

Day-of-Week Patterns for High-Risk Toxic Air Contaminants Los Angeles Area, 1989 through 1998

Jeff Austin
Planning and Technical Support Division
California Air Resources Board

September 29, 1999

Introduction

In support of our ongoing analysis of weekday-weekend differences in air pollution, we computed means by day of week for six high-risk toxic air contaminants (TACs): 1,3-butadiene, benzene, acetaldehyde, formaldehyde, carbon tetrachloride, and perchloroethylene. We performed statistical tests to provide an indication of the magnitude of the systematic differences between days of the week relative to random day-to-day variation. The following is a brief summary of our findings.

Data

Data consisted of ambient concentrations of six of the highest-ranked TACs in terms of total estimated excess lifetime cancer cases in California. The study covered the Los Angeles area, from 1989 through 1998. During that period, six ARB TAC network sites in the Los Angeles were operational: Burbank, Los Angeles, North Long Beach, Rubidoux, Simi Valley, and Upland. Not all sites and compounds are represented for the entire study period. All data were taken from the ARB's official TAC database. For information on this dataset, and for data requests, please see the ARB Air Quality web site: <http://www.arb.ca.gov/aqd/aqd.htm>.

Other high-risk TACs which were not analyzed include diesel particulate matter, para-dichlorobenzene, and hexavalent chromium. Diesel particulate matter was not considered because ambient data is not available. Para-dichlorobenzene and hexavalent chromium were dropped from the analysis because most of the ambient measurements were below the limit of detection (LOD).

Methodology

- (1) The first step was data quality screening. Two obvious outliers, out of 7,001 observations, were dropped from the analysis. Where data were reported as below LOD, we substituted 2/3 of the LOD, based on a linear approximation to the left tail of the data distribution. Investigations suggest that for a variety of assumed data distributions, this is a better choice than the conventional 1/2 LOD.
- (2) We adjusted for seasonality and trend by taking residuals (differences between actual and fitted values) from a smoothing spline. Splines have an advantage over other smoothing methods when applied to complex datasets in that their degree of smoothness is locally adaptive, rather than being uniform over the range of the data. The degree of smoothness was initially selected by generalized cross validation, then fine-tuned by eye, compound by compound, to yield a curve which followed the seasonal pattern and trend without excessive roughness (Figure 1, upper half). The residuals from the spline fit, henceforth referred to as the seasonally adjusted data, were largely uncorrelated and symmetrically distributed (Figure 1, lower half).
- (3) Treating the seasonally adjusted data for different days of week as independent, we computed group means and standard errors. The assumption of independence is reasonable because TAC samples are collected 12 days apart, long compared to the time scale of meteorological events which impact atmospheric concentrations.
- (4) We compared days of the week by examining error bar charts of mean seasonally adjusted concentration by day-of-week (Figure 2). The width of the error bars was set to a 97.5% confidence interval to yield an approximate 95% confidence level for pairwise comparisons between days.
- (5) To ensure that the seasonal adjustment procedure did not introduce artifacts, we computed "raw" means without seasonal adjustment. While there were minor differences between the means of the raw and seasonally adjusted data, it made no difference to the overall conclusions.
- (6) To confirm the statistical results, we used the SAS GLM (general linear model) procedure to perform analysis of variance on day-of-week means, including fixed effects for month crossed with year. In order to stabilize the error variance and reduce the effect of extreme observations, we transformed the data according to the relationship $y = \log(x + offset)$, where *offset* varied from compound to compound. We compared the GLM significance levels for pairwise comparisons of days of week against the error bar charts. GLM tends to report slightly higher significance levels than the charts, which is expected since the confidence bounds used to generate the charts are conservative.

Results

1,3-butadiene	Sunday is the lowest day of the week at all sites, often significantly different from mid-week. The Saturday mean concentration is comparable to weekdays, but generally slightly lower than Friday. Many sites show a slight dip on Wednesday.
benzene	Site-by-site results are virtually identical with 1,3-butadiene.
acetaldehyde	Data only available for 1996-1998. No discernible pattern.
formaldehyde	Data only available for 1996-1998. No discernible day-of-week pattern, but site-by-site results are virtually identical with acetaldehyde.
carbon tetrachloride	No discernible pattern.
perchloroethylene	Sunday is strikingly lower and less variable than other days of the week. Saturday is lower than weekdays, except at Long Beach, where it is comparable to weekdays.

