

## Controls on urban ozone production rate as indicated by formaldehyde oxidation rate and nitric oxide

Robert B. Chatfield<sup>a,\*</sup>, Xinrong Ren<sup>b</sup>, William Brune<sup>c</sup>, James Schwab<sup>d</sup>  
Queens College and PMTACS Measurement Teams

<sup>a</sup> Earth Sciences, NASA Ames Research Center, MS 245-5, Moffett Field, CA 94035, USA

<sup>b</sup> Rosenstiel School of Marine and Atmos. Science, Univ. of Miami, Miami, FL 33149, USA

<sup>c</sup> Department of Meteorology, Pennsylvania State University, University Park, PA 16802, USA

<sup>d</sup> Atmospheric Sciences Research Center, University at Albany, State University of New York, Albany, NY 12203, USA

### ARTICLE INFO

#### Article history:

Received 1 August 2009

Received in revised form

27 August 2010

Accepted 28 August 2010

#### Keywords:

Tropospheric ozone

Smog

Ozone abatement

Atmospheric chemical mechanisms

New York City

### ABSTRACT

Several strong statistical relationships quantifying local ozone generation are found which use only easily measured variables: nitrogen oxides ( $\text{NO}_x$ ), formaldehyde (HCHO), its photolysis (i.e., UV), and temperature ( $T$ ). A parameterized regression developed for rural air was adapted to central Queens, New York City, i.e., considerable fresh emissions. Measurements of the radicals  $[\text{HO}_2]$  and  $[\text{OH}]$  were available. These provided explicit reference estimates of the predominant terms for chemical ozone production,  $P_o(\text{O}_3) = k[\text{HO}_2][\text{NO}]$ , of the predominant chemical loss of nitrogen oxides,  $L(\text{NO}_2) = k[\text{OH}][\text{NO}_2]$ , and also their ratio. (This is termed a production efficiency for  $\text{O}_3$ .) Chemical modeling supports a robust extension from  $P_o(\text{O}_3)$  to total chemical production,  $P(\text{O}_3)$ . The two regression variables,  $[\text{NO}]$  and  $j_{\text{HCHO}} = \text{rads} \times [\text{HCHO}]$ , which best explain  $P_o(\text{O}_3)$ , have low correlation,  $R \sim 0.2$  (variable, interacting urban plumelets?). In our analysis,  $R^2$  for  $P_o(\text{O}_3)$  (and an estimate for its rate-determining  $[\text{HO}_2]$ ) was in the range 0.48–0.81.

Signally, the method suggests a quantitative and very local application of descriptions of “VOC limitation” or “ $\text{NO}_x$  limitation” to  $P(\text{O}_3)$  and  $L(\text{NO}_2)$ , expressed as dimensionless sensitivity variables. Unexpected sources, transport, or chemistry may be highlighted using only HCHO,  $\text{NO}_x$ , and UV radiation. More complex relationships are needed in a focused analysis of intermediate polluted situations, where timescales or individual sources may give trouble. Here, we find that  $T$  is informative, and cooperates with  $j \times [\text{HCHO}]$  in defining  $[\text{HO}_2]$ . Sensitivities for radicals and  $\text{NO}$  for  $P_o(\text{O}_3)$  are similar  $\sim 0.4$ , but sensitivities for radicals and  $\text{NO}_2$  for  $L(\text{NO}_2)$  emphasize  $\text{NO}_2$ . Remaining variability in the statistical estimates of  $P_o(\text{O}_3)$  and  $L(\text{NO}_2)$  is modulated by incompletely understood, slowly varying gain factors. Understanding of these gain factors promises a better empirical indicator for  $P_o(\text{O}_3)/L(\text{NO}_2)$ . Complete 3-d simulations are not replaced, but this view helps separate sub-problems in the estimation of  $\text{HO}_2$  and  $P(\text{O}_3)$ .

© 2010 Elsevier Ltd. All rights reserved.

### 1. Introduction

The origin of high levels of boundary layer smog  $\text{O}_3$  is a concern for health and agricultural economic reasons; concern has motivated a new, lower regulatory standard (U.S. EPA, 1987, 2008). Lofting and long-distance transport of ozone and its precursors also cause concern regarding  $\text{O}_3$  as a greenhouse gas and an intercontinental pollutant (Jacobson, 1998).

The processes determining levels of smog ozone are complex, involving the chemical production, destruction, surface uptake, and

transport of ozone, complicated all the more by similar processes determining its precursors. The chemical production term for ozone is perhaps the most difficult to describe from first principles, since it depends on the levels of peroxy radicals, both  $\text{HO}_2$  and organic peroxy radicals  $\text{RO}_2$ .

The  $\text{O}_3$  gross production rate is described by a principal term and additional ones:

$$\begin{aligned} \text{O}_3: & \quad \text{HO}_2 + \text{NO}^- \rightarrow \text{OH} + \text{NO}_2 \\ P_o(\text{O}_3) & = k_{\text{HO}_2+\text{NO}}[\text{HO}_2][\text{NO}] \\ P(\text{O}_3) & = k_{\text{HO}_2+\text{NO}}[\text{HO}_2][\text{NO}] + \sum_{i=1,n} k_{\text{RO}_2+\text{NO}}[\text{R}_i\text{O}_2][\text{NO}] \end{aligned}$$

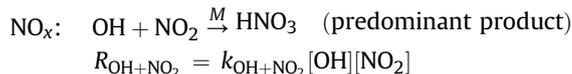
The number  $n$  of organic peroxy radicals can run to the hundreds and thousands. The “principal” production rate  $P_o(\text{O}_3)$  due to  $\text{HO}_2$

\* Corresponding author.

E-mail address: [robert.b.chatfield@nasa.gov](mailto:robert.b.chatfield@nasa.gov) (R.B. Chatfield).

radicals is typically ~60% of the total production rate  $P(\text{O}_3)$  in this and other datasets (see Section 3.4). Peroxy radicals are measured only in specialized field campaigns, due to the technology required (Eisele et al., 1997; Brune et al., 1998), so they are commonly estimated by calculating the interaction of intricate, branching organic reactions (Seinfeld and Pandis, 1998).

Nearly as important is a second model calculation involving a radical species and  $\text{NO}_x$ :



This dominates the removal of  $\text{NO}_x$ , and  $\text{NO}_x$  in turn controls  $\text{NO}$  and  $\text{HO}_2$ . The concentrations of all secondary organic compounds are determined by myriad reactions (e.g. MCM v3.1, Jenkin et al., 1997), with most kinetic processes inferred by analogy with a few measured ones (Atkinson and Arey, 2003; Atkinson et al., 2005, 2006).

In work leading to Chatfield et al. (submitted for publication), we suggested a diagnostic method for understanding smog ozone throughout mostly rural areas of the Eastern US. In a similar situation where  $\text{HO}_2$ ,  $\text{NO}$ , etc. were measured on an aircraft, we found remarkably simple empirical functional forms allowing estimation of  $P_0(\text{O}_3)$  and  $\text{HO}_2$ . The variables were  $[\text{NO}]$  and the product  $j_{\text{HCHO} \rightarrow \text{rad}} \times [\text{HCHO}]$ ; the latter helps quantify the organic oxidation rate which is linked to  $[\text{HO}_2]$  and  $[\text{RO}_2]$ . We term the general method POGO, for Production of Ozone by Gauging (organic) Oxidation. A function of these two variables reproduced the  $P_0(\text{O}_3)$ , method reminiscent of the breakthrough “contour plot” or “empirical kinetic model approximation” method used in the 1980s (Dodge, 1977; Shafer and Seinfeld, 1985; Seinfeld and Pandis, 1998, p. 301), and of several analyses since (e.g., Sillman and He, 2002). HCHO is our reasonable choice to trace organic oxidation and the production of  $\text{HO}_2$  radicals because it is a near-terminal breakdown product in reaction chains resulting from a vast majority of VOC emissions. Its few-hour lifetime makes it a good local gauge of VOC oxidation activity.

The method is best introduced if we may assume that daytime HCHO is not too far from a quasi-steady state (QSS) determined by VOC's. HCHO destruction rates vary strongly with UV radiation; indeed HCHO can persist all night away from surface. If one forms the product  $j \times [\text{HCHO}]$ , where  $j = j_{\text{HCHO} \rightarrow \text{CHO}+\text{H}}$  is the photolysis of HCHO to radicals, one creates a gauge of the flowrate of organics, a gauge that should be related to the production and thus the concentration of reactive  $\text{HO}_2$  and  $\text{RO}_2$ . This particular  $j$  is also a good choice since it responds to more energetic UV, and so, and so correlates with harder UV processes creating radicals, those involving  $\text{O}(^1\text{D})$ , peroxides and other aldehydes beyond HCHO. Economical measurements of a UV/Vis photolysis rates using sun angle, total ozone column (e.g., from satellite measurements), can easily provide this  $j$ . In moving to an urban sampling site, we explored whether urban emission could establish HCHO at informative levels, and whether primary HCHO would compensate or overcompensate the short time for a QSS downwind of large fresh urban emissions.  $j \times [\text{HCHO}]$  and  $\text{NO}$  (as we shall abbreviate them below) again proved useful and suggested generalizations.

Our approach to an empirical description concentrates on a local description rather than the time-integrated  $P(\text{O}_3)$  from a hypothesized dominating source. Two previous approaches describing  $P(\text{O}_3)$  are particularly useful as they develop themes that complement ours. One approach relied more on kinetic models applied in idealized situations (a single typical VOC– $\text{NO}_x$  mix characterizing a city or region) (Kleinman et al., 2002). Another described controls on  $P(\text{O}_3)$  using correlations of relevant  $\text{HO}_2$  and  $\text{OH}$  radical reactions with those producing other oxidation by-products (e.g.,  $\text{H}_2\text{O}_2$  and  $\text{HNO}_3$ ) (Sillman and He, 2002).

The goodness of fit that we find using HCHO in our procedure turned out to hide some complications that are revealed when we focus on pollutant levels of  $\text{NO} < 8$  ppb; we will describe successive levels of approximation to  $P_0(\text{O}_3)$ . Generally similar relationships held, but there seemed to be a slow modulation of a gain factor. We seek some explanation for this variation. Similar slow variability is found in fits made for  $L(\text{NO}_2) \sim R_{\text{OH}+\text{NO}_2}$ . These may track some measures of VOC reactivity revealed by measurements of VOC.

## 2. Sampling location and instrumentation

Sampling was at a small tower located at Queen's College, in central Flushing, New York City; the sponsoring PM2.5 Technology Assessment and Characterization Study (PMTACS) study focused on aerosols from the city and regionally. Gas-phase concentrations suggested both accumulated and local sources, lower when fresher Atlantic air flowed a short distance across the city. The site has commercial, residential, and industrial sources, and is ringed by freeways and arterials. Further details are in Ren et al. (2003a).

