

5. Comparing Modeled to Monitored Data: Evaluation of Confidence in Ambient Modeled Data

5.1. Overview

The 1999 monitored values made it possible for DEQ to compare the initial set of modeled concentration values predicted by the CALPUFF measurements from the same year. This step was extremely valuable in highlighting deficiencies in the data inputs or in the model. The model to monitor comparison resulted in reassessment and in some cases, revision of some emissions estimates prior to the final model runs. The primary objective in doing the monitor to model comparison was to determine if model estimates could be used reliably as an indication of air toxics concentrations throughout the Portland metropolitan area.

5.1.1 Graphs showing model to monitor comparisons

For each pollutant (except acrolein and diesel PM which were not monitored) a graph was constructed to show how the predicted model concentration compared to measured values at each of the five monitoring locations. Figure 5.1 below shows an example of a model to monitor comparison. The measured values are presented on the graph as a set of three horizontal lines. The middle line represents the annual average of the measured concentrations. This is calculated by first replacing all concentrations reported as below the minimum quantitation level (MQL) with half the MQL, for each 24-hour sample, and then averaging all 24-hour sample values over the period that the samples were measured. The number of 24-hour samples for any given pollutant ranged from 11 to 56. For many of the pollutants in this study, a substantial number of the measurement values were below the MQL. These included 1, 3-butadiene, arsenic, chloroform, chromium, nickel, and tetrachloroethylene. The influence of the MQL on the comparisons is discussed in the pollutant-specific sections below.

The top and bottom horizontal lines represent the upper and lower bounds, respectively, of the 95% confidence interval for the measured annual average concentrations. They are calculated as follows:

1. Average all the 24-hour sample values, after replacing the concentrations reported as below the MQL with zero for the lower bound and with the MQL for the upper bound: X_0 and X_1 .
2. Calculate the standard deviations, S_0 , $S_{0.5}$, and S_1 , corresponding to substitution for values reported below the MQL of 0, half MQL, and MQL, respectively. Select the highest (S_{MAX})
3. Calculate the bounds as

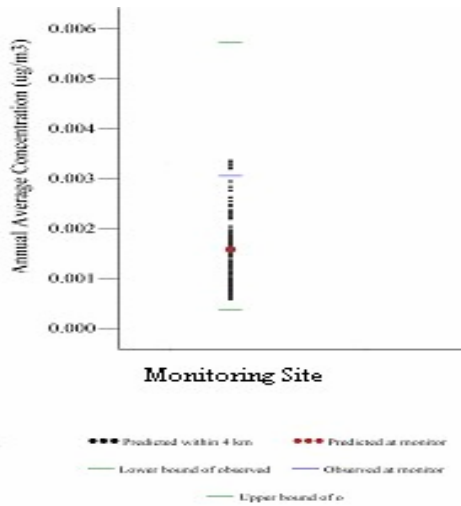
$$\text{Lower bound} = X_0 - t * S_{MAX} / \text{sqrt}(N)$$

$$\text{Upper bound} = X_1 + t * S_{MAX} / \text{sqrt}(N)$$

where t = the t-statistic = 1.96 to 2.23, depending on the number of samples, and N = the number of samples = 11 to 56.

Thus, the 95% confidence interval of the monitored annual average increases as the fraction of samples below the MQL increases, and as the standard deviation (i.e., variability) of the sample values increases.

Figure 5.1 Example of model to monitor comparison



The large red dot in the graph represents the annual average concentration model predictions at the exact location of the monitor. However, because the monitors are considered neighborhood-scale, according to US EPA monitoring guidelines the measurements are expected to be representative for a distance of 0.5 to 4 kilometers. Therefore, for comparison purposes, all the receptor concentration model predictions within 4 kilometers of the monitoring location are also indicated on the figure with small black dots.

5.1.2 Adjustment for the Downtown monitor height

Because the modeled concentrations from CALPUFF are predicted at ground level, an adjustment was made to the measured concentrations at the Downtown monitor, which was on the roof of a 15 meter building. Because the bulk of emissions in the PATA domain are released at ground level, ground level concentrations would be expected to be higher than those above ground. EPA's SCREEN3 air dispersion model was used to estimate the ratio between concentrations at ground level and concentrations at 15 meters elevation for a ground level release. The Downtown monitor measurements were multiplied by 1.37 for the model to monitor comparison to account for this difference in height.

5.2 Data summary

Table 5.1 presents the annual average concentrations from the measurements, the model predictions, and the model to monitor ratios.

Table 5.1 Annual average concentrations ($\mu\text{g}/\text{m}^3$) derived from measurements and predicted by CALPUFF for 5 locations in the PATA modeling domain.

