2000

Analysis of Remote Sensing Data for Development of I/M Program Evaluation Protocols

Sajal S. Pokharel

Gary A. Bishop

Donald H. Stedman

Analysis of Remote Sensing Data for Development of I/M Program Evaluation Protocols

Sajal S. Pokharel, Gary A. Bishop, Donald H. Stedman

Department of Chemistry and Biochemistry
University of Denver
Denver, CO 80208

January 2000

Prepared for:

EPA
Attn. Dr. Jim Lindner
2000 Traverwood St.
Ann Arbor, MI 48105
Order number: 9A-0633-NAEX

INTRODUCTION

Many cities in the United States are in violation of the air quality standards established by the Environmental Protection Agency (EPA). Carbon monoxide (CO) levels become elevated primarily due to direct emission of the gas, and ground-level ozone, a major component of urban smog, is produced by the photochemical reaction of nitrogen oxides (NO_x) and hydrocarbons (HC). As of 1996, on-road vehicles were the single largest source for the major atmospheric pollutants, contributing 60% of the CO, 29% of the HC, and 31% of the NO_x to the national emission inventory.[1]

According to Heywood,[2] carbon monoxide emissions from automobiles are at a maximum when the air/fuel ratio is rich of stoichiometric, and are caused solely by a lack of adequate air for complete combustion. Hydrocarbon emissions are also maximized with a rich air/fuel mixture, but are slightly more complex. When ignition occurs in the combustion chamber, the flame front cannot propagate within approximately one millimeter of the relatively cold cylinder wall. This results in a quench layer of unburned fuel mixture on the cylinder wall and in crevices, which is scraped off by the rising piston and sent out the exhaust manifold. With a rich air/fuel mixture, this quench layer becomes more concentrated in HC, and thus more HC is sent out the exhaust manifold by the rising piston. There is also the possibility of increased HC emissions with an extremely lean air/fuel mixture, when a misfire occurs and an entire cylinder of unburned fuel mixture is emitted into the exhaust manifold. Nitric oxide (NO) emissions are maximized at high temperatures when the air/fuel mixture is slightly lean of stoichiometric, and are limited during rich combustion by a lack of excess oxygen and during extremely lean combustion by low flame temperatures. In most vehicles, practically all of the on-road NO_x is emitted in the form of NO.[2] Properly operating modern vehicles with three-way catalysts are capable of partially (or completely) converting engine-out CO, HC and NO emissions to CO_2, H_2O and N_2.[2]

The remote sensor used in this study was developed at the University of Denver for measuring the pollutants in motor vehicle exhaust, and has previously been described in the literature.[3,4] The instrument consists of a non-dispersive infrared (IR) component for detecting carbon monoxide, carbon dioxide (CO_2), and hydrocarbons, and a dispersive ultraviolet (UV) spectrometer for measuring nitric oxide. The source and detector units are positioned on opposite sides of the road in a bi-static arrangement. Colinear beams of IR and UV light are passed across the roadway into the IR detection unit, and are then focused onto a dichroic beam splitter, which serves to separate the beams into their IR and UV components. The IR light is then passed onto a spinning polygon mirror, which spreads the light across the four infrared detectors: CO, CO_2, HC and reference.

The UV light is reflected off the surface of the beam splitter and is focused into the end of a quartz fiber-optic cable, which transmits the light to an ultraviolet spectrometer. The UV unit is then capable of quantifying nitric oxide by measuring an absorption band at

226.5 nm in the ultraviolet spectrum and comparing it to a calibration spectrum in the same region.

The exhaust plume path length and density of the observed plume are highly variable from vehicle to vehicle, and are dependant upon, among other things, the height of the vehicle's exhaust pipe, wind, and turbulence behind the vehicle. For these reasons, the remote sensor can only directly measure ratios of CO, HC or NO to CO_2. The ratios of CO, HC, or NO to CO_2, termed Q, Q' and Q'' respectively, are constant for a given exhaust plume, and on their own are useful parameters for describing a hydrocarbon combustion system. This study reports measured emissions as %CO, %HC and %NO in the exhaust gas, corrected for water and excess oxygen not used in combustion. However, these percent emissions can be directly converted into mass emissions by the equations shown below.

$$\text{gm CO/gallon} = 5506 \cdot \%\text{CO}/(15 + 0.285 \cdot \%\text{CO} + 2.87 \cdot \%\text{HC})$$
$$\text{gm HC/gallon} = 8644 \cdot \%\text{HC}/(15 + 0.285 \cdot \%\text{CO} + 2.87 \cdot \%\text{HC})$$
$$\text{gm NO/gallon} = 5900 \cdot \%\text{NO}/(15 + 0.285 \cdot \%\text{CO} + 2.87 \cdot \%\text{HC})$$

These equations indicate that the relationship between concentrations of emissions to mass of emissions is quite linear, especially for CO and NO and at typical concentrations of HC. Thus, the percent difference in emissions calculated from the concentrations of pollutants reported here are equivalent to a percent difference in the mass emissions of the pollutants.

Another useful conversion is from percent emissions to g pollutant per kg of fuel. This conversion is achieved directly by first converting the pollutant ratio readings to the moles of pollutant per mole of carbon in the exhaust from the following equation:

$$\frac{\text{moles pollutant}}{\text{moles C}} = \frac{\text{pollutant}}{\text{CO} + \text{CO}_2 + 3\text{HC}} = \frac{(\text{pollutant}/\text{CO}_2)}{(\text{CO}/\text{CO}_2) + 1 + 3(\text{HC}/\text{CO}_2)}$$

Next, moles of pollutant are converted to grams by multiplying by molecular weight (e.g. 44 g/mole for HC since propane is measured), and the moles of carbon in the exhaust (the denominator) are converted to kilograms by multiplying the denominator by 0.014 kg of fuel per mole of carbon in fuel, assuming gasoline is stoichiometrically CH_2.

Quality assurance calibrations are performed twice daily in the field unless observed voltage readings or meteorological changes are judged to warrant more frequent calibrations. A puff of gas containing certified amounts of CO, CO_2, propane and NO is released into the instrument's path, and the measured ratios from the instrument are then compared to those certified by the cylinder manufacturer (Praxair). These calibrations account for day-to-day variations in instrument sensitivity and variations in ambient CO_2 levels caused by local sources, atmospheric pressure and instrument path length. Since propane is used to calibrate the instrument, all hydrocarbon measurements reported by the remote sensor are as propane equivalents.

Studies sponsored by the California Air Resources Board and General Motors Research Laboratories have shown that the remote sensor is capable of CO measurements that are correct to within ±5% of the values reported by an on-board gas analyzer, and within ±15% for HC.[5,6] The NO channel used in this study has been extensively tested by the University of Denver, but we are still awaiting the opportunity to participate in an extensive blind study and instrument intercomparison to have it independently validated. Tests involving a late-model low-emitting vehicle indicate a detection limit (3ó) of 25 ppm for NO, with an error measurement of ±5% of the reading at higher concentrations. Appendix A gives a list of criteria for valid or invalid data.

The remote sensor is accompanied by a video system to record a freeze-frame image of the license plate of each vehicle measured. The emissions information for the vehicle, as well as a time and date stamp, are also recorded on the video image. The images are stored on videotape, so that license plate information may be incorporated into the emissions database during post-processing. A device to measure the speed and acceleration of vehicles driving past the remote sensor was also used in this study. The system consists of a pair of infrared emitters and detectors (Banner Industries) which generate a pair of infrared beams passing across the road, 6 feet apart and approximately 2 feet above the surface. Vehicle speed is calculated from the time that passes between the front of the vehicle blocking the first and the second beam. To measure vehicle acceleration, a second speed is determined from the time that passes between the rear of the vehicle unblocking the first and the second beam. From these two speeds, and the time difference between the two speed measurements, acceleration is calculated, and reported in mph/s. Appendix B defines the database format used for a typical data set.

Remote Sensing Devices (RSD) are able to obtain independent I/M program evaluation information, since actual on-road emissions can be monitored. The first study evaluates the RSD data collection and analysis methods. Several experiments were proposed for this purpose. We have investigated the correlation between on-road RSD data and local IM240 program data when the fleet is averaged by model year. To the extent that linear correlations are obtained then proportionate effects shown by RSD can be scaled directly to IM240. To the extent these correlations pass through the origin then the percent effects are the same. Null experiments were conducted; one in which two cities with similar socioeconomic and fleet characteristics are compared in order to quantify the residual uncertainties due to vehicle manufacturer, prior fueling practices, differences in vehicle type, driving, maintenance practices, etc. In the other null experiment, the RSD data set from a certain location and year is randomly divided into two fleets, and the emissions of these two fleets are compared in order to quantify the variability in RSD fleet measurements. A second study, what is termed the "step method", estimates the effect of a change in an I/M program by measuring fleet average emissions using RSD data from a point half way through the program change. The third study, called the LBNL method, tracks remote sensing over time so that a change in fleet emissions after the I/M test can be measured. In what we have termed the impact study, analysis of international data sets are performed to determine the potential benefits of I/M programs even in the absence of emission control technology on the fleet.

RESULTS AND DISCUSSION

Correlation studies

To measure the correlation between on-road RSD data and local IM240 program data, pollutant measurements from both methods were converted to common units, namely g of pollutant per kg of fuel. The conversion of RSD measurements to these units is described above. IM240 data, including all "fast pass" estimates, are reported in g/mi, which is easily converted by multiplying first by the miles per gallon of the vehicle (measured during IM240) and then by the inverse of the density of gasoline (0.33 gal/kg).

This analysis showed that fleet averaged on-road remote sensing data correlate very well versus fleet average IM240 data. We have demonstrated this with three data sets from Denver: RSD January 1999, RSD January 1997 and RSD January 1996 correlated versus IM240 for the whole year in 1998, 1996 and 1995, respectively. The figures (1-3) show average emissions for each measured model year. There are many more cars in the newest model years. The plots illustrate that, though the slopes of the correlations are not all one, the relationships are mostly linear. Furthermore, the reproducibility of the excellent correlations (r^2 in every case greater than 0.95) during the three separate years of study is evident.

In each of the three years of study, the whole of the IM240 database, including the calculated FAST-PASS emissions, was used. Thus, in each case, that gave approximately 1,000,000 IM240 measurements. The remote sensing data consisted of about 25,000 measurements in 1999 and 1996 and about 35,000 in 1997.

There is a slight curvature in the NO data which may be due to temperature effects. Since all RSD data were obtained during the winter when ambient temperatures are low and oxygenated fuel is mandated in Denver, a correlation study was done with the RSD data and IM240 data from January and the first half of February. In this way the temperature and oxy-fuel difference is accounted for, and the curvature of the NO correlation plot is diminished (Figure 4).

The CO plots show negligible intercepts. The HC and NO plots do show an intercept. The intercept does not detract from the excellent correlations but does mean that the relationship needs to be treated with this intercept in mind for each species separately. The intercepts may arise from different driving modes or from a remote sensing measurement offset, which applies to all vehicles regardless of emissions or model year.

Similar correlation plots were constructed using IM240 and RSD data from Phoenix, Arizona in 1998. The resulting slopes and intercepts for the three pollutants were similar to those in the three years in Colorado, including the significant offset for HC. The r^2 values were somewhat lower, however, and this may be attributed to the smaller sample number in both the RSD data set (N=12,000) and in the I/M data set, which consisted of a 2% random sample of the entire Phoenix 1998 set.

