NCHRP Project 25-55

NCHRP Research Report 1058
Assessing Air Pollution Dispersion Models for Emissions Regulation

Appendix C
Dispersion Model Intercomparison Method and Results

NCHRP 25-55 Appendix C: Dispersion Model Intercomparisons Methodology and Results

To:	Trey Wadsworth, Program Officer, NCHRP, 500 Fifth St. NW, Washington, D.C. 20001
From:	Ed Carr, Seth Hartley, Chris Holder, Samuel Kovach, Sam Pournazeri, ICF
CC:	NCHRP 25-55 Panel Members
Re:	Dispersion Model Inter-Comparisons Methodology and Results Summary (Task 7 and 8 Findings Technical Memorandum)

1. Summary of Work Scope

1.1. Purpose and Summary

This memo summarizes ICF's research evaluating existing models available for near-roadway air pollution assessments against co-located measurements of roadway air pollutants and a unique tracer dataset created under this project. We used both existing and custom project measurements as summarized in the Task 6 report.[1] We use these as a basis of comparison for near-roadway dispersion model performance. Our objective is to comprehensively evaluate the dispersion models that support hotspot assessments of vehicle air pollution in regulatory applications, assess the best performing models, and improve confidence in modeling assessments of near-roadway air pollution. Use of both the co-located and tracer data also allows for the assessment of the regulatory modeling chain (traffic, emissions and dispersion including the determination of background concentrations) against pollutant-monitoring data and the comparison of those results to those for the dispersion model evaluations against tracer data to determine whether the dispersion models or other steps in the regulatory modeling chain contributed the most to inaccuracy and uncertainty in the modeling results.

The focus of this analysis is on one of the most important facility types for hotspot analysis – high traffic volume freeways. A wide variety of models are available for different applications; the U.S. Environmental Protection Agency's (EPA) currently preferred and recommended models are available on its Support Center for Regulatory Atmospheric Modeling website. In the regulatory setting, models used for transportation projects are applied for project-level conformity analysis—based on Clean Air Act Section 176(c) (42 United States Code 7506(c)) and EPA regulations at 40 Code of Federal Regulations (CFR) 51 and 93—and environmental impact analysis under the National Environmental Policy Act (NEPA) and similar state laws. Such analysis typically follows EPA's transportation conformity guidance documents for quantitative hotspot analyses for CO and fine particulate matter ($PM_{2.5}$ and PM_{10}). Significant changes have occurred in EPA's prescribed approach, including

- Use of EPA's Motor Vehicle Emission Simulator (MOVES; California continues to use its EMFAC emission model). Emission models are used to determine the source strength, a required input to dispersion models.

---

[1] NCHRP 25-55: Assessment of Regulatory Air Pollution Dispersion Models to Quantify the Impacts of Transportation Sector Emissions: *Task 6: Initial Report on Tracer Experiments*, March 2022. Included here as Appendix B.

- The 2017 *Guideline on Air Quality Models* (Appendix W) update replacing CALINE3 with AERMOD for refined air dispersion modeling. CALINE3 (and the related CAL3QHC and CAL3QHCR models) has a long history of use and acceptance in the transportation community, requires fewer detailed inputs than AERMOD, and takes less time to set up and use than the AERMOD system. CAL3HQC may still be used for CO screening and nonregulatory purposes.

Air monitoring in this research includes both tracer studies and ambient air pollutant monitoring. Tracer studies involve the controlled release of a gas that is uncommon in the air (or has very low background levels) and easily measured downwind of the release location. Measurements of meteorological conditions are also conducted during the tracer release. The same conditions are then simulated with air quality models and the observed and predicted concentrations are compared. Two tracer datasets were used in EPA's technical support document for the replacing CALINE3 with AERMOD for transportation-related air quality analyses:

- The California DOT (Caltrans) State Route (SR) 99 study, performed in a mostly rural to suburban location near Sacramento, CA with an annual average daily traffic (AADT) rate of just 35,000 and with the closest receptor at 50 meters (m) from the roadway
- The Idaho Falls line-source tracer study representing a rural highway (but without vehicles) at Idaho Falls National laboratory with the nearest receptor at 15 m

This technical memorandum presents the results of the Interstate 80 Freeway Tracer Experiment dataset conducted in Berkeley, CA to evaluate model performance in a near-road, high-volume freeway setting. The analysis includes an evaluation using tracer data, using on-site air quality data collected during the field experiment and against historical data collected in the near roadway environment.

The remainder of this section summarizes the purpose and work competed under Tasks 7 and 8 of the 25-55 project. It first presents a detailed table of the contents of the memo. It then discusses the air quality models we evaluated and how the evaluation was performed. Sections 2 through 4 provide details on our technical approach. Section 2 provides details on the dispersion modeling approach, while section 3 summarizes the observational dataset used in the evaluation. Section 4 discusses the approach for emissions modeling used in this evaluation. We note that this work is not focused on evaluation of emissions models and, as the project takes place in California, it relies on the EMFAC model, not the US EPA's MOVES model for estimating vehicle emissions. Section 5 presents the results of the dispersion modeling evaluation. In particular, Section 5.2 presents side-by-side comparisons of the performance of the different models. Finally, Section 6 provides a discussion on determining representative background concentrations.

1.2. Contents

1.3. Task Summary

Task 7 is focused on the performance of dispersion modeling and inter-comparison of results across models. Task 8 was focused on dispersion models strengths and weaknesses. However, Task 8 scope of work was minimized in order to conduct tracer experiments from moving as completed in Task 6.[2] The following summarizes the scope of Tasks 7 and 8.

1.3.1. Task 7: Perform Dispersion Model Inter-Comparisons

The focus of Task 7 was to develop a robust set of statistics for each model, evaluating performance of the models in terms relevant to NAAQS evaluation, graphically illustrating performance that facilitates understanding of the models, and to the extent possible assessing if the models are finding the right answer for the right reason.

The analysis performed in this task is based on the dispersion models identified for comparison in Task 2[3] and the methods for evaluation determined in Task 3[4] along with the data collected in Task 6[5]. These results are used in a model inter-comparison assessment, one of the key outcomes of this research project. Finally, we evaluated each model using the field data for the I-80 Freeway experiment, with facility-specific geometry, land use, and on-site meteorology. This also included comparison among modeling options, including the volume, area, and RLINE source formulations available for AERMOD. We developed a set of modeling inter-comparison statistics for each of the models assessed. This included the tracer dataset, the near-road air quality monitor for 1-hour and 8-hour CO standard and the 24-hour and annual $PM_{2.5}$ standard using one-year of representative hourly meteorological, traffic data, emission factor model (EMFAC) and representative background concentration, along with the short-term CO and $PM_{2.5}$ concentrations observed as part of the I-80 Freeway experiment.

1.3.2. Task 8: Identify Strength and Weakness in the Dispersion Models

The original purpose of Task 8 was to analyze the results of Task 7 to identify the best performing models and, where possible, the underlying basis for its better performance, and identify strengths and weaknesses for each model in the studied facility. To accommodate the additional cost associated with the use of a mobile source tracer release platform, the Phase II work plan removed this task, however the model performance statistics and graphical results which are the basis for the strength and weakness analysis are prepared and included as part of Task 7. We present in this memorandum performance statistics and a discussion on quantitative results across models, along

---

[2] Discussed further in the Phase II work plan memorandum: *NCHRP 25-25, Assessment of Regulatory Air Pollution Dispersion Models to Quantify the Impacts of Transportation Sector Emissions:*
Proposed Work for Phase II, Memorandum to NCHRP 25-55 Advisory Panel from Edward Carr, Seth Hartley, Mike Brady, and John Zamurs, October 31st, 2019.

[3] Task 2 – Identify Air Quality Models and Field Studies with Suitable Datasets for Model Intercomparison for Transportation Projects, Edward Carr, Seth Hartley, Mike Brady, Elliott Wezerek, and John Zamurs, October 26, 2018.

[4] Delivered as part of the Phase I Interim Report – Research Plan and Design, Edward Carr, Seth Hartley, Chris Holder, Elliott Wezerek, Mike Brady; John Zamurs; Jeffrey Collett, and Arsineh Hecobian, April 2019.

[5] NCHRP 25-55: Assessment of Regulatory Air Pollution Dispersion Models to Quantify the Impacts of Transportation Sector Emissions: Task 6: Initial Report on Tracer Experiments, March 2022. Included here as Appendix B.

with a methodology for determining representative background concentration needed for NAAQS modeling.

1.4. Air Quality Models Included

Two types of model categories were evaluated:

- refined
- screening

The three *refined* are:

- ADMS-Roads
- AERMOD-RLINE
- AERMOD

The two *screening* modeling applications are:

- CAL3QHC
- AERSCREEN (AERMOD in screening mode)

1.4.1. Screening Model Details

AERSCREEN is EPA's recommended screening model for single sources in most kinds of terrain, and EPA regularly updates and maintains the software. AERSCREEN uses the same AERMOD algorithm to model the dispersion processes, but rather than using historical meteorological data it develops a suite of meteorological conditions based on land-use setting. It is, designed to determine a maximum 1-hour-average air concentrations for each downwind distance specified. The 1-hour maximum can be converted to other averaging times through use of persistence (scaling) factors.

EPA allows the use of CAL3QHC as a screening model for carbon monoxide (CO) evaluations, although EPA cautions that its boundary layer dispersion science is considered outdated. CAL3QHC uses CALINE3's dispersion algorithms but expands its functionality by including calculations for traffic queueing, producing maximum 1-hour-average air concentrations (from "worst-case" meteorology) which can be converted to 8-hour averages for comparison to the typical more stringent CO air quality standard, using persistence factors.

For this screening model evaluation, we ran these two screening models for CO following regulatory guidance to the extent possible. This is consistent with the intention of screening models to produce results that are conservative—that is, overstated by a margin designed to give the user confidence that in no case could higher concentrations be predicted when using a refined model or against observed values. The model evaluations presented here are designed such that comparing CO screening outputs to measured outputs ensures confidence in this margin.

1.4.2. Refined Model Details

Table 1-1 summarizes the refined models used in this evaluation.

AERMOD is EPA's preferred model for near-field air quality dispersion. It offers flexibility in terms of specifying terrain, meteorology, source setup, and outputs, which can include hourly air concentrations if needed, though it has not been validated against field data for transportation features like barriers and depressed roadways. EPA regularly updates the model. AERSURFACE (current version released in 2020), while not a regulatory program in the AERMOD system, can help supply modeling values for surface roughness (z_o), albedo (α), and Bowen ratio (Bo).

An update in 2019 incorporated the RLINE model's treatment of dispersion from emissions from near-surface transportation sources. The model formulation is considered a true line source treated as a series of point sources whose dispersion results are integrated. However, use of the RLINE feature in this model update is limited to flat terrain settings (as complex terrain cannot be used). The line source formulation as used in RLINE is a beta option in the current version of AERMOD (v21112) and in EPA's current PM Hot-spot Guidance. The beta option is considered to have had sufficient review to meet requirements for consideration as an alternative model when there is no preferred model.[6]

ADMS-Roads is regularly updated and maintained by a consulting firm (Cambridge Environmental Research Consultants) in the United Kingdom. Its dispersion algorithms are similar to those of AERMOD, though it also has parameterizations for roadway barriers and street canyons. Like AERMOD, its boundary-layer algorithms are sophisticated, with increasing accuracy as additional observed parameters are included. Both AERMOD and ADMS-Roads can produce hourly-resolved concentration estimates at user-specified receptors. A model update in 2020 allows users to supply both vehicle emission rates and vehicle activity data (traffic flows), where the latter engages the model algorithms for traffic-induced turbulence (previously, users had to use the model's UK-based traffic emissions calculators to engage the turbulence generator for most pollutants, thus making it difficult to incorporate custom emission values while including the traffic-based turbulent mixing).

---

[6] Per the updated PM Hotspot Guidance (EPA-420-B-21-037, October 2021):
The latest version of AERMOD now also includes two additional source types to represent line sources: "RLINE" and "RLINEXT" (for RLINE-extended). RLINE is a Beta feature, meaning its use requires alternative model approval (see Section 7.3.3), and RLINEXT is an Alpha feature, meaning it is for research purposes only.
In limited cases, an alternate model for use in a PM hot-spot analysis may be considered. … However, should a project sponsor seek to employ a new or alternate model for a particular transit or highway project, that model must address the criteria set forth in Section 3.2 of Appendix W. Determining model acceptability in a particular application is an EPA Regional Office responsibility involving consultation with EPA Headquarters, when appropriate.
https://www.epa.gov/system/files/documents/2021-10/420b21037.pdf

Table 1-1. Refined Models Selected for Evaluation

Selected Model	Guidance Status	Technical Notes	Other Notes
AERMOD	EPA's preferred model for near-field (including mobile sources) for 50 km.[a] Includes PM and CO hotspot analyses, NO2 near-monitor siting, PM SIP attainment determinations, and PSD.[b] Model updated regularly (last updated in 2021).	Accounts for low-wind plume meander Gaussian plume diffusion (vertical diffusion follows bi-Gaussian PDF during convective conditions). terrain interactions, building downwash, and urban. Can produce hourly concentrations. Also other time averages (including some customized to the NAAQS).	Customizable for receptor locations, hourly meteorology, time-varying emissions, and source type (point/area/volume), along with initial plume characterization. Lagrangian models (e.g., CALPUFF[c]) may perform better in unique scenarios with complex terrain and winds, but with reduced spatial resolution. Area and volume sources can be valid representations of roadways, but parameterizations for roadway barriers and depressed roadways are not available.
AERMOD-RLINE (AERMOD with line source as integrated point sources)	RLINE source type is a beta (pre-promulgation) option, with barriers and depressed roadways as alpha (incompletely-evaluated) options.[d]	Accounts for low-wind plume meander (already available for AERMOD volume sources but not for area and line sources).[d] Approved for flat terrain only.[d]	Accounts for effects of roadway barriers and depressed roadways (alpha options, not rigorously tested), which AERMOD currently cannot do with area- and volume-source parameterizations.[d]
ADMS-Roads (Cambridge Environmental Research Consultants, UK)	ADMS overall is used widely in the UK. EPA lists ADMS (an industrial-source model) as an alternative model (currently available version is ADMS 5). Model updated regularly (last updated in 2020).	Dispersion formulations similar to those of AERMOD. Can produce hourly concentration if needed, or other time averages and statistics (including some customized to NAAQS forms).	Roadways decomposed into elements that are parameterized according to where receptors are located. Accounts for effects of roadway barriers and street canyons. Offers several options for meteorological data depending on the parameters measured.

Notes: EPA = U.S. Environmental Protection Agency, disp. = dispersion, km = kilometers, PM = particulate matter, CO = carbon monoxide, NO2 = nitrogen dioxide, SIP = State Implementation Plan, PSD = Prevention of Significant Deterioration, PDF = probability density function, NAAQS = National Ambient Air Quality Standards, NA = not available, .

[a] From Appendix W – EPA Guideline on Air Quality Models – FR 82, No. 10, pp 5182-5235 (January 17, 2017).

[b] From "Technical Support Document (TSD) for Replacement of CALINE3 with AERMOD for Transportation Related Air Quality Analyses". EPA-454/B-16-006 (December, 2016).

[c] See https://www.epa.gov/scram/air-quality-dispersion-modeling-alternative-models#calpuff.

[d] From "User's Guide for the AMS/EPA Regulatory Model (AERMOD)". EPA-454/B-21-001 (April 2021).

1.5. Database Used in the Intercomparison

We used the observed concentration data from the I-80 Freeway tracer experiment dataset in the model inter-comparison. This experiment included 10 hourly experiments over 3 days. It measured concentrations of an inert tracer released from moving vehicles at multiple locations downwind of the freeway. It also measured air pollutants emitted from all vehicles on the freeway at the same time, but in fewer downwind locations.

We also include a comparison for the models be made against long-term monitored data from the near-road SLAMS air quality monitor at the same location for both $PM_{2.5}$ and CO. This comparison with the NAAQS standard (design values; DV) is for 1-hour and 8-hour CO standard and the 24-hour and annual $PM_{2.5}$ standard using 1-year of representative hourly meteorological, traffic data, emission factor model (in CA EMFAC) and representative background concentration. Table 1-2 summarizes these datasets.[7]

Originally, it had been proposed to evaluate the air quality models using tracer data collected during the GM Sulfate Experiment, however it was decided that the I-80 Freeway Tracer Experiment would benefit more from the addition of upwind air quality monitoring, additional on-site traffic data collection, and the addition of an additional tracer release vehicle.

Table 1-2. Datasets used for Inter-comparison Study

Study Name and Location	Facility Type	Comment
I-80 Freeway Tracer Experiment	High volume Freeway	Mobile platform release, use of multiple tracers, high volume freeway (10-lane 280,400 AADT). Co-located measurements of CO, $PM_{2.5}$, and BC made during the tracer experiment.
Aquatic Park Berkeley air quality monitoring data	High volume Freeway	Long-term, near-road air monitoring of CO, $PM_{2.5}$, and BC collected and maintained by BAAQMD.[a]

[a] Bay Area Air Quality Management District. The project area's local air management agency.

1.6. Types of Comparisons Made

There are three sets of comparison approaches employed in this Task. Each provides a dataset used to compare the different models against observed concentrations in the near roadway setting. All were obtained in the I-80 Freeway Tracer experiment. The three applications are:

- An application for the inert tracer released from moving vehicles during the experiment (*tracer-based observations*)
- An application for air quality impacts of traditional vehicle pollutants from all vehicles on the road during the experiment (*pollutants-based observations*)
- An analysis of air quality in a long-term, regulatory application at the same site as the experiment (*regulatory application*)

---

[7] Further detailed evaluations of near-road monitors have been recently conducted as part of the Near Road Pooled Fund Study Brown et al., Conditions Leading to Elevated PM2.5 at Near-Road Monitoring Sites: Case Studies in Denver and Indianapolis, Int. J. Environ. Res. Public Health 2019, 16, 1634.

Evaluations were performed for the three evaluation approaches.

- Quantitative
- Graphical
- Qualitative

Section 2.3 provides more information on these approaches.

All five models are evaluated in all three applications. However, not all applications are used identically for each evaluation approach. Specifically, qualitative evaluations only apply to the regulatory approach (this focuses on inputs as used for regulatory or similar model applications) and statistics are limited in cases when the applications or model categories provide limited datasets. Quantitative and graphical evaluations include:

- Average residual (bias), absolute average residual
- Standard deviation of residual
- Root Mean Square Error (RMSE)
- Coefficient of determination (r^2)
- Robust Highest Concentration (RHC)
- Fractional bias
- QQ plots
- Time-series plots
- Paired scatter plots

while qualitative evaluations include subjective consideration of items including:

- User interface,
- Ease of use,
- Model transparency
- Output resolution, and
- Specific functional capabilities.

It should be noted that these series of evaluations (tracer, pollutant, and regulatory) are a critical element in assessing the performance of the complete traffic, emission and dispersion modeling chain (including the determination of background concentrations) for their intended regulatory applications, i.e., showing compliance with the applicable NAAQS. If the dispersion models perform well when compared to the observed tracer concentrations in this application, but the full roadway pollutant modeled concentrations perform relatively less well against near-road monitoring data, then the conclusion may reasonably be drawn that other parts of the modeling and evaluation chain (traffic, emissions modeling, and/or assessment of near-road concentrations from the road source only including determination of background concentrations) are contributing to the increased error. Findings such as from this study may be useful in determining priorities for future improvements to the traffic, emissions, and dispersion modeling chain.

We also implemented a qualitative scoring system and report on each of these categories for all models. All five models are employed in this comparison and applied to all three applications. However, CAL3QHC is not applied to PM evaluations as EPA does not recognize use of CAL3QHC for $PM_{2.5}$ screening. Also, as screening models produce single values relying on screening rather than actual meteorology, only peak-to-peak evaluations are included.

1.6.1. Tracer-Based Model Intercomparison Summary

This evaluation uses the 10 one-hour experiments conducted using three different tracers released from mobile platforms. Three vehicles served as tracer release platforms, one distinct tracer per vehicle. Appendix C discusses these tracers and the experiment. The multiple platforms were used primarily to assess if a sufficient number of vehicles was employed to achieve steady concentrations. The location of the receptors and the strength of the emissions vary in each experiment. The model evaluation uses a full set of statistical performance measure to compare models. Each hour observation is independent from other hours. Note that model-to-model comparisons are based on comparing different model's performance against the observed values, not directly between the models.