Pollutant	Site	Monitored Concentration			CALPUFF Prediction			Model-to-Monitor Ratios*	
		Mean	Std Error**	% <MQL	At Site	Within 4 Km of Site		Mean	Std Dev
						Mean	Std Dev		
1,3-Butadiene	Beaverton	0.11	0.06	97%	0.31	0.30	0.08	2.6	1.6
	Downtown	0.19	0.08	85%	0.84	0.35	0.16	1.8	1.1
	N Roselawn	0.14	0.06	84%	0.23	0.29	0.11	2.0	1.1
	NW Post Office	0.14	0.06	84%	0.37	0.38	0.17	2.6	1.6
	SE Lafayette	0.14	0.05	79%	0.23	0.27	0.09	1.8	0.9
Acetaldehyde***	Beaverton	1.9	0.1	0%	0.7	0.6	0.2	0.3	0.1
	Downtown	3.0	0.2	0%	2.2	1.0	0.6	0.3	0.2
	N Roselawn	2.4	0.2	0%	0.5	0.6	0.5	0.3	0.2
	NW Post Office	2.1	0.1	0%	0.7	1.1	0.7	0.5	0.3
	SE Lafayette	2.0	0.1	0%	0.5	0.5	0.2	0.3	0.1
Arsenic	Beaverton	0.0021	0.0023	90%	0.0003	0.0002	0.0001	0.1	0.1
	Downtown	0.0018	0.0012	100%	0.0007	0.0003	0.0002	0.2	0.2
	N Roselawn	0.0023	0.0012	93%	0.0002	0.0002	0.0001	0.1	0.1
	NW Post Office	0.0015	0.0008	89%	0.0004	0.0004	0.0002	0.2	0.2
	SE Lafayette	0.0029	0.0030	75%	0.0002	0.0002	0.0001	0.1	0.1
Benzene	Beaverton	1.4	0.1	0%	6.1	5.1	1.1	3.7	0.8
	Downtown	2.6	0.2	0%	7.7	4.7	1.4	1.8	0.6
	N Roselawn	2.1	0.2	0%	4.0	4.1	1.1	1.9	0.6
	NW Post Office	1.9	0.1	0%	4.7	5.1	1.4	2.6	0.7
	SE Lafayette	2.5	0.2	0%	3.7	4.5	1.0	1.8	0.4
Benzene (June-August)	Beaverton	1.1	0.1	0%	2.2	2.2	0.5	1.9	0.5
	Downtown	1.9	0.1	0%	5.4	2.7	1.0	1.4	0.5
	N Roselawn	1.4	0.1	0%	1.9	2.1	0.7	1.4	0.5
	NW Post Office	1.4	0.1	0%	2.8	2.7	1.0	1.9	0.7
	SE Lafayette	1.5	0.1	0%	2.1	2.3	0.6	1.5	0.4
Chloroform	Beaverton	0.24	0.14	100%	0.09	0.09	0.00	0.4	0.2
	Downtown	0.34	0.20	100%	0.09	0.09	0.00	0.3	0.1
	N Roselawn	0.25	0.14	100%	0.09	0.09	0.00	0.3	0.2
	NW Post Office	0.26	0.14	98%	0.09	0.09	0.00	0.3	0.2
	SE Lafayette	0.24	0.14	100%	0.09	0.09	0.00	0.4	0.2
Chromium	Beaverton	0.0007	0.0005	100%	0.0080	0.0068	0.0028	9.9	7.7
	Downtown	0.0017	0.0007	75%	0.0045	0.0033	0.0015	1.9	1.1
	N Roselawn	0.0014	0.0004	69%	0.0020	0.0023	0.0011	1.6	0.9
	NW Post Office	0.0024	0.0005	56%	0.0063	0.0036	0.0015	1.5	0.7
	SE Lafayette	0.0016	0.0005	64%	0.0016	0.0022	0.0008	1.3	0.6

Pollutant	Site	Monitored Concentration			CALPUFF Prediction			Model-Monitor Ratios*	
		Mean	Std Error**	% < MQL	At Site	Within 4 km of site		Mean	Std Dev
						Mean	Std Dev		
Formaldehyde***	Beaverton	1.9	0.1	0%	1.5	1.5	0.3	0.7	0.2
	Downtown	3.8	0.3	0%	4.7	2.2	1.1	0.6	0.3
	N Roselawn	3.2	0.5	0%	1.4	1.7	0.7	0.5	0.2
	NW Post Office	2.5	0.2	0%	1.6	2.4	1.2	0.9	0.5
	SE Lafayette	2.4	0.1	0%	1.4	1.5	0.3	0.6	0.1
Nickel	Beaverton	0.0009	0.0004	93%	0.0018	0.0017	0.0009	1.9	1.4
	Downtown	0.0032	0.0010	42%	0.0017	0.0015	0.0007	0.5	0.3
	N Roselawn	0.0021	0.0004	48%	0.0010	0.0013	0.0006	0.6	0.3
	NW Post Office	0.0044	0.0006	26%	0.0033	0.0017	0.0007	0.4	0.2
	SE Lafayette	0.0027	0.0006	42%	0.0008	0.0010	0.0004	0.4	0.2
POM (as 16-PAH)	Beaverton	0.022	0.003	n/a****	0.889	0.455	0.196	20.7	9.4
	Downtown	0.076	0.008	n/a****	0.528	0.459	0.191	6.1	2.6
	N Roselawn	0.042	0.005	n/a****	0.611	0.467	0.194	11.1	4.8
	NW Post Office	0.068	0.009	n/a****	0.195	0.455	0.228	6.6	3.4
	SE Lafayette	0.036	0.004	n/a****	0.434	0.477	0.134	13.2	4.0
POM (as 16-PAH) (June-August)	Beaverton	0.012	0.003	n/a****	0.014	0.012	0.003	1.0	0.3
	Downtown	0.061	0.010	n/a****	0.115	0.020	0.019	0.3	0.3
	N Roselawn	0.024	0.007	n/a****	0.010	0.014	0.008	0.6	0.4
	NW Post Office	0.055	0.007	n/a****	0.016	0.022	0.022	0.4	0.4
	SE Lafayette	0.023	0.004	n/a****	0.010	0.011	0.004	0.5	0.2
Tetrachloro-ethylene	Beaverton	0.34	0.19	98%	0.22	0.22	0.12	0.6	0.5
	Downtown	0.53	0.26	93%	0.19	0.21	0.02	0.4	0.2
	N Roselawn	0.38	0.18	92%	0.20	0.20	0.02	0.5	0.3
	NW Post Office	2.29	0.33	28%	0.29	0.20	0.03	0.1	0.0
	SE Lafayette	0.46	0.20	91%	0.23	0.24	0.03	0.5	0.2