Null Experiments

Two Similar Cities

The first of the null experiments is a comparison of two cities with similar fleet characteristics. The two cities would need to have similar socioeconomic and fleet characteristics and the same I/M program for the same amount of time. Denver and Phoenix is a pair of cities that may fall into this category. The optimal conditions would be to take readings at similar locations in the two cities so that the driving mode and load on the vehicles are the same. In the absence of this condition, however, one may group the emissions as a function of vehicle specific power (VSP) so that differing load and driving modes is, at least approximately, accounted for. VSP is a measure of load on a vehicle developed by Jose Jimenez at MIT. It is given by:

$$VSP = 4.364*\sin(\text{slope})*v + .22*v*a + 0.0657*v + 2.7*10^{-5}*v^3$$

Here, slope of the road is in degrees, velocity (v in mph) and acceleration (a in mph/s) are that of the vehicle, and the units for VSP are kW/tonne.[7]

Such a comparison was done on the RSD data sets from Denver and Phoenix in the January of 1999 and November of 1998, respectively. All measurements with valid gas emissions data, model year and VSP inputs were used. For Denver N is 16776, and for Phoenix is 7239. The data were divided into VSP and model year bins. The data indicate that Denver has consistently higher levels of CO and NO emissions but that HC emissions are noisy but higher in Denver for older vehicles and lower for newer ones. Small n values for data points had to be included because dividing the data set up by both model year and VSP yielded a large number bins, some of which contained few vehicles. The smallest n value for a single data point is 23. Data in which acceleration was less than –14 mph/s or greater than 13 mph/s were also discarded.

Figure 5 indicates the expected general trends of emissions as a function of VSP and of model year. There is a marked increase in all three gas emissions as model year gets older. Furthermore, NO increases and HC decreases with increasing VSP, while CO remains relatively constant. It can be seen that CO and NO emissions in Phoenix (light lines) are consistently lower than those in Denver (dark lines) regardless of model year or VSP. This may indicate that the I/M program in Phoenix has been more successful than in Denver. However, other factors may be affecting the data. The winter oxy-fuel in Denver differs from that in Phoenix in November. More vehicles registered outside the I/M program may drive by the RSD site in Denver than in Phoenix. Maintenance practices may differ in the two cities, and average temperatures during the measurement periods were approximately 60 °F in Phoenix and 45 °F in Denver.

An attempt was made to correct for differences in vehicle specific power and fleet age between the two cities. The question asked was, "What would Phoenix's emissions look like had it had Denver's fleet?" In order to answer this question, the Phoenix average emissions for each model year and VSP bin was multiplied by the number of vehicles seen in Denver for that particular bin. The products for every bin were then added together and divided by the total number of vehicles measured in Denver. The resulting average emission is then the Phoenix average emission given a fleet and driving pattern (VSP) identical to that of Denver. Such an analysis showed that average CO and HC emissions in Phoenix would not change significantly had the city had a fleet the same age and with the same driving pattern as that of Denver. However, fleet average NO emission would increase by 38% making it closer to the actual Denver average.

A similar comparison was done with the Austin, Texas data set from July 1998 versus, again, Denver from January 1999. The results (see Figure 6) differ from the Phoenix comparison. There are slightly higher CO emissions in Austin, and the difference occurs mostly at higher VSP. The HC and NO data, on the other hand, show much greater differences between the two cities, with Austin vehicles having significantly higher levels of hydrocarbons and lower levels of NO. Austin has no I/M program nor oxygenated fuels; however, the measured data in Austin were also obtained at 90-110° F temperatures compared to Denver's 25-60° F.

This effect of temperature and oxygenated fuels has also been investigated. The Denver IM240 1998 data were analyzed for differences in emissions between oxygenated and non-oxygenated fuels as a function of ambient temperature. Since oxygenated fuel is used in Denver from mid-November to mid-February, oxygenated fuel measurements consisted of all IM240 tests occurring after November 30th and before February 15th while non-oxygenated fuel measurements are those taken after February 28th and before November 15th. The data set contained 825,191 tests of which 152,728 were oxy fuel within the temperature range analyzed and 514,555 were non-oxy fuel also within the temperature range. There was not a significant fleet age difference between the oxygenated and non-oxygenated fuel fleets. The oxy fuel 25 to 35 °C data point consisted of the fewest number of tests with only 2475 readings because during the oxy fuel season the ambient temperature rarely went above 25 °C. Plots for the three pollutants measured are in Figure 7. The CO plot shows the greatest effect with a significant slope in emissions as a function of temperature for both oxy and non-oxy fuels. Furthermore, there is a 11±0.8%[i] difference in average emissions between oxy and non-oxy fuels, with non-oxy fuels giving higher levels of CO emissions.

Though not as discernable as with CO, HC emissions also increase with increasing temperature, and there is a 5.8±1.6% greater emission with non-oxy fuel. NO_x emissions seem to remain constant, if not decrease slightly (especially in the non-oxy fuel case), with increasing temperature. There is a significantly higher level of NOx emission (9.7±2.8%) with oxy fuel, however. Thus, the increased HC and decreased NO emissions in Austin compared to Denver are somewhat accounted for by temperature and oxy-fuel differences. In the case of CO emissions in the two cities, only a small difference is seen. The analysis is given in more detail below.

[i] All uncertainties are standard errors of the mean unless reported otherwise.

In order to correct quantitatively for temperature and oxy-fuel effects, the Denver 1998 IM240 data set was used to obtain scaling factors. Using IM240 data to scale RSD data is valid here because we have shown that IM240 and RSD fleet averaged data correlate very well (See Correlation Studies).

The same Denver 1998 IM240 data set was used to test the correlation between emission readings when the ambient temperature is low and when it is high. Low ambient temperatures were defined to be between –5 and 5 °C, and high ambient temperatures were defined to be between 30 and 40 °C. Correlation plots (Figure 8) indicate a clear effect of temperature, with increased CO and HC and decreased NOx emissions at higher temperatures. The slopes of these plots incorporate an oxy-fuel effect since from November to February, when most of the cold temperature readings occurred, oxygenated fuel is mandated in the Denver area. During March through October, when most of the warm temperature readings occurred, the fuel is not oxygenated. Thus, the slopes of these plots are appropriate scaling factors since the temperature difference between the Denver readings in January and the Austin readings in July was approximately the difference between the defined "cold" and "warm" readings in the IM240 data. The scaling factors are 1.28, 1.13 and 0.82 for CO, HC and NO, respectively. Furthermore, the incorporation of the oxy-fuel effect is appropriate since Denver vehicles were operating with oxygenated fuels at the time of the on-road testing while the Austin vehicles were not.

The connected points in Figure 9 indicate an apparent difference in average vehicle emissions in Austin and Denver by model year. There is greater CO and HC in Austin, while there is less NO. However, the temperature and oxygenated fuel differences have not been taken into account. The individual points on the plot not connected by a line represent the corrected Austin average emissions values. In order to correct the Austin data to the equivalent at lower temperatures and with oxy-fuels, the average values were divided by the correction factor – the slope of the Denver IM240 temperature correlation plots.

The corrected values for CO fall right on top of the Denver points, and a paired t-test shows that CO emissions from vehicles in Austin by model year are not significantly different from that in Denver, i.e. Austin %CO = Denver %CO±0.04. The corrected HC values, though not falling on top of the Denver values, are closer and more parallel to the Denver data. This indicates that the HC emissions profiles of the two cities are indeed similar but that an HC offset may exist in the instrumentation. This effect in the HC data has been observed before.[8] The corrected NO data are closer to the Denver values, but the data sets do not become parallel. This result indicates that there is somewhat of a difference in NO emissions between the two cities. The data have not been corrected for vehicle load. NO emissions increase with load, and this effect may contribute to the difference in NO emissions in the two cities because cars at the Austin site are at slightly higher load than ones at the Denver site.

Temperature correction was also attempted with the Phoenix data discussed previously. Since the temperature difference during the time of on-road measurement in Denver and Phoenix was only 45% of that between Denver and Austin, the scaling factors for Phoenix were obtained by reducing the Austin factors to reflect this 55% decrease in temperature difference. The resulting scaling factors are 1.13, 1.06 and 0.92 for CO, HC and NO, respectively. Applying these scaling factors caused the difference in average CO emission between the Denver and Phoenix to increase by 6% and the NO difference to decrease by 9%.

<u>Random Halves</u>

The concept in this null study is to carry out an analysis in which no detectable I/M effect is possible but also carry out the same type of analysis as in the step method (See below). Thus, the data were binned into even and odd model year pairs, then each pair divided randomly by VIN. If the last digit of the vehicle's VIN was odd it was put into one group. If it was even the vehicle was assigned to the other group. Each day's mean data were treated as independent measures of program "effect", the same method used by Stedman *et al*. 1997 and 1998.[9,10] The difference in average emission between the two fleets in each of the days was then calculated. Finally, a statistical analysis ("Descriptive Statistics" on MS Excel) was run on the values of differences to obtain the standard error, and the percent standard error of the mean was calculated. As expected, the apparent "effect" is indistinguishable from zero. What is important for this study is that for fleet sizes of twenty to thirty thousand, differences of less than 3% for CO and NO and 5% for HC will not be detectable. Table 1 summarizes the data for two cities: Denver January, 1996 and Chicago September, 1998.

Table 1: Random fleet effects

	CO	**HC**	**NO***
Denver 1996 (N=30,659)	.007±.019 %	.0006±.0011 %	-.003±.018 %
Denver (% std error of mean)	3.5	4.6	2.8
Chicago 1998 (N=22,877)	.015±.008 %	-.0009±.0015 %	-.006±.0009 %
Chicago (% std error of mean)	2.0	5.9	2.1

* Nitric oxide measurements in 1996 were made using a non-dispersive ultraviolet absorption nitric oxide channel. See Zhang *et al*.[11]

There is an improvement in the uncertainty for CO and NO as expected since the equipment was improved in subsequent years. The HC data do not show this trend, however, and this was most likely due to a faulty HC detector cooler, discovered soon after the Chicago measurements. Plots of the percent emissions of the randomly constructed fleets binned my model year are included as Figures 10 & 11. The slopes and intercepts of the correlation plots are not distinguishable from one and zero, respectively, within a 95% confidence interval. These null studies indicate that for RSD fleet sizes of order 20,000, if half of the vehicles are in a non-I/M control category and the other half in the program, fleet effects any smaller than 2-5%, depending upon pollutant, are unlikely to be observable.

The Step Method

There are several reasons for on-road emissions decreasing with the passage of time independent of an I/M program. New technology vehicles are lower emitting for a given fleet age than older technology vehicles. Depending on local and national economic factors, the fleet age itself may be changing (newer = lower emitting), and it is possible that public education/willingness to carry out required maintenance, and the auto repair industry capability are improving irrespective of the presence/absence of an I/M program. All these factors make it important not only to measure the on-road emission reductions of the I/M fleet, but also to measure the emissions of a well matched control fleet, preferably differing only in I/M status.

The step method is an on-road evaluation of new or changed I/M programs using a built-in representative control group. On-road emissions are the parameter which I/M programs are intended to control. Most I/M programs emphasize testing of fully warmed-up exhaust emissions. If I/M exhaust emissions failure is followed up by successful repair or scrapping the vehicle or relocating it to a region from which it rarely (or never) drives in the program area, then the program should show on-road exhaust emission reductions. When a new I/M program starts or when there is a major program change, then there is a window of opportunity to evaluate the effectiveness of that change in reducing on-road emissions. That window arises when the new (or changed) program has impacted about 50% of the local fleet. If an annual program starts, then the window is after about six months. In a biennial program the window is after the first year. The concept behind this evaluation is that the untested fleet serves as the representative control group for the tested fleet.

<u>How Used</u>

Colorado had various versions of decentralized idle/2500 tests since the early 1980s and switched in the Denver metro area to a biennial centralized IM240 based program on Jan.1 1995. Because the program is biennial, by January of 1996, roughly half the measured fleet (odd MY) had been through the new I/M program and the other half (even MY) had missed a year of their old annual program. On-road monitoring was carried out for five days in January of 1996 at a single heavily trafficked site. Approximately 26,000 valid, plate-matched records were obtained.

Several factors obscured the clean 50/50 split between untested even MY and tested odd MY. Many 1994 MY vehicles were tested in 1995. 1995 and later MY new vehicles obtained a four-year I/M waiver. All vehicles had to take the I/M test upon change of ownership regardless of MY. Many of these potentially confounding factors can be corrected.

Working at a freeway off-ramp, we essentially eliminated cold-start vehicle emissions. We did not control for vehicle load, reasoning: a) tightly curved uphill ramps have little off-cycle power-enrichment, b) the tested and untested MY are randomly interspersed and subject to the same loads thus making for a valid comparison independent of load.

