

COMPARISON OF MODEL-PREDICTED VOC CONCENTRATIONS WITH MEASUREMENTS

Prakash Doraiswamy^{1,†,*}, Christian Hogrefe^{1,2}, Winston Hao², Kevin Civerolo²,
Robert Henry², J.-Y. Ku² and Gopal Sistla²

¹Atmospheric Sciences Research Center, University at Albany, Albany, NY

²Bureau of Air Quality Analysis and Research (BAQAR), New York State Department of Environmental Conservation (NYSDEC), Albany, NY

[†]On assignment to BAQAR, NYSDEC

1. INTRODUCTION

Volatile organic compounds (VOCs) play an important role in ground-level ozone formation. The research group at the New York State Department of Environmental Conservation (NYSDEC) has been conducting daily air quality forecasting of ozone and fine particles (PM_{2.5}) using the Community Multi-scale Air Quality Model (CMAQ) on a pilot-scale, in collaboration with EPA and NOAA, since June 2005. As part of the ongoing evaluation of model performance for ozone and its precursors, this work presents a comparison of the archived model-predicted concentrations of selected VOCs from the forecast-based simulations against measured concentrations during June through August of 2005. It expands on the earlier work (Doraiswamy et al. 2007) that compared isoprene concentrations, and the application of the forecasting system as a diagnostic tool.

2. MODEL AND OBSERVATIONAL DATABASE

2.1. Model Set-Up and Archived Database

Archived model-predicted concentration fields from the air quality forecast simulations for the summer of 2005 were used in the analysis. A detailed description of the modeling system used during that time period is available elsewhere (Hogrefe et al. 2006; 2007). Briefly, it consisted of the ETA meteorological model, the PREMAQ (Otte et al. 2004; 2005) emissions and meteorology preprocessor, and the CMAQ (Byun and Ching 1999) photochemical model (v4.4), operated in a forecasting mode. The carbon-bond IV (CB4)

chemical mechanism (Gery et al. 1989) was applied. Biogenic emissions were estimated using the Biogenic Emissions Inventory System (BEIS3.12) incorporated into the PREMAQ processor. Each day, the model simulation was performed for 48 hours starting at 12:00 Greenwich Mean Time (GMT), initialized using modeled concentration fields from the previous day. Time-invariant clean boundary conditions were used for all days. Model predictions from June 1st to 11th were excluded from the analysis to eliminate the effect of initial conditions. The modeling domain covered almost the entire Eastern U.S with a 12 km horizontal grid resolution. The surface layer is ~35 m thick.

2.2. Observational Database

Hourly concentrations of various VOCs measured at the Photochemical Assessment Monitoring Stations (PAMS) were obtained from the EPA Air Quality System (AQS) for all monitors within the domain. This analysis presents comparisons at monitors within the New England and Mid-Atlantic State region comprising of Maine, Vermont, New Hampshire, New York, Massachusetts, Connecticut, Rhode Island, Pennsylvania, New Jersey, Maryland, Delaware and District of Columbia.

Table 1 lists the sites with hourly measurements that were used in this analysis.

3. DATA ANALYSIS, RESULTS AND DISCUSSION

Model predictions of the following CB4 species grouping (Gery et al. 1989) were extracted from the archived simulations: single-bonded one carbon surrogate PAR representing alkanes, the double-bonded two carbon surrogate OLE representing alkenes, seven-carbon aromatic hydrocarbon species TOL representing monoalkylbenzene structures such as toluene, the eight-carbon species XYL representing di- and tri-alkylbenzenes such as xylene, and the explicitly