- Ratios of CALPUFF annual average concentration from receptors within 4 km of the monitoring site to the average monitored concentration.

** When there are no concentration measurements below the MQL, the standard error of the mean is calculated as

$$\text{std error} = \text{std dev} / \sqrt{N}.$$

When there are concentration measurements below the MQL, 3 annual averages and std deviations are calculated.

- X0 is the annual average, and S0 is the std dev, assuming all measurements below the MQL = 0.
- X0.5 is the annual average, and S0.5 is the std dev, assuming all measurements below the MQL = 0.5 * MQL.
- X1 is the annual average, and S1 is the std dev, assuming all measurements below the MQL = MQL.

In that case the mean is specified as X0.5, and the std error is approximated as

$$\text{std error} = \sqrt{\frac{\max^2(S0, S0.5, S1)}{N} + \frac{(X1 - X0)^2}{12}}$$

The second term on the right-hand side represents the variance of a uniform distribution between X1 and X0, and therefore accounts for the uncertainty due to the concentration measurements below the MQL.

*** CALPUFF predictions for acetaldehyde include the primary component only, and for formaldehyde include the primary and background components only.

**** The number of concentration measurements below the MQL varies among the 16 PAH constituents of POM.

5.3 Model Performance

According to US EPA (2001), studies of the performance of long-term air quality models suggest that 90 percent of the estimated concentrations should be within a factor of two of those observed. This should be the case if the emissions are well characterized, chemical reactions that may form or remove pollutants do not occur, and the meteorological data are representative. Another factor to consider in performance evaluation, as seen in the PATA study, is that when measured pollutant concentrations are frequently below the detection limit, the large confidence intervals in the measured annual means make true comparisons difficult. In general, the results indicate that the CALPUFF model did a very credible job of predicting ambient concentrations where there were measurements for comparison.

5.3.1 Metals

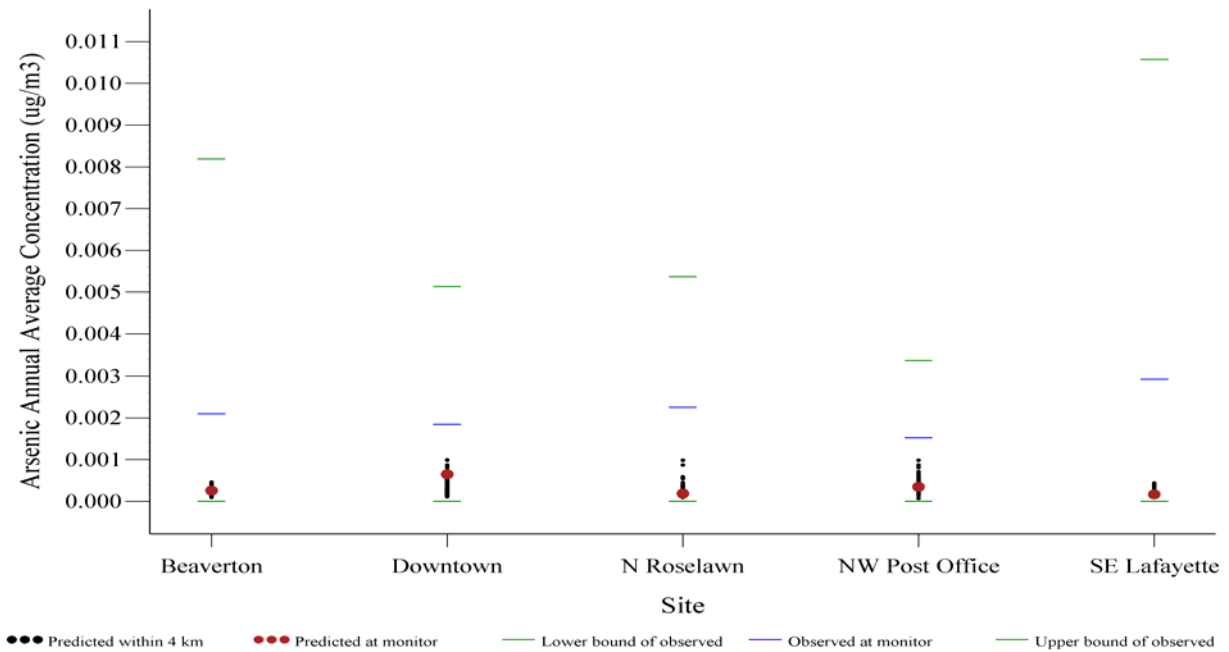
Arsenic, chromium, and nickel were the metals included in this study. One common factor for all the metals is that the modeling is based on an inventory of metals emissions with no specification for the size of the metal-containing particles. Monitoring, on the other hand, was restricted to only the PM_{2.5} size fraction of the total ambient particulate. In all cases, this should result in the model over-predicting the measured values, although the amount of over-prediction is not known.