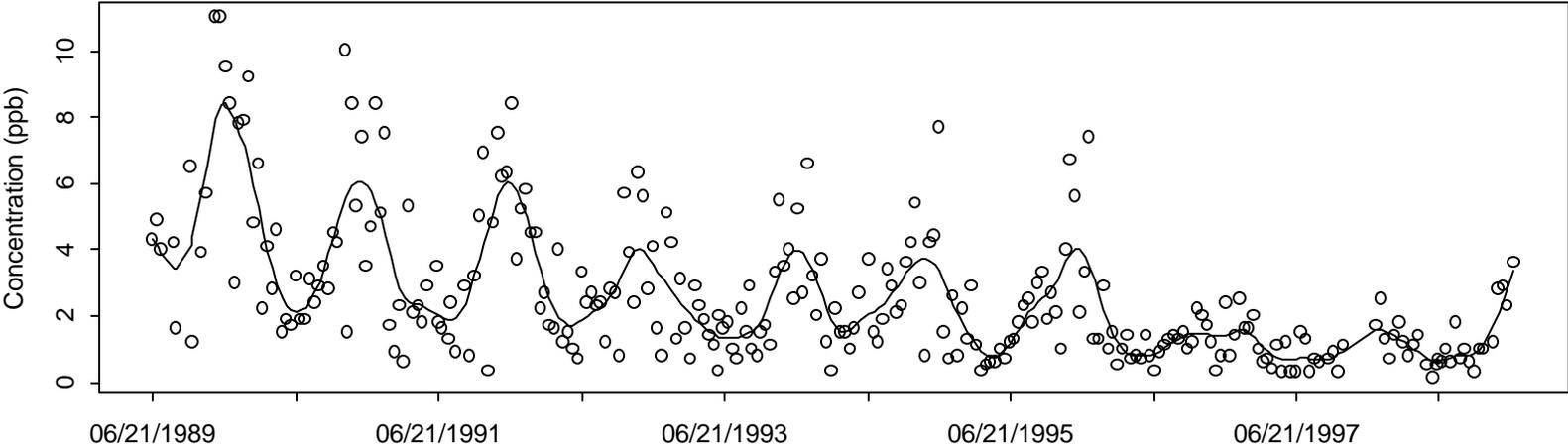
Explanation of Error Bar Charts

The attached error bar charts (Figure 2) show the means of the seasonally adjusted concentrations (residuals from the smoothing splines) by day of week, with error bars showing 97.5% confidence intervals for individual means. If the error bars for two days do not overlap, the means are significantly different at the 95% level of confidence. If two error bars do overlap, the result is inconclusive; it does not necessarily mean that the two means are the same, only that the difference between them is small relative to day-to-day variation. Note that the vertical scale is different from one compound to another. Numbers next to the means indicate sample sizes.