DMV records give us county of registration, I/M eligibility and most recent I/M status (P, F, W). Individual emission data bases are not normally distributed; however, if one treats the means from each measurement day as an independent sample then these sub-samples can be analyzed using normal statistics. For a fleet of about 26,000 vehicles we found that the uncertainty in the apparent emissions benefits is $\pm$2%. This error would be reduced with a larger fleet size provided that approximate equality between tested and untested vehicles could be maintained.

Our first analysis was "eligible and certainly tested" versus "eligible in the future but not tested" giving 7$\pm$2% apparent CO benefit. We recognized that many vehicles should have been tested but were not, so a second analysis was "should have been tested" versus "not tested". This reduced the apparent benefit to 6$\pm$2%. We also measured at the one site chosen about 1300 vehicles registered in locations not required to take the I/M test. These vehicles showed higher average on-road emissions but also showed an alternation of emissions by MY as if the I/M program had caused failing vehicles to be reregistered to outlying counties but yet continue to be driven in Denver. A follow up study a year later confirmed that indeed this effect is happening and, when included for that site, reduces the apparent benefit by 2%. The contribution of these "repair avoidance" cheaters to the basin wide fleet emissions cannot be determined from one freeway interchange site, but their emissions were large enough that at the measurement site the 6$\pm$2% apparent I/M CO benefit was reduced to 4$\pm$2%. Because we could show that these vehicles were "escaped" from the I/M program it is appropriate to apply their measured emissions to the program "benefit" because they would otherwise have been in the "should have been tested" emissions category.

The same database actually allowed for two other I/M benefit tests of lower precision. Using only the even MY vehicles we evaluate the on-road emissions of those tested versus those untested. This resulted in a 5$\pm$3% difference. We were also able to evaluate the difference in on-road emissions between vehicles of all MY tested within four months before the measurement time and two months after. This analysis gave rise to an apparent 8$\pm$6% benefit for CO. On-road benefits for HC and NO were not significant.

Systematic Errors

A major advantage of a single-site, single time I/M evaluation study is that instrument calibration and vehicle load/speed are irrelevant since both fleets are subject to the same set of conditions. A second advantage is the measured and the control fleets are perfectly matched socioeconomically. A third advantage is that the evaluation can be carried out with only a single week of work to within 2% accuracy levels, and the fleet average remote sensing data has been shown to correlate very well with fleet average IM240 data (see Correlation Studies).

Two disadvantages are apparent; one that the window of opportunity is only when a new program starts up or a program change which is predicted to have measurable effect is initiated; the second is that the reference group of untested vehicles may not be a correct reference.

There is some evidence that change of ownership vehicles have higher emissions than the average of the same MY. This effect would cause the average of the untested even MY vehicles (the control group) to be biased low and thus cause an underestimate of apparent I/M benefit. It is possible to attempt to correct for this bias as Slott did in 1998.[12] He also eliminated the large sample of 1994 MY vehicles which had been tested because they were very numerous and certainly a few months older than the untested (last quarter) of the 1994 MY. These two effects both lower the apparent emissions of the untested fleet, thus increasing the apparent I/M benefit from the previous 4%-7% range to 8%-11% with the same $\pm$2% error. The last two analyses are not affected by these corrections and remain at 5$\pm$3% and 8$\pm$6% apparent IM benefit for CO.

There had been an annual I/M program in place in Denver for more than ten years. The odd MY fleet took the old test in 1994 and the new in 1995. The untested even MY fleet skipped testing in 1995 because their scheduled IM240 was in 1996. If the old program had no benefit, then this skip introduces no bias. If the old program had emissions benefits which last a long time (long repair lifetimes as in the EPA I/M model) then no bias is introduced. But, the apparent benefit is the benefit of the new program relative to the older one; not relative to a "no I/M" baseline. To the extent that repair lifetimes are not as long as modeled by EPA and the old program did lead to reduced emissions, then the skipped annual test moves the control group back toward the no I/M line, thus overestimating the I/M benefit relative to the previous program but with the upper limit being relative to no I/M.

To correct for this bias, one needs to estimate both the emission reductions from the previous (idle/2500) program and the apparent repair lifetime. This is not straightforward. If from the DMV records one can determine which tested odd MY vehicles were not changing ownership, then the even MY bias is removed and the study measures the apparent I/M benefit for the fleet which does not change ownership. This is the basis of the Slott analysis.[12]

If the I/M program causes high emitting vehicles to have more obscure license plates, paper plates, etc., then this would cause a bias in that readable DMV plates would underestimate the emissions of I/M tested cars. This bias would also cause an overestimate of I/M benefit.

To avoid sampling bias, one should eliminate all but the first measurement on vehicles measured more than once.

LBNL Method

Following this method of RSD data analysis,[13] emission measurements of a fleet of vehicles are plotted as a function of time before and after its I/M test. Thus, one can observe the trend in fleet average emissions in the weeks leading up to and after the I/M test. A decrease in overall fleet emissions after the I/M test would indicate the emission reduction concurrent with the I/M test. This difference is the apparent initial percent reduction of emissions due to the I/M program. For this type of analysis a large number of measurements and vehicles is needed. Furthermore, the RSD data cannot be confined to a certain time during the year as seasonal variability in emissions can affect the data. The Denver Smart Sign[14] data set was used for this study because it contained the largest number of measurements that spanned more than a year from 5/1/96 to 5/15/97. Only CO data were available from this set, however. They were coupled to 1996 and 1997 Denver I/M records, from which only initial test data are used. This yielded total of 37,876 Smart Sign measurements on vehicles that had I/M records in 1995 or 1996.

Figure 12, with data binned into thirty day groups excluding the I/M test day, shows a gradual increase in CO emissions in the time leading up to the I/M testing. This trend is expected as vehicles deteriorate with time so that as a vehicle gets closer to its scheduled I/M test, and further from its last one, CO emissions increase. There is an observable decrease in emission right after the I/M test which is most likely due to adjustments and repairs on the failures (and maybe even anticipated failures). The difference of 13±6 %[ii] in CO levels before and after I/M is significant; the error bars on the intercept do not intersect. The plot is somewhat noisy, however, due to non-optimal number of measurements. Data in which RSD measurement and IM test took place on the same day were discarded because the order of the two tests could not be verified easily. The slight negative slope of the line after I/M is a result of decreasing fleet age in the limited RSD data set.

An attempt was made to study the on-road emissions of vehicles which failed their first I/M test. However, the number of observations per data point went down to less than 20, and since vehicle emissions do not follow a normal distribution the noise was high enough to render the results statistically insignificant. Further measurements (more than 500,000 valid readings so that there are more than 200 observations per data point) should be taken in the Denver area in order to do additional analysis of this kind on the data set.

Average Emissions in Various Cities Relative to Denver

In one approach to RSD data analysis, it is assumed that average emissions of the newest model year vehicles will be essentially unaffected by differences among cities, such as I/M program, fuel type, driving mode, and other factors mentioned earlier. By assuming that emissions from the three newest model year cars should be relatively the same at different locations, a scaling factor is determined which when applied to all model years gives the emission distribution of a subject area relative to a reference area. Here Denver January 1999 is chosen to be the reference area. Figure 13 illustrates the results where the projected equivalent CO emissions from Austin, Phoenix and Chicago for older model year cars are all below that of Denver. Thus, different cities can have significantly different levels of emission even when many pertinent factors are accounted for by this scaling method. One problem with this method, however, is that offset differences pose a difficulty. Offsets can be caused by a small fraction of power enriched or cold start vehicles in the fleet. Such things tend to offset all model year emissions about equally[15] and are, thus, not properly taken into account by this scaling technique.

[ii] Uncertainty here is the root mean square of the standard errors of the two intercepts.

When a small offset exists between two cities, the scaling factor becomes artificially large since a small difference in a small value (which is the case for emissions from the newest cars) can be a large factor. Thus, projections for older model years can be skewed. We believe that this method can only be used if potentially offsetting variables are more closely controlled than in the data available to us.

Distributional Statistics

The emissions distribution of a fleet of vehicles is very close to a γ-distribution[16] where most vehicles emit very little, while a few high emitters contribute a large part of the overall fleet emissions. In order to judge the frequency of high emitters, the distribution of vehicles over the amount of emission can be studied. One may look at the mean fleet emission as compared to the median.

Table 2: Distribution Ratios – Mean to median and top ten percent to mean

	Mean to Median			**Top Ten Percent to Mean**		
	CO	HC	NO*	CO	HC	NO*
Denver 1996	4.8	1.6	1.5	6.4	5.8	3.8
Denver 1997	5.7	1.4	1.5	6.7	4.8	4.4
Denver 1999	5.0	1.8	2.5	6.6	6.1	4.6
Chicago 1998	2.6	1.5	2.9	6.0	5.7	4.7
Austin 1998	4.1	1.8	2.9	6.7	4.8	5.3
Phoenix 1998	4.0	2.1	3.0	7.1	6.6	5.6

* Nitric oxide measurements in 1996 and 1997 were made using a non-dispersive UV absorption nitric oxide channel. See Zhang *et al*.[11]

In each case, the mean is higher than the median resulting in a mean to median ratio greater than unity. This result indicates the presence of a few very high emitters, which raise the mean but not the median. The same conclusion is reached if one looks at the ratio of the average of the top 10% of polluters to the overall average, where the value is high enough to indicate a definite concentration of pollutants from the top 10% emitters. Again, as a tool to investigate I/M effectiveness it is not apparent that this tool has any power to discriminate.

Impact Study

An additional evaluation of an I/M program might be a study of its benefits even when there is no emissions control technology in the fleet. Such fleets are encountered in other countries. Figure 14 shows percent CO emissions as a function of model year for three locations. The filled squares are data from 1991 in Los Angeles. New vehicles have low average emissions. As the vehicles get older, the average emissions increase. Notice that there is no discernable break in 1974 or 1980 when new technologies (catalysts, 1974; closed-loop computer systems, 1980) were introduced. The line close to the L.A. data was obtained in 1991 in Sweden. Sweden introduced catalysts in 50% of vehicles 1987 and 100% in 1988. The break is clearly discernable, and Swedish catalyst-equipped cars have lower average emissions (by half) than similarly equipped vehicles in Los Angeles.

If good maintenance is even more important than catalysts, then as L.A. cars age, one might expect to see the (apparently badly maintained) catalyst-equipped cars in L.A. having higher emissions than non-catalyst cars in Sweden. This effect is observed in the 1975-81 model years. Contrasting with the lower two lines is the upper line of data from the United Kingdom in 1992. The U.K. introduced catalysts in 1990, but it is apparent the nation suffers from a combination of both poor technology and poor maintenance.

CONCLUSION

Several studies have been conducted to validate and develop RSD data collection methods. These studies indicate that on-road remote sensing is an accurate method for monitoring vehicle emissions, with a 2% margin of error on fleet averaged emissions after a week's worth of data collection. Furthermore, the correlations observed between RSD data and IM240 data indicate that RSDs can be an excellent tool for I/M program evaluation.