However, the generally good agreement between model estimates and measured ambient values suggests that with some improvements in the emissions inventory, modeling should be a very effective way to predict ambient metals concentrations.

5.3.1.1 Arsenic

The model to monitor comparison of arsenic predicts consistently very low levels all across the city. The largest emission sources of arsenic in the PATA inventory were combustion of distillate oil and natural gas used for heating homes and businesses, which would be expected to be widespread. Annual average measured arsenic ranged from about 0.0015 to 0.0030 µg/m³. This data is graphically presented in Figure 5.2.

Figure 5.2 Model to monitor comparisons for arsenic



The wide 95% confidence intervals around the monitored annual averages in the figure indicate that the bulk of measured arsenic concentrations, ranging from 70% to 100%, were below the MQL at all monitors. The figure shows that the model predictions are fairly consistent with the measurements. In all cases, the range of annual average predictions within 4 kilometers of the monitor location fell within the wide 95% confidence interval of the average of the measured values. Table 5.1 shows that for all the monitor locations, the ratio of the mean concentration within 4 kilometers to the mean measured concentration ranged from 0.07 to 0.24. Given the uncertainties in the inventory and the measured values, the true ratios may be higher or lower.

5.3.1.2 Total Chromium

Agreement between model predictions and measurements was not quite as good for chromium, where the model tended to over-predict slightly. However at the Beaverton site, the model greatly over-predicts. The largest emission source of chromium in the PATA inventory was surface coating of plastic parts. It appears that the spatial allocation of these emissions resulted in an unlikely high emissions density for this pollutant in the Beaverton area. Figure 5.3 shows the comparisons for chromium.

As a result of a correction in chromium emissions from surface coating facilities, the modeled concentrations were adjusted in the area just west of State Highway 217 in Beaverton, after the initial CALPUFF modeling. The surface coating facilities are located between 1.2 and 1.5 miles from the Beaverton toxics monitor. If the change in modeled concentrations were reflected in the model to monitor graph (Figure 5.3), the modeled values would be somewhat lower and in better accord with the monitored value at the Beaverton site.