*Corresponding author: Prakash Doraiswamy, BAQAR, New York State Department of Environmental Conservation, 625 Broadway, Albany, NY 12233-3259; Phone: (518) 402-8402; Fax: (518) 402-9035; Email: pdorai@gmail.com

treated species, two-carbon compound, ethylene (ETH) and the five-carbon compound, isoprene (ISOP). The CB-4 groupings FORM (formaldehyde) and ALD2 (acetaldehyde and higher aldehydes) were not considered because measurements of the constituent VOC species were not available at hourly time resolution at these PAMS sites. In order to compare model predictions with measurements, the measured concentrations of the various VOCs were grouped into the above CB4 groupings, as per the mapping procedure presented by Yarwood et al. (see Table 4-1 of Yarwood et al.(2003)). Table 2 summarizes the VOCs considered in the analysis. About 47 PAMS species are mapped to PAR, and hence are not elaborated in that table. All species concentrations were maintained in parts per billion carbon (ppbC). Thus, the model predictions were converted to ppbC using the respective number of carbon atoms indicated above for each surrogate. Measurements were not always available for all the compounds listed in Table 2 when attempting to group them into the appropriate CB4 classes. Hence, the mapped concentration was calculated utilizing the available measurements only. Thus, the CB4 grouping concentrations derived from measurements may be underestimated depending on the compounds, if any, that were missing. Diurnal profile comparisons are presented at selected sites.

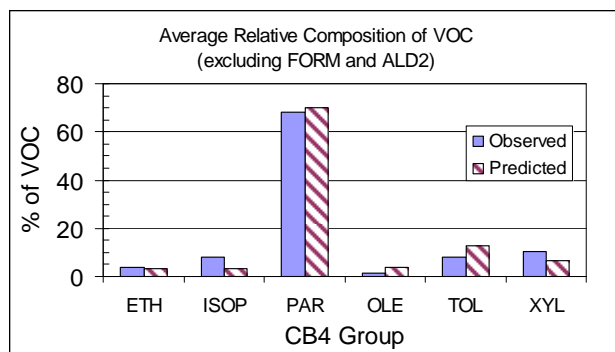


Figure 1. Relative Observed and Predicted Composition of Ambient VOC Aggregated in Terms of CB4 Classes Averaged over PAMS #2 Sites

Table 3 summarizes the average observed and predicted concentrations of the six CB4 surrogate species considered in this analysis. In general, with the exception of a few of these sites, the model typically over-predicted ETH, PAR and TOL concentrations by ~1.5 to 5 times, and OLE by 2 to more than 10 times. Observed OLE concentrations were low, often less than 1 ppbC. XYL was over-predicted except at 4 sites. ISOP concentrations were under-predicted at roughly half

of the sites in the northeast. Figure 1 shows the relative composition of VOC (without considering FORM and ALD2) at sites categorized as #2 (emissions impacted sites) in the PAMS network. The PAR constituted ~68% of the VOC, followed by XYL (~10%), ISOP (~8%), TOL (~8%), ETH (~4%) and OLE (~2%) based on measurements. Noticeable discrepancies in the modeled composition include the larger contribution of TOL (~13%) and the smaller contribution of ISOP (~3%) relative to observations.

In order to explore the observed and predicted diurnal profiles, measured and predicted hourly concentrations were averaged by hour over all days from June 12-Aug 31, 2005. Figure 2 presents the average observed and predicted diurnal profiles of each of the aforementioned six CB4 VOC species at the following three sites: an urban site in NY (360050083) located at NY Botanical Gardens (NYBG), Bronx, NY, a suburban site in NJ (340210005) located at Rider College (NJRC) and a rural site in PA (420010001), which is a NARSTO site located at Arendtsville, PA (PANARSTO). Table 4 lists the species for which measurements were not available at these sites. Please note the predicted concentrations are plotted on the right ordinate, while the measured concentrations are plotted on the left ordinate, as they often differ by a factor of two or more. The intent here is to examine the shape of the profile. Overall, the shape of the observed and predicted profiles was similar at the three sites, except for isoprene at the NYBG site. The typical diurnal profile for ETH, PAR, OLE, TOL and XYL consisted of a morning and an evening peak with a trough in the afternoon. The evening peak was more pronounced in the modeled profile. For isoprene, the predicted profile at NYBG showed an afternoon trough, while the measured profile showed a bell-shaped curve. One possible explanation is that the measurements are influenced by the localized isoprene emissions from the vegetation in the botanical gardens. On the other hand, the 12-km model grid covers a much larger urban area resulting in a lower average isoprene concentration across the grid which was insufficient to offset the dilution due to expansion of the boundary layer. At the NJRC and PANARSTO sites, the isoprene profiles showed a pronounced evening peak, which was reasonably replicated by the model. However the isoprene emissions were likely overestimated resulting in an over-prediction of isoprene concentrations at these two sites.

