CEEPR **Center for Energy and Environmental Policy Research**

The effect of variability in industrial emissions on ozone formation in Houston, Texas

by

Mort Webster, Junsang Nam, Yosuke Kimura, Harvey Jeffries, William Vizuete, David T. Allen

07-008 **August 2007**

A Joint Center of the Department of Economics, Laboratory for Energy and the Environment, and Sloan School of Management

The effect of variability in industrial emissions on ozone formation in Houston, Texas

Mort Webster[1,*] Junsang Nam[2], Yosuke Kimura[2], Harvey Jeffries[3], William Vizuete[3], David T. Allen[2]

[1]Massachusetts Institute of Technology, Department of Earth, Atmosphere, and Planetary Sciences, E40-408, 77 Massachusetts Avenue, Cambridge, MA 02139

[2] University of Texas, Center for Energy and Environmental Resources, 10100 Burnet Road, M/S R7100, Austin, TX 78758

[3]University of North Carolina, Department of Environmental Sciences and Engineering, School of Public Health, Chapel Hill, NC 27599

*author to whom correspondence should be addressed, fax 617-253-9845 mort@mit.edu

Abstract

Ambient observations have indicated that high concentrations of ozone observed in the Houston/Galveston area are associated with plumes of highly reactive hydrocarbons, mixed with NO_x, from industrial facilities. Ambient observations and industrial process data, such as mass flow rates for industrial flares, indicate that the VOCs associated with these industrial emissions can have significant temporal variability. To characterize the effect of this variability in emissions on ozone formation in Houston, data were collected on the temporal variability of industrial emissions or emission surrogates (e.g., mass flow rates to flares). The observed emissions variability was then used to construct region-wide emission inventories with variable industrial emissions, and the impacts of the variability on ozone formation were examined for two types of meteorological conditions, both of which lead to high ozone concentrations in Houston. The air quality simulations indicate that variability in industrial emissions has the

potential to cause increases and decreases of 10-52 ppb (13-316%), or more, in ozone concentration. The largest of these differences are restricted to regions of 10-20 km^2, but the variability also has the potential to increase region wide maxima in ozone concentrations by up to 12 ppb.

Keywords: Photochemical Grid Model, highly reactive volatile organic compounds (HRVOC), ozone, uncertainty analysis, Monte Carlo simulation.

Introduction

Ambient observations have indicated that ozone formation in the Houston/Galveston (HG) area is faster and more efficient, with respect to NO_x consumed, than other urban areas in the United States. This results in highly localized but extreme ozone events, frequently in excess of the National Ambient Air Quality Standards (NAAQSs) for ozone. It is believed that these unique characteristics of ozone formation in the Houston metropolitan area are associated with plumes of highly reactive hydrocarbons, which have been observed in airborne measurements (Kleinman et al., 2003; Ryerson et al., 2003) over or near the industrial Houston Ship Channel area. Thus, accurate quantification of industrial emissions, particularly of reactive hydrocarbons, is critical to effectively address the rapid ozone formation and the consequent high ozone events in the Houston metropolitan area.

Industrial emissions of hydrocarbons, from non-electricity generating units (NEGUs), have traditionally been assumed to be continuous at constant levels for the State Implementation Plan (SIP) development and photochemical modeling purposes. However, observational data collected during the Texas Air Quality Study in 2000 showed that industrial emissions of hydrocarbons from NEGUs have significant temporal variability (Murphy and Allen, 2005; Vizuete, 2005).

Variability in industrial emissions of hydrocarbons can be ascribed to the occurrence of both episodic emission events and variable continuous emissions. Murphy and Allen (2005) have investigated characteristics of emission events in the HG area with a focus on highly reactive volatile organic compounds (HRVOCs; defined in Texas air quality regulation as ethylene, propylene, isomers of butene and 1,3-butadiene); characteristics of emission events examined included event magnitude, event duration and source types. These emissions events

are discrete non-routine emissions events, of more than permitted amounts, with reporting required under Texas law. Since 2003, reporting is required of emissions events of over 100 lbs. of a specific compound, or over 5000 lbs of VOCs if composed of less than 2% of individual highly reactive species. Murphy and Allen (2005) showed that depending on the time, location, and magnitude of the emission event, ozone concentrations can be increased by as much as 100 ppb. Nam et al (2006) performed additional analysis using a subdomain model, and showed that 1.5% of emission events would produce an additional 10 ppb of ozone, and 0.5% of events would produce more than 70 ppb of additional ozone.

Vizuete (2005) investigated the physical and chemical processes of ozone formation and accumulation in the HG area during a small number of large magnitude emission events. Nam et al. (2006) developed computationally efficient photochemical models and examined the impact, on ozone formation in the HG area, of emission events using a stochastic characterization of the emission events described by Murphy and Allen (2005). While the previous studies provided important information on impacts of emission variability on ozone formation in the HG area, they are limited in the sense that the episodic emission events account for just a part of emission variability and contribute to just 10% of the mass of annual HRVOC emissions. As described in later sections of this paper, data from emissions monitors for several industrial sources show that there is considerable variability in routine emissions that is not high enough to require reporting as an event, but may significantly impact ozone formation. If so, representing this variability may be critical in developing ozone control strategies, since different strategies will target different parts of the probability distribution of emissions.

The overall goal of this work is to estimate potential changes in ozone formation and accumulation in the HG area due to variability in continuous hydrocarbon emissions. Because

the variability may be an important consideration in developing control strategies, this study focuses on establishing the methodology and the importance of emissions variability on ozone formation. The impact on control strategies is explored in a subsequent study. Variability in VOC emissions will be simulated based on observations of emission variability from a group of industrial sources in the HG area, and a stochastic emission inventory generator, described in the Methods section. The characterization of VOC emissions variability and the impacts of the variability on ozone formation in the Houston area are described in the Results.

Methods

The stochastic emissions inventory generator

Observations from various emission sources and ambient measurements indicate that industrial emissions of VOC have significant temporal variability. For example, Figure 1 shows the hourly measurements of the mass flow rate to a typical flare at an industrial facility in the Houston area over the course of a year. Variability in mass flow to a flare represents the variability in emissions if combustion efficiency is constant and likely represents a lower bound on emission variability if combustion efficiency decreases at high or low flows. Although the annual average mass flow rate (blue horizontal line at 2.93 kilo-lb/hr) is below the permitted annual average mass flow rate (purple line at 3.43 kilo-lb/hr), the significant temporal variability in mass flow leads to frequent exceedances of the annual average emission rate. While temporal variability is large compared to the mean, this variability does not necessarily result in a reportable emission event. The maximum allowable flow to the flare of Figure 1, when averaged on a daily basis, is 34,700 lb/hr. Only a few of the instances of high flow rates exceed this

amount and are reportable as emission events. Data from other emission sources also exhibit high variability, with different temporal patterns.

The motivating question for this study is: does the variability in VOC emissions from point sources contribute to ozone exceedences in Houston-Galveston? If, for example, the occasional spikes of extremely high VOC emissions as seen in Figure 1 are a major cause for ozone exceedences, then this has a strong implication for control strategies. Rather than lower the annual average of VOC emissions, it may be much more effective to eliminate the high emissions spikes. Before this can be explored, however, a method for representing the stochastic emissions process must be developed.

A probabilistic model of the stochastic emissions generating process was constructed. To simulate the observed temporal patterns, sampling from a simple probability density function (PDF) is not sufficient; samples from a single PDF cannot characterize the flare flow. Standard time series methods are similarly unable to capture the characteristic patterns in the VOC emissions. For example, in Figure 2, the cumulative distribution function (CDF) is shown for the observations from Flare 1 and for a simulation of an auto-regressive moving-average (ARMA) with all significant terms. While the ARMA fit approximates the mean, it does not reproduce the standard deviation or the tails of the observations. Another application area where standard time series approaches have proven insufficient is in forecasting electricity prices in deregulated markets. Time series of electricity prices show similar volatility and extreme jumps with low frequency. Recent work in that area (Johnson & Barz, 1999; Knittel and Roberts, 2001) has shown jump-diffusion and markov process models to be superior, and a similar approach is used here.