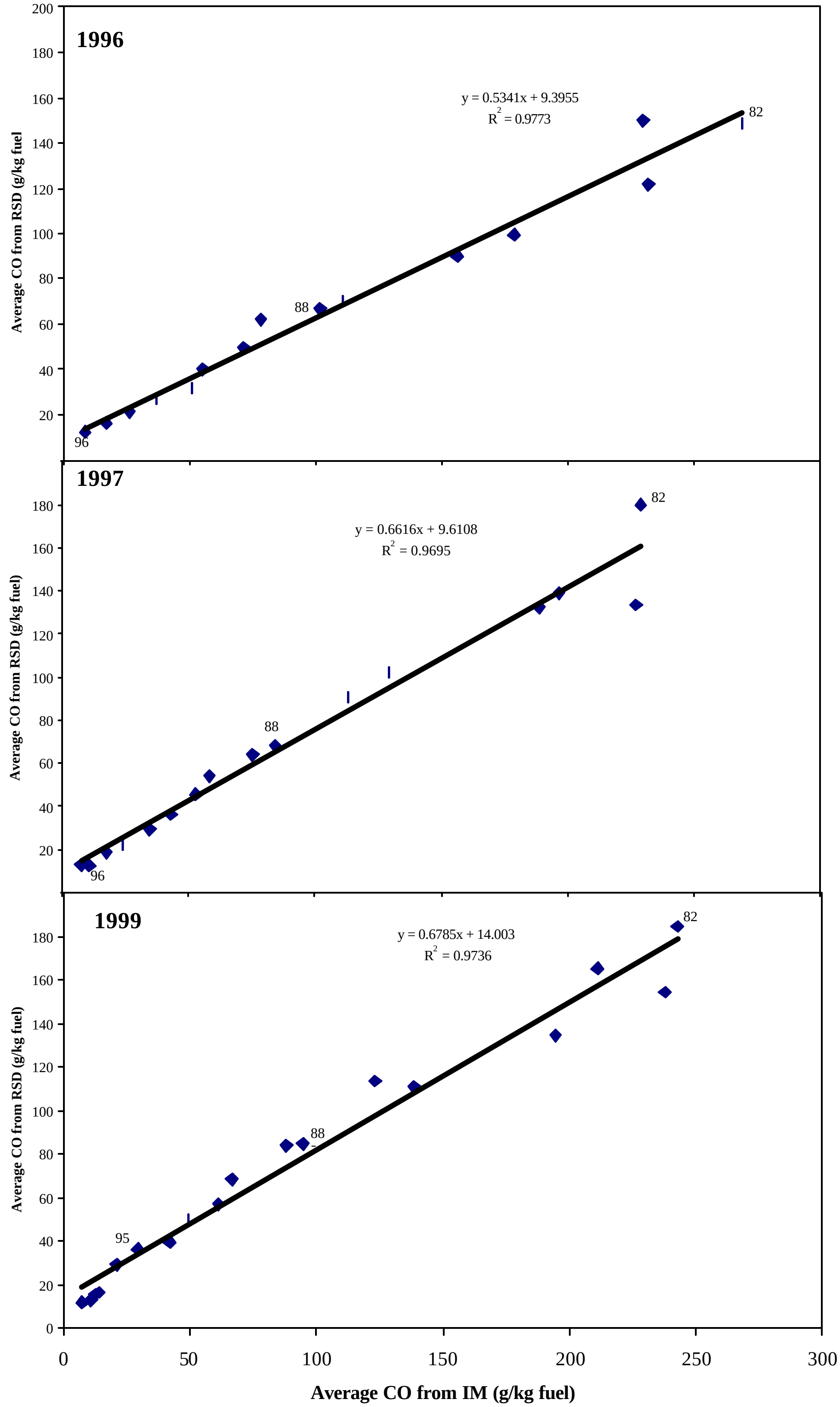

Figure 1: Correlation plots for CO between Denver IM240 and RSD for three separate years. The IM240 data are from a whole year of testing before the RSD data collection, which consisted of a week of measurements in the January of the year labeled on the plot. Each point represents a model year. Three model years are labeled.

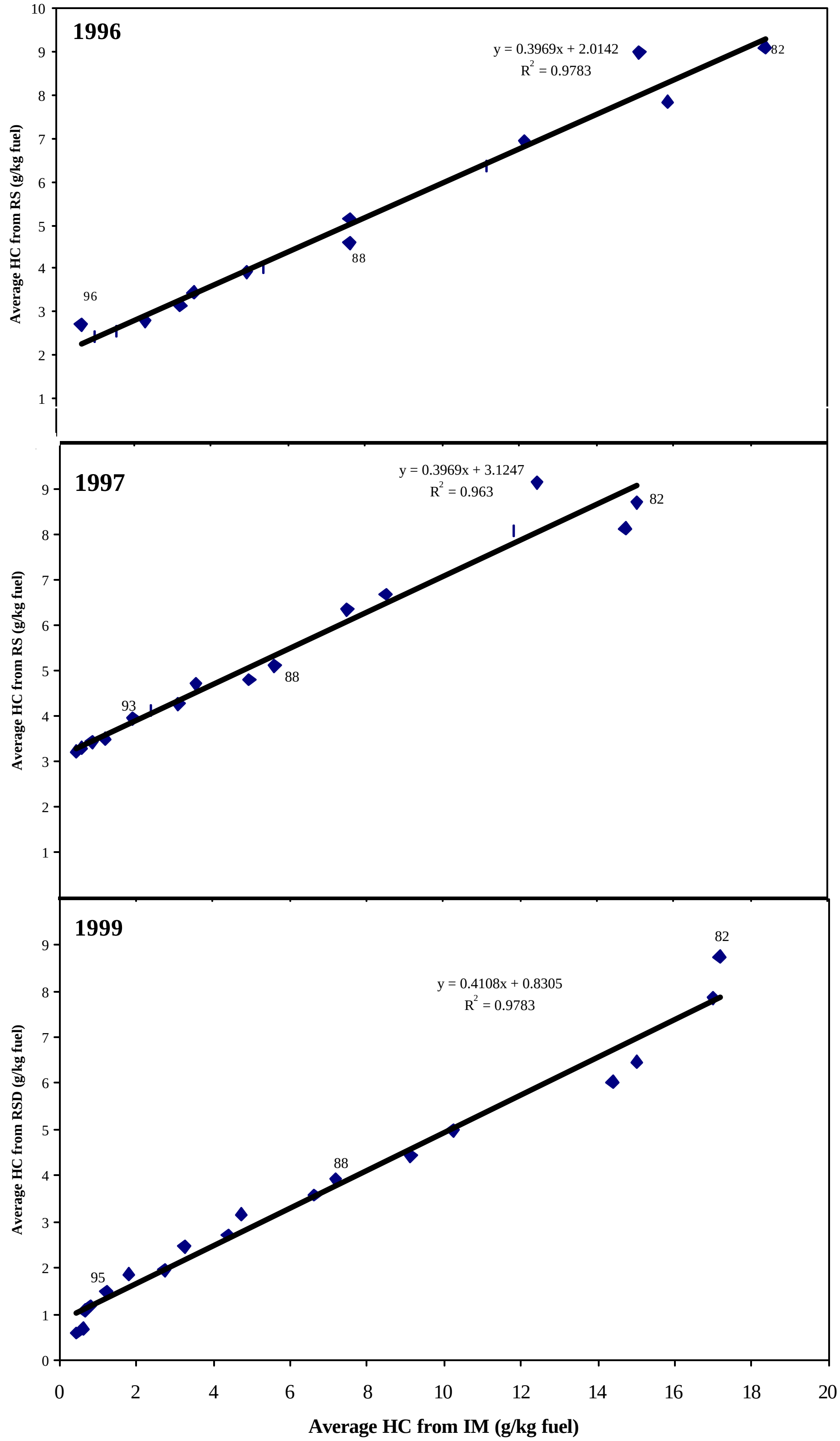

Figure 2: Correlation plots for HC between Denver IM240 and RSD for three separate years. The IM240 data are from a whole year of testing before the RSD data collection, which consisted of a week of measurements in the January of the year labeled on the plot. Each point represents a model year. Three model years are labeled.

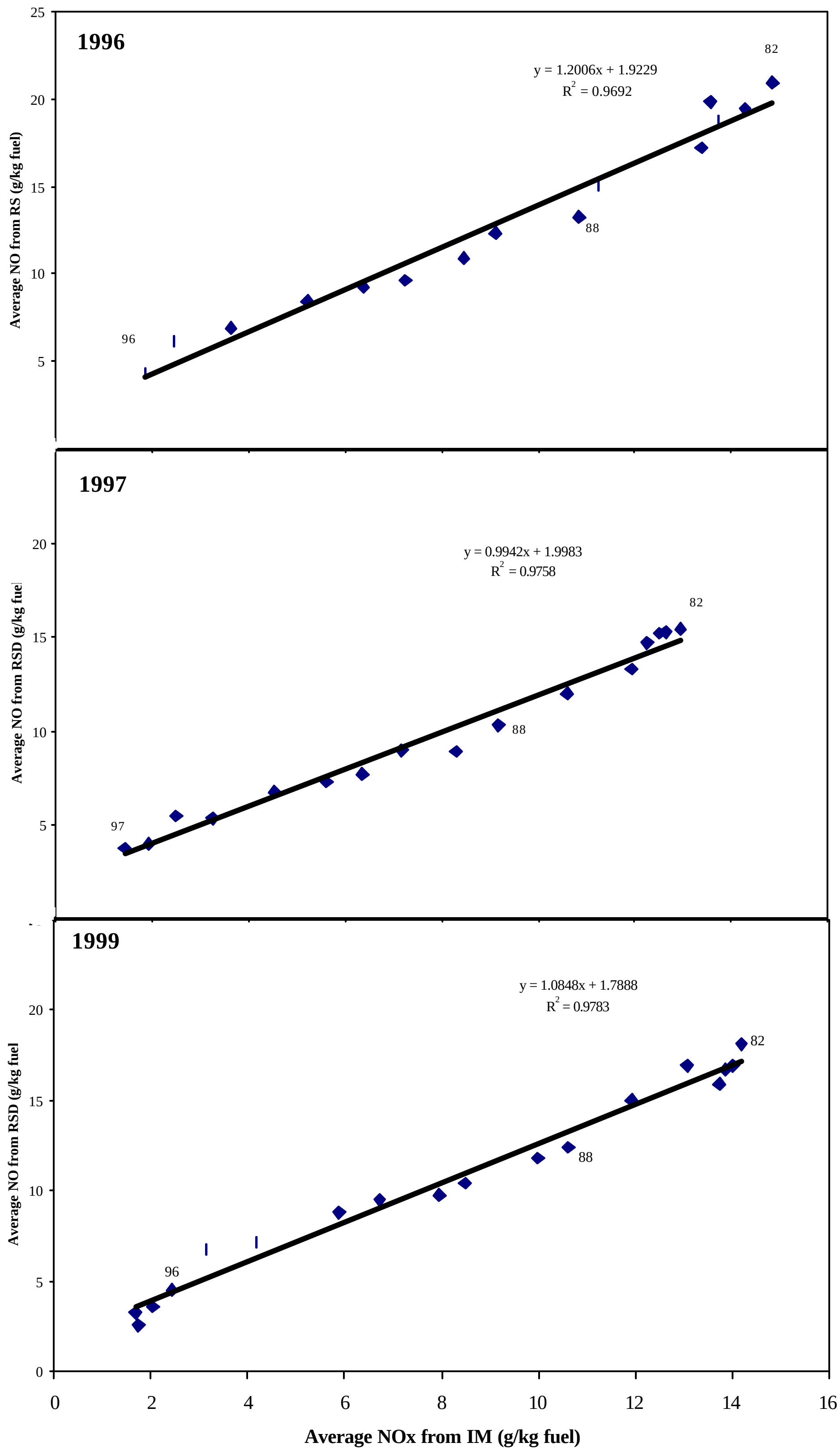

Figure 3: Correlation plots for NO between Denver IM240 and RSD for three separate years. The IM240 data are from a whole year of testing before the RSD data collection, which consisted of a week of measurements in the January of the year labeled on the plot. Each point represents a model year. Three model years are labeled.