Figure 1. Smoothing Spline

DRAFT

Burbank Benzene concentrations with smoothing spline



Histogram of residuals

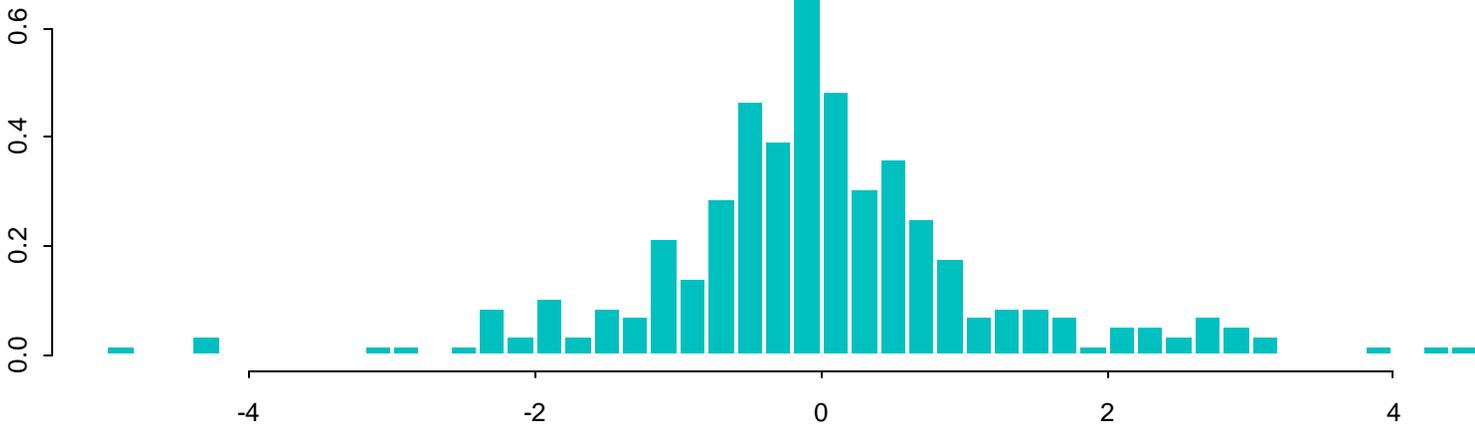
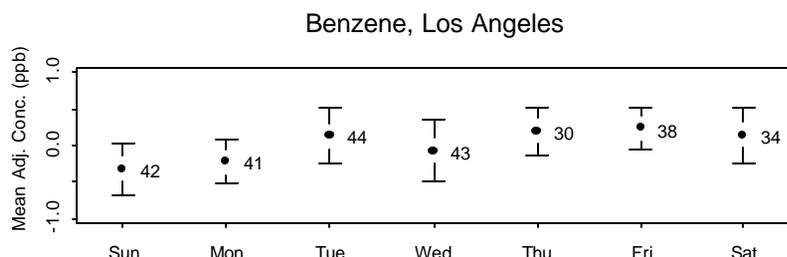
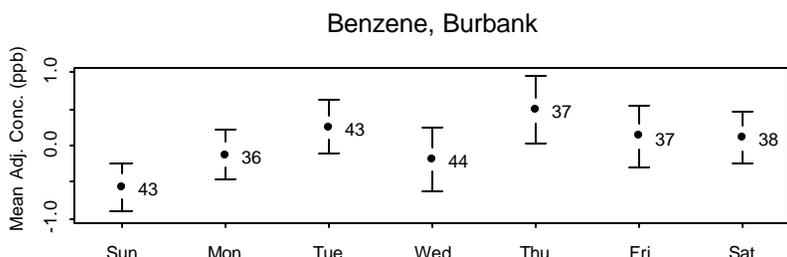
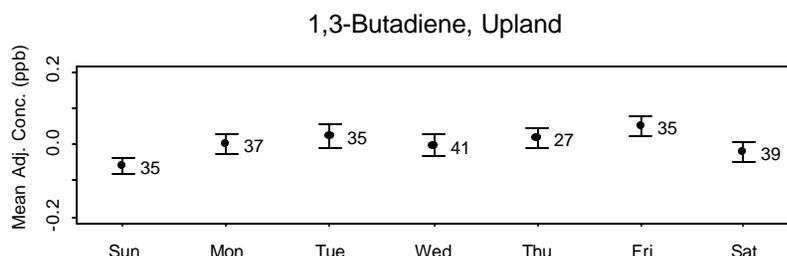
