Final Report

REGRESSION MODELING OF OXYFUEL EFFECTS ON AMBIENT CO CONCENTRATIONS

SYSAPP-96/78

January 8, 1997

Prepared for

Renewable Fuels Association
1 Massachusetts Avenue, N.W.
Washington, DC 20001

and

Oxygenated Fuels Association
1300 N. 17th Street, Suite 1850
Arlington, VA 22209

Prepared by

Gary Z. Whitten
Jonathan P. Cohen
Alex M. Kuklin

Systems Applications International, Inc.
101 Lucas Valley Road
San Rafael, California 94903
510-507-7100

REGRESSION MODELING OF OXYFUEL EFFECTS ON AMBIENT CO CONCENTRATIONS

SUMMARY

There has been a substantial interest in the impact of the oxyfuel program on ambient concentrations of carbon monoxide (CO). This study reports a substantial (14 percent reduction) and statistically significant (± 4 percent with 95% confidence) association between the use of oxyfuels and monitored CO concentrations. Starting with the available large EPA database of hourly observations of CO concentrations, regression modeling techniques were used here to investigate the relationship between the winter use of oxyfuels and winter ambient carbon monoxide (CO) concentrations at over 300 monitoring sites which were designated as oxyfuel or non-oxyfuel sites by the Environmental Protection Agency (EPA) and also met appropriate data completeness criteria.

This investigation builds on the EPA's recent statistical analysis of quarterly CO averages reported by Cook, Enns, and Sklar (1996). Other review of the EPA analysis have recommended a focus on December–January data rather than quarterly data and that further analysis using data from individual monitoring sites should improve the estimate of oxyfuel impacts on ambient CO concentrations. Based on these suggestions, we evaluated a set of 53 alternative regression models fitted to the bimonthly (December–January for winter and June–July for summer) average hourly CO and average daily eight-hour maximum CO. The models differ in how they account for variations in CO concentration by season, by site, and by oxyfuel status.

The models which include a universal downward trend augmented by oxyfuels at appropriate sites and include a local starting concentration specific to each site tend to best explain the database. Such models can explain 70–90 percent of the variability in the observed CO concentrations. These models consistently show a statistically significant relationship such that, when wintertime oxyfuel is used, the corresponding ambient CO concentrations are reduced by approximately 14 percent on average, with a 95 percent confidence interval of ±4 percent. Considering that the EPA recently estimates that on-road vehicles contribute only 62 percent of total CO emissions, this new analysis indicates that oxyfuels may be reducing mobile CO emissions by as much as 20 percent, a value consistent with recent tunnel data and emissions tests.

BACKGROUND

Recent reviews of the winter oxyfuel program by the Office of Science and Technology Policy and the National Research Council (NRC) suggested the need for a thorough, statistically defensible analysis of ambient CO data. Subsequent to these reviews, the U.S. EPA released a statistical study by Cook et al. (1996) that used ambient CO data from about 300 monitoring sites. However, Cook et al. used quarterly averages that contain measurements that are often outside the winter period when oxyfuels are fully in use. The highest CO concentrations are typically recorded during December and January and these two months are included in virtually all winter oxyfuel programs. Unfortunately, quarterly data, by definition, contain either December or January, so that the precision of the oxyfuel signal in the data is masked not only by inclusion of some data taken when oxyfuels may not be present, but by data taken when CO concentrations are lower and may even fall below the lower detection limit of 0.5 ppm generally available at most monitors. Also, the Cook et al. study did not statistically account for effects due to individual site variability. Their statistical study gives a rather wide range for the estimate of the oxyfuel impact: between 3.1 and 13.6 percent additional ambient CO reductions due to oxyfuels beyond what would have been expected from the long-term trend due to other control strategies. The authors themselves, and others, have noted that a follow-on study could or should use the available hourly database to create bimonthly averages for only December–January and that a new study could or should account for individual site-effects in the statistical analysis.

Critical reviews of the oxyfuel program (e.g., NRC) have also noted that the use of a control population, where oxyfuels are not in place, should be part or any new analyses. Marnino and Etzel (1996) used summer data along with winter data at the same monitoring sites to address this issue. Therefore, the present study includes June and July data to represent summers when it is assumed that oxyfuel programs are not in place. For 1995, when federal reformulated gasoline was widely used, and for some areas which use substantial percentages of ethanol blended gasoline, it is realized that this assumption is not a guarantee that fuels containing oxygen will not be used during the summer months. We did not use 1996 data for the analysis due to the even more widespread year-round use of oxygenated fuels.

The new study reported here, cofunded by the Oxygenated Fuels Association (OFA) and the Renewable Fuels Association (RFA), addresses the two recommended additions to the Cook et al. analysis of about 300 monitoring sites and utilizes the summer data for some control tests.

DATABASE PREPARATION

The entire hourly CO database was downloaded from the Aerometric Information Retrieval System (AIRS). These data span the timeframe of 1979 through July 1996 from nearly 1000 monitoring sites. The full data set occupies about 600 megabytes of storage. For this study the data were first trimmed to include only U.S. sites for the months of January, June, July and December. Next the data were all converted (where necessary) to the same units of parts per million. The lowest available Pollutant Occurrence Code (POC) value was used in cases where there were multiple measurements for the same site, year and month[1]. For each season we required 2 or more years of data from 1986 to

---

[1] The EPA Aerometric Information Retrieval System (AIRS) database used for these analyses reports all hourly concentrations for the site, month, and POC in each row of the database. Warren Freas, of the EPA Office of Air Quality Planning and Standards, recommended the use of the lowest POC for each month, since this value usually denotes the primary instrumental method in cases where multiple measurements using different instruments are taken at the same time. In some cases this approach may not be strictly valid because the state reporting agencies do not always correctly report the POC codes, and because the primary method may change during the bimonthly period, but it should be sufficiently accurate for this analysis since the use of alternative CO measurements will not usually bias the results although it may slightly increase the overall uncertainty.

1992, 2 or more from 1993 to 1995, and 5 or more from 1986 to 1995. These criteria alone reduced the database to about 400 monitoring sites.

For the bimonthly mean analysis, an estimated background value of 0.2 ppm was subtracted from all hourly concentrations of 0.25 ppm or above; values less than 0.25 ppm were reset to 0.05. The minimum hourly non-background value of 0.05 ppm was selected to allow for logarithmic transformations. Subtracting a 0.2 ppm background value was done for two reasons: first to help reduce instrumental variability at the lowest concentrations and second to make the database more relevant to local CO emissions. As expected, the impacts of making this background adjustment mainly affected the analyses of the summer data[2]. As discussed below, we believe that this background adjustment was especially helpful for identifying a potential effect due to I/M programs that would be masked by the oxyfuels in the winter data but more evident in the summer data at oxyfuel sites. Since the oxyfuel sites generally have higher CO concentrations than non-oxyfuel sites, the oxyfuel sites would be expected to have more of a need for an I/M program than the non-oxyfuel sites but the oxyfuel sites would presumably not have oxyfuels in the summer. Nevertheless, the impacts of I/M programs on CO levels should exist throughout the year and the percentage changes in CO due to I/M may be at least similar between winter and summer.

The database analyzed consisted of the summer (June, July) and winter (December, January) bimonthly means or maximum daily 8-hour concentrations from 1986 to 1995 (e.g., Winter 1995 means December 1994 and January 1995). Bimonthly means were computed by averaging in turn across days and months for each site, season and year combination (after the background subtraction). Maximum rolling-eight-hour averages were also computed for each day of the bimonthly periods (without a background subtraction). By using either the hourly averages or only the highest 8-hour averages, we could check to see if the results were affected in any way by focusing only on the higher concentrations. One anticipated way to affect the results was just discussed above concerning the removal of background concentrations and the influence on low concentrations due to the lower detection limits of some monitors. The federal standard and human exposure also concerns the highest 8-hour averages, so that is why 8-hour averages were used to represent the higher concentrations.

Using information supplied from the EPA, sites were identified as either nonoxy sites or oxy sites, with the latter also identified by the year oxyfuels were introduced. One year of the Denver sites could not be used because the oxyfuel program was introduced in January. Our database had sufficiently complete data for 147 oxy sites and 174 nonoxy sites in the winter and for 143 oxy sites and 170 nonoxy sites in the summer. This is a slightly bigger set of sites than was used in the latest EPA analysis, since we included sites that entered the program before 1992/93 or left the program before 1994/95. Data for 1996 were excluded because several areas entered the reformulated fuels program in 1996, and we believe that it would have complicated the analysis and added significant

---

[2] The background subtraction did not change the statistical analysis results very much. Compared to analyses of the raw data without the estimated background subtraction, the R squared values given in Table 2 below were increased by about 0.05 in the summer and by about 0.01 in the winter. The estimated oxyfuel log CO reductions computed from the best-fitting of the models were slightly larger, by about 0.02 in the summer and 0.01 in the winter.

uncertainty to also model the impacts of year-round reformulated (and oxygenated) gasoline on ambient CO.

