ANALYSIS OF WEEKEND-WEEKDAY DIFFERENCES OF OZONE USING THE ANALYSIS TOOLS: PAW AND ROOT

Shaheen R. Tonse*
Lawrence Berkeley National Lab., Berkeley, CA, USA

1. INTRODUCTION

Urbanized regions tend to have higher concentration of ozone on weekends than weekdays. This "weekend effect" appears paradoxical since it is well known that the total anthropogenic emissions of NO_x and are typically lower on weekends. Currently most of the hypotheses can be summarized as: reductions in the mass of emitted NO_x (in relation to the NO_x/VOC ratio of the region); change in the timing of NO_x emission; carryover of O_3 from the previous day; increases in the mass of emitted NO_x during weekends; change in aerosol scattering due to decreases in soot emissions, which would accelerate various photolysis reactions.

In this study we have conducted CMAQ simulations of weekdays and weekends, outputting concentrations and process analysis variables. We have analysed the results with a focus on the use of an analysis and data-interrogation tool, PAW. We also touch on the use of another, more sophisticated tool, Root.

PAW (Physics Analysis Workstation) and Root are tools developed at CERN (European Center for Nuclear Research). Both are open source and also free. PAW was developed in the late 1980's and extensively used worldwide by the nuclear and high-energy physics fields for data analysis. Root first made an appearance in he late 1990's as an object-oriented approach to data analysis and was seen as being the best way to handle complicated data hierarchies and structures in a sane manner. The original intent was for Root to replace PAW, however PAW's ease of use has resulted in its continued support by CERN.

2. SIMULATION SETUP AND PROCEDURE

The simulation is conducted using CMAQ 4.3 with the SAPRC99 gas-phase mechanism. The mechanism has been enhanced by the addition of methyl-butenol, a biogenic emission, along with 4 associated reactions. The simulated domain is shown in Fig.1.

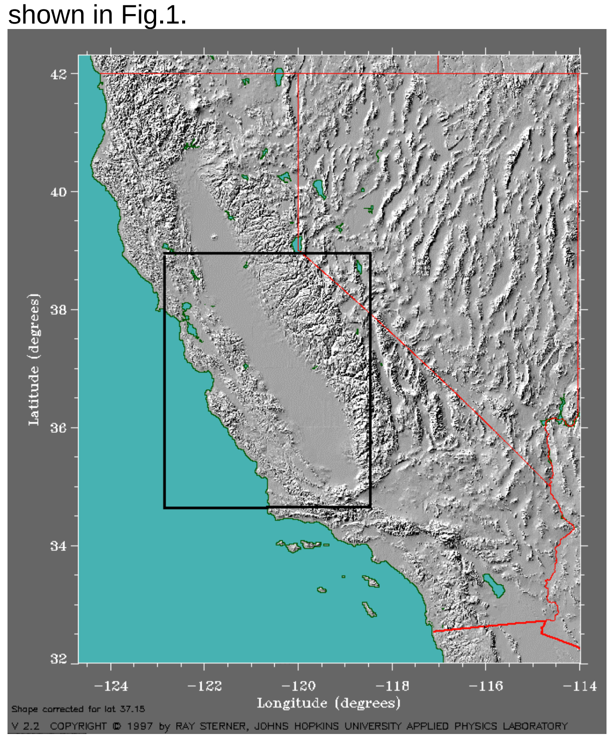

Figure 1 The domain covers the California central coast and the Central Valley, and both the San Francisco and Sacramento metropolitan areas.