Mass flows to the flare in Figure 1 are composed of various components in magnitude, including nearly constant, routinely variable, and allowable episodic mass flow rates (Figure 3). Therefore a more appropriate model for this process is a mixture of multiple PDFs (Cornell, 2002), each accounting for one of the components of the mass flow variability. To determine the form of the PDF (normal or log-normal) that should be used for the nearly constant, routinely variable and allowable episodic mass flows, normal probability plots are used (Figure 4). A normal probability plot (Hogg and Ledolter, 1992, p. 137) graphs the quantiles of the observations against the standardized normal scores, defined as $\left(z = \frac{Y - \overline{Y}}{s} \right)$. $\overline{Y}$ is the mean emission rate and *s* is the standard deviation of the normal function that best fits the emission rate. If the observations are normally distributed, the plot will form a straight line with a slope of 1/s and will intersect the point ($\overline{Y}$, 0). In addition, the log of the sorted observations are plotted against the standardized normal scores; if the relationship is linear between the inverse normal and the logarithm of the mass flow rate, it is reasonable to assume that the mass flows are lognormal. An alternative would be to graph the sorted observations against the quantiles of an assumed distribution, known as a ”q-q plot”. The advantage of the normal probability plot used here is in identifying breaks between different components, each with different means and variances. The q-q plot, by contrast, is best suited to assessing whether the data is drawn from a single theoretical distribution.

By looking for regions of the observations that are linear with constant slope and intercept, the different components of the distribution can be identified. Continuing with the example emissions from Figure 1, a total of three components of variable mass flows were identified: one normal distribution and two lognormal distributions, as indicated with the colored lines in Figure 4. In addition, the mean and standard deviation of the distribution were obtained

from the midpoint and slope of the fitted line, respectively. Figure 5 shows resultant PDFs for each component identified (Figure 5a) and the normalized PDFs based on the proportion of mass flow associated with each component (Figure 5b) for the mass flow rates shown in Figure 1. When these distributions of mass flows are used for other flares in the HG area, the distributions are scaled such that the mean of the distribution exactly equals the deterministic value for hourly emissions, and that the coefficient of variation, defined as the standard deviation normalized by the mean, is preserved.

Once PDFs for each emission mode are calculated, the duration of emissions in one mode, before transition to another mode, is simulated using an exponential distribution function. Exponential distributions are probability distributions widely used to model the time between events. Each hour's emissions is identified as belonging to one of the three component PDFs determined as above. Thus the time series is converted into a series of the number of hours in a given state before a jump to the next state (a Poisson model). The mean time for an exponential distribution for each component is then calculated. For the flare flow in Figure 1, mean time was calculated as 7.41, 4.41 and 1.15 hour for the nearly constant, routinely variable, and allowable episodic emission components, respectively, and the distribution of the duration in each component is shown in Figure 6.

Overall, the algorithm for generating emission samples is to (1) randomly sample the mode of emissions using the proportions of each emission mode, (2) randomly sample the number of hours to remain in the current mode using the exponential distributions for the selected emission mode, and (3) randomly sample the emission rate for each hour based on the PDF of emission rates for the current mode. In sampling emissions in step (3), an autocorrelation of 0.99 with the emission rate of the previous hour is imposed. Hourly emissions

from a source are sampled as a random draw from a standard normal distribution, and then transformed to a sample from the current component normal or lognormal distribution. The emissions from each source for each hour is sampled by imposing correlation with the previous hour's emissions from that source, using a standard algorithm for sampling correlated normal variates (Press et al, 1992). The autocorrelation of 0.99 was estimated from the observations, and is a reasonable model of a continuous industrial process, since the best predictor of one hour's activity level and operating conditions is the previous hour's. Based on the number of hours selected in step (2), step (3) is repeated. For example, if the first component, nearly constant, was selected in step (1) and two-hour duration was selected in step (2), then emission rates would be randomly sampled from the PDF of the first component for two hours before randomly selecting the next component. This method produces samples of emissions that closely match the cumulative distribution properties of the observations. Note that this model of emissions includes only permit allowable emission rates. Emission events above permitted levels are not included. The nature and impacts of these emission events have been described by Murphy and Allen (2005) and Nam, et al. (2006).

Application of stochastic emission inventory to the HG area

A set of observations similar to that shown in Figure 1 were obtained from various flares and cooling towers in the HG area. The models to simulate emission variability were developed using the process described in the previous section. Table 1 summarizes these observations and parameters for the models developed. Emissions from different sources exhibit different patterns of variability, but only limited data on that variability is available. For this study, therefore, individual flares, cooling towers and other point sources in the HG area were assigned one of the

unit operation models developed and overall time-varying emissions from the point sources were simulated with the model. Because of the limited data available, we randomly assign a variability pattern to each source in the emissions inventory. Cooling towers were randomly assigned one of the cooling tower models. Flares were randomly assigned one of the flare models. Stack and fugitive emissions were assumed to be lognormally distributed, as shown in Table 1. The mean of the selected model for simulating emissions is scaled to be equal to the average emission rates in the inventory for each emission source. In addition, the standard deviation of the models was scaled to preserve the coefficient of variation (the ratio of the standard deviation to the mean) shown in Table 1. For emission sources other than flares, all emissions except VOCs were kept intact throughout the simulation. For flares, both VOC and NO_x emissions were assumed to scale with flow rate, so for flares, the same pattern of variability was assumed for both VOC and NO_x emissions. The composition of the emissions from all sources was assumed to be constant, with only the temporal variability in the magnitude of emissions changing.

Because so few emissions sources had hourly emissions monitoring data available, it is not clear whether the variability of emissions in the region as a whole is over- or underestimated. As a first study of the importance of this variability, the available data and random assignment have been used as a type of bootstrapping. However, there are reasons to believe that this may be a reasonable first-order estimate: the sample emissions data come from among the largest sources in the HG area, and it has been shown (UNC, 2004) that the largest few sources dominate the VOC emissions. For example, the top 20 flares are responsible for 45% of all flare VOC emissions. Nevertheless, data from more point sources are needed to better characterize the variability, and this is one area that future analysis could focus.

Table 2 compares the mean, standard deviation, and several fractiles of the observations with simulations of 10,000 hours of the corresponding stochastic emission model. Figure 7 compares the cumulative distribution function (CDF) of the observations with the CDF of emissions for four of the flares.

Air quality modeling

The impact of industrial point source variability on ozone formation was assessed using the Comprehensive Air Quality Model with extensions (CAMx) (Environ, 2004). In this work, a computationally efficient version of CAMx, referred to as a sub-domain model and described by Nam, et al (2006), was used. The overall strategy in developing the sub-domain model was to (1) identify a geographical region (sub-domain), from a full, 3-D photochemical model, (2) create a computationally efficient photochemical model of the sub-domain, and (3) analyze many scenarios or snapshots of variable emissions using the sub-domain model. Steps 1 and 2 in the development of the model are analogous to the methods used by Nam, et al (2006) and are only summarized here. Step 3 is described in the results section.

The geographical region (sub-domain) to be modeled is the HG 1 km domain, shown as the region in red in Figure 8. CAMx simulations using the full domain, shown in Figure 8, were used to develop boundary and initial conditions for the sub-domain. Details of the meteorological modeling and the VOC and NOx emission inventory development for simulation of the full domain are available from the TCEQ (2006) and from Nam, et al (2006). Briefly, meteorological inputs were based on results from the NCAR/Penn State Mesoscale Meteorological Model version 5 (MM5). Emission inventories were prepared by the Texas Commission on Environmental Quality (TCEQ). The TCEQ data includes a MOBILE6-based

inventory developed for on-road mobile source emissions, and emissions for non-road mobile and area sources developed using the U.S. EPA's NONROAD model, using local activity data when available. Biogenic emission inventories were estimated using the GLOBEIS emission model with locally developed land cover data. Point source emissions data were developed with TCEQ's point source database and special inventory. Approximately 150 tons/day of reactive olefin species were added to approximately 100 point sources in the domain, based on ambient measurements made by aircraft (Ryerson et al., 2003). These point source inventory additions are commonly referred to as the imputed inventory, since the added emissions were estimated based on ambient measurements rather than reported inventories. The imputed point source inventory and the other components of the emission inventory, described above, were used as the base case in this work and will be collectively referred to as the imputed inventory. In short, TCEQ's standard emissions inventory was used for all sources without change except for selected VOC emissions (and VOC and NOx emissions from flares) from point sources. The mean of emissions from these sources in the stochastic runs are equal to the TCEQ emissions in its inventory, as described above. Both the sub-domain modeling and the full domain modeling in the region with industrial emissions were performed at a 1 km spatial resolution.

The full domain model was used to establish initial conditions and time varying boundary conditions for the sub-domain model. Calculations reported by Nam, et al (2006) indicate that the sub-domain model responds to temporal variability in industrial emissions in a manner that correlates (r^2>0.96) with the response of the full domain model.

The sub-domain model was run for two episode days: 25 August and 30 August, 2000. These two days were selected because there was rapid ozone formation on both days and distinctly different meteorological conditions on the two days had the potential to lead to

different processes for ozone formation and accumulation. Details of the meteorological conditions on these two days have been reported by Nam, et al (2006).