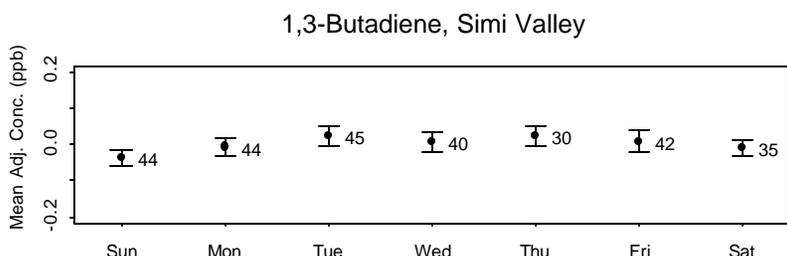
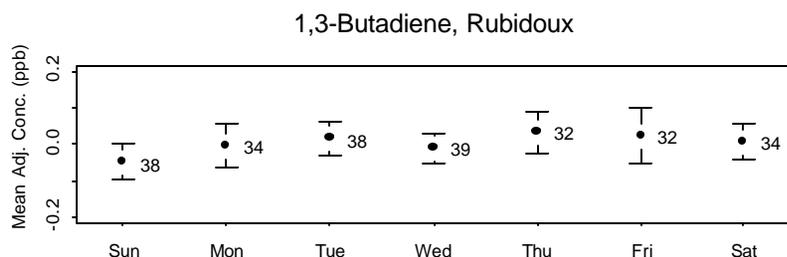
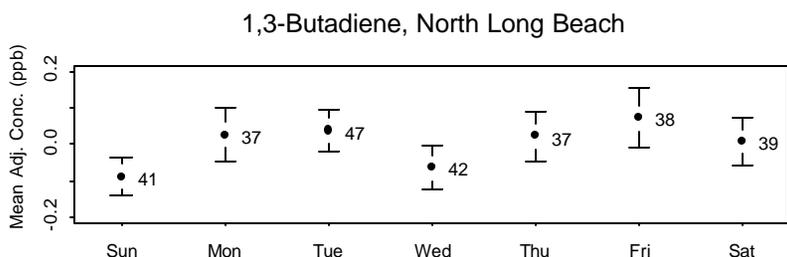
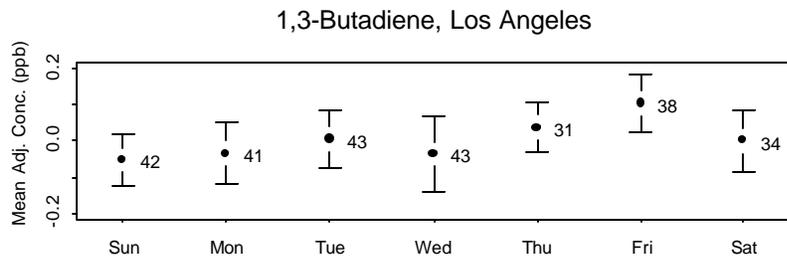
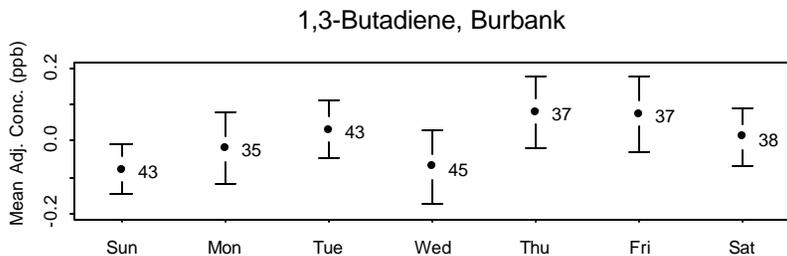
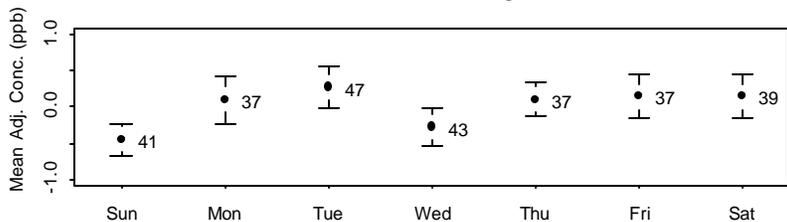