1.6.2. Pollutants-Based Model Intercomparison Summary

The evaluation uses the co-located measurements of BC, CO, and $PM_{2.5}$ made during the same days as the tracer experiments. These pollutants are released from all vehicles traveling on the roadway. We used the difference in the 1-hour averaged downwind and upwind concentration to isolate the contribution from vehicle emissions on the freeway only. Meteorology and other location specific data are the same days as the tracer experiment, however slightly different hours were captured for the air pollutants on the experimental days. To account for the time differences, a separate set of meteorological hours is used in the air pollutant evaluation. While the Tracer-based intercomparison isolates the dispersion component of the model chain, this intercomparison also evaluates impacts of the emissions modeling and simultaneously observed up- and down-wind concentrations.

We do not include NO, NO_2, and NOx in this evaluation as the measured data were not considered reliable.

The on-road vehicle emissions were determined from applying California's Air Resources Board (ARB) EMFAC2021 emission factor model and observed traffic. Traffic was determined from Caltrans Performance Measurement System (PeMS) values except for hours during experiment #2 when subcontractor Wiltec made direct vehicle counts and vehicle fleet mix. Wiltec data provided a much more granular counts by vehicle classes which helped improve emissions estimates, however, we have not quantified how much it helped improve the overall model performance. In all cases, emissions modeling relied on county-specific defaults for vehicle age and fuel distributions since no measurements were made for these variables. Real-time truck and all-vehicle VMT speed values were determined from Caltrans' PeMS data (from loop detectors located within the proposed freeway segment of the tracer study), and from the Wiltec traffic data. The quality of these traffic data to support emission and dispersion modeling exceeds that typically available for regulatory assessments. That is, any error attributable to the traffic data in this analysis is expected to be less than that for future year transportation modeling used for regulatory analyses for NEPA and conformity. Conversely, the greater error for transportation modeling compared to measured traffic data contributes to increased errors in emission and dispersion modeling in regulatory applications compared to that reported in this study.

1.6.3. Regulatory Application Model Intercomparison

This evaluation uses the co-located measurements of pollutants available from the Bay Area Air Quality Management District (BAAQMD)-operated, near-road, SLAMS air quality monitoring trailer at this site for a 1-year period. Our regulatory evaluation includes measurements of CO, BC, and $PM_{2.5}$ made during a 1-year period of observations with 1 year of meteorological data for the same year measured nearby. All data cover the year 2019.

As with the Pollutants-Based Intercomparison, these three pollutants are released from vehicles traveling on the roadway. However, here we do not have corresponding, on-site upwind observations, so rely on nearby, upwind monitoring sites. Consistent with regulatory evaluations, we report here design values (DV) consistent with the form of the NAAQS standards for 1-hour and 8-hour CO and 24-hour and annual $PM_{2.5}$. We exclude BC here as there is no NAAQS for it. Traffic data is based on local traffic data collection at the same location as for the pollutant evaluation but for year 2019 as available through Caltrans PeMS records. Along with the observed air quality values measured at the Berkeley near-road site, we pair background concentrations from the nearby San Pablo air monitoring station (AIRS ID: 06-013-1004) located just south of Carquinez bridge. The site is on the west side of I-80 so has frequent exposure to the relatively clean air over San Pablo Bay. Both sites use the same instrument types to measure CO and $PM_{2.5}$.

Similar to the case with the pollutant-based model intercomparison, the quality of these traffic data to support emission and dispersion modelling exceeds that typically available for regulatory assessments. That is, any error attributable to the traffic data in this analysis is expected to be less than that for future year transportation modelling used for regulatory analyses for NEPA and conformity. Conversely, the greater error for transportation modelling compared to measured traffic data contributes to increased errors in emission and dispersion modelling in regulatory applications compared to that reported in this study.

This configuration is designed to mimic how a regulatory evaluation would be performed, but with additional information so that detailed model evaluation statistics can be calculated. Thus, it includes both a DV and statistical evaluation involving all modeled hours. That is, the graphical and quantitative evaluations include 365 data points for 24-hour average $PM_{2.5}$, 8,760 hourly data points for 1-hour CO, and 1,095 8-hour CO values, along with a comparison of the (single) representative design values for each pollutant and averaging period.

2. Dispersion Model Approach

This section provides details on the modeling approach and tools used for the evaluation. It also discusses the model settings we used in the evaluations, and how we used the outputs in the evaluations (evaluation results are provided in Section 0).

- Section 2.1 covers Model Versions and Overall Modeling Approach
- Section 2.2 discusses Model Settings and Inputs
- Section 2.3 discusses Model Evaluation Approach and Metrics

Note: all modeling is performed in local standard time, LST. That is, no conversion was made for Daylight Saving Time, consistent with AERMOD modeling for annual periods.

Table *2-1* is a "quick reference" to the details of the modeling. Some additional nuances are missing in the table which are described in greater detail in the remainder of Section 2.

2.1. Model Versions and Overall Modeling Approach

The most current version of the air dispersion models is used in the analysis. That is

- AERMOD (21112),
- ADMS-Roads v5.0
- AERSCREEN (21112) and
- CAL3QHC v04244.

Note that CAL3QHC is only approved as a screening model only for CO. We followed regulatory guidance where it exists and provide reasons if/when we selected a different course of action. We evaluated regulatory screening models separately from regulatory refined models.

2.2. Model Settings and Inputs

In this section, we generally do not provide the specific parameter values for each model run; rather, we provide the approach we followed for how we developed those values, which includes careful consideration of applicable regulatory guidance as well as the specific characteristics of the project site.

Table 2-1. Summary of the three modeling analyses conducted at the I-80 freeway experiment site, Berkeley, California

	(1) Tracer-based	(2) Pollutant-based	(3) Regulatory application
Pollutants	3 tracers (PMCP, PDCB, PMCH), plus the sum of the 3[8]	CO, $PM_{2.5}$, BC	CO, $PM_{2.5}$
Time Period	5 separate hours: 5/3/2021 at 8–9am, 6/6 and 6/24 at 8–9am and 1040–1140am	9 separate hours: 5/3/2021 at 8am, 6/6 and 6/24 at 8-11am	1 year (hourly) (2019)

---

[8] See Appendix B for details of the experimental approach and tracers employed.

	(1) Tracer-based	(2) Pollutant-based	(3) Regulatory application
Receptors and Sampling Locations	All time periods have 3 receptors at north end (same coordinates each day). 6/6/2021 and 6/24 have 6 at south end. 5/3 has 12 at south end. These receptors are at increasing distances from road.	Single receptor located at BAAQMD monitoring site	Single receptor located at BAAQMD monitoring site
Number of Receptor*Time Outputs (per pollutant)	51 per tracer, 3 tracers plus sum, so 204 total outputs	9	8760
Background	Upwind on-site monitor	Upwind on-site monitor	San Pablo (AIRS 06-013-1004), just south of Carquinez Bridge
On-site Meteorology	Met One AIO sonic anemometer (logged 1-min average winds - calculated sigma theta (standard deviation of horizontal wind direction).	Met One AIO sonic anemometer (logged 1-min average winds - calculated sigma theta (standard deviation of horizontal wind direction).	None, nearest representative site based on consultation with BAAQMD – Oakland Sewage Treatment Plant (STP) and Oakland NOAA National Data Buoy (OKXC1) Oakland Berth 34 for missing data
Other Met Needs	OAK airport for other surface parameters (cloud cover); OAK twice-daily radiosonde for upper -air.	OAK airport for other surface parameters (cloud cover); OAK twice-daily radiosonde for upper -air.	OAK airport for cloud cover, OAK twice-daily radiosonde for upper - air.
Terrain	AERMOD, ADMS-Roads, CAL3QHC: utilize terrain. AERSCREEN: area source cannot utilize terrain	AERMOD, ADMS-Roads, CAL3QHC: utilize terrain. AERSCREEN: area source cannot utilize terrain	AERMOD, ADMS-Roads, CAL3QHC: utilize terrain. AERSCREEN: area source cannot utilize terrain
Emissions	Tracer release rates	EMFAC emission factors with corresponding traffic data. Stratify $PM_{2.5}$ exhaust, brake/tire, and road dust, and stratify BC exhaust, brake/tire.	EMFAC emission factors with corresponding traffic data (2019). Stratify $PM_{2.5}$ exhaust, brake/tire, and road dust.
Traffic	Not needed except for certain ADMS-Roads and CAL3QHC inputs. PeMS and Wiltec counts. Default county age and fuel	PeMS and Wiltec counts. Default county age and fuel distributions PeMS truck counts and all-vehicle speed.	Annual data from on-site PeMS (2019)

	(1) Tracer-based	(2) Pollutant-based	(3) Regulatory application
	distributions. PeMS truck traffic counts and all-vehicle speed.		
Models	AERMOD (volume, area, RLINE); ADMS-Roads (line); AERSCREEN (area); CAL3QHC (free-flow link)	AERMOD (volume, area, RLINE); ADMS-Roads (line); AERSCREEN (area); CAL3QHC (free-flow link; only for CO)	AERMOD (volume, area, RLINE); ADMS-Roads (line); AERSCREEN (area); CAL3QHC (free-flow link; only for CO)
Sources[9]	Cover both freeway directions as one source (area and line sources) or many sources (volume). Incorporate some ramp areas near the monitors.	Model the two freeway directions as separate sources where the model allows (some models allow only a single source; many sources for modeling with volume sources); same for ramps that may have substantive impacts on the monitor	Model the two freeway directions as separate sources where the model allows (some models allow only a single source; many sources for modeling with volume sources); same for ramps that may have substantive impacts on the monitor
Quantitative and Graphical Evaluation Measures	**Refined models:** Average Residual Absolute Average Residual Standard Dev. of Residual Root Mean Square Error R-squared Fractional Bias Robust Highest Concentration – Observed Robust Highest Concentration – Modeled Scatter plots Q-Q plots **Screening models:** Peak modeled to peak observed	**Refined models:** Average Residual Absolute Average Residual Standard Dev. of Residual Root Mean Square Error R-squared Fractional Bias Robust Highest Concentration – Observed Robust Highest Concentration – Modeled Scatter plots Q-Q plots **Screening models:** Peak modeled to peak observed	**Refined models:** Design Value (DV) comparison Scatter plots Q-Q plots Time series plots **Screening models:** Peak modeled to peak observed

---

[9] See the following discussion on Emission Sources for details on the selected approach.

2.2.1. Upwind Monitoring

Just prior the tracer experiments, measurements were made and in all cases were well below detection limits indicating no background levels of the tracer gases.

For the pollutant-based evaluation, we measured CO, $PM_{2.5}$, and black carbon (BC) on the upwind side of I-80. These served as the background monitored concentrations. The downwind BAAQMD near-road monitor was compared against the background monitored + modeled roadway concentrations.

For the regulatory evaluation, the offsite San Pablo CO and $PM_{2.5}$ monitors (AIRS 06-013-1004) served as the background monitored concentrations (further discussion on this choice can be found in Section 6). Both the CO and $PM_{2.5}$ monitors have some missing data, and those times were not included in model-monitor evaluations.

2.2.2. Emission Sources

2.2.2.1. General Source Discussion

In this section we provide some higher-level descriptions of the sources modeled, with further details in Sections 2.2.2.2–2.2.2.4.

All dispersion modeling covered the roadway segment of I-80 between Hearst Avenue in the north and 65th Street in the south, which runs past the Aquatic Park in Berkeley, California (approx. 2.4 km or 1.5 miles in length) (Figure 3-1). For **AERMOD**, we modeled with AREAPOLY (complex polygon), RLINE (a line source that AERMOD models as a series of point sources with their dispersion curves integrated), and volume sources (three-dimensional sources). We modeled line sources with **ADMS-ROADS** and free-flow links with **CAL3QHC**. With **AERSCREEN**, there is a limitation of only allowing one emission source per run, and so the most realistic option was to model the overall freeway area as a single area source. When we refer to line sources generically, that includes AERMOD RLINE sources unless otherwise stated.

For the **tracer-based evaluations**, we modeled hourly-resolved tracer emission rates (for each of the three tracers, since the tracer study indicated that each had sufficient strength to justify the use of each vehicle, plus the aggregate of the three tracers). The source geometry was characterized as a continuous line source of tracer releases. We parameterized turbulent mixing from the freeway traffic where the models allowed, and we paid additional attention to mixing near the air monitoring locations (on the north and south end of the Aquatic Park) by including some areas of ramp lanes near the monitors as emission sources. With the exception of volume sources, we modeled both freeway traffic directions within the same source(s) (not as separate sources per direction); that is, an area or line source covering both sides of freeway. This simplifies the assumptions of the combined turbulence and limited tracer releases into one overall, general road source. The two directions were aggregated here to accommodate the relatively few release vehicles. As noted in the body of the report, we found sufficient vehicle passes to determine steady hourly concentrations. However resolving separate directions reduces the number of vehicle passes per direction in half, which we did not feel was a level of resolution supported by the experiment. We modeled limited sections of ramps directly adjacent to the free-flowing freeway lanes (e.g., the first parts of off-ramps where traffic exits the freeway), where ramp traffic is assumed to contribute to the well-mixed vehicle turbulence—but only for AERMOD areapoly, not for other sources or models.

For the **pollutant-based and regulatory evaluations**, we modeled hourly-resolved pollutant emission rates estimated from available traffic data, resolved on each directional side of the freeway and with ramp traffic. We included emissions from exhaust, brake and tire wear, and road dust,[10] depending on the pollutant. Since we used directional data, the two directions of the freeway were modeled as separate sources where the models allowed (e.g., an AERSCREEN run allows only one area source), and ramps were also modeled as separate sources (only used on the north end of the project, where the air monitors were located). We parameterized turbulent mixing from the traffic where the models allowed. Directions were resolved for pollutant-based and regulatory evaluations, but not the tracer evaluation, given the high source strength (number of vehicle passes) in each direction.

ADMS-Roads allows input of emission rates (tracer, or pollutants estimated with EMFAC) alongside data on traffic flows—the latter used to estimate traffic-induced turbulence. CAL3QHC incorporates traffic flows and vehicle emission factors to calculate total emissions.

2.2.2.2. Source Geometry, Release Heights, and Parameterizations of Traffic-induced Turbulence

The freeway is 5 lanes in each direction with each being 12 feet (ft), totaling a 60-ft (approx. 18-m) source width in each direction, but with extended merge lanes in some locations. For polygon and line sources, one source covered the freeway from far-right lane westbound[11] shoulder to far right lane eastbound shoulder (for the **tracer evaluation**, see Figure 2-1 and Figure 2-2), or we used separate sources from westbound travel and eastbound travel (for the **pollutant and regulatory evaluations**, see Figure 2-3 and Figure 2-4). The line and area source specification necessitated a number of adjustments for adjacent sources because the freeway does not run straight, and where necessary we split sources (lengthwise) so that the source aspect ratios did not exceed 100:1. To cover the modeled area for AERMOD volume sources, we used numerous adjacent sources, with six volume sources across the full width of the freeway (three for each direction of the freeway) and many sources along the length of the freeway (see Figure 2-5 and Figure 2-6, and also Figure 2-7 which is a close-up view of the sources on the northern end so that volume source geometry can more clearly be seen). This setup of volume sources struck a balance between the resolution of the traffic and emissions data (which were by overall traffic direction and not by lane) and the need to ensure that none of the receptors/monitors fell within the model exclusion zone for volume sources (which would happen if the sources were too large).[12]

We generally did not model full ramps. For the **tracer evaluation**, the tracer was not released on ramps, but we assumed some ramp traffic contributed to the well-mixed vehicle turbulence. As such, for modeling with polygons (i.e., AERMOD areapoly sources), we included limited sections of ramps directly adjacent to the free-flowing freeway lanes (e.g., the first parts of off-ramps where traffic exits the freeway). However, for line sources (i.e., AERSCREEN area, AERMOD RLINE, ADMS-Roads, and CAL3QHC) and AERMOD volume sources, those areas cannot be reasonably incorporated into the freeway sources, and it is not practical to create

---

[10] Guidance from BAAQMD for projects in the Bay Area of CA. Email from Stephen Reid (BAAQMD) sent Wednesday, November 10, 2021 11:47 AM to Edward Carr, RE: BAAQMD - PM2.5 hotspot modelling analysis.

[11] I-80 overall is an east-west bound freeway however over the area over the experiment the freeway is oriented in a near north-south direction.

[12] See Section J.3.1 of: Transportation Conformity Guidance for Quantitative Hot-spot Analyses in PM2.5 and PM10 Nonattainment Maintenance Areas. EPA-420-B-21-037 (October 2021). https://www.epa.gov/state-and-local-transportation/project-level-conformity-and-hot-spot-analyses#pmguidance.

sources just for those ramp areas (and AERSCREEN does not allow multiple area sources in a run, see Figure 2-8).

For the **pollutant and regulatory evaluations**, ramp traffic can contribute to mobile emissions *and* near-freeway traffic turbulence. We did not model any ramps at the southern end of the project because they are nearly 2 km away (and generally not upwind) from the monitors, which were on the north end; but, at the north end of the project, we modeled a limited section of the eastbound offramp to University Ave., plus the westbound offramp to University Ave. and a limited section of the University Ave. onramp to the freeway—as these areas are generally upwind of the monitors.

Figure 2-1. "Areapoly" source geometry used in tracer-based evaluations (with AERMOD)

Figure 2-2. Line source geometries used in tracer-based evaluations (with AERMOD, ADMS-Roads)

Figure 2-3. "Areapoly" source geometries used in pollutant- and regulatory-based evaluations (with AERMOD)

Figure 2-4. Line source geometries used in pollutant- and regulatory-based evaluations (with AERMOD, ADMS-Roads)

Figure 2-5. Volume source geometries used in tracer-based evaluations (with AERMOD)

Figure 2-6. Volume source geometries used in pollutant- and regulatory-based evaluations (with AERMOD)

Figure 2-7. Volume source geometries used in pollutant- and regulatory-based evaluations (with AERMOD), zoomed in on a section to more clearly show the source geometries

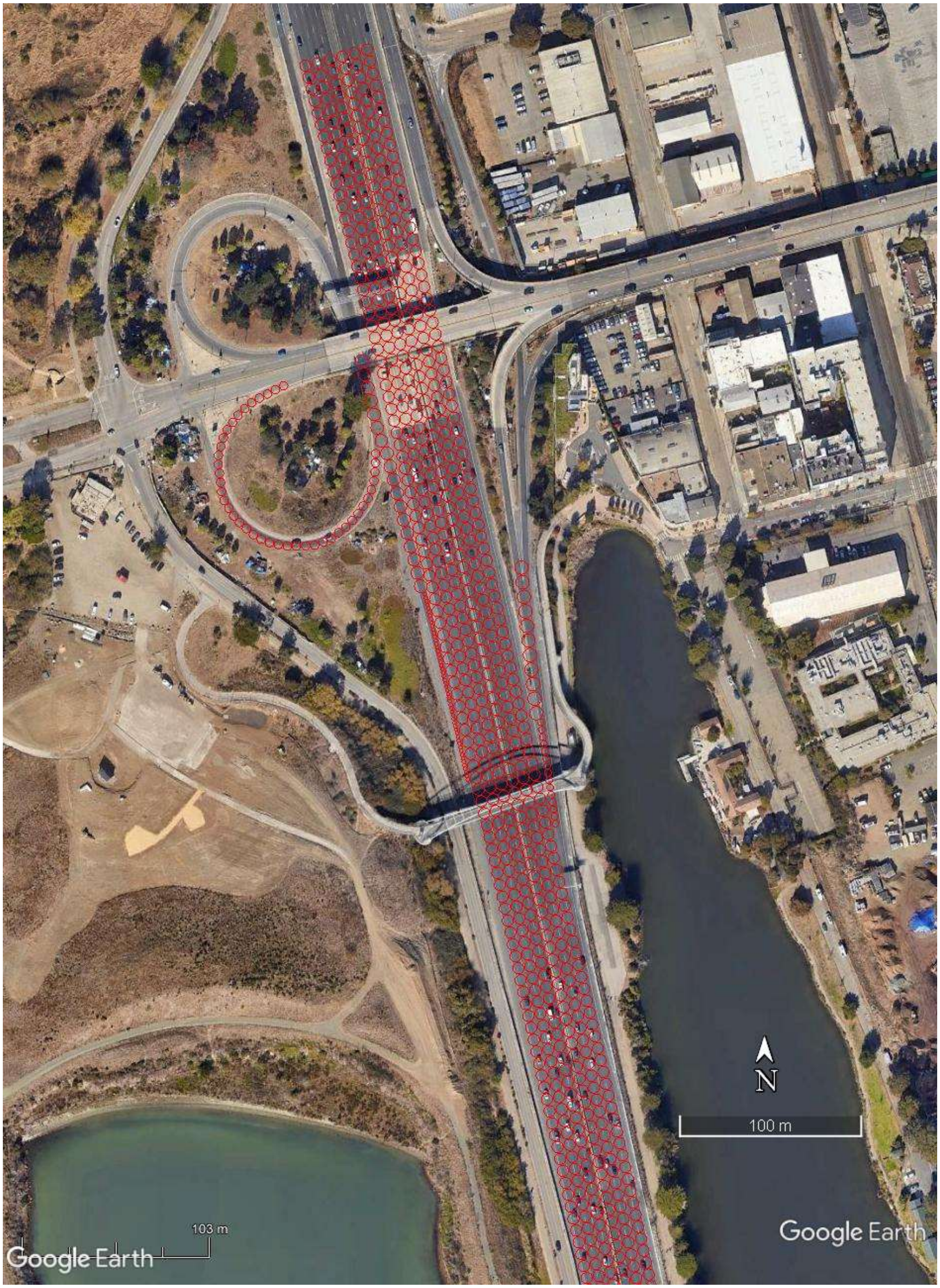


Figure 2-8. Area source geometries used in tracer-, pollutant-, and regulatory-based evaluations (with AERSCREEN)

Note: The AERSCREEN area sources (and the receptors in AERSCREEN) are not actually geospatially referenced (i.e., are not placed at real coordinates). The area source shown here was modeled as a simple rectangle, and here we have artificially displayed it on the project map.