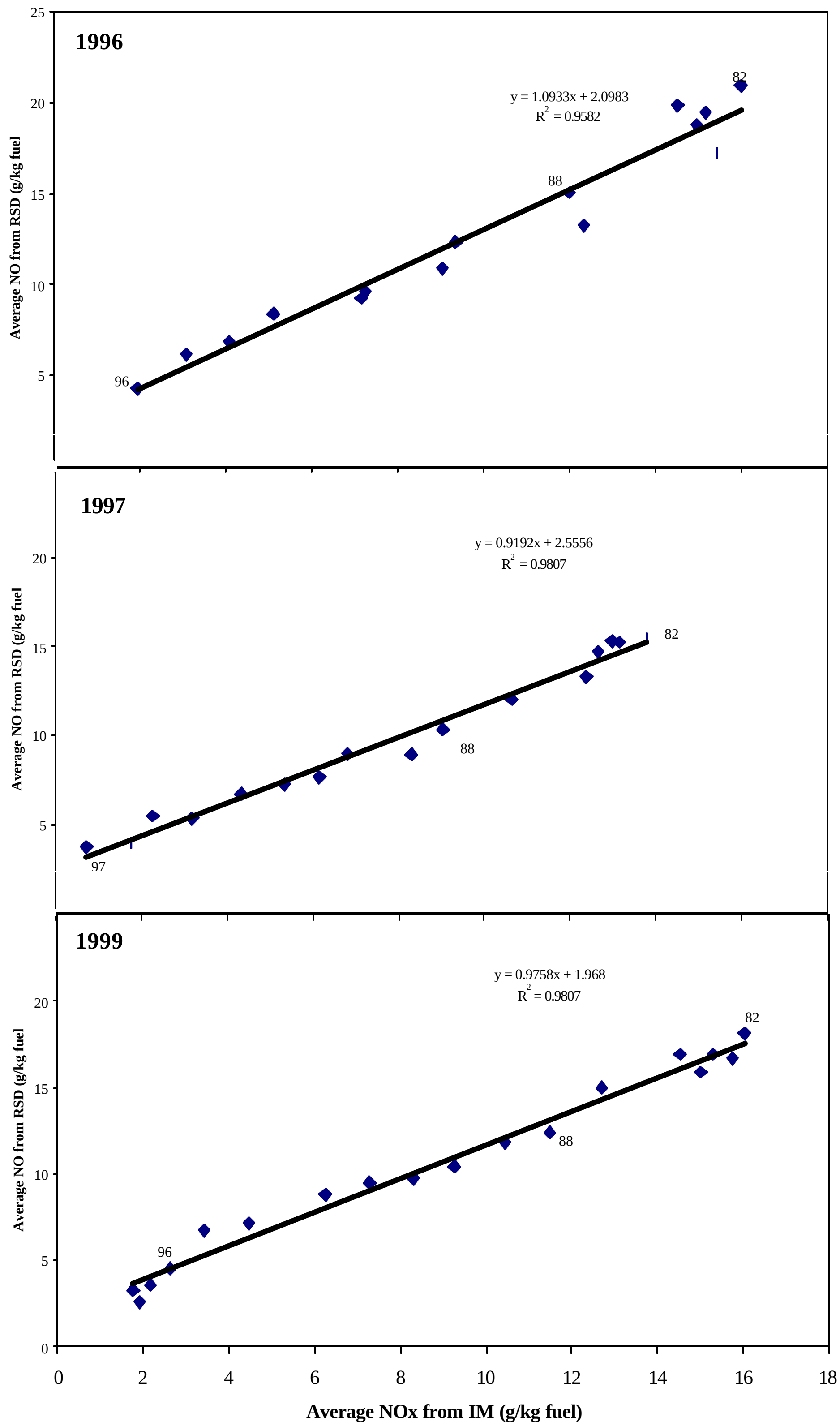

Figure 4: Correlation plots for NO during cold months between Denver IM240 and RSD for three separate years. The IM240 data are from January and February before the RSD data collection, which consisted of a week of measurements in the January of the year labeled on the plot. Each point represents a model year. Three model years are labeled.

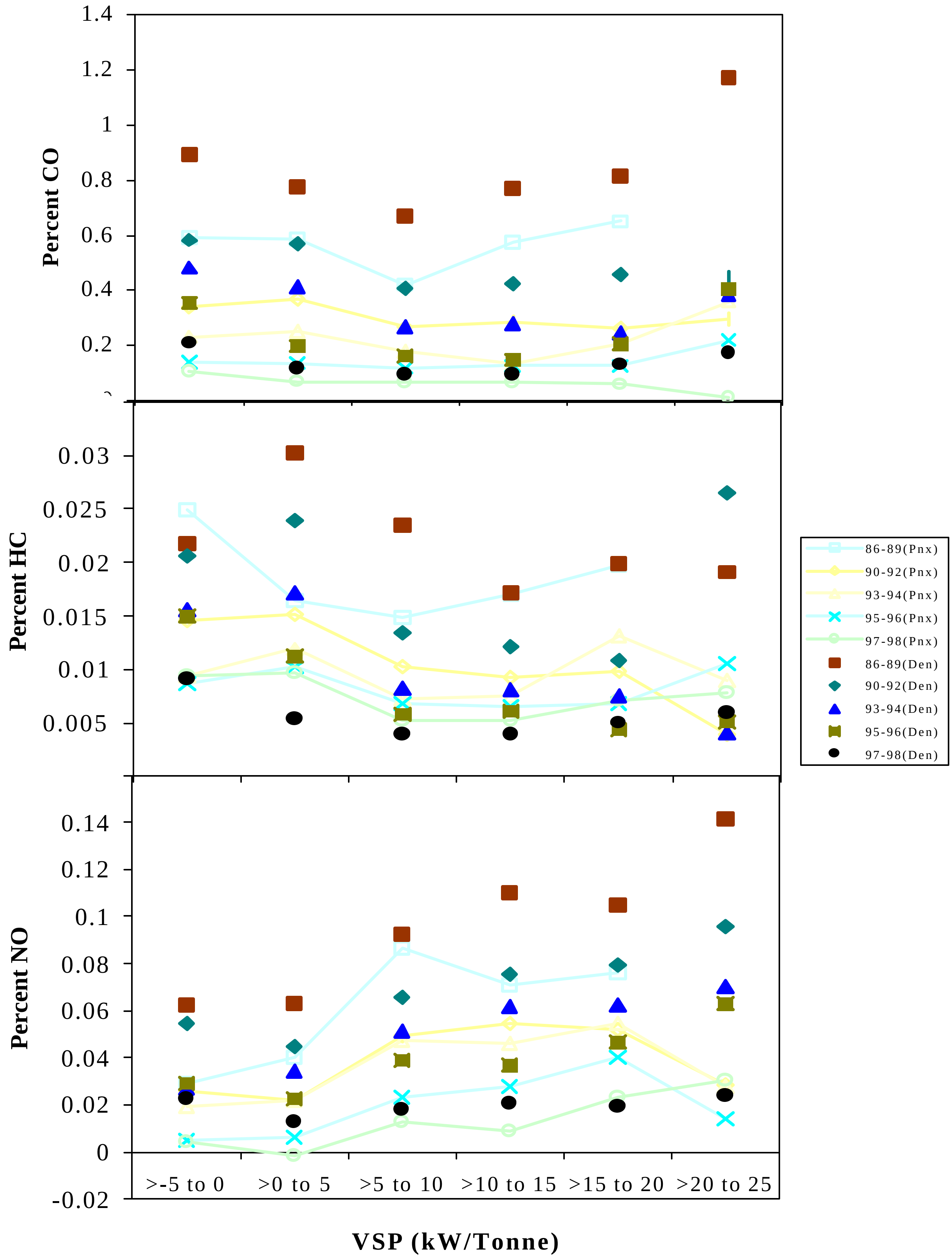

Figure 5: Comparison of emissions measured in Phoenix (November, 1998) to Denver (January, 1998). Each data point corresponds to the indicated model year and VSP bin.

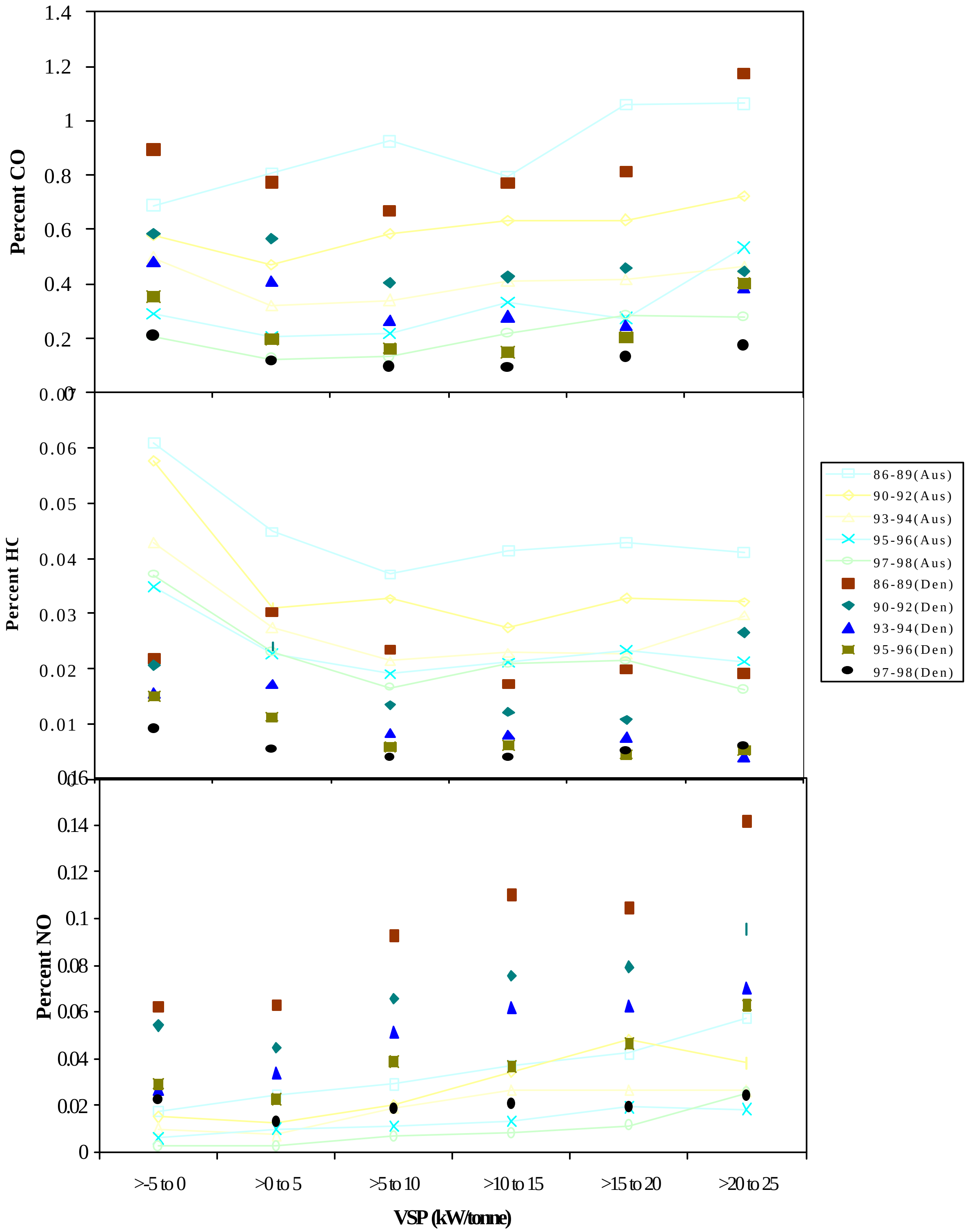

Figure 6: Comparison of emissions measured in Austin (July, 1998) to Denver (January, 1998). Each data point corresponds to the indicated model year and VSP bin.