The afternoon trough noted in the profiles of other species is likely due to loss by photochemical reactions combined with expansion of the boundary layer. As mentioned above, the model typically over-predicted these species. It could be due to a combination of reasons: a) the VOC emissions may be overestimated in the model and/or b) the observations may not include all the compounds within the respective CB4 grouping. Although forecasted emissions segregated by source category were not available for this specific period, emission inventories utilized by NYSDEC in other modeling efforts were examined to determine the relative contributions of source categories that emit each species. As expected, nearly 100% of ISOP is emitted by biogenic sources. The county-wide emissions at an urban location such as Bronx County, NY (where NYBG is located) show that area sources contributed more than 50% of PAR, TOL and XYL emissions, while ETH was dominated by on-road (43%) and non-road (27%) mobile sources. At Mercer County, NJ (which includes NJRC), biogenic sources contributed ~30% or more to ETH, PAR and OLE emissions, while at the rural Adams County, PA (which covers PANARSTO) biogenic contribution to these species was more than 60%. More than 50% of XYL was emitted by area sources at both the suburban and rural counties.

Examining the profiles by site, the following features are evident:

- At the NYBG site, the model profile showed a sharper trough than observations, particularly for TOL, XYL and PAR, although higher in overall magnitude (except ISOP). The over-prediction is likely due to overestimated VOC emissions (except ISOP), particularly from area sources, as suggested by the large over-prediction in TOL and XYL, which have relatively larger contribution from area sources. In CB4 mechanism, the only reaction contributing to a loss of TOL and XYL is the reaction with the hydroxyl radical. This may point to a possible excess generation of OH. As shown in Doraiswamy et al. (2007), NO₂ concentrations were also over-predicted at this site, suggesting an overestimation of NO_x emissions. The combination of overestimated VOC and NO_x emissions could have led to an increased OH generation by increasing the OH chain length.
- At the NJRC site, similar over-predictions were found. In addition, the evening peak was pronounced in the model predictions, while the observations showed no such evening peak.

While the model was ~1.3 times higher than observations for XYL, it was ~2.5 times for PAR and ~3 times higher for ETH. The larger over-prediction for ETH, a species for which the area source contribution is lower than that for XYL, suggests that this overestimation is likely not from area sources, and hence could be attributed to mobile and biogenic sources. This, in combination with a drop in boundary layer height could explain the evening peak noted in model predictions.

- At the PANARTSO site, modeled isoprene profiles were similar to measured, although it was ~3.5 times higher. While the profiles for other species were similar in general, few minor discrepancies are noticeable. They include: a decrease in concentration of ETH, PAR, OLE, TOL and XYL between midnight and 5 am in the predicted profile, when the observed profile shows an increase; a nearly flat observed profile for OLE not noticed in predictions; and an increase in the evening concentration to a level comparable to morning peak, while the observed profile showed an increase to only half of the morning peak followed by a slight decrease. It is unclear if the nearly flat observed profile for OLE is realistic (likely representing an influence of biogenic emissions at this rural site) or if it is an artifact of instrument noise due to low concentrations. Although model over-predictions were found for ETH, PAR, OLE and TOL, XYL predictions were similar in magnitude to observed concentrations. XYL is primarily emitted from anthropogenic sources, suggesting that the primary VOC emissions, particularly from area sources, are reasonable at this site. The over-predictions of other species are likely resulting from overestimated biogenic emissions, which were the largest contributor to each of those species.

