PERGAMON

Atmospheric Environment 36 (2002) 4767–4775

ATMOSPHERIC ENVIRONMENT

www.elsevier.com/locate/atmosenv

Statistical analysis of primary and secondary atmospheric formaldehyde

Stephen Friedfeld[a], Matthew Fraser[a,*], Kathy Ensor[b], Seth Tribble[b], Dirk Rehle[c], Darrin Leleux[c], Frank Tittel[c]

[a] *Department of Environmental Science and Engineering, Rice University, Houston, TX 77005, USA*
[b] *Department of Statistics, Rice University, Houston, TX, USA*
[c] *Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA*

Received 10 February 2002; accepted 3 July 2002

Abstract

Regression models coupled with time series data were used to analyze the contribution of primary and secondary sources to formaldehyde (HCHO) concentrations, as determined by statistical analogy to primary (carbon monoxide, CO) and secondary (ozone, O_3) compounds measured simultaneously in Houston, TX. Time series analyses substantiated the need for statistical methods of analysis, given the complexity of the data and the rapid fluctuations that occur in atmospheric concentrations. A positive relationship was found for both the auto-correlation function (ACF) and partial auto-correlation function (PACF) of HCHO with either CO or O_3. Regression models used to distinguish primary and secondary contributions included a simple linear regression of the three compounds (one lag unit of time, 5 min) on current HCHO concentrations, resulting in a ratio of secondary formation to primary emission of 1.7. A second, more robust model utilized auto-correlated error processes to approximate the true nature of the linear regression; this model also indicates the ratio of secondary to primary contribution at 1.7 as the mean of ten model simulations. From the error processes model, one lag unit of time was most significant for CO predicting HCHO, while simultaneous measurements (lag 0) were most significant for O_3 predicting HCHO. Outlying O_3 and HCHO concentrations were shown not to affect the results.

Keywords: Formaldehyde; Statistical modeling; Secondary formation; Houston; TX

1. Introduction

Volatile organic compounds (VOCs) and oxides of nitrogen (NO_x) react in the presence of sunlight to yield ozone (O_3), one of six criteria pollutants regulated under the Clean Air Act. Elevated concentrations of tropospheric O_3, the primary component in photochemical smog, are commonly measured in urban centers and across regional airsheds, impacting human health and causing damage to crops, forests, and materials (Moussiopoulos, 1990; Berntsen et al., 1997; Seinfeld and Pandis, 1998). Continuous exposure to substantial levels of O_3 irritates the eyes and lungs at low persistent levels, and causes respiratory illness at higher concentrations (Johnson et al., 1981; Pryor, 1998). For those regions in violation of the federal O_3 standard, oxidant control strategies for attainment purposes require an understanding of O_3 formation precursors, including VOCs. In particular, it is important to differentiate between primary (directly emitted from sources) and secondary (formed by atmospheric chemical reactions) VOCs to direct policy for the control of these compounds.

One VOC of interest is formaldehyde (HCHO), a highly reactive compound found in urban and rural atmospheres; primary HCHO sources include vehicular

*Corresponding author. Tel.: +1-713-348-5883; fax 1-713-348-5203.
E-mail address: mpf@rice.edu (M. Fraser).

1352-2310/02/$ - see front matter
PII: S1352-2310(02)00558-7

exhaust and fugitive industrial emissions (Altshuller, 1993). Secondary formation of HCHO from the breakdown of primary compounds, also contributes to ambient formaldehyde concentrations. Primary sources and secondary formation must both be included in calculating VOC loadings in the troposphere for purposes of predicting concentrations of O_3 or other species. However, differentiating between primary and secondary sources of HCHO is important for regulatory and control purposes.

In this study, primary and secondary contributions to HCHO concentrations are estimated through statistical methods of analysis of simultaneous real-time measurements of HCHO, carbon monoxide (CO), and O_3. Significant previous research has focused on HCHO, as the photolysis of HCHO during daylight hours produces atmospheric radicals that drive photochemical smog formation (National Research Council, 1991). Past research (Altshuller, 1993; National Research Council, 1991; Possanzini et al., 1996; Kawamura et al., 2000) has investigated the production of HCHO from both primary and secondary sources by comparing estimates of HCHO from vehicular emissions with estimates of HCHO from alkene and alkane atmospheric reactions. The advantage of the current study is that by using simultaneous real-time measurements and rigorous statistical analysis in determining these relative contributions, the statistical accuracy of the estimations can be quantified.

