Supplementary Material for

Controls on Pollution Ozone Production Measurable from Surface, Aircraft, and Satellite Monitors

Chatfield, Ren, Brune, Crawford, Esswein, Fried, Olson, Shetter

Sampling.

The International Consortium for Atmospheric Research on Transport and Transformation contained several land-based and aircraft-borne research programs, including INTEX-NA, the Intercontinental Transport Experiment – North America [Singh et al., 2007]. NASA's main contributions to ICARTT were satellite measurements and measurements aboard the DC-8 aircraft, including OH (hydroxyl), HO_2 (hydroperoxyl), and NO (nitric oxide) radicals. (M1), photolysis rates [Shetter et al., 2002], and HCHO Fried et al., 1997]. These were used in our analyses. Only daytime data with [HCHO] > 30 ppt and [NO] > 10 ppt for the region 0-1300 m characterizing Continental North America south of were used. Figure S1 indicates the broad sampling region. The sampling period was not one of remarkably high regional ozone. An indirect estimate of NO was also available from highly accurate measurements of NO_2, but was not needed in this analysis, since our estimates tend to have canceling effects at very low NO values where the indirect estimates might be superior. We found no difficulties suggested by the histogrms describing the statistical distribution of direct low NO that would urge a change to indirect measurements based on NO_2. Surface-based data exist which characterize near-urban environments, but such samples must be examined more carefully.

Background Regarding Ozone Production.

Jacobson and Seinfeld and Pandis introduce smog ozone chemistry and mathematical approaches to its simulation. Chatfield and Delany [1990] McKeen et al. [1991], Olszyna, et al., 1994] among many others describe inherent difficulties in numerical simulation of smog. analyses describe the difficulty of the problem. Factors determining smog ozone may be partially composed with this summary formula:

$$d\mathrm{O}_3/dt = P_{\mathrm{O}_3} - L_{\mathrm{O}_3}[\mathrm{O}_3] + Tr + Dep + Het,$$

describing chemical production, chemical loss (expressed as proportional to ambient ozone concentrations), transport, deposition, and (aerosol-cloud) heterogeneous removal. In our experience, major controversies concerning air pollution control confront the difficulty of distinguishing relatively local chemical production from more distant transport (e.g., interstate).

Thornton et al. [2002], Sillman et al. [2002], Kleinman [2005a], and Kleinman et al[2005b] relate urban and regional smog to observational evidence, concentrating on the production term, P_{O_3}. HO_2 reactions with RO_2 and other HO_2 radicals are the primary ways that peroxy radical production does not lead to ozone production. There are various self-consistent ways to describe production of oxidant, e.g., one variant is to subtract the rate destruction of NO_2 by OH. We consider this to be a destruction term for oxidant, one that can be of comparable magnitude to the destruction of ozone by HO_2 and of excited atomic oxygen by water.

Variety of Conditions Sampled

The statistical estimation we employed was based on all samples in the Eastern and Midwestern United States (plus a few in Canada) during the INTEX-NA sampling period. The locations of the simultaneous measurements used to make the estimate are shown in Figure S1.

Estimation of Production Functions F_2

The purely statistical 2-D relationships were estimated using two variables, α = $\log10(\phi \bullet \nu)$ and a compact bounded expression of the ratio, viz $\beta = \arctan(\phi / \nu)$. We then used

the *R* language routine *gam* in the *mgcv* package by Simon Wood [2004] to estimate a version of $\log_{10}(\boldsymbol{F_2})$ as a function of α and β. α was allowed 3 degrees of freedom, and β was allowed 4. Contours in each sector of the graph are allowed individual fitting by *mgcv*. It was appropriate to use logarithms to cover the wide range of predictors and ozone production, and also since the quantities tended to have errors proportional to the quantity [Ren et al., 2008]. The transform allowed the curves to emphasize a hyperbolic relationship, reveal deviations from that relationship directly, and avoid problems with very low values of j_{HCHO} [HCHO] or [NO]. The relationship was then transformed back to be graphed on the familiar axes shown, which represent predominantly the effects of VOC's and NO on ozone primary production rate. For the quantity $\log_{10}(\boldsymbol{F_2})$, $r = 0.96$ and $r^2 = 0.92$, an impressive fit. For $\boldsymbol{F_2}$ itself, $\boldsymbol{F_2}$, $r = 0.88$ and $r^2 = 0.78$; note that this fitting method respected relative error in the measurements and did not attempt unrealistically to fit high ozone production rates, in view of the likelihood of high measurement error. Consequently, as estimated by *mgcv*, $\boldsymbol{F_2}$ did not explain appreciably more of the variance than f_1. A wider study might allow greater variance explained. A red dashed line in Figure 2a indicates a perfect hyperbola, and many of the contours approximate a hyperbola fairly well.