STATISTICAL MODEL DEVELOPMENT

Statistical regression and general linear models work by first defining a line (using an equation for a straight line or a simple curve) that a computer will then determine the best possible "fit" through some data set. The best fit occurs when the sum of the squares of the distances between each of the data points (observed value) and the "line" (predicted value) is minimized. Each model that is developed is in thus a new definition of how the line is to be drawn through the database. A wide variety of statistical models were developed in this study so that several alternative statistical representations of the annual, seasonal, oxyfuel, and site-to-site effects on ambient CO could be evaluated and compared.

It was first assumed that a general downward trend would be present at all sites. However, it was not known if the trend might best be considered to be the same at all sites, if it should best be fitted with a straight line (through the logarithm of the CO concentrations) or a curve of some kind. Furthermore, it was not known if the trend might somehow be different in the winter than in the summer or if it would make any difference if it were assumed that the trend might somehow be different at non-oxyfuel sites than at the known oxyfuel sites. Thus, different models were developed that tried each of these possibilities or combinations of them.

Site-to-site effects were accounted for in several of the models by allowing each site to have its own starting point (i.e., intercept) for 1986 for the line to be fitted to the data from that site. In the study by Cook et al. the oxyfuel sites were grouped to fit a single generalized intercept and likewise for the non-oxyfuel sites. When a statistical model is forced to fit such a single intercept, the site-to-site variability shows up in the error standard deviation and the quality of overall fit of the model, that is as if these site differences were "noise" in the data. By allowing each site to have a separate intercept, but using a common trend and a common oxyfuel effect for all sites, then much of this local site-to-site "noise" is thus "suppressed" or filtered out and the "signal" of an oxyfuel effect can be seen more clearly. Furthermore, the addition of extra parameters (i.e., the local site intercepts) enable the statistical models to better fit the data.

Some extra parameters would not be expected to be important to the signal of the oxyfuel effect. However, by using various extra parameters and still seeing similar results for an oxyfuel effect, we should gain some measure of confidence that the results are not critically dependent on a specific model definition. For example, the general downward trend (presumably due mainly to fleet turnover) in various model definitions is allowed to vary by season, by site, by year (i.e., using a curve instead of a straight line), and by oxyfuel status. Other extra parameters that should not be statistically important would be to assume that non-oxyfuel sites have an oxyfuel effect or that oxyfuel sites in the summer might have an oxyfuel effect. In one type of model the latter is used to make a rough estimate of a special I/M effect that might be seen only at oxyfuel sites, since it is reasonable to assume that oxyfuel sites would have more of a need of an I/M program to supplement the use of oxyfuels. However, such an estimate of a supplemental I/M impact to the oxyfuel effect suffers from the very reason that we use a variety of model definitions for the winter oxyfuel signal, namely that the I/M estimate might be critically dependent on the particular model definition used.

Detailed definitions of the 53 models used and their individual results are given in the Appendix. Here we will present an overall discussion of the results.

RESULTS AND DISCUSSION

Due to the many variations of model definition employed in this study, accurate counts of the number of models actually used and the number of data points involved are not clearly definable, but some explanation is given here. For both the winter and the summer data sets, there were essentially two separate databases used: one for the 24-hour averages and the other for the means of the daily 8-hour maxima. With 321 winter sites (and 313 summer sites) over a period of 10 years, the winter and summer databases had nearly 3000 data points each (not all sites had data for all 10 years). For 25 of the model variations used in this study the data were the natural logarithms (log) of the bimonthly observed CO averaged concentrations. For one type of model (described as Group 8 in the Appendix), which had 3 variations, the data used were the differences in log CO between winter and summer[3].

For the 25 of the model variations using CO observations an average temperature adjustment team could also used for each data point. Hence, a total of 53 "models" (28 without temperature and 25 with temperature) can be counted. Technically however, 5 of these model variations (described as Models 1 through 5 in the Appendix) fitted a complete and separate statistical "model" to each and every monitoring site so one could argue that hundreds of models were in fact used. Nevertheless, these 5 model variations had a maximum of only 10 data points (one for each year, though many sites did not have data for all 10 years) to work with for each site. Consequently, with so few data points per "model" these particular variations did not generally show the narrow statistical confidence intervals possible from some of the rest of the model variations which all used databases with nearly 3000 data points each, and hence provide much more accurate estimates of effects common to large numbers of sites.

For the 3000 point databases without temperature, there were, thus, some 23 model variations utilized. The single most important factor to fitting the given database well turned out to be using a site-specific starting point (or intercept) for the type of trend assumed. In all 13 model variations where this site-to-site effect was included, the models all fit their given databases quite well as judged by R^2, a statistical measure of quality where a "perfect" fit would score a 1.0. Conversely, in all 10 model variations which did not include a local site-specific parameter the models did not perform well, always scoring R^2 values less than 0.4.

As long as a given model did include a site-specific term for the trend starting point, the oxyfuel impact determined by every model variation consistently fell in the range corre-

---

[3] When logarithms are subtracted, it corresponds to division. In the present case the winter concentrations were essentially divided by the corresponding summer concentrations for a given year and site. Any impact on the summer concentrations such as from an I/M program would thus be automatically factored out of the data for that site for that year. This assumes that the control program CO reduction in each season is proportional to the CO concentration. Any oxyfuel impact remaining would be over and above the I/M impact. However, any unusual meteorological or other unusual differences between summer and winter for the given year would be also be factored into the data by this technique, so some caution is required so as not to over-emphasize the results from this technique.

sponding to an ambient CO concentration reduction of between 11 and 17 percent. This range of CO reduction is also consistent with the results from the models separately fit to data from a single site. Although, with only a maximum of 10 data points for each of these models R^2 performance quality cannot be as usefully applied, it is noteworthy that the mean oxyfuel impacts for these sets of nearly 300 sub-models each were still consistent with the 11 to 17 percent range of CO reduction from oxyfuels. One might also note that these 300 sub-model sets, by definition, are also consistent with having a local site-specific trend intercept parameter.

Including a temperature parameter with each data point appears to not affect the statistical analysis results found in this study. That is, neither R^2 performance nor the oxyfuel impacts changed significantly when the temperature parameter was added to the statistical models. During this project a new study by Glen, Zelenka, and Graham was published in *Atmospheric Environment* (Vol. 30, pp. 4225-4233, 1996) that does successfully use monthly averages of local mixing height and wind-speed to account for meteorological variances in monthly CO concentrations. Perhaps such local mixing-height and wind-speed data might reduce some of the variance in the data used for the present study better than the temperature parameter that was used here. If this were to be the case, the 300 sub-model sets would be expected to show the most significant improvements in variance reduction. However, the average oxyfuel impact determined here for the 300 sites should not be affected by such reduced variability unless a general non-random meteorological trend effect is present in the data specific only to the oxyfuel sites, which seems highly unlikely.

Figure 1 shows the performance of the oxyfuel program and how one version of the statistical models used in this study track the observed average data. In this figure a significant decrease in CO is seen for the oxyfuel sites between 1992 and 1993, when most of the sites entered the program. Less noticeable in the figure are some small deviations from a straight line for the oxyfuel sites at years prior to 1992. These small deviations are caused by the model tracking some of the earlier oxyfuel programs. Had all sites started oxyfuels in 1993, the trend line before would be perfectly straight and the sharp decrease in 1993 would be even more pronounced.

Addition of 1996 Data

As a further sensitivity test we added the 1996 winter data to the database and repeated the first 25 model variations on the 24 hour data without a temperature effect. The original results using the data up to 1995 are given in Table 2 and the results using the 1996 data are given in Table 2a for comparison. We also show a graphical comparison in Figures 1 and 1a.

Adding the 1996 data makes a significant change in the overall database, because 46 new oxyfuel sites enter the database. Almost all of the new oxyfuel sites come from the non-oxyfuel site data, but three more were added because they could meet the criteria for 5 total years. The winter total becomes 324 with 189 oxy sites and 130 non-oxy sites.

The additional dip in CO concentrations that occurs between 1995 and 1996 for the upper regression line is new, expected and attributed to RFG (which is treated as an oxyfuel). The performance of the models is approximately the same, although the “oxyfuel effect” and R^2-values often drop (e.g., for model 12 [24-hr averages] the R^2 drops from 0.87 to

0.86 and the "oxyfuel effect" drops from about 14 percent to 13 percent. The standard deviations, however, are generally smaller for the 1996 data, so the uncertainty is reduced somewhat apparently because of the additional post-implementation data. The corresponding values using 8-hr maximum average concentrations were qualitatively similar to the previous results excluding the 1996 data.)[4]

CONCLUSIONS

By assuming that a general downward trend exists in all CO observations, but some monitoring sites experience a sudden additional reduction in ambient CO concentrations when oxyfuels are known to be present, it was possible to account for 70–90 percent (as measured by the statistical R^2 factor) of the observed variability in winter data provided each local site was assigned a starting point for the general trend specific to that site.