The domain has 96 cells horizontally, 117 cells vertically, with each cell side being 4km. The vertical extent of the domain is up to pressure of 100 mbars, about 16km above sea level. The meteorology is that of a 5-day epsiode from 29th July to 3rd Aug. 2000. The 50 original vertical layers of the MM5 simulation have been consolidated to 27 by MCIP 2.3. Emissions come from Area, Point, Biogenic and Motor Vehicle (MV) sources. The MV emission inventory contains 4 distinct days: Weekday (Mon-Thu), Friday, Saturday, and Sunday. Friday emissions differ from the other weekdays. The time dependence of the MV emissions shows significant differences on weekends, due to absence of morning and evening rush hours, but a still substantial amount of activity, albeit later in the day. Also the contribution from diesel truck traffic is very different on weekends. The temperature

*Corresponding author: Shaheen R. Tonse, MS 51R0208, Lawrence Berkeley Lab., Berkeley, CA, 94720; e-mail: tonse@lbl.gov

dependence of the biogenic emissions has been derived from the meteorology of the actual day. Fig. 2 and 3 show time series of weekend and weekday VOC and NO_x emissions respectively.

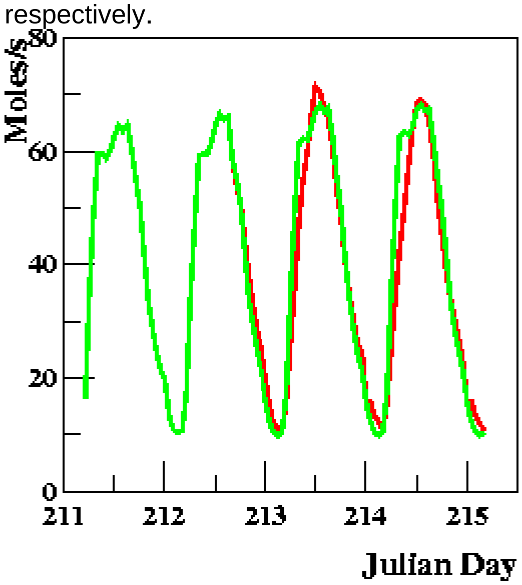

Figure 2 Summed VOC emissions for weekdays (green) and weekends (red) for the first 4 days of the simulation.

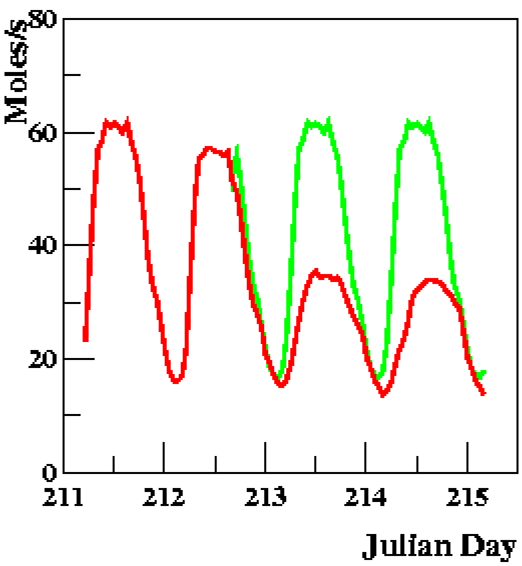

Figure 3 Summed NO_x emissions for weekdays (green) and weekends (red) for the first 4 days of the simulation.

Two hypothetical 5-day emission scenarios were simulated:

- A weekend simulation consisting of Weekday-Friday-Saturday-Sunday-Sunday
- A weekday simulation consisting of Weekday-Friday-Weekday-Weekday-Weekday

Identical meteorology and boundary conditions were used in both. Both runs were initiated from a 48-hour spinup, and in addition to that the first 2 days (Weekday, Friday) are identical. Outputs consisted of an average hourly concentration file, and IPR and IRR process analysis integrated rates. The Process Analysis feature of CMAQ was crucial to the analysis of the results, so as to track the formation and consumption of species in the NO_x and HO_x cycles. The analysis focuses on the 3rd and 4th days of the simulation.

3. PAW

PAW Capabilities: The most commonly used data form in PAW is the Ntuple, which is simply a collection of data lines, each consisting of N data fields. In our analysis a simplified example would contain: Date, Time, IX, IY, IZ, [O3wkend], [NO2wkend], [O3wkday], [NO2wkday].