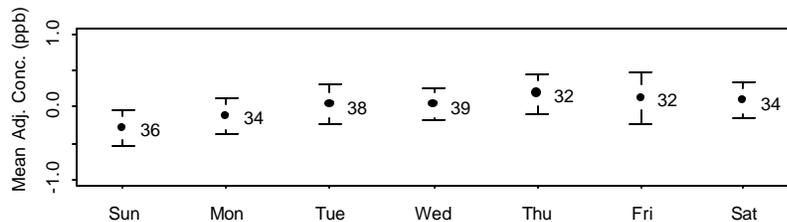
Figure 2. Day-of-week means by Compound



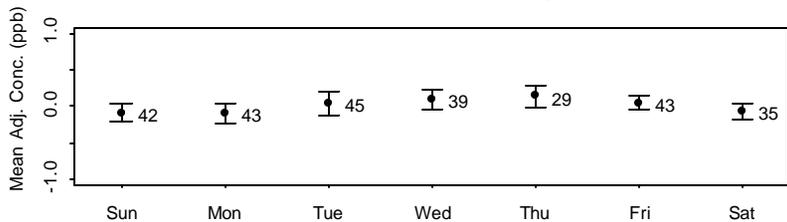
Benzene, North Long Beach



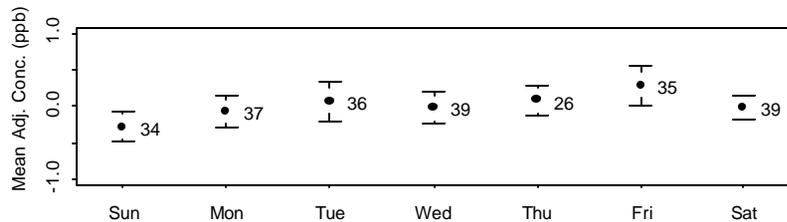
Benzene, Rubidoux



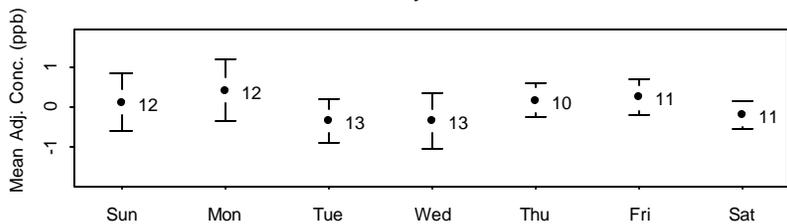
Benzene, Simi Valley



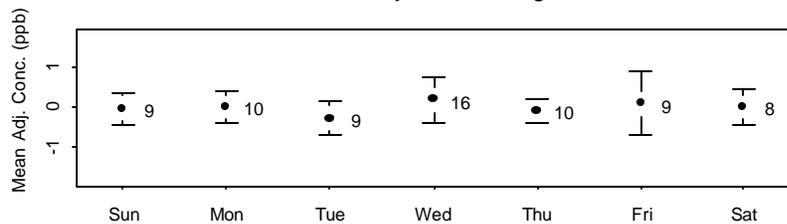
Benzene, Upland



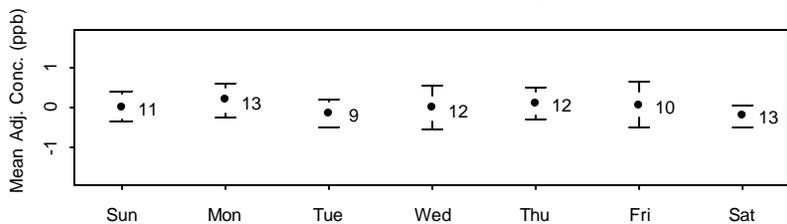
Acetaldehyde, Burbank



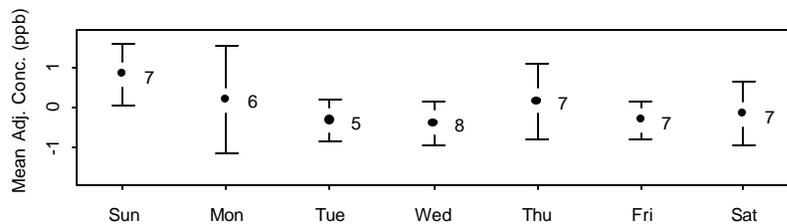
Acetaldehyde, Los Angeles