4. CONCLUSIONS

This study compared hourly predictions of volatile organic compounds grouped by CB4 classes. Overall, the model appeared to track the diurnal profile at most sites. Although the model over-predicted concentrations of most species, the relative distribution of these species appeared to be reasonable, except for the lower ISOP contributions in the predictions. The analyses revealed possible overestimation of VOC and NO_x emissions at the NYBG site. In addition, it appears that biogenic VOC emissions are overestimated at the NJRC and PANARSTO sites.

5. DISCLAIMER

This work was funded in part by the U.S. EPA under Cooperative Agreement CR83228001 and the New York State Energy Research and Development Authority (NYSERDA) under agreement #10599. In addition, Prakash Doraiswamy was partly supported by the Air Pollution Educational and Research Grant (APERG) funded by the Mid-Atlantic States Section of the Air and Waste Management Association (MASS-A&WMA). The results presented here have not been reviewed by the funding agencies. The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of NYSDEC or those of the sponsoring agencies.

6. REFERENCES

- Byun, D. W. and J. K. S. Ching (1999). Science algorithms of the EPA MODELS-3 community multiscale air quality (CMAQ) modeling system. Washington, DC, Office of Research and Development, U.S. Environmental Protection Agency.
- Doraiswamy, P., C. Hogrefe, W. Hao, K. Civerolo, J.-Y. Ku, R. Henry and G. Sistla (2007). Use of Air Quality Forecasting as a Diagnostic Tool over the Northeastern US. 6th Annual Community Modeling and Analysis System (CMAS) Conference. Chapel Hill, NC.
- Gery, M. W., G. Z. Whitten, J. P. Killus and M. C. Dodge (1989). "A photochemical kinetics mechanism for urban and regional scale computer modeling." Journal of Geophysical Research **94**: 12925-12956.
- Hogrefe, C., W. Hao, K. Civerolo, J.-Y. Ku and G. Sistla (2006). Exploring Approaches to Integrate Observations and CMAQ Simulations for Improved Air Quality Forecasts. Models-3 Users' Workshop. Chapel Hill, NC.
- Hogrefe, C., W. Hao, K. Civerolo, J.-Y. Ku, G. Sistla, R. S. Gaza, L. Sedefian, K. Schere, A. Gilliland and R. Mathur (2007). "Daily Simulation of Ozone and Fine Particulates over New York State: Findings and Challenges." Journal of Applied Meteorology **46**(NOAA/EPA Golden Jubilee Special Issue): 961-979.
- Otte, T. L., G. Pouliot and J. E. Pleim (2004). PREMAQ: A New Pre-Processor to CMAQ for Air Quality Forecasting. Models-3 Conference. Chapel Hill, NC.
- Otte, T. L., G. Pouliot, J. E. Pleim, J. O. Young, K. L. Schere, D. C. Wong, P. C. S. Lee, M. Tsidulko, J. T. McQueen, P. Davidson, R. Mathur, H. Y. Chuang, G. DiMego and N. L. Seaman (2005). "Linking the Eta Model with the Community Multiscale Air Quality (CMAQ) Modeling System to Build a National Air Quality Forecasting System." Weather and Forecasting **20** (3): 367-384.
- Yarwood, G., C. Tran, S. Lau and E. Fujita (2003). Impact of Updates to On-Road Mobile Source Emission Factor Models (EMFAC) for the Los Angeles Region: Ozone Model Sensitivity and Ambient/Inventory Reconciliation: Final Report Novato, CA, Environ International Corporation: 128.