Results and Discussion

The Results and Discussion will be presented in two parts. The first part summarizes the stochastic emission inventories; the second part describes the air quality modeling based on those inventories.

Stochastic emission inventories

Figure 9 shows probability distributions for 500 random realizations of one day's emissions (12,000 samples of hourly emissions), and numerical values of the mean, standard deviation, and selected percentiles are given in Table 3. Note that the total of all VOC emissions from the industrial point sources show relatively little variation in any given hour. This is a consequence of the Law of Large Numbers: the variance of a sum is significantly smaller than the variance of any individual component. Similarly, the sum of all hydrocarbon flares in Houston Galveston shows a relatively small variance. The mean of these distributions is virtually equal to the values from the deterministic emissions inventory.

As the area and the time period over which emissions are reported decrease, however, emission variability becomes more evident. The VOC emissions from two 1-km by 1-km grid cells near the Ship Channel, where significant VOC sources are concentrated, have 95% probability bounds that span more than a factor of two (Figure 9c; Table 3). A third grid cell near downtown Houston has 95% bounds that span a factor of 1.6. The three largest flares in the Houston Galveston region have 95% probability bounds that span factors of 9, 4, and 16, respectively (Table 3).

In addition to variability of any individual source, which can be quite large, it is important to explore other statistical properties of the combination of point sources. Figure 10 shows the probability, in any given hour, that at least N sources have emissions greater than or equal to a factor of 2, 5, and 10 times its annual average emissions, examining N over the range from 1 to 9. Only the 50 largest point sources, in terms of annual average emission rates, are considered. These sources comprise 20% of the total VOC emissions. For example, there is a 40% chance in any one hour that 8 or more of these large sources are emitting at greater than twice their average value. There is an 8.6% chance that at least 2 sources will emit more than five times their average in the same hour, and there is a 12% chance that in any hour, one of these sources will be emitting 10 times their average rate.

Impacts of VOC emission variability on ozone formation

A total of 50 sets of stochastic emission inventories were randomly generated with the models described in the Methods section and simulations representing 25 August and 30 August, 2000 were performed using these inventories, for a total of 100 simulations. Figure 11 shows the differences in ozone concentration on 25 August between using the 45th stochastic inventory and the deterministic imputed inventory. The 45th stochastic inventory showed the largest increase in ozone concentration for the August 25 meteorology, as shown in Figure 12. Since the stochastic inventory has both higher and lower VOC emissions across the HG area over the course of the day, ozone concentrations using the stochastic inventory are both higher and lower than using the imputed inventory without VOC emission variability depending on time of day and location. At conditions that lead to maximum difference in ozone concentration, ozone concentrations predicted using the stochastic inventory are approximately 82 ppb higher than using the imputed

inventory without variable VOC emissions. Ozone concentrations are also up to 6 ppb lower using the stochastic inventory than using the imputed inventory with constant industrial emissions.

Figure 12 summarizes the maximum changes in ozone concentrations that were observed, both positive and negative, when all 50 stochastic emission inventories were used for simulations on 25 August and 30 August, 2000. Specifically, the quantity presented is the maximum difference in ozone concentration between using the stochastic inventory and the deterministic inventory. On the top and bottom of each column, the ozone concentration for the stochastic inventory, at the time when the maximum difference occurred, is indicated in ppb. In the simulations of 25 August, the maximum difference in ozone concentration is largest when the 45^{th} stochastic emission inventory was used; the ozone concentrations are 24 ppb and 106 ppb when the deterministic and the stochastic inventory were used, respectively, at conditions that lead to the maximum increase in ozone concentration. In the simulations of 30 August, the largest maximum difference in ozone concentration occurred when the 48^{th} stochastic inventory was used; a 43 ppb decrease in ozone concentration was predicted, relative to the deterministic imputed inventory. The probability distributions of maximum difference in ozone concentration for both days are shown in Figure 13. The maximum increase in ozone concentration at any hour and location from including the variability in VOC point-source emissions has 90% bounds of 11-52 ppb for the August 25 meteorology and 10-23 ppb for the August 30 meteorology. In percentage terms, the maximum difference in ozone concentration between the stochastic and deterministic version had a 90% probability of being between 13% and 316% for August 25, and between 10% and 85% on August 30. In order to give a sense of the range of uncertainty, Figure 14 shows the hourly spread of ozone concentrations for two grid cells discussed above, (39,55)

and (40, 43). Note that for (39, 55), the uncertainty is large enough to affect whether or not ozone exceeds the one-hour standard of 120 ppb. The early morning peak in Figure 14b is a result of the meteorology for August 25, when there was a period of stagnation from 6am to 11am, and a low mixing height. Then winds increased and advected the high ozone to the west. This is a typical meteorology pattern for Houston, and is the most common pattern for extreme ozone exceedence days.

Simulations of the two episode days exhibit different responses of ozone formation due to variable point source emissions. For example, the 45th stochastic inventory led to a maximum difference in ozone concentration of 82 ppb in a grid cell which is the 40th to the east and the 54th to the north from the southwest corner of the region in red, shown in Figure 11, at 7am 25 August. At the same time of day and location on 30 August, the ozone concentration was not affected by variable point source emissions. Distinctly different meteorological conditions on the two days led to these different behaviors of ozone formation for the same stochastic emission inventory.

This change in ozone concentration did not always increase the peak ozone concentration in the 1-km domain over the course of the day. For example, on 25 August the daily maximum ozone concentration using the stochastic inventory is up to 11.9 ppb higher and up to 6 ppb lower than when the non-stochastic imputed inventory was used, depending on the stochastic inventory used. The average and standard deviations for the increases and decreases in the sub-domain wide daily maximum ozone concentrations were 3.3±2.9 and 2.4±1.7, respectively. These results can be contrasted with a maximum increase of 82 ppb and a maximum decrease of 56 ppb, shown in Figure 12, and average increases and decreases of 24±15 and 17±14, respectively. A total of 31 out of the 50 sets of stochastic inventories led to increases in daily maximum ozone

concentration in the sub-domain. For 30 August, the daily maximum ozone concentration increased for 37 sets of stochastic inventories. The maximum increase in the peak ozone concentration in the 1-km domain was approximately 10.7 ppb and maximum decrease was approximately 4.5 ppb.

In summary, variability in continuous industrial emissions has the potential to have a significant impact on ozone formation in the Houston-Galveston area. Increases and decreases of 10-52 ppb or more in ozone concentration are possible as a result of emission variability. The largest of these differences are restricted to regions of 10-20 km^2 (see Figure 11), but the variability also has the potential to increase region wide maxima in ozone concentrations by to 12 ppb.

These results raise important questions about effective ozone control strategies for ozone in the Houston Galveston region, and perhaps other regions with significant petrochemical industrial facilities such as Baton Rouge or New Jersey. If some of the high ozone episodes are the result of less frequent higher than average emission rates, rather than the mean emissions, control strategies that lower the annual average may prove ineffective at reducing ozone exceedances. Further, strategies that target the upper tails of the distribution may have very different, perhaps lower, compliance costs than traditional approaches. This is an important area for future inquiry.

References

Cornell, J., 2002. *Experiments with mixtures: designs, models, and the analysis of mixture data*. Wiley, New York.

Environ International Corporation (Environ), 2004. User's Guide: Comprehensive Air Quality Model with Extensions (CAMx), Version 4.03, Document and model are available online at http://www.camx.com

Environ International Corporation (Environ), 2006. Comprehensive Air Quality Model with Extensions (CAMx) Pre-processors, Accessed January 2006 at http://www.camx.com/down/support.php

Hogg, R. V., Ledolter, J., 1992. *Applied Statistics for Engineers and Scientists*. Macmillan Publishing Company, New York.

Johnson, B. Barz, G., 1999. "Selecting Stochastic Process for Modelling Electricity Prices." *Energy Modelling and the Management of Uncertainty*, Risk Publications.

Kleinman, L.I., Daum, P.H., Imre, D., Lee, Y.N., Nunnermacker, L.J., Springston, S.R., Weinstein-Lloyd, J., Rudolph, J., 2002. Correction to "Ozone production rate and hydrocarbon reactivity in 5 urban areas: A cause of high ozone concentration in Houston". Geophys. Res. Lett. 30 (12), 1639.

Knittel, C. R., Roberts, M., 2001. "An Empirical Examination of Deregulated Electricity Prices," POWER WP-087, University of California Energy Institute.