Ren et al. (2003a) detail the laser-induced fluorescence measurements of  $\text{HO}_2$  and  $\text{OH}$  that were used to quantify  $P_0(\text{O}_3)$  using GTHOS (ground-based tropospheric hydrogen oxides sensor).  $\text{NO}$ ,  $\text{NO}_2$ ,  $\text{O}_3$ , and other species were measured at a PAM site (EPA, 1994). HCHO was measured based on the Hantzsch reaction (Dasgupta et al., 1988). Hourly VOC measurements were made 2.5 km away. Histograms of pertinent variables to this paper are shown in Fig. 1(a). Fig. 1(b) describes a “moderate pollution” subset. Ren et al. (2003b) describe the VOC and  $j$ 's based using a shadowband UV radiometer, and compare measured and modeled  $\text{HO}_2$  and  $\text{OH}$ .

Ren et al.'s (2003b) simulations tracked  $\text{HO}_2$  and  $\text{OH}$  well over the period July 10–August 2, 2001. Ren did note systematic phenomena: Modeled  $\text{HO}_2/\text{OH} > \text{HO}_2/\text{OH}$  ratios for  $\text{NO} < 1$  ppbv, but reversed for  $\text{NO} > 8$  ppbv. Ren et al. (2003b) summarized: “This difference is consistent with measured  $\text{OH}$  being greater than modeled  $\text{OH}$  at low  $\text{NO}$ , while measured  $\text{HO}_2$  is much greater at high  $\text{NO}$ .”

## 3. Results and discussion

### 3.1. Organic reactivity, $\text{NO}$ , and empirical estimates of $P_0(\text{O}_3)$

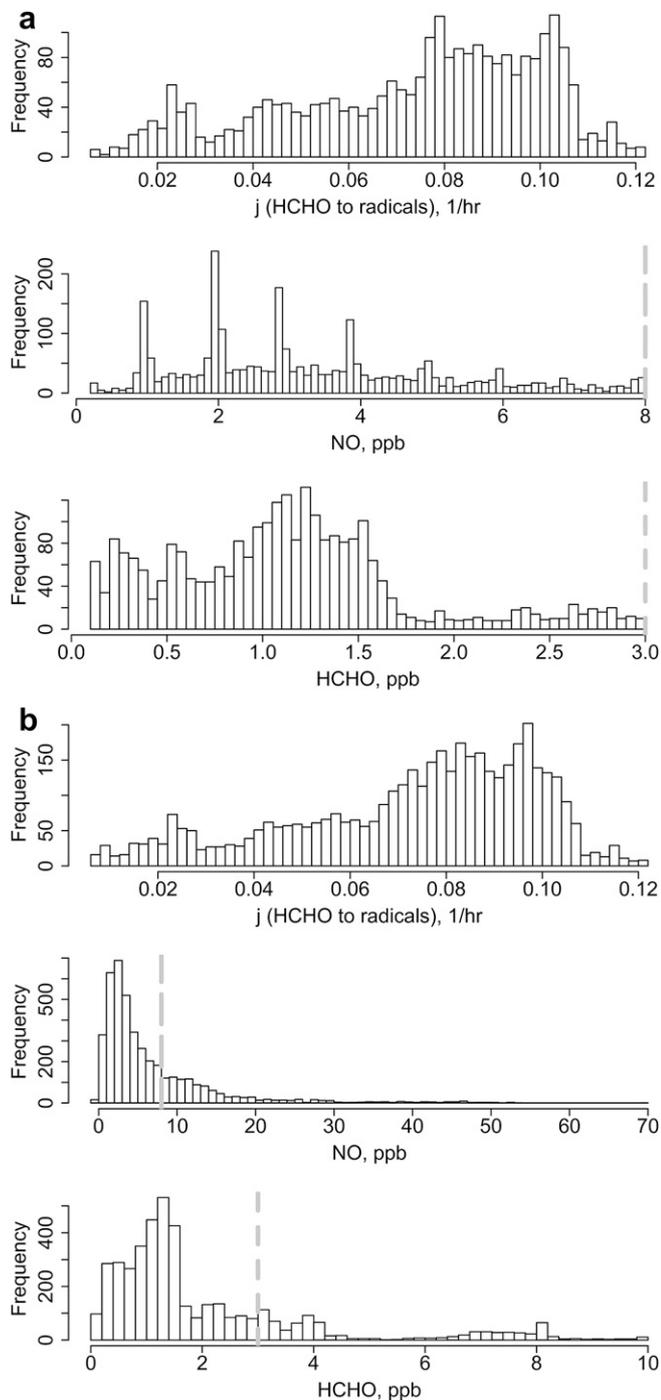
Following ideas from work on regional smog ozone (Chatfield et al., submitted for publication), we sought relationships between  $[\text{NO}]$ , VOC reactivity, i.e.,  $j \times [\text{HCHO}]$ , and  $P_0(\text{O}_3)$ . A very simple relationship  $\text{constant} \times (j \times [\text{HCHO}] \times [\text{NO}])^a$  was useful, where  $a \sim 0.4$ , and the product  $j[\text{HCHO}] \times [\text{NO}]$  was below  $\sim 0.4 \text{ ppb}^2 \text{ h}^{-1}$  and  $P_0(\text{O}_3)$  was  $< 10 \text{ ppb h}^{-1}$ . It seemed reasonable that organic reactivity and  $\text{NO}_x$  might influence  $P_0(\text{O}_3)$  differently in cities, so we sought a more general relation:

$$P_0(\text{O}_3) = C(j \times [\text{HCHO}])^a [\text{NO}]^b$$

or,

$$\log P_0(\text{O}_3) = \log C + a \log(j \times [\text{HCHO}]) + b \log[\text{NO}]$$

The latter form allows a simple linear regression, and is a preferable form for least-squares regression, since residuals are most nearly normally distributed. A non-parametric interpolation of the dependence of  $P_0(\text{O}_3)$ , shown later, supports the choice. The residuals also tend to have variance constant with the magnitude of  $\log P_0(\text{O}_3)$ , i.e., errors are proportional to  $P_0(\text{O}_3)$ . A negative aspect of this proportionality is that large errors for a small percentage of observations (at highest values) must limit overall variance explained; a positive is that the relationship works well in detail at low  $P_0(\text{O}_3)$ .



**Fig. 1.** Histograms of observed distributions of quantities relevant to  $P_0(O_3)$  sampled at Queens College, (a) in all complete samples including highly polluted cases and (b) in a selection of “moderate pollution” cases specified by the condition that NO concentrations be less than 8 ppb (vertical long dashed line). This constrained the range of HCHO also. The NO concentrations in (b) show a preference for even-ppb values, considered an artifact of recording. This artifact may have limited  $R^2$ .

If  $j \times [\text{HCHO}]$  is considered to reflect the oxidation rate of volatile organics, the formula should be most appropriate when HCHO change satisfies a timescale based on the sink terms,  $\tau_{\text{HCHO}} = 1/(k_{\text{OH}+\text{HCHO}}[\text{OH}] + j_{\text{HCHO} \rightarrow \text{rads}} + j_{\text{HCHO} \rightarrow \text{H}_2})$ . All the sink terms are at maximum near midday. To allow for a close, not greatly time-delayed, relationship between radicals and  $P_0(O_3)$ , we made empirical statistical models of the data observed between 10 AM and 4 PM local daylight time (approximately 9 AM and 3 PM sun

time). Since pollution levels, especially [NO], were occasionally extremely high, it seemed appropriate to consider two situations, “All Pollution,” and “Moderate Pollution,” the former testing usefulness over an extended range, the latter focusing on details.

### 3.2. Results and discussion: $P_0(O_3)$ including heavily polluted situations

We begin by considering all pollution samples. Somewhat counter-intuitively, parameterizations of the whole dataset, not subsets, yield the highest  $R^2$  for  $P_0(O_3)$ . The large range of  $P_0(O_3)$ , given the same amplitude of noise can raise  $R^2$ , if the relation is robust. The maximum NO mixing ratio of 25 ppb suggests recent pollution was often sampled. It is doubtful that the HCHO timescale condition was satisfied unless older emissions VOC emissions still had a large role. Nevertheless, the estimated empirical fit was good, with

$$a = 0.430 \pm 0.006 \text{ for HCHO activity}$$

$$b = 0.483 \pm 0.006 \text{ for NO}$$

$$C = 1.5 \pm 1.17, c = 0.18 + 0.07$$

$$R^2(\log \text{ vs } \log) = 0.73, R^2(\text{non-log, molec cm}^{-3} \text{ s}^{-1} \text{ or ppb hr}^{-1} \text{ terms}) = 0.71$$

$F = 5075$  on 2 and 4911 dof (extremely significant), residuals normally distributed

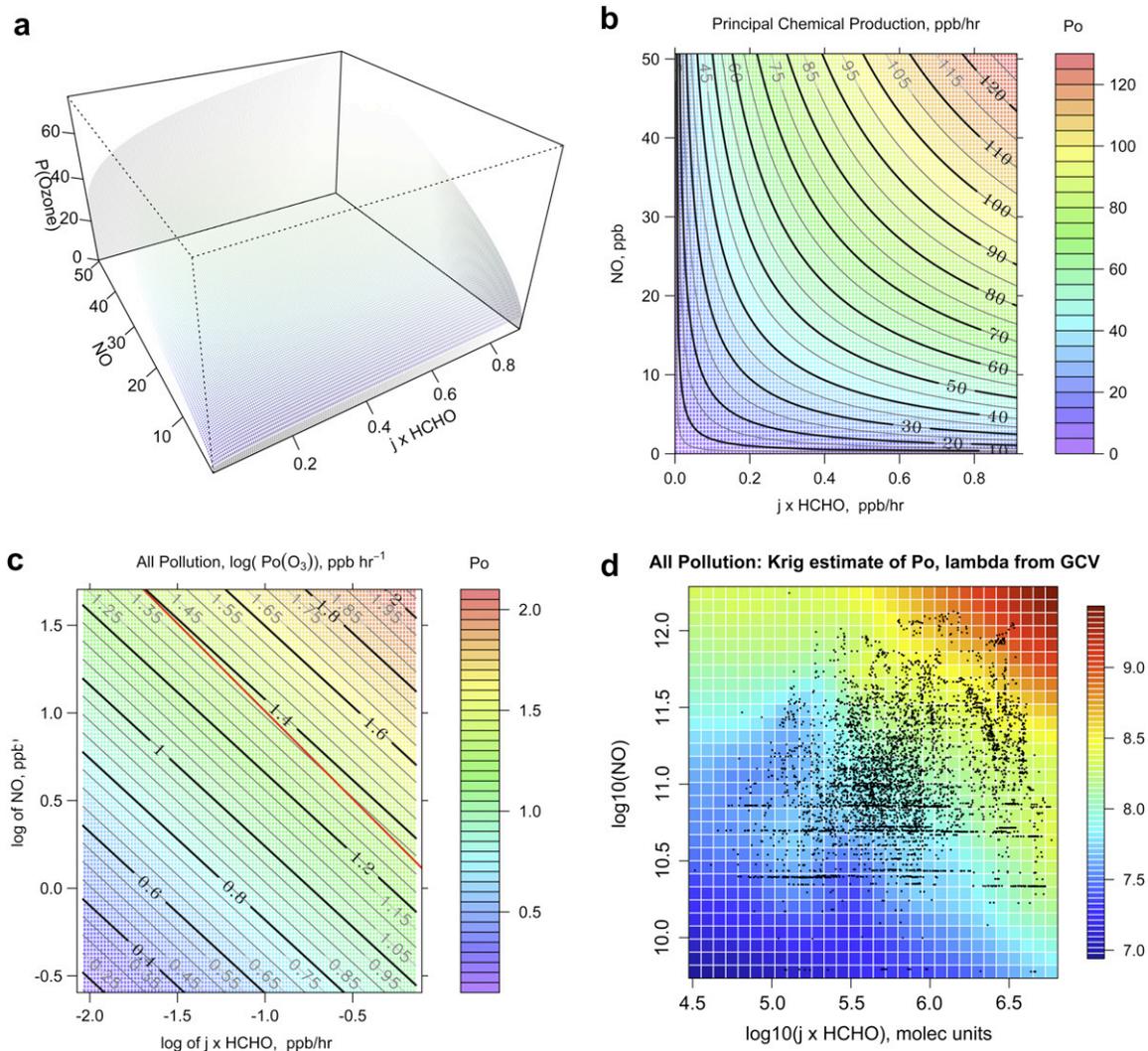
All estimates were formally determined to a  $p$ -value of  $< 2 \times 10^{16}$ . These estimates for  $a$  and  $b$  are relatively close to estimates made for airborne sampling of urban, rural, and some remote boundary layers in separate Central-US/East-Coast and California sampling (Chatfield et al., submitted for publication). Fig. 2 shows the relationship defining  $P_0(O_3)$  to  $j \times [\text{HCHO}]$  and NO, in several ways: a 2-d surface, a contour plot, and a graph drawn when all three quantities are expressed in terms of  $\log_{10}(\text{quantity})$ , and an interpolation independent of the regression formula. Note the simplicity of the log-log plots.