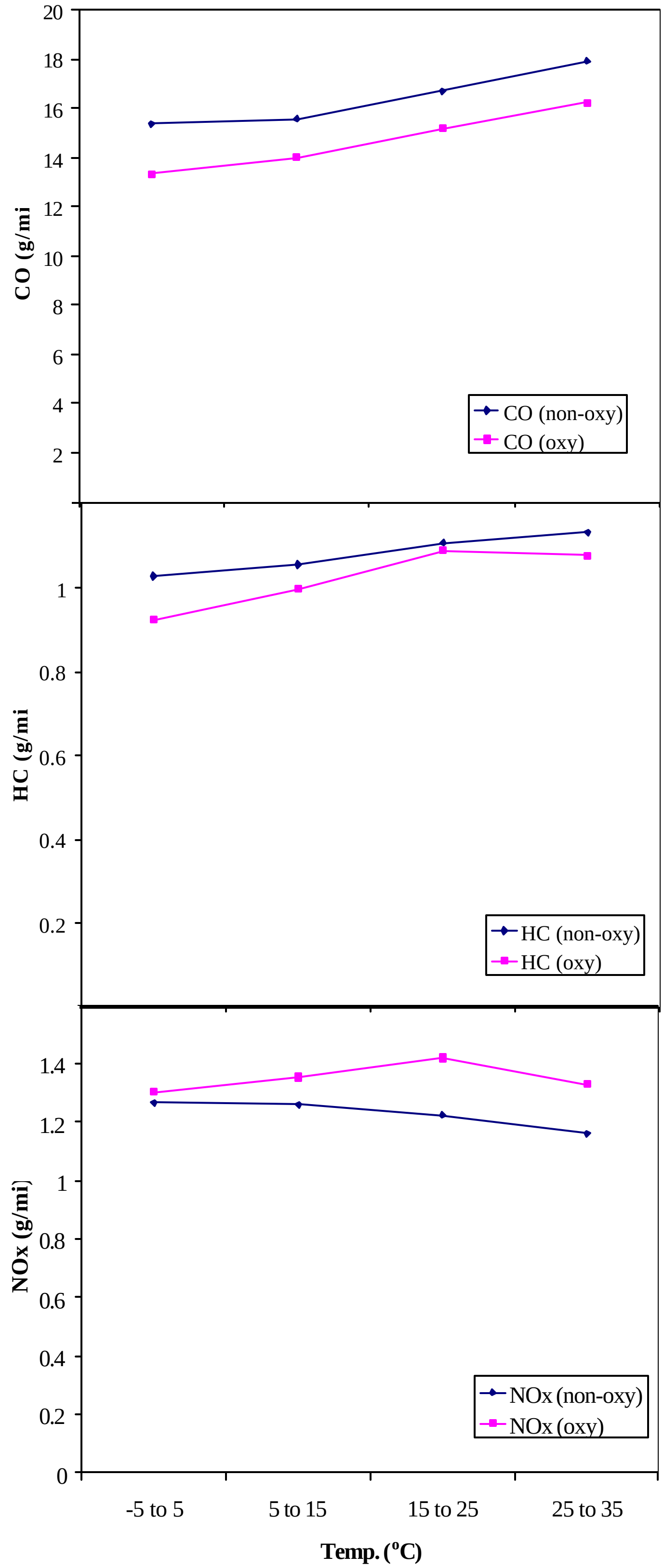

Figure 7: The effect of oxy-fuels on vehicle emissions as a function of temperature. Data are from the Denver IM240 1998 data set. Data points are in ten degree bins.

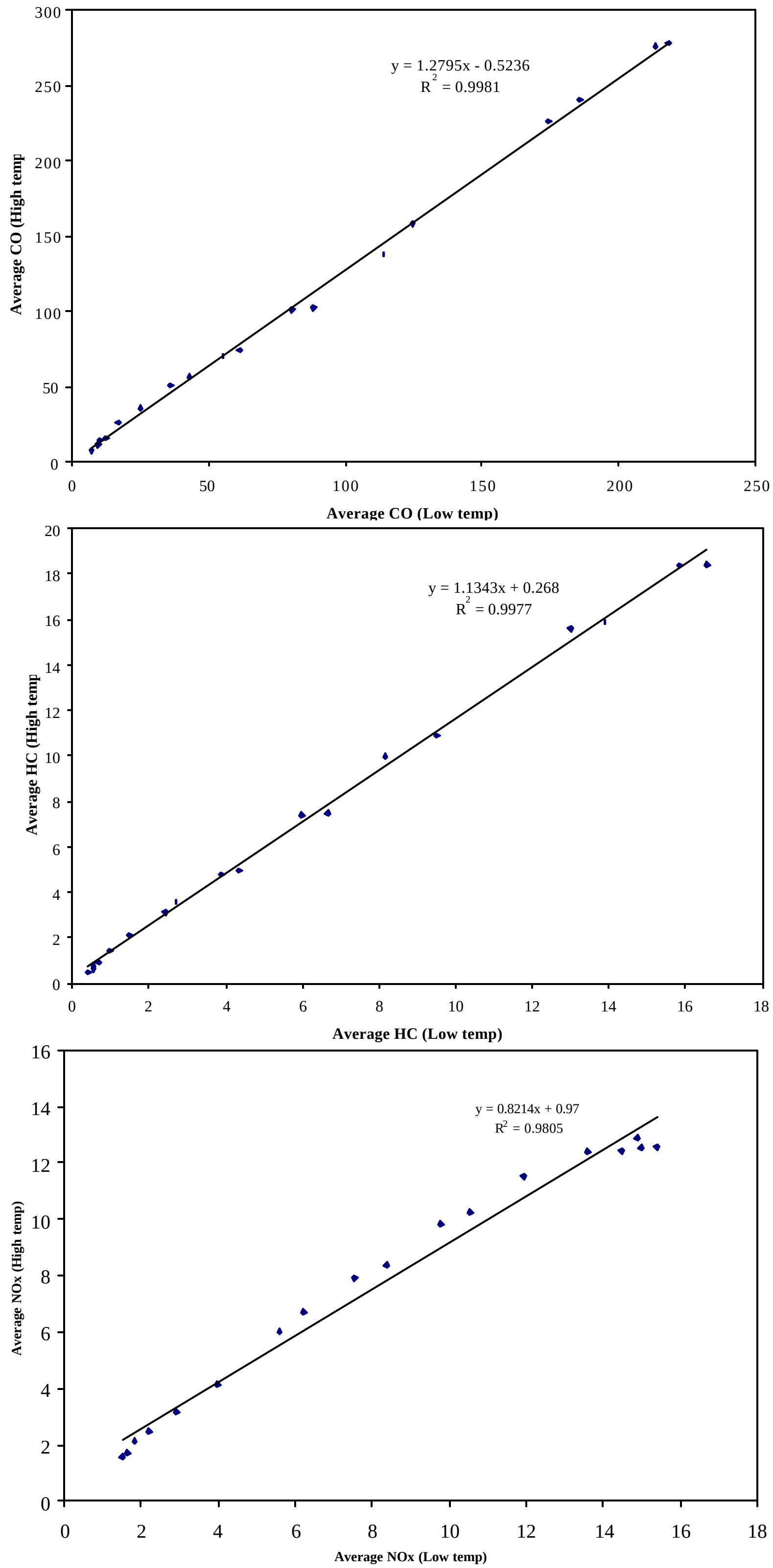

Figure 8: Correlation reported in g/kg fuel of emissions between two temperature ranges from the entire Denver 1998 IM240 data set. Each point represents a model year bin.

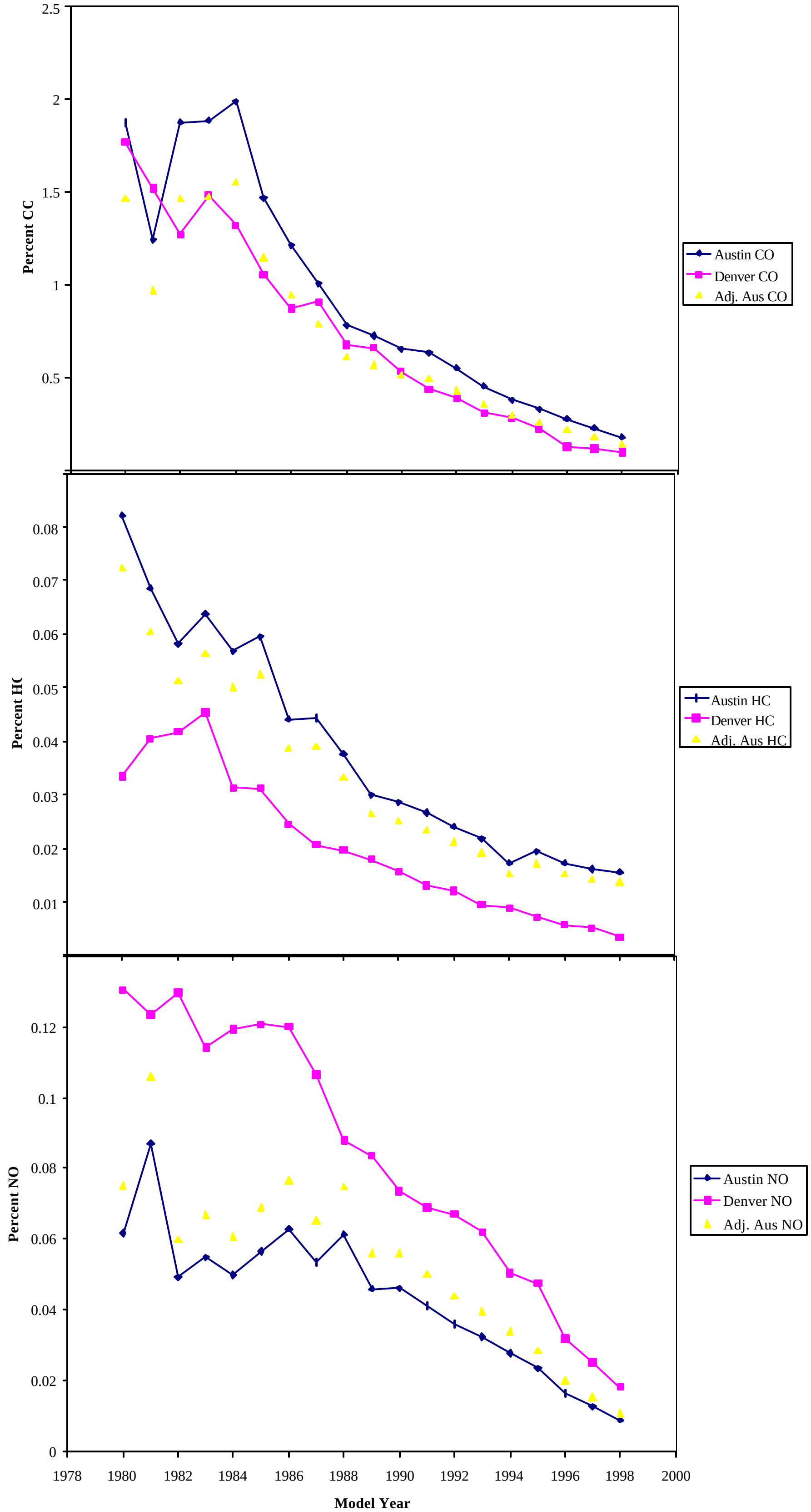

Figure 9: Average emissions by model year for Austin July 1998 and Denver January 1999 RSD data. Points not connected by a line are temperature and oxy-fuel corrected Austin data.