We followed **CAL3QHC** guidance[13] by adding 3 m width to either side of line sources to create a well-mixed zone (see Figure 2-9 and Figure 2-10), with link heights set equal to actual roadway terrain heights (see Section 2.2.3 for additional discussion on terrain). We followed EPA guidance with **AERMOD and AERSCREEN** to calculate source release height (average vehicle height X 1.7 ÷ 2, assuming truck heights are 4.0 m and car heights are 1.53 m, with the average vehicle height calculated with weighting based on the observed hourly truck and car traffic mix), the initial lateral dimension of the source plume (needed only for volume sources: length of side ÷ 2.15, since they are adjacent sources), and initial vertical dimension (average vehicle height X 1.7 ÷ 2.15).[14,15] For **ADMS**, the model takes traffic data as an input to account for traffic-induced turbulence (see Section 2.2.2.4). The 1.7 factor or "injection height" noted above is based on limited data (mostly passenger vehicles) under stable conditions[16] and likely is considerably higher under unstable conditions; the injection height is likely a function of the shape of the vehicle and the angle of wind direction relative to the motion of the vehicle.

---

[13] See Section 3.2.1 of: User's Guide to CAL3QHC Version 2.0: A Modeling Methodology for Predicting Pollutant Concentrations Near Roadway Intersections. EPA-454/R-92-006, Revised (September 1995). https://gaftp.epa.gov/Air/aqmg/SCRAM/models/preferred/cal3qhc-r/cal3qhcug.pdf.

[14] See Table 3-2 and Section 3.3.2.5 of: User's Guide for the AMS/EPA Regulatory Model (AERMOD). EPA-454/B-21-001 (April 2021). https://gaftp.epa.gov/Air/aqmg/SCRAM/models/preferred/aermod/aermod_userguide.pdf.

[15] See Section J.3.1 of: Transportation Conformity Guidance for Quantitative Hot-spot Analyses in PM2.5 and PM10 Nonattainment Maintenance Areas. EPA-420-B-21-037 (October 2021). https://www.epa.gov/state-and-local-transportation/project-level-conformity-and-hot-spot-analyses#pmguidance.

[16] J.A.Gillies,.V. Etyemezian H. Kuhns, D.Nikoli, D.A.Gillettec Effect of vehicle characteristics on unpaved road dust emissions, Atm Env, 39 (2005), 2341-2347. https://doi.org/10.1016/j.atmosenv.2004.05.064

Figure 2-9. Line source geometries used in tracer-based evaluations (with CAL3QHC)

Figure 2-10. Line source geometries used in pollutant- and regulatory-based evaluations (with CAL3QHC)

2.2.2.3. Traffic Data and Emissions

Section 4 of this document describes in detail the traffic and emissions modeling for **pollutant and regulatory evaluations**. The section also provides details on how tracer release data were analyzed.

Based on historical data we estimated that ramp traffic was 8% (by volume) of the freeway traffic at the University exit. Also note that since **CAL3QHC** and **AERSCREEN** do not receive hourly inputs or export hourly outputs, for the tracer and pollutant evaluations we modeled each experiment hour separately (with those hours' correct source inputs). For the regulatory evaluation, we identified the worst-case hour from the 2019 modeled year, defined as the hour with the largest combined westbound and eastbound emission rates, which was Saturday 21 September 2019 at 1pm for CO and Friday 7 June 2019 at 9am for $PM_{2.5}$.

2.2.2.4. Additional Source Discussion—ADMS-Roads

The ADMS model internally estimates traffic-induced initial lateral and vertical turbulence. These are parameterized within the model based largely on traffic volume and fleet characteristics. To our knowledge these values are not being reported in the outputs of the model and therefore cannot be compared against AERMOD.

The latest model allows the user to enter both the vehicle emissions and the traffic data, the latter used to estimate vehicle-induced turbulence. As with the other models, for ADMS-Roads we used the tracer-release data for emissions input for the tracer evaluation, and we used emissions derived from EMFAC for emissions input for the pollutant-based and regulatory evaluations. We used observed traffic data for the inputs to the model for average vehicle speeds and traffic counts per vehicle category. Considering that the model only includes a single vehicle count and speed for each run, we used an average traffic flow for light- and heavy-duty vehicles for each evaluation.

2.2.2.5. Additional Source Discussion—CAL3QHC and AERSCREEN

Since **CAL3QHC** and **AERSCREEN** do not receive hourly inputs or export hourly outputs, for the tracer and pollutant evaluations we modeled each experiment hour separately (with those hours' correct source inputs). For the regulatory evaluation, we selected the approximate worst-case hour from the 2019 target year, defined as the hour with the largest combined westbound and eastbound emission rates, which was 21 September 2019 at 1pm for CO and 7 June 2019 at 9am for $PM_{2.5}$.

2.2.3. Terrain

We used terrain information in the models allowing it for the source types we modeled—**AERMOD (areapoly and volume sources; not yet a beta feature for RLINE sources), ADMS-Roads, and CAL3QHC**. We utilized National Elevation Dataset (NED) DEM data within the AERMOD terrain preprocessor, AERMAP (version 18081).

Running AERMAP provided us the heights above sea level for all sources and receptors (receptors discussed in Section 2.2.5). The freeway emission sources all were between 4 and 5 m above sea level, while the clover-leaf ramp from the westbound freeway to University Ave. reached almost 8 m at University Ave. After examining the receptor terrain results, we manually lowered the terrain heights for three of the receptors in the tracer evaluation that were close to the freeway—it appeared that the resolution of the terrain data did not adequately capture the instantaneous drop in height from the freeway to the access roads directly beside it. For the two receptors on the south end of the project that were closest to the freeway during the June 6 and June 24 experiments, we lowered their heights from 3.19 and 3.43 m, respectively, to 2.6 m, and for the receptor on the north end that was closest to the freeway and used in all three tracer experiment days, we lowered its height from 3.62 m to 2.5 m. After these changes to receptor terrain heights, all receptor heights used for the models listed above were between 2 and 3 m above ground.

ADMS-Roads accounts for complex terrain through input of a grid of elevation data, which we provided to the model as the same DEM data used for AERMAP (converted to a point dataset). We selected pre-set grid resolution of 64x64 for the model to interpolate terrain data.

We entered terrain heights for **CAL3QHC** as part of the z-coordinate for receptors (terrain height + monitor inlet height) and as the source height for the roadway links.

2.2.4. Urbanization

Analysis of the 2019 land-cover data in the vicinity of the model indicates that half the area is medium-to-highly developed land, with about 41% being water, and the 2019 population density within 3 km was about 2,275/km^2 (well above the 750 /km^2 threshold of urbanization according to EPA guidance[17]).

We chose to use the urban modeling setting in **AERMOD** (except for RLINE sources, where it is an alpha testing option only) and **AERSCREEN**, with a 2019 population of 1,116,884 (calculated within 15 km of the project). This setting results in greater dispersion of emissions in the immediate area around the sources, due to the higher surface roughness length that the model assumes with the urban setting. We allowed the default setting of urban surface roughness length of 1 m, reflecting high-intensity development[18]. We also used AERMOD's NOTURBST setting so that on-site meteorological measurements of turbulence are not used during stable conditions in urban environments (to avoid overestimating turbulence from urban heat islands)[19]. Likewise, for **CAL3QHC** we selected Pasquill-Gifford stability class "D" as recommended for urban areas.[20]

Section 2.2.6.4 provides additional information on meteorology for ADMS-Roads, which is related to urbanization.

2.2.5. Receptors

All of the models, except AERSCREEN, allow the user to specify the exact receptor location of the receptors. The locations and inlet heights above ground were recorded and provided by UCB/LBNL team (for the tracer and pollutant based modeling) and by BAAQMD (for the pollutant based and regulatory modeling)[21], with

---

[17] See Section 7.2.1.1.b.ii of: Appendix W – EPA Guideline on Air Quality Models – FR 82, No. 10, pp 5182–5235 (January 17, 2017). https://www.epa.gov/scram/2017-appendix-w-final-rule.

[18] Note that this urban roughness length is used to adjust dispersion during nighttime urban heat island. As both tracer and pollutant-based modeling only occurred during the daytime hours this has no effect on the modeling results.

[19] See Section 3.3 of: AERMOD Implementation Guide. EPA-454/B-21-006 (July 2021). https://gaftp.epa.gov/Air/aqmg/SCRAM/models/preferred/aermod/aermod_implementation_guide.pdf

[20] See Section 6.4.1 of: User's Guide to CAL3QHC Version 2.0: A Modeling Methodology for Predicting Pollutant Concentrations Near Roadway Intersections. EPA-454/R-92-006, Revised (September 1995). https://gaftp.epa.gov/Air/aqmg/SCRAM/models/preferred/cal3qhc-r/cal3qhcug.pdf. Also see Section 4.7.1 of Guideline for Modeling Carbon Monoxide from Roadway Intersections. EPA-454/R-92-005 (November 1992). https://www.epa.gov/sites/default/files/2020-10/documents/coguide.pdf.

[21] See Section 5.1 of: 2020 Air Monitoring Network Plan. Bay Area Air Quality Management District (BAAQMD). July 1, 2021. https://www.baaqmd.gov/~/media/files/technical-services/2020-network-plan-draft-202100526-pdf.pdf?la=en.

terrain heights as discussed in Section 2.2.3, and flagpoles matching the inlet heights of the monitors. For the tracer evaluation, the inlet heights (model flagpole heights) were not consistent across all monitors: inlet heights were as high as 3 m at the receptors closest to the road, with most other receptors near 0.7 m. For the pollutant and regulatory evaluations, the inlet heights were 6 m for the CO monitor, 5 m for the $PM_{2.5}$ monitor, and 4 m for the BC monitor.

With **AERSCREEN**, we estimated the distance of each monitor to the nearest edge of the modeled sources, and then consolidated those distances to a maximum of 10 distances to provide the model (a limitation of the model). Each AERSCREEN run also allows only one flagpole height for all receptors; for model runs for the tracer evaluation, we utilized a 3-m flagpole for all receptors (the approximate inlet height of receptors closest to the freeway, where concentrations were highest), while for the pollutant and regulatory evaluations we modeled each pollutant separately, which enabled use of the correct flagpole heights for the different pollutant sensor inlet heights.

2.2.6. Meteorology

2.2.6.1. Tracer and Pollutant-Based Experiment Days

For the tracer and pollutants-based modeling studies on experiment days, we used the on-site meteorology data collected by UCB/LBNL during the experiment. As the tracer and air pollutants had different periods of data collection, two separate datasets were prepared for ease in modeling. Both the tracer and pollutant-based datasets included data from the May 3, 2021, June 6, 2021, and June 24, 2021, experiment days. For all datasets the data were averaged for a one-hour period consistent with the temporal resolution of the air quality models.

At the south Aquatic Park location, a 3-D sonic anemometer was used to measure the three-dimensional winds with a 1-second reporting frequency. Hourly averages were determined for the:

- vector-averaged horizontal wind speed and direction,
- standard deviation of the vertical wind speed (“σ_w”, a measure of turbulence), and
- standard deviation of the horizontal wind direction (“σ_Θ”, which is another measure of turbulence).

The 3-D sonic was located at approximately 20 m from the nearest travel lane. A careful review of the vertical wind component revealed unrealistically high vertical wind speeds, making the calculated σw results unrealistically large and unsuitable as a measure of the near-road turbulence[22]. We also reviewed use of the vector-average horizontal wind but found it to be unsuitable for representing the ambient horizontal wind speed and direction. In the near-roadway vicinity there are two components of induced turbulence:

- a “vehicle wake turbulence”—these are the perturbations following the vehicles passing and are found behind the passing vehicle; they are the random turbulence motions found over a variety of scale lengths; and
- a “wake passing effect”—this is the air that is dragged by the vehicles and is most pronounced at this location in a northerly direction from the nearby eastbound traffic (which at south Aquatic Park is headed north).

---

[22] The suspected cause for this issue was with the instrument leveling.

Eskridge[23] showed in numerical simulations—for a two-lane divided north-south freeway with a westerly wind of 0.94 m/s, and 1,366 vehicles per hour in each direction, under stable conditions, and with a mean vehicle speed of 80 km/hr—that the wake passing effects extend tens of meters beyond the freeway before falling to background levels. These conditions are similar to the meteorological and traffic conditions during our experiment, although our freeway has much higher volumes and so may have an even stronger wake turbulence effect. Therefore, we concluded the 3-D sonic measurements were "seeing" both the wake turbulence and the wake passing effect, with the wake passing effect biasing observations in the direction of the vehicle movement. As a result, the sonic anemometer was not reporting the "ambient" wind direction or wind speed and was not representative of the upstream winds needed for input to dispersion models.

As a result, the 2-D sonic anemometer was used for the ambient measure of wind speed and direction for both the north and south Aquatic Park dispersion modeling. The 2-D sonic anemometer was located at a background location upwind of I-80. The 2-D sonic recorded data every minute from which hourly averages were determined for the

- horizontal wind speed and direction, and
- standard deviation of the horizontal wind direction, σ_Θ.

The north Aquatic Park 2-D sonic anemometer data was located on the upwind side of the freeway and placed about 200 m to the west of the freeway, so it is free from the vehicle wake and turbulence effects.

Additional meteorological and surface characteristic parameters needed by the dispersion models but not measured on-site are:

- surface air temperature,
- cloud cover,
- upper-air temperature,
- Bowen ratio (ratio of sensible heat to latent heat),
- albedo, and
- surface roughness.

Meteorological data from the Oakland Airport Automated Surface Observing System (ASOS) and upper-air radiosondes were used for the first three bullets above. The latter three bullets were derived using land cover information as described below.

Land Cover Processing and Additional Meteorology

For the AERMOD modeling, the land-cover pre-processor (AERSURFACE[24]) was used in determining surface characteristics (Bowen ratio, albedo, and surface roughness). While AERSURFACE is not a regulatory mandated modeling methodology for developing surface characteristics, it is widely used and maintained by EPA. AERSURFACE uses lookup tables to estimate the surface characteristics based on land cover within a certain distance of the meteorological site (10-km radius for the albedo and Bowen Ratio; 1-km radius for surface roughness). AERSURFACE by default uses inverse-distance weighting for calculations of surface roughness ("ZORAD").

---

[23] Eskridge, R and Catalano, 1987. ROADWAY – *A Numerical Model for Predicting Air Pollutants near Highways*. User's Guide, Atmospheric Sciences Research and Development, EPA/ORD, EPA/600/9-98/010, March 1987.

[24] AERSURFACE User Guide for the AERMOD Tool, EPA-454/B-20-008, EPA/OAQPS, February 2020

We used the latest version of AERSURFACE (20060), which accommodates more recent versions of available land-cover data available from USGS NLCD (years 2001, 2006, 2011, and 2016)—we used the 2011 data, as the Aquatic Park location has not undergone any major changes in land-cover characteristics over the past ten years. We used site-specific information and professional judgment to estimate the surface characteristics following the AERSURFACE methodology. We followed the local air district's (BAAQMD) recommendations for seasonal land-use settings consistent with a Mediterranean climate (dry, cool temperatures during the summer and fall and wet winters), with surface moisture conditions characterized using 2019 data derived from Berkeley monthly precipitation data (site 04069) compared with 30-year climatological means for determination of the Bowen Ratio.

Because AERMOD estimates of concentrations are particularly sensitive to the treatment of surface roughness (Long et al. 2004[25]), we used the full spatial resolution permitted by AERMOD for determining values of surface roughness by direction (separate values for each of twelve radial directions, with 30° spacing) using the default inverse weighting scheme (ZORAD).

A threshold hourly average horizontal wind speed of 0.2828 m/s was used in the model for "on-site" winds, although data collected by the sonic with speeds below this level were used in computing the hourly averages and turbulence. The 0.2828 m/s threshold wind speed (determined as $(2*0.2*0.2)^{0.5}$) is based on AERMOD's default values of 0.2 m/s for σ_v and σ_w (was not changed in this study). AERMOD's AERMET[26] meteorological pre-processor (version 21112) was used to integrate the on-site and off-site data. Because the processors operate on a minimum of 24-hours data, we processed an entire day at a time.

2.2.6.2. Regulatory Modeling

The regulatory approach follows EPA and local BAAQMD guidance on how to prepare meteorological data for use in air dispersion models, as would be the case in an EIS/EIR and/or a NEPA assessment. For a regulatory application, the most representative one-year set of high-quality hourly meteorological data available was from the 2019 Oakland Sewage Treatment Plant (STP) archived by the BAAQMD. It is located approximately 4.5-km south of the BAAQMD near-roadway air quality monitoring trailer, with similar orientation to the Bay. However, the observational dataset has frequent gaps in the wind measurements. An automated weather station with a similar land-water exposure is operated by the National Buoy Center (OKXC1) located at Pier 34, which is only 3.8-km southwest of the STP site, and its hourly wind and temperature data were used for substitutions, bringing the completeness up to 99.6%, 99.0%, 99.8%, and 93.1% for each quarter. Only 41 hours out of the 8,760 hours have key data missing for which AERMOD is unable to determine a concentration. This dataset is representative of a complete one-year dataset with available on-site turbulence parameter of σ_Θ, that would be used in a typical regulatory application. No adjustment for friction velocity was needed since we used meteorology data that have on-site measurements of turbulence.

Similar to the tracer experiment, we used EPA's AERSURFACE (v20060) with 2016 land-cover data and the geographical coordinates of the Oakland STP site to determine surface characteristics needed by AERMOD,

---

[25] Long, G.E.; Cordova, J.F.; Tanrikulu, S. An Analysis of AERMOD Sensitivity to Input Parameters in the San Francisco Bay Area. In Proceedings of the 13th Joint Conference on the Applications of Air Pollution Meteorology with the Air & Waste Management Association; A&WMA: Pittsburgh, PA, 2004; pp 203-206.

[26] AERMET User Guide for the AERMOD Meteorological Preprocessor (AERMET), EPA-454/B-21-004, EPA/OAQPS, April 2021

including twelve radial sectors within 1 km for estimation of surface roughness. The site is neither arid nor at an airport. The monthly seasonal profile used is shown in Table 2-2.

Table 2-2. Monthly Seasonal Profile

Months	Season
November, December, January	Late autumn after frost and harvest, or winter with no snow
February, March	Transitional spring with partial green coverage or short annuals
April, May, June, July	Midsummer with lush vegetation
August, September, October	Autumn with unharvested cropland

AERSURFACE was run separately specifying dry, average, and wet surface moisture, with the results later used to create composite surface characteristics by month. The rainfall data for the 30-year period ending 2019 were gathered and 30-year monthly averages computed for Berkeley, CA (cooperative # 040693). The monthly statistics were not used in the average if more than five days were missing in a given month. Data for Oct 2019 was missing, so the total representative data was the average of September and November. The months were then classified as “wet” if the precipitation total was in the upper 30th percentile, “dry” if the precipitation was in the lower 30th percentile, and “average” if precipitation was in the middle 40th percentile. Table 2-3 provides the summary information for the moisture classification of the region.

Table 2-3. Moisture Classification at the Berkeley Site

	Precipitation (in)											
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2019	avg	wet	wet	dry	wet	avg	avg	avg	wet	avg	dry	avg

The Oakland Sewer Treatment Plant (STP) onsite data were used for the ONSITE portion. Data from the CMAN buoy OKXC (9414776 - Oakland Berth 34) were used to fill-in for missing time periods for wind speed and direction. The MODIFY option in AERMET was used to allow Oakland Airport cloud cover data and temperature to be used when needed for the SURFACE portion of the processing (the CCVR SUB_CC and TEMP SUB_TT options). The missing flags were set as shown in Table 2-4 and Table 2-5.

