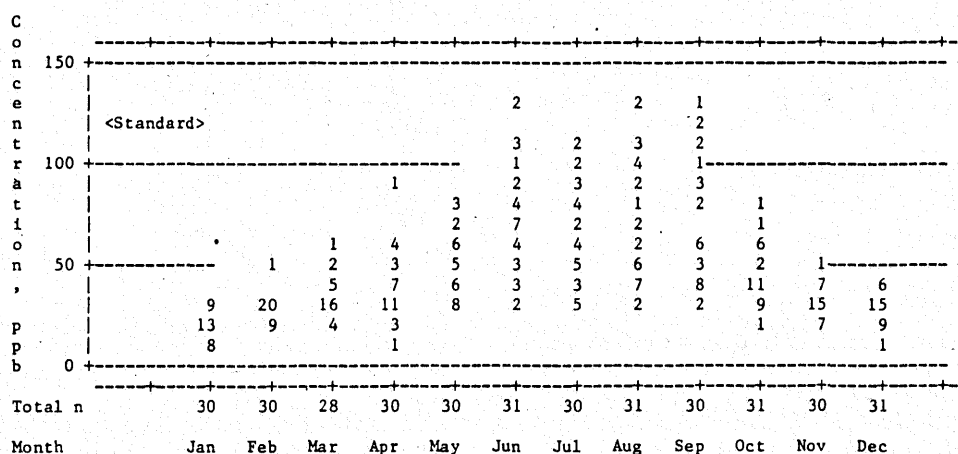


Figure 5. Number of daily maximum hours in various concentration classes for ozone, pollutant-method 4420111, by month of year 1976 at site 335760004F01, Rochester, NY.

Data for years having samples for less than half the summer days were not processed since that few samples may not have adequately characterized air quality.

A total of 1039 samples, including 526 summer samples, are available in the NADB for years 1975 through 1977 for the Rochester site (Figure 6c). Therefore, the design frequency (Eq. 6) is 0.2774%.

The computer program interpolates between two adjacent points to calculate the design concentration expected at the design frequency lying between them (Eq. 5). If the design frequency is lower than that for the highest concentration class, as it just barely is in this case, then the program extrapolates by using data for the two highest concentration classes. Extrapolation is needed infrequently. It was needed here because six samples occurred in the top concentration class and the plotting rank was thus 3.5 (Eq. 2). Only one or two samples are usually in the top concentration class. This site was selected because of its need to use extrapolation instead of interpolation. Even so, the calculated design concentration equals that of the class containing these six samples, 130 ppb, 109% of the standard. A concentration reduction of 8% is needed to lower this design concentration to the standard.

is used in this method, greater year to year variations in calculated design concentrations are expected than would be expected if all of the data for 3-yr periods were fitted to the best fitting of several candidate distributions.<sup>14</sup> This expected greater variability should be kept in mind when using this approach.

As mentioned earlier, the accuracy of Eq. 2 depends on how far the data depart from being an independently identically distributed random variable. Horowitz and Barakat have treated this problem.<sup>15</sup> They have used example ozone data for year 1976 measured during the Regional Air Pollution Study (RAPS) at Station 109 near East St. Louis, IL. They showed that these data are highly autocorrelated and thus not independent but that this high autocorrelation has little effect on the calculated expected annual maximum concentration.

The monthly distribution of ozone concentrations at Station 109 looks much like Figure 5. The mean and spread of concentrations in the summer varies markedly from that in the winter. Therefore, the data are not identically distributed. Horowitz and Barakat found that this non-stationarity could have an important effect on fitting the data to a distribution