Adding variables related to radicals significantly improves the model's prediction. Consider the sequence of three regressions using NO only, NO with  $j$ , and NO with  $j \times [\text{HCHO}]$ . The  $R^2$  values calculated with ppb (not log) quantities improve from 0.48 to 0.58 to 0.73 as information is added.

A reviewer asked us to seek a regression allowing separate exponents for  $j$  ( $a_1$ ) and for [HCHO] ( $a_2$ ), with the same meaning for  $b$  and  $c$ . This could highlight photolysis producing  $\text{HO}_2$  by reactions uncorrelated with HCHO. We find  $a_1 = 0.778$ ,  $a_2 = 0.335$ ,  $b = 0.541$ ,  $c = 2.19$ . The  $R^2$  for separate  $a_1$  and  $a_2$  rises to 0.81; given 4908 degrees of freedom (dof), the  $F$  statistic appears very significant. Since  $a_1 > a_2$ , it seems that  $P_0(O_3)$  is formally more sensitive to photolysis rate ( $j$ ). However, the inclusion of [HCHO] is more *informative*, raising  $R^2$  from 0.58 ( $j$  but no HCHO) to 0.73 (unseparated  $a$ ) or perhaps 0.81 (separated  $a_1$ ,  $a_2$ ). The relative difference between each method's predictions amounts to  $\leq 1.5\%$  on only during restricted periods (day 205, the afternoon of 207), one half hour early (208, 10 AM), and one hour late (214, 4 PM). This also suggests some autocorrelation in time, i.e.,  $\ll 4908$  dof. Variation in sources and transport most likely determines the useful generalization of our estimates, not sample-to-sample instrumental errors. Separating  $a_1$  and  $a_2$  exponents may well be helpful; however, more varied conditions of observation are required.

### 3.3. Re-thinking sensitivities of $P(O_3)$ — proportional precursor changes

The use of log values for predictors of  $P_0(O_3)$  simplifies not only the regression technique, but also display and interpretation. They



**Fig. 2.**  $P_o(O_3)$  as a function of organic activity as measured by HCHO photolysis and of NO, portrayed in several ways (all cases). (a) Contours shown on a three-dimensional surface, emphasizing the nature of diminishing  $P_o(O_3)$  as a function of each coordinate: units are in ppb and hours; i.e.,  $P_o(O_3)$  in ppb/h, NO in ppb, and HCHO activity in ppb/h. (b) The same information displayed in the contour-plot style used most commonly with Empirical Kinetic Model (EKMA) methods, also in ppb units. (c) A plot using logarithmic quantities, with the abscissa and ordinate shown in equal aspect ratio, i.e., equal relative sensitivity  $\partial \log_{10} P_o(O_3) / \partial \log_{10}(\text{coordinate})$ . Molecular units: all rates and statistical fits. (d) A Krig interpolation of  $\log_{10} P_o(O_3)$ , molecular units.

express logarithmic-derivative “sensitivities” in a technical use of the term:  $\partial \log P_o(O_3) / \partial \log(j \times [\text{HCHO}])$  and  $\partial \log P_o(O_3) / \partial \log \text{NO}$ . The slope of the isopleths in the log–log plots summarizes the ratio of the sensitivity to NO and HCHO activity. A sketched diagonal 1:–1 lines on the log–log plots in this report represent equal exponents or equal sensitivities for the control variable selected, *independent of units*. Spacing of the contours also yields an easily calculated function of the power laws obeyed.

Comparing the linear and log depiction of  $P_o(O_3)$  the log depiction is more intuitive and easy to read. However, the prominent feature of a central mound indicating “most efficient mixtures for  $P_o(O_3)$ ” disappears. This is only appearance. While equal horizontal (or vertical) increments in Fig. 2(a) can be interpreted as equal additions or subtractions of ppb s<sup>-1</sup> or ppb, equal horizontal (vertical) increments in Fig. 2(b) are equal *relative* changes in concentration (i.e., 20% or 60% reduction). (Physical conditions for  $j$  are assumed constant.) Traditional terms, “optimum oxidant production regions” may be seen in Fig. 2(a,b). These lie away from the origin along the axes: maximum sensitivity to increased NO along the horizontal axis, maximum for VOC additions along the

vertical. Parameters are further discussed in considering HO<sub>2</sub>, Section 3.5.

A note on O<sub>3</sub> budgeting. Ozone is best described by the common generalized “oxidant” quantity conserved for longer times,  $\text{O}_3 + \text{NO}_2 + \text{O}(\text{}^3\text{P}) + \text{O}(\text{}^1\text{D}) + 2\text{NO}_3$ . This oxidant quantity has main destruction reactions like  $\text{OH} + \text{NO}_2$ ,  $\text{O}_3 + \text{alkene}$ ,  $\text{O}_3 + \text{HO}_2$  or  $\text{O}(\text{}^1\text{D}) + \text{H}_2\text{O}$ . NO “titration” of O<sub>3</sub> is ignored. An NO/NO<sub>2</sub> steady state is not strictly required. Unlike other descriptions of smog formation, the parameterization never decreases with increasing amounts of [NO] or  $j \times [\text{HCHO}]$ . Fig. 2’s upper corners along the NO axis show continued increase. This behavior is supported by a Krig interpolation of the  $P_o(O_3)$  surface shown in Fig. 2(d). Krig analysis uses variability of nearby point values of  $P_o(O_3)$  in a region to limit the fitting and allows for a background variability described by 2nd order polynomials. (Documentation for the R routine Krig in the *fields* package describes our choices.)

Added chemical parameters tended not to add much significance. Separating  $a_1$  and  $a_2$  as above, illustrates a useful direction, refining radical sources to specific conditions. Exact parameters might vary for observations that contain very high pollution, high

pollution, or regional dilute pollution; variable-exponent models seem reasonable, given the data.

Once costs of emission control are brought into consideration, a diagram will allow an asymptotic estimate of an optimal control strategy for a locale. True determinations of the control strategy of course depend (a) on the effect of upwind  $\text{NO}_x$  emissions on local NO, (b) of VOC emission mixtures on HCHO, and (c) other processes affecting the integral production of  $\text{O}_3$  upwind.

### 3.4. Substantial variability in radical-precursor/NO mixes

The previous section implied evaluation over many different combinations of ( $j \times [\text{HCHO}]$ ) and NO. However, it is natural to think of urban pollution as having very correlated sources of VOC and NO, e.g., traffic sources. Correlated precursors would imply that our contour graphs would have experimental relevance only along a diagonal band of concentrations. We found instead a remarkable variability of the precursor mixes for this city site. Fig. 3(a,b) shows views of a 2-d histogram of the simultaneous pairs of ( $j \times [\text{HCHO}]$ ) and NO. A logarithmic scale was necessary; even log plots portrayed little correlation. Similar histograms obtain for HCHO vs NO.

Fig. 3(c) gives a view of the course of the ratio ( $j \times [\text{HCHO}]/\text{NO}$ ) throughout each day of sampling. The ratio can change by a factor of 10 over 3–6 h, or even faster. The ratio typically increases from 10 AM to the afternoon, with a possible decrease in late afternoon. Changes could be due to the rapid decline of morning-rush  $\text{NO}_x$  from traffic, and to high sun angle and UV favoring radical-producing photolytic rates. Day-to-day variations might reflect wind and weather conditions; shorter-term variability might reflect local sources.

### 3.5. Estimates of $\text{HO}_2$ for model and mechanism improvement

While estimates of  $P_0(\text{O}_3)$  may give an overview of the air pollution process, the most practical use of our statistical analysis technique is as a comparison check to simulated  $\text{HO}_2$  and  $\text{RO}_2$ . Regressions estimates  $[\text{HO}_2]$  can be evaluated just as for  $P_0(\text{O}_3)$ :

$$[\text{HO}_2] = 2.019 \times 10^{11} (j \times [\text{HCHO}])^{0.435} [\text{NO}]^{-0.522}$$

i.e.  $c_h = 11.31$ ,  $C_h = 10^{c_h} = 2.019 \times 10^{11}$ ,  $a_h = 0.435$ , and  $b_h = -0.522$

$$R^2(\log \text{ vs } \log) = 0.70, R^2(\text{non-log molec cm}^{-3} \text{ terms}) = 0.79$$

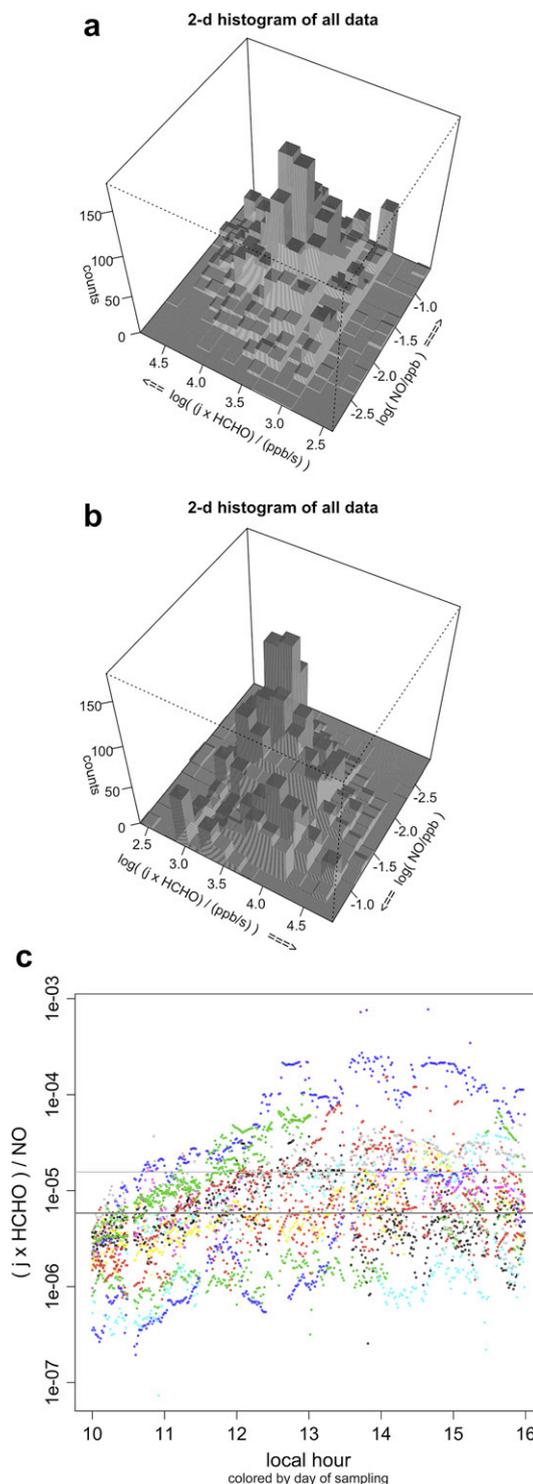
$F = 5075$  on 2 and 4278 dof (extremely significant),  
residuals normally distributed