By using as many as 53 different combinations of assumptions, it was found that it made no significant difference in the size (about 14 ± 4 percent with 95 percent confidence) of this estimated sudden oxyfuel impact, whether the general downward trend (not associated with the oxyfuel program) was assumed to be linear (through logarithms of the observed concentrations), or curved, or the same, or different between summer and winter, or the same or different between known oxy sites and nonoxy sites. Even when a separate general trend was allowed at each of the nearly 300 monitoring sites, the average oxyfuel impact over the 300 sites was still found to be consistent with the 14 ± 4 percent range of oxyfuel impacts estimated using a consistent trend at all sites. Additionally, the size of the estimated oxyfuel effect does not appear to be sensitive to whether the data were prepared from full 24 hour averages or if only higher concentrations were used such as the highest daily 8-hour averages. That is, data variability due to low values near the detection limits of the instruments, did not seem to affect the results of this study (an assumed background of 0.2 ppm was subtracted from all bimonthly average hourly data and this subtraction also did not appear to alter the results significantly).

For one type of analysis the winter data were, in effect, divided by the summer data at each site for each year in an attempt to make an estimate of the importance of I/M programs in place all year at the oxyfuel sites. This filtering technique did result in a somewhat lower estimate for the oxyfuel impact, but it was still within the 95 percent confidence interval found in the estimates without using this data-filtering technique. Hence, it appears that the impact of any I/M programs at oxyfuel sites is small compared to the 14 percent average impact due to oxyfuels estimated elsewhere in this study.

---

[4] The additional dip over the 1995-1996 period for the oxysite regression line, as noted, is attributed to the inclusion of 46 new "oxysites", which arises from the introduction of RFG during 1995. These new oxysites also show a reduction in the running mean values, indicating that these new sites were generally lower in CO concentrations. What also looks very interesting is the dip in the (lower) non-oxysite regression line. One reason for this could be that many of these "non-oxysites" have not been "properly" identified as "RFG-sites" in the file sent to us by the EPA or that the oxyfuel effect occurs over a wider region than the specific oxyfuel area.

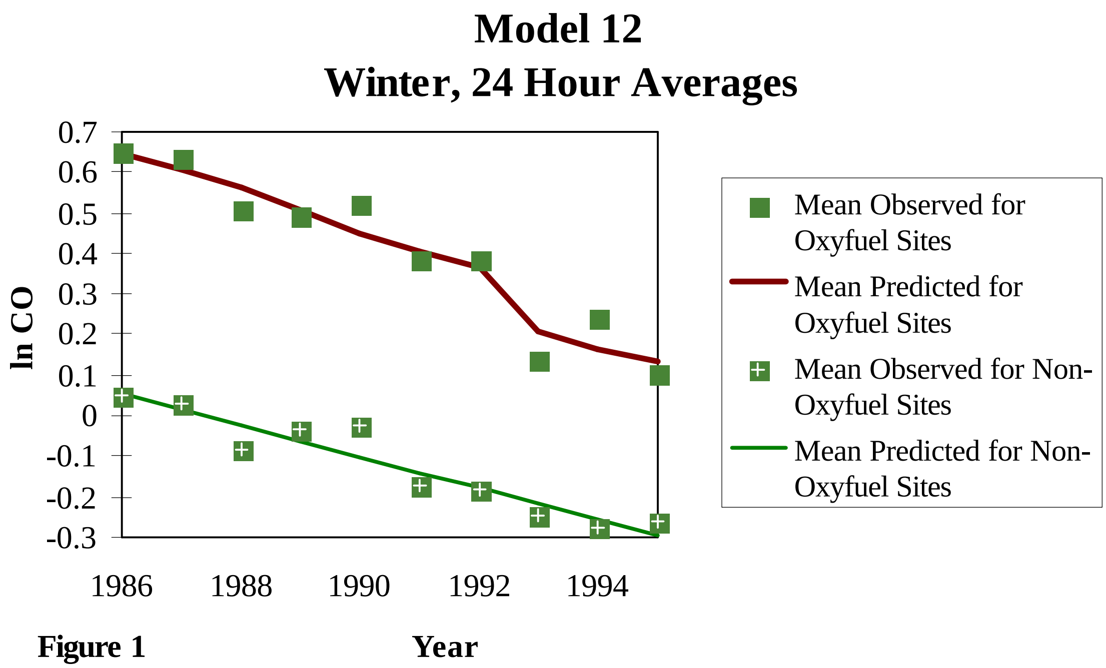

Figure 1

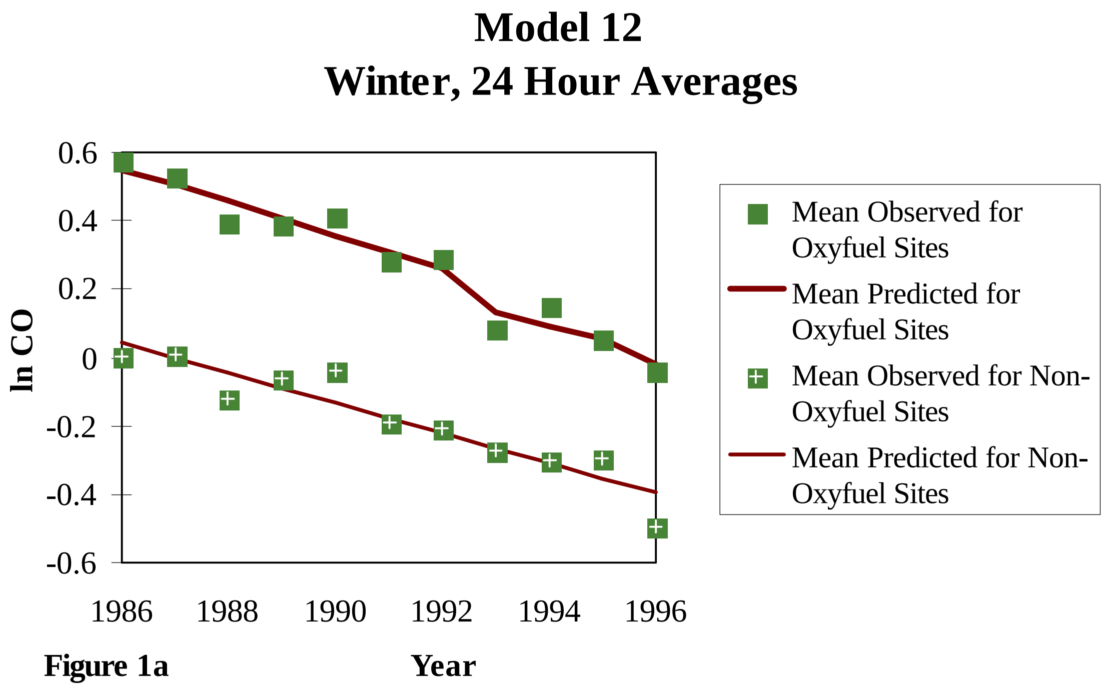

Figure 1a

Sources of variability in the data due to meteorological effects may be important. However, when an average temperature was associated with each data point, the results seen in this study were not significantly altered even though virtually every model variation was run both with and without the temperature parameter.

The statistically robust oxyfuel effect of 14 ± 4 percent reduction in ambient CO concentrations appears to result from two key factors not found in previous studies: first, the use of December–January data ensure that no uncertainties (i.e., masking variabilities) are introduced from data taken during transition periods before and after local oxyfuel programs are clearly in place; and second, the justifiable use of site-specific starting points for a common trend across the sites seems to account for site-to-site variability that would otherwise be part of the statistical uncertainty in the estimated oxyfuel effect.

APPENDIX

The analyses can be conveniently divided into eight groups, depending upon how the site/season data are grouped together and upon whether the base log CO emissions levels can vary by site or just by oxyfuel program status (i.e. whether or not the site was in the oxyfuel program for one or more years from 1986 to 1995). The various models are described in Table 1. The results from fitting those models are detailed in Table 2. In Table 2, for each group of sites and seasons, we give the mean, standard deviation, minimum, and maximum estimated oxyfuel effect. The oxyfuel effect is the estimated change in the natural logarithm of CO when all other factors in the model are unchanged. This value is close to the estimated relative change in CO due to the oxyfuel effect (the latter approximation is most accurate for small changes in log CO). The standard deviation of the oxyfuel effect is provided if only one model is fitted to the group of sites and seasons. A 95% confidence interval for the oxyfuel effect is the estimated effect ("Mean") plus or minus 1.96 standard deviations. The minimum and maximum across the site/season values are provided only if separate estimates are computed for individual sites.

Group 1, model 1, separately fits an intercept, trend and oxyfuel effect for each site and season. In effect, this gives 300 separate "models" (i.e., one for each site). In this model the oxy program is hypothetically assumed to continue in the summer, and the nonoxy sites are hypothetically assumed to have possible oxyfuel effects for years 1993 to 1995. These hypothetical effects do not affect the analyses for the oxy program sites in the winter. The model shows an average log CO reduction of 0.17 in the winter for oxy sites during program years. For oxy sites during the summer the hypothetical effect is an average 0.10 reduction, and nonoxy sites show a hypothetical average increase of 0.01 in the summer and a reduction of 0.04 in the winter. The predictions vary greatly across sites and seasons and are often not statistically significant, particularly for the hypothetical cases.