Concentration Class	150																				
		80000					80000					80006					82887				
		5 1					6906 1389														
		<Std> 80002 2 80000					83155 16575 80000					4					7700				
		7 10 4					17350 33149 8333					21					19731				
	100	3 8 6					33123 58011 22222					17					38017				
		8 11 10					50473 84254 44444					29					60154				
		10 15 13					78864120166 76389					38					92397				
		7 14 18					105678160221119444					39					129451				
		32 33 38					167192225138197222					103					197786				
Plotting Rank	50	29 31 49					263407313536318056					109					299808				
		66 63 77					413249443370493056					206					451396				
		65 114 83					619874687845715278					262					676612				
		51 46 38					802839908840883333					135					867661				
		34 10 21					936909986188965278					65					963908				
	0	3 2					995268 997222					5					997594				
Total n		317	362	360											1039						
Summer n		165	182	179											526						
Design conc		121	140	122											130						
% of standard		100	117	102											109						
% reduct needed		0	15	2											8						
Year		1974	1975	1976	1977	1978	1974	1975	1976	1977	1978	1974	1975	1976	1977	1978	1974	1975	1976	1977	1978
		a					b					c					d				

Figure 6. Number of daily maximum hours in each concentration class (a) for each year (c) plus preceding year and following year, and cumulative plotting frequency (in samples per million) of median sample in each concentration class (b) for each year (d) plus preceding year and following year, for pollutant-method 4420111, ozone, at site 335760004F01, Rochester, NY. (The concentration class that contains the design concentration has been designated by adding 80000 to it.)

## Comparison of Daily Maximum and All Hours Standards

The contrast between a daily maximum hour standard and an all hours standard can be seen by comparing Figures 1 and 7. Five hours had concentrations above 330 ppb for years 1972 through 1974 at this Los Angeles site. All of these hours occurred on one day in 1973. For an all hours standard and a design frequency of 1/8760, this one particular day would have raised the required concentration reduction from 61% to 72%. This example illustrates the greater stability of the daily standard. Studies indicate<sup>1</sup> that a daily standard affords adequate protection. Therefore, other factors, such as stability, were considered in choosing a daily form for the standard.

## Variability in Calculated Design Concentrations

As was mentioned at the beginning, the *Guideline* discusses four methods that can be used to determine a design concentration for a site. The method presented here has been to use the empirical frequency distribution of 3-yr periods of data. Since only the high concentration tail of the distribution

and calculating the expected maximum. They developed a method to correct for non-stationarity and calculated an expected annual maximum concentration for this site of 194 ppb.

They also said that the effect of non-stationarity could be minimized by analyzing data for only a part of the year when the data were more nearly identically distributed, such as for a few summer months when the maximum was expected to occur. When the data for either all of 1976 or only the summer months (April through September, as used here) were processed by the computer program that produced Figure 6, a design concentration of 205 ppb was calculated. These results are identical because only the highest samples in the upper tail of the distribution were used. This 205 ppb is close to the expected annual maximum concentration of 194 ppb calculated by Horowitz and Barakat. Thus, for this particular case, their expected maximum and the calculated design concentration are similar. Additional studies at other sites could give more comprehensive results of how well the two methods agree.

## Concentration Reductions Needed to Achieve the New Ozone Standard

The degree of concentration reduction needed to achieve the new ozone standard in various areas of the United States is shown in Figure 8. All of the ozone data in the NADB for years 1975 through 1977 were processed for each site in the same manner as for Figure 6. The degree of concentration reduction for the site with the highest concentrations in each county has been plotted in Figure 8.

Precursors emitted from an urban area often produce ozone concentrations that are highest many miles downwind from

that area, sometimes in a county other than the source county. For instance emissions from New York City are probably an important factor in the highest ozone concentrations shown in Figure 8 for Connecticut.<sup>16</sup>

Some of the values indicated by Figure 8 will vary from those calculated by a different method and used by a State in their implementation plan. Also, as would be expected for this large a data base, some errors, such as a misplaced decimal place resulting in an ozone concentration of 2.4 ppm, are contained in the NADB. The FORTRAN computer programs employed have overlooked all concentrations above 0.9 ppm. The effects of some lower, but still erroneously high values