Parameters correspond as expected from regressions for  $P_0(\text{O}_3)$ . Note that the value  $a_h = 0.435$  here is near the value  $a = 0.430$  for  $P_0(\text{O}_3)$  and that the value  $b_h = -0.522$  is nearly  $b = 0.483 - 1 = -0.517$ . Differences for  $a$  and  $a_h$ ,  $b - 1$  and  $b_h$  were  $\sim 0.005$ . We will describe  $P_0(\text{O}_3)$  estimates for other situations: estimates of  $[\text{HO}_2]$  do change consistently.

In this local formulation of sensitivity, the shape of the  $P_0(\text{O}_3)$  isopleth curves is completely determined by the dependence of the  $[\text{HO}_2]$ -indicator on  $[j \times \text{HCHO}]$  and  $[\text{NO}]$ . A simple explanation of the fall-off behaviors given by the fractional-power laws ( $a$ ,  $a_h$ ,  $b$ ,  $|b_h| < 1$ ) comes from considering an  $\text{HO}_2$  QSS:

$$[\text{HO}_2]: [\text{HO}_2] \sim P(\text{HO}_2) / \left( k_{\text{HO}_2+\text{NO}}[\text{NO}] + \sum_{i=1,n} k_{\text{HO}_2+\text{XO}_2}[\text{XO}_2] \right)$$

where  $P(\text{HO}_2)$  is the gross chemical production rate of  $\text{HO}_2$ , and  $\text{XO}_2$  has the special meaning of any species besides NO that reacts with  $\text{HO}_2$ . Commonly X may be H, R (organic radical), N (i.e.,  $\text{NO}_2$ ), or O (i.e.,  $\text{O}_3$ ). OH can play a similar role. Discussion of the chemistry may



**Fig. 3.** (a) and (b) two views of the 2-d histogram of the joint statistics of ( $j \times [\text{HCHO}]$ ) and NO, shown in logarithmic scales,  $R$  for logs  $\sim 0.17$ . Similar HCHO and NO analyses also had.  $R$  for logs was 0.28. (c) Rapid variation of the ratio of ( $j \times [\text{HCHO}])/\text{NO}$  during the days of sampling (colors) in a time-series plot with a log scale. Legend indicates day–number of the year. Sample median (dark gray line) and mean (light gray line) are shown. Note frequent variations of the ratio by  $3 \times$  to  $10 \times$  over periods of hours.

be found in Liu et al. (1988), Seinfeld and Pandis (2005), Jacobson (1998) and especially the several Sillman and Kleinman papers referenced elsewhere.

Certain features of the  $[\text{HO}_2]$  and  $P_0(\text{O}_3)$  power laws are now explained. As NO increases, removal of  $\text{HO}_2$  by  $\text{HO}_2 + \text{NO}$  is

counterbalanced by regenerated HO<sub>2</sub> from rapid VOC oxidation by OH formed in that reaction. At even higher NO, both HO<sub>2</sub> and chain-propagating OH are removed by NO<sub>2</sub> and perhaps HNO<sub>3</sub> and PAN, all associated with the increased NO levels. Consequently, the [NO] dependence of P<sub>o</sub>(O<sub>3</sub>) has  $a < 1$ . Lower P<sub>o</sub>(O<sub>3</sub>) at higher total NO<sub>x</sub> values are the common in other, similar VOC/NO<sub>x</sub> analyses. In our analysis indirect effects (on radicals) never outweigh the direct effect of higher NO in the rate  $k_{\text{HO}_2+\text{NO}}[\text{HO}_2][\text{NO}]$ .

As  $(j \times [\text{HCHO}])^a$  increases, there is progressively less increase in HO<sub>2</sub> and P<sub>o</sub>(O<sub>3</sub>), as species accompanying strong oxidation, HO<sub>2</sub>, RO<sub>2</sub>, and NO<sub>2</sub>, etc, more efficiently remove HO<sub>2</sub> from the system. To the extent that  $[\text{XO}_2] \propto [\text{HO}_2]$ , the QSS equation for [HO<sub>2</sub>] above suggests that  $[\text{HO}_2] \times [\text{XO}_2] \sim P(\text{HO}_2) = \text{constant}$ , and so  $a_{\text{h}}$  and  $a \sim 0.5$  are reasonable.

It is striking that  $b \sim a$  in many situations (Chatfield et al., submitted for publication); we understand this only partially. It could be that a perfect predictor of [HO<sub>2</sub>] would have  $a = b$ . Since  $j \times [\text{HCHO}]$  is an imperfect predictor,  $a$  may vary by situation. Since we assume a power law, the sensitivities  $a$  and  $b$  must satisfy the conditions for an isopleth. Isopleths of P<sub>o</sub>(O<sub>3</sub>) are  $P_o(\text{O}_3) = k[\text{HO}_2][\text{NO}] = C(j \times [\text{HCHO}])^a[\text{NO}]^b = \text{constant}$ . This may constrain  $a$ . The Kleinman and Sillman methodologies should be helpful to extend this analysis. Similar behavior is seen in analysis of P<sub>o</sub>(O<sub>3</sub>) and [HO<sub>2</sub>] in local box models that the authors have run, but the simulations do not help determine  $a$  and  $b$ . Ren et al. (2003b, 2008) discuss features of similar box modeling calculations.

Estimates for RO<sub>2</sub> depend on the use box-model analysis like those made by Ren et al. (2003b) for this dataset. Fig. 9 of that paper plots modeled HO<sub>2</sub> and RO<sub>2</sub>, and these provide an average ratio [HO<sub>2</sub>]/[RO<sub>2</sub>] of 1.5–1.7. This then suggests  $P(\text{O}_3) \sim 1.6P_o(\text{O}_3)$  to appropriate accuracy. RO<sub>2</sub> should respond to increasing pollution like HO<sub>2</sub> in the discussion above; P(O<sub>3</sub>) and P<sub>o</sub>(O<sub>3</sub>) should behave similarly also.

### 3.6. P<sub>o</sub>(O<sub>3</sub>) in more complex, less polluted situations

We focused in greater detail at a significant subset of the same Queens dataset, and relationships describing [HO<sub>2</sub>] levels became more complex. The focused analysis restricted the samples to [NO] ≤ 7 ppb (Fig. 1). If the lack of very high concentrations means that the emissions had more reaction time, one would expect VOC mixing (more even HCHO yield) and better QSS for HCHO. Both ideas imply statistical fits would be more accurate. Instead, we found multiple similar relationships that suggest a remaining, elusive parameter. Fig. 5 shows a series of correlation plots for successively more complex statistical relationships, extending the basic power-law formula we have used so far. Fig. 5(a) is correlation plot for modeled and observed P<sub>o</sub>(O<sub>3</sub>) based only on the HCHO activity and NO.

Table 1 gives a summary of attempts to improve regression results. This table adds temperature  $T$ , and also a “specific situation”

**Table 1**  
P<sub>o</sub>(O<sub>3</sub>) for moderately severe pollution precursor conditions.

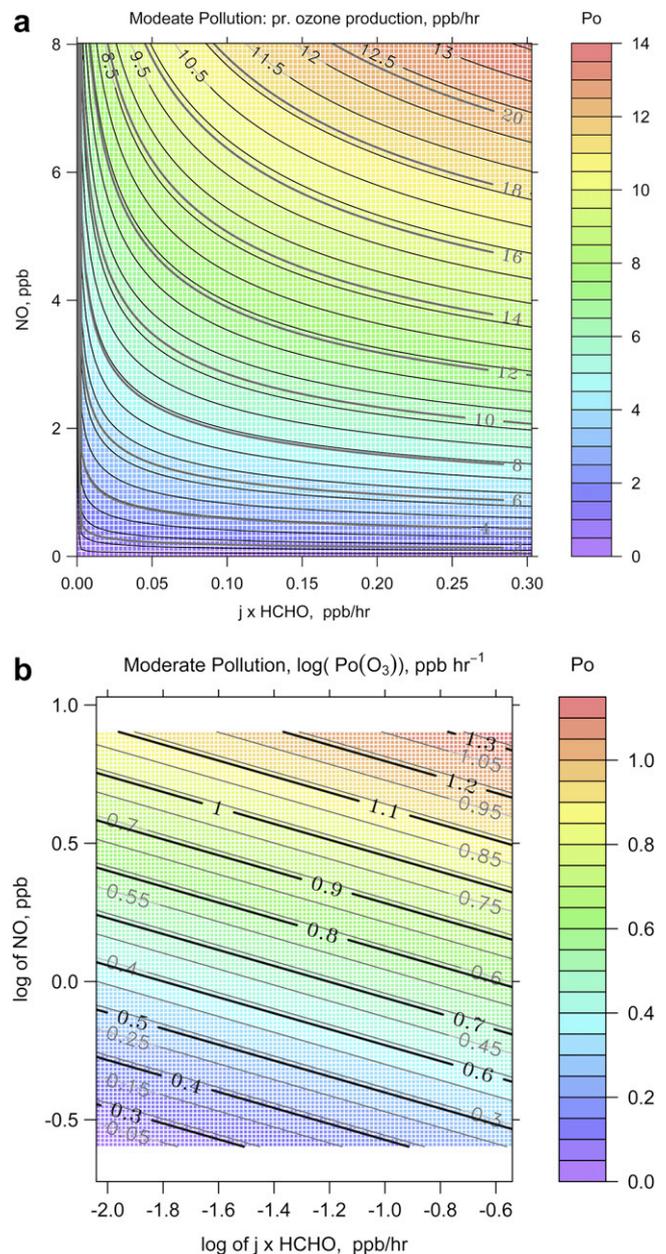
	Regression	(a, b, c)	(a, b, c, d)	(a, b, c, f(t))	(a, b, c, d, f(t))
$c$	10 <sup>c</sup>	1.16 ± 0.14	−16.2 ± 0.12	−0.54 ± 0.12	−13.1 ± 0.63
$a$	$(j \times [\text{HCHO}])^a$	0.36 ± 0.01	0.17 ± 0.01	0.28 ± 0.01	0.17 ± 0.01
$b$	(NO) <sup>b</sup>	0.41 ± 0.01	0.58 ± 0.01	0.60 ± 0.01	0.64 ± 0.01
$d$	10 <sup>d(T/298)</sup>		16.51 ± 0.38		12.82 ± 0.64
$f(t)$	Day-dependent 10 <sup>f(t)</sup>			9.0 eq. dof	9.0 eq. dof
	R <sup>2</sup> in log terms	0.65	0.75	0.75	0.79
	R <sup>2</sup> in ppb terms	0.46	0.56	0.64	0.67

indicator  $f$  which is characterized a slow (several-day) variability with time. Table 1 shows that, at first glance, ozone production is only modestly well predicted by our VOC activity and NO variables.