Figure 5.3 Model to monitor comparison for total chromium

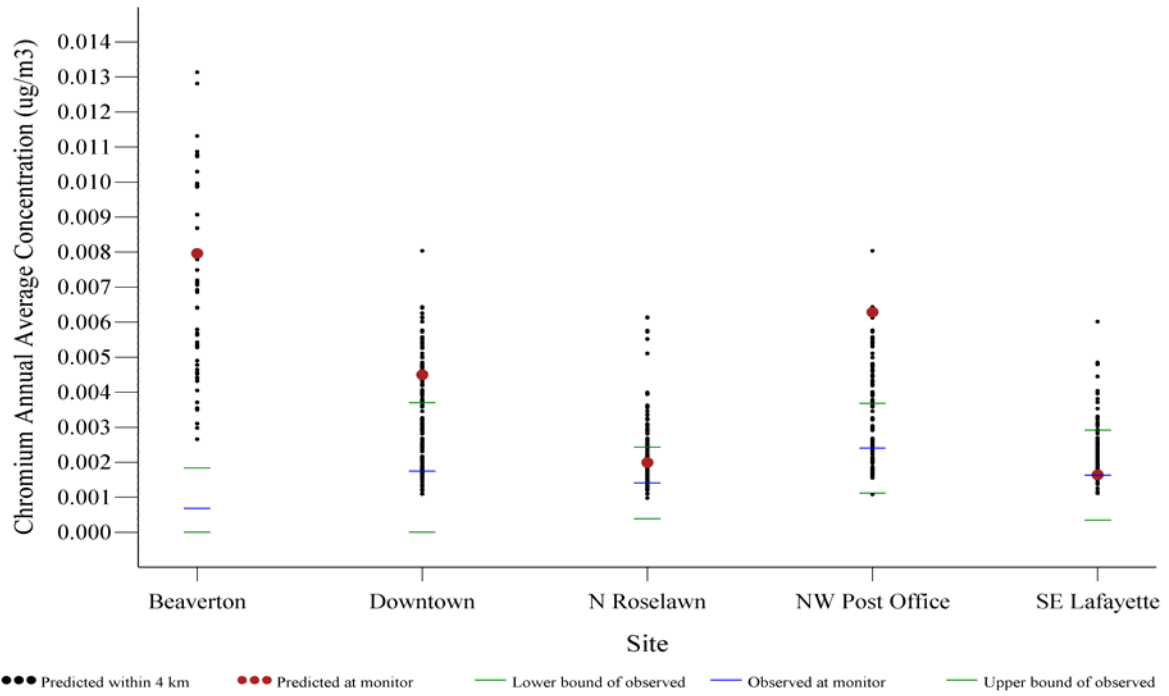
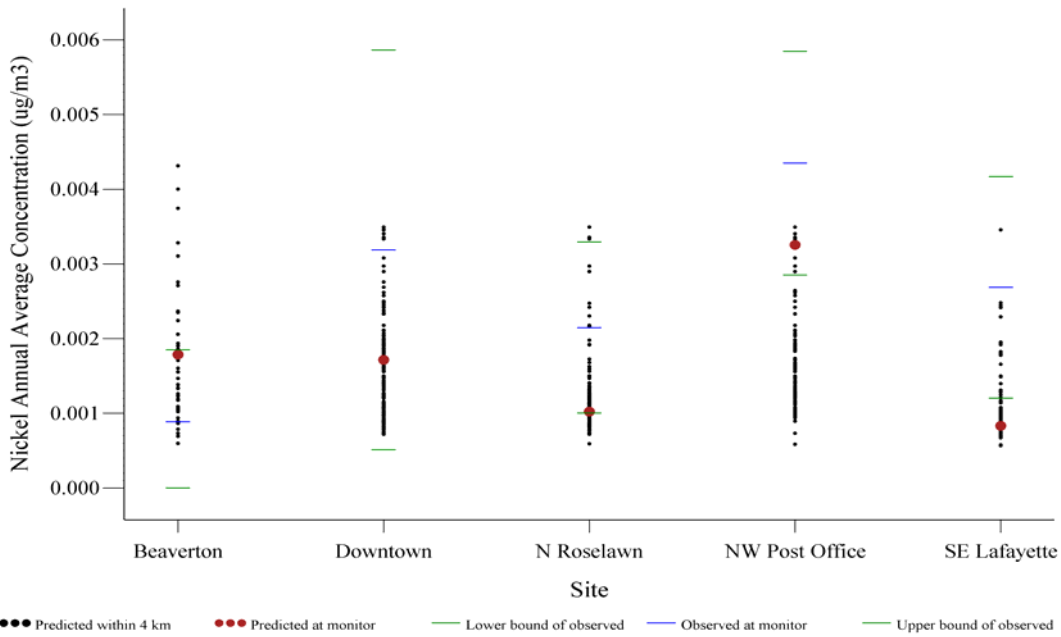


Figure 5.4 Model to monitor comparison for nickel



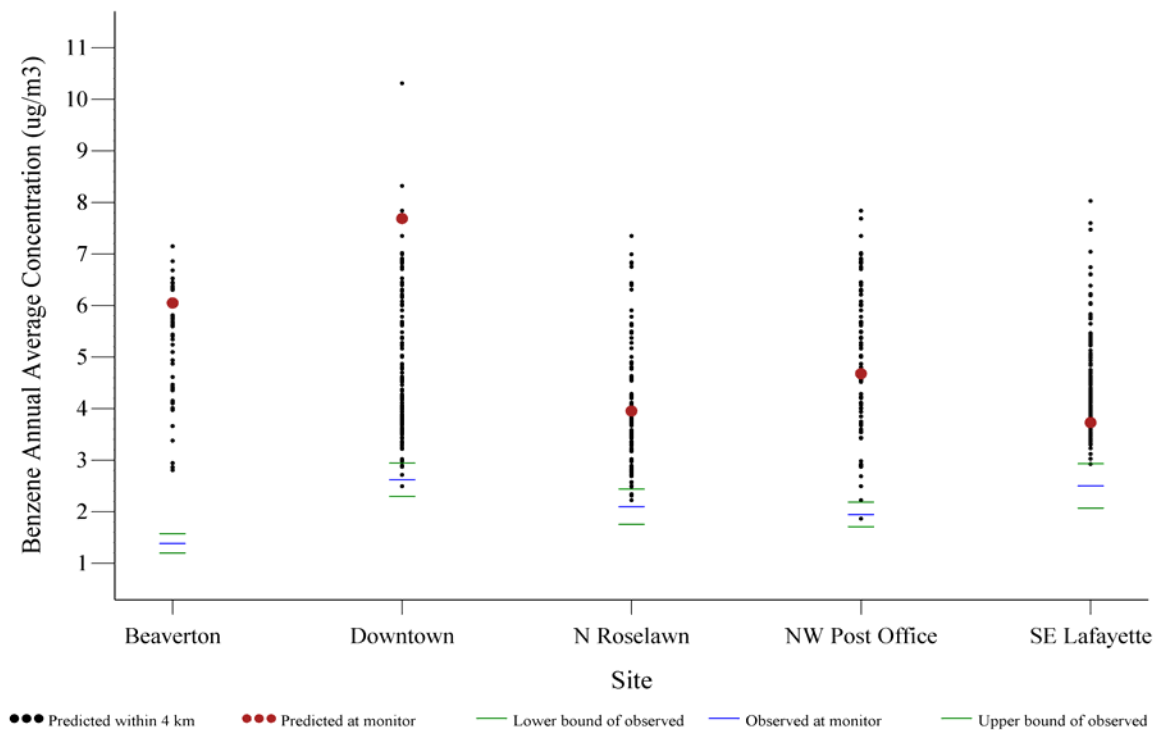
5.3.2 Volatile Organic Compounds

Several volatile organic compounds (VOC) were both modeled in PATA and measured in 1999. They were benzene, 1,3-butadiene, and perchloroethylene.