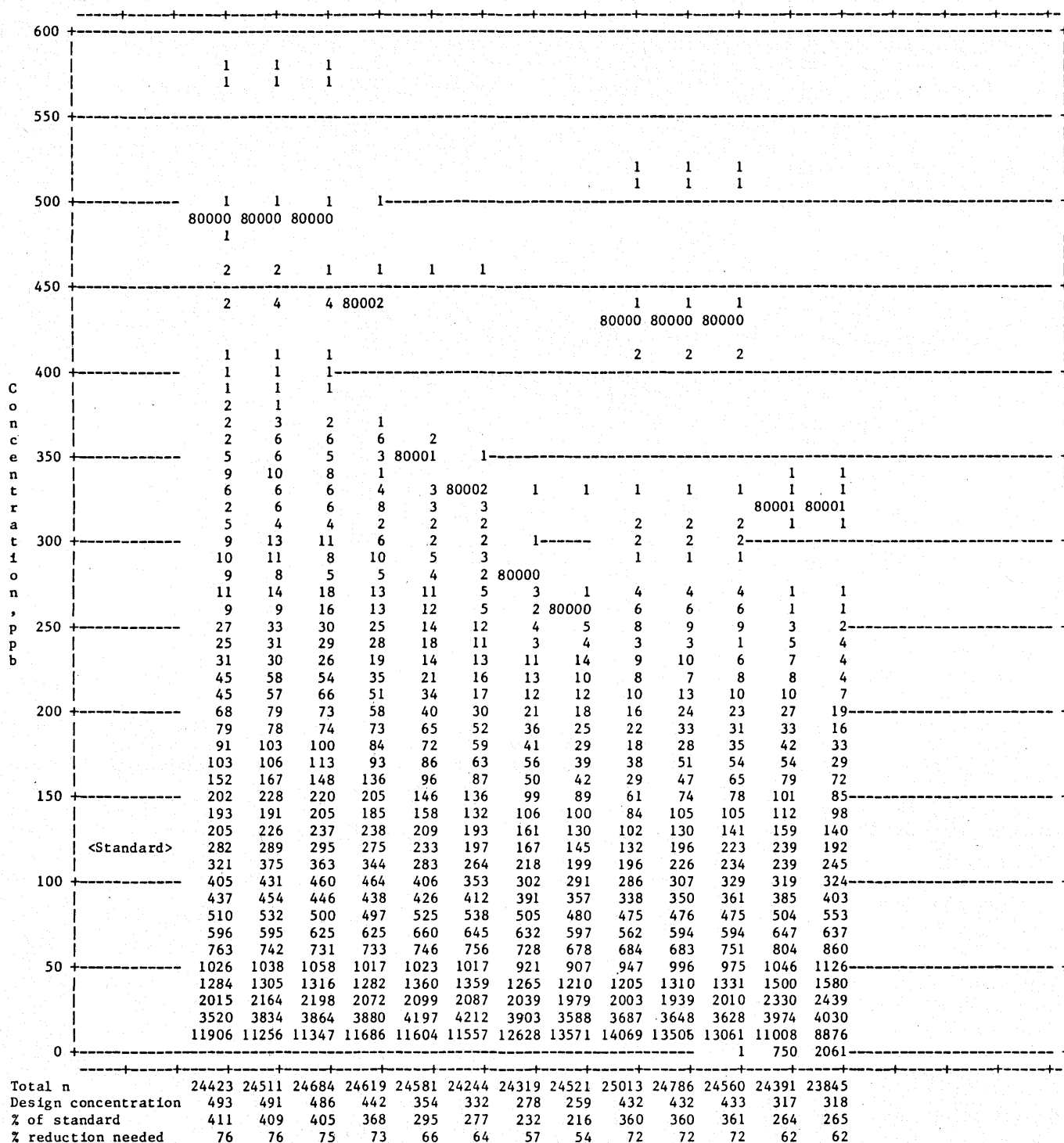


Figure 7. Number of hours in various concentration classes for pollutant-method 4410114, total oxidants, for each stated year plus preceding year and following year at site 054180001101, 434 S. San Pedro Street, Los Angeles, CA. (The concentration class that contains the design concentration has been designated by adding 80000 to it.)



may still show in Figure 8. The emphasis here is on the method, not the end results. The first author will be glad to provide any state air pollution control agency with Figure 6-type plots for any or all of their ozone sites so that erroneous data can be spotted and corrected.