Murphy, C.F., Allen, D.T., 2005. Hydrocarbon emissions from industrial release events in the Houston- Galveston area and their impact on ozone formation. Atmospheric Environment 39 (21), 3785-3798.

Nam, J., Kimura, Y., Vizuete, W., Murphy, C.F., Allen, D.T., 2006. Modeling the impacts of emission events on ozone formation in Houston, Texas. Atmospheric Environment 40 (28), 5329-5341.

Press, W. H., Teukolsky, S. A., Vetterling, W. V., Flannery, B. P., 1992. *Numerical recipes in C.* Cambridge University Press, Cambridge, England; New York, N.Y.

Ryerson, T.B., Trainer, M., Angevine, W.M., Brock, C.A., Dissly, R.W., Fehsenfeld, F.C., Frost, G.J., Goldan, P.D., Holloway, J.S., Hubler, G., Jakoubek, R.O., Kuster, W.C., Neuman, J.A., Nicks, D.K., Parrish, D.D., Roberts, J.M., Sueper, D.T., Atlas, E.L., Donnelly, S.G., Flocke, F., Fried, A., Potter, W.T., Schauffler, S., Stroud, V., Weinheimer, A.J., Wert, B.P., Wiedinmyer, C., Alvarez, R.J., Banta, R.M., Darby, L.S., Senff, C.J., 2003. Effect of petrochemical industrial emissions of reactive alkenes and NOx on tropospheric ozone formation in Houston, Texas. J. Geophys. Res. 108 (D8), 4249.

Texas Commission on Environmental Quality (TCEQ), 2006. Houston-Galveston-Brazoria Ozone SIP Mid-Course Review Modeling, Accessed January 2006 at http://www.tceq.state.tx.us/implementation/air/airmod/data/hgb1.html

University of North Carolina, 2004. Stochastic Emission Inventories of Continuous Emissions. Houston Advanced Research Center, H13.2003 Final Report Appendix B, available at http://files.harc.edu/Projects/AirQuality/Projects/H013.2003/H13AppendixB.pdf

Vizuete, W., 2005. Implementation of Process Analysis in a three dimensional air quality model, PhD thesis, University of Texas.

Source name (number of observations)	Type	Component	Number of observations	Normalized proportion	Mean time	Normal or Lognormal	Standard deviation	Mean or Mean (LN Value)
Flare 1	Flare	1	4913	0.599	7.41	N	0.34	2.43
(8208)		2	3223	0.393	4.41	LN	0.36	1.23
		3	68	0.008	1.15	LN	0.8	2.19
Flare 2	Flare	1	360	0.501	3.34	N	0.8	1.89
(720)		2	331	0.460	3.5	N	2.5	4.99
		3	28	0.039	2.63	LN	0.33	2.45
Flare 5	Flare	1	1757	0.486	3.09	N	225.52	758.83
(3624)		2	1706	0.472	5.36	N	1197.5	1721.7
		3	147	0.041	1.04	LN	0.71	8.08
HC Flare	Flare	1	128	0.071	0.52	LN	1.12	0.07
(1800)		2	1625	0.903	15.28	LN	0.26	0.9
		3	46	0.026	1.35	LN	0.3	1.58
Olefins Flare	Flare	1	1078	0.599	14.18	N	0.53	1.6
(1800)		2	700	0.389	10.58	N	3.1	4.91
		3	21	0.012	6.5	LN	0.42	2.95
FCCU	Flare	1	12383	0.743	195.31	N	3	20
(17533)		2	4270	0.256	17.85	N	11.93	29.38
		3	12	0.001	12.6	LN	1.2	3.97
Merox Flare	Flare	1	307	0.056	10.65	N	874.78	4.63
(17543)		2	15705	0.914	200.07	N	100.19	494.09
		3	1152	0.03	45.6	LN	0.001	6.63
Low Pressure	Flare	1	200	0.011	9.37	LN	0.48	3.04
Flare		2	16830	0.971	229.58	LN	0.05	3.23
(17543)		3	314	0.018	132.63	LN	0.52	3.46
General Service	Flare	1	16892	0.964	155.47	N	1.66	21.45
#1		2	405	0.023	2	LN	0.38	3.26
(17543)		3	233	0.013	5.78	LN	0.43	3.47
General Service	Flare	1	17322	0.67	508.65	N	1.31	17.8
#2		2	141	0.25	4	LN	0.08	3.05
(17543)		3	70	0.08	2.9	LN	0.7	3.26
Cooling Tower 1	Cooling	1	243	0.779	5.72	LN	0.2	-2.69
(314)	Tower	2	64	0.221	0.8	LN	0.3	-0.73
Cooling Tower 2	Cooling	1	98	0.291	4.26	N	6.3	0.23
(340)	Tower	2	148	0.439	2.61	N	0.39	0.67
		3	91	0.270	2.29	LN	0.68	0.28
Stacks/ Fugitives		1	NA	1	NA	LN	0.56	-0.15

Table 1. Observations from emission sources and fitted parameters

Flare Name		mean	stdev	0.05	0.1	0.25	0.5	0.75	0.9	0.95
Flare 1	Actual	2.92	1.65	2.09	2.21	2.40	2.66	3.10	3.81	4.43
	Simulation	2.90	0.95	2.07	2.17	2.38	2.70	3.10	3.63	4.29
Flare 2	Actual	3.71	2.64	1.09	1.34	1.97	2.81	4.78	7.42	9.64
	Simulation	3.73	3.02	1.09	1.32	1.84	2.91	4.33	7.36	11.07
Flare 5	Actual	1398.85	1278.19	614.25	682.75	819.50	1042.51	1703.33	2514.84	2696.56
	Simulation	1311.27	698.98	629.79	704.06	832.58	1024.56	1571.65	2469.56	2669.95
HC Flare	Actual	2.47	0.74	1.44	1.72	2.01	2.43	2.86	3.28	3.69
	Simulation	2.54	0.71	1.53	1.72	2.09	2.51	2.93	3.32	3.58
Olefins Flare	Actual	3.11	2.77	1.07	1.29	1.60	1.92	4.29	5.85	6.92
	Simulation	2.98	2.41	1.00	1.20	1.51	1.94	4.30	5.77	6.81
FCCU	Actual	21.87	7.90	14.10	17.42	19.02	20.63	23.98	29.58	34.76
	Simulation	22.23	5.41	16.10	16.92	18.85	21.17	23.71	31.25	34.51
Merox Flare	Actual	500.21	136.97	284.82	342.40	432.96	501.94	582.40	660.94	737.14
	Simulation	488.84	150.79	42.98	351.78	435.32	512.64	572.60	630.95	684.81
Low Pressure	Actual	25.54	1.89	23.67	24.10	24.54	25.05	26.41	27.59	28.20
	Simulation	25.56	1.87	23.17	23.49	24.39	25.20	26.62	27.70	28.59
General Service #1	Actual	21.70	2.11	18.96	19.50	20.44	21.67	22.85	23.49	24.05
	Simulation	21.86	3.07	18.40	18.96	19.85	20.89	22.17	23.90	24.68
General Service #2	Actual	17.85	1.45	15.86	16.17	16.79	17.65	19.09	19.42	19.61
	Simulation	17.68	1.37	15.59	15.97	16.62	17.41	19.07	19.34	19.54
Cooling Tower #1	Actual	0.18	0.24	0.02	0.03	0.06	0.09	0.20	0.50	0.59
	Simulation	0.19	0.23	0.02	0.03	0.06	0.09	0.20	0.53	0.66
Cooling Tower #2	Actual	0.75	0.57	0.16	0.21	0.30	0.68	0.94	1.33	1.72
	Simulation	0.79	0.72	0.18	0.22	0.32	0.68	0.95	1.49	1.89

Table 2. Comparison of mean, standard deviation, and selected fractiles of VOC emissions between actual emissions samples (see Table 1) and 10000 simulated hours from the stochastic emission generator.

Percentile	All Sources	All Flares	(40, 43)	(18, 51)	(39, 55)	3 Flares with Highest Avg Emi.		
0.01	6.77	1.21	0.015	0.023	0.023	0.043	0.049	0.005
0.025	6.83	1.24	0.016	0.024	0.024	0.052	0.061	0.010
0.05	6.87	1.27	0.016	0.026	0.026	0.062	0.077	0.023
0.25	7.02	1.35	0.018	0.031	0.031	0.107	0.096	0.056
0.5	7.14	1.41	0.019	0.035	0.035	0.156	0.101	0.069
0.75	7.27	1.48	0.021	0.041	0.041	0.229	0.106	0.083
0.95	7.48	1.61	0.024	0.051	0.051	0.391	0.216	0.131
0.975	7.57	1.66	0.026	0.055	0.055	0.469	0.249	0.157
0.99	7.69	1.73	0.028	0.061	0.060	0.558	0.283	0.202
mean	7.15	1.42	0.020	0.037	0.037	0.183	0.112	0.074
stdev	0.19	0.11	0.003	0.008	0.008	0.109	0.042	0.045

Table 3. Mean, standard deviation, and selected fractiles of point source VOC emissions (tons/hr) in Houston Galveston.