5.3.2.1 Benzene

Benzene provides an example where measurements have a great deal of certainty and model predictions are clearly overestimated. Measured annual averages ranged from 1.4 $\mu\text{g}/\text{m}^3$ in Beaverton to 2.6 $\mu\text{g}/\text{m}^3$ Downtown. Figure 5.5 shows for benzene how the model predictions at the sites, and within the 4 kilometer radius, compare to the measured annual averages, with their associated 95% confidence intervals. Table 5.1 also gives the ratios for each of the sites, ranging from 1.8 to 3.7. Both show a substantial over-prediction by the model.

Figure 5.5 Model to monitor comparison for benzene



The modeling runs provided other data that helps to explain the comparison results. At the monitor locations about 30% to 70% of the total concentration is predicted to come from on-road mobile sources. Another 10% to 40% is contributed by area sources. The remainder is a combination of non-road mobile, special non-road mobile, and background. Because on-road mobile emissions of other HAPs, such as 1,3-butadiene and primary acetaldehyde, do not appear to be significantly overestimated it seemed unlikely that the on-road mobile contribution was the cause of the discrepancy. On the other hand, POM concentrations also appear to be over-predicted and POM concentrations are dominated by residential wood combustion emissions. This led to examination of the wood combustion category.

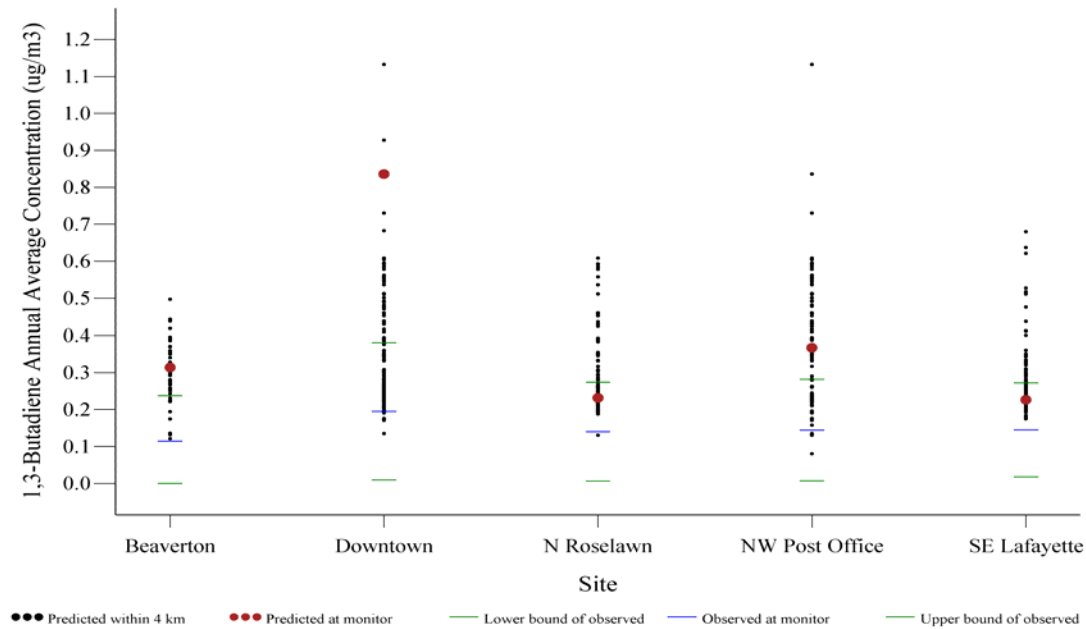
Since no residential wood combustion emissions were temporally allocated to the third quarter of the year (June-August), DEQ examined the model performance for benzene during this period when residential wood combustion contributions were absent. The

results of this analysis for the third quarter indicate much improved model performance. Table 5.1 shows that for the third quarter, the ratios of the mean predicted concentration within 4 kilometers to the mean measured concentration now ranged from 1.4 to 1.9. These findings suggest that benzene emissions from residential wood combustion are likely to be overestimated in the PATA inventory. Since the benzene emissions factor for residential wood combustion was recently updated and confirmed through a literature review, it seems most likely that the over-prediction is a result of overestimating the amount of residential wood heating being done in the Portland area. A new survey of residential woodstove use should lead to more accurate emissions estimates.

5.3.2.2 1,3 Butadiene

Although many of the 1, 3-butadiene measurements were below the MQL, there is substantial overlap with the model predictions as shown in figure 5.6. This pollutant is primarily an indicator of motor vehicle impact. CALPUFF tends to over-predict slightly but consistently across the metropolitan area.