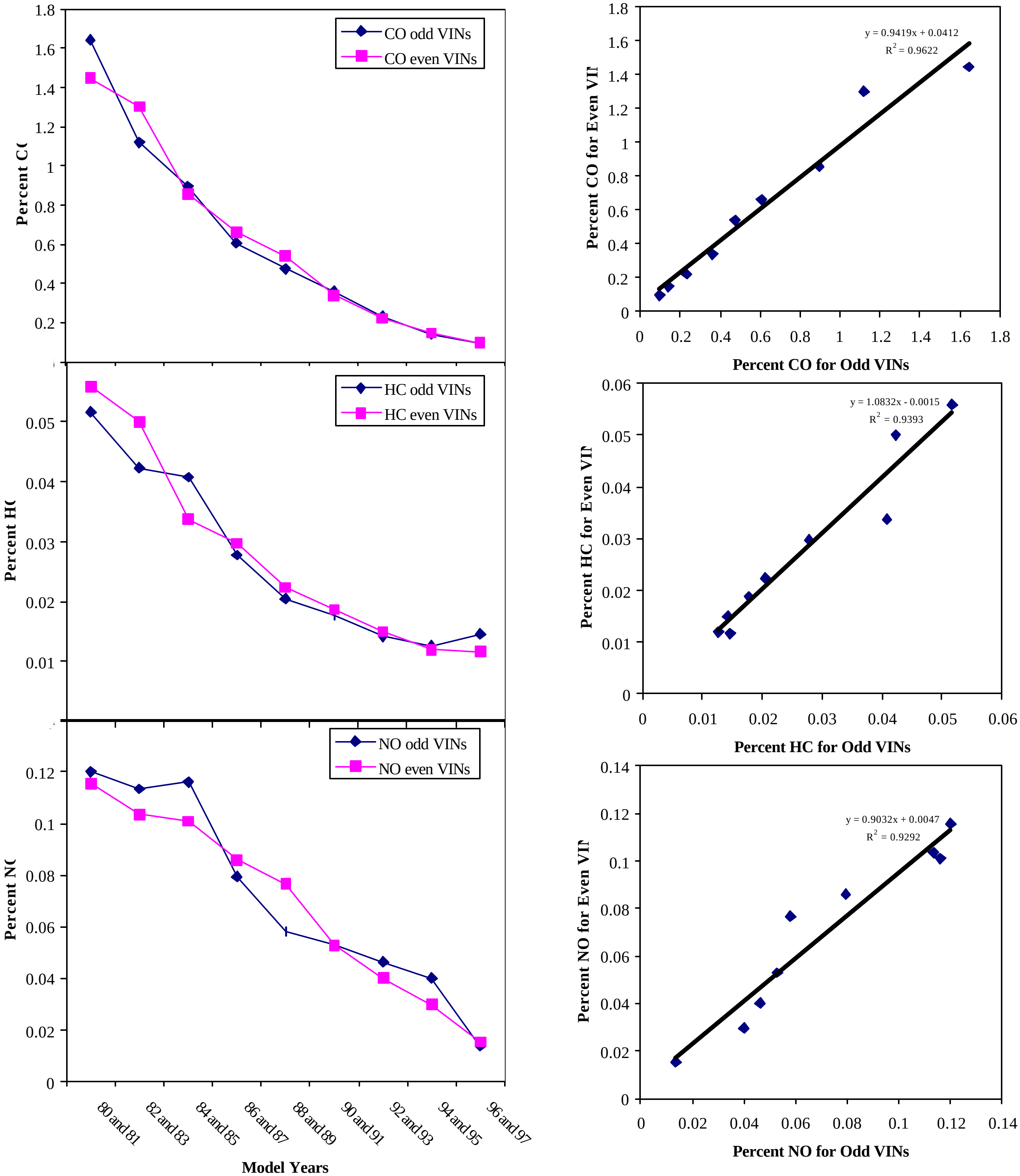

Figure 10: Comparison and correlation plots of two randomly constructed fleets from Denver RSD data (January, 1996). Each point represents one odd and even model year pair bin.

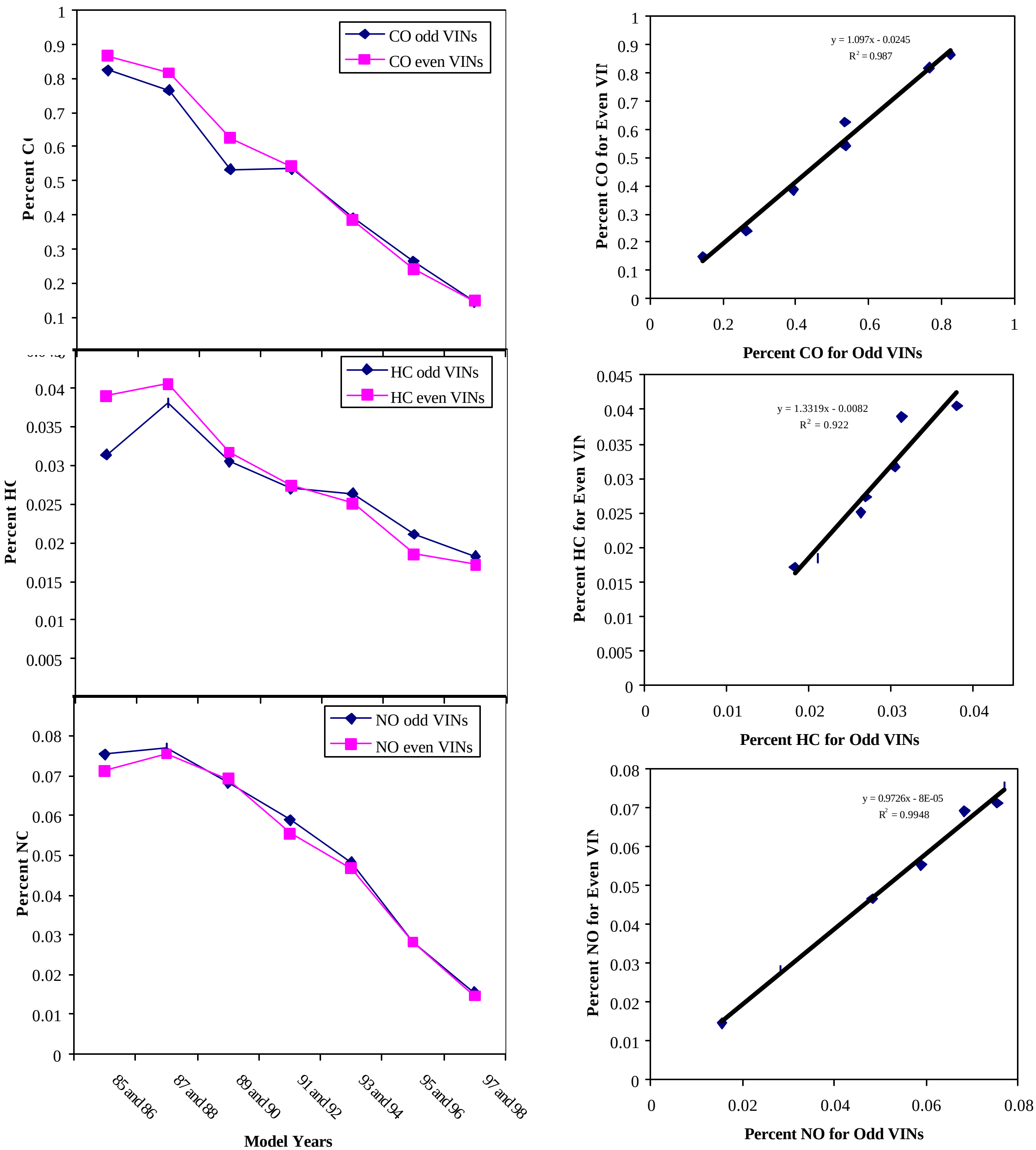

Figure 11: Comparison and correlation plots of two randomly constructed fleets from Chicago RSD data (September, 1998). Each point represents one odd and even model year pair bin.

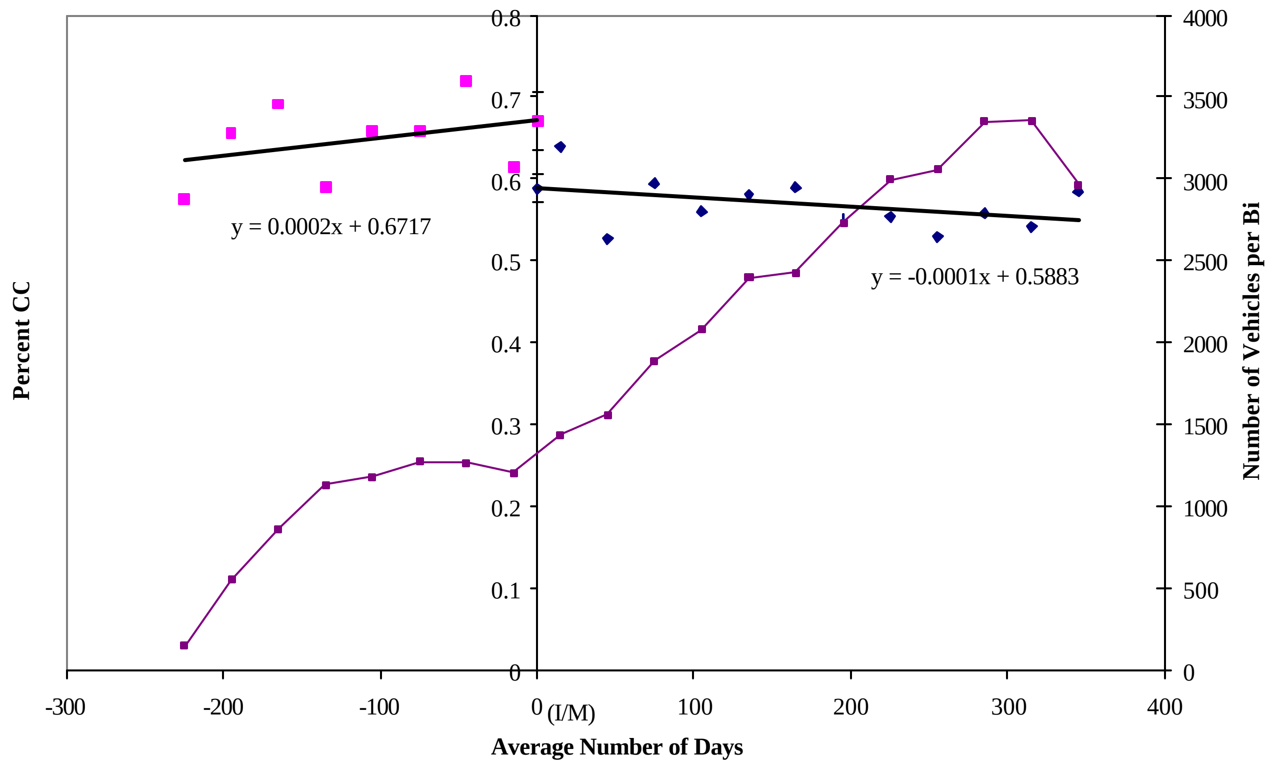

Figure 12: LBNL analysis of Denver (1996-7) Smart Sign Data. Average CO measurement from the RSD is plotted as a function of the number of days before and after the IM240 test. Data points are thirty day bins. A plot of the number of vehicles per data point is also included (right hand axis).

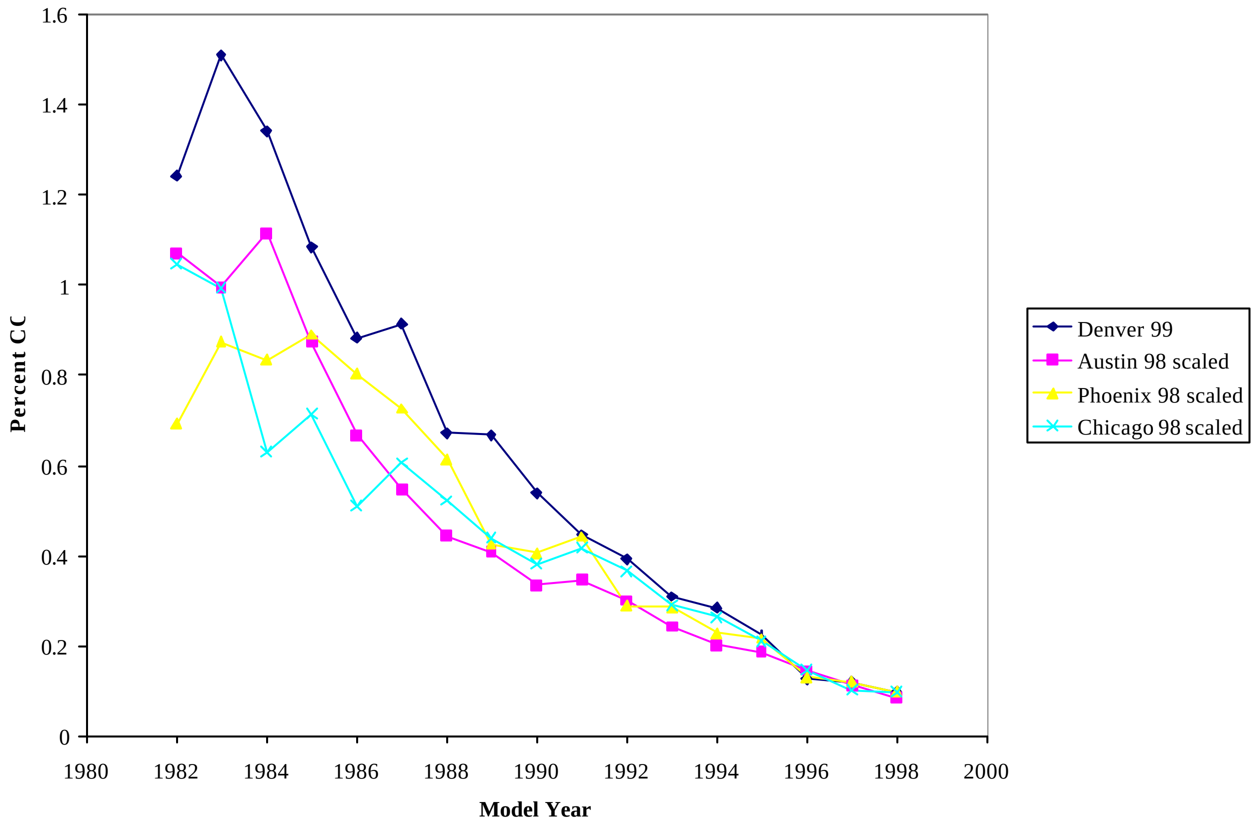

Figure 13: CO measurements from RSD for several cities. Data are scaled so that the three newest model year vehicles in other cities have the same average emissions as the fleet in Denver. Data points represent model year bins.