Table 1. Sites Used in Model Evaluation

State	County	Site ID	Latitude (degrees)	Longitude (degrees)	Site Description (PAMS Site Type)	Land Use	Location Setting
Connecticut	Fairfield	090019003	41.1183	-73.3367	Sherwood Island State Park (#3/#1)	Forest	Rural
Connecticut	Hartford	090031003	41.7847	-72.6317	McAuliffe Park (#2)	Residential	Suburban
Connecticut	New Haven	090090027	41.3011	-72.9028	1, James Street (#2)	Commercial	Urban city center
<i>District of Columbia</i>		110010043	38.9189	-77.0125	S.E. End McMillian Reservoir (#2)	Commercial	Urban city center
Maine	Cumberland	230052003	43.5608	-70.2078	Two Lights State Park (#4)	Residential	Rural
Maine	Hancock	230090102	44.3517	-68.2272	Top of Cadillac Mountain (#4)	Mobile	Rural
Maine	York	230313002	43.0833	-70.7500	Frisbee School, Goodsoe Rd (#2)	Residential	Suburban
Maryland	Baltimore	240053001	39.3108	-76.4744	Woodward and Franklin Roads, Essex (#2)	Residential	Suburban
Massachusetts	Essex	250092006	42.4744	-70.9725	390, Parkland (#2)	Commercial	Urban city center
Massachusetts	Essex	250094004	42.7894	-70.8092	Sunset Blvd (#3)	Residential	Suburban
Massachusetts	Hampden	250130008	42.1945	-72.5557	Anderson Rd, AFB (#2)	Commercial	Suburban
Massachusetts	Hampshire	250154002	42.2983	-72.3347	Quabbin Summit (#3)	Forest	Rural
New Hampshire	Hillsborough	330111011	42.7204	-71.5231	Gilson Road (#1)	Residential	Suburban
New Jersey	Camden	340070003	39.9228	-75.0972	Copewood & E. Davis Streets, Trailer (#2A)	Residential	Suburban
New Jersey	Mercer	340210005	40.2828	-74.7467	Rider College, Lawrence Township (#3)	Residential	Suburban
New Jersey	Middlesex	340230011	40.4619	-74.4298	R.U. Veg Research Farm, 3 Ryders Ln (#1)	Agricultural	Rural
New York	Bronx	360050083	40.8659	-73.8808	200th Street and Southern Blvd (#2)	Commercial	Urban city center
Pennsylvania	Adams	420010001	39.9200	-77.3100	NARSTO Site, Arendtsville (#1)	Residential	Rural

Table 2. VOC Considered for Analysis

CB4 Group	Compounds considered, if available in measurements*	Compounds not considered
ETH	Ethylene	None
ISOP	Isoprene	None
PAR	47 PAMS compounds mapped to PAR in Table 4-1 of Yarwood et al. (2003)	Non-PAMS compounds and carbonyl compounds in Table 4-1 of Yarwood et al. (2003), which includes higher aldehydes, ketones, isomers of methyl butene, methyl pentene, dimethyl hexane etc.
OLE	Propene, 1-butene, 1-pentene, 2-methyl-1-pentene, styrene	i-butene, 1,3-butadiene, 3-methyl-1-butene, 2-methyl-1-butene, 4-methyl-1-pentene, 1-hexene, 2,4,4-trimethyl-1-pentene
TOL	Toluene, ethyl benzene, styrene, isopropyl benzene, n-propyl benzene,	None
XYL	m/p-xylene, o-xylene, m-ethyl toluene, p-ethyl toluene, 1,3,5-trimethyl benzene, o-ethyl toluene, 1,2,4-trimethyl benzene, 1,2,3-trimethyl benzene, m-diethyl benzene, p-diethyl benzene	None

* see Table 4-1 of Yarwood et al. (2003) for mapping factors

Table 3. Average Hourly Concentration of CB4 VOC Classes (June 12 – August 31, 2005)*

