

OZONE STANDARD EXCEEDANCE DAYS IN THE SOUTH SAN JOAQUIN VALLEY

JAN DE LEEUW AND SHUOJUN WANG

ABSTRACT. This note analyzes ozone standard exceedance days in Bakersfield, Arvin, Oildale, and Shafter since 1989. We have plotted violations of both one-hour and eight-hour federal and state standards (data from CARB monitoring stations). We have also drawn linear regression lines to estimate trend. They show a modest long term improvement (around 20%) in violations of one-hour state standards, but generally a smaller increase in violations of eight-hour state standards. They also show larger improvement in urban Bakersfield than in the rural areas.

Date: September 23, 2007.

I know that many people have already written about these matters, I fear that I shall be considered presumptuous in writing about them, too, the more so because in treating this subject I depart from the rules set down by others. But since it is my intention to write something of use to those who will understand, I deem it best to stick to the practical truth of things rather than to fancies.

NICCOLO MACHIAVELLI, *THE PRINCE*, 1513

| | |
|-----------------|---|
| OZONE STANDARDS | 3 |
|-----------------|---|

CONTENTS

| | |
|-----------------|----|
| List of Tables | 3 |
| List of Figures | 3 |
| 1. Introduction | 4 |
| 2. Results | 6 |
| 3. Trend Slopes | 10 |

LIST OF TABLES

| | |
|-------------------------|----|
| 1 Slopes of Trend Lines | 11 |
|-------------------------|----|

LIST OF FIGURES

| | |
|---------------|---|
| 1 Arvin | 6 |
| 2 Bakersfield | 7 |
| 3 Oildale | 8 |
| 4 Shafter | 9 |

1. INTRODUCTION

The San Joaquin Valley Air Pollution Control District, on its website at www.valleyair.org gives itself an A^+ on its own *Air Quality Report Card*. See the details at

www.valleyair.org/brochures/docs/ReportCard081706.pdf.

Closer investigation of one particular case suggests that perhaps a grade change is in order.

For our example we look at ozone, the District claims an 82% improvement from 1990 to 2005. This is because there were 45 days above standard in 1990 and 8 days above standard in 2005. In later communications (not in the Report Card), the district uses three-year running averages and finds an 80% improvement. Several comments are in order.

Looking at only two years (or even two three year running averages) gives an estimate that is unreliable statistically. Picking two other years, or two other windows, even if they are close to 1990 and 2005, can give quite different measures of improvement. If 2007 comes out to be a particularly bad year for whatever climatological reason, then will the District say that there has been a 50% deterioration since 1990, and give themselves an F? Probably not. Using only two years gives, at the very least, the impression of what is generally known as “data snooping” or “data mining”. It’s not exactly fraud, but it’s not exactly honest data analysis either.

Second, days above standard is a rough measurement, it would be much better to look at actual concentrations or peak concentrations. People in the Valley breathe actual air, not federal or state standard air. Rounding data to the yes-no exceedance format throws away useful information, and introduces unnecessary sampling errors into the estimates. Moreover, there are four different standards to pick from: State 1 hour, State 8 hour, Federal 1 hour, and Federal 8 hour. These create more possibilities for “data mining”.

Third, the data are aggregated to such an extent that they become almost meaningless (except to a regulating agency). We are interested in ozone

levels at certain fixed measurement stations, such as Bakersfield or Arvin. Again, nobody in the district breathes average air. Also, we should not aggregate over months or days, unless we correct for seasonal variation. And we cannot directly compare ozone levels unless we take meteorological variation over years into account. Some years are hotter than others, some are wetter, some have more wind, and so on. All this is known to influence ozone levels. It is unclear if a three-year running average corrects for this, because obviously droughts and warming produce cycles which are much longer than three years.

Basically, this note is about accuracy in reporting and about appropriate ways to summarize pollution data of this form. We do not make any statements about the health effects of ozone. CES is looking at many more analyses of these data, at the daily level, using concentrations instead of violations, using other pollutants such as CO, PM 2.5 and NO^2 , using space-time models, and correcting for seasonality, autocorrelation, geography, and meteorology. The outcome will hopefully give better and more interesting descriptions of how air quality in the Valley is developing.