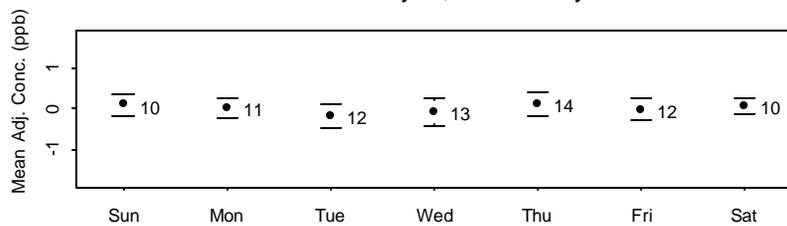
Acetaldehyde, North Long Beach



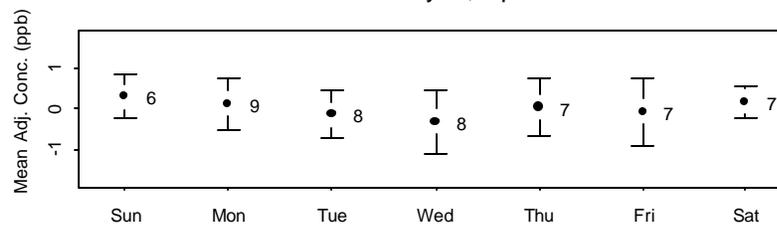
Acetaldehyde, Rubidoux



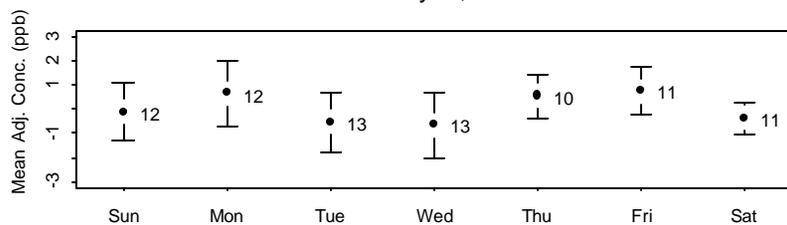
Acetaldehyde, Simi Valley



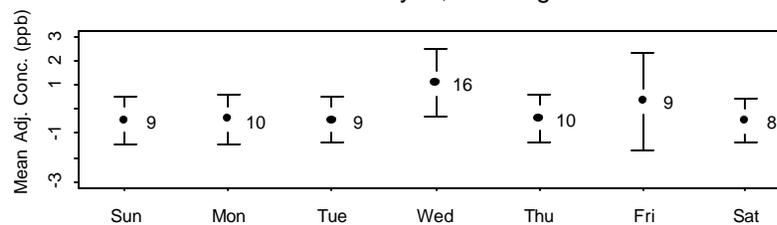
Acetaldehyde, Upland



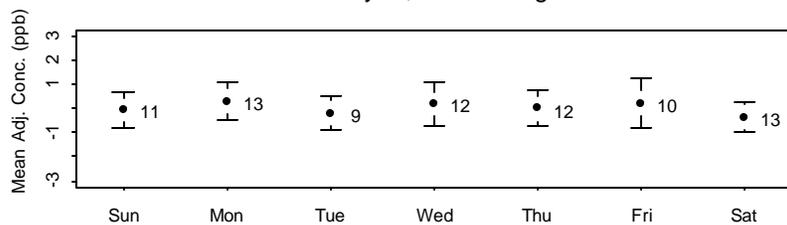
Formaldehyde, Burbank



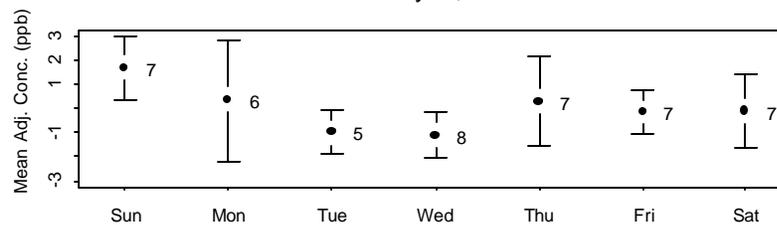
Formaldehyde, Los Angeles



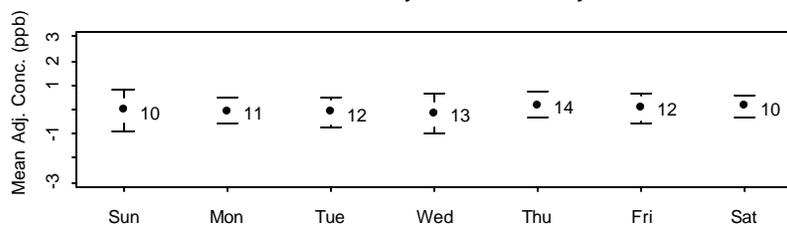
Formaldehyde, North Long Beach