Formaldehyde concentrations are governed by both primary emissions and secondary formation. The simplest, most reactive, and most abundant atmospheric carbonyl, HCHO levels typically range from about 10–20 ppbv in most urban settings (Carlier et al., 1986). Primary sources of HCHO include vehicle exhaust (Grosjean et al., 1993) and stationary sources (Grosjean and Swanson, 1983). Radical-driven atmospheric chemistry of a wide variety of VOCs is responsible for secondary HCHO in urban atmospheres, and isoprene interactions account for a large percentage of secondary HCHO from hydrocarbons emitted from biogenic sources (Altshuller, 1993; Harder et al., 1997). During a recent sampling campaign (Rehle et al., 2001), HCHO concentrations in Houston reached almost 45 ppbv in a heavily industrialized section of the city.

Since chemical quantification cannot differentiate between primary emissions and secondary formation of atmospheric formaldehyde, an alternative approach must be taken. In this work, statistical time-series analysis is used to separate the concentration of primary formaldehyde from secondary formaldehyde using the information contained in the fluctuating concentration of other pollutant gases (Box et al., 1994). Carbon monoxide is a stable, colorless and odorless gas with an atmospheric lifetime of 2–4 months. Anthropogenic primary emissions dominate the sources of atmospheric CO. In Houston, mobile sources contributed approximately 90% of the CO, and stationary point sources contributed 8% (Houston-Galveston Area Council, 1999). For the purposes of this study, statistical analogy of HCHO to CO measurements represents the primary contribution to HCHO formation. In contrast to CO which is dominated by primary emissions, tropospheric ozone exists solely as a secondary compound, formed and destroyed in the atmosphere via a cyclical series of reactions involving VOCs, NO_x and sunlight. For this study, statistical analogy of HCHO to O_3 is used to estimate the secondary formation of HCHO.

The objectives of the study were to determine the relative contribution of primary emissions and secondary formation to the atmospheric HCHO levels based on the statistical relationship between ambient concentrations of HCHO, CO and O_3. The temporal relationship between these pollutants was sought to determine the relative staging of pollutant concentrations. While this work was conducted in Houston area, the results can generally be extrapolated given certain limitations, including (1) likely contributors to both primary and secondary HCHO formation will only be generally approximated by CO and O_3 concentrations; (2) source contributions vary in importance, which cannot be elucidated when using only one primary and only one secondary compound in the analogy (Duarte-Davidson et al., 1997); (3) atmospheric lifetimes (τ) of aldehydes are on the order of hours, where loss occurs mostly via reaction with OH radicals and photolysis (Tanner et al., 1988); and (4) the diurnal variation in solar intensity will lead to changing ratios of primary emissions to secondary formation for HCHO during the course of the day, while statistical time series analysis can only estimate the average ratio of primary to secondary formaldehyde (Box et al., 1994)

2. Methods

2.1. Sampling site

Over 30 Continuous Air Monitoring Stations (CAMS) are operated by the Texas Natural Resource Conservation Commission (TNRCC), the City of Houston, and the Houston Regional Monitoring Corporation throughout Houston and the surrounding counties (see http://www.tnrcc.state.tx.us/cgi-bin/monops/select_summary). These sites collect continuous measurements of gaseous pollutants such as ozone, carbon monoxide, oxides of nitrogen, and meteorological variables. The compound data in the present study were collected between 14 June and 21 June 2000 at Deer Park CAMS 35 (operated by the TNRCC) located south of the heavily industrial Houston Ship Channel, and northwest of Galveston Bay and the Gulf of Mexico.

During the 1-week sampling period, temperature and relative humidity ranged from 23°C to 33°C and 48% to 98% RH. The period included several hours of precipitation, from midday 15 June to mid-morning 16 June 2000 during which HCHO data were not recorded. All instrumentation was kept inside the CAMS trailer, with tubing leading to the rooftop for ambient air collection. Sampling from the 4 m rooftop allowed free airflow in all directions. The collection instruments were maintained at 26°C in an air-conditioned monitoring station.

2.2. Data acquisition

Ozone measurements were made by using a Dasibi 1008AH instrument with precision and detection limits of 1 ppbv. Carbon monoxide was measured using a Teco Model 48 instrument with precision and detection limits of 10 ppbv. Though TNRCC-operated instruments at Deer Park record O_3 and CO concentrations every 5 min, the data are reported publicly as hourly averages, reflecting federal standards such as NAAQS. To understand more clearly the regional atmospheric chemistry, however, requires accurate measurements at shorter intervals to account adequately for rapid fluctuations in atmospheric concentrations (Fried et al., 1998). Traditionally, HCHO concentrations were measured hourly or daily (Salas and Singh, 1986; Possanzini et al., 1996; Gilpin et al., 1997) due to a lack of technology for accurate real-time collection and analysis, in addition to relative simplicity with time-integrated methods. The present study employed a relatively new spectroscopic technique based on difference frequency generation (DFG), designed to measure HCHO on a real-time basis with an estimated precision of ~ 0.5 ppbv (Lancaster et al., 2000). In a recent study, the DFG sensor was shown to be a robust and accurate method for HCHO quantification (Friedfeld et al., 2000). Details of the DFG sensor are reported in Rehle et al., 2001.