Figure 5.6 Model to monitor comparisons for 1, 3 butadiene



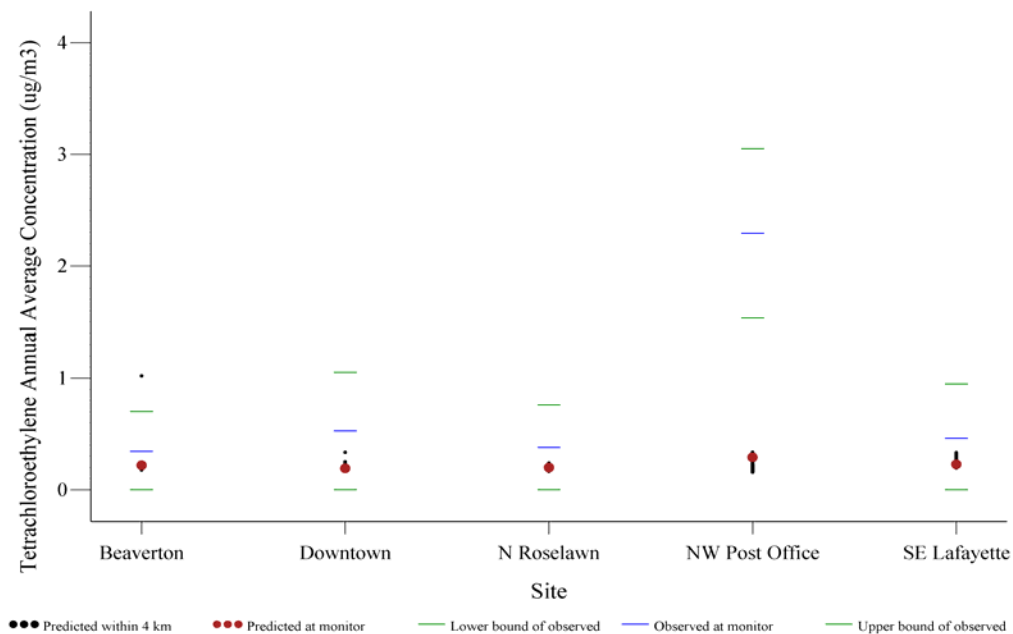
5.3.2.3 Perchloroethylene

Dry cleaning impact was also expected to be consistent across the city as evidenced by direct observation and the model's perchloroethylene predictions. Because the exact locations of dry cleaners were known and used in the model, it is straightforward to identify the sources with major impacts.

Measurements at the NW Portland site, however, were higher than the other four sites, and substantially higher than the model estimates. The NW Portland monitoring site is located within 50 yards of the exhaust from a neighborhood dry cleaner suggesting local impacts from dry cleaning operations not necessarily reflected by the model. Figure 5.7 shows model to monitor comparisons for perchloroethylene. As a result of a correction in perchloroethylene emissions after the initial CALPUFF modeling, the modeled concentrations were adjusted adjacent to a drycleaner located in Beaverton. The

drycleaner is located about 1.9 miles from the Beaverton toxics monitor and the adjustment in modeled concentrations is not significant at this monitor. If the change were reflected in the model to monitor graph (Figure 5.6), the slightly lower modeled value would be in better accord with the monitored value at the Beaverton site.

Figure 5.7 Model to monitor comparisons for perchloroethylene (also known as tetrachloroethylene)



5.3.3 Carbonyls

Acetaldehyde and formaldehyde are similar carbonyl compounds both measured by the same method. Both are common combustion products but ambient concentrations are comprised of primary and secondary components because both are readily formed in the atmosphere by chemical reactions. Due to limitations of the CALPUFF model only the primary components were simulated in this study and therefore it was expected that the model would under-predict.

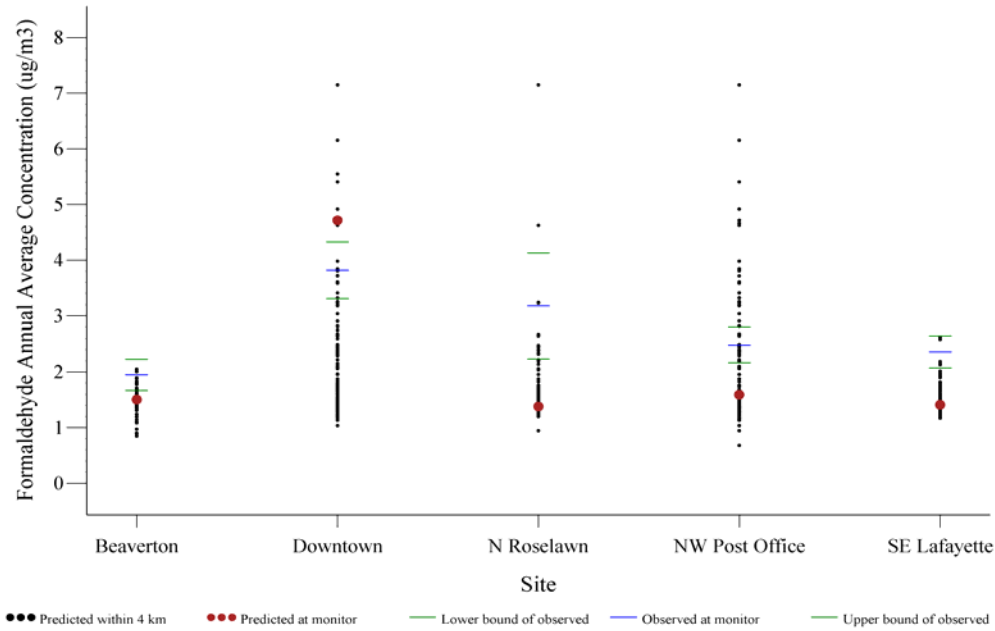
Although it is not possible to determine the secondary fraction of either ambient carbonyl's concentrations directly from monitoring data, some estimates are available from previous modeling studies. One such study compared the [UAM-Tox] modeled results in Houston and Baltimore-Washington to ambient concentration for 1988, 1995, and 1999 (Ligoeki et al., 1992). The results for Baltimore-Washington (which is more similar to Portland than is Houston with respect to air toxics emissions and climate) suggest that the secondary fraction is approximately 70%-80% in winter and 90%-95% in summer. EPA has also found that the ASPEN model predicts approximately 60% of total carbonyl as secondary. More recent modeling in the Northwest using OZIPPR also predicts a secondary contribution of about this magnitude.