A closer inspection of the results for this model 1 system indicates that of the total 147 oxyfuel sites 119 showed winter CO decreases when oxyfuels were in use, but only 11 sites showed statistically significant decreases at the five percent level. Surprisingly, 10 of these 11 sites were in California, where a somewhat lower fuel-oxygen percentage is used than in the other states. Of the 28 sites giving CO increases using this modeling approach, none of these increases were statistically significant. We note that the Marnino and Etzel (1996) study was criticized in the NRC review for not explaining CO increases modeled at some oxyfuel sites, and suggest that the statistical significance or those potential CO increases may also have been low in the similar statistical model used by Marnino and Etzel.

Group 2, models 2 and 3, separately fit a seasonal effect, annual trend, and oxyfuel effect for each site (combining data from both seasons at each site). In both models the oxy fuel effect is again assumed to apply in both seasons during program years, and the program is again assumed to start in 1993 for nonoxy sites. In model 2 the annual trend is assumed the same for both seasons. In model 3 the annual trend can differ between the summer and winter averages. Average estimated reductions in log CO due to the oxy program are

0.14 for oxy sites and 0.02 for nonoxy sites. (The results for the hypothetical 1993–1995 effects in nonoxy sites do not affect the results for the oxy sites).

Group 3, models 4 and 5, is similar to group 2, models 2 and 3, except that in this case the oxy fuel effect is only assumed to occur in the winter. Average estimated reductions in log CO due to the oxy program are 0.13 or 0.17 for oxy sites and 0.02 or 0.03 for nonoxy sites. (The results for the hypothetical 1993–1995 effects in nonoxy sites do not affect the results for the oxy sites).

Group 4, models 6 to 9, are variants of the model fitted in the EPA analysis of winter quarter data (Cook et al., 1996). Each model is fitted separately to each season. The analysis of the summer season data, for comparison purposes only, does not affect the results for the winter season that are of primary interest. Each model assumes a possible reduction in log CO for program years (in both seasons). Each model assumes that the 1986 base CO level for each season depends only upon whether the site is a program site or not, and does not vary across each group of sites. Any site to site variability in the 1986 base level is therefore included in the error variance and is not adjusted for. Model 6 assumes a linear annual trend and is the same as the EPA model “B” except that the term “POST92” is excluded. (The POST92 term allows for a hypothetical nationwide reduction in log CO that occurs for the same years as the oxy program, in most cases, but is not attributable to the oxyfuel program. The POST92 effect was not statistically significant in the EPA analysis. If the POST92 effect really exists, then its exclusion from the model would indeed confound the estimated oxyfuel effect.) Model 7 assumes a quadratic annual trend instead of a linear trend. Model 8 allows for different linear trends for oxy sites and nonoxy sites (similar to EPA’s model A). Model 9 allows for different quadratic trends for oxy sites and nonoxy sites.

The results in Table 2 show that all four models in group 4 fit poorly, with R squared values of 0.26 for the winter and 0.09 for the summer. For a general linear model, the R squared is a measure of fit that ranges from 0 (no fit) to 1 (perfect fit). While there is no specific acceptable level for the R squared statistic, models with R squared less than 0.5 can usually be treated as unacceptable, since the model then explains less than 50 percent of the variability in the data. The predicted winter log CO reductions are 0.07 for the linear trend models (comparable to the EPA analysis results) and –0.01 or +0.01 for the quadratic trend models. However these estimated effects are likely to be inaccurate because of the poor overall fit. The poor fit is because site-to-site variability in the base levels is not taken into account.

Group 5, models 10 to 13, correspond to group 4, models 6 to 9, but fit the data much better because the base level is in each case assumed to vary by site, instead of by the program status (oxy site or nonoxy site). These models have R squared values of 0.87 for the winter and 0.76 for the summer. The oxyfuel effects on log CO show estimated reductions from 0.04 to 0.09 (corresponding to 4 to 8 percent reduced CO) in the summer and from 0.14 to 0.16 in the winter (corresponding to 13 to 15 percent reduced CO). The winter percentage effects are correct to within about 4 percentage points, based on a 95% confidence level.

Group 6, models 14 to 19, and group 7, models 20 to 25, are alternative models that combine data from both seasons and all sites. All models in these groups assume that the oxyfuel effect occurs only at oxy sites in program years in the winter. The group 6 models each assume that the base level depends upon whether or not the site is an oxy program site and upon the season. Model 14 assumes a linear annual trend, the same for both seasons. Model 15 assumes a quadratic annual trend, the same for both seasons. Model 16 assumes a linear annual trend that may vary between oxy and nonoxy sites and between seasons. Model 17 assumes a quadratic annual trend that may vary between oxy and nonoxy sites and between seasons. Model 18 assumes a linear annual trend that may vary between seasons, but not between oxy and nonoxy sites. Model 19 assumes a quadratic annual trend that may vary between seasons. Just as for group 4, the group 6 models fit poorly, with R^2 values of 0.33, since they do not allow for site-to-site variation in the base levels (other than between oxy and nonoxy sites).

The group 7 models 20 to 25 are versions of the corresponding group 6 models 14 to 19. The assumptions about the annual trend and oxyfuel effects are the same. For group 7, the base level is assumed to vary by site and season. The much better fitting group 7 models have R squared values of 0.85 and show oxyfuel log CO reductions ranging from 0.12 ± 0.04 to 0.16 ± 0.04 (corresponding to 11 ± 3 to 15 ± 4 percent reduced CO). The ± uncertainties are at a 95% confidence level.

For Groups 1-7 the error terms for the two seasons in each site/year combination are assumed to be independent and to have equal variances. For Group 8 these assumptions were dropped, since the Group 8 analyses were of the winter log CO minus the summer log CO. The Group 8 models apply under the weaker assumptions that the *differences* in log CO are independent and have equal variances across sites and years. In each case the base level varies by site and the oxyfuel effect applies only for oxysites in the winter during program years. In Model 26 the annual trend in log CO is assumed to be the same for both seasons, and all sites, but is otherwise arbitrary (e.g., linear, quadratic, or more complicated formulations). Thus, the Model 26 assumptions hold under the more restrictive conditions of Models 20 and 21, but Model 26 applies in more general situations. Similarly, Models 27 and 28 are more generally applicable versions of Models 24 and 25, respectively. Under Model 27 the seasonal difference in log CO is assumed to have a linear trend. In Model 28, the seasonal difference in log CO is assumed to have a quadratic trend. The Group 7 and 8 models share the advantage that the summer trends can be used to adjust for annual trends in winter CO unrelated to the oxyfuel program (e.g., I/M effects or fleet turnover effects). The Group 8 models have the additional advantage of not requiring equal variances and independence across the two seasons, although the Group 8 models will also apply if these distributional assumptions hold.

The results for the group 8 models show smaller estimated CO reductions than the other models. The model 26, that assumes no difference between the summer and winter trends for each site, shows an oxyfuel log CO reduction of 0.07 ± 0.04, corresponding to a percentage CO reduction of 7 ± 4 percent. The statistically significantly better fitting models 27 and 28, that assume that the yearly site differences vary with linear or quadratic trends, find oxyfuel log CO reductions of 0.12 ± 0.05, corresponding to percentage CO reductions of 11 ± 5 percent.

In summary, the results show that the best fitting groups of models are groups 1, 2, 3, 5, 7, and 8, since groups 4 and 6 have very low R^2 values. The results for groups 1 to 3 show wide variability in the oxyfuel and other effects across site, although it should be noted that the individual site effects cannot be estimated very precisely since only a few annual statistics are used for each site, leaving few degrees of freedom for error. The models in groups 5 and 7 assume the same oxyfuel effects and trends across large groups of sites (and, in some cases, seasons) and fit the data reasonably well. The estimated winter oxyfuel effect is very consistent among all these group 5 and 7 models, showing CO reductions of 11 to 15 percent due to the winter oxy fuel program. These estimates are accurate within 4 percentage points at the 95 percent confidence interval. Nevertheless, a more detailed review of the statistical results (not shown here) shows that the group 5 models are the better fitting, primarily because they allow for different variances for the two seasons; the fitted error variances for each season are statistically significantly different (by a factor of two or more), and the total log-likelihoods for the group 5 models are much lower.

It is not clear why the less restrictive group 8 models show smaller oxyfuel effects than the reasonably well-fitting models in groups 5 and 7. Since the group 8 analysis is a paired comparison, using one data point for each year instead of the two seasonal values, the R squared statistics for these groups cannot be meaningfully compared. The approach for the analyses in groups 7 and 8 adjust for impacts of year-round control programs and fleet turnover effects by using the combined summer and winter data to estimate the part of the trend unrelated to the oxyfuel program. However this approach is subject to the greater uncertainty and variability in the summer ambient CO concentrations compared to the much higher winter ambient CO concentrations; the summer concentrations are closer to the uncertain background levels. The group 5 models that analyze each season separately may give more accurate results for the winter season, under the assumption that the assumed linear or quadratic trends capture most of the impacts of year-round control programs and fleet turnover.