2. RESULTS

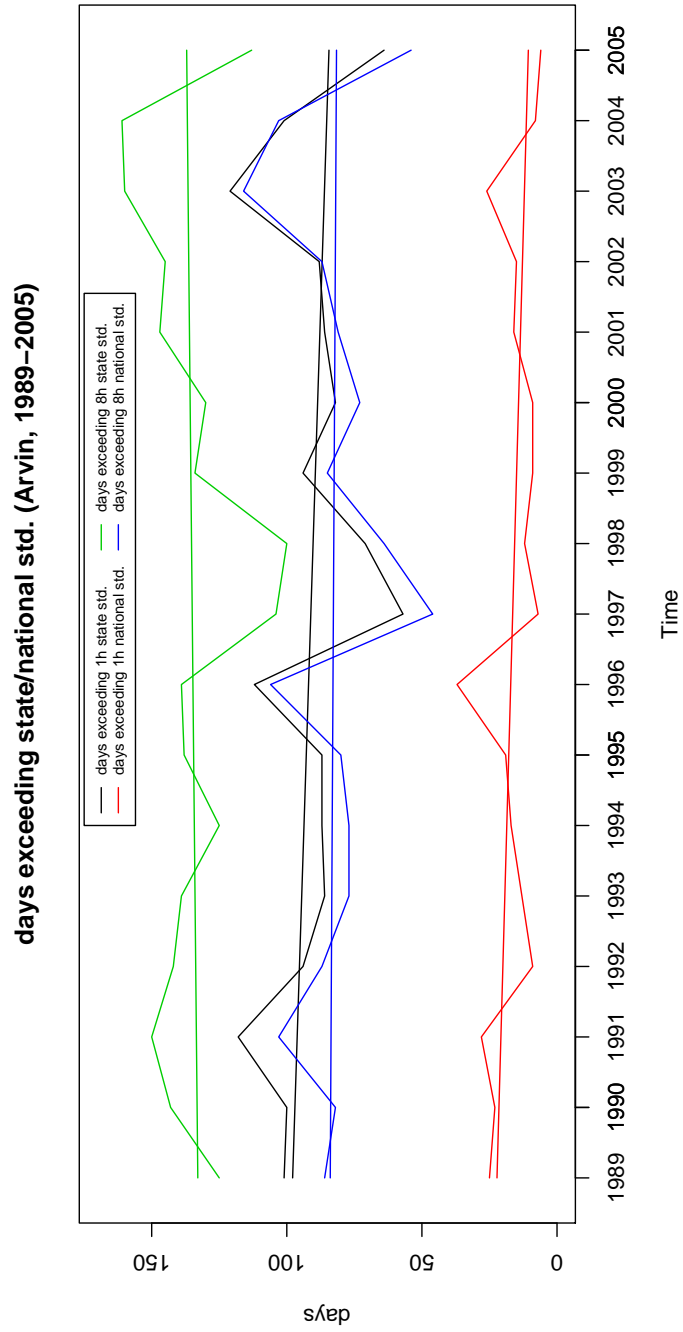


FIGURE 1. Arvin

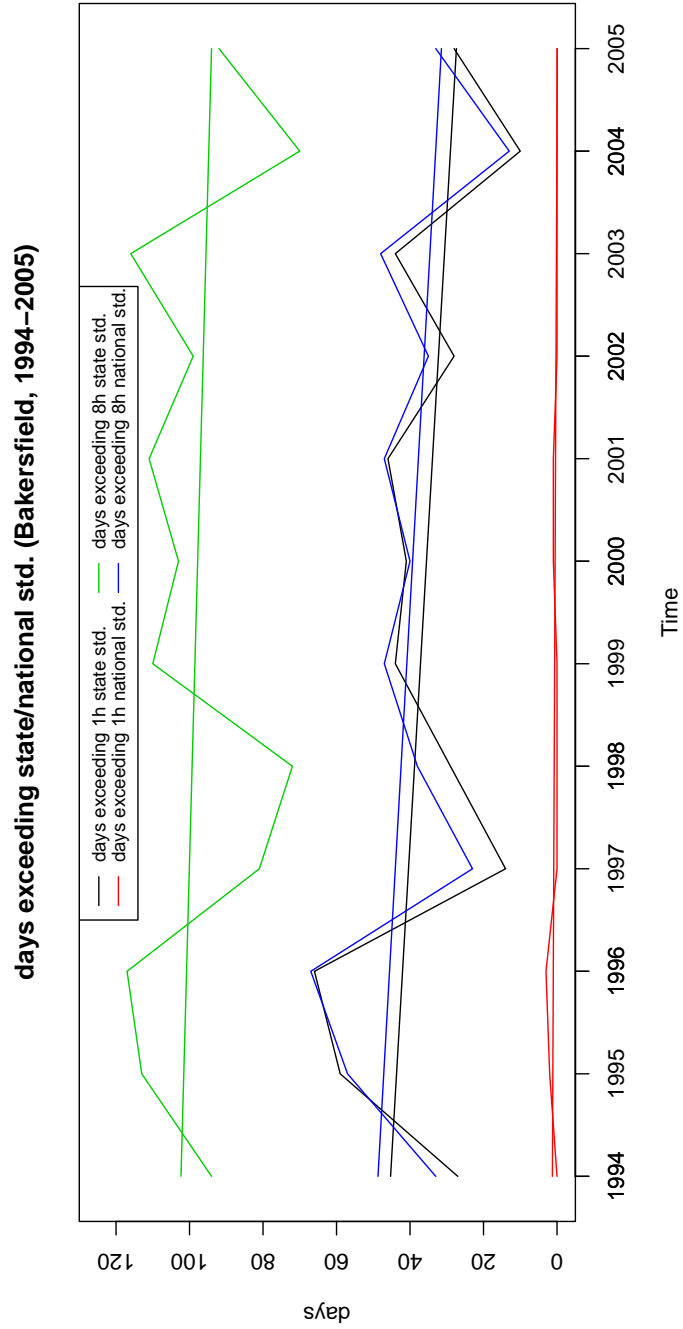


FIGURE 2. Bakersfield

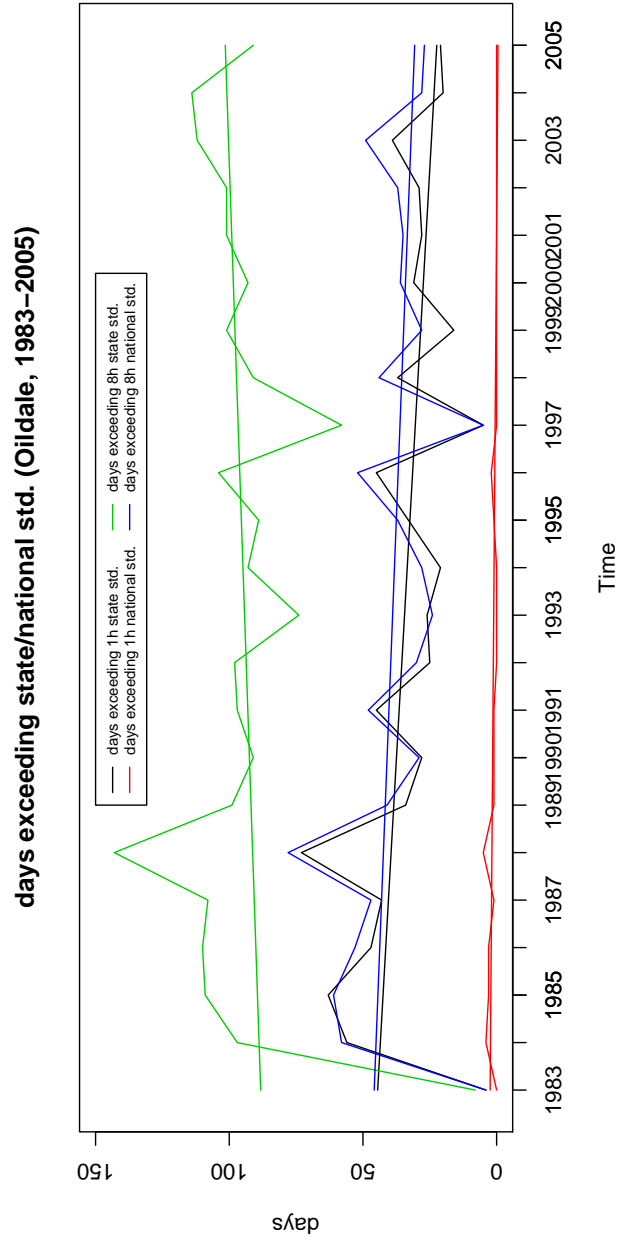


FIGURE 3. Oildale

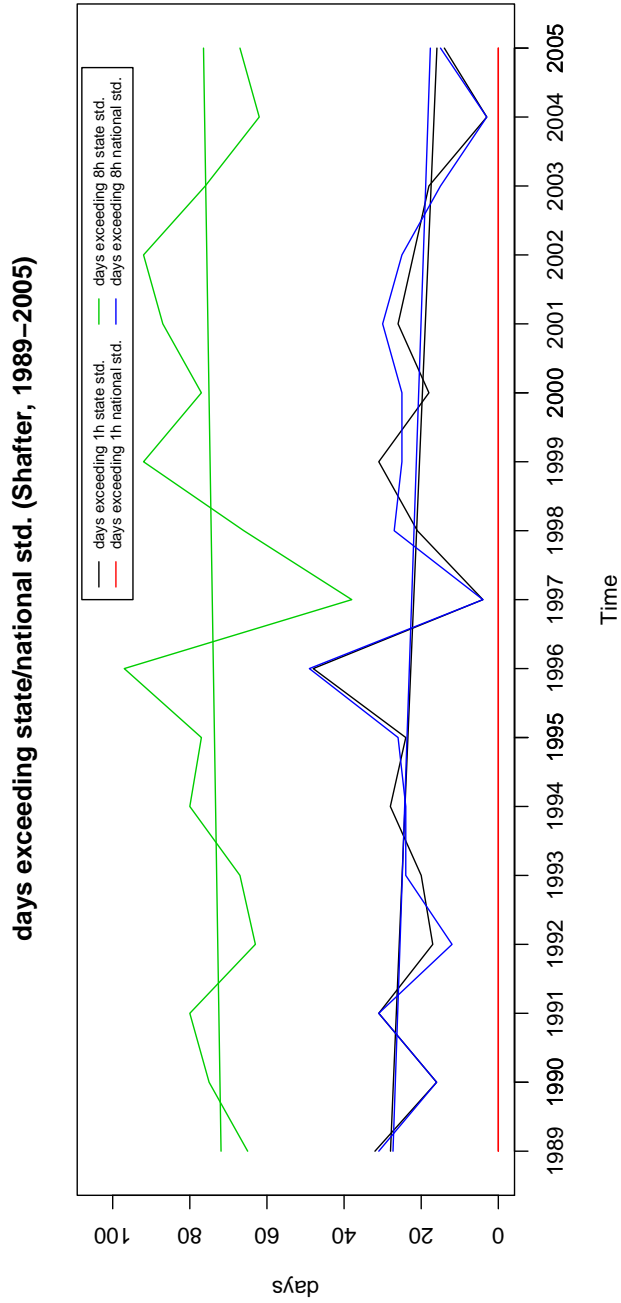


FIGURE 4. Shafter

3. TREND SLOPES

In Table 1 we give the slopes of the regression lines in the plots, which are rough measures of trend. Slopes can be interpreted as estimated number of days increase or decrease in a year. Thus Arvin is estimated to increase non-compliance days on the state eight year standard by 0.2574 days per year, i.e. by one day every additional four years. Bakersfield, in the same way, will decrease non-compliance by three days every four years and on the federal eight hour standard by three days every two years. We see modest gains in most places, but slopes for the state eight hour standard are positive in Arvin, Oildale, and Shafter.