## Summary

The new ozone National Ambient Air Quality Standard specifies that the expected number of days per calendar year that the 1 hour average ozone concentration can exceed 0.12 ppm must be equal to or less than 1. This paper describes a quick method to calculate design frequency, design concentration, and the percentage concentration reduction necessary to achieve this standard. The method adjusts for the relative frequency of summer (April through September) and winter (October through March) sampling at each site. The number of summer samples and the total number of samples available for a particular site are entered into Eq. 6 to calculate design frequency. The cumulative frequency distribution for each of the highest ozone concentration classes for a site is then plotted on semi-logarithmic graph paper by using Eq. 2 to determine plotting frequency. The plotted points are connected with straight lines. The design concentration is the point where one of the lines intersects the design frequency. Instead of using this graphical approach, Eq. 5 can be used to calculate the design concentration. Example calculations are shown for two air sampling sites. Concentration reductions required by the new ozone standard are calculated for each site in the National Aerometric Data Bank having data available for at least half of the days in one summer of years 1975 through 1977. The degree of reduction calculated for the site with the highest concentrations in each county is indicated by shading on a map of the United States.

## References

1. "National primary and secondary ambient air quality standards; revisions to the National Ambient Air Quality Standards for photochemical oxidants," *Federal Register* 44: 8202 (Feb. 8,

- 1979).
2. "Preparation, adoption, and submittal of implementation plans; revisions to implementation procedures related to photochemical oxidants (ozone)," *Federal Register* 44: 8234 (Feb. 8, 1979).
3. R. I. Larsen, "A new mathematical model of air pollutant concentration, averaging time, and frequency," *JAPCA* 19: 24 (1969).
4. "National primary and secondary ambient air quality standards," *Federal Register* 36: 8186 (April 30, 1971).
5. *Guideline for Interpretation of Ozone Air Quality Standards*, EPA-450/4-79-003, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1979.
6. T. C. Curran and W. M. Cox, "Data analysis procedures for the ozone NAAQS statistical format," *JAPCA* 29: 532 (1979).
7. *AEROS Manual Series Vol. V: AEROS Manual of Codes*, EPA-450/2-76-005, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1976.
8. *Directory of Air Quality Monitoring Sites Active in 1977*, EPA-450/2-78-048, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1978.
9. G. J. Hahn and S. S. Shapiro, *Statistical Models in Engineering*, John Wiley and Sons, New York, 1967, p. 294.
10. R. I. Larsen, *A Mathematical Model for Relating Air Quality Measurements to Air Quality Standards*, AP-89, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1971.
11. E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians, Vol. 1*, Cambridge University Press, New York, 1962.
12. H. E. Neustadter and S. M. Sidik, "On evaluating compliance with air pollution levels 'not to be exceeded more than once a year,'" *JAPCA* 24: 559 (1974).
13. T. C. Curran and N. H. Frank, "Assessing the Validity of the Lognormal Model when Predicting Maximum Air Pollution Concentrations," Paper No. 75-51.3, 68th APCA Annual Meeting, 1975.
14. T. Fitz-Simons and D. M. Holland, *The Maximum Likelihood Approach to Probabilistic Modeling of Air Quality Data*, EPA-600/4-79-044, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1979.
15. J. Horowitz and S. Barakat, "Statistical analysis of the maximum concentration of an air pollutant: effects of autocorrelation and non-stationarity," *Atmos. Environ.* 13: 811 (1979).
16. R. A. Rubino, L. Bruckman, and J. Magyar, "Ozone transport," *JAPCA* 26: 972 (1976).