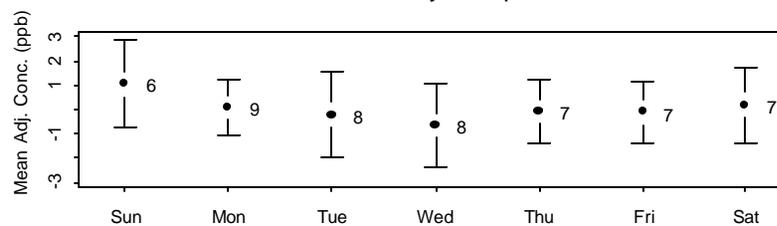
Formaldehyde, Rubidoux



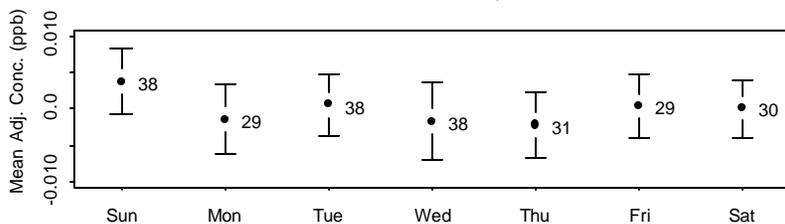
Formaldehyde, Simi Valley



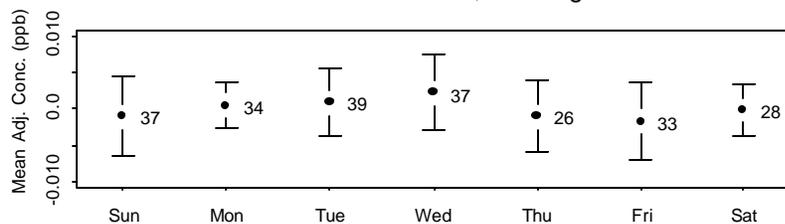
Formaldehyde, Upland



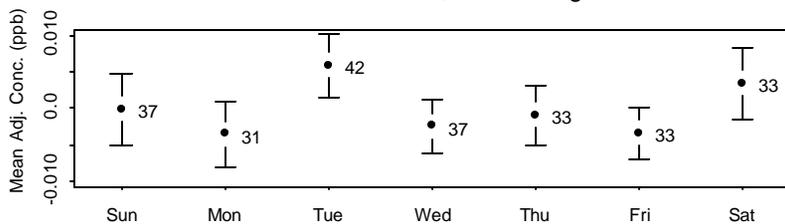
Carbon Tetrachloride, Burbank



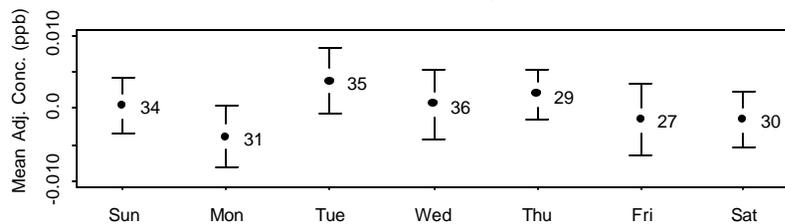
Carbon Tetrachloride, Los Angeles



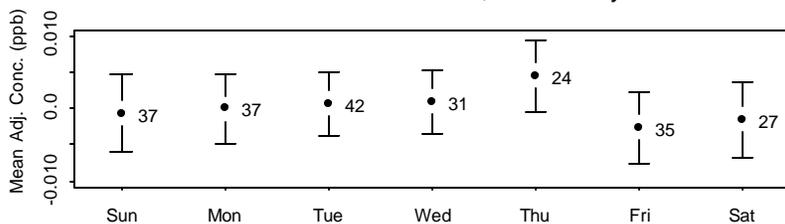
Carbon Tetrachloride, North Long Beach



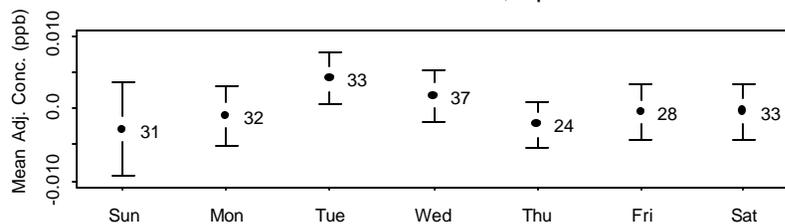