For an alternative interpretation of the slopes, the last column of the table estimates the year of full compliance if current rates of improvement continue. i.e. the intersection of the regression line with the horizontal axis. Of course current rates of improvement cannot possibly continue indefinitely, because that would eventually imply a negative number of exceedance days in most places and more than 365 days of exceedance per year in Arvin. Also, we cannot expect rates of improvement to be independent between our four locations, because ozone measured in Arvin and Shafter is largely produced in urban Bakersfield. Looking at four locations separately does not take any type of transport into account.

| Site | Agency | Standard | Slope | Full Compliance |
|-------------|---------|----------|---------|-----------------|
| Arvin | State | 1 hour | -0.8407 | 2105 |
| | | 8 hour | +0.2574 | ∞ |
| | Federal | 1 hour | -0.7255 | 2019 |
| | | 8 hour | -0.1446 | 2570 |
| Bakersfield | State | 1 hour | -1.6290 | 2030 |
| | | 8 hour | -0.7552 | 2130 |
| | Federal | 1 hour | -0.1224 | 2004 |
| | | 8 hour | -1.5700 | 2025 |
| Oildale | State | 1 hour | -1.0049 | 2027 |
| | | 8 hour | +0.6018 | ∞ |
| | Federal | 1 hour | -0.1334 | 1998 |
| | | 8 hour | -0.6877 | 2050 |
| Shafter | State | 1 hour | -0.7525 | 2012 |
| | | 8 hour | +0.2843 | ∞ |
| | Federal | 1 hour | ----- | ----- |
| | | 8 hour | -0.6054 | 2034 |

TABLE 1. Slopes of Trend Lines

CENTER FOR ENVIRONMENTAL STATISTICS, DEPARTMENT OF STATISTICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 90095-1554

E-mail address, Jan de Leeuw: deleeuw@stat.ucla.edu

URL, Jan de Leeuw: <http://gifi.stat.ucla.edu>