Data were collected at 5-min intervals from 14 June 2000 (1400 CDT) to 21 June 2000 (1100 CDT) with several periods of data missing for instrument recalibration and maintenance. Examination of the data demonstrated a few observations missing at random until an 11-h rainstorm (15 and 16 June 2000) and disruption of DFG collection, followed by a reasonably contiguous period for the remainder of the week with some further random missing data. Due to a large data pool following the 11 h disruption of HCHO collection, we focused analysis on this latter period; overall, nearly 5900 measurements were collected simultaneously for the three compounds of concern. To treat missing data from either instrument recalibration or maintenance of the DFG monitor, a univariate auto-regression (AR) model was fit to each series separately, where the sample mean value was assigned to all missing data points. The missing observations were imputed by simulating random values random values based on the sampling distribution of the estimator, at that point in time. The new imputed observations do not incorporate cross-correlative structure, as only univariate models were considered; thus, resulting estimates of possible interactions between pollutants are conservative.

2.3. Diagnostic statistics

All statistical analysis was performed with the software package S-Plus© (S-Plus, 2000). Time series analyses are generally used to understand the stochastic mechanism that gives rise to an observed series or to predict future values based on series history (Box et al., 1994). The time series data collected at Deer Park allow for statistical investigation of primary and secondary inputs to atmospheric chemical reactions. Modeling the time series using the chosen methodologies, however, requires a transformation of each series to correct for the skewness of the data. Quantile plots were thus developed to assess the symmetry and tail behavior of various data transformations of the series, including comparing the original data with natural log, square root, and negative inverse transformations of the three series.

To remove the auto-correlation *within* each series so that we can identify lagged correlations between series, a fitted auto-regressive equation was determined for each series, and the residuals (error terms) from each equation obtained. Only cross-correlation effects *between* series therefore remained. Both the auto-correlation function (ACF) and the partial auto-correlation function (PACF) were evaluated using the residuals of the model fit to the square-root transformed data to attribute both the cumulative and individual influence of primary and secondary compounds on HCHO (through analogy with CO and O_3 measurements). Recall, only data following the rainstorm on 15 and 16 June were included in the analyses; any few remaining missing values were accounted for as previously described.

The ACF represents the sample correlation between sets of ordered data pairs; the sample correlation was computed for HCHO at time t ($HCHO_t$) and CO at time $t-1$, $t-2$, …, $t-25$, as well as O_3 at the same lagged values. Recall in this study the time unit is 5 min, thus CO_{t-25} and $O_{3,t-25}$ indicate concentrations measured approximately 2 h prior to $HCHO_t$. By definition, the ACF indicates the cumulative effect of all lag periods on the present concentration. The PACF, on the other hand, measures the correlation between the present concentration and the concentration at one lag only.

2.4. Statistical modeling

The complexity of the data—~5900 simultaneous measurements—requires rigorous statistical analyses to discern a lag–lead relationship and primary vs. secondary HCHO precursor contributions. Two linear regression models were applied to the HCHO, CO, and O_3 series. The first model uses a basic linear regression of HCHO at time t, with explanatory variables CO_{t-1}, $O_{3,t-1}$, and $HCHO_{t-1}$, that is, at 1 lag unit of time as per the equation

$$HCHO_t = \beta_0 + \beta_1 CO_{t-1} + \beta_2 O_{3,t-1} + \beta_3 HCHO_{t-1} + \varepsilon_t, \quad (1)$$

where the coefficients β_1 and β_2 approximate the true nature of the contribution of primary and secondary lag concentrations on HCHO as measured by a linear relationship and ε_t is the error term that is assumed to be independent with constant variance and mean zero. This simple linear regression model provides a preliminary assessment of the relationship between HCHO and the primary and secondary variables (CO and O_3).

Improving on the previous model, we consider a regression model relating lagged values of CO and O_3 to HCHO but also accounting for the serial correlation in the error process. More specifically, the second model considered is

$$HCHO_t = \beta_0 + \beta_1 CO_{t-j} + \beta_2 O_{3,t-k} + z_t, \quad (2)$$

where j and k equal lag $(0, 1, 2, \ldots)$ and j may or may not equal k. The error process z_t is assumed to follow an auto-regressive process of order p; the order of the process is estimated from the observed series. In this analysis, the constant term, β_0, is estimated as the sample mean of the HCHO series. The Cochran–Orcutt method (25) was used iteratively to obtain estimates of the remaining regression coefficients and the parameters of the auto-regressive error process.