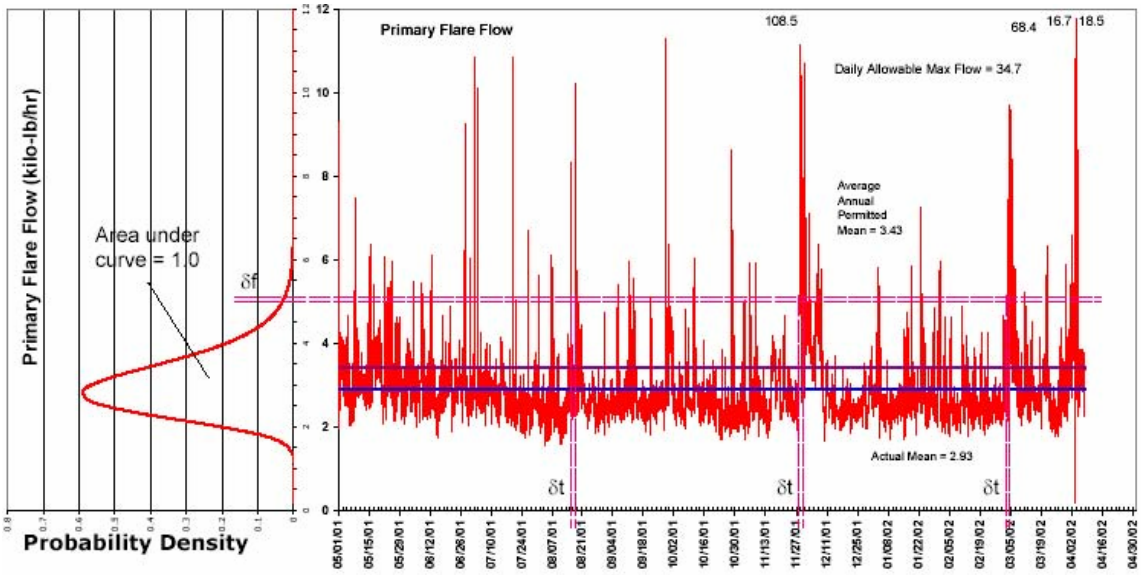

Figure 1. Probability distribution function (PDF) and time series of mass flow rate to a flare at an industrial facility in the HG area.

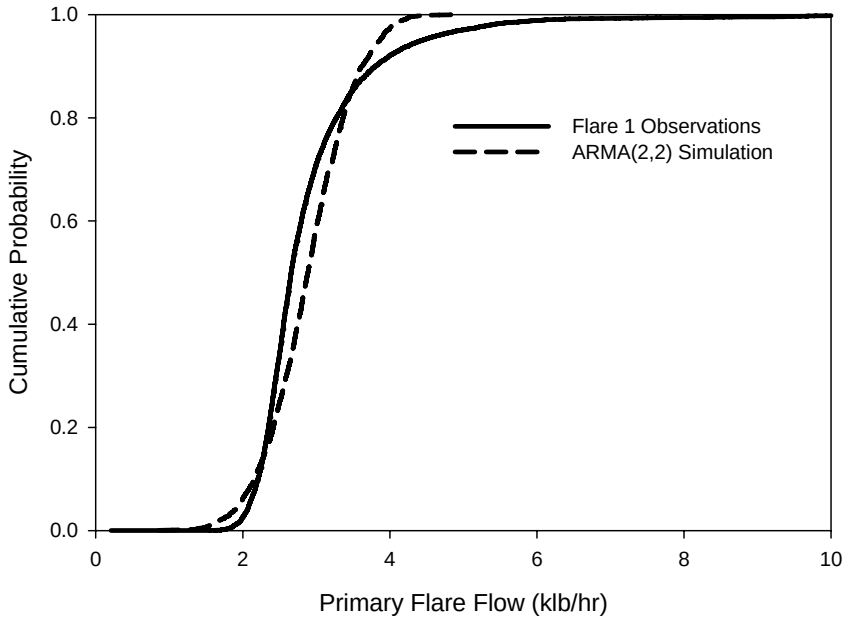

Figure 2. Cumulative Distribution Functions for observations from Flare 1 and from an auto-regressive moving-average (ARMA) model.

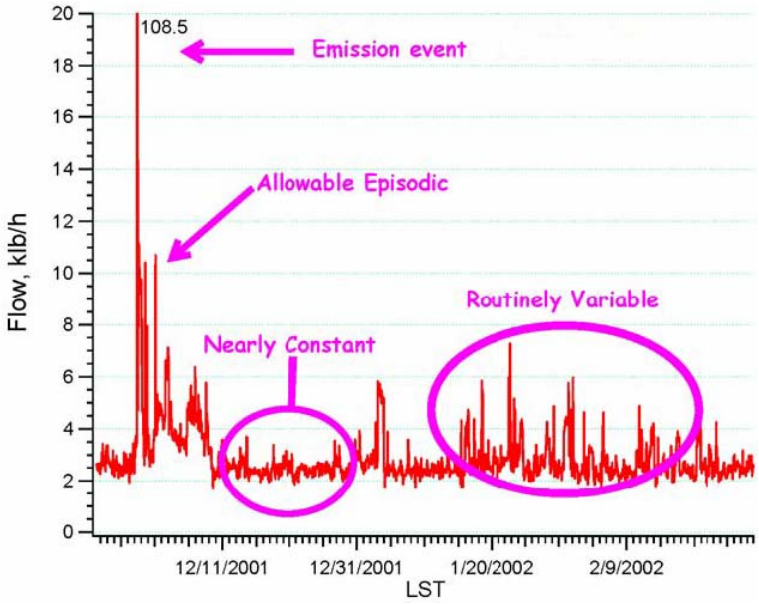

Figure 3. Components of mass flows: emission event, allowable episodic, routinely variable and nearly constant.

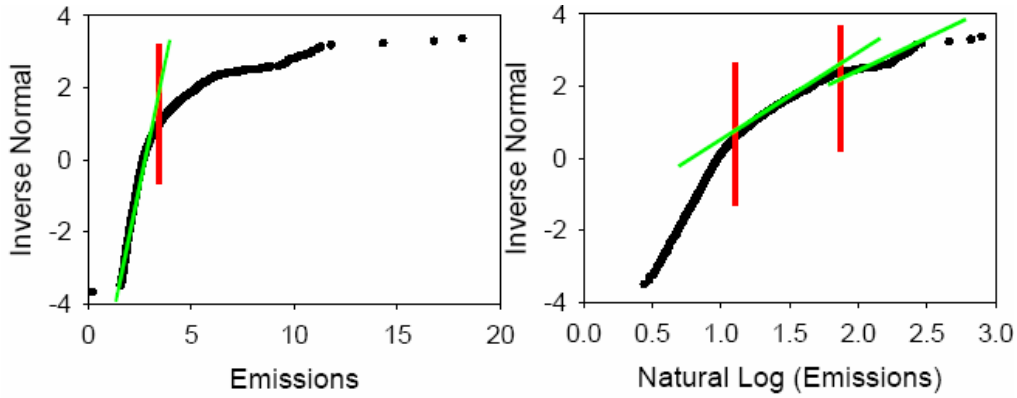

Figure 4. Decomposition into three PDFs of emissions from flare shown in Figure 1. Inverse normal is expressed in the standard form (z) in relation to the mean and standard deviation $\left(z = \frac{Y - \overline{Y}}{s} \right)$.