Bimonthly Means Reanalysis Adjusting for a Possible Temperature Effect

The analyses of the bimonthly mean concentrations described above did not directly take into account meteorological effects on ambient concentrations. These possible impacts were incorporated into the error variability and it is implicitly assumed that the averaging over a long, two month period would remove much of the impact due to meteorological variability. To evaluate the importance of meteorological variability on the results, we evaluated whether incorporating an adjustment for temperature would change the estimated oxyfuel impacts. Monthly mean temperature data were downloaded from the National Climatic Data Center internet database for the years studied. These data were converted into degrees Kelvin (absolute temperature) and averaged across the two months for each site, year, and season. All but one of the CO monitoring sites were within 50 km of the nearest meteorological site used for this analysis. Each of the group 1 to 7 regression models described in Table 1 were then redefined to add a linear temperature term as one of the independent variables and the estimated oxyfuel effects were recomputed. These revised estimates are therefore adjusted for a linear temperature effect. This procedure was not applied to the group 8 models for the seasonal change in log CO since for those two models the summer and winter temperatures would both have possible impacts on the log CO difference that might be difficult to model.

The results of the analysis of the oxyfuel impacts on the bimonthly mean CO adjusted for temperature are shown in Table 3. Note that the R^2 values in Table 3 should not be directly compared with the values in Table 2 because the dataset with nonmissing log CO and temperature values was smaller than the dataset with nonmissing log CO values used for the Table 2 analysis (a total of 5643 data points with temperature values compared to

6058 with or without temperature values). The same remark applies for the R^2 values in Tables 4 and 5 described below. For the group 4 and 6 models 6 to 9 and 15 to 19, the temperature adjustment made a small, but statistically significant, improvement in the model fit, but these slightly improved models still have too low an R^2 value to be reliable. For the group 5 models 10 to 13, the temperature adjustment made a small improvement in the model fit, but the temperature effect was not statistically significant for either season. For those models, the adjusted estimated oxyfuel effect was a little larger than the unadjusted estimate, particularly for the summer season. Adjusted winter oxyfuel effects were log CO reductions of 0.14 ± 0.05 to 0.17 ± 0.03, depending upon the selected model (corresponding CO percentage reductions are 13 ± 4 to 16 ± 3 percent). Similarly, the estimated oxyfuel effects using the group 7 models 20 to 25, that combined results from both seasons, showed slightly greater log CO reductions of 0.13 ± 0.04 to 0.17 ± 0.04 (corresponding CO percentage reductions are 12 ± 4 to 16 ± 4 percent).

Oxyfuel Impact on Average Daily Maximum Eight-Hour Average CO

Using the same hourly concentration database as in the bimonthly mean analyses, and the same set of regression models, we analyzed the winter oxyfuels program impact on the average daily maximum eight-hour average ambient CO concentration. For each site, year, and day, and for each hour of the day starting at midnight, we computed the average of the concentrations for that hour and the following seven hours. We then computed the daily maximum eight hour average concentration for each site and day. These daily maxima were then averaged across days in each of the two month seasons (December to January, and June to July). To avoid the need to analyze extra months of data, we excluded the January 31 and July 31 daily maxima from these averages, since their inclusion would require analyzing the first seven hours of the months of February and August.

The results of the regression analyses are shown in Tables 4 (no temperature adjustment), and 5 (oxyfuel effects adjusted for a linear temperature effect). These results are very similar to the corresponding results in Tables 2 and 3 for the bimonthly mean statistics. As before, the best fitting of the non-site specific models are in groups 5, 7, and 8, with R^2 values ranging from 0.7 to 0.9. The models that do not allow for a base level that varies by site fit poorly, as before. The models for the average daily maximum consistently fit better than the models for the bimonthly means, as might be expected due to the larger influence of the uncertain background concentrations on the mean concentrations. The estimated winter oxyfuel log CO reductions for the best-fitting group 5 and 7 models for the average daily maximum, range from 0.13 to 0.14, with uncertainties of ± 0.03 to ± 0.05, depending upon the model formulation. These reductions correspond to percentage CO reductions of 12 or 13 percent, with uncertainties of about ± 4 percent, at the 95 percent confidence level. For the group 8 models, the estimated reductions are again smaller, showing log CO reduced by 0.12 ± 0.04, and CO reduced by 11 ± 3 percent.

TABLE 1. Statistical models for Log CO.

Group	Data Subsets[1]	Model Number	Trend Varies by[2]	Type of Trend	Base Level (intercept) varies by[2]	Oxyfuel Effect[3]
1	site, season	1	site, season	linear	site, season	oxyyear
2	site	2	site	linear	site, season	oxyyear
		3	site, season	linear	site, season	oxyyear
3	site	4	site	linear	site, season	oxyind[2]
		5	site, season	linear	site, season	oxyind[2]
4	season	6	season	linear	season, oxysite	oxyind[3]
		7	season	quadratic	season, oxysite	oxyind[3]
		8	season, oxysite	linear	season, oxysite	oxyind[3]
		9	season, oxysite	quadratic	season, oxysite	oxyind[3]
5	season	10	season	linear	season, site	oxyind[3]
		11	season	quadratic	season, site	oxyind[3]
		12	season, oxysite	linear	season, site	oxyind[3]
		13	season, oxysite	quadratic	season, site	oxyind[3]
6	all	14	constant trend	linear	season, oxysite	oxyind
		15	constant trend	quadratic	season, oxysite	oxyind
		16	season, oxysite	linear	season, oxysite	oxyind
		17	season, oxysite	quadratic	season, oxysite	oxyind
		18	season	linear	season, oxysite	oxyind
		19	season	quadratic	season, oxysite	oxyind
7	all	20	constant trend	linear	season, site	oxyind
		21	constant trend	quadratic	season, site	oxyind
		22	season, oxysite	linear	season, site	oxyind
		23	season, oxysite	quadratic	season, site	oxyind
		24	season	linear	season, site	oxyind
		25	season	quadratic	season, site	oxyind
8[4]	all	26	constant trend	zero trend difference	season, site	oxyind
		27	season	linear trend difference	season, site	oxyind
		28	season	quadratic trend difference	season, site	oxyind

Notes:

1. A separate statistical model is independently fitted to each data subset. The error variance depends on the data subset.
2. Oxysite = 1 if the site was ever in the oxyfuel program between 1986 and 1995.
Oxysite = 0 otherwise.
3. Oxyyear = 1 for nonoxy sites in years 1993-1995 and for oxysites in the years when the winter oxyfuel program was in effect. (For example, oxyyear = 1 for all sites and both seasons in 1993).
Oxyind = 1 if oxysite = 1, oxyyear = 1, and season = winter.
Oxyind2 = 1 if oxyyear = 1 and season = winter.

TABLE 1. Statistical models for Log CO.

Group	Data Subsets[1]	Model Number	Trend Varies by[2]	Type of Trend	Base Level (intercept) varies by[2]	Oxyfuel Effect[3]

Oxyind3 = 1 if oxyyear = 1 and oxysite = 1.
Otherwise the indicator = 0.

4. For Groups 1 to 7 the dependent variable is log CO. For Group 8 the dependent variable is the winter log CO minus the summer log CO.

TABLE 2. Results for bimonthly mean ambient CO. No temperature effect assumed.