Using both a partial auto-correlation (which is dependant on measurements of CO, O_3 and HCHO made in the previous time period) and an auto-correlation function (which is dependant on measurements of CO and O_3 made over the previous 25 time periods) evaluates the effects of chemical reactions on model predictions. In the partial auto-correlation function, relating the present HCHO concentration to the pollutant concentration measured in the previous time period minimizes biases in the model from chemical reaction of HCHO as well as O_3 and CO. On the other hand, the auto-correlation function calculates the present HCHO concentration as a function of CO and O_3 measured over the previous 25 time periods, which could be subject to biases from chemical reaction of the pollutant gases. Using both calculations allows empirical evaluation of whether chemical reaction of pollutant gases affects the model predictions.

3. Results

3.1. Time series analysis

Separate time series plots were created for all three compounds of interest (Fig. 1a). The x-axis is marked by 1200 CDT on each sampling day, while the y-axis represents atmospheric concentration in ppbv (HCHO and O_3) or ppmv (CO); each graph is scaled individually to emphasize diurnal patterns. For CO, the overall average, minimum, and maximum concentrations are 0.37, 0.28, and 0.62 ppmv. The CO concentrations follow a relatively cyclical pattern, with an average daily maximum of 0.51 ppmv occurring at around 0600 CDT during the weekdays (and at 2000 CDT on the weekend, 17 and 18 June); the mean daily minimum was 0.30 ppmv, with varied times throughout the week.

The O_3 concentrations fluctuate in a less distinct diurnal cycle highlighted by extreme values, with a weekly minimum of 0.4 ppbv and maximum of 77.3 ppbv. The average O_3 concentration was 18.4 ppbv, and the average daily maximum was 38.0 ppbv, and the highest levels occurred between 1000 and 1300 CDT almost daily. From the figure, there is a noticeable spike on 18 June at 1300 CDT, where the concentration reached 77 ppbv. The missing data for HCHO during the rainstorm on 15–16 June are shown as a straight line simply for graphical purposes and do not reflect interpolated values. Formaldehyde concentrations ranged from 0.2 to 19.6 ppbv, with a weeklong average of 4.9 ppbv. Typical diurnal fluctuations remained limited between 1.2 and 9.5 ppbv, the daily minimum and maximum averages. As with O_3, a concentration spike up to ~20 ppbv occurred in early afternoon on June 18. The smaller and more rapid fluctuations of HCHO highlight the importance of real-time measurements in studying regional atmospheric chemistry.

Fig. 1b contains an overlay plot of HCHO and O_3, as these compound concentrations follow a more irregular pattern than CO levels. The regular CO pattern presumably arises from the primary nature of CO sources, devoid of external interferences such as wind patterns, meteorology, temperature, and other compound concentrations and reaction conditions. The CO cycle changes during the weekend, reflecting different industrial production and vehicular usage. Secondary compounds are all subject to similar external influences; the 18 June spike for both secondary compounds indicates the presence of such an influence (Fig. 1b). While similar cyclical patterns exist throughout the week, the small range of HCHO concentration fluctuations requires statistical considerations for further analysis.