Table 2-4. AERMET Single-value and Date/time Variable Descriptions and QA Values

AERMET Name	Description	Missing Indicator	Lower Bound	Type	Upper Bound
OSDY	Day	-9	1	<=	31

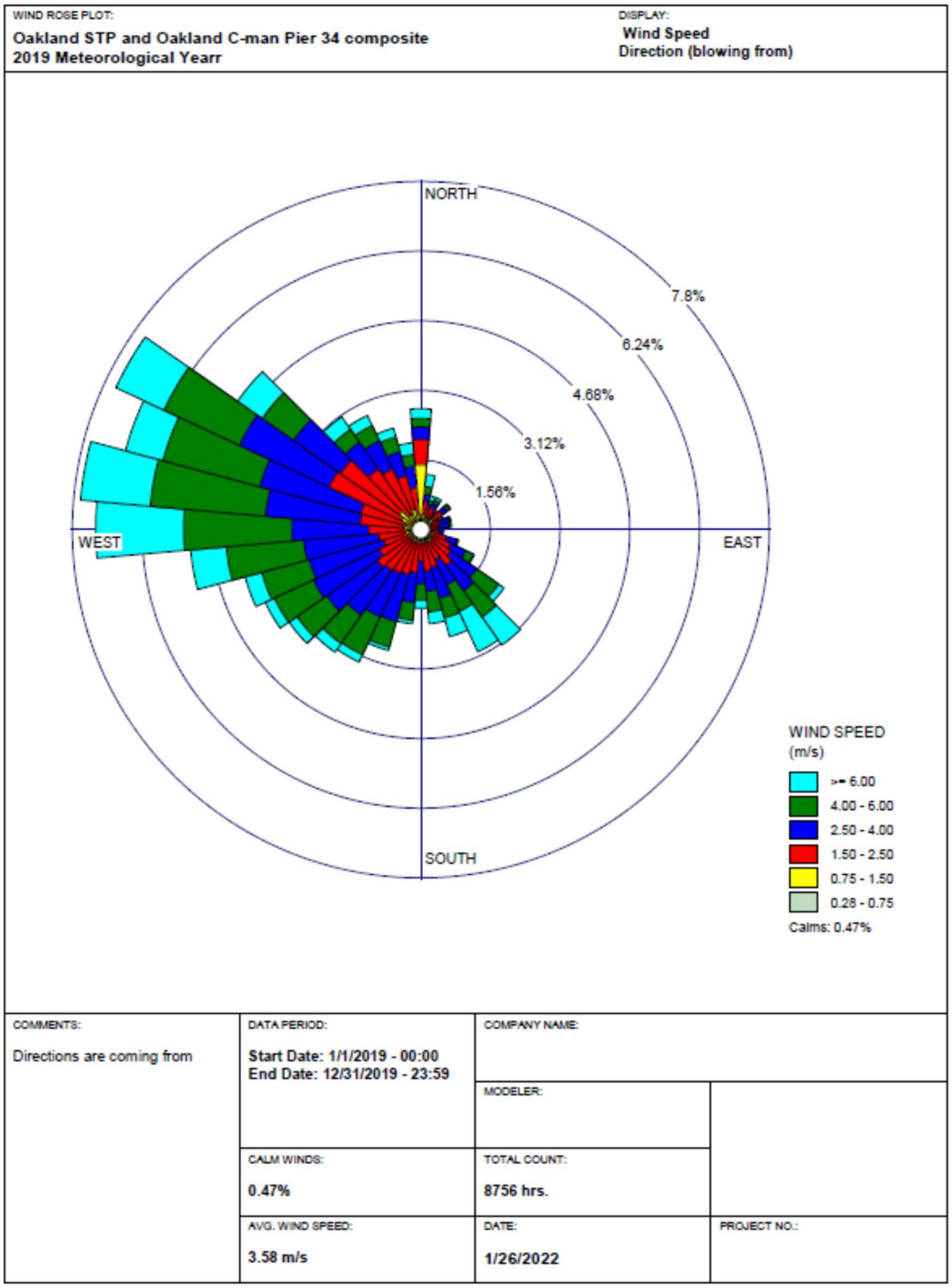
AERMET Name	Description	Missing Indicator	Lower Bound	Type	Upper Bound
OSMP	Month	-9	1	<=	12
OSYR	Year	-9	0	<=	99
OSHR	Hour	-9	0	<=	24

Table 2-5. AERMET Multi-value Variable Descriptions and QA Values

AERMET Name	Description	Units	Missing Indicator	Lower Bound	Type	Upper Bound
TT01	Temperature	deg C	99	-30	<	46
WS02	Wind Speed	meters / second	999	0	<	50
WD02	Wind Direction	degrees from north	999	0	<=	360
SA02	Std Dev Horiz Wind	degrees	999	0	<	104

Figure 2-11 shows the 2019 wind rose created from the AERMET-generated surface winds output. The wind rose shows wind direction from which the wind is coming, and the percentage of time coming from each direction. Winds were predominantly from the west and west-northwest, indicating predominate onshore flow from the San Francisco Bay.

Figure 2-11. 2019 Wind Rose



WIND ROSE PLOT:
Oakland STP and Oakland C-man Pier 34 composite
2019 Meteorological Yearr

DISPLAY:
Wind Speed
Direction (blowing from)

COMMENTS:
Directions are coming from

DATA PERIOD:
Start Date: 1/1/2019 - 00:00
End Date: 12/31/2019 - 23:59

COMPANY NAME:

MODELER:

CALM WINDS:
0.47%

TOTAL COUNT:
8756 hrs.

AVG. WIND SPEED:
3.58 m/s

DATE:
1/26/2022

PROJECT NO.:

2.2.6.3. Additional Meteorology Processing - AERSCREEN

We used AERSCREEN’s MAKEMET program to build a matrix of possible meteorological conditions. Temperature bounds are based on climatological extremes. For tracer and pollutant evaluation modeling, the temperature bounds were 42–88° F on May 3, 46–81° F on June 6, and 47–99° F on June 24 (from Western

Regional Climate Center climate summary for 1948–9 June 2016)[27]. For regulatory modeling, we used the all-time extreme values for the Berkeley area (25–104° F, from Western Regional Climate Center minimum and maximum monthly values for 1948–15 February 2022)[28].

For tracer and pollutant evaluation modeling, the minimum wind speed was 0.2828 m/s (the threshold hourly average windspeed used during AERMET processing of the data for modeling with refined models; see Section 2.2.6.1) and a 1.7-m anemometer height, while for regulatory modeling the minimum wind speed was set at 0.5 m/s with an anemometer height of 16.3 m, and we used the ADJ_U* option.

We selected the land use type of water, since it makes up 49% of the land cover within 1 km of the project site, and average moisture for the dominant surface profile; the model then uses internal lookup tables to assign seasonal values of surface roughness, albedo, and Bowen ratio. As with all models that allow it, we utilized the urban option.

2.2.6.4. Additional Meteorology Discussion—ADMS-Roads

To minimize differences in the meteorological inputs to the various dispersion models, we used the AERMET/AERSURFACE-generated meteorological variables for input to ADMS-Roads, wherever possible. This includes using values of surface roughness and albedo representative of the facility site (in addition to those representatives of the meteorological site). While ADMS-Roads offers the option to also specify surface roughness values in a grid encompassing the modeling domain (as part of the algorithms for complex terrain), the hourly surface roughness values are already based on inverse distance weighting of upwind surface roughness values (they are effectively a summary of upwind surface roughness); therefore, we did not use the gridded surface roughness parameterization.

We allowed the ADMS-Roads default option to calculate the minimum Monin-Obukhov length (L). This has the effect of limiting the stability of the modeled site, which can be important in urban locations where the heat island prevents high stability, but since all of our experiments were conducted during daylight hours this point is moot for the tracer experiment. As part of the dispersion calculation, a minimum turbulence value was applied, varying between 0.01 m/s to 0.2 m/s depending on the minimum value of L. While the user has the option to overwrite these minimum turbulence values, we have selected the default option of allowing the model to use minimum L to set minimum values of σ_θ and σ_w. While the model also allows customization of the Priestley-Taylor parameter, which affects modeled evapotranspiration and is used in calculations of surface heat flux and L, we used a default model value of 1.0 for that parameter.

Finally, ADMS-Roads does not generate results when wind speeds are below 0.75 m/s. When the model encounters such conditions, it sets the wind speed to 0.75 m/s and the wind direction to that associated with the next or previous hour with wind speed 0.75 m/s or above. If no such hour exists, then this hour is marked as invalid. This could lead in some cases to significant differences in maximum modeled concentrations under the regulatory option, since for AERMOD we used a minimum wind speed of 0.2828 m/s. For the regulatory

---

[27] Western Regional Climate Center, Oakland International Airport, 1 January 1948 to 9 June 2016, Period of Record Monthly Climate Summary,: https://wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca6335.

[28] Western Regional Climate Center, Oakland International Airport, 1 January 1948 to 15 February 2022, Minimum of Minimum Temperature (https://wrcc.dri.edu/WRCCWrappers.py?sodxtrmts+046335+por+por+mint+none+mmin+5+01+F) and Maximum of Maximum Temperature (https://wrcc.dri.edu/WRCCWrappers.py?sodxtrmts+046335+por+por+maxt+none+mmax+5+01+F).

modeling, a total of 97 hours of wind speeds were between 0.75 m/s and 0.2828 m/s. During the tracer and pollutants-based modeling studies, hourly average wind speeds were all above 0.75 m/s.

The following lists meteorological parameters that were extracted from AERMET/AERSURFACE-generated data and reformatted for input to ADMS-Roads:

- Extracted from AERMET's .SFC file:
- surface heat flux
- inverse of L
- albedo
- surface roughness
- wind speed
- wind direction
- boundary layer height (we used the maximum of mechanically and convective boundary layer height)
- surface temperature
- relative humidity
- Extract from AERMET's .PFL file:
- σ_Θ (this was only used for pollutant and tracer evaluations; for regulatory evaluation, we allowed the model to calculate the σ_Θ)

2.2.6.5. Additional Meteorology Discussion—CAL3QHC

We followed regulatory guidance by selecting a mixing height of 1,000 m, wind speed 1 m/s, wind directions every 10 degrees, and a stability class of D (neutral) for urban modeling.[29] We utilized a surface roughness coefficient of 0.01 cm, reflecting the largely open-water landscape in the project area.

2.2.7. Other Settings

For the regulatory evaluation, we utilized AERMOD's averaging settings to output 8-hour averages for CO (in addition to the model's minimum timestep of 1 hour) and to output 24-hour and annual averages for $PM_{2.5}$. The CO outputs of interest were the second-highest 1-hour and 8-hour averages, since the form of the CO National Ambient Air Quality Standard (NAAQS) is that the 1-hour concentration cannot exceed 35 ppm more than once per year and the 8-hour concentration cannot exceed 9 ppm more than once per year (i.e., "cannot exceed more than once" means that the second-highest value cannot be above the standard value). The $PM_{2.5}$ outputs of interest were the annual average (the form of the NAAQS is that the standard of 12.0 $\mu g/m^3$ cannot be exceeded) and the 8th-highest 24-hour average (the form of the NAAQS is that the 8th-highest concentration cannot exceed 35 $\mu g/m^3$ when using one year of modeling).

We used the default CAL3QHC persistence factor of [maximum 1-hour concentration * 0.7] to estimate the maximum 8-hour concentration. There are no AERSCREEN persistence factors for area sources.

---

[29] See Sections 4.2 and 6.4.1 of: User's Guide to CAL3QHC Version 2.0: A Modeling Methodology for Predicting Pollutant Concentrations Near Roadway Intersections. EPA-454/R-92-006, Revised (September 1995). https://gaftp.epa.gov/Air/aqmg/SCRAM/models/preferred/cal3qhc-r/cal3qhcug.pdf. Also see Section 4.7.1 of Guideline for Modeling Carbon Monoxide from Roadway Intersections. EPA-454/R-92-005 (November 1992). https://www.epa.gov/sites/default/files/2020-10/documents/coguide.pdf.

2.3. Model Evaluation Approach and Metrics

This section describes the metrics we employed to assess the performance of the dispersion models considered in this study quantitatively, graphically, and qualitatively.

2.3.1. Quantitative Measures

Quantitative comparisons included statistical measures between models and observations. We conducted statistical analyses on several groups of data, grouped as follows.

- By site, using all concentration pairs (paired in time and space)
- By site, using a subset of data with the highest observed concentrations (paired in time and space)
- By site, using a subset of data with the highest observed and predicted concentrations (unpaired), rank ordered (this is one metric used by EPA)

We used the first bulleted item to evaluate overall performance of the models at a specific site. We then used the second and third bulleted items to evaluate model performance in predicting peak concentrations at a given site. These latter two data groupings are the most important when using model results for determining exceedances of a short-term regulatory standard, while the first bullet provides insight as to whether the model is likely getting the right answer or not.

Statistical procedures developed by Cox and Tikvart provide an objective comparison of the performance between models and with observations (Cox and Tikvart, 1990[30]). To evaluate the models' performance, we use the following statistical measures in this research: average residual, absolute average residual, root mean square error (RMSE), Pearson correlation coefficient (r-squared), robust highest concentration (RHC), and fractional bias.

The average residual is the average difference between the observed and predicted values. This indicates the accuracy of model predictions and is calculated as:

$$Average\ Residual\ (\bar{r}) = \frac{\sum C_{observed} - C_{modeled}}{n} \qquad [1]$$

The absolute average residual provides a greater indication of true deviation of the model from the measured data, since positive and negative differences do not cancel out in the calculation of the average as is calculated as:

$$Absolute\ Average\ Residual = \frac{\sum \overline{C_{observed} - C_{modeled}}}{n} \qquad [2]$$

The Root Mean Squared Error or RMSE indicates the spread of the model accuracy relative to measured values. These measures provide indication of both the model's precision and accuracy. The RMSE helps determining if a model is likely to be a good fit and with what consistency this is likely to be true. This metric is calculated as:

$$RMSE = \sqrt{\frac{\sum (r_i - \bar{r})^2}{n}} \qquad [3]$$

---

[30] Cox, W.M., & Tikvart, J.A. (1990). A statistical procedure for determining the best performing air quality simulation model. Atmospheric Environment. Part A. General Topics, 24, 2387-2395.

The Pearson Correlation Coefficient is a statistical metric that measures the statistical relationship, or association, between two continuous variables. We use the Pearson correlation coefficient to evaluate the bivariate correlation between the modeled and measured data. This provides an initial estimation, prior to more in depth analyses, indicating whether or not the two data sets (modeled and measured) are aligned. The correlation coefficient is calculated as:

$$r = \frac{\sum(C_{Obs} - \overline{C_{Obs}})(C_{Modeled} - \overline{C_{Modeled}})}{\sqrt{\sum(C_{Obs} - \overline{C_{Obs}})^2 \sum(C_{Modeled} - \overline{C_{Modeled}})^2}} \qquad [4]$$

The Robust Highest Concentration (RHC), is a calculation designed to combat the high variability in peak concentrations typically found in measured data and often not well reflected in models. The RHC is calculated by factoring in the probability of high concentration values occurring along with the average concentration of these concentrations above a specifically set threshold (such as the 90th percentile). This allows for the evaluation of a single number which considers both the value and likelihood of these high concentrations that may otherwise skew (such as with a mean) or not be represented (such as with a median) in reported values. Representation of these peak concentrations is especially important for applications such as this, where the model should be flexible enough to predict concentrations that may hit peak concentration limits. This metric therefore provides a percentile that is meant to be representative of the upper tail/ peak concentrations. To calculate the RHC, the modeled values are first ranked and the RHC is determined using the Equation 5 below:

$$RHC = x_n + (\overline{x_{n-1}} - x_n)\ln\left(\frac{3n-1}{2}\right) \qquad [5]$$

where x_n is the nth largest value, and $\overline{x_{n-1}}$ is the average of the n-1 largest values. A value of n = 6 was used for this project.

Fractional bias of means is a measure of model performance detailed in Cox and Tikvart (1990)[31]. The fractional bias calculates a fraction using the RHC of both data sets (modeled and monitored) that is used to represent the overall fit of the model to the data with respect to these peak concentrations. This metric is a unitless value between -2.0 and 2.0, meant to predict model fit to the observed data, where a value closer to 0 indicates more accurate predictions of the model.

$$Fractional\ Bias = \frac{\sum C_{observed} - C_{modeled}}{\frac{1}{2}\sum C_{observed} + C_{modeled}} \qquad [6]$$

These latter two statistics (RHC and fractional bias) are the most succinct measures of model performance. Due to the importance of peak values in air quality data and air quality conformity determinations, these values are likely to be the most relevant in this assessment. They consider both peak concentrations and the spread of the data in the reporting of a single value that is easily comparable across several models.

We also use cumulative frequencies to describe the data. These values, which are reported at the highest concentration percentiles, as a sum of all concentrations up to the set cutoff, can give some detail on the distribution of the data and how heavily the upper tail of the distribution is weighing on the data. These values are being presented both graphically (in a histogram) and in a table.

---

[31] Cox, W.M., & Tikvart, J.A. (1990). A statistical procedure for determining the best performing air quality simulation model. Atmospheric Environment. Part A. General Topics, 24, 2387-2395.

The assessments described above focus on the tracer and pollutant evaluations. The regulatory evaluation relies on the design value comparisons and the ability of the model to predict the design values for near-road CO and $PM_{2.5}$.

2.3.2. Graphical Measures

Graphical measures, such as scatter plots of observed and predicted concentrations, paired in both time and space, are useful in providing insight into model behavior. We include quantile-quantile plots (Q-Q) plots, time-series plots, and scatter plots as part of our graphical evaluation metrics to compare model performances against each other.

Q-Q plots involve plotting the distribution quantile (minimum, maximum, and 25th-, 50th-, and 75th-percentile) values of both the observed and modeled data. By plotting these values on the same graph, we can easily evaluate the distributions of both sets of data to determine how well they correspond. If the model distribution is very different from observed distribution, it indicates that model is not reliable. The ideal Q-Q plot is a 1:1 line from one corner of the graph to the other which is not very common due to inherent uncertainties in modeled and observed data, but the closer the plot is to linear, the more normal the data distribution. Most of the variation in this plot is typically found at the extremes of the data range, though mid-range departures from the line can indicate overall data skew.

Time-series plots graph the modeled (and measured) data over time to evaluate any time-dependent trends that may be present and need to be considered in the interpretation of the model. For pollutants modeling, the resolution depends on the averaging periods in the graph. We use this graphical representation to evaluate how well the modeled data follows the measured data over time.

Scatter plots display the measured data against the modeled, pairing the data over space and time, to determine relationships that may be otherwise masked. This is intended to elucidate trends in the data that are specific to a single location or a specific time-period.

2.3.3. Qualitative "Ease of Use" and Utility Evaluations

In addition to the quantitative measures of model performance, we evaluate qualitative model properties. These include ease of use, data input requirements, first principal differences between the models, accessibility, where the model is being maintained and improvements are supported, functional capabilities for use in regulatory applications, and limitations to the user of the model in regulatory applications. Each of these measures of the model aid in the model's overall functionality and use in the regulatory environment. By evaluating which models are widely available and easily accessible, data requirements, and level of expertise needed to run we can reliably make recommendations about the applicability for general use.

We base this evaluation on available model features relevant to the setting studied here, including:

- **User Interface:** We evaluate the models user interface (UI) including features that facilitate the model setup and the preparation of model input data. Additionally, we describe features available through the model to visualize and make sense of the output data.
- **Ease of Use**: While being a subjective metric, we describe the challenges in setting up and running the dispersion model, all the way to extracting and post processing the data. While user experience can vary depending on their skillset, we highlight areas where we felt further improvement can enhance users

experience and make the model more accessible to both senior professional as well as intermediate air quality modelers.

- **Model Transparency**: We evaluate the availability of the model source codes, easy to understand users' guide, technical documentations, and case studies for understanding the model formulation, preparation of model inputs, proper setup of the model, describing the model outputs, and assessing the methodological differences across various models.
- **Runtime:** We provide insights on the time for running the model under the three evaluation scenario studies in this project (i.e., tracer, pollutant, and regulatory evaluation).
- **Output resolution:** We describe the resolution of the output concentration data from the model, and whether the model has the ability to post-process the modeled concentrations to produce air quality metric needed for regulatory evaluation and conformity determinations.
- **Transportation Conformity and Hot-Spot Analysis**: We provide a qualitative discussion on the suitability of the model for determining the air quality impact of vehicular related emissions. For example, whether the emissions inputs into the model requires significant amount of pre-processing, or whether the model has features that can facilitate determining source geometry for roadways.
- **Suitability of boundary-layer parameterization**: We describe the inputs that the model takes to parameterize the boundary layer conditions, and the sensitivity of the model in reflecting those parameters.
- **Initial Plume Dispersion:** We describe how the model accounts for traffic induced turbulence in characterizing the initial plume dispersion,
- **Model maintenance:** We consider who maintains and updates the model and how frequently.
- **Cost:** We describe whether these models are publicly available at no cost, or if there is a cost (or subscription) associated with acquiring and subsequent use

3. Observational Dataset

The tracer experiment location is along Interstate 80 in Berkeley, CA. This area includes a 10-lane freeway (five in each direction), separated only by a concrete jersey barrier median. Caltrans reported this section of freeway had an AADT of 280,400 (pre-pandemic, 2017) and 4.8% truck fraction (13,460 AADT). It is located near the eastern shoreline of San Francisco Bay. This section of the I-80 freeway has high traffic volume that runs between the San Francisco-Oakland Bay Bridge to just south of the City of Richmond. We focused along the portion of I-80 between University Ave and CA Highway 13 (Ashby Ave.). The BAAQMD (the local air agency) reports that this site has the fifth highest Fleet Equivalent AADT (FE-AADT) in the San Francisco Bay Area and is ranked first for highest traffic congestion by the Bay Area's Metropolitan Transportation Commission.