PAW operates through commands at a terminal window, and graphics are plotted on a separate window. Entries on the same ntuple line can be plotted against one another as 2D and 3D plots. Mathematical operations (including trignometric, exponential, and log) are allowed. Conditional Boolean cuts on variables are allowed. Examples of plotting from an ntuple, (numbered Ntuple#1) :

- NTUPLE/PLOT 1.O3WKEND (plot a 1D histogram (distribution) of [O3])
- NTUPLE/PLOT 1.O3WKEND DATE==213 (as above, but for Julian day 213)
- NT/PL 1.O3WKEND%TIME DATE==213 (time series of [O3])
- NT/PL 1.O3WKEND%NO2wkend (scatterplot of [O3] vs [NO2])
- NT/PL 1.O3WKEND%IY%IX DATE==213 (spatial map of [O3])
- NTU/PL 1.IY%IX DATE==213&&TIME=1500&&O3WKEND>O3WKDAY (spatial map of [O3] for cells with weekend [O3] greater than the weekday O3 at 3pm on Day 213)

In general, such operations are permitted between elements of the same ntuple line. Other features:

- Multiple input PAW files can be opened at the same time, appearing like Unix-style directories.
- Multiple PAW commands can be put into macro files and executed. The macros also allow simple IF and DO constructs, and variable-naming.

  PAW contains a built-in Fortran 77 interpreter for more complicated operations than is possible in the macros. In general any ntuple variable can be passed to the Fortran.

- Graphical cuts can be drawn on the graphics window and saved; useful for interrogating anomalous areas on a scatterplot.
- Ntuple variable values can be written out to ascii file.
- There are numerous graphics options for color, text fonts, plot type
- The MINUIT fitting package is part of PAW.
- Help on any command is available by simple typing HELP or USAGE.
- Hardcopy output is printed as Postscript files.
- Availability: PAW can be downloaded from http://wwwasd.web.cern.ch/wwwasd/paw/ Precompiled versions exist for many platforms.

PAW's data input is in a binary format and is prepared by our IOAPI to PAW converter, Ioapi2ahb. Ioapi2ahb is written in Fortran 90 and requires linking only to the IOAPI and Netcdf libraries. Ioapi2ahb is driven by a simple ascii input file which allows one to select variables from multiple IOAPI files. It consolidates variables with the same timestep and grid cell coordinates onto the same ntuple line. User-specified limits on X,Y,Z, timestep range are permitted. In addition to writing output for every X,Y,Z cell, Ioapi2ahb also performs weighted averaging of the variables over the column of Z layers over each cell and for the domain as a whole. PAVE-style alphabetical suffixing of variables is an option. Variable name translation is also an option.

4. APPLICATION OF PAW TO CMAQ RESULTS

The ntuple can of course contain whatever the user desires, but in general, our PAW ntuple made from IOAPI files usually contains a few emissions, (NO_x, VOC's), average hourly (ACONC) values for a handful of select species, and numerous variables from the IRR output files. IRR output was usually of the "Partial IRR" type, containing user-defined family species. One "Full IRR" run was conducted with all 200-odd SAPRC reaction rates output for limited XY-domain, so that in the event of an isolated reaction rate being needed, a CMAQ re-run would not be necessary, and the processing could be handled in PAW instead. Selected weekday and weekend variables are output to the same ntuple line, with the respective suffices D and E appended (weekDay, weekEnd). In addition the date, time, X, Y, Z indices are included as variables on every ntuple line. The IOAPI to PAW programme also, for every X,Y cell, performs a layer-weighted column average for every ntuple variable and writes out a line with the Z index set to –1. Our column-averaging is over the bottom 11 layers, about 250m high. Similarly a domain averaged quantity is calculated with all 3 of the X,Y,Z indices set to –1. These features allow for less "busy" plotting, while still retaining the capability to check the degree to which a domain or column average represents the values in multiple individual cells.

Fig.4 shows spatial dependence of weekend O_3 excess. Over most of the domain weekend O_3 is less than weekday O_3, however, in select areas the reverse is true. This was observed on all of the 3 days for which the 2 simulations differ. Our analysis from this point will show data from a 16x16 cell sub-domain east of San Francisco Bay, near Pittsburgh.