Varying the number of degrees of freedom did not significantly alter the basic relationship, the slightly deformed hyperbola, nor did direct computation using $\log(\phi)$ and $\log(\nu)$; Rather than resembling $\boldsymbol{Q_2}$ or modeled ozone production (each described below), fits with more parameters simply made exceptions for very localized features which appeared around the sparsely populated periphery of the dataset; we conclude that a very simple description fits the data. Since Figure 1b uses logarithmic axes, and the dashed red line indicates a perfect hyperbolic, $\phi \bullet \nu$ = constant, relationship. Perfect hyperbolas imply that f_1 (whose estimation is described below) is a statistically complete explanation of the F_2 relationship. These allow us to seek (expectable) deviations from f_1 behavior. HO_2, NO, and HCHO are quoted to have an approximately 10% accuracy; both the local smoothness of both f_1 and F_2 suggest that random error in the precision of these variables has a (surprisingly) small effect. In this situation we surmise that further analysis of possible structure in F_2 is not statistically warranted by our sample; rather more data points should be added in a search for behavior of estimated functions with even more varied VOC and NO_x mixtures.

Estimation of the Production Function f_1

f_1 was first estimated using the projection pursuit regression routine ppr in the R statitical language. This technique suggested nothing more than a simple spline fit of $\phi \bullet \nu$. A spline-fit method with generalized cross-validation was employed; non-decreasing functions were required to avoid-overfitting the top 5% situations of ozone production. Other terms may be added using ppr. Terms involving the production of radicals involving ozone photolysis; a large variety of terms expressing in terms of simple measurable species, e.g., reactions of the excited product $O(^1D)$ with water, and reactions of OH radicals with CO were attempted; these did not add significant amounts of explained variance. A natural question is the sufficiency of the best possible fit using only NO, let us call it $f_0([NO])$, and the best fit using both NO and UV, f_{00} (j_{HCHO}[NO]). These successfully explained 50–60% and 63–69% of the variance, short of the 79–84% possible by using all three factors. The higher values quoted for $f_0([NO])$ and f_{00} (j_{HCHO}[NO]) had features that made them appear accidental or misleading, i.e., with wiggles or notable decreasing behavior as NO approached higher values. It is interesting to note the correlation of the quantities $f_1(\alpha)$ / [NO] with $P_{O_3}^{\infty}$ / [NO], to address the fact that we are correlating [NO} with a function of [NO]. The correlation is $r = 0.67$ ($r^2 = 0.45$), indicating a substantial variance explained associated with this portion of P_{O_3}, independent of [NO] self-correlation.

By excluding only 5 points from α-based fit with significantly poorer fit than 839 others, the estimated $f_1(\alpha)$ could achieve a variance explained of 84%. However, we could not exclude

the possibility that the five points represented genuine if anomalous plumes with special composition in the rural atmosphere sampled.

Estimation of the Kinetics-based Function Form Q_2

Useful insight into the character of f_1 is available by statistical estimation using a different approach. We attempt to mimic the mathematics of the basic kinetics, while continuing to use only commonly available observational data. It is best to describe a three-parameter fit of the HO_2 radical concentration, although one parameter must be set. Since HO_2 radicals persist only a few seconds in daylight, we make a quasi-steady state approximation to the HO_2 radical concentration ($d[HO_2]/dt$ << chemical sink and source terms). The quadratic relationship

$$2\eta k_{HH}\,[HO_2]^2 + (1-\rho)k_{NH}[NO]\,[HO_2] = \gamma\, j_{HCHO}[HCHO]$$

equates the loss terms due to peroxy radical self reaction and NO reaction with the a source term. k_{NH} and k_{HH} refer to the reaction rate coefficients of HO_2+NO and HO_2+ HO_2 respectively.

Each term contains a parameter modifier (η,ρ, or γ) that attempts to broaden the scope of kinetic processes contemplated while keeping a simple form quadratic in $[HO_2]$. The factor γ on the right-hand side expresses the multiplier on $j_{HCHO}[HCHO]$ that expresses the instantaneous source of *new* HO radicals. If only HCHO photolysis provides radicals, $\gamma = 2$.

The sink of HO_2 due to self-reaction is of course $2k_{HH}\,[HO_2]^2$. Hydroperoxy radicals may also be destroyed by reaction with organic peroxies, RO_2, and RO_2 frequently correlates with HO_2 closely, since RO production often results from or creates a reaction that also produces HO_2. Statistical fits will pick this up. The rate coefficient differs from that of HO_2 self reaction, and the reaction only destroys one HO_2. All of these complexities must be summarized by one parameter, η.