Table 1-2 summarized the datasets used in this evaluation. The Task 6 report[32] described the Berkeley Freeway Tracer Experiment, including the three tracers and both tracer and co-located air pollutant observations. Please see that report for details on the experimental approach and results, including traffic observations collected during the experiment. Traffic data as used here is described In Section 4.1. For reference, we have also included Figures 1 and 2 from the Task 6 report here as Figure 3-1 and Figure 3-2 to emphasize source and receptor locations.

In addition to the I-80 Freeway Tracer Experiment, we rely on co-located measurements of pollutants available from the BAAQMD. This site is located at a near-road SLAMS air quality monitoring trailer that records continuous measurements of: NO, NO_2, CO, continuous (hourly) $PM_{2.5}$, black carbon (BC), and ultrafine PM. It includes Federal Equivalent Method (FEM) for $PM_{2.5}$. Continuous monitoring began at the site on July 1, 2016. The distance from the road to probe inlet is 8 meters. The BAAQMD operates and maintains the trailer. Table 3-1 summarizes the air quality monitor station information. Figure 3-3 shows its location, as viewed from the freeway.

The site does not include on-site meteorology nor traffic measurements, but Caltrans operates a suite of PeMS measurements throughout the area continuously monitoring traffic and speed. The Task 6 report described the PeMS and other available traffic data. Section 2.2.6 described the on-site and nearby meteorological observations used here.

---

[32] NCHRP 25-55 Assessment of Regulatory Air Pollution Dispersion Models to Quantify the Impacts of Transportation Sector Emissions: Task 6: Initial Report on Tracer Experiments. Chelsea Preble, Thomas Kirchstetter, Michael Sohn, Edward Carr, Seth Hartley, Elliott Wezerek, Moses Wilson, March 2022. Included here as Appendix B.

Figure 3-1. The study area along I-80 in Berkeley, CA, with greater detail for the boxed area indicated in (a) shown in subpanel (b). The mobile platform turnaround points are marked by the red circles with a white X. The tracer release zone is shown by the green line bounded by the open green circles. The two experimental areas noted by the two boxes in (b). Source: Task 6 Report. (Appendix B)

Figure 3-2. The two experimental areas. (a) The North Experiment Area, where 60-min integrated samples were collected at three sites (navy diamonds) within the dark purple box to the right of I-80. Downwind air quality (blue star) was measured at the BAAQMD monitoring station shown within that dark purple box. Upwind meteorological (blue circle) and air quality (blue star) sampling location shown in the light blue box to the left of I-80. (b) The South Experiment Area, where 10-min integrated samples were collected at six sites (purple circles). Source: Task 6 Report. (Appendix B)

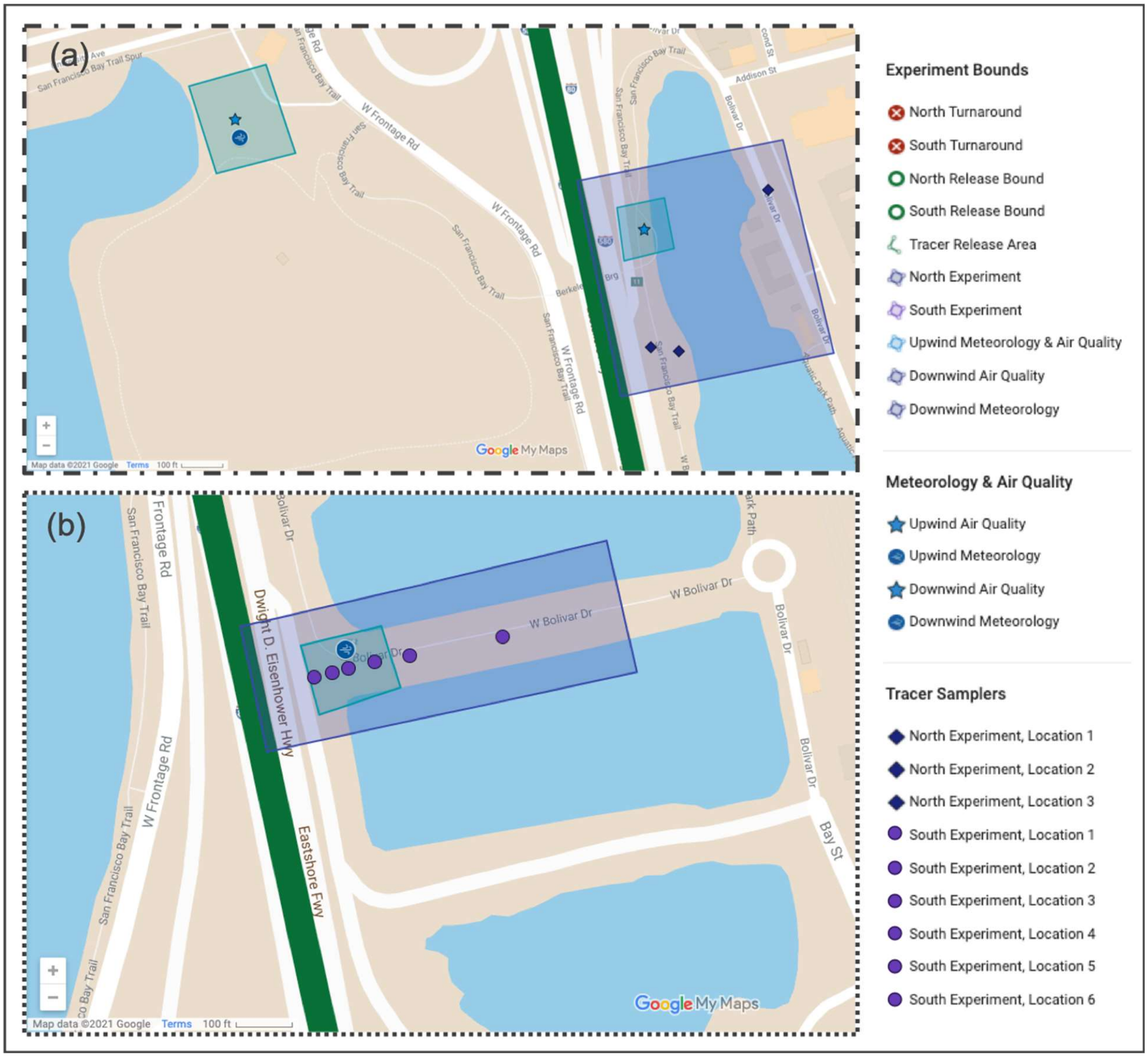
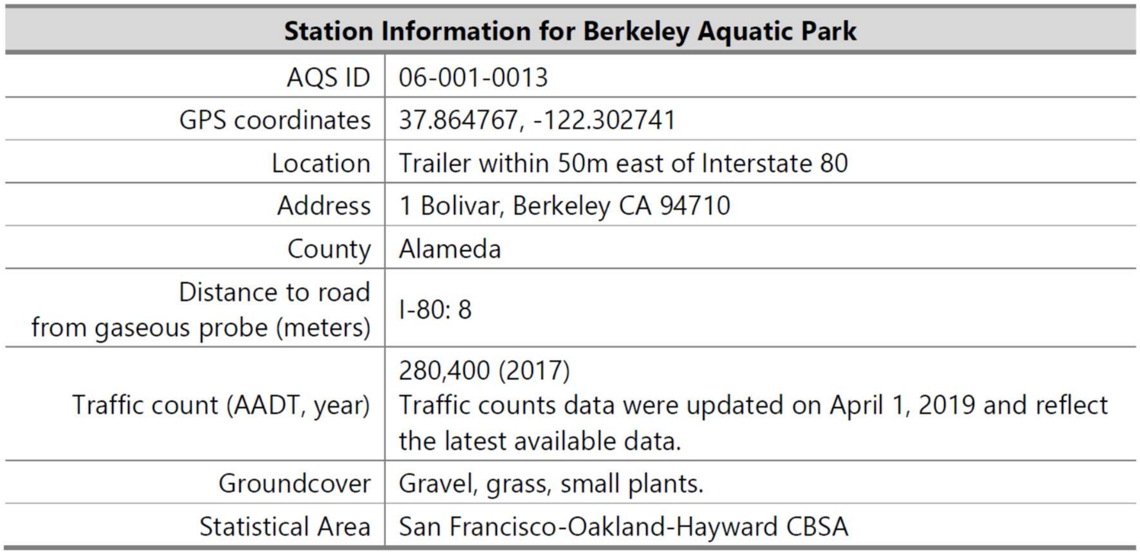


Table 3-1. Berkeley Near Road Monitor Specifics[33]



Station Information for Berkeley Aquatic Park	
AQS ID	06-001-0013
GPS coordinates	37.864767, -122.302741
Location	Trailer within 50m east of Interstate 80
Address	1 Bolivar, Berkeley CA 94710
County	Alameda
Distance to road from gaseous probe (meters)	I-80: 8
Traffic count (AADT, year)	280,400 (2017) Traffic counts data were updated on April 1, 2019 and reflect the latest available data.
Groundcover	Gravel, grass, small plants.
Statistical Area	San Francisco-Oakland-Hayward CBSA

Figure 3-3. Google Street View Image of the Berkeley Near Road Monitor[34]

---

[33] BAAQMD, 2018 Air Monitoring Network Plan, July 1, 2019.

[34] Note that the pedestrian overpass appears above the station in this figure due to camera angle, but is, in fact, east of the monitoring trailer. This provides unobstructed flow from the edge of the freeway (hence the site selection by the local air quality management district).

4. Emissions modeling

4.1. Traffic Data

To estimate the emission rates of various pollutants (e.g., CO, PM, BC) used for evaluating the regulatory applications of the air quality models, we relied on various sources of traffic data that were used as inputs into the California Air Resources Board's (CARB) emission factor model (EMFAC). These traffic data included vehicle counts collected by Wiltec, a professional traffic engineering firm that specializes in the collection and analysis of traffic, transportation, transit and parking data; the traffic volume extracted from Caltrans Performance Measurement System (PeMS); and truck traffic counts collected by Caltrans Weigh-in-Motion (WIM) sensors. These data sources are further described in this section.

4.1.1. Wiltec traffic survey

To collect accurate traffic data by various vehicle classifications, ICF subcontracted Wiltec Inc. to conduct manual observation surveys of vehicular traffic flow in both directions of the I-80 freeway near the Aquatic Park Berkeley. The surveys were conducted between midnight Saturday and midnight Sunday, June 6, 2021. To collect the traffic count data, Wiltec installed video traffic monitors to collect video footage for a 24-hour period. The recording cameras were located to allow views of vehicle axles and for safety of the installation location with regards to passing vehicles and the safety of the video equipment. During advance field reconnaissance Wiltec determined that they could not install any of the equipment on the pedestrian overpass because it would have hung over the freeway below. Instead, Wiltec installed the equipment on other poles located on the ramp leading up to the overpass. The view from this location is shown in Figure 4-1.

Figure 4-1. View from Wiltec traffic cameras on the I-80 E, with the BAAQMD Trailer in the Foreground

For the west bound direction, the ramp on the other side led away from the overpass and Wiltec could not install the equipment there and instead the camera was installed on a light pole down near the Ashby off-ramp. Considering that there were no on or off-ramp between these two locations, the traffic volumes were expected to be the same. The view from this location is shown in Figure 4-2

Figure 4-2. View from Wiltec traffic cameras on the I-80 W

Following the collection of video recordings, these files were then reviewed by trained staff from which the traffic data were then extracted. The outcome of this effort was 24-hour vehicle counts of the total freeway in both directions with results broken down in 15-minute increments according to the standard vehicle classification format as shown in Figure 4-3. Figure 4-4 illustrates the traffic counts collected by Wiltec on June 6, 2021 on both west- and east-bound of I-80 near N. Aquatic Park.

Figure 4-3. FHWA vehicle classifications

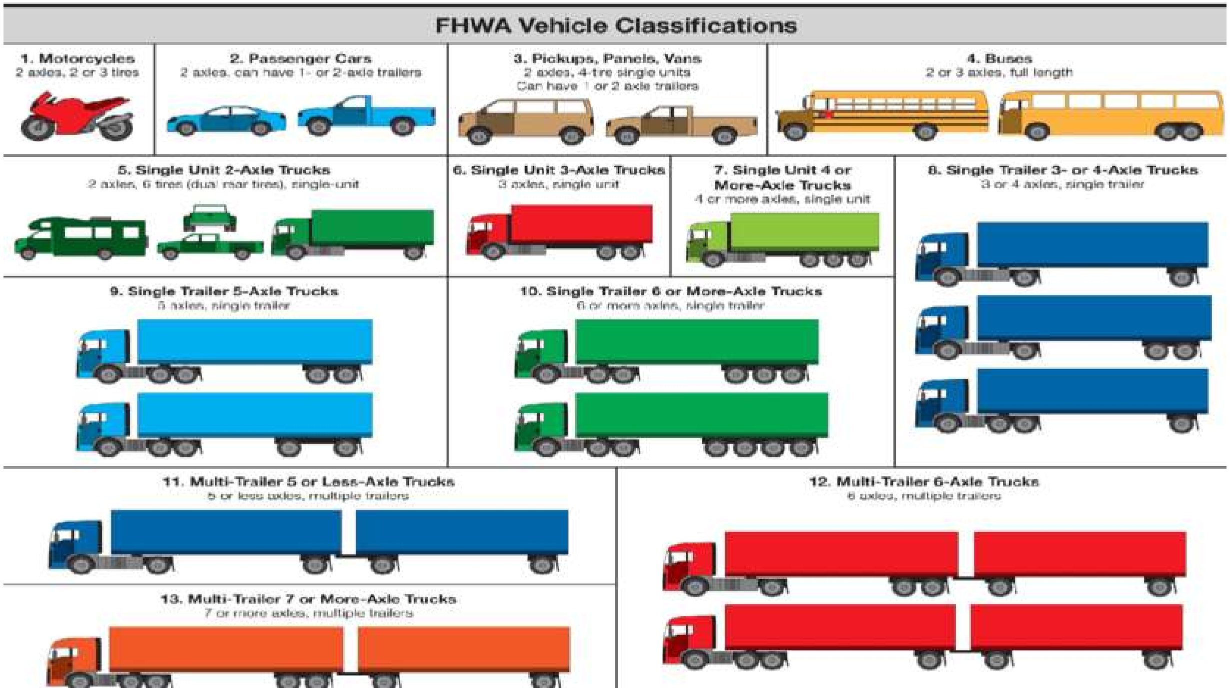
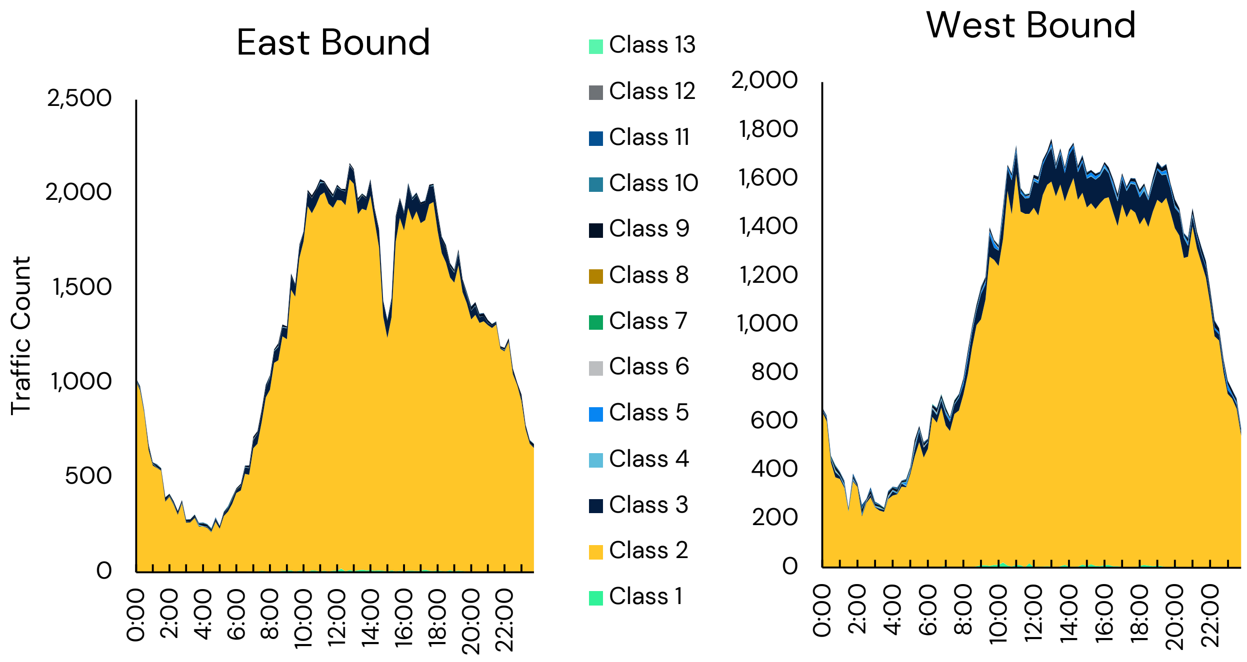


Figure 4-4. Wiltec Collected Traffic Count Data



4.1.2. PeMS

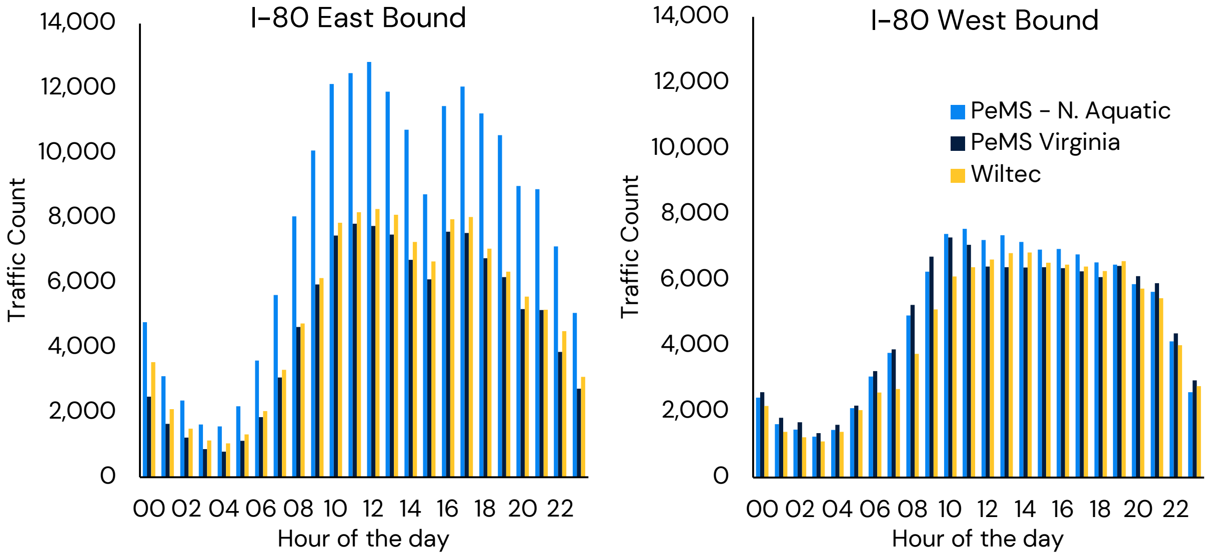
Traffic volume (vehicles per hour) were collected from Caltrans' online PeMS[35]. PeMS allows users to query measured traffic data from over 39,000 individual detectors for more than 10 years of historical data. These sensors span the freeway system across all major metropolitan areas of the State of California. PeMS can output vehicle counts, vehicle miles traveled (VMT), vehicle hours traveled (VHT), flow, occupancy, speed, truck flow, truck proportion (percent), Q (VMT/VHT), truck VMT, truck VHT, and other metrics on an hourly (and even every 5 min) interval by freeway, direction (i.e., I-80-E), and lane from individual sensors.