StCySite	Count	ETH (ppbC)		ISOP (ppbC)		PAR (ppbC)		OLE (ppbC)		TOL (ppbC)		XYL (ppbC)	
		Obs	Pred	Obs	Pred	Obs	Pred	Obs	Pred	Obs	Pred	Obs	Pred
090019003	1616	0.99	1.97	2.39	1.15	14.84	48.02	0.44	2.09	2.38	7.58	1.66	4.61
090031003	490-1738	1.50	2.51	2.48	1.32	30.59	65.44	0.89	3.78	5.76	13.22	7.22	7.51
090090027	1865	1.77	1.79	1.30	1.42	98.60	44.89	1.42	2.06	10.77	7.44	10.75	4.02
110010043	1518-1523	1.33	3.44	2.01	4.80	26.11	87.22	0.44	4.29	0.16	14.30	1.75	8.16
230052003	1831	0.28	1.50	0.92	1.09	8.25	41.23	0.17	2.84	0.84	7.16	0.76	6.73
230090102	1630	0.04	0.59	1.56	0.53	3.04	17.25	0.05	1.67	0.44	1.19	1.21	0.47
230313002	1602	0.80	1.84	2.17	1.91	14.10	37.15	0.45	3.09	1.85	5.88	3.10	3.12
240053001	803	1.84	1.80	2.15	2.62	51.87	43.09	0.95	2.29	5.42	6.09	7.53	3.37
250092006	1454	1.22	2.41	5.00	2.34	20.20	56.07	0.74	3.10	3.38	14.17	3.51	7.29
250094004	1606	0.70	1.06	2.45	0.98	11.17	28.49	0.45	1.50	1.62	4.48	1.72	2.06
250130008	1453-1625	1.34	2.27	4.09	2.23	17.23	49.09	0.37	3.41	2.89	9.60	3.74	4.57
250154002	824	0.45	1.47	9.94	2.63	8.21	34.58	0.29	2.75	1.26	4.94	1.53	2.05
330111011	1512-1515	0.55	2.31	4.43	3.82	13.98	45.15	0.43	3.76	2.40	7.40	5.21	3.66
340070003	1430-1810	1.54	4.21	1.98	5.90	46.27	75.31	1.43	3.78	5.44	12.57	4.00	6.52
340210005	1788-1796	0.92	2.60	2.19	3.48	17.96	55.08	0.55	2.90	2.44	8.55	2.47	4.18
340230011	273-1584	2.48	3.67	3.91	5.70	19.43	58.62	0.93	3.45	4.20	9.37	2.82	4.68
360050083	1526	2.74	5.29	3.84	1.50	39.03	149.52	0.42	4.77	6.19	32.10	8.28	18.00
420010001	1720-1724	0.61	1.45	2.39	6.85	9.53	35.21	0.41	2.48	1.04	3.35	1.17	1.27

* Note that there were a total of 1944 possible hours. Sites with number of pairs less than 1458 imply fewer than 75% of total possible dataset. Number of pairs varied between species groupings, and values in italics indicate those with pairs less than half the maximum count at that site.

Table 4. Missing VOC Species (among those considered) at the three sites

CB4 Group	SI: 360050083	SI: 340210005	SI: 420010001
OLE	Propene, 2-methyl-1-pentene	2-methyl-1-pentene	2-methyl-1-pentene
TOL	None	None	None
XYL	1,2,4-trimethyl benzene, 1,2,3-trimethyl benzene	m-ethyl toluene, p-ethyl toluene	None
PAR	Propene, n-undecane, 1,2,4-trimethyl benzene, 1,2,3-trimethyl benzene	2-methyl-1-pentene, m-ethyl toluene, p-ethyl toluene	2-methyl-1-pentene