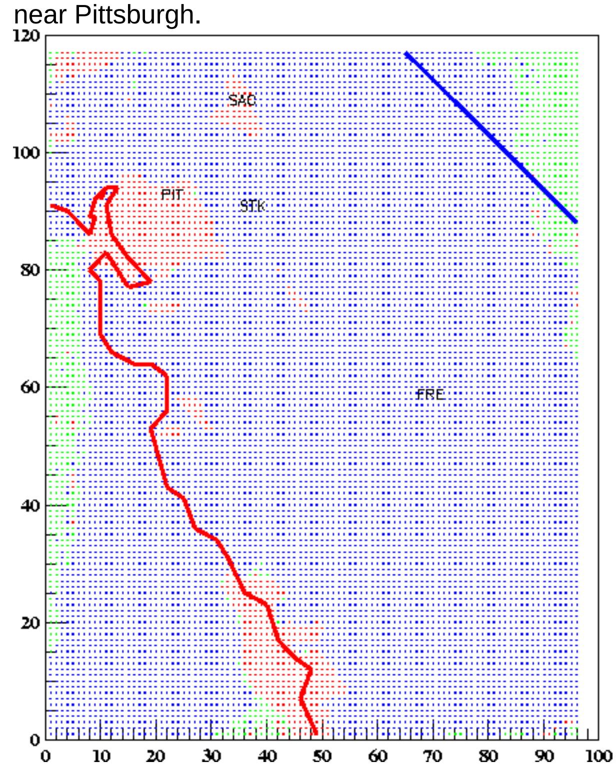

Figure 4 The domain at 3pm. Grid cells are colored depending on whether the weekend O_3 is greater than (red), equal (green) or less than (blue) weekday O_3. The coast is marked red, the CA-NV border is blue, and the cities of Pittsburgh, Sacramento, Stockton and Fresno are shown.

The time-dependence of weekend/weekday O3 excess is shown inthe histogram of Fig.5. Note that this is a sample chosen from a geographical region in which weekend O_3 is generally larger than weekday O_3.

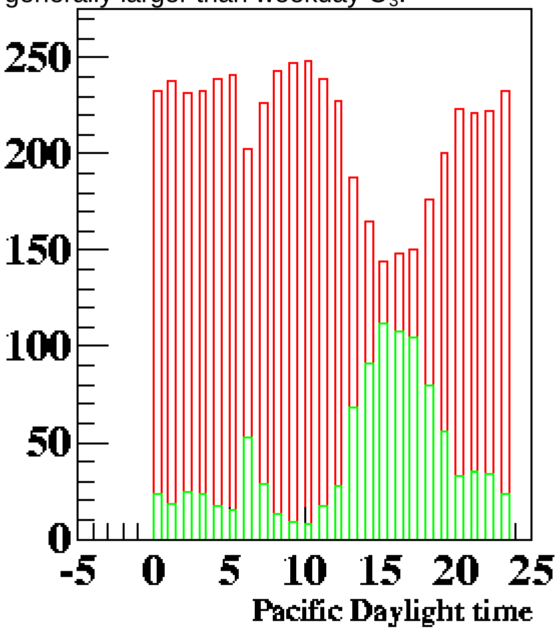

Figure 5 Time-dependence of the number of cells for which weekend [O_3] is greater than (red) or less than (green) weekday [O_3]. The red and green histograms are superimposed, not stacked, and the total number of entries is the number of xy cells times number of timesteps.

Fig.6 shows a profile plot of the mean difference of the excess O_3 for the same criteria. The Y-axis units are ppm, and clearly show that when weekend excesses are seen, they can be large, about 0.01-0.02 ppm. Since it is already known there are lower NO_x emissions on the weekend, the weekend excess could be a trivial result from a lower rate of titration via eqn.1.

$$O_3 + NO \Rightarrow NO_2 \qquad (1)$$

That the rate of titration is smaller on weekends is seen directly from process analysis information, and we display it as a lifetime in Fig.7.