Multiple datasets are available in PeMS, including total VMT, total VHT, truck VMT, and truck VHT. Primary information that were gathered from the PeMS data included vehicle counts, and vehicle speed from sensor 401198, the North Aquatic Park Sensor, and the North of University (Virginia St) locations (403192 and 403192). These sensors are located directly in the observation zone. The North of University sensors are just north of the measurement locations, but within the driving loop. Upon initial evaluations of traffic volume data extracted from these sensors, we noticed that the traffic volume data from sensor 401198 located in the North Aquatic Park are unexpectedly higher than those collected by Wiltec Inc., and the ones from the Virginia St sensors. As shown in Figure 4-5, vehicle count data from the North Aquatic sensor on I-80E were consistently higher, by almost 50 percent in certain hours, than traffic counts from Virginia St. and those collected by Wiltec.

---

[35] https://merritt.cdlib.org/d/ark:%2F13030%2Fm5154j4b/2/producer%2FPRR-2009-25.pdf

Figure 4-5. Traffic count data comparison between N. Aquatic, Virginia St. PeMS sensors and Wiltec traffic survey

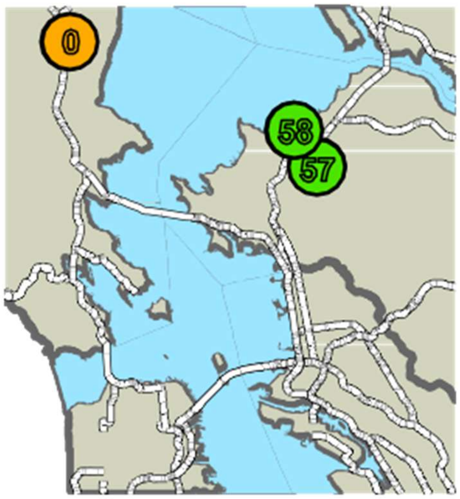


Considering that Wiltec data were in close agreement with the Virginia St. PeMS data, we decided to use the Virginia St. PeMS for modeling data where Wiltec data was not available. Also as shown in Figure 4-5, the issue with N. Aquatic sensor was only for the east bound sensor, and data extracted from the West Bound sensor was in close agreements with the other two data sources.

4.1.3. WIM

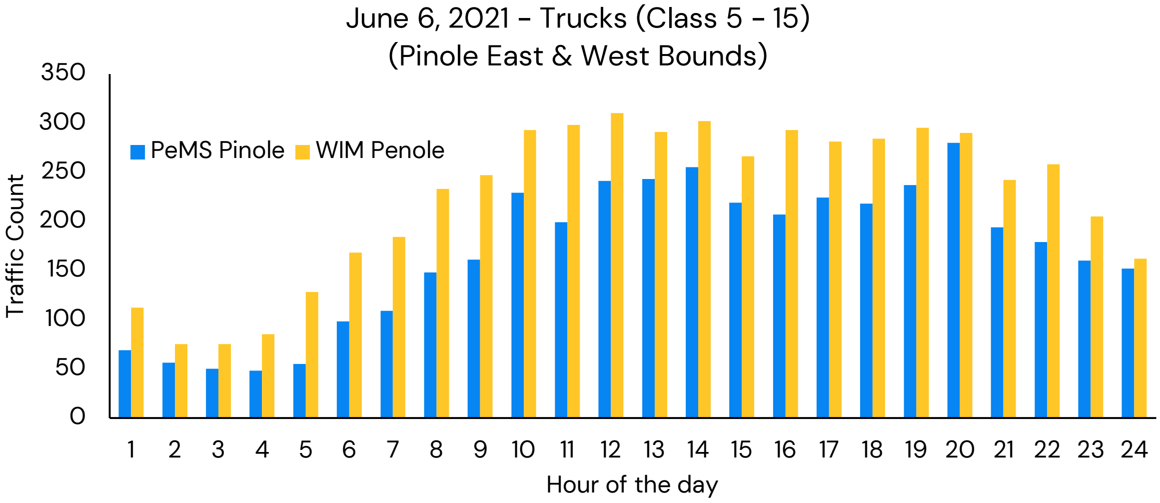
Caltrans also collects WIM data, with one set of detectors along I-80 approximately 9 road miles eastbound (north) of the proposed monitoring site in the city of Pinole. These are shown as sites 57 and 58 in Figure 9. WIM provides 24-hour traffic data including axle weights and gross weight, axle spacing, vehicle classification, speed, and vehicle overall length. Considering that the WIM data is north of the I-80/I-580 split outside of the study area, it was only used mainly to examine the accuracy of PeMS data for determining the truck counts and separating those from passenger vehicle counts, and it is not used in modeling.

Figure 4-6. Caltrans District 4 Weigh in Motion Data Sites in the Proposed Study Area



Comparing the vehicle counts between PeMS and WIM stations in Pinole, we have seen that while the total vehicle counts are very consistent, it was difficult to find a perfect match for truck count. WIM data are being collected across 15 different classes of vehicles whereas PeMS reports vehicle counts over two classes. It is not very clear on how the definition of trucks from PeMS can be translated to truck definition for WIM stations. For this exercise, ICF assumed that truck classification from PeMS is associated with vehicles classes of 5 through 15 from WIM data. Using that mapping, ICF compared the truck count data from PeMS to WIM as shown in Figure 4-7. As shown, PeMS truck count are on average 25 percent lower than class 5 – 15 truck counts from WIM. This could be partly due to the fact that some of the pick-up trucks that are captures by WIM as being Class 5 vehicles might be classified as passenger vehicles in the PeMS data. Considering that the temporal profile of both PeMS and WIM for truck counts are closely correlated, ICF decided to continue using PeMS for emissions modeling, except for June 6, 2021 where Wiltec data was available.

Figure 4-7. WIM vs. PeMS truck counts at Pinole Station



4.2. Emission Model

The EMission FACtors model[36] (EMFAC2021) was used to determine the emission rates of CO, $PM_{2.5}$, and BC for use in regulatory evaluation of air quality models. EMFAC2021 is California's official on-road mobile source emission inventory model for use in California and recommended by U.S. EPA for use in regional conformity determination as well as project level/hot spot analysis. The model reflects latest understanding of fleet composition based on the California vehicle registration database, activity data from California Smog Check program, and emission rates from years of laboratory and real-world emission testing conducted by CARB. Additionally, the model reflects California's unique policies for controlling emissions from on-road vehicles (e.g., EMFAC reflect the California Truck & Bus regulation). The following section describe the input data, as well as the methodology that were utilized to covert emission data extracted from EMFAC into emission rate data that can be used as inputs into air quality models.

4.3. Model Inputs for EMFAC

Existing sources of monitored traffic data provide all the traffic information required for the air dispersion modeling conducted here but with some parameters not as well-known as others. Table 4-1 presents an overview of the parameters needed for EMFAC2021 to determine the hourly emission factors of CO, $PM_{2.5}$, and BC.

Table 4-1. Emission Factor Model Variables and Inputs.

Model Variable	Options	Input	Needs Traffic Data?	If Used, Source of Traffic Data
Region type	Statewide, air basin, air district, MPO, county, sub-area	County	No	
Region	-	Alameda County	No	
Calendar year	2000 to 2050	2021	No	
Vehicle category scheme	EMFAC202x EMFAC2011 EMFAC2007	EMFAC2007*	No	WilteC Survey Virginia PeMS
Model year	Aggregate, select, range	Aggregate	No	
Speed	5 to 90 mph in 5-mph increments	Select all	Yes	PeMS
Fuel	Aggregate, by fuel	Aggregate	No	
Season	Summer/Winter/Annual	Summer for pollutant evaluation Annual for regulatory evaluation	No	

* This is a naming convention that EMFAC model uses for different sets of vehicles categories, and it is not representative of the version of the model used in emissions calculations.

The output of the EMFAC2021 simulations is a single, flat file summary of pollutant emissions (short tons per day), along with vehicle population, and VMT that are used to estimate emission rates in unit of g/veh/sec.

---

[36] https://arb.ca.gov/emfac/

These emission factors are then combined with the vehicle count data from Wiltec surveys and PeMS to determine the emission factor (g/s) for the regulatory air dispersion modeling.

4.4. Traffic Emission Modeling: Methodology

To determine the emission factors that can be used for air quality dispersion modeling, we first calculated an aggregated distance-based emission rates per vehicles in the unit of grams per vehicle-miles. As described earlier, EMFAC2021 provides total emissions (in unit of short tons per day) as well as the VMT by:

- Calendar year
- Vehicle class
- Fuel type
- Speed

Using these emissions data, we first calculated aggregated emission factors for the vehicle categories where traffic data was available. This means six categories of motorcycles, light duty vehicles, light heavy duty trucks, buses, medium-heavy, and heavy-heavy duty trucks for June 6, 2021 where Wiltec survey data was available, and two categories of light- and heavy-duty vehicles for other days where we relied solely on PeMS data to determine the emission rates. To aggregate the emission factors at these categories, we used the vehicle class mapping as shown in Table 4-2 .

Table 4-2. Vehicle Class Mapping for Emission Rate Aggregation

EMFAC2007 Vehicle Class	Vehicle Class - Wiltec	Vehicle Class - PeMS
MCY	Motorcycles	Light Duty Vehicles
LDA	Light Duty Vehicles	
LDT1		
LDT2		
MDV		
LHDT1	Light Heavy Trucks	Heavy Duty Vehicles
LHDT2		
MH	Motorhome	
MHDT	Medium-Heavy Truck	
HHDT	Heavy-Heavy Trucks	
OBUS	Buses	
SBUS		
UBUS		

Once the emission rates are aggregated by speed, the traffic data (i.e., vehicle counts at every 15 min intervals) along with traffic speed (at 15 min intervals) are then used to calculate an aggregated distance-based emission factor at each 15 min interval through the following equation. This is also illustrated in Table 4-3.

$$\text{Aggregated Emission Rate}\left(\frac{\text{g}}{\text{veh.mi}}\right)=\frac{\sum_{\text{Vehicle Class}}\text{Emission Rate (Vehicle Class,Speed)}\times\text{Vehicle Counts (Vehicle Class)}}{\sum\text{Vehicle Counts}}\qquad[7]$$

Table 4-3. Example Emission Rate Aggregation Using Wiltec Traffic Counts and PeMS Vehicle Speed data

Vehicle Class	Speed*	CO (g/mi)	Vehicle Counts
Motorcycles	40 mph	13.484	11
Light Duty Vehicles		1.315	2067
Light Heavy Trucks		0.560	73
Motorhome		1.907	3
Medium-Heavy Truck		0.331	1
Heavy-Heavy Trucks		0.926	10
Aggregated Emission Rate (g/veh/mi)			1.350

* We assumed at every given time interval, the vehicle speed of different categories are the same

Upon calculation of the aggregated emission rates, they were then multiplied by the total traffic count (in unit of vehicles per second) and the length of the freeway (in unit of miles) to calculate emission rates (in unit of grams per second).

$$\text{Emission Factor}\left(\frac{\text{g}}{\text{s}}\right) = \text{Aggregated Emission Rate}\left(\frac{\text{g}}{\text{veh.mi}}\right) \times \text{Vehicle Counts}\left(\frac{\text{Veh}}{\text{s}}\right) \times \text{Source Distance (mi)} \qquad [8]$$

4.5. Speciation

EMFAC model does not speciate particulate matter emissions into organic carbon (OC) versus elemental carbon (EC). We used CARB's speciation profiles[37] for on-road mobile sources to apportion the PM10 emission factors extracted from EMFAC into EC and OC emission factors. Table 4-4 shows these speciation profiles, and the suggested fraction of elemental carbon. Each number represents the portion of emissions from the specific profile that are considered to be BC.

Table 4-4. EC Fraction of PM10 – CARB's PM Speciation Profiles

Profile Description	EC Fraction
2021 Heavy-Duty Diesel Truck-Cruise (2016 Update)	19.9%
2021 School Bus-Transient (2016 Update)	7.8%
2021 Transit Bus-Transient (2016 Update)	14.4%
Diesel Vehicle Exhaust	26.4%
Gasoline Vehicle Equipped with Catalytic Converter	25.2%
Brake Wear	2.6%
Tire Wear	22.0%
Running Exhaust from CNG -TWC on UDDS Cycle	20.3%

These were then mapped to the appropriate EMFAC2007 vehicle class and fuel types to calculate the BC emission factors. Table 4-5 shows the resulting EC fractions.

---

[37] https://ww2.arb.ca.gov/speciation-profiles-used-carb-modeling

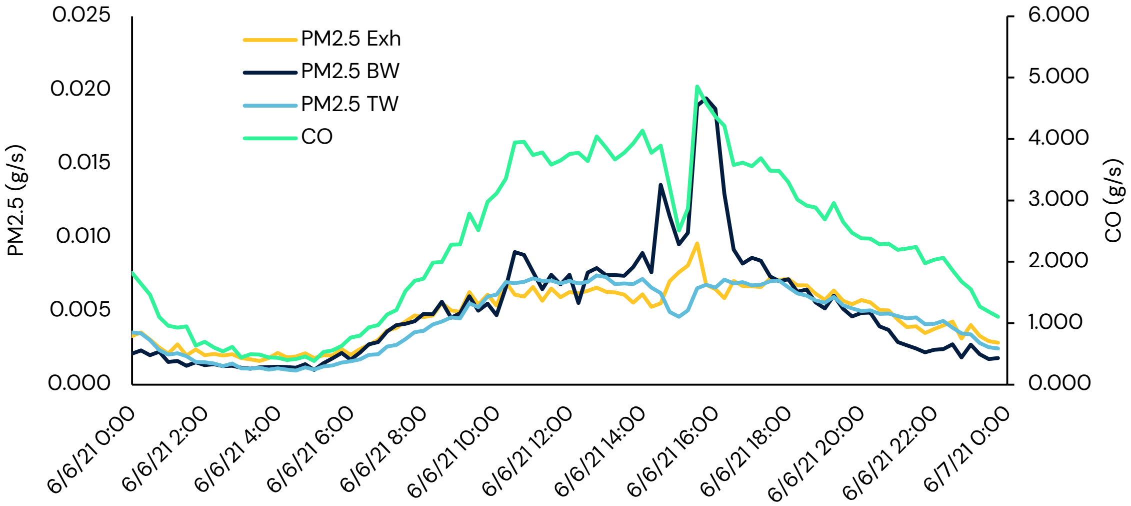
Table 4-5. EC Fraction as a Function of Vehicle Class and Fuel Type

Vehicle Category	Fuel	Exhaust	Brake Wear	Tire Wear
LDA	Electricity	N/A	2.6%	22.0%
LDT1				
LDT2				
MDV				
UBUS				
SBUS	Diesel	7.8%		
UBUS		14.4%		
HHDT		19.9%		
MHDT				
OBUS				
HHDT	Natural Gas	20.3%		
MHDT				
OBUS				
SBUS				
UBUS				
HHDT	Gasoline	25.2%		
LDA	Gasoline			
LDA	Plug-in Hybrid			
LDT1	Gasoline			
LDT1	Plug-in Hybrid			
LDT2	Gasoline			
LDT2	Plug-in Hybrid			
LHDT1	Gasoline			
LHDT2	Gasoline			
MCY	Gasoline			
MDV	Gasoline			
MDV	Plug-in Hybrid			
MH	Gasoline			
MHDT	Gasoline			
OBUS	Gasoline			
SBUS	Gasoline			
UBUS	Gasoline			
LDA	Diesel	26.4%		
LDT1	Diesel			
LDT2	Diesel			
LHDT1	Diesel			
LHDT2	Diesel			
MDV	Diesel			
MH	Diesel			

4.6. Emission Rate Results

Utilizing the methodology described, we calculated emission factors (g/s) for every 15-min interval on May 3, June 6, and June 24, 2021. Figure 4-8 illustrates the temporal variation of the calculated emission factors due to changes in traffic counts, vehicle speeds, and car/truck distribution on June 6, 2021. Please note that the temporal variation would be different for May 3 and June 24. June 6, 2021 was a weekend, whereas May 3, and June 24, were representing a weekday operation. The reduction of emission factors in the early hours of the day (between 0:00 – 03:00) is consistent with reduction in traffic volume within those hours. Also, the sudden drop in CO emission factor between 14:00 – 15:00 is mainly caused by a sudden increase in traffic congestion which reduced the traffic flow by almost 30 percent and the traffic speed from 53 mph to 13 mph. This congestion resulted in a sudden and transient change to the emission factors. Note that during this period of congestion, the car-truck share also changed (percent truck increased from 3.5 percent to almost 7 percent at 15:00) which increased the average $PM_{2.5}$ exhaust emission factor per vehicle. Therefore, the drop in overall traffic volume was offset by the increase in average vehicle $PM_{2.5}$ exhaust emission factor (g/s/veh) and did not let the overall $PM_{2.5}$ exhaust emission factor to change abruptly.

Figure 4-8. Emission Factors of CO, PM (exhaust, brake, and tire wear) on I-80 E



The resultant emissions factors (in unit of grams per second) for the three experimental days are shown in Table 4-6. Please note that the emission factors in Table 4-6 include emissions on both the east as well as the west bound of the freeway, whereas data illustrated in Figure 4-8 is only for the east bound. Additionally, the $PM_{2.5}$ emission factors shown in Table 4-6 includes emissions from exhaust, brake and tire wear.

Table 4-6. Emission Factors for the nine experiments conducted between May 3 and June 24, 2021

Date & Hour	CO (g/s)	BC (g/s)	$PM_{2.5}$ (g/s)	HD Vehicle Counts	LDV Vehicle Counts	Speed (mph)
05/03 08	6.773	0.011	0.077	819	12867	50
06/06 08	3.540	0.006	0.040	611	7854	67
06/06 09	4.895	0.008	0.049	701	10517	65

06/06 10	6.577	0.009	0.060	807	13129	57
06/06 11	7.091	0.009	0.062	673	13878	50
06/24 08	6.367	0.011	0.075	826	12390	52
06/24 09	6.844	0.011	0.079	822	12927	49
06/24 10	6.567	0.011	0.076	801	12450	50
06/24 11	6.336	0.010	0.071	654	11624	47

In addition to emission factors for three experiment days, we also utilized the same methodology to calculate emission factor of CO, $PM_{2.5}$, and BC for all 8,760 hours in 2019. We used the annual setting of EMFAC2021 for calculating emission factors in 2019.

4.7. Road Dust Emission Factors

While EMFAC model provides emissions factors for all vehicle related emissions such as exhaust, evaporative, brake, and tire-wear emissions, the model does not have the ability to provide emission rates for re-entrained road dust. This category includes emissions of particulate matter (PM) from resuspended road surface material that are entrained by vehicular travel on public and industrial paved roads. For the purpose of this project, we used CARB's methodology which is based on the Final Section of the Fifth Edition, Volume 1, Chapter 13.2.1, of the U.S. EPA's AP-42 document (January 2011). The AP-42 emission factor equation used to estimate paved road dust emissions for PM10 is provided below, followed by a description of the inputs to the equation:

$$E=[k(SL)^{0.91} \text{ x } (W)^{1.02}] \text{ x } (1-P/4N) \qquad [9]$$

Where:

- E=the particulate emission factor in units of pounds of particulate matter per VMT,
- k=the U.S. EPA AP-42 particle size multiplier (PM10=0.0022 lb/VMT)
- SL=the roadway-specific silt loading in grams/square meter (g/m2)
- W=the average weight of vehicles traveling the road (California statewide default=2.4 tons)
- P=number of "wet" days, when at least one site per county received at least 0.01 inch of precipitation during the annual averaging period, and
- N=the number of days in the annual averaging period (default=365).

Assuming a silt loading of 0.015 (g/m^2) for freeways and 64 days of annual rainfall in Alameda County, CARB estimated an emission factor of 112.4 lbs. PM10 per million miles as shown in Table 4-7. Utilizing this emission factor, we calculated respective $PM_{2.5}$ road dust emission factor assuming a size fraction of 0.15 for $PM_{2.5}$.

Table 4-7. Silt Loadings (SL, g/m2) and PM10 Emission Factors (EF; lbs PM10/106 VMT) for Alameda County Entrained Paved Road Dust

Air Basin	County	Air District	Freeway		Ave. Vehicle Weight (tons)
			SL	EF	
SF	Alameda	BA	0.015	112.4	2.4

4.8. Emission Factors for On- & Off-Ramps

The methodology presented earlier was mainly used to determine the emission factors for the I-80 freeway. To determine the emission rates for on- and off-ramps, we assumed same fleet composition and average gram per mile emission factors apply to on- and off-ramps. We also conduct an analysis, illustrated in Figure 4-9, using the available data for on- and off-ramps near the study region, and demonstrated that on average the traffic on ramps is approximately 8 percent (by volume) of the main freeway. Therefore, to determine the emission factors for ramps, we multiplied the gram per second emission rates for main freeway (per unit of miles of the freeway) by 0.08 to get per unit mile emission factors for on-ramps. Knowing the length of the ramps, we were able to convert those to the grams per second emission rates.