Model Number[1]	Season	R^2	Oxy Site?	Mean Oxyfuel Effect[2]	Standard Deviation[3]	Minimum Oxyfuel Effect[2,4]	Maximum Oxyfuel Effect[2,4]
1	Summer	NA	No	0.009	NA	-1.333	1.819
1	Summer	NA	Yes	-0.103	NA	-2.541	0.876
1	Winter	NA	No	-0.035	NA	-0.874	1.367
1	Winter	NA	Yes	-0.171	NA	-1.023	0.260
2	NA	NA	No	-0.015	NA	-0.891	1.237
2	NA	NA	Yes	-0.144	NA	-1.782	0.350
3	NA	NA	No	-0.015	NA	-0.913	1.237
3	NA	NA	Yes	-0.145	NA	-1.782	0.350
4	NA	NA	No	0.021	NA	-0.656	0.800
4	NA	NA	Yes	-0.125	NA	-0.929	0.483
5	NA	NA	No	-0.034	NA	-0.874	1.367
5	NA	NA	Yes	-0.171	NA	-1.023	0.260
6	Summer	0.091	NA	-0.119	0.042	NA	NA
6	Winter	0.256	NA	-0.074	0.032	NA	NA
7	Summer	0.091	NA	-0.127	0.043	NA	NA
7	Winter	0.256	NA	-0.071	0.033	NA	NA
8	Summer	0.092	NA	-0.190	0.055	NA	NA
8	Winter	0.257	NA	-0.005	0.042	NA	NA
9	Summer	0.093	NA	-0.219	0.059	NA	NA
9	Winter	0.258	NA	0.015	0.045	NA	NA
10	Summer	0.758	NA	-0.041	0.025	NA	NA
10	Winter	0.872	NA	-0.157	0.015	NA	NA
11	Summer	0.758	NA	-0.047	0.025	NA	NA
11	Winter	0.872	NA	-0.157	0.015	NA	NA
12	Summer	0.758	NA	-0.068	0.034	NA	NA
12	Winter	0.872	NA	-0.148	0.021	NA	NA
13	Summer	0.759	NA	-0.086	0.037	NA	NA
13	Winter	0.872	NA	-0.144	0.023	NA	NA
14	NA	0.330	NA	-0.054	0.035	NA	NA
15	NA	0.330	NA	-0.054	0.035	NA	NA
16	NA	0.330	NA	-0.005	0.049	NA	NA
17	NA	0.331	NA	0.015	0.052	NA	NA
18	NA	0.330	NA	-0.074	0.037	NA	NA
19	NA	0.330	NA	-0.071	0.038	NA	NA
20	NA	0.845	NA	-0.122	0.019	NA	NA
21	NA	0.845	NA	-0.124	0.019	NA	NA
22	NA	0.846	NA	-0.148	0.028	NA	NA
23	NA	0.846	NA	-0.144	0.030	NA	NA
24	NA	0.845	NA	-0.157	0.021	NA	NA
25	NA	0.846	NA	-0.157	0.021	NA	NA
26	NA	0.622	NA	-0.071	0.021	NA	NA
27	NA	0.624	NA	-0.119	0.025	NA	NA
28	NA	0.624	NA	-0.115	0.025	NA	NA

1. The models are defined in Table 1.
2. Type of oxyfuel effect modeled is defined in Table 1. Reported value is change in natural logarithm of ambient CO attributable to oxyfuel effect, keeping all other factors fixed. The corresponding percentage change in CO is approximately given as 100 times the reported value.
3. Not available for Models 1 to 5 where different effects are calculated for each site.
4. Range of oxyfuel effects across groups of sites and seasons. Not reported for Models 6 to 28 where a national oxyfuel effect is calculated for each season or for both seasons.
5. NA = not available or not applicable.

TABLE 2a. Results for bimonthly mean ambient CO. 1996 data included.

Model Number[1]	Season	R^2	Oxy Site?	Mean Oxyfuel Effect[2]	Standard Deviation[3]	Minimum Oxyfuel Effect[2,4]	Maximum Oxyfuel Effect[2,4]
1	Summer	NA	No	-0.011	NA	-1.333	1.371
1	Summer	NA	Yes	-0.074	NA	-2.541	0.876
1	Winter	NA	No	-0.044	NA	-0.866	1.367
1	Winter	NA	Yes	-0.162	NA	-1.072	0.433
2	NA	NA	No	-0.034	NA	-0.900	1.063
2	NA	NA	Yes	-0.132	NA	-1.782	0.480
3	NA	NA	No	-0.032	NA	-0.908	1.063
3	NA	NA	Yes	-0.138	NA	-1.782	0.480
4	NA	NA	No	-0.014	NA	-0.918	0.800
4	NA	NA	Yes	-0.119	NA	-1.072	0.444
5	NA	NA	No	-0.044	NA	-0.866	1.367
5	NA	NA	Yes	-0.162	NA	-1.072	0.433
6	Summer	0.090	NA	-0.037	0.038	NA	NA
6	Winter	0.225	NA	0.098	0.029	NA	NA
7	Summer	0.090	NA	-0.038	0.039	NA	NA
7	Winter	0.227	NA	0.110	0.029	NA	NA
8	Summer	0.091	NA	-0.013	0.044	NA	NA
8	Winter	0.231	NA	0.188	0.034	NA	NA
9	Summer	0.091	NA	-0.010	0.045	NA	NA
9	Winter	0.233	NA	0.208	0.035	NA	NA
10	Summer	0.759	NA	-0.045	0.024	NA	NA
10	Winter	0.862	NA	-0.137	0.015	NA	NA
11	Summer	0.760	NA	-0.050	0.041	NA	NA
11	Winter	0.862	NA	-0.129	0.015	NA	NA
12	Summer	0.760	NA	-0.023	0.029	NA	NA
12	Winter	0.862	NA	-0.136	0.018	NA	NA
13	Summer	0.760	NA	-0.025	0.030	NA	NA
13	Winter	0.862	NA	-0.131	0.019	NA	NA
14	NA	0.313	NA	0.087	0.030	NA	NA
15	NA	0.314	NA	0.095	0.031	NA	NA
16	NA	0.316	NA	0.188	0.038	NA	NA
17	NA	0.317	NA	0.208	0.040	NA	NA
18	NA	0.314	NA	0.098	0.033	NA	NA
19	NA	0.314	NA	0.110	0.033	NA	NA
20	NA	0.842	NA	-0.113	0.017	NA	NA
21	NA	0.842	NA	-0.111	0.017	NA	NA
22	NA	0.843	NA	-0.136	0.023	NA	NA
23	NA	0.843	NA	-0.131	0.024	NA	NA
24	NA	0.842	NA	-0.137	0.019	NA	NA
25	NA	0.843	NA	-0.129	0.019	NA	NA
26	NA	NA	NA	NA	NA	NA	NA
27	NA	NA	NA	NA	NA	NA	NA
28	NA	NA	NA	NA	NA	NA	NA

1. The models are defined in Table 1.
2. Type of oxyfuel effect modeled is defined in Table 1. Reported value is change in natural logarithm of ambient CO attributable to oxyfuel effect, keeping all other factors fixed. The corresponding percentage change in CO is approximately given as 100 times the reported value.
3. Not available for Models 1 to 5 where different effects are calculated for each site.
4. Range of oxyfuel effects across groups of sites and seasons. Not reported for Models 6 to 28 where a national oxyfuel effect is calculated for each season or for both seasons.
5. NA = not available or not applicable.

TABLE 3. Results for bimonthly mean ambient CO. Model includes a linear temperature effect.

Model Number[1]	Season	R^2	Oxy Site?	Mean Oxyfuel Effect[2]	Standard Deviation[3]	Minimum Oxyfuel Effect[2,4]	Maximum Oxyfuel Effect[2,4]
1	Summer	NA	No	0.014	NA	-1.354	1.907
1	Summer	NA	Yes	-0.109	NA	-2.535	0.946
1	Winter	NA	No	-0.030	NA	-1.000	1.375
1	Winter	NA	Yes	-0.131	NA	-1.015	0.418
2	NA	NA	No	-0.011	NA	-0.891	1.218
2	NA	NA	Yes	-0.132	NA	-1.782	0.334
3	NA	NA	No	-0.013	NA	-0.911	1.218
3	NA	NA	Yes	-0.132	NA	-1.776	0.329
4	NA	NA	No	0.023	NA	-0.679	0.835
4	NA	NA	Yes	-0.094	NA	-0.945	0.443
5	NA	NA	No	-0.034	NA	-0.895	1.338
5	NA	NA	Yes	-0.142	NA	-1.010	0.316
6	Summer	0.128	NA	-0.140	0.046	NA	NA
6	Winter	0.284	NA	-0.096	0.034	NA	NA
7	Summer	0.129	NA	-0.151	0.047	NA	NA
7	Winter	0.284	NA	-0.094	0.035	NA	NA
8	Summer	0.131	NA	-0.240	0.059	NA	NA
8	Winter	0.286	NA	-0.038	0.044	NA	NA
9	Summer	0.132	NA	-0.292	0.065	NA	NA
9	Winter	0.286	NA	-0.017	0.048	NA	NA
10	Summer	0.773	NA	-0.063	0.027	NA	NA
10	Winter	0.877	NA	-0.165	0.016	NA	NA
11	Summer	0.774	NA	-0.070	0.027	NA	NA
11	Winter	0.877	NA	-0.163	0.016	NA	NA
12	Summer	0.774	NA	-0.087	0.036	NA	NA
12	Winter	0.877	NA	-0.152	0.022	NA	NA
13	Summer	0.774	NA	-0.112	0.040	NA	NA
13	Winter	0.877	NA	-0.143	0.024	NA	NA
14	NA	0.355	NA	-0.070	0.037	NA	NA
15	NA	0.355	NA	-0.071	0.037	NA	NA
16	NA	0.356	NA	-0.041	0.052	NA	NA
17	NA	0.356	NA	-0.022	0.056	NA	NA
18	NA	0.356	NA	-0.098	0.040	NA	NA
19	NA	0.356	NA	-0.097	0.041	NA	NA
20	NA	0.854	NA	-0.130	0.020	NA	NA
21	NA	0.854	NA	-0.130	0.020	NA	NA
22	NA	0.854	NA	-0.152	0.029	NA	NA
23	NA	0.854	NA	-0.143	0.033	NA	NA
24	NA	0.854	NA	-0.165	0.022	NA	NA
25	NA	0.854	NA	-0.163	0.022	NA	NA

1. The models are defined in Table 1.
2. Type of oxyfuel effect modeled is defined in Table 1. Reported value is change in natural logarithm of ambient CO attributable to the oxyfuel effect, keeping all other factors fixed. Corresponding percentage change is approximately given as 100 times the reported value.
3. Standard deviation of the oxyfuel effect. Not available for Models 1 to 5 where different effects are calculated for each site.
4. Minimum or maximum oxyfuel effect gives the range of oxyfuel effects across groups of sites and seasons. Not reported for Models 6 to 25 where a national oxyfuel effect is calculated for each season or for both seasons.
5. NA = not available or not applicable.