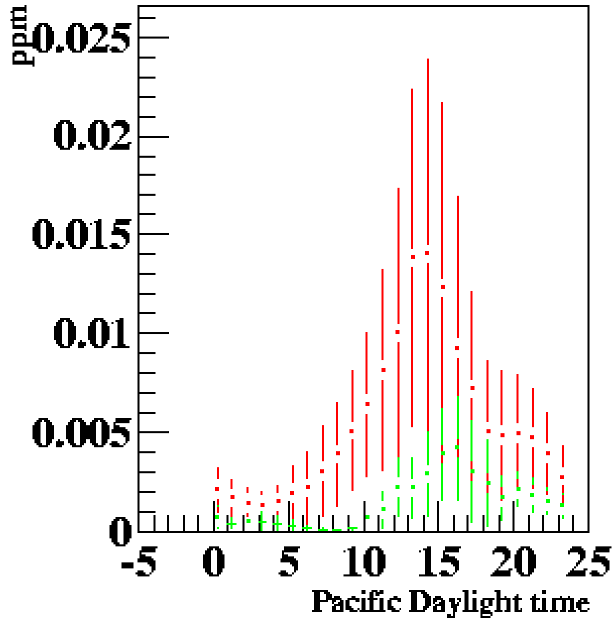

Figure 6 Profile plot of the mean difference of the excess O_3 as a function of time for all column-averaged cells. (This is simply a scatterplot in which each channel in the Y direction has been replaced by the mean and spread of its data points. The bars do not signify errors but indicate the spread in values over the entries in each Y column.)

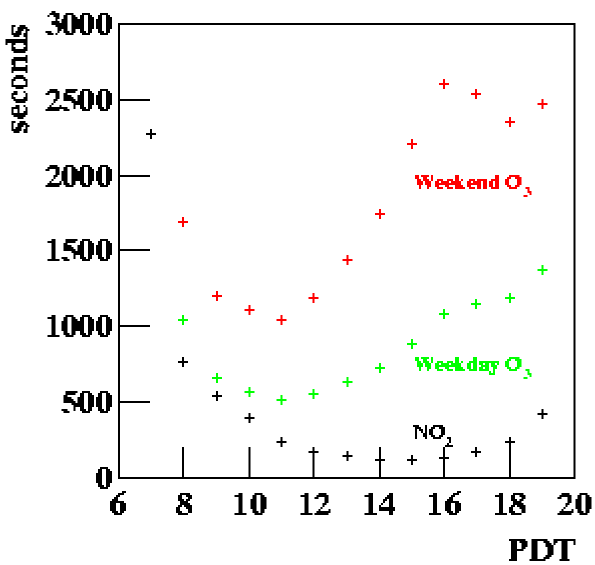

Figure 7 Lifetimes of O_3 and NO_2 as a function of local time.

Weekend O_3 is much longer lived than weekday O_3, due to scarcity of NO with which to titrate. NO_2 lifetime, also shown, is only a function of sunlight intensity and is identical for the 2 simulations.

While much of the O_3 difference is a titration effect, there is also a difference in the odd oxygen (O_3+NO_2). If there is no difference in odd oxygen, then we expect an excess O_3 on weekends to be compensated by a deficit of NO_2. Fig.8 shows a scatterplot of weekday minus weekend NO_2 vs. weekend minus weekday O_3. If titration were the only cause for the weekend O_3 excess, then the points would lie close to the X=Y black line, (actually, since NO_2 emissions are higher on weekdays, we would expect the points in general be slightly above the black line.) This is the case for nighttime (blue) and early morning (cyan) points, but during the day, we see many entries below the black line, indicating that the increased $[O_3]$ is not merely a result of reduced weekend titration through eqn.1.