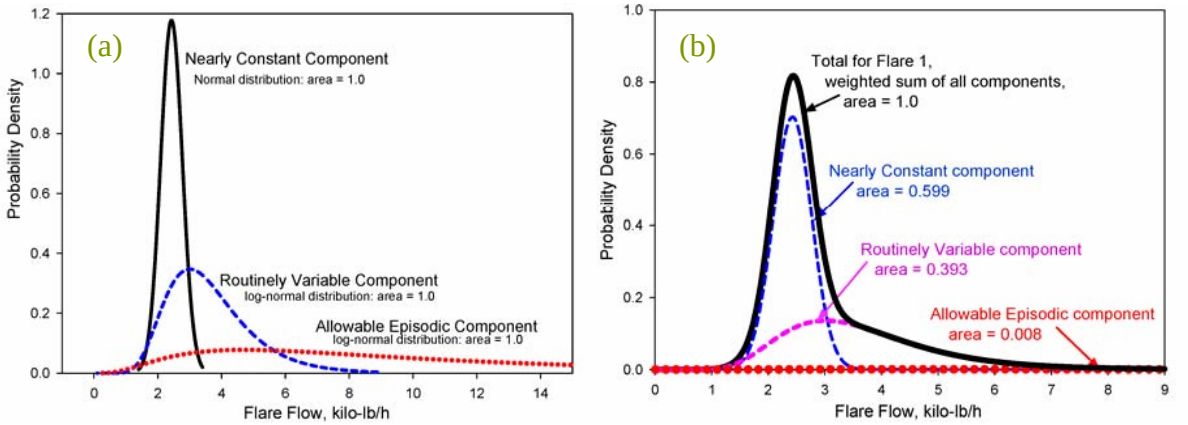

Figure 5. (a) PDFs for each component of emissions from the flare in Figure 1 (b) normalized PDFs of three components based on the proportion of each emission component

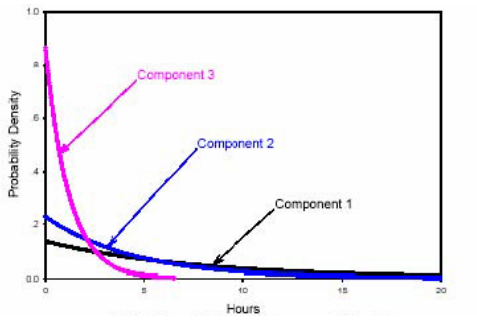

Figure 6. Time within each emission component before transition to next component

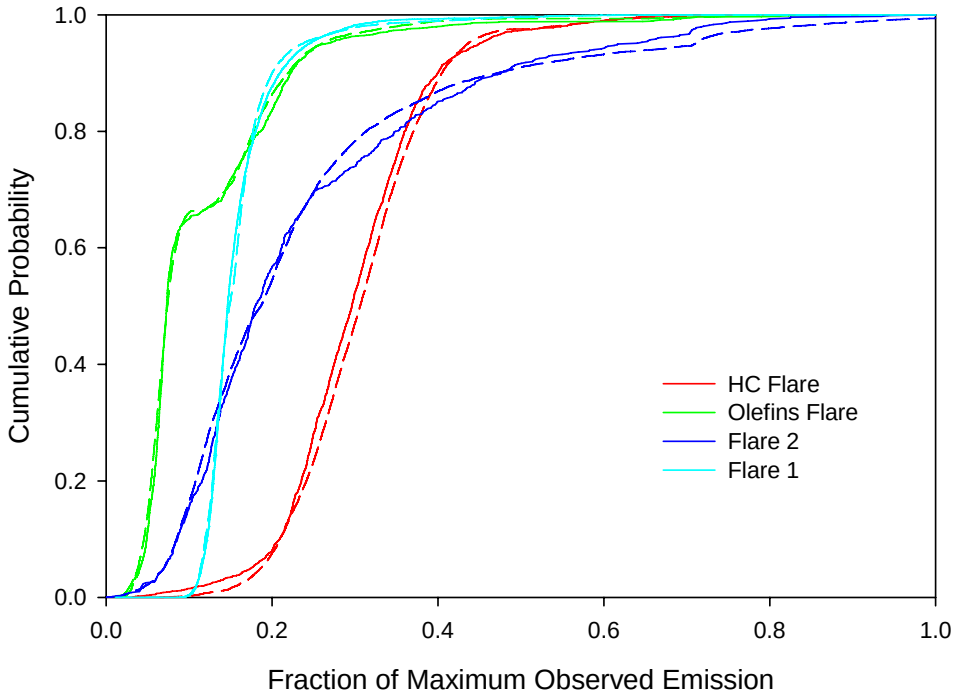

Figure 7. Comparison of cumulative probability distribution functions for observed (solid lines) and 10,000 simulated hours (dashed lines) for four flares. X-axis shows emissions as a fraction of the maximum observed emission.

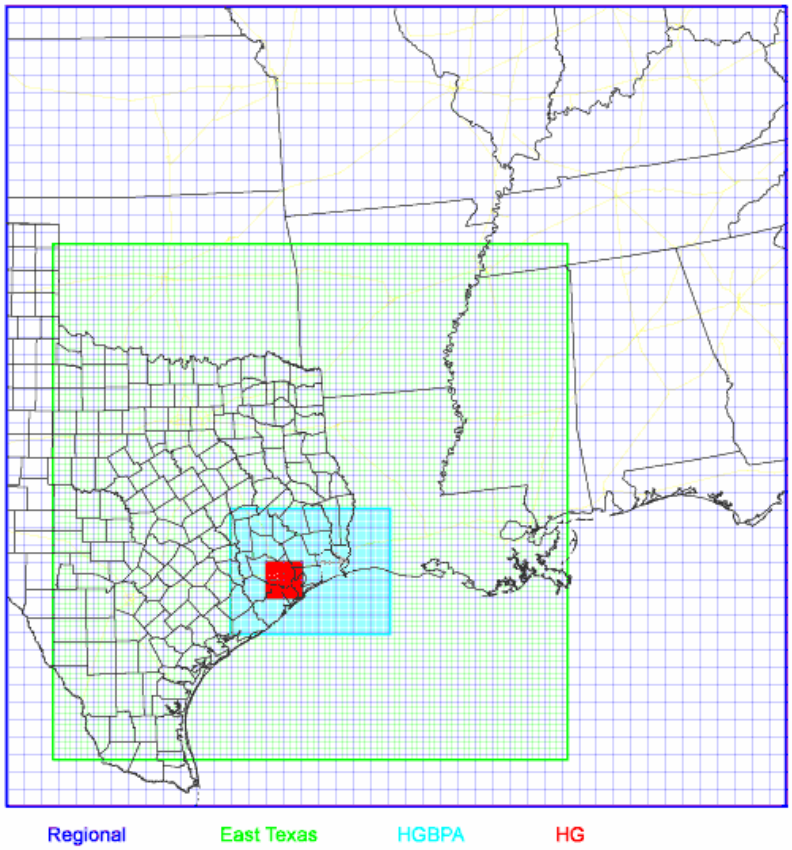

Figure 8. Modeling domain used in the study. The Regional, East Texas, Houston-Galveston-Beaumont-Port Arthur (HGBPA), and Houston Galveston (HG) nested domains had 36, 12, 4 and 1 km resolution, respectively.

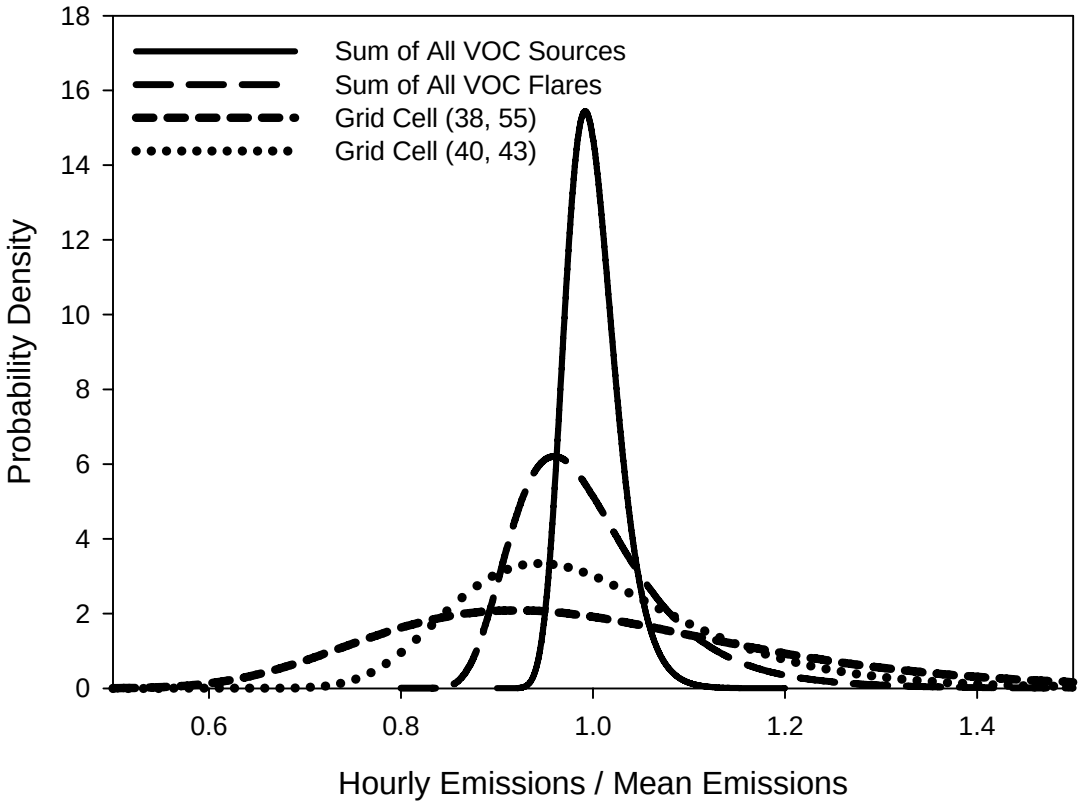

Figure 9. Variability of hourly VOC emission rates for all industrial point sources in the HG domain (solid line), all industrial flares in the HG domain (long dash), and for the industrial point sources within two representative 1 km^2 grid cells.