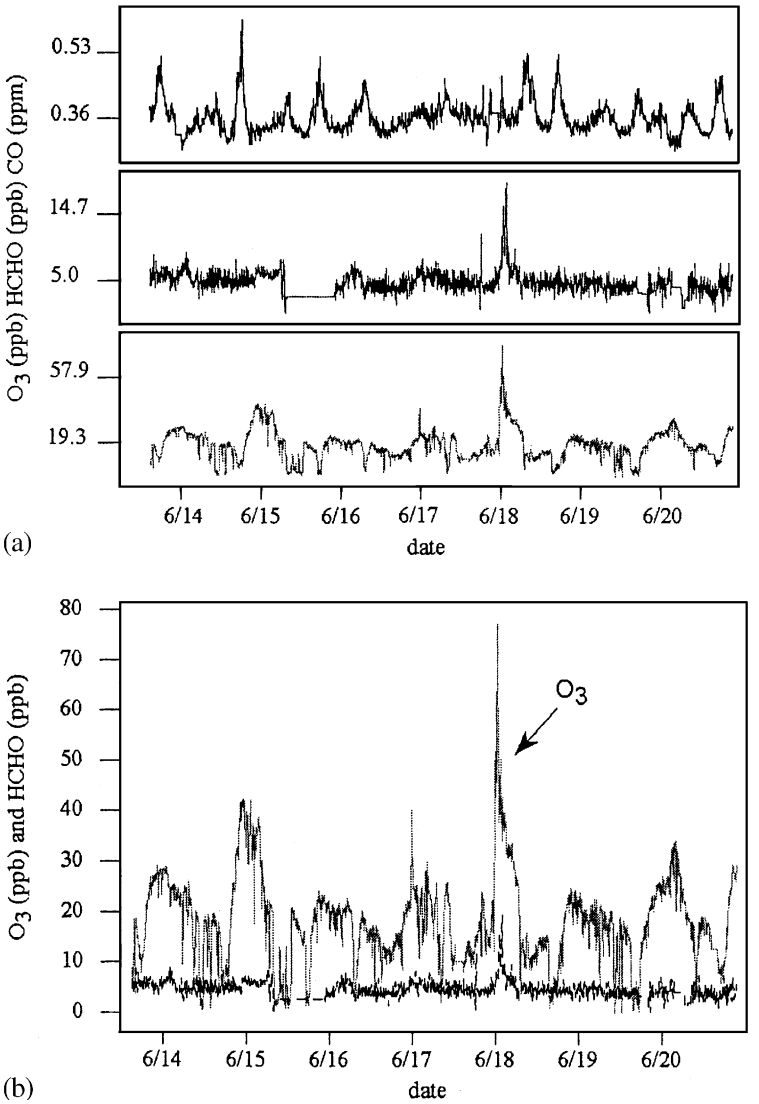

Fig. 1. (a) Time series plots of CO, HCHO, and O_3 measured at Deer Park. The x-axis is time, marked at 1200 CDT. The y-axes are the concentration in ppmv for CO and ppbv for HCHO and O_3. (b) Overlay of O_3 and HCHO time series plots.

3.2. Quantile plots

Fig. 2 contains quantile plots of the non-transformed data along with the natural log, square root, and negative inverse transformations. Each row represents a time series (HCHO, O_3, and CO, respectively) and each column represents a transformation (original data (x), $\ln(x)$, $\sqrt{x}$, and x^{-1}, respectively). The x-axis indicates the percentile of a value in the series centered about 0 and the y-axis represents each value. Visually, the most well-behaved data occur with the square-root function. While all transformations appeared generally symmetric with the CO data, the log and negative inverse functions lose symmetry about the median for HCHO and O_3 values. The square-root transformation is not perfectly normal for all three data sets, but appears the most symmetric transformation nonetheless.

3.3. ACF and PACF

The ACF and PACF were plotted against time (in lag units of 5 min) for CO and O_3 correlated with HCHO. The ACF as a function of lagged time for both CO and O_3 correlated with HCHO indicates an overall strongly positive relationship. Similarly overall positive partial auto-correlations are observed as well. Although the

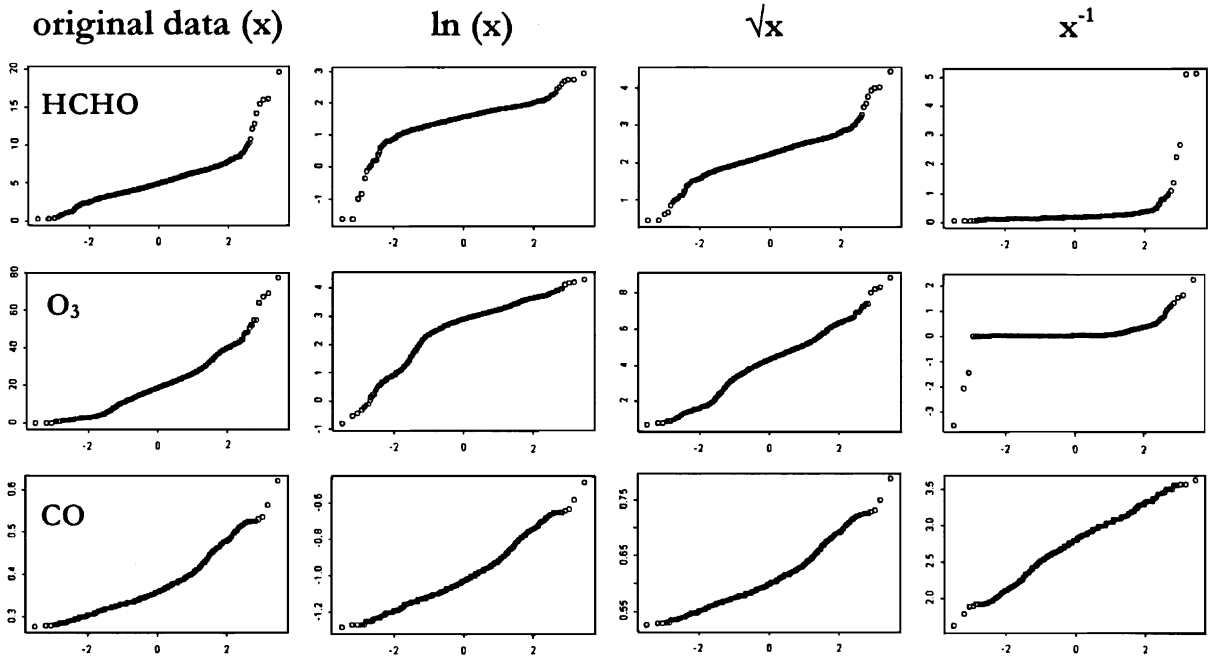

Fig. 2. Quantile plots of original and transformed data series for HCHO (top row), O_3 (center), and CO (bottom). The transformations include: original data, natural log, square root, and negative inverse. The x-axis indicates the percentile of a value in the series centered about the median (0); the y-axis represents the transformed value. The square-root transformation is considered to be the most symmetric.