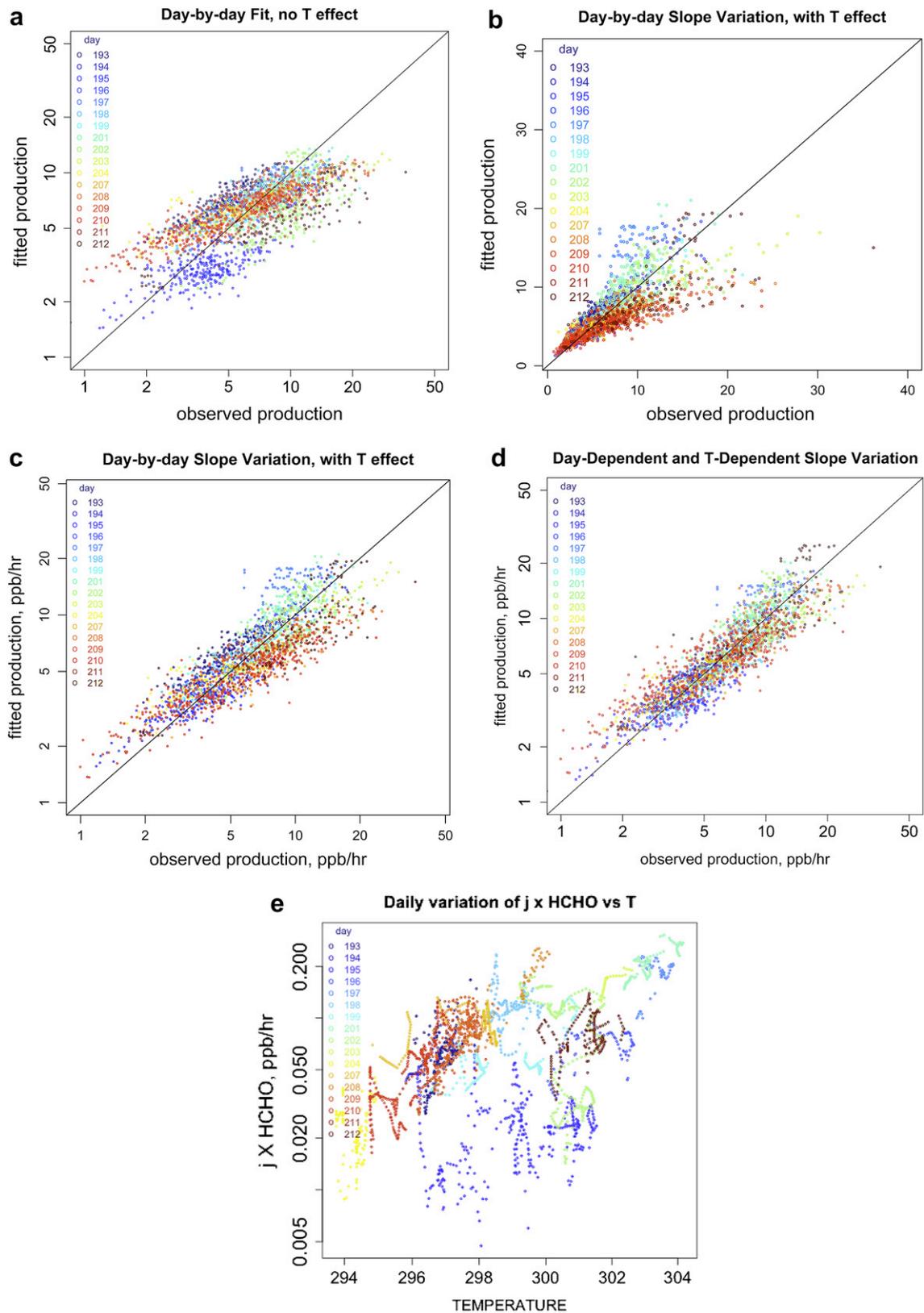
$$P_o(\text{O}_3) = c(j \times [\text{HCHO}])^a[\text{NO}]^b$$

$R^2 = 0.41$  is indicated for a relationship based on VOC activity and NO alone. Despite this low  $R^2$ , there are consistent patterns for periods of one half to a few days, as Fig. 5 shows. P<sub>o</sub>(O<sub>3</sub>) relationships vary regularly day-by-day, shown as layers in Fig. 5(a).

Temperature was the single variable most helpful in increasing explanative power. Why should including  $T$  help? High temperatures have long been associated with high ozone concentrations



**Fig. 4.** P<sub>o</sub>(O<sub>3</sub>) as a function of organic activity as measured by HCHO photolysis and of NO, portrayed in several ways. Focused-analysis moderate pollution data (NO ≤ 8 ppb). (a) Contours shown in linear terms. (b) Contours shown in log terms; axes are set at a 1:1 ratio. Heavy, color-filled contours (and side-scale) indicate the fit at  $T = 298$  K, approximately the median temperature. Gray contours indicate much greater P<sub>o</sub>(O<sub>3</sub>) at the 90th-quantile of  $T$ .



**Fig. 5.** (a) Comparison of prediction with only  $j \times [\text{HCHO}]$  and NO to observations. (b) Comparison of prediction including also  $T/298$ , linear scale exhibiting different slope on different days. (c) Same prediction including  $T/298$ , but log scale. (d) Comparison of prediction including a slowly changing  $f(t)$ , log scale. (e) Complex relationship of  $T$  and  $\log_{10}(j \times [\text{HCHO}])$ . Colors again note the year–day.

and ozone production (Vukovich et al., 1977; Sillman and Samson, 1995). Chamber studies suggest that this is more important in cases of higher  $\text{NO}_x$  additions compared to VOC additions (Carter et al., 1997). We considered the relationship

$$P_o(\text{O}_3) = c(j \times [\text{HCHO}])^a [\text{NO}]^b 10^{d(T/298)}$$

In introducing a temperature effect, the expression  $10^{d(T/298)}$  was chosen as the numerically most tractable. Formulas like  $\exp(-T_{\text{Arrh}}/T)$  or  $T^d$  worked poorly numerically and gave no advantage. A resulting fit (graphed with  $T$  held at one value) is shown in Fig. 4(a) and (b).

Fig. 5(b) indicates how much correlation graphs tend to narrow at low values when temperature is included. There are linear relationships with different slopes from day to day. Since the upper end of the log plot splays wider than in panel Fig. 5(a), the poor fitting at high values of  $P_o(\text{O}_3)$  limits the  $R^2$  (non-log) to about 0.57.

When explanatory variables are somewhat correlated, emphasis in the regression may shift among variables. For example in Table 1, when temperature is added, the coefficient  $b$  for NO rises from 0.41 to 0.58, while  $a$  for HCHO activity drops from 0.36 to 0.17. Fig. 5(e) shows the clear but complex relationship between  $T$  and  $\log(j \times [\text{HCHO}])$ , even though the correlation evaluates to  $\sim 0.4$ . The correlation of  $j$  with  $j \times [\text{HCHO}]$  is about 0.5.

Possible explanations include these: (a) Days with different temperatures may have other variations in chemical composition. (b) Very fresh emissions contain higher aldehyde products, or fresh alkenes or similar compounds that have yet to be processed to produce the high quantities of HCHO that they would have in a few-hours steady state. (c)  $T$  is linked to local reactivity, and  $(j \times [\text{HCHO}])$  to species emitted earlier.

If NO is considered in isolation (affecting peroxyes only by direct reaction, and  $j \times [\text{HCHO}]$  and  $T$  are considered to be joint controls on  $\text{HO}_2$  radicals), we recover the symmetry and near 1:–1 behavior of Fig. 2. Fig. 6(a) shows the effect of each term in isolation. Note that the (sensitivity) slopes for  $j \times [\text{HCHO}]$  and  $T$ , give sensitivity like NO. The full interaction of  $j \times [\text{HCHO}]$  and  $T$  vs NO for is displayed in Fig. 6(b) non-parametric local Krige fit (again we used fields in  $R$ ) of the interaction of  $(j \times [\text{HCHO}]$  and  $T$ ) vs NO for  $P_o(\text{O}_3)$ . The 0.4 power law looks generally appropriate at higher concentrations, although there is lots of variation, presumably reflecting the fresh emissions. Rural North American analyses also gave an exponent of  $\sim 0.4$ .

### 3.7. Day-by-day non-predictive analyses of $P_o(\text{O}_3)$

The colors of points (i.e., sampling day) in Fig. 5(a)–(c) show a comparison of observed and predicted  $P_o(\text{O}_3)$ . The striking thing to note is that different near-linear relationships are observable for different days. This led us to a statistical analysis which takes us beyond an obvious predictive capability. Based on our observations of different slopes on different days, we allow a slow progression with time, with the variability limited by the analysis method of *mgcv*.

$$P_o(\text{O}_3) = c(j \times [\text{HCHO}])^a [\text{NO}]^b 10^{d(T/298)} f_{\text{spline}}(t), \text{ i.e.,}$$

The fitting parameters for the regression fit shown are shown in Table 1, yielding

$$P_o(\text{O}_3) = 10^{-0.54} (j \times [\text{HCHO}])^{0.28} [\text{NO}]^{0.60} f_{\text{spline}}(t)$$

$$P_o(\text{O}_3) = 10^{-13.1} (j \times [\text{HCHO}])^{0.17} [\text{NO}]^{0.64} 10^{12.82(T/298)} f_{\text{spline}}(t),$$

where  $t$  is measured in days and  $f_{\text{spline}}(t)$  is a spline function which minimizes residuals but is simple enough to satisfy a generalized cross-validation condition. An additive model from the  $R$  package