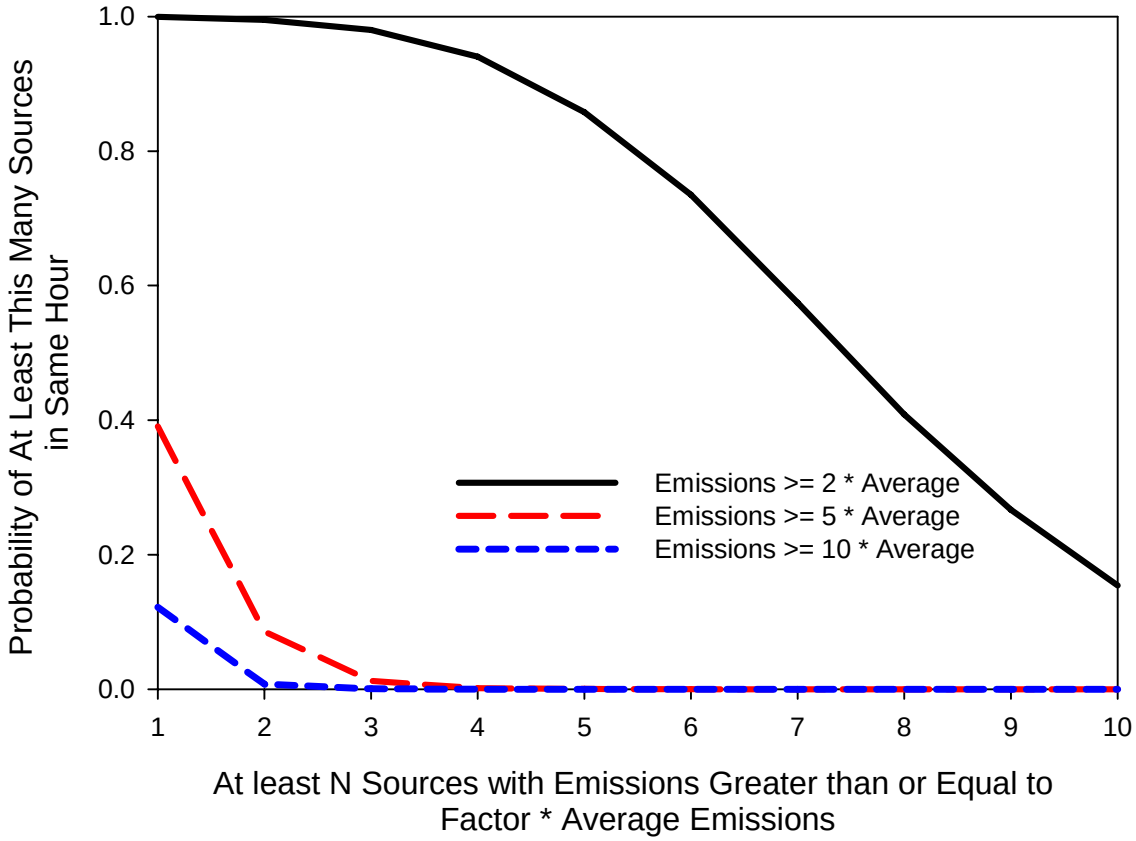

Figure 10. The probability that N or more point source emissions are a factor of 2, 5, and 10 times their annual average within the same hour.

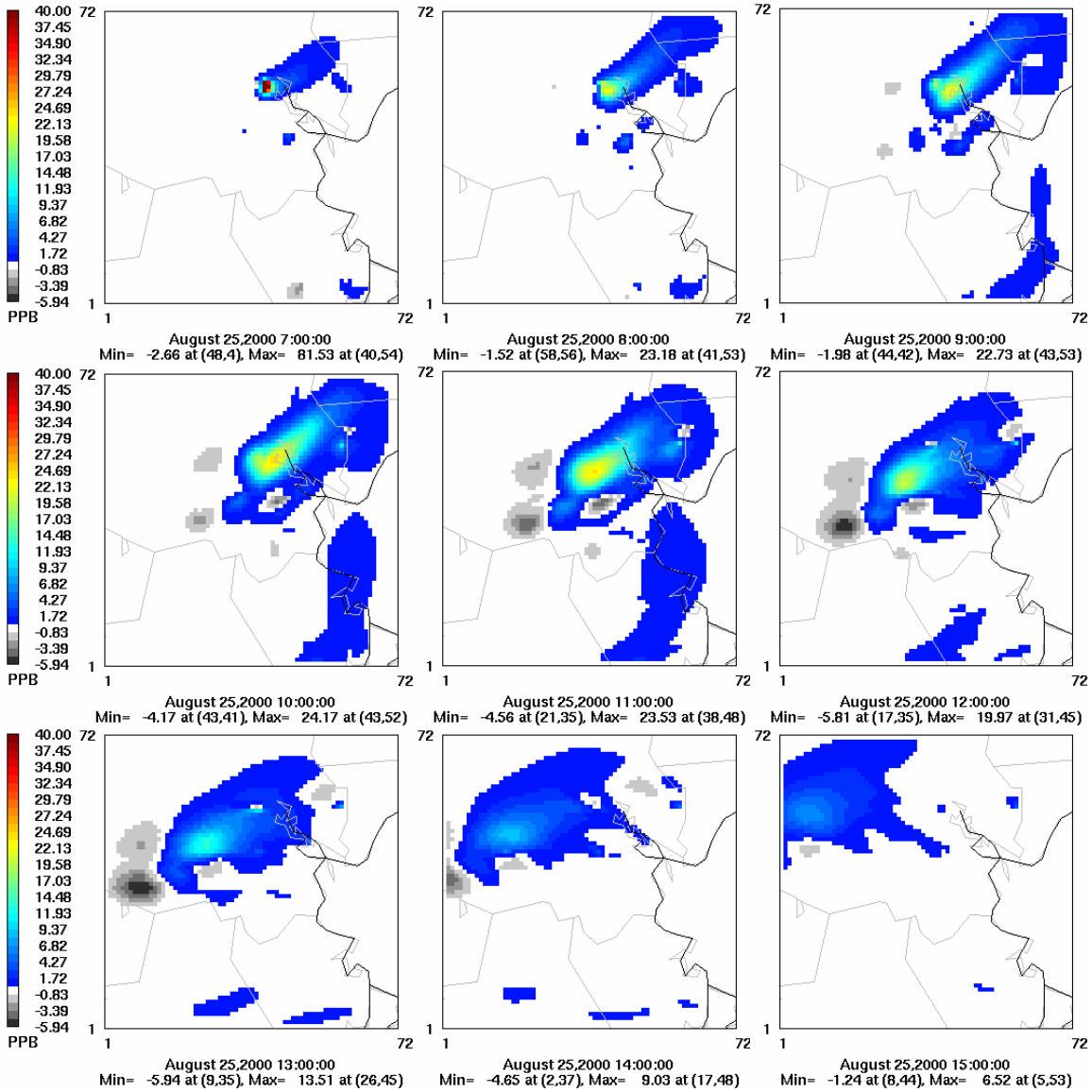

Figure 11. Difference in ozone concentration for 25 August from 700 hr to 1500 hr between using the 45th stochastic imputed emission inventory and using the deterministic imputed inventory.