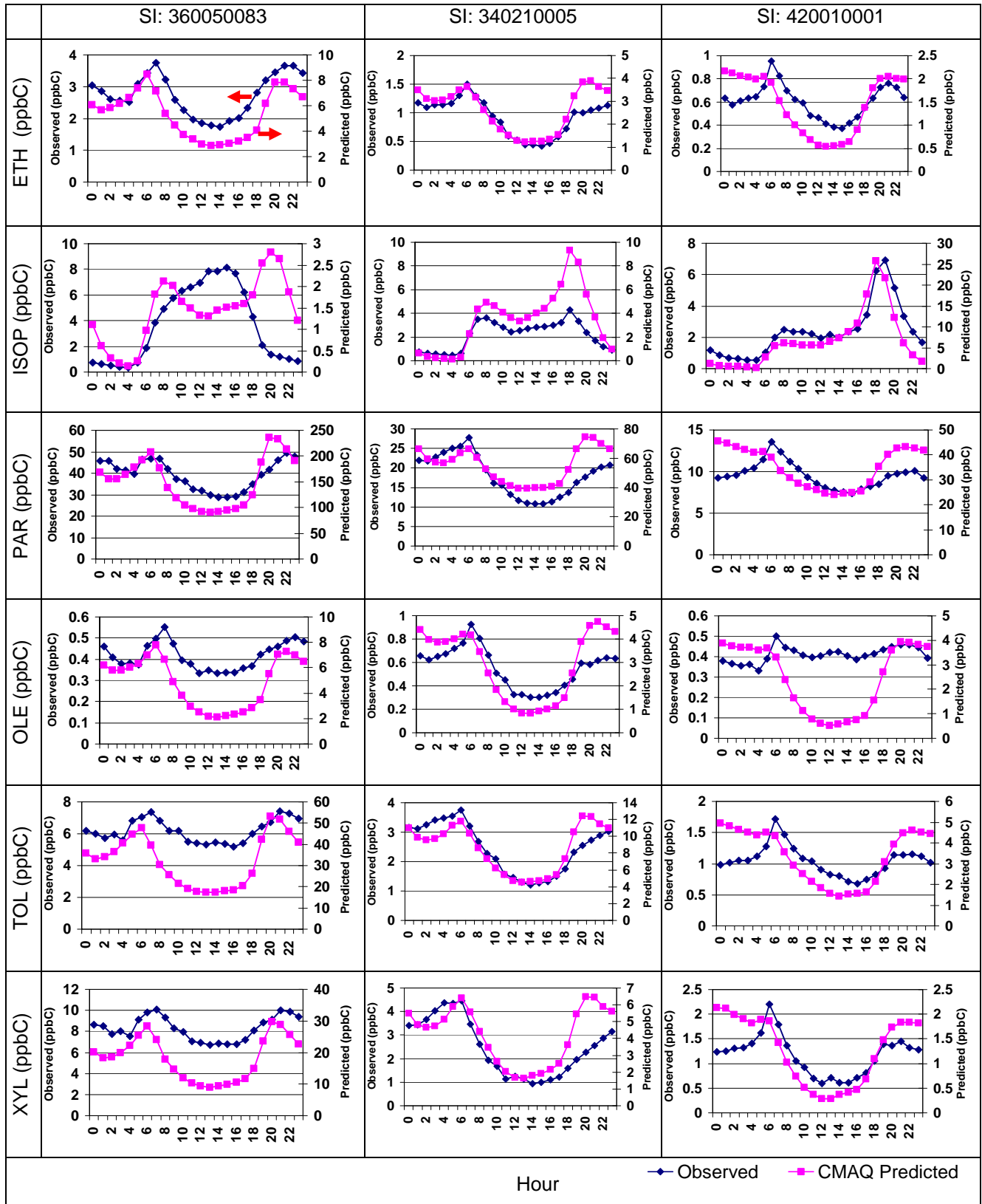


Figure 2. Average Diurnal Profile of CB4 VOC Classes during June 12-August 31, 2005. Note that the observed concentration is plotted on the left ordinate, while the predicted concentration is plotted on the right ordinate (often on different scales), as directed by the arrows in the top left panel.