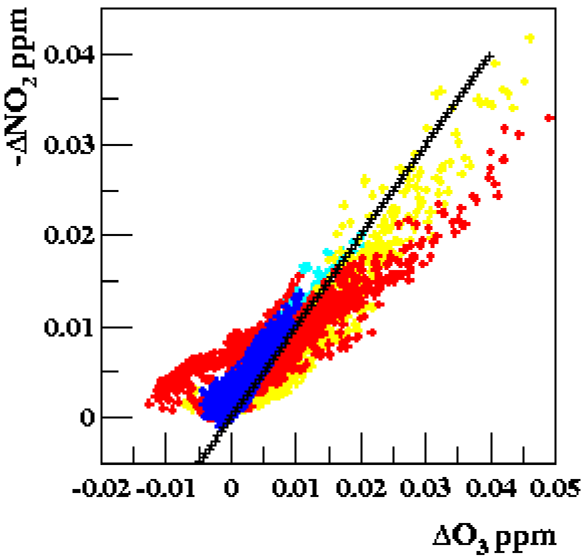

Figure 8 A scatterplot of minus weekend minus weekday NO_2 vs. weekend minus weekday O_3 for every column-averaged grid cell. The color coding shows time of day: Cyan 6am-9am, Yellow: 9am-Noon Red: Noon-6pm, Blue: 6pm onward.

We shall study this further through examination of precursors, the production of OH and HO_2 radicals and the OH chain length. Production of "new" (i.e. OH not formed by the HOx –cycling reaction eqn.4) is 2-3 percent higher on the weekend. The relative contributions from different reactions to new OH production differ, some higher, some lower, eg.

$O3 \Rightarrow O^1D$ (photolysis) (2a)

$O^1D + H_2O \Rightarrow 2\ OH$ (2b)

is slightly higher on weekends because of the higher overall O_3 concentration.

Once produced, most OH reacts either with VOC's or terminates with NO_2 through

$OH+NO_2 \Rightarrow HNO_3$ (3)

Because of a factor 2 higher termination rate on weekdays, weekend OH has a 5-7% higher probability of reacting with a VOC. Fig.9 shows this probability, $f_{(OH+VOC)}$ for weekend vs weekday as a domain-averaged comparison

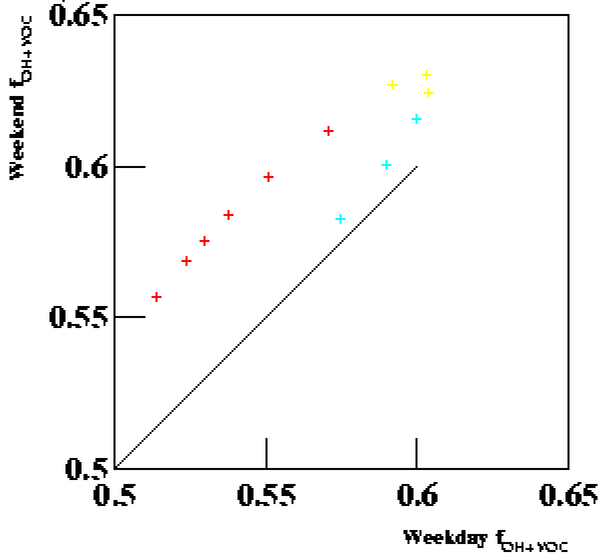

Figure 9 Fraction of domain-averaged OH which reacts with VOC's for weekends vs weekdays. The color coding shows time of day: Cyan 6am-9am, Yellow: 9am-Noon Red: Noon-6pm, Blue: 6pm onward.

HO_2 is ultimately produced as a result of OH+VOC oxidation reactions, with an efficiency of about 100%. The HO_2 will then either react through eqn.4

$HO_2+NO \Rightarrow NO_2+OH$ (4)

or will terminate through other reactions and leave the HO_x cycle. The fraction of HO_2 which reacts with NO, ($f_{(HO2+NO)}$), times $f_{(OH+VOC)}$, gives the probability that an OH makes it though the HO_x cycle, and possibly reacts with another VOC molecule and so on. Fig.10 shows $f_{fHO2+NO)}$ for weekends vs weekdays. After noon the weekend ($f_{(HO2+NO)}$) is about 10% lower due to increased competition for NO from a higher O_3

concentration.