The role of ρ may seem indirect, but a significant value, countering the HO_2+NO reaction, is absolutely required for the quadratic relationship to have the right behavior as a function of NO. The alternative of an incorrect, very low measured direct rate k_{HN} has been repeatedly studied and seems very improbable [JPL, 2006]. Of course we have not yet accounted for recycling of radicals, e.g., OH+CO -> H + CO_2 -> HO_2. Our best estimate for this is simply to parameterize this based on the HO_2+NO rate; in very long radical-chain situations, ρ approaches 1. Note that since the parameterization contains the factors [NO][HO2], it is placed on the left-hand side of the quadratic. The estimation does not work well without a consequential value of ρ, for $\rho = 0$ the behavior of HO_2 with increasing NO cannot be fit. We surmise that the effect of ignoring statistical correlations of NO with some process like this may be noted in at the bottom of Figure 7 of Ren et al. [2008]; possibly the model and observations have differing chain lengths. (For values of NO below 50 ppt in that figure, HO_2 self reaction dominates and there is no effect.) Thus ρ represents all rapid processes more dependent on NO than on photolysis, and might be used to derive the chain length of the rapid OH-HO_2 process.

The solution to the quadratic can be written,

$$[HO_2] = \gamma\, j_{rads}[HCHO] \Big/ \left((1-\rho)k_{NH}[NO] + \sqrt{(1-\rho)^2 k_{NH}{}^2[NO]^2 + 4k_{NH}\gamma\, j_{rads}[HCHO]} \right)$$

and the primary production rate of ozone is then obtained by multiplying by $k_{NH}[NO]$. The three parameters must η, $(1-\rho)$, and γ all vary in a proportion. Assuming $\eta = 1.5$, we have $\rho = 0.77$ (seems high) and $\gamma = 13.8$. Proportionally lower values may be more realistic. All of these are reasonable values, but the set (η, $1-\rho$, and γ) are linearly dependent in this analysis, and can each move up or down by the same factor to values that the reader may prefer.

POGO-FAN

We have named the estimation technique "Production of Ozone by Guaging of (organic) Oxidation: Formaldehyde and Nitric oxide (POGO-FAN) in recognition of the difficulty and opportunity that the control of smog ozone poses for our society. Smog-producing ways, probably unnecessary, seem to be embedded in of our daily habits. The invented character Pogo is famous for opining that "we have met the enemy, and he is us.," a wry

reunderstanding of Commander Oliver Parry's triumphal message during the War of 1812: "We have met the enemy and they are ours."

References beyond those presented in the main text

Chatfield, R. B. and Delaney, A. C.: Convection links biomass burning to increased tropical ozone: However, models will tend to overpredict O3, J. Geophys. Res., 95, 18 473–18 488, 1990.

Fried, A., S. Sewell, B. Henry, B. P. Wert, T. Gilpin, and J. R. Drummond (1997), Tunable diode laser absorption spectrometer for ground-based measurements of formaldehyde, J. Geophys. Res., 102(D5), 6253–6266.

Shetter, R., Shetter, R. E., L. Cinquini, B. L. Lefer, S. R. Hall, and S. Madronich, Comparison of airborne measured and calculated spectral actinic flux and derived photolysis frequencies during the PEM Tropics B mission, J. Geophys. Res., 107, 8234, doi:10.1029/2001JD001320, (2002).

Wood, S.N., Stable and efficient multiple smoothing parameter estimation for generalized additive models, Journal of the American Statistical Association, 99, 673-686, (2004).

POGO information currently available at http://www.pogo-fan-club.org/faq.html and likely visible by internet search. See also introduction to W. Kelly, Potluck Pogo (Simon and Schuster, New York, 1955).

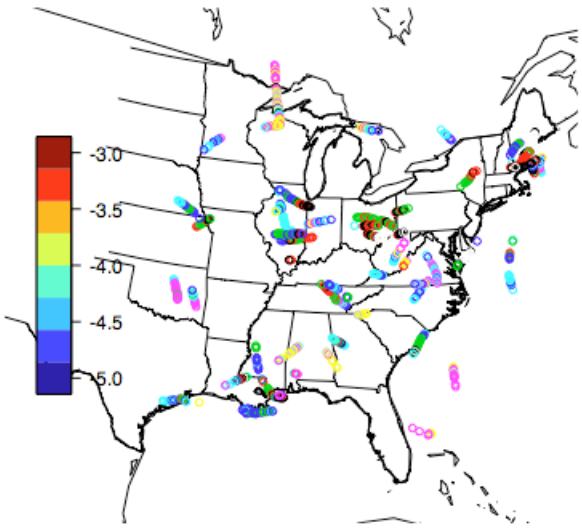

Figure S1. Widely differing smog-production conditions as sampled in the continental boundary layer ozone by the NASA DC-8 during INTEX-NA. Color scale refers to log10 of the "formaldehyde activity" divided by the NO_x concentration, in ppt units. Formaldehyde activity is taken to be the formaldehyde concentration times its photolysis rate (s^{-1}) to radicals and CO, and is described below to be one measure of VOC weighted by reactivity. Approximately 1800 samples over the populated regions of Central and Eastern North America were included, or samples potentially just downwind. Altitudes up to 1300 m ASL were included since the DIAL observations suggested that they were relatively similar in ozone and aerosol characteristics, e.g., influenced by cloud mixing. We conclude that a wide variety of NO_x-limited and VOC-(radical-production-) limited areas were sampled.