Figure 4-9. Traffic counts on mainline as well as on- and off-ramp near University Ave.

4.9. Tracer Study Release Rates

Appendix B defines and provides details on the three tracers. While the emission factors described in this section are used for the pollutant and regulatory evaluation of various dispersion models studied in this project, the emission factors for the tracer study are based on the release rates discussed earlier in the Task 6 report.[38] It needs to be noted that the release rates ($\dot{m}_{i,avg}$) shown in Table 5 of Appendix B, are the average instantaneous release rate and do not correct for the time when source bag was not releasing tracer. For example, while Table 5 shows an average instantaneous PDCB emission rate of 0.06 g/s for release #2, however not all of the tracer mass was released during the time in which the ambient air was being measured. We recalculated the average 1-hour emission rate (from the high-resolution data) corresponding to the ambient air measurement time, and in this example the recalculated average emission rate of PDCB is 0.0176 g/s. The average tracer release rates corresponding to the air being collected are shown in Table 4-8.

---

[38] 38 NCHRP 25-55: Assessment of Regulatory Air Pollution Dispersion Models to Quantify the Impacts of Transportation Sector Emissions: *Task 6: Initial Report on Tracer Experiments*, March 2022. Included here as Appendix B.

Table 4-8. Tracer release rates used in air quality dispersion modeling

Start Time	End Time	Emission Rate (g/s)		
		PDCB*	PMCP	PMCH
5/3/2021 8:00	5/3/2021 9:00	0.01763	0.01615	0.02104
6/6/2021 8:00	6/6/2021 9:00	0.03211	0.03548	0.02797
6/6/2021 10:40	6/6/2021 11:40	0.01684	0.01860	0.02510
6/24/2021 8:00	6/24/2021 9:00	0.03525	0.04044	0.02376
6/24/2021 10:40	6/24/2021 11:40	0.03701	0.04008	0.02295

* On June 6, 2021, no data was recorded for the PDCB release logger after the first 30 min of the release. For 8-9 am release, we assumed same release rate as the first 30 min. We then used the ratio of 10:40 - 11:40 vs. 8-9 am release for PMCP to estimate the release rate for PDCB within 10:40 - 11:40 am. The magnitude of error from this correction, if any, is unknown.

5. Dispersion Modeling Results

This section provides the modeling results for the tracer, pollutants, and regulatory evaluations. Each model is evaluated using a series of metrics, each of which are defined in Section 2.3. In summary, for each model under each evaluation (i.e., tracer, pollutant based, and regulatory), we employ the following metrics:

- Average Residual
- Absolute Average Residual
- Standard Dev. of Residual
- Root Mean Square Error
- R-squared
- Robust Highest Concentration – Observed
- Robust Highest Concentration – Modeled
- Fractional Bias
- Regulatory design values

All are applied to tracer and pollutant-based evaluations except the regulatory design values, which are applied only for regulatory evaluation. Additionally, the modeled concentrations are also compared graphically through:

- Scatter plots
- Q-Q plots
- Time series plots

With the first two for tracer, pollutant based, and regulatory evaluations, and the last only for regulatory evaluation.

Section 5.1 provides detailed results and charts for each model individually. Section 5.2 compared models side-by-side and describes the relative performance of the models compared to each other. Section 5.3 provides an evaluation of the refined models with respect to the qualitative metrics (Section 2.3.3.)

Here we present results compared for each of the three tracers individually, as well as the total of all. As each vehicle released a single tracer, the total (sum from all tracer release vehicles) provides the most continuous mass release at any given location. However, the performance of individual tracers is similar to that of the total, supporting the idea that each tracer gas was released consistently and that use of a single to two tracer vehicles are sufficient for determining hourly average concentration. We present all cases in this report.

5.1. Quantitative and Graphical Results by model

5.1.1. AERMOD-AREAPOLY

5.1.1.1. Tracer Evaluation

Table 5-1 provides evaluation metrics for AERMOD with the polygon area source configuration (referred to as AERMOD-AREAPOLY) compared against observed tracer concentrations. As indicated by the residual values, AERMOD-AREAPOLY slightly underestimates on average the concentrations of PDCB, PMCP, and the total of all three tracers, while it showed relatively good performance for PMCH. The fractional bias of means are between -0.67 and 0.67 illustrating that all modeled concentration are within factor of two of observations for all three tracers. The correlation coefficient (R-squared) ranged between 0.74 to 0.83 across the tracers.

Table 5-1. AERMOD-AREAPOLY Quantitative Evaluation Metrics for Tracer Study (ppt unless otherwise stated)

Metric	PDCB	PMCH	PMCP	Total
Average Residual	29.6	-0.38	15.9	45.0
Absolute Average Residual	35.1	11.5	23.5	62.8
Standard Dev. of Residual	26.1	15.3	23.5	59.1
Root Mean Square Error	39.4	15.3	28.3	74.3
R-squared (unitless)	0.77	0.74	0.83	0.82
Fractional Bias (unitless)	0.398	-0.009	0.217	0.236
Robust Highest Concentration – Observed	194.8	114.2	261.1	563.8
Robust Highest Concentration – Modeled	204.3	144.4	234.2	585.3

Figure 5-1 and Figure 5-2 are the scatter and Q-Q plots of AERMOD-AREAPOLY modeled concentrations against observed data. Blue dots represent concentrations at near-field receptors (less than 50 m from the roadway) while red dots represent concentrations at far-field receptors (≥50 m receptors). As shown, the model tended to have a better performance for near-field receptors as compared to far-field receptors. As values are unpaired in space, the Q-Q plots don't distinguish these distances. At the higher concentrations, points lie relatively close to the 1:1 line, indicating good agreement where concentrations are relatively high, consistent with the relatively good agreement between the RHC-Observed and RHC-Modeled.

Figure 5-1. Scatter Plots of Observed vs. AERMOD-AREAPOLY Modeled Concentrations for Tracer Study (n=51)

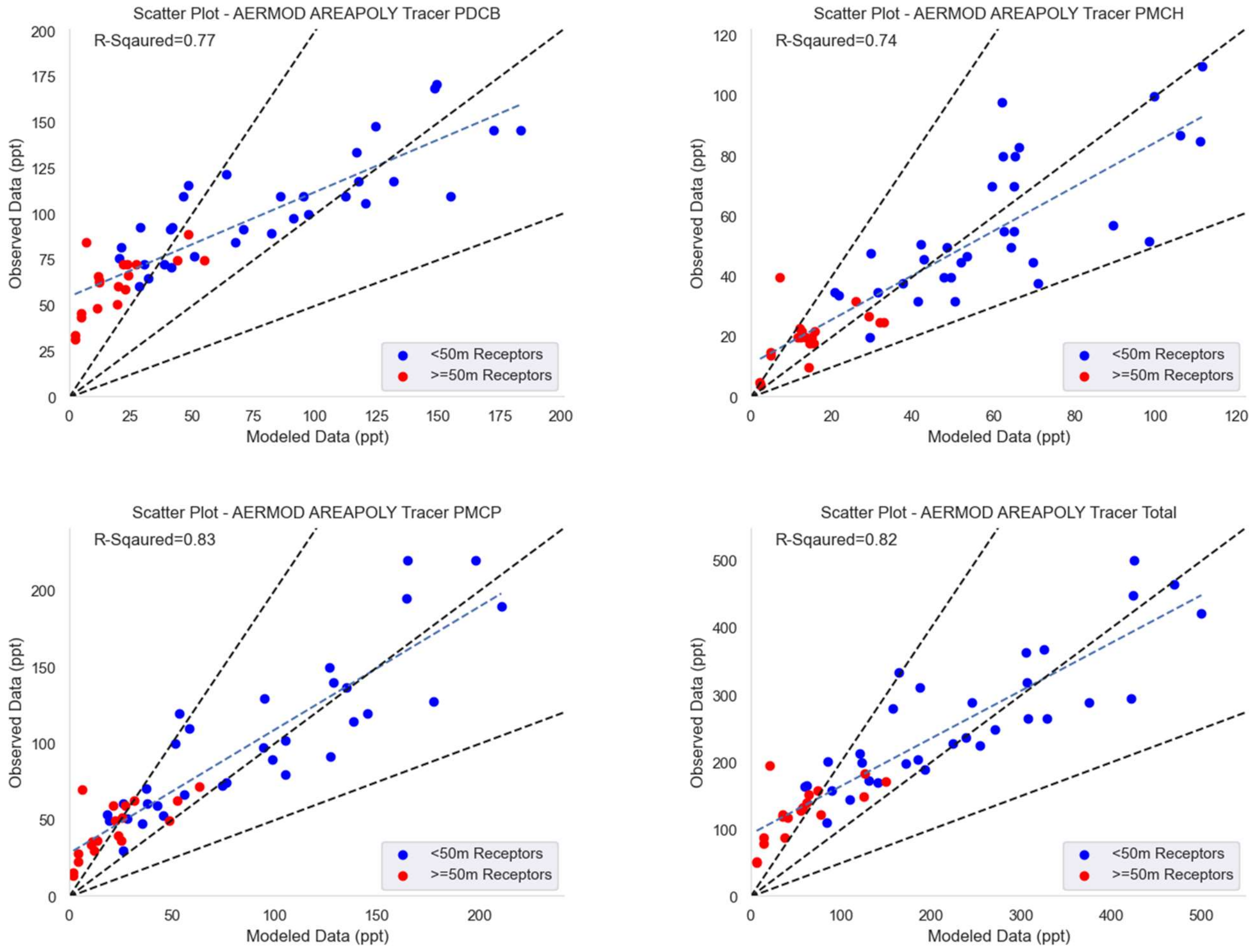
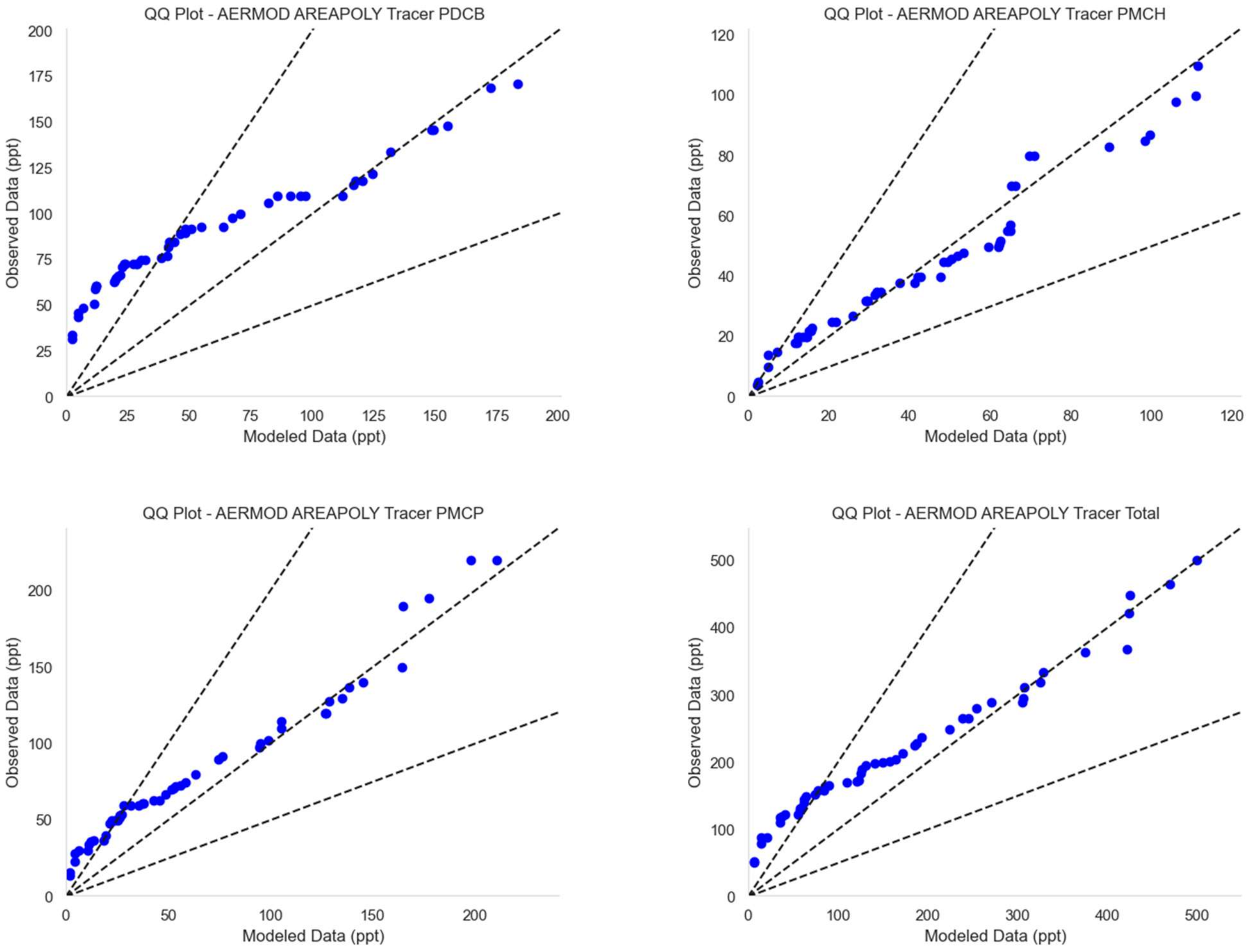


Figure 5-2. Q-Q Plots of Observed vs. AERMOD-AREAPOLY Modeled Concentrations for Tracer Study (n=51)



5.1.1.2. Pollutant Evaluation

Table 5-2 provides evaluation metrics for AERMOD-AREAPOLY when compared against observed concentrations for the pollutants-based evaluation for CO, $PM_{2.5}$, and BC. As shown by the Fractional Bias (FB), the model tends to underestimate concentrations for CO and BC, while slightly overestimating concentrations for $PM_{2.5}$. While for $PM_{2.5}$ the FB value is between -0.67 and 0.67, for the other two pollutants (CO and BC), the FB values are greater than 0.67 illustrating that modeled concentration for these pollutants are underestimated by more than factor of two.. Also, the weak correlation (i.e., R-squared) between the modeled and observed concentrations indicates little relationship between observed and predicted values when using estimated pollutant emission rates for CO, $PM_{2.5}$, and BC.

Table 5-2. AERMOD-AREAPOLY Quantitative Evaluation Metrics for Pollutant Evaluation

Metric	CO (ppm)	$PM_{2.5}$ (µg/m^3)	BC (µg/m^3)
Average Residual	0.18	-0.45	0.43
Absolute Average Residual	0.18	1.94	0.43
Standard Dev. of Residual	0.09	2.94	0.23
Root Mean Square Error	0.20	2.98	0.49
R-squared (unitless)	0.22	0.08	0.02
Fractional Bias (unitless)	0.894	-0.236	1.060
Robust Highest Concentration – Observed	0.55	4.92	1.13
Robust Highest Concentration – Modeled	0.19	3.55	0.30

Figure 5-3 and Figure 5-4 are the scatter and Q-Q plots of AERMOD-AREAPOLY modeled concentrations against observed air pollutant data. Consistent with the statistical results presented above, the estimated concentration was significantly lower than observed data for CO and BC, while for $PM_{2.5}$ the modeled concentrations were within factor of two. Unlike with the tracer evaluations we relied on the emission factor data from CARB's EMFAC2021 model as well as background concentrations from an upwind monitor to calculate the on-road vehicle emissions contributions to the pollutant concentrations downwind.[39] Therefore, these comparisons represent the performance of the dispersion models used in this study coupled with emissions models and background concentration. All of which introduce uncertainties that propagate through these comparisons. Given the relatively high correlation and relatively low fractional bias (between -0.67 and 0.67) found in the tracer measurements and the use of nearly identical meteorology and measure on-site background concentration we suspect the root cause of the poor performance are the emission factors.

---

[39] See Appendix B for description of the experimental setup, including use of combined up- and down-wind monitors.

Figure 5-3. Scatter Plots of Observed vs. AERMOD-AREAPOLY Modeled Concentrations for Pollutant Evaluation (n=9)

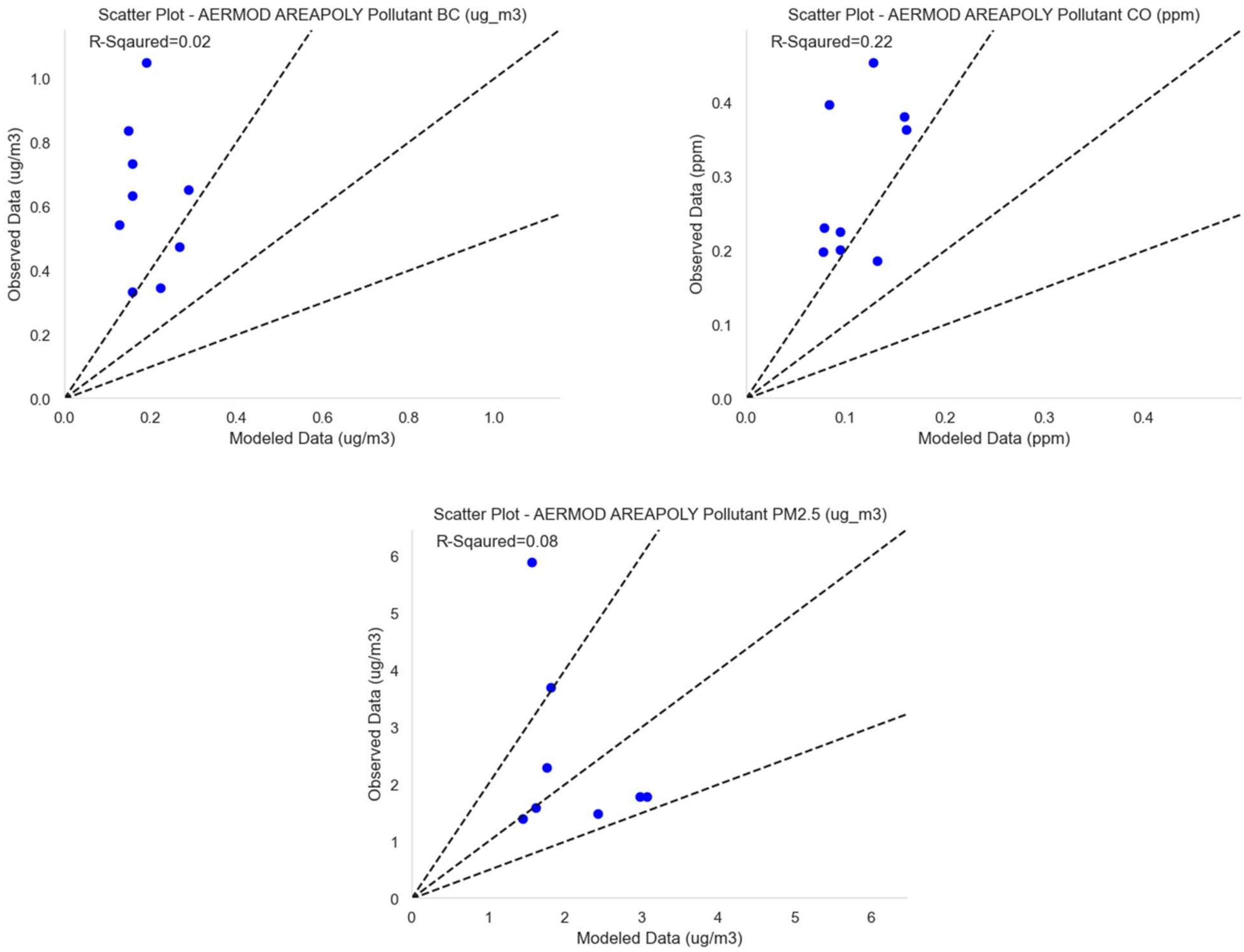
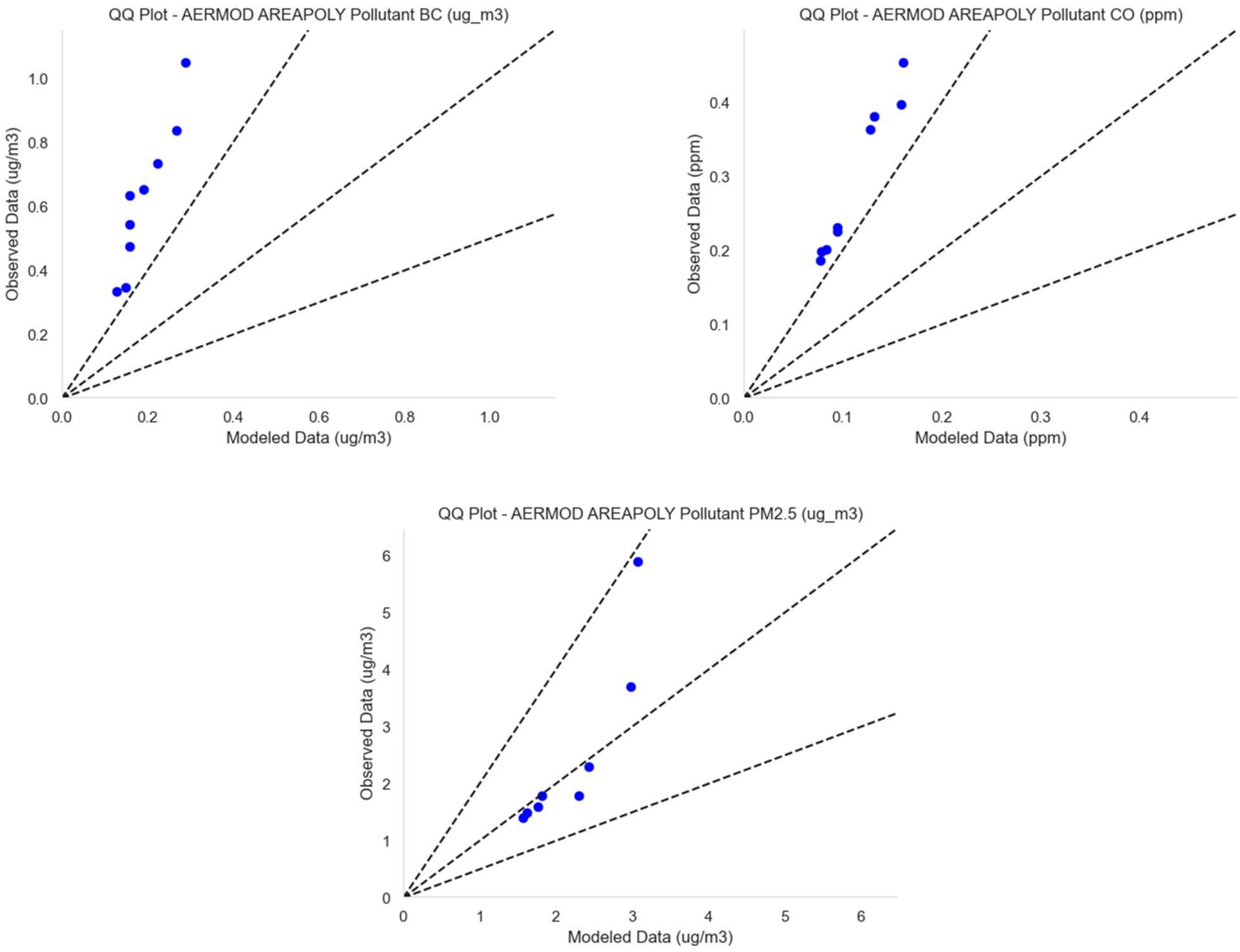