5.3.3.1 Primary Formaldehyde

Measured annual average formaldehyde ranged from 1.9 $\mu\text{g}/\text{m}^3$ in Beaverton to 3.8 $\mu\text{g}/\text{m}^3$ Downtown. Figure 5.8 shows how the model predictions for formaldehyde at the sites, and within the 4 kilometer radius, compare to the measured annual averages with

their associated 95% confidence intervals. A much smaller difference, on the order of 30% of the total, is apparently contributed by secondary formation in the Portland area. Table 5.1 also gives the ratios for each of the sites, ranging from 0.5 to 0.9.

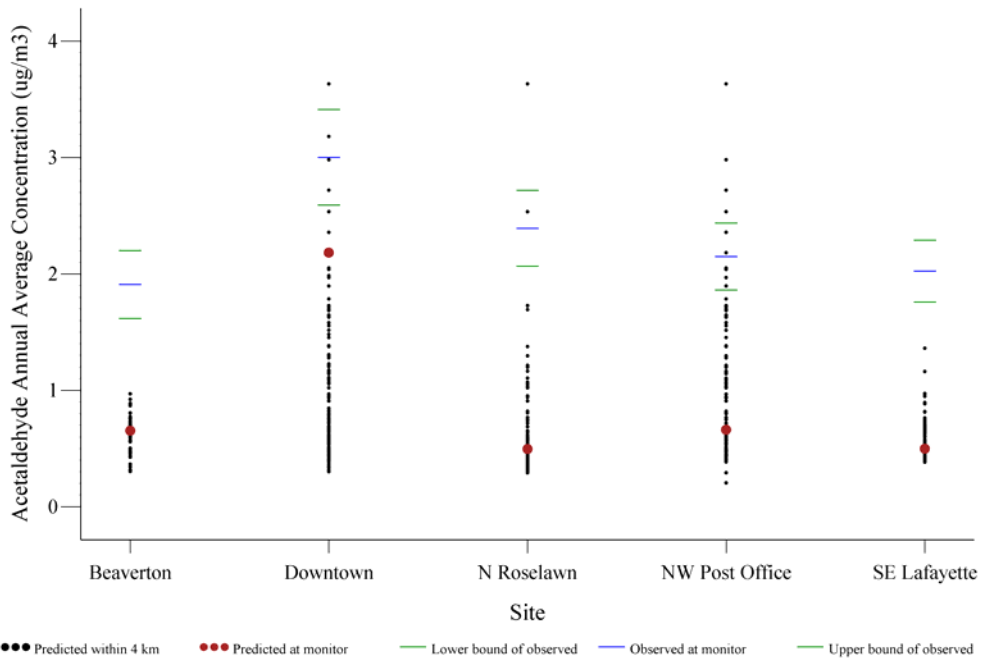
Figure 5.8 Model to monitor comparisons for formaldehyde



5.3.3.2 Primary acetaldehyde

Figure 5.9 shows that compared to monitored concentrations, acetaldehyde was generally underpredicted by the model. Model predictions include only the primary component of total acetaldehyde, while the measured values also include secondary components. Table 5.1 shows that for all the monitor locations, the ratio of the mean primary acetaldehyde concentration within 4 km to the mean measured total acetaldehyde concentration ranged from 0.26 to 0.50.

Figure 5.9 Model to monitor comparisons for acetaldehyde



5.3.4 Chloroform and POM

Little was gained from the comparisons of measured chloroform or POM to the CALPUFF concentration estimates. Table 5.1 contains the monitored and modeled annual averages for both these pollutants.

In the case of chloroform, the model predicts very low ambient concentrations arising largely from background. All of the ambient levels were below the minimum quantitation level.

As was already discussed with respect to benzene, the model greatly over-predicted POM concentrations. POM and benzene concentrations are dominated by residential wood combustion emissions. DEQ examined model performance for benzene during the June-August period when residential wood combustion contributions were absent. The results of this analysis indicate much improved model performance for benzene, and the same would be true for POM.