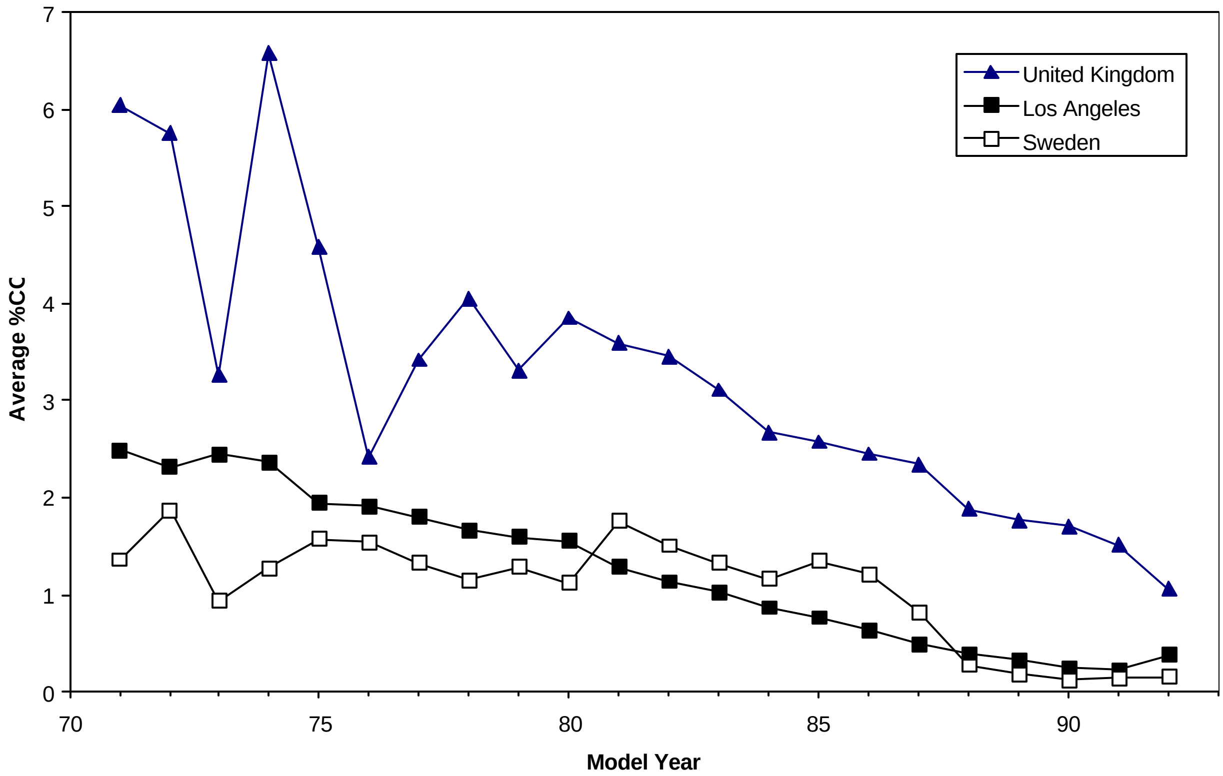

Figure 14: A graph of carbon monoxide versus model year, with results from three countries: Los Angeles, U.S.A. (1991); Gothenberg, Sweden (1991); and various locations in the U.K (1992).

REFERENCES

1. United States Environmental Protection Agency. *National Air Pollution Emission Trends, 1900-1996*. EPA-454/R-97-011. December 1997.

2. Heywood, J.B. *Internal Combustion Engine Fundamentals*. McGraw-Hill: New York, 1988.

3. Bishop, G.A.; Stedman, D.H. *Acc. Chem. Res.* **1996**, *29*, 489.

4. Popp, P.J.; Bishop, G.A.; Stedman, D.H. *J. Air & Waste Manage. Assoc*. In Press.

5. Lawson, D.R.; Groblicki, P.J.; Stedman, D.H.; Bishop, G.A.; Guenther, P.L. *J. Air & Waste Manage. Assoc.* **1990**, *40*, 1096.

6. Ashbaugh, L.L.; Lawson, D.R.; Bishop, G.A.; Guenther, P.L.; Stedman, D.H.; Stephens, R.D.; Groblicki, P.J.; Parikh, J.S.; Johnson, B.J.; Haung, S.C. "On-road remote sensing of carbon monoxide and hydrocarbon emissions during several vehicle operating conditions." Presented at Environmental Source Controls, Phoenix, AZ, 1992.

7. Jimenez, J.L.; McClintock, P.; McRae, G.J.; Nelson, D.D.; Zahniser, M.S. In *Proceedings of the 9th CRC On-Road Vehicle Emissions Workshop*, San Diego, CA, 1999.

8. Popp, P.J.; Bishop, G.A.; Stedman, D.H. "*On-Road Remote Sensing of Automobile Emissions in the Chicago Area: Year 2*." Coordinating Research Council, Inc., 1999, Interim Report.

9. Stedman, D.H.; Bishop, G.A.; Aldrete, P.; Slott, R.S. *Environ. Sci. Technol*. **1997**, *31*, 927-931.

10. Stedman, D.H.; Bishop, G.A.; Slott, R.S. *Environ. Sci. Technol*. **1998**, *32*, 1544-1545.

11. Zhang, Y.; Stedman, D.H.; Bishop, G.A.; Beaton, S.P.; Peterson, J.E.; McVey, I.F. *J. Air & Waste Manage. Assoc*. **1996**, *46*, 25-29.

12. Slott, R.S. "The use of remote sensing measurements to evaluate control strategies: measurements at the end of the first and second year of Colorado's biennial enhanced I/M program." Proceedings of the 8th CRC on-road vehicle emissions workshop. San Diego, CA. April 20-22, 1998.

13. Wenzel, T. "*Evaluation of Arizona's Enhanced I/M Program*." presentation at the 9th CRC On-Road Vehicle Emissions Workshop, San Diego, CA, April, 1999.

14. Bishop, G.A.; Stedman, D.H. "Final Technical Report for ITS for Voluntary Emission Reduction: An ITS Operational Test Using Real-Time Vehicle Emissions Detection."

15. Zhang, Y.; Stedman, D.H.; Bishop, G.A.; Beaton, S.P.; Guenther, P.L. Environ. Sci. Technol. **1996**, 30, 1445-1450.

16. Zhang, Y.; Bishop, D.A.; Stedman, D.H. *Environ. Sci. Technol*. **1994**, *28*, 1370-1374

17. Popp, P.J.; Stedman, D.H.; Bishop, G.A. *"On-Road Remote Sensing of Automobile Emissions in the Denver Area: Year 1*." Coordinating Research Council, Inc., 1999, Interim Report.

18. Bishop, G.A.; Stedman, D.H. *"On-Road Remote Sensing of Automobile Emissions in the Phoenix Area: Year 1*." Coordinating Research Council, Inc., 1999, Interim Report.

19. Popp, P.J.; Bishop, G.A.; Stedman, D.H. *"On-Road Remote Sensing of Automobile Emissions in the Chicago Area: Year 2*." Coordinating Research Council, Inc., 1999, Interim Report.

APPENDIX A: FEAT criteria to render a reading “invalid” or not measured.

Not measured:

1) apparent vehicle with less than 0.5 seconds clear to the rear. Often caused by elevated pickups and trailers causing a “restart” and renewed attempt to measure exhaust. The restart number appears in the data base.
2) vehicle which drives completely through during the 0.4 seconds “thinking” time (relatively rare).

Invalid :

1) Insufficient plume to rear of vehicle relative to cleanest air observed in front or in the rear; at least five, 10ms averages >160ppmm CO_2. Often HD diesel trucks, bicycles.
2) too much error on CO/CO_2 slope, equivalent to $\pm$20% for %CO. >1.0, 0.2%CO for %CO<1.0.
3) reported %CO , <-1% or >21%. All gases invalid in these cases.
4) too much error on HC/CO_2 slope, equivalent to $\pm$20% for HC >2500ppm propane, 500ppm propane for HC <2500ppm.
5) reported HC <-1000ppm propane or >40,000ppm. HC “invalid”.
6) too much error on NO/CO_2 slope, equivalent to $\pm$20% for NO>1500ppm, 300ppm for NO<1500ppm.
7) reported NO<-700ppm or >7000ppm. NO “invalid”.

Speed/Acceleration valid only if at least two blocks and two unblocks in the time buffer and all blocks occur before all unblocks on each sensor and the number of blocks and unblocks is equal on each sensor and 100mph>speed>5mph and 14mph/s>accel>-13mph/s and there are no restarts, or there is one restart and exactly two blocks and unblocks in the time buffer.

APPENDIX B: Explanation of the RSD databases.

The following is an explanation of the data fields found in this databases:

License	Illinois license plate
Date	Date of measurement, in standard format.
Time	Time of measurement, in standard format.
Percent_co	Carbon monoxide concentration, in percent.
Co_err	Standard error of the carbon monoxide measurement.
Percent_hc	Hydrocarbon concentration (propane equivalents), in percent.
Hc_err	Standard error of the hydrocarbon measurement.
Percent_no	Nitric oxide concentration, in percent.
No_err	Standard error of the nitric oxide measurement
Percent_co2	Carbon dioxide concentration, in percent.
Co2_err	Standard error of the carbon dioxide measurement.
Opacity	Opacity measurement, in percent.
Opac_err	Standard error of the opacity measurement.
Restart	Number of times data collection is interrupted and restarted by a close-following vehicle, or the rear wheels of tractor trailer.
Hc_flag	Indicates a valid hydrocarbon measurement by a "V", invalid by an "X".
No_flag	Indicates a valid nitric oxide measurement by a "V", invalid by an "X".
Opac_flag	Indicates a valid opacity measurement by a "V", invalid by an "X".
Max_co2	Reports the highest absolute concentration of carbon dioxide measured by the remote sensor; indicates the strength of the observed plume.
Speed_flag	Indicates a valid speed measurement by a "V", an invalid by an "X", and slow speed (excluded from the data analysis) by an "S".
Speed	Measured speed of the vehicle, in mph.
Accel	Measured acceleration of the vehicle, in mph/s.
Lic_type	Unknown.
Reg_month	Indicates the month the current registration expires.
Reg_year	Indicates the year the current registration expires.
Address_2	Indicates the city, state, and zip code of the registrants' address.
Year	Model year of the vehicle.
Make	Manufacturer of the vehicle.
Body_style	Type of vehicle.
Vin	Vehicle identification number.
Owner_code	Unknown.
Make_abrv	Abbreviated manufacturer.