ACF and PACF values are small, they are highly significant in general.

3.4. Statistical modeling results

Table 1 contains a summary of the results and the statistical significance of the fitted simple linear regression for Eq. (1). The p-value (denoted by $Pr(>|t|)$) indicates that there is approximately zero probability (truncating at four significant digits) that with any of the estimates of the coefficients, the true nature of the relationship between the gases for the equation considered is 0. Each coefficient can therefore be considered not only statistically significant, but in fact a good estimate of the true (theoretical) value. The p-value of the F-statistic is 0, further demonstrating that there exists a non-zero relationship between the response and explanatory variables. Considering *solely* primary and secondary contributions in this model, 64% of $HCHO_t$ concentrations is correlated with $O_{3,t-1}$ (coefficient β_2, 0.22) and 36% with CO_{t-1} (coefficient β_1, 0.13), or a ratio of 1.7. It is prudent for us to note some obvious limitations of this modeling strategy, given that CO and O_3 concentrations would not fully represent accurately and completely the primary and secondary compounds. Many other reactive compounds exist and contribute to the complexity of the atmospheric composition and chemistry.

Table 1
Regression coefficients and related statistics from a simple linear regression model for Deer Park data, June 2000. The term $Pr(>|t|)$ indicates that the coefficients on the regressors are acceptable at the 95% confidence interval. The ratio of coefficients (used to estimate the ratio of secondary to primary formaldehyde) is 1.7

Compound	Coefficient	Standard error	t-value	$Pr(>\|t\|)$
CO	0.13	0.025	5.15	0.00
O_3	0.22	0.026	8.60	0.00
HCHO	0.33	0.025	13.1	0.00

Table 2 contains summary results for the fitted model of Eq. (2). To determine the lag periods for which CO and O_3 most significantly lead to the formation of HCHO, the contributions to a multiple regression model for varying lags in the presence of one another were compared. The most significant lags in the regression, when compared with other lags, include lag 1 (5 min) for CO and lag 0—that is, simultaneous concentrations—for O_3. These lag periods (0 and 1) emphasize that hourly or daily time-integrated measurements do not

accurately represent the true nature of the atmospheric chemistry of a region; a significant simultaneous secondary concentration (HCHO and O_3) highlights the need for historical data with statistical analyses to predict peak O_3 concentrations.

As described above, missing values were replaced with random imputed values; thus, the iterative process was simulated ten times each for CO_{t-1} and $O_{3,t}$, the most significant lag periods. The mean coefficient for both compounds is listed in Table 2; the mean p-value suggests that the results are significant for CO at 94.5% confidence and O3 at 99.9% confidence. The auto-correlated error-processes model produced an identical result to the basic linear regression model, with a ratio of $\beta_2(O_3)$ to $\beta_1(CO)$ of 1.7.

As described in the time series analysis above, a noticeable spike in concentration occurred for both HCHO and O_3; the remaining data appeared to follow repetitive diurnal cycles. In Fig. 3a plot of the residuals vs. each of the predictors was created to assess the effect of outliers, such as that on 18 June, on the fitted auto-correlated error processes model. Although there are a few moderate outliers, they do not overly influence the fitted models.

Table 2
Mean values from ten simulations of the auto-correlated error processes model. The $Pr(>|t|)$ terms indicate that the CO coefficient is acceptable at the 95% confidence interval, and the O_3 coefficient is acceptable at the 99% confidence interval. The ratio of coefficients (contribution of secondary to primary) is 1.7

Compound	Coefficient	$Pr(>\|t\|)$
CO	0.09	0.055
O_3	0.15	0.001

3.5. Implications

Both modeling strategies attribute nearly two-thirds (ratio 1.7) of HCHO formation to secondary VOC reactions, and only one-third to primary emissions. Earlier studies have been conducted similarly investigating the contribution of primary and secondary precursors to HCHO formation. As previously stated, these studies often relied on estimates of vehicular exhaust and assumptions of reactant concentrations and reaction rate constants.