Carbon Tetrachloride, Rubidoux



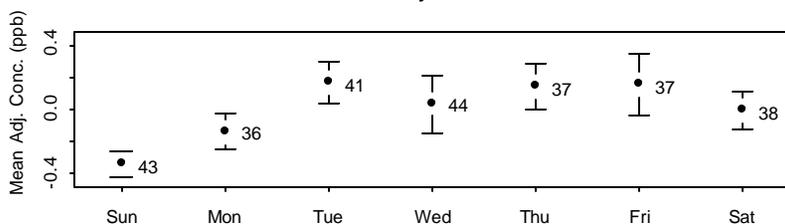
Carbon Tetrachloride, Simi Valley



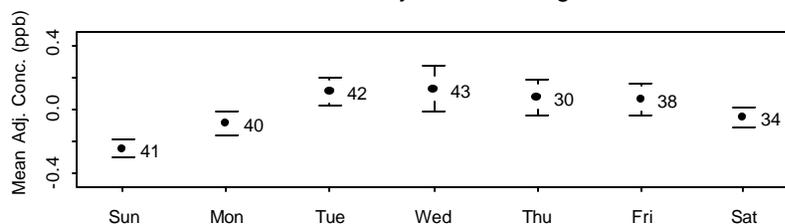
Carbon Tetrachloride, Upland



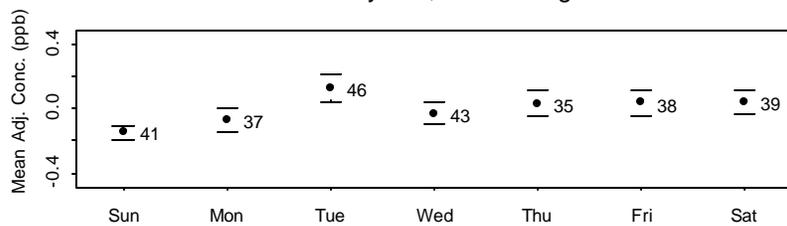
Perchloroethylene, Burbank



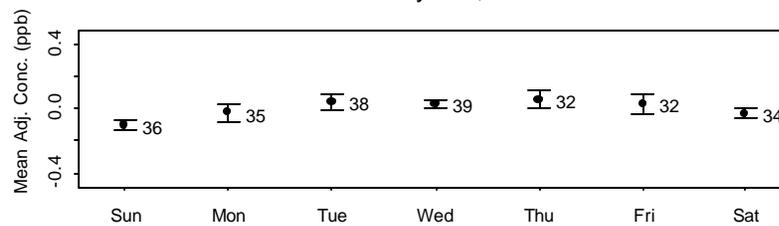
Perchloroethylene, Los Angeles



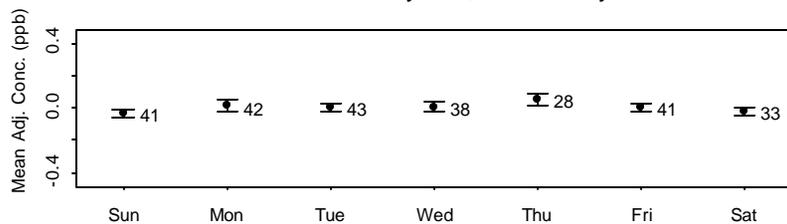
Perchloroethylene, North Long Beach



Perchloroethylene, Rubidoux



Perchloroethylene, Simi Valley



Perchloroethylene, Upland