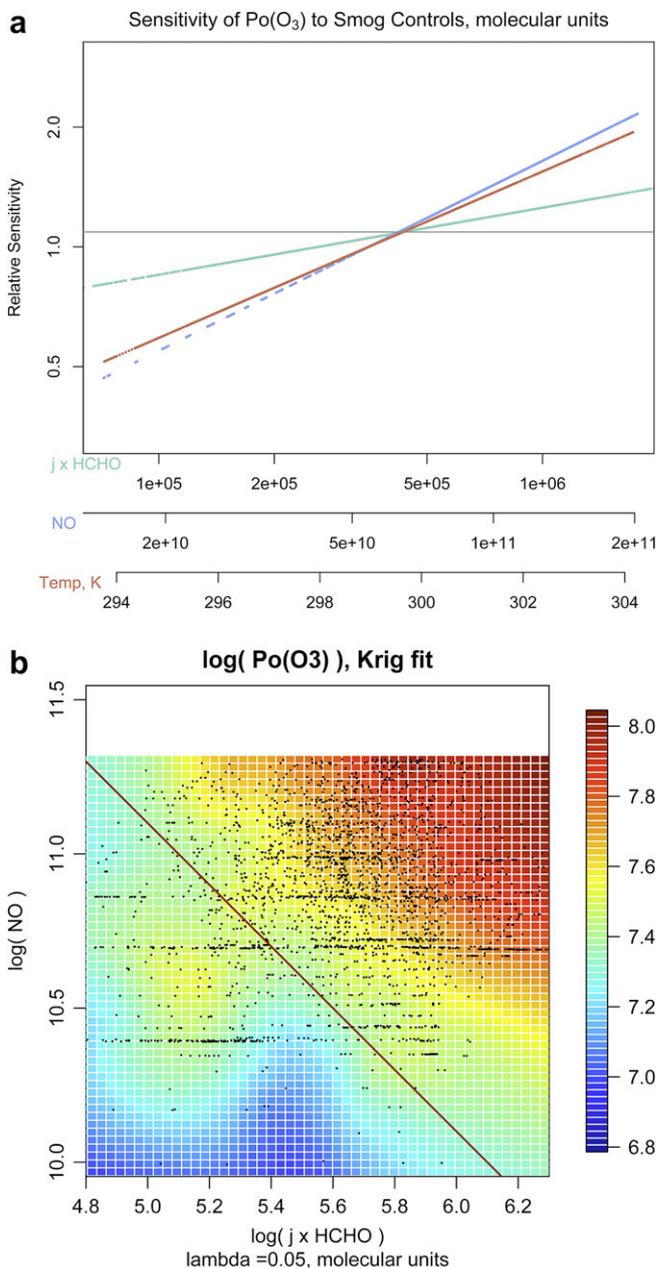


Fig. 6. (a) Response of  $P_o(\text{O}_3)$  to identified controlling variables; range of variables shown includes 98% of observations. (b) a completely empirical Kriging depiction of observed  $P_o(\text{O}_3)$  in terms of  $j \times [\text{HCHO}]$  and NO, log terms and using molecular units. Dots are data points.

*mgcv* was used (Wood, 2006). (The confidence intervals described in Table 1 were estimated using generalized cross-validation.) Fig. 5 (d) shows time, in a method that recognizes this slow variation. To what extent does this represent “over-fitting”? The function described has three maxima and *mgcv* describes  $\sim 9$  equivalent dof over 23 days. This is extremely small compared to the 4261 observations. We have examined the variation, and unfortunately, predictivity has been lost for this particular parameterization. We hope the analysis might stimulate future investigations of such uniformities in terms of chemistry.

Many attempts were made to explain the slow variation  $f_{\text{spline}}(t)$ . They are somewhat instructive in what hypotheses they discourage. It has no clear relation to wind direction or obvious categories of VOC, and the correlation with  $(T/298)$  is low. It was

natural to seek relationships with hydrocarbon reactivity and tendency to produce HO<sub>2</sub> or RO<sub>2</sub> radicals. We sought this with a partial least squares, PLS, (*L*<sup>1</sup> or “lasso” variety) regression fit (Hastie et al., 2001; Goeman, 2010). Groups of variables, components, were sought that were most strongly related to the function  $f_{\text{spline}}(t)$ . The procedure resembles principle components analysis, except that orthogonal groups are identified and ranked according to their contribution in explaining  $f_{\text{spline}}(t)$ . PLS analyses deal skillfully with correlated variables and allow generalized cross-validation (Hastie et al., 2001). Fig. 7(a) shows the pattern fitted using a dataset selected which could include VOC species. There was significant contribution of the first component, but little understanding: the compound identified as *trans*-2-butene was the major positive contributor (more O<sub>3</sub> production), while *cis*-2-butene (similar reactivity) was distinguishably negative; 3-methyl pentane was the strongest negative contributor. Few compounds contributed much correlation (factor loadings) to this group. We draw few conclusions: the VOC are likely acting as tracers of other conditions, e.g. wind. A regression against wind direction explained little. The PLS relationships are suggestive enough that we return to them for Fig. 7(b).

### 3.8. The OH + NO<sub>2</sub> loss term for oxidant

Earlier we noted that many ozone/oxidant destruction terms are relatively easily checked with measurements of stable species, but that the OH + NO<sub>2</sub> term can also require chemical modeling. Not only does it remove oxidant, it can also be extremely important in limiting the radical recycling of OH, HO<sub>2</sub>, and RO<sub>2</sub> which amplifies the effect of radical sources in the HO<sub>2</sub>-determining coordinate of our parameterization (Kleinman, 2004; Seinfeld and Pandis, 2004). Even more fundamentally extremely important loss term for NO<sub>x</sub> and consequently [NO] needed for P<sub>o</sub>(O<sub>3</sub>). Since it is often the major loss term for NO<sub>x</sub> and radicals, it has also been used to construct an “ozone production efficiency” OPE (Liu et al., 1987; Lin et al., 1988; Kleinman et al., 2002). This is the rate of oxidant produced by an NO

molecule divided by the NO<sub>x</sub> loss rate: it integrates to the oxidant produced from NO<sub>x</sub> emission to removal.

Consequently, we estimated regression involving NO<sub>2</sub> loss similar to P<sub>o</sub>(O<sub>3</sub>), assuming that OH reaction is the main loss for NO<sub>x</sub>.

$$L(\text{NO}_x) \sim R_{\text{OH}+\text{NO}_2} = k_{\text{OH}+\text{NO}_2}[\text{OH}][\text{NO}_2]$$

[NO<sub>2</sub>] rather than [NO] since it can be measured directly or estimated from [NO] and [O<sub>3</sub>] (Leighton, 1961). Using logs once again to allow linear regression estimates,

$$\log R_{\text{OH}+\text{NO}_2} = \log C + a \log(j \times [\text{HCHO}]) + b \log[\text{NO}_2] + d \log(T/298) + f(t)$$

We derived estimates (compared in Table 2 accompanied by approximate error estimates) such as the following

$$\begin{aligned} R_{\text{OH}+\text{NO}_2} &= 10^{0.42} (j \times [\text{HCHO}])^{0.17} [\text{NO}_2]^{0.90} \\ R_{\text{OH}+\text{NO}_2} &= 10^{1.51} (j \times [\text{HCHO}])^{0.11} [\text{NO}_2]^{0.70} 10^{0.048(T/298)} \\ R_{\text{OH}+\text{NO}_2} &= 10^{1.61} (j \times [\text{HCHO}])^{0.21} [\text{NO}_2]^{0.53} 10 f_{\text{spline}}(t) \\ R_{\text{OH}+\text{NO}_2} &= 10^{1.93} (j \times [\text{HCHO}])^{0.16} [\text{NO}_2]^{0.52} 10^{0.018(T/298)} \\ &\quad 10 f_{\text{spline}}(t) \end{aligned}$$

The estimates *a*, *b*, *C*, *d*, *f* follow generally similar patterns to those for P<sub>o</sub>(O<sub>3</sub>). In this urban situation, HCHO activity and NO<sub>2</sub> provide modest R<sup>2</sup> (R<sup>2</sup> = 0.41 ppb) for L(NO<sub>x</sub>) ~ R<sub>OH+NO<sub>2</sub>. Using *T* improves the relationship (0.56), but fitting a day-to-multiday function considerably raises the R<sup>2</sup> to 0.89. Fig. 8(a,b) gives an overview of sensitivities to radical production and NO<sub>2</sub> concentration in practical ppb units and also with a more abstract scale in which log quantities are used. The main contours indicate response to radical activity and to NO<sub>2</sub>, while the lighter contours indicate comparison to the ozone production relationships just shown. (NO was rescaled NO/NO<sub>2</sub> ~ 0.21). Considering *T* and *f* terms makes the NO<sub>2</sub> sensitivity somewhat lower.</sub>

Introducing a separate, slowly varying “gain” function *f* again (dramatically) increased R<sup>2</sup>. As Fig. 7(b) shows, this dimensionless *f* has a range of ~0.6 units, considerably greater than the ~0.25 range for P<sub>o</sub>(O<sub>3</sub>). In this case, the *L*<sup>1</sup> penalized PLS suggested that there were strong negative effects on L(NO<sub>2</sub>) that were associated with high reactivity VOC compounds. Possibly OH was being reduced by other reactants (e.g., VOCs) than NO<sub>2</sub>. Further studies, e.g., using reactivity weighted VOC’s, promise to define R<sub>OH+NO<sub>2</sub> better.</sub>

### 3.9. Approximate efficiency of ozone production

The culminating search for regression relationships that use common measurements is for OPE. Previous sections defined the components of P<sub>o</sub>(O<sub>3</sub>)/L(NO<sub>2</sub>). The values derived from the measurement OH and HO<sub>2</sub>, etc, suggest OPE between 0.5 and 18, with one outlier at ~45. Fig. 9(a) suggests that naive linear regression estimates of P<sub>o</sub>(O<sub>3</sub>)/L(NO<sub>2</sub>), without extra information are limited. The fit shown gives R<sup>2</sup> = 0.51, although explanatory