TABLE 4. Results for bimonthly mean daily maximum 8-hr average ambient CO. No temperature effect assumed.

Model Number[1]	Season	R^2	Oxy Site?	Mean Oxyfuel Effect[2]	Standard Deviation[3]	Minimum Oxyfuel Effect[2,4]	Maximum Oxyfuel Effect[2,4]
1	Summer	NA	No	0.037	NA	-0.843	2.591
1	Summer	NA	Yes	-0.072	NA	-2.696	1.224
1	Winter	NA	No	-0.029	NA	-0.807	0.798
1	Winter	NA	Yes	-0.153	NA	-0.752	0.273
2	NA	NA	No	0.002	NA	-0.770	1.327
2	NA	NA	Yes	-0.117	NA	-1.666	0.398
3	NA	NA	No	0.003	NA	-0.784	1.317
3	NA	NA	Yes	-0.119	NA	-1.666	0.405
4	NA	NA	No	-0.006	NA	-1.404	0.652
4	NA	NA	Yes	-0.138	NA	-1.204	0.344
5	NA	NA	No	-0.029	NA	-0.807	0.798
5	NA	NA	Yes	-0.153	NA	-0.752	0.273
6	Summer	0.083	NA	-0.058	0.036	NA	NA
6	Winter	0.264	NA	-0.045	0.028	NA	NA
7	Summer	0.083	NA	-0.068	0.036	NA	NA
7	Winter	0.264	NA	-0.043	0.028	NA	NA
8	Summer	0.083	NA	-0.096	0.047	NA	NA
8	Winter	0.266	NA	0.028	0.036	NA	NA
9	Summer	0.084	NA	-0.119	0.050	NA	NA
9	Winter	0.267	NA	0.047	0.038	NA	NA
10	Summer	0.774	NA	-0.024	0.020	NA	NA
10	Winter	0.901	NA	-0.141	0.012	NA	NA
11	Summer	0.775	NA	-0.035	0.020	NA	NA
11	Winter	0.901	NA	-0.142	0.012	NA	NA
12	Summer	0.775	NA	-0.044	0.028	NA	NA
12	Winter	0.901	NA	-0.135	0.016	NA	NA
13	Summer	0.775	NA	-0.068	0.030	NA	NA
13	Winter	0.901	NA	-0.138	0.017	NA	NA
14	NA	0.359	NA	-0.048	0.029	NA	NA
15	NA	0.359	NA	-0.051	0.030	NA	NA
16	NA	0.360	NA	0.028	0.041	NA	NA
17	NA	0.360	NA	0.047	0.044	NA	NA
18	NA	0.359	NA	-0.045	0.032	NA	NA
19	NA	0.360	NA	-0.043	0.032	NA	NA
20	NA	0.870	NA	-0.130	0.015	NA	NA
21	NA	0.870	NA	-0.135	0.015	NA	NA
22	NA	0.870	NA	-0.135	0.022	NA	NA
23	NA	0.870	NA	-0.138	0.024	NA	NA
24	NA	0.870	NA	-0.141	0.016	NA	NA
25	NA	0.870	NA	-0.142	0.017	NA	NA
26	NA	0.677	NA	-0.111	0.017	NA	NA
27	NA	0.677	NA	-0.126	0.020	NA	NA
28	NA	0.677	NA	-0.118	0.021	NA	NA

1. The models are defined in Table 1.
2. Type of oxyfuel effect modeled is defined in Table 1. Reported value is change in natural logarithm of ambient CO attributable to the oxyfuel effect, keeping all other factors fixed. Corresponding percentage change is approximately given as 100 times the reported value.
3. Standard deviation not available for Models 1 to 5 where different effects are calculated for each site.
4. Minimum or maximum oxyfuel effect gives range of oxyfuel effects across groups of sites and seasons. Not reported for Models 6–25 where national oxyfuel effect is calculated for each season or for both seasons.
5. NA = not available or not applicable.

TABLE 5. Results for bimonthly mean daily maximum eight-hour average ambient CO. Model includes a linear temperature effect.

Model Number[1]	Season	R^2	Oxy Site?	Mean Oxyfuel Effect[2]	Standard Deviation[3]	Minimum Oxyfuel Effect[2,4]	Maximum Oxyfuel Effect[2,4]
1	Summer	NA	No	0.042	NA	-1.011	2.561
1	Summer	NA	Yes	-0.056	NA	-2.706	1.317
1	Winter	NA	No	-0.016	NA	-0.791	1.011
1	Winter	NA	Yes	-0.111	NA	-0.635	0.350
2	NA	NA	No	0.008	NA	-0.773	1.311
2	NA	NA	Yes	-0.095	NA	-1.687	0.400
3	NA	NA	No	0.004	NA	-0.786	1.306
3	NA	NA	Yes	-0.097	NA	-1.673	0.405
4	NA	NA	No	0.002	NA	-1.071	0.790
4	NA	NA	Yes	-0.104	NA	-1.193	0.304
5	NA	NA	No	-0.021	NA	-0.812	0.785
5	NA	NA	Yes	-0.121	NA	-0.654	0.332
6	Summer	0.120	NA	-0.071	0.039	NA	NA
6	Winter	0.297	NA	-0.062	0.029	NA	NA
7	Summer	0.121	NA	-0.083	0.040	NA	NA
7	Winter	0.297	NA	-0.062	0.030	NA	NA
8	Summer	0.121	NA	-0.136	0.050	NA	NA
8	Winter	0.299	NA	0.004	0.038	NA	NA
9	Summer	0.123	NA	-0.175	0.055	NA	NA
9	Winter	0.299	NA	0.025	0.041	NA	NA
10	Summer	0.787	NA	-0.035	0.022	NA	NA
10	Winter	0.906	NA	-0.144	0.012	NA	NA
11	Summer	0.787	NA	-0.046	0.022	NA	NA
11	Winter	0.906	NA	-0.144	0.012	NA	NA
12	Summer	0.787	NA	-0.045	0.029	NA	NA
12	Winter	0.906	NA	-0.134	0.016	NA	NA
13	Summer	0.787	NA	-0.068	0.032	NA	NA
13	Winter	0.906	NA	-0.130	0.018	NA	NA
14	NA	0.386	NA	-0.059	0.031	NA	NA
15	NA	0.386	NA	-0.062	0.032	NA	NA
16	NA	0.387	NA	0.002	0.044	NA	NA
17	NA	0.387	NA	0.022	0.048	NA	NA
18	NA	0.386	NA	-0.063	0.034	NA	NA
19	NA	0.386	NA	-0.063	0.034	NA	NA
20	NA	0.876	NA	-0.134	0.016	NA	NA
21	NA	0.877	NA	-0.138	0.016	NA	NA
22	NA	0.877	NA	-0.135	0.023	NA	NA
23	NA	0.877	NA	-0.133	0.026	NA	NA
24	NA	0.877	NA	-0.144	0.017	NA	NA
25	NA	0.877	NA	-0.145	0.018	NA	NA

1. The models are defined in Table 1.
2. Type of oxyfuel effect modeled is defined in Table 1. Reported value is change in natural logarithm of ambient CO attributable to the oxyfuel effect, keeping all other factors fixed. Corresponding percentage change is approximately given as 100 times the reported value.
3. Standard deviation not available for Models 1 to 5 where different effects are calculated for each site.
4. Minimum or maximum oxyfuel effect gives range of oxyfuel effects across groups of sites and seasons. Not reported for Models 6–25 where national oxyfuel effect is calculated for each season or for both seasons.
5. NA = not available or not applicable.

Responses to Some Critical Comments

1 The study is inconclusive because models 1 and 6-13 are not consistent with the conclusion that winter oxyfuel effect is 14 percent. In some cases the summer effects appear to be greater than the winter impact of oxyfuels.

Response: Our conclusions are based on models which fit the data with R^2 greater than at least 0.5. Model variations were performed to test the assumptions used in defining various regression models. Models which show low R^2 values are useful because they highlight poor assumptions. That is, such sets of assumptions produce regression equations that cannot properly represent the data. Because these models do not adequately describe the database neither their results nor any oxyfuel implications from such models are meaningful.