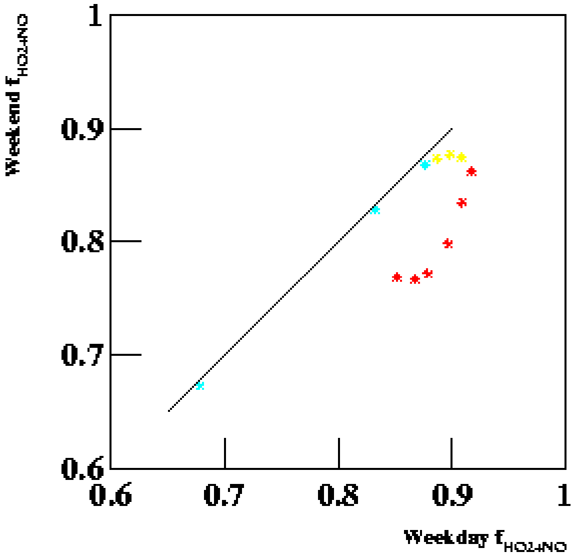

Figure 10 Fraction of domain-averaged HO_2 which reacts with NO to produce NO_2, for weekends vs weekdays. The color coding shows time of day: Cyan 6am-9am, Yellow: 9am-Noon Red: Noon-6pm, Blue: 6pm onward.

We see that the larger weekday NO and NO_2 concentrations have canceling effects in Figs. 9 and 10. There is more termination of OH with NO_2 on weekdays, resulting in a lower $f_{(OH+VOC)}$, but there is also a higher probability of reaction of Eqn.4 on a weekday. The convolution of the 2 probabilities gives Pr(OH) the propagation probability of an OH, which is related to the chain length by:

Chain length = 1/(1 + Pr(OH)) (5)

We do not show this as a plot, however, we observed that domain-averaged Pr(OH) was about equal for the 2 simulations until noon, and was then generally lower in the weekend simulation, by 2-5%, for the afternoon hours.

This is still work-in-progress and at this stage in the analysis we cannot conclusively say why the excess of odd oxygen occurs, due to conflicting effects:

- Higher new OH production on the weekend, about 3%
- Larger $f_{(OH+VOC)}$ on the weekend, due to a lower rate of eqn.3, 5-7%
- Smaller $f_{(HO2+NO)}$ for the weekend due to a lower rate of eqn.4, 10%
- Lower NO_2 emissions on the weekend, 50%.

We expect to continue this analysis with more process analysis variables, in order to isolate the source of odd oxygen deficit in the weekday case.

5. ROOT

Since PAW was sufficient, in the end we did not use Root in this analysis so far. A summary of Root's features follows:

Root, similar to PAW in principle, is a more powerful successor, geared to handling larger data sets and more complicated operations. Its C++ based command line interface, combined with a windows interface allows complicated interactive and automated analysis to be conducted entirely within Root. It provides a set of frameworks with the functionality needed to handle and analyse large amounts of data in a very efficient way. Specialised storage methods are used to get direct access to selected objects, without having to touch the bulk of the data. The graphics is superior to that of PAW. Data interrogation capabilities are superior, in essence one has the advantage of scripting in a high-level language without having to worry about programming for data input and graphics output. The command-line interface and syntax is more difficult than PAW's, and consists of C++ commands. We shall be conducting an analysis with Root in the next fiscal year, and will evaluate whether a community effort to build utility scripts for CMAQ (or other) analysis can make the use of Root much easier.

6. ACKNOWLEDGEMENTS

We would like to thank:

- The California Air Resources Board for funding the analysis work, and for providing Area and Point Emissions.
- Assistant Secretary for Fossil Energy, Office of Natural Gas and Petroleum Technology, through the National Petroleum Technology Office, under U.S. Department of Energy contract No. DE-AC03-76SF00098.
- Allison Steiner (U.C.Berkeley) for providing biogenic emissions
- Rob Harley and Linsey Marr (U.C.Berkeley) for providing motor vehicle emissions
- Jim Wilczak et al (NOAA) for providing MM5 meteorological input
- U.S. Dept. of Energy for funding purchase of a Linux Parallel Cluster Computer.