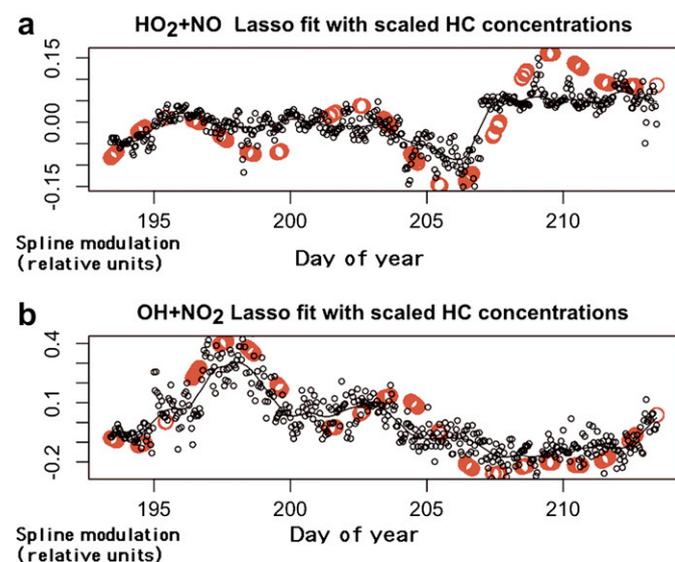


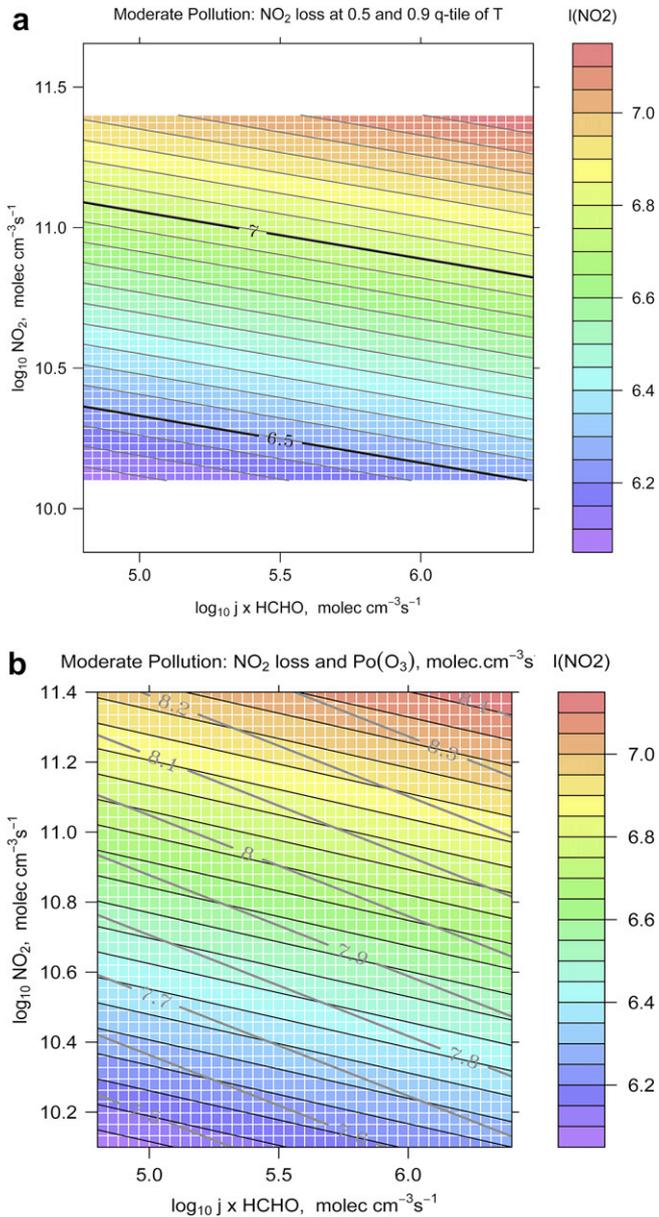
Fig. 7. (a) Red circles: Slow variation of *f*(*t*) for P<sub>o</sub>(O<sub>3</sub>) which greatly improves R<sup>2</sup>. Ordinates of the fit are nondimensional units (Wood, 2006). Smaller dark circles indicate variation of a best-fitting linear combination of hydrocarbon measurements chosen to mimic the trend of the red circles. (b) A similar depiction of a fit of the loss rate of NO<sub>2</sub> due to OH. Red circles are fitted smooth *t*-variation, dark circles are best-fit linear combination of hydrocarbon concentrations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

OH + NO<sub>2</sub> rate: moderately severe pollution precursor conditions.

Regression	( <i>a</i> , <i>b</i> , <i>c</i> )	( <i>a</i> , <i>b</i> , <i>c</i> , <i>d</i> )	( <i>a</i> , <i>b</i> , <i>c</i> , <i>f</i> ( <i>t</i> ))	( <i>a</i> , <i>b</i> , <i>c</i> , <i>d</i> , <i>f</i> ( <i>t</i> ))	
<i>c</i>	10 <sup><i>c</i></sup>	<b>0.42</b> ± 0.10	<b>-21.8</b> ± 0.47	<b>1.61</b> ± 0.06	<b>-4.5</b> ± 0.49
<i>a</i>	( <i>j</i> × [HCHO]) <sup><i>a</i></sup>	<b>0.17</b> ± 0.02	<b>0.11</b> ± 0.01	<b>0.21</b> ± 0.01	<b>0.16</b> ± 0.01
<i>b</i>	(NO <sub>2</sub> ) <sup><i>b</i></sup>	<b>0.90</b> ± 0.03	<b>0.69</b> ± 0.02	<b>0.53</b> ± 0.01	<b>0.52</b> ± 0.01
<i>d</i>	10 <sup><i>d</i>(<i>T</i>/298)</sup>		<b>20.37</b> ± 0.48		<b>4.78</b> ± .51
<i>f</i> ( <i>t</i> )	Day-dependent			<b>9.0</b> eq. dof	<b>9.0</b> eq. dof
	10 <sup><i>f</i>(<i>t</i>)</sup>				
	R <sup>2</sup> in log term	0.47	0.68	0.91	0.91
	R <sup>2</sup> in ppb terms	0.41	0.56	0.89	0.88

All regressions highly significant by *F* test.



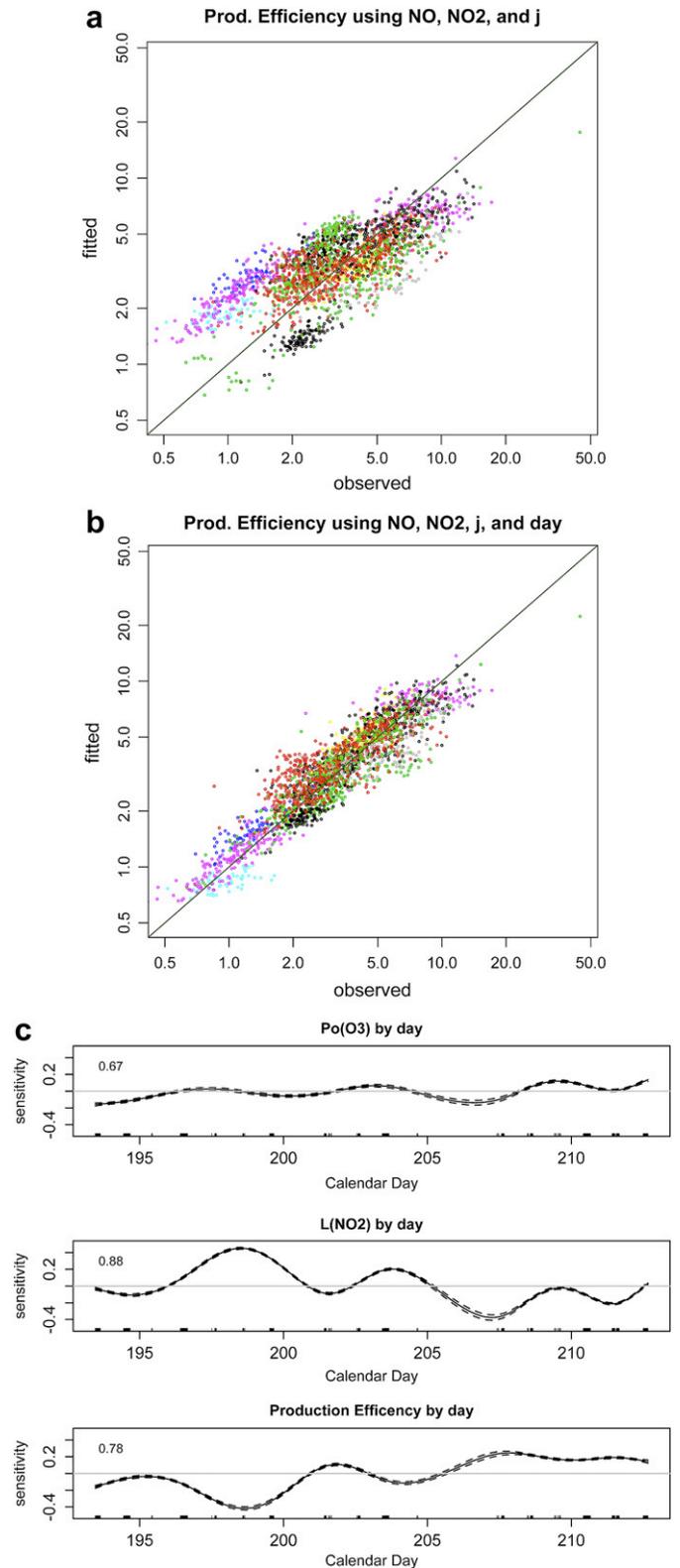
**Fig. 8.**  $L(\text{NO}_2)$  by OH + NO<sub>2</sub> reaction, color-filled contours. (a) Graph showing a large sensitivity to NO<sub>2</sub> and a moderate sensitivity to  $T$ . The parallel heavy lines show the dependence at 301.5 K, about five contour lines displaced from the color-graph scale, corresponding to  $T = 298$  K. (b) Graph does not have 1:1 aspect ratio. The gray contours indicate for comparison the sensitivity of the production rate of ozone; to do this, NO was estimated from NO<sub>2</sub> using a simple rescaling of the ratio of means.  $L(\text{NO}_2)$  is very NO<sub>x</sub>-sensitive.

power of  $f$  ( $R^2 = 0.78$ ) suggests better estimates can be discovered. One would expect that the NO/NO<sub>2</sub> ratio dominates OPE, with a smaller role for other variables. The radicals should track the NO<sub>x</sub> species in the central ratio  $[\text{HO}_2][\text{NO}]/[\text{OH}][\text{NO}_2]$  and both should rise with rising  $P(\text{HO}_2)$ . Indeed, NO and NO<sub>2</sub> dominate. The best estimate we found for this ratio included only  $j$ , not HCHO. Accounting for slow multiday variation with a simple spline “sensitivity” helped considerably.

$$\text{Ozone Production Efficiency} = 10^{1.830j} 1.830 \text{NO}^{0.81} \text{NO}_2^{-0.95},$$

$$R^2 = 0.51$$

$$\text{or } 10^{-8.2j-0.2} \text{NO}^{0.82} \text{NO}_2^{-0.94} 10^{9.6(T/298)} f_{\text{spline}}(t), \quad R^2 = 0.78$$



**Fig. 9.** (a) Calculated and fitted dimensionless  $P_o(\text{O}_3)$  as described by  $P_o(\text{O}_3)$ ,  $L(\text{NO}_2)$ . Day-to-day variation shown in different colors. (b) Similar comparison using also  $f(t)$ . (c) Slow, multiday variation of  $f(t)$  for  $P_o(\text{O}_3)$ ,  $L(\text{NO}_2)$ , and the corresponding OPE  $P_o(\text{O}_3)/L(\text{NO}_2)$ . Relative units are used (Wood, 2006). OPE variation mirrors the variation for  $L(\text{NO}_2)$ .

Notice in Fig. 8(c) that  $f_{\text{spline}}(t)$  used to improve this estimate is nearly the opposite of the function describing  $\text{NO}_2$  loss. The  $R^2$  values are encouraging, given strong compensating tendencies in the ratio. Section 3.7 suggests  $L(\text{NO}_2)$  (and OH) have strong  $f(t)$  dependences so is quite possible to learn more from hypothesized variables involving VOC composition or VOC reactivity.

### 3.10. Comparisons to previous work

There are many similarities of these relationships to those described by He. He describes relationships for ozone produced along a parcel trajectory with power laws somewhat similar to our own. Also links OPE to  $[\text{NO}]/[\text{NO}_2]$ . In these relationships, Kleinman makes a useful distinction between “original” sources of  $\text{HO}_x$  radicals and cycling of OH to  $\text{HO}_2$ . We have not distinguished these roles; Kleinman also notes some overlap in the concepts, especially concerning HCHO. HCHO is both an important source of  $\text{HO}_x$  radicals and also a significant measure of the rate of organic reactivity that converts OH to  $\text{HO}_2$ . Our research is also difficult to compare in detail to that of Sillman (Sillman and He, 2002, and referenced previous work) due to the near-source sampling. Sillman’s analysis moved forward the longitudinal studies of ozone, nitrogen oxides, and some pollution tracers in the work that followed the Liu et al. (1987) studies. Ren’s point-modeling work on the Queens situation is closer in spirit, and the Sillman and the Ren references are quite relevant to similar further studies. Still, the greatest utility of this work is to facilitate comparisons to regional model simulations, particularly for  $\text{HO}_2$  at high  $\text{NO}_x$  concentrations. This style of analysis may be greatly improved by VOC reactivity measurements wherever possible (Mao et al., 2009).