Possanzini et al. (1996) measured HCHO and CO at 1-h time-averaged intervals and obtained vehicular emissions from a previous source. Using average O_3 and OH reaction concentrations for estimating secondary reactions with alkenes, the group found between 44 and 49% of the HCHO formation could be attributed to secondary formation. Others used similar procedures, but obtained results closer to those in this study, where secondary emission precursors predominate HCHO formation. Altshuller (1993), for example, investigated aldehyde formation from primary and secondary sources during night and early morning hours; the investigation entailed an emissions inventory estimate

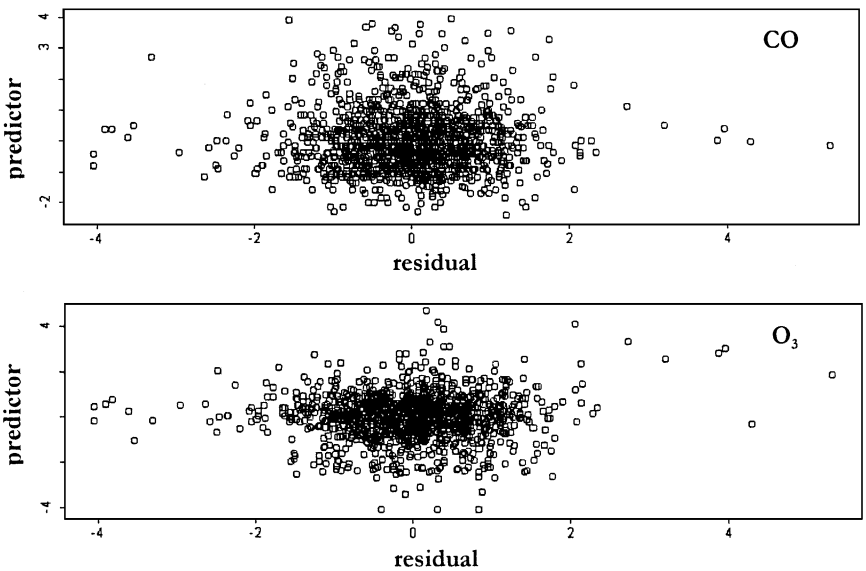

Fig. 3. Plots of the residuals vs. predictors for both CO and O_3 indicate that outliers, such as the spike in concentration of O_3 and HCHO on 18 June, do not affect the results of the auto-correlated error processes model, and thus the model is considered valid in assessing the primary and secondary contributions to HCHO concentrations.

from 1980s vehicles and typical diurnal concentrations and rate constants for alkenes, NO_3, and O_3. The mean result for that investigation led to ~78% HCHO attributable to secondary reactions and 22% to primary emissions for 0600–0900 h. A high ratio of secondary/primary compounds was also calculated by Kawamura et al. (2000). Kawamura et al. (2000) used historic 1980s Los Angeles data for both vehicular exhaust and compound concentrations and concluded that secondary photochemical reaction contributed approximately 87% to aldehyde production during daylight hours.

The present study incorporates simultaneous measurements of primary and secondary compounds at relatively short intervals over a statistically significant time period and therefore provides the most accurate assessment of HCHO formation. These studies are all complicated by the fact that measured HCHO concentrations need to account for both formation and destruction, that is, net production. Altshuller (1993) emphasizes only HCHO production, but states also that HCHO reaction with O_3 is extremely slow; thus, the short (5 min) time intervals do not need to account for that reaction. Formaldehyde reaction with OH does serve as an important mechanism of HCHO destruction; however, this reaction is considered implicitly in the regression models, since any estimate of the destruction (e.g., rate constant $k_{OH \cdot HCHO}$ and OH concentration) would be subtracted equally from each HCHO concentration in the data series prior to statistical modeling, and subsequently not affect the resultant coefficients β_1 and β_2. Kawamura et al. (2000) simply accounted for net production by assuming a conservative change of 2/3 in concentration.

The assumptions used in this study are important in interpreting the results. Many sources emit primary HCHO, and many reactions and precursors contribute to secondary HCHO production. Thus, the analogy using only CO and O_3 concentrations serves only as an approximation; nonetheless, the results of this study can be extrapolated to other urban areas with similar O_3 problems more readily than can the results from previous studies, as both real-time and simultaneous measurements greatly decrease the gross assumptions made by others. Because the models showed that nearly two-thirds of HCHO production is attributable to secondary reactions in the atmosphere, legislative efforts to reduce primary emissions of higher aldehydes and ketones—those compounds that break down in the atmosphere to form HCHO—will more effectively contain HCHO production, and ultimately O_3 production as well. A future shift in gasoline composition to methanol or natural gas could result in higher primary emissions of HCHO and thus create a likewise shift in the secondary/primary contribution of HCHO. Obviously, decreasing primary emissions of HCHO along with a simultaneous decrease in emissions of higher carbonyl compounds as part of an implementation plan would most effectively reduce O_3 concentrations.

Acknowledgements

This work was funded by the Gulf Coast Hazardous Substance Research Center, the Texas Advanced Technology Program, NASA, the Welch Foundation and the National Science Foundation. The authors would like to acknowledge Jairam Nadkarny of the TNRCC for his assistance with this project. The authors are especially grateful to the statistical analyses and results from the Rice University Statistical Consulting Laboratory.

References

Altshuller, A., 1993. Review: natural volatile organic substances and their effect on air quality in the United States. Atmospheric Environment 27A, 2131–2165.