The authors are all with the U.S. Environmental Protection Agency, Environmental Sciences Research Laboratory, Research Triangle Park, NC 27711.

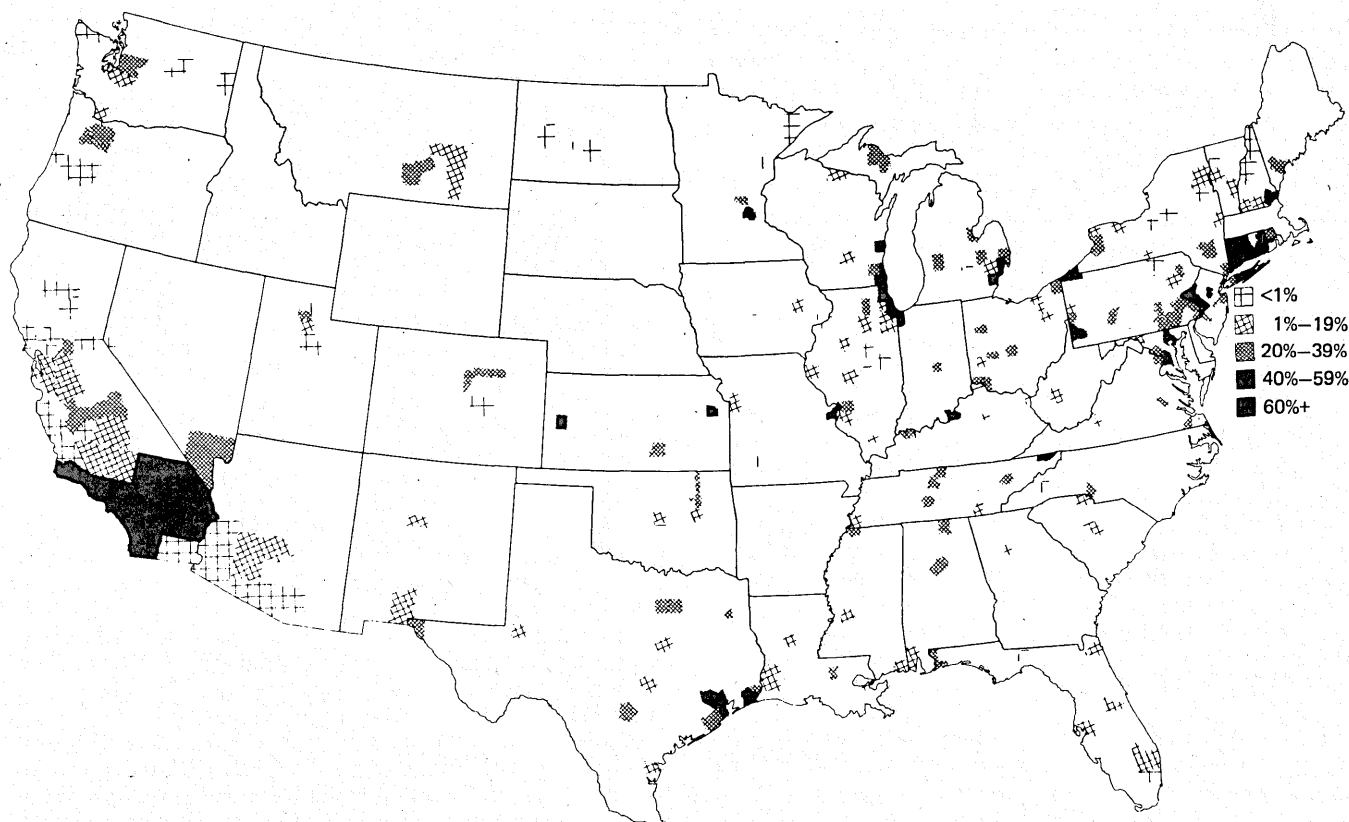


Figure 8. Percent concentration reduction needed to achieve ozone standard.