## 4. Conclusions

We propose and illustrate an analysis of local ozone production processes, interpreting measurements of the highly important photochemical terms,  $[\text{HO}_2][\text{NO}]$  and the  $[\text{OH}][\text{NO}_2]$ . We found that regressions involving easily measured non-radical species (HCHO) and physical conditions (UV, temperature) can be a surrogate for  $\text{HO}_2$  and so explain considerable variance for the  $\text{HO}_2 + \text{NO}$  rate, termed  $P_0(\text{O}_3)$ . The total rate  $P(\text{O}_3)$ , which includes  $\text{RO}_2 + \text{NO}$ , should closely parallel  $P_0(\text{O}_3)$ .  $\text{RO}_2$ , HCHO, and  $\text{HO}_2$  are closely related in sequential reactions in VOC oxidation, and so we propose POGO: Production of Ozone by Gauging (organic) Oxidation. POGO’s simplest description relates  $P_0(\text{O}_3)$  to a power-law expression,  $P_0(\text{O}_3) = C(j \times [\text{HCHO}])^a [\text{NO}]^b$ , found useful in rural atmospheres. In the Queens College, NYC, atmosphere, which was subject to local and urban-regional scale sources, this formulation worked well over  $\sim 100$ -fold variation in  $P_0(\text{O}_3)$ :  $R^2 \geq 0.7$ . Estimated exponents  $a$  and  $b$  were  $> \sim 0.4$  as in other rural studies. Broad testing was possible since  $\log(j \times [\text{HCHO}])$  and  $\log[\text{NO}]$  were surprisingly poorly correlated,  $R = 0.2$ .

The power-law form chosen suggested logarithmic sensitivities like  $\partial \log P_0(\text{O}_3) / \partial \log(j \times [\text{HCHO}])$  and also plots using logarithms of quantities. Our regressions suggest their usefulness over broad ranges of  $\text{O}_3$ -precursor concentrations. The sensitivity view suggests a focus on relative emissions reductions, provided that we can extend this local outlook to emission-to-receptor relationships.

A focus on a “moderate pollution” subsample revealed some complexity we had expected in the urban Queens dataset. Temperature seemed to play an important role in defining radical sensitivities, at least in several low-HCHO circumstances. An overall sensitivity  $a \sim 0.4$  was still observed.

Techniques similar to  $P_0(\text{O}_3)$  were applied to estimate the  $\text{OH} + \text{NO}_2$  reaction taken as defining the major loss  $L(\text{NO}_2)$  of  $\text{NO}_x$ . The log sensitivity to  $\text{NO}_2$  was greatest, but the definition of radicals

by  $j \times [\text{HCHO}]$  was also important.  $R^2$  was considerably raised by employing an empirical function  $f(t)$  changing slowly on timescales of 6–48 h. A similar  $f(t)$  was also useful for  $P_0(\text{O}_3)$ . When we combined the estimates to form an ozone production efficiency  $P_0(\text{O}_3)/L(\text{NO}_2)$ , these parameterizations helped considerably, raising  $R^2$  from 0.51 to 0.78. We speculate that these functions describe VOC composition or VOC reactivity, especially for  $L(\text{NO}_2)$ . There is some evidence for this from an  $L^1$  partial least-squares analysis.

For modelers, the main fruit of this work is likely a description of  $[\text{HO}_2]$  that underlies the discussion of  $P_0(\text{O}_3)$ . Models and other analyses are still required to link together our local descriptions of  $P_0(\text{O}_3)$  and  $L(\text{NO}_2)$  with the evolution of VOC’s and  $\text{NO}_x$ . This work may contribute to decomposition of the smog process and some checks on its simulation.

## Acknowledgements

We appreciate funding from Lawrence Friedl and NASA Applications Rapid Development program, the PMTACS team, anonymous reviewers, Laura Iraci, special editor Ian Barnes, and audience feedback at the Conference on Atmospheric Chemical Mechanisms, Davis, CA, USA, 2008.

## References

- Atkinson, R., Arey, J., 2003. Gas-phase tropospheric chemistry of biogenic volatile organic compounds: a review. *Atmospheric Environment* 37 (Suppl. 2), 197–219.
- Atkinson, R., Baulch, D.L., Cox, R.A., Crowley, J.N., Hampson, R.F., Hynes, R.G., Jenkin, M.E., Rossi, M., Troe, J., 2006. Evaluated kinetic and photochemical data for atmospheric chemistry: Volume II D. Gas phase reactions of organic species. *Atmospheric Chemistry and Physics* 6, 3625–4055.
- Brune, W.H., Faloona, I.C., Tan, D., Weinheimer, A.J., Campos, T., Ridley, B.A., Vay, S.A., Collins, J.E., Sachse, G.W., Jaeglé, L., Jacob, D.J., 1998. Airborne in situ OH and  $\text{HO}_2$  observations in the cloud-free troposphere and lower stratosphere during SUCCESS. *Geophysical Research Letters* 25, 1701–1704.
- Carter, W.P.L., Cocker III, D.R., Fitz, D.R., Malkina, I.L., Bumiller, K., Sauer, C.G., Pisano, J.T., Bufalino, C., Chen, S., 1997. Smog chamber studies of temperature effects in photochemical smog. *Environmental Science and Technology* 13, 1094–1100.
- Chatfield, R.B., Ren, Xinrong, Brune, William, Fried, Alan. Controls on regional ozone production: contrasts between California and the Eastern States. *Journal of Atmospheric Chemistry and Physics*, submitted for publication.
- Dasgupta, P.K., Dong, S., Hwang, H., Yang, H.-C., Gensa, Z., 1988. Continuous liquid phase fluorometry coupled to a diffusion scrubber for the determination of atmospheric formaldehyde, hydrogen peroxide, and sulfur dioxide. *Atmospheric Environment* 22, 949–964.
- Dodge, M.C., 1977. Combined use of modeling techniques and smog chamber data to derive ozone-precursor relationships. In: *Proceedings, International Conference on Photochemical Oxidant and Its Control*, Research Triangle Park, NC, USA.
- Eisele, F.L., Mount, G.H., Tanner, D., Jefferson, A., Shetter, R., Harder, J.W., Williams, E.J., 1997. Understanding the production and interconversion of the hydroxyl radical during the tropospheric OH photochemistry experiment. *Journal of Geophysical Research* 102, 6457–6465.
- Goeman, J., 2010.  $L^1$  and  $L^2$  Penalized Regression Models. <http://cran.r-project.org/web/packages/penalized/vignettes/penalized.pdf>.
- Hastie, T., Tibshirani, R., Friedman, J., 2001. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Verlag.
- Jacobson, M., 1998. *Fundamentals of Atmospheric Modeling*. University Press, Cambridge.
- Jenkin, M.E., Saunders, S.M., Pilling, M.J., 1997. The tropospheric degradation of volatile organic compounds: a protocol for mechanism development. *Atmospheric Environment* 31, 81–104.
- Kleinman, L.L., Daum, P.H., Lee, Y.-H., Nunnermacker, L.J., Springston, S.R., Weinstein-Lloyd, J., Rudolph, J., 2002. Ozone production efficiency in an urban area. *Journal of Geophysical Research* 107 (D23), 4733. doi:10.1029/2002JD002529.
- Leighton, P.A., 1961. *Photochemistry of Air Pollution*. Academic Press, San Diego.
- Lin, X., Trainer, M., Liu, S.C., 1988. On the nonlinearity of tropospheric ozone production. *Journal of Geophysical Research* 93, 15,879–15,888.
- Liu, S.C., Trainer, M., Fehsenfeld, F.C., Parrish, D.D., Williams, E.J., Fahey, D.W., Hubler, G., Murphy, P.C., 1987. Ozone production in the rural troposphere and the implications for regional and global ozone distributions. *Journal of Geophysical Research* 92, 4191–4207.
- Mao, J., Ren, X., Brune, W.H., Olson, J.R., Crawford, J.H., Fried, A., Huey, L.G., Cohen, R.C., Heikes, B., Singh, H.B., Blake, D.R., Sachse, G.W., Diskin, G.S.,

- Hall, S.R., Shetter, R.E., 2009. Airborne measurement of OH reactivity during INTEX-B. *Atmospheric Chemistry and Physics* 9, 163–173.
- Ren, X., Harder, H., Martinez, M., Leshner, R.L., OligierShirley, A.T., Adams, J., Simpás, J.B., Brune, W.H., 2003a. HO<sub>x</sub> concentrations and OH reactivity observations in New York City during PMTACS-NY2001. *Atmospheric Environment* 37, 3627–3637.
- Ren, X., Harder, H., Martinez, M., Leshner, R.L., OligierSimpas, J.B., Brune, W.H., Schwab, J.J., Demerjian, K.L., He, Y., Zhou, X., Gao, H., 2003b. OH and HO<sub>2</sub> chemistry in the urban atmosphere of New York City. *Atmospheric Environment* 37, 3638–3639.
- Ren, X., Olson, J.R., Crawford, J.H., Brune, W.H., Mao, J.Q., Long, R.B., Chen, Z., Chen, G., Avery, M.A., Sachse, G.W., Barrick, J.D., Diskin, G.S., Huey, L.G., Fried, A., Cohen, R.C., Heikes, B., Wennberg, P.O., Singh, H.B., Blake, D.R., Shetter, R.E., 2008. HO<sub>x</sub> chemistry during INTEX-A 2004: observation, model comparison and comparison with previous studies, 2008. *Journal of Geophysical Research* 113, D05310. doi:10.1029/2007JD009166.
- Seinfeld, J.H., Pandis, S.N., 1998. *Atmospheric Chemistry and Physics*. J. Wiley, Hoboken.
- Shafer, J.T., Seinfeld, J., 1985. Evaluation of Chemical Reaction Mechanisms for Photochemical Smog. Part 3. Sensitivity of EKMA (Empirical Kinetic Modeling Approach) to Chemical Mechanism and Input Parameters. U.S. Environmental Protection Agency, Washington, D.C. EPA/600/3-85/042 (NTIS PB85210888).
- Sillman, S., He, D., 2002. Some theoretical results concerning O<sub>3</sub>-NO<sub>x</sub>-VOC chemistry and NO<sub>x</sub>-VOC indicators. *Journal of Geophysical Research* 107. doi:10.1029/2001JD001123.
- Sillman, S., Samson, P.J., 1995. The impact of temperature on oxidant formation in urban, polluted rural and remote environments. *Journal of Geophysical Research* 100, 11497–11508.
- U.S. EPA, 1987. Guidelines of Use of the City-Specific EKMA in Preparing Post-1987 SIP's. Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina.
- Vukovich, F.M., Bach, W.D., Crissman, B.W., King, W.J., 1977. On the relationship between high ozone in the rural surface layer and high pressure systems. *Atmospheric Environment* 11, 967–983.
- Wood, S., 2006. *Generalized Additive Models: An Introduction with R*. Chapman and Hall.