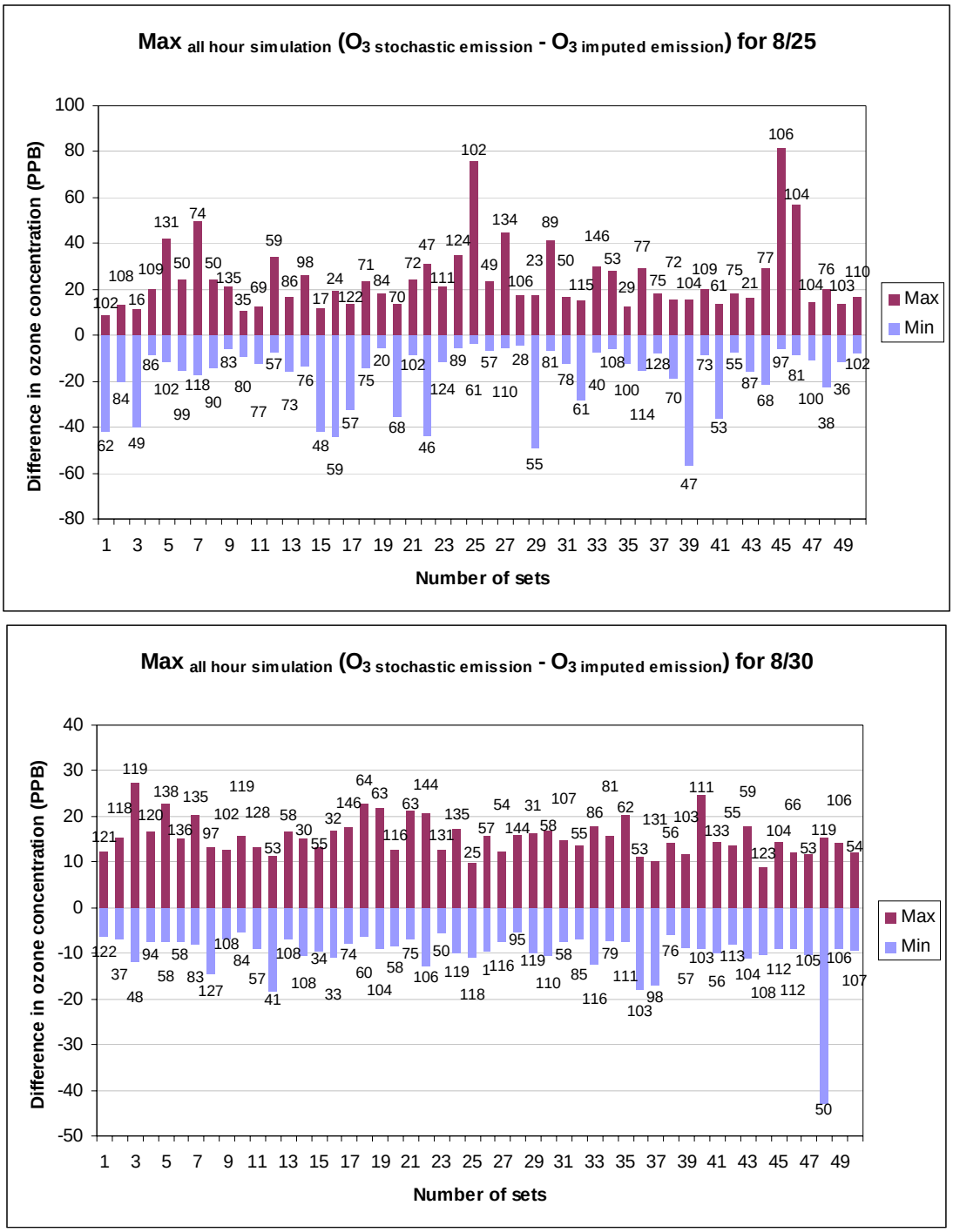

Figure 12. Maximum difference in ozone concentrations in one day simulations representing 25 August and 30 August, 2000. The difference is taken between the imputed inventory with constant industrial emissions and the stochastic inventory for 10 instances of the stochastic

inventory. Ozone concentrations (ppb), using the stochastic inventory, at the time the maximum difference was observed are indicated on the top and bottom of each column.

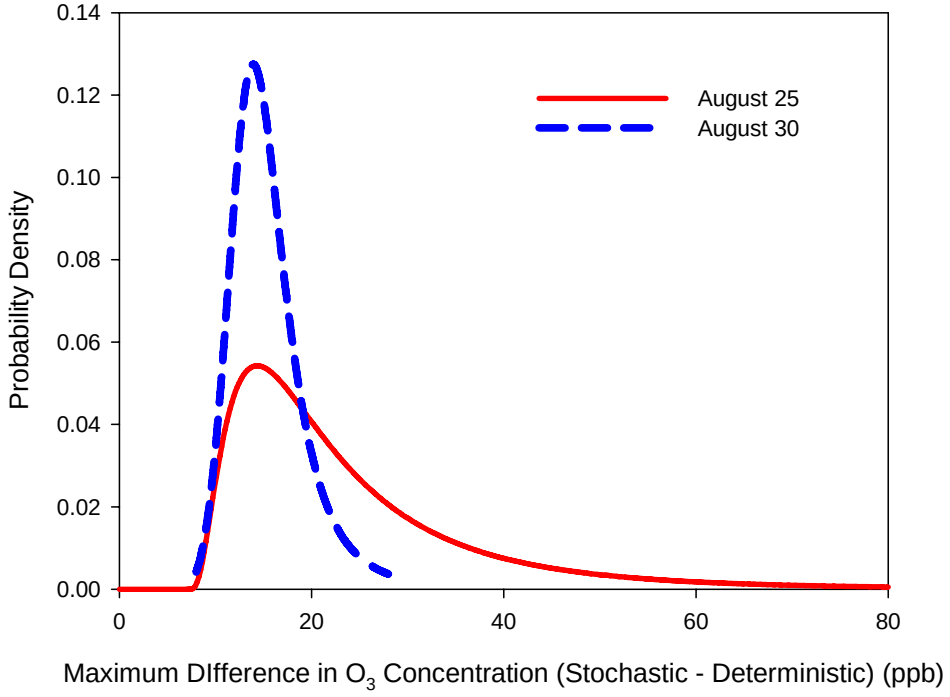

Figure 13. Probability distributions of maximum difference in ozone concentrations in one day simulations representing 25 August and 30 August, 2000.

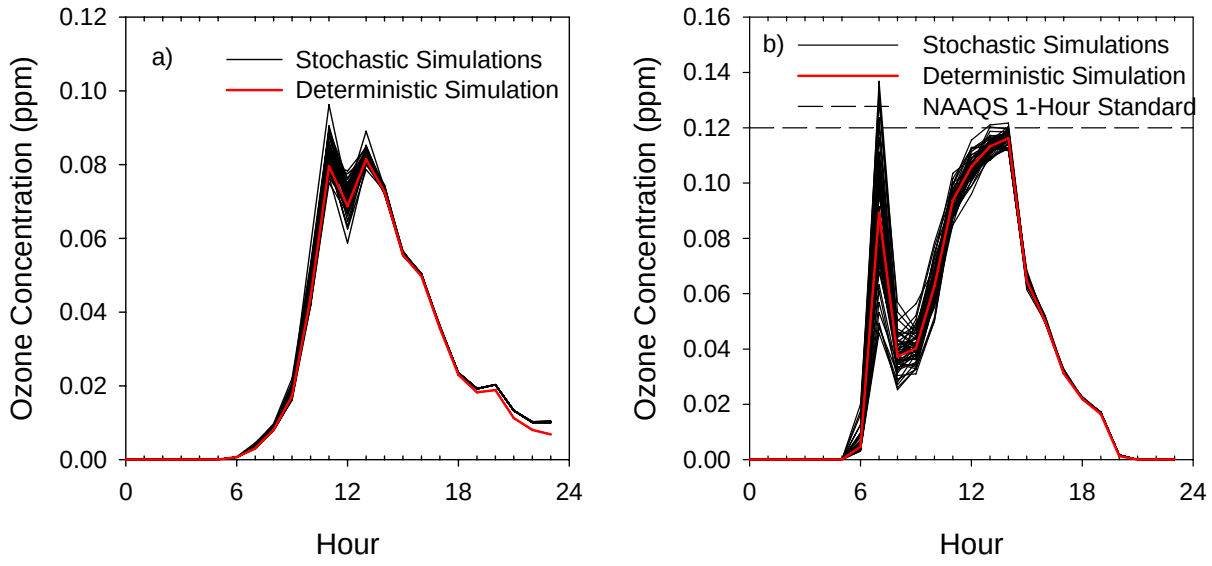

Figure 14. Hourly ozone concentrations from the 50 simulations for two grid cells on August 25: a) grid cell (40, 43), and b) grid cell (39, 55). Black lines show all 50 simulations with stochastic emissions inventory, and red line shows simulation with deterministic emissions inventory.