Berntsen, T., Isaksen, I., Myhre, G., Fuglestvedt, J., Stordal, F., Larsen, T., Freckleton, R., Shine, K., 1997. Effects of anthropogenic emissions on tropospheric ozone and its radiative forcing. Journal of Geophysical Research 102, 28101–28106.

Box, G., Jenkins, G., Reinsel, G., 1994. Time Series Analysis: Forecasting and Control, 3rd Edition. Prentice-Hall, Englewood Cliffs, NJ.

Carlier, P., Hannachi, H., Mouvier, G., 1986. The chemistry of carbonyl compounds in the atmosphere—a review. Atmospheric Environment 20, 2079–2099.

Duarte-Davidson, R., Sewart, A., Alcock, R., Cousins, I., Jones, K., 1997. Exploring the balance between sources, deposition, and the environmental burden of PCDD/F in the UK terrestrial environment: an aide to identifying uncertainties and research needs. Environmental Science and Technology 31, 1–11.

Fried, A., Henry, B., Wert, B., Sewell, S., Drummond, J., 1998. Laboratory, ground-based and airborne tunable diode laser systems: performance characteristics and applications in atmospheric sciences. Applied Physics B 67, 317–330.

Friedfeld, S., Fraser, M., Lancaster, D., Leleux, D., Rehle, D., Tittel, F., 2000. Field intercomparison of a novel optical sensor for formaldehyde quantification. Geophysical Research Letters 27, 2093–2096.

Gilpin, T., Apel, E., Fried, A., Wert, B., Calvert, J., Genfa, Z., Dasgupta, P., Harder, J., Heikes, B., Hopkins, B., Westberg, H., Kleindienst, T., Lee, Y.-N., Zhou, X., Lonneman, W., Sewell, S., 1997. Intercomparison of six ambient [CH2O] measurement techniques. Journal of Geophysical Research 102, 21161–21188.

Grosjean, D., Swanson, R., 1983. Carbonyls in Los Angeles air: contribution of direct emissions and photochemistry. Science of the Total Environment 29, 65–85.

Grosjean, E., Williams II, E., Grosjean, D., 1993. Ambient levels of formaldehyde and acetaldehyde in Atlanta, Georgia. Journal of the Air Waste Management Association 43, 469–474.

Harder, W., Fried, A., Sewell, S., Henry, B., 1997. Comparison of tunable diode laser and long-path ultraviolet/visible spectroscopic measurements of ambient formaldehyde. Journal of Geophysical Research 102, 6267–6282.

Houston-Galveston Area Council, 1999. Air Quality Reference Guide for the Houston-Galveston Area, Houston, TX.

Johnson, L., Joseffson, B., Marstorp, P., 1981. Determination of carbonyl compounds in automobile exhausts and atmospheric samples. International Journal of Environmental Analytical Chemistry 9, 7–26.

Kawamura, K., Steinberg, S., Kaplan, I., 2000. Homologous series of C1-C10 monocarboxylic acids and C1–C6 carbonyls in Los Angeles air and motor vehicle exhausts. Atmospheric Environment 34, 4175–4191.

Lancaster, D., Fried, A., Wert, B., Henry, B., Tittel, F., 2000. Difference frequency based tunable absorption spectrometer for detection of formaldehyde. Applied Optics 39, 4436–4443.

Moussiopoulos, N., 1990. Influence of power plant emissions and industrial emissions on the leeward ozone levels. Atmospheric Environment 24A, 1451–1460.

National Research Council, 1991. Rethinking the Ozone Problem in Urban and Regional Air Pollution. National Academy, Washington, DC.

Possanzini, M., Di Palo, V., Petricca, M., Fratarcangeli, R., Brocco, D., 1996. Measurements of lower carbonyls in Rome ambient air. Atmospheric Environment 30, 3757–3764.

Pryor, S., 1998. A case study of emission changes and ozone responses. Atmospheric Environment 32, 123–131.

Rehle, D., Leleux, D., Erdelyi, M., Tittel, F., Fraser, M., Friedfeld, S., 2001. Ambient formaldehyde detection with a laser spectrometer based on difference frequency generation in PPLN. Applied Physics B 72, 947–952.

Salas, L., Singh, H., 1986. Measurements of formaldehyde and acetaldehyde in the urban ambient air. Atmospheric Environment 20, 1301–1304.

Seinfeld, J., Pandis, S., 1998. Atmospheric Chemistry and Physics: from Air Pollution to Climate Change. Wiley Interscience, New York.

S-Plus, 2000. User's Manual, Insightful Software, Seattle, WA.

Tanner, R., Miguel, A., de Andrade, J., Gaffney, J., Streit, J., 1988. Atmospheric chemistry of aldehydes: enhanced peroxylacetyl nitrate formation from ethanol-fueled vehicular emissions. Environmental Science and Technology 22, 1026–1034.