Based on the more complete analysis with the 1996 data included, the lnCO reduction in the summer for oxyfuel program years at oxysites is estimated to be between 2 and 5 percent for models 10-13, (the corresponding winter effect is between 13 and 14 percent), all of which fitted the data in both seasons reasonably well ($R^2 > 0.75$). If the summer effect can be interpreted to represent CO reductions from year-round effects unrelated to the oxyfuels program, which is an arguable hypothesis, the net winter effect would still be between 8 and 11 percent based on those models.

The individual site models (1 through 5) do not have tabulated R^2 values since each site has its own R^2, but these models were useful to compare their overall consistency with the results of the ensemble models used in the rest of the study. Taken as a whole the average results of the site models (1 through 5) are found to be consistent with the conclusion of a 14 percent ambient winter impact from oxyfuels. However, it is true that many individual site models do show strange results (e.g., large summertime oxyfuel impacts), but as noted in the main text, the strange results from these models are always statistically insignificant. (See discussion on meteorology below).

2 The summary should not use any CO emissions contribution percentage from motor vehicles that is not supported in the main text. The 62 percent mobile contribution quoted is annual, the urban winter contribution may be as high as 95 percent.

Response: We merely quote a contribution percentage published by the U.S. EPA. We see no reason to support such a number in the main text. Some critics of our study claim that 95-100 percent of urban ambient CO is from mobile emissions. We have no question that at some monitoring sites this could be true. However, our study deals with the average of all available ambient monitoring sites and the EPA number of a 62 percent mobile contribution is the average for all CO emissions.

We looked further into EPA's Emissions Trends Reports for 1994 and 1995 (available on the internet at the OAQPS web site). First of all we found that the 1995 report shows that for 331 CO monitoring sites only 55 percent are classified as urban with 45 percent suburban and 9 percent rural. The average concentrations at the suburban sites are only about 10 percent or so less than the concentrations at the urban sites for the last ten years. Also we found that for the last ten years the mobile contribution has been dropping from 67 percent down to the 62 percent for 1994. Admittedly we might have included non-road CO emissions and focused on the winter season only. The combined on-road and non-road contribution to CO emissions has apparently been stable over the past decade at about 78 percent. Strangely, the winter total mobile contribution was reported to be less than the year round contribution ten years ago, but appears to be slightly greater in 1994. The winter on-road contribution was the same as the annual on-road contribution ten years ago but the winter contribution has not shown the drop seen in the annual value noted above. Cooler temperatures can increase vehicle CO emissions, but some stationary CO sources operate primarily during the winter months (e.g., wood-burning stoves).

3 The 1996 data should be added to the winter database.

<u>Response:</u> This was done (See new main text and response 5 below on meteorology).

4 A straight line can be fitted through the oxyfuel site mean results shown in Figure 1.

<u>Response:</u> We agree that a straight line can be so fitted but its slope would clearly be steeper than the line through the non-oxyfuel sites. We are unclear as to how such a line should be interpreted. A crucial issue in interpreting our graph is that the plotted points are averages across large numbers of sites. Our models show a statistically significant oxyfuel impact when fitted to all the sites, rather than simply being fitted to just the averages. Our analysis shows that fitting separate parallel lines to the pre-oxy program and the oxy program years gives a statistically significant fit to the ambient CO data which is better than a single trend line.

One interpretation might be that no oxyfuel impact exists, but sites in the oxyfuel program somehow show a faster decline in CO concentrations (Mannino and Etzel, 1996, reported this type of result). When the 1996 data were used, this would also lead to the strange conclusion that the 42 sites which were moved from previously non-oxy sites in the database ending in 1995 to become oxy sites, suddenly had their overall rate of CO decline increased all the way back to 1986. It seems more sensible to us to assume instead that some general and apparently consistent background decline in CO exists and that this decline is essentially the same for both oxy and non-oxy sites due mainly to fleet turnover moderated by an increasing VMT, and then when an oxyfuel site is known to begin the use of oxyfuels a jump downward in CO concentrations occurs (i.e., the oxyfuel effect) but the background decline returns to essentially the same slope at a different level as long as oxyfuels remain in use.

5 There is too much influence and variability in local meteorological conditions to make meaningful conclusions without correcting for those influences. Meteorology is the single most significant factor affecting the year to year variations.

Response: Critics of our study claim that meteorology must be accounted for to produce a conclusive study of the ambient air impacts on CO due to the oxyfuel program. Although we agree that a follow-on study that adjusts for meteorology should go a long way towards bringing these critics into the “conclusive” camp, we will further (than what is already in the text of the report itself) describe here why we expect that the inclusion of meteorology may not produce a result substantially different than the 10-18 percent decrease in ambient CO found in our study.

First we believe that the use of the 1996 data strengthens the conclusiveness of our study. These new data not only added another year to the meteorological span, especially after 1993. The new year also leads to a large shift in the overall database by adding some 42 new sites to the oxyfuel regression equations with nearly as many sites, of course, being removed from the non-oxy equations (the numbers are not exactly the same because four of the new oxyfuel sites added had previously failed the five-years-of-data completeness criteria). All the data from these 46 new sites were included back to 1986. Yet this significant shift in the database produced results which were marginally lower, but substantially the same, for all the model variations with an R^2 greater than 0.5. That many of these new sites actually had RFG with only 2 percent oxygen (instead of the 2.7 percent oxygen required at most non-attainment CO sites) is consistent with a slightly lower average oxyfuel impact than seen previously when the 1996 data were not included.

Second the conclusiveness of statistics-based studies rely on randomness and independence assumptions, and we believe there is little evidence to support the claim that the meteorological effects are so correlated with the oxyfuel effects that they significantly bias the results towards too high of an implied oxyfuel impact. Critics have noted that most oxyfuel programs were initiated between 1992 and 1993 and that 1992 meteorology tended to elevate CO levels while 1993 meteorology tended to suppress CO. Figures 1 and 1a in our report clearly show that the mean observed CO concentrations for 1992 were above the trend line for our Model 12 and that the 1993 observations were much lower than the regression line. Yet it can be seen in these same figures that the mean 1991 and 1994 data were nearly equivalent to the 1992-1993 pair but apparently providing a bias in the opposite direction.

A regression model uses a computer to find the line (consistent with the assumptions given) which gives the best fit to the data. Every data point is given equal weight along the regression line. That is, a 1994 data point counts as much as a 1993 data point even if the oxyfuel program was initiated between 1992 and 1993. For example, in Figure 1 the 1993 data can be thought of as pulling the regression lines down just as much as the 1994 data might be considered to be pulling the regression lines up. As a sensitivity test, we selectively removed all the 1992 and 1993 data (but kept all other data including the 1996 data). Since these particular data by themselves clearly pull the regression lines toward a higher implied oxyfuel impact, the expected result is a lower implied value for the oxyfuel effect. Indeed, this intentional biasing of the data reduced the computed oxyfuel impact down to approximately 10 percent from the approximately 13 percent seen when the 1996 data were included. Judging by Figure 1a we would expect a similar increase to about 16 percent if we were to selectively remove the 1991 and 1994 data that clearly bias the implied oxyfuel effect in the opposite direction to the 1992 and 1993 data. The point here is that there appears to be enough "balance" in the data to reduce the probability that meteorology will be so non-random and so correlated with the oxyfuel sites that a significantly different oxyfuel result than 10-18 percent would be obtained.

Accounting for meteorology should be advantageous for some of the models used in this study. In particular, models 1 through 5 almost always produced results which were not statistically significant, apparently due to meteorological "noise," and only 9 or 10 data points per model. These models were for individual sites. Having statistically valid local models might be useful to nearby communities to evaluate their own local need for continued oxyfuel use. If several sites are within the same oxyfuel program then suppression of the meteorological variance in a "cluster-of-sites" model should work even better than the individual site models. For the models in our study that already show good representation of the full database (i.e., $R^2 > 0.5$), accounting for meteorological variances may show even better representations. Moreover, accounting for the meteorological variance can be expected to produce a narrower 95 percent confidence interval of results than the present range of 10-18 percent that is seen when meteorological variations are ignored. Meteorological effects will not bias our estimates unless they are correlated with the trend or oxyfuel effects.

6 The EPA CO monitors may not be reliable in the 1-2 ppm range.

Response: We addressed this potential problem in our study. Footnote 2 on page 3 states that subtracting 0.2 ppm from all hourly data greater than 0.2 produced no significant change in results. The results based on 24 hour averages (with or without the background subtraction) were essentially the same as the results based on the maximum daily high 8 hour values. If the monitors were in fact generating noise at the 1 ppm level, one would expect to see differences between the results when only the highest concentrations are used compared to including (or not) much lower values. The slight changes in R^2 that were obtained are consistent with some noise in the lowest data, but these changes were not considered significant enough to impact the results. The fact that the background slope for the oxyfuel sites consistently came out virtually the same as the slope for the non-oxyfuel sites further suggests that the non-oxyfuel site data are adequately above the noise. Finally the results again came out to be similar when the 1996 data were added even though some 42 non-oxy sites were removed from the non-oxy database.