Figure 5-4. Q-Q Plots of Observed vs. AERMOD-AREAPOLY Modeled Concentrations for Pollutant Evaluation (n=9)



5.1.1.3. Regulatory Evaluation

Table 5-3 provides a comparison of the design values (DVs) for CO and $PM_{2.5}$ concentrations against the those calculated based on the BAAQMD's near-road monitoring station. A DV is a statistic that describes the air quality status of a given area relative to the level of the National Ambient Air Quality Standards (NAAQS). Here in this study, the 1-hour and 8-hour CO DVs are calculated as the second highest 1-hour, and 8-hour averaged concentrations over the course of the year 2019. For 24-hr $PM_{2.5}$, the DV is calculated as the 98th percentile of 24-hr average $PM_{2.5}$ concentration throughout the year. Two sets of comparisons are provided: a) DVs with outliers, and b) DVs without outliers. For the purpose of this study, outliers are defined as any concentration values exceeding average concentrations plus 3 standard deviations. Upon analyzing the observed CO data, we noticed unexpectedly high values of CO concentrations (5.6 ppm) on November 7th at 2 AM and November 8th at 1 AM. Considering the low traffic pollution at these hours, we speculate that these values are not correctly measured and thus removed them. That is one of the main reasons why we implemented an outlier exclusion method to draw a reasonable comparison between modeled and observed data under regulatory evaluations. As shown in Table 5-3, when outliers were removed, modeled DVs calculated through AERMOD-AREAPOLY were very consistent with DVs calculated from observed data. For 1-hour CO, DVs were different by 0.1 ppm; for 8-hour CO, DVs were the same; for 24-hour $PM_{2.5}$, the modeled DV was 1 μg/m^3 lower than observed, and for annual $PM_{2.5}$, there was 0.5 μg/m^3 difference.

Table 5-3. AERMOD-AREAPOLY Design Values for Regulatory Evaluation[40]

Averaging Period and Pollutant	With Outliers		Without Outliers	
	Modeled	Observed	Modeled	Observed
1-hour CO (ppm)	2.2	5.6	1.2	1.1
8-hour CO (ppm)	1.3	1.2	1.0	1.0
24-hour $PM_{2.5}$ (μg/m^3)	20	20	18	19
Annual $PM_{2.5}$ (μg/m^3)	9.0	9.5	8.8	9.3

Figure 5-5, Figure 5-6, and Figure 5-7 are scatter, Q-Q, and time-series plots, respectively, comparing modeled vs. observed 1-hour and 8-hour average CO as well as 24-hour average $PM_{2.5}$ concentrations[41]. As shown, the modeled concentrations (with background concentration added) were well correlated with observed data.

We note that background concentration contributes more than 80 percent (across various averaging periods) to the total CO and PM concentrations at the near-road site. Therefore, when compared, it is no surprise to see relatively low bias between modeled and observed concatenations. Similar to pollutant-based evaluation, the Q-Q plot demonstrates modeled CO concentrations smaller than the observed data. As described earlier, such underestimation cannot be solely attributed to the performance of the dispersion model, due to relatively high uncertainties with modeled emission factors (e.g., emissions deterioration, in-use performance,

---

[40] Consistent with the form of the NAAQS and reporting requirements, annual PM2.5 are reported to one decimal place 24-hr PM2.5 values are rounded to nearest whole number. The form of the CO NAAQS is also integer values, however this not specified in EPA's CO Hotspot Guidance and U.S. EPA reporting of design values are typically done to one decimal place. We have repeated that here for 1-hr and 8-hr CO DVs.

[41] Only data without outliers are presented in these charts.

impact of ambient conditions). Please note that when comparing the data scatter in Figure 5-5, 1-hr and 8-hr CO plots show a higher scatter than 24-hr $PM_{2.5}$ concentration. This is partly due to the temporal averaging for the metric shown, which inherently reduces scatter for longer-term averages.

Figure 5-5. Scatter Plots of Observed vs. AERMOD-AREAPOLY Modeled Concentrations for Regulatory Evaluation (n=8,760 for 1 hour CO, n=1,095 for 8-hour CO, and n=365 for 24-hr $PM_{2.5}$)

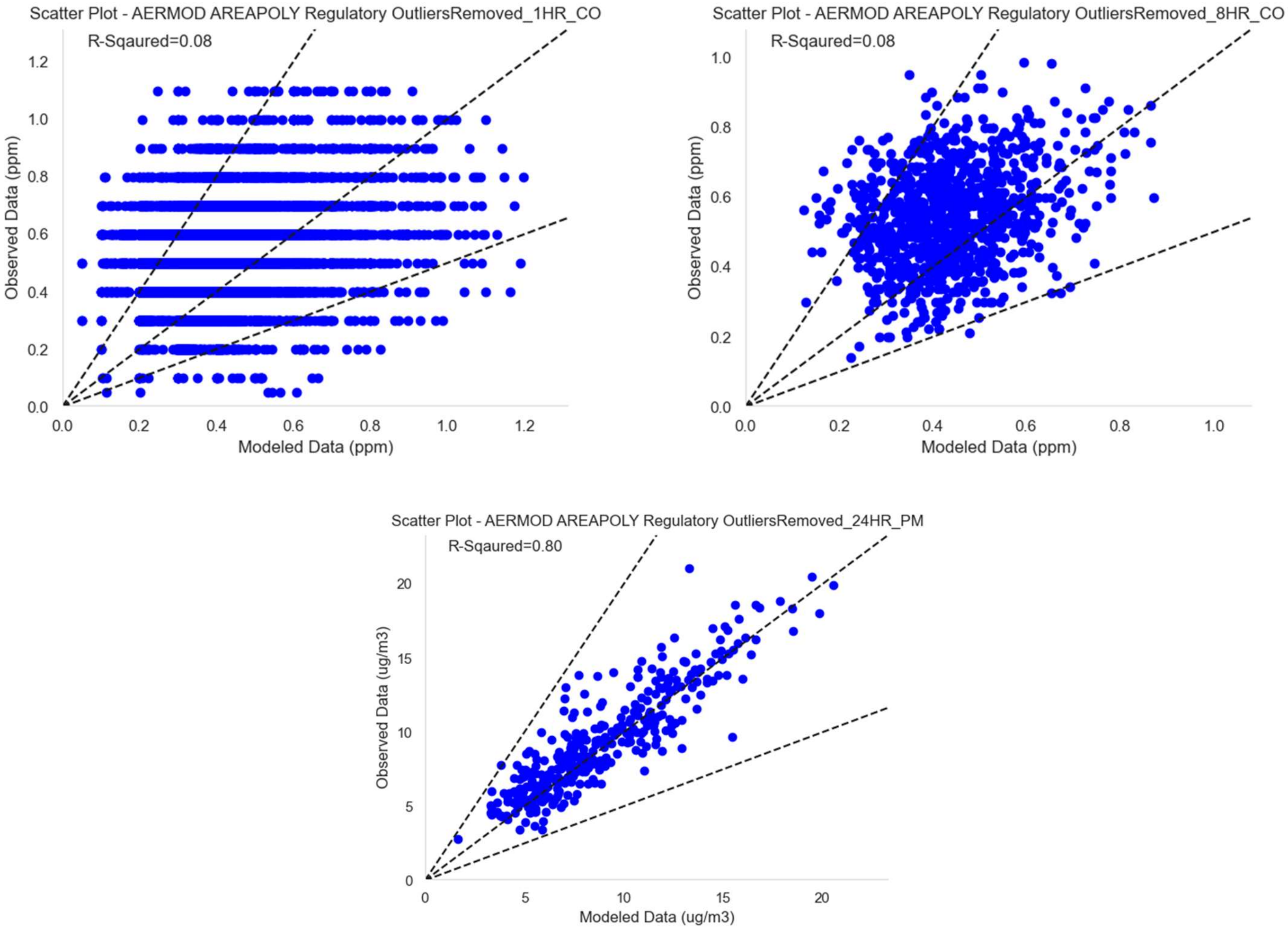


Figure 5-6. Q-Q Plots of Observed vs. AERMOD-AREAPOLY Modeled Concentrations for Regulatory Evaluation (n=8,760 for 1 hour CO, n=1,095 for 8-hour CO, and n=365 for 24-hr $PM_{2.5}$)

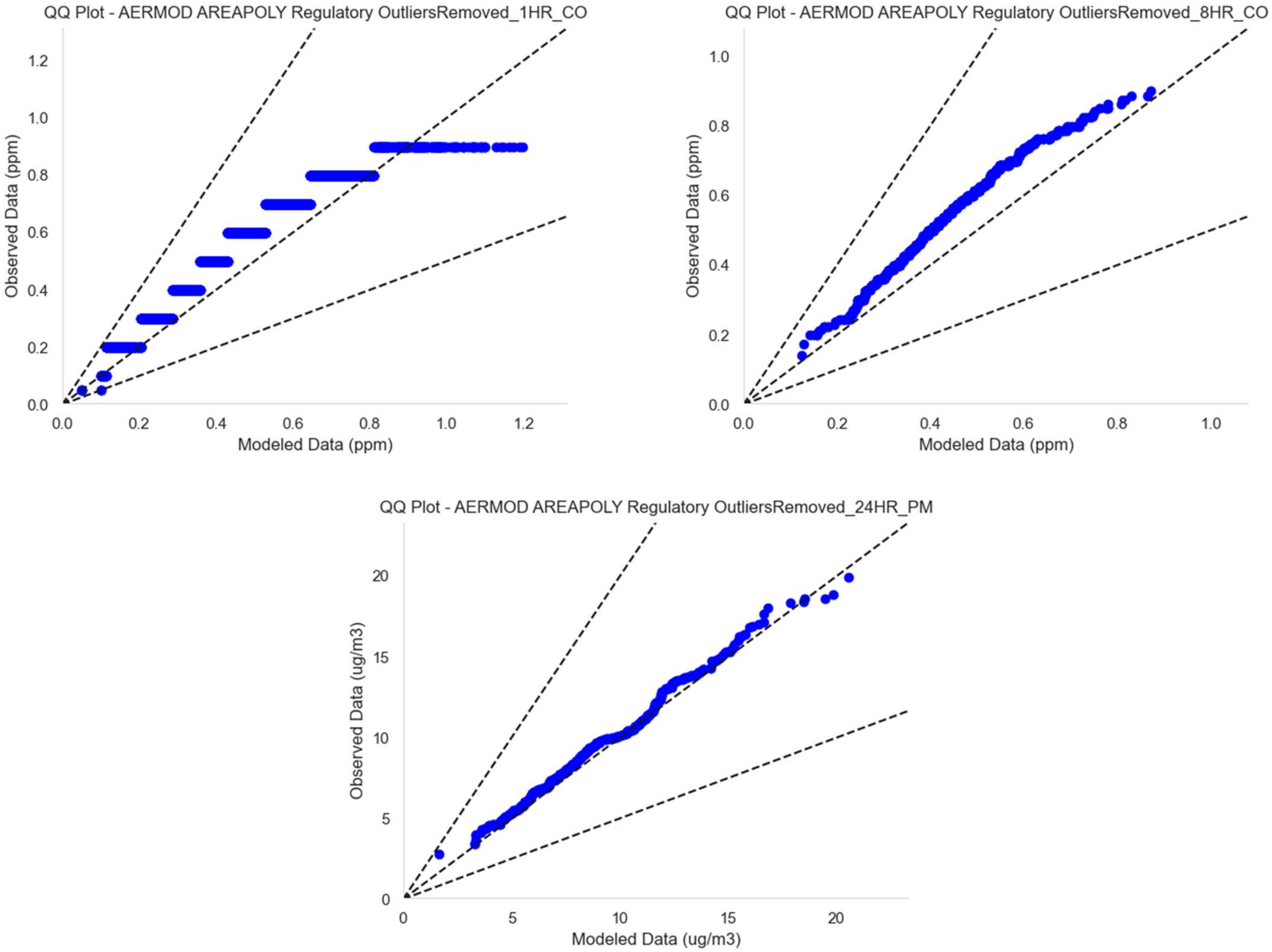
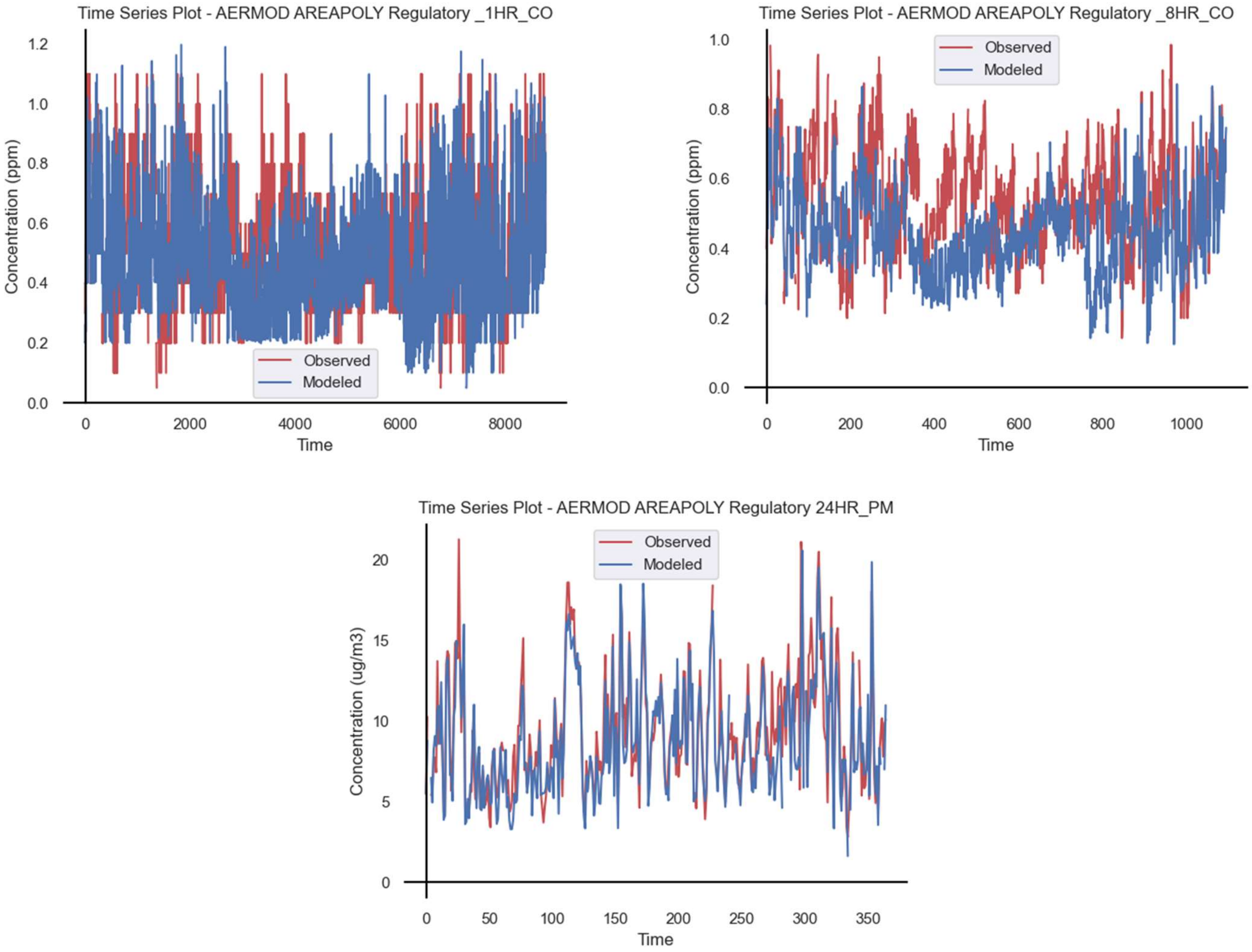


Figure 5-7. Timeseries Plots of Observed vs. AERMOD-AREAPOLY Modeled Concentrations for Regulatory Evaluation for 2019. For CO, x-axis represents 1-hr increments for 1-hr averaged CO and 8-hour increments for 8-hr averaged CO, while for $PM_{2.5}$, the x-axis represents 24-hr increments.



5.1.2. AERMOD-Volume

5.1.2.1. Tracer Evaluation

Table 5-4 provides evaluation metrics for AERMOD with the volume source configuration (referred to as AERMOD-Volume) when compared against observed data under the tracer evaluation. As indicated by the residual values, the AERMOD-Volume tended to underestimate the concentrations of all three tracers. The correlation coefficient (R-squared) ranged between 0.72 to 0.81 across the tracers. The fractional bias of means are between -0.67 and 0.67 illustrating that all modeled concentration are within factor of two of observations for all three tracers.

Table 5-4. AERMOD-Volume Quantitative Evaluation Metrics for Tracer Evaluation (ppt unless otherwise stated)

Metric	PDCB	PMCH	PMCP	Total
Average Residual	42.3	9.4	29.7	81.4
Absolute Average Residual	42.7	12.5	31.19	82.5
Standard Dev. of Residual	20.7	13.7	21.9	49.9
Root Mean Square Error	47.07	16.6	36.9	95.4
R-squared (unitless)	0.72	0.72	0.81	0.79
Fractional Bias (unitless)	0.623	0.247	0.449	0.472
Robust Highest Concentration – Observed	194.8	114.2	261.1	563.8
Robust Highest Concentration – Modeled	154.4	97.8	181.5	425.9

Figure 5-8 and Figure 5-9 are the scatter and Q-Q plots when AERMOD-Volume modeled concentrations are compared against observed values. Blue dots represent concentrations at near-field receptors (less than 50 m from the roadway) while red dots represent concentrations at far-field receptors (≥50 m receptors). Similar to AERMOD-AREAPOLY, the model tended to have better performance for near-field receptors as compared to far-field receptors. In general model performance was not as good as AREAPOLY which showed slightly higher r-squared values and better agreement between the RHC observed and modeled.

Figure 5-8. Scatter Plots of Observed vs. AERMOD-Volume Modeled Concentrations for Tracer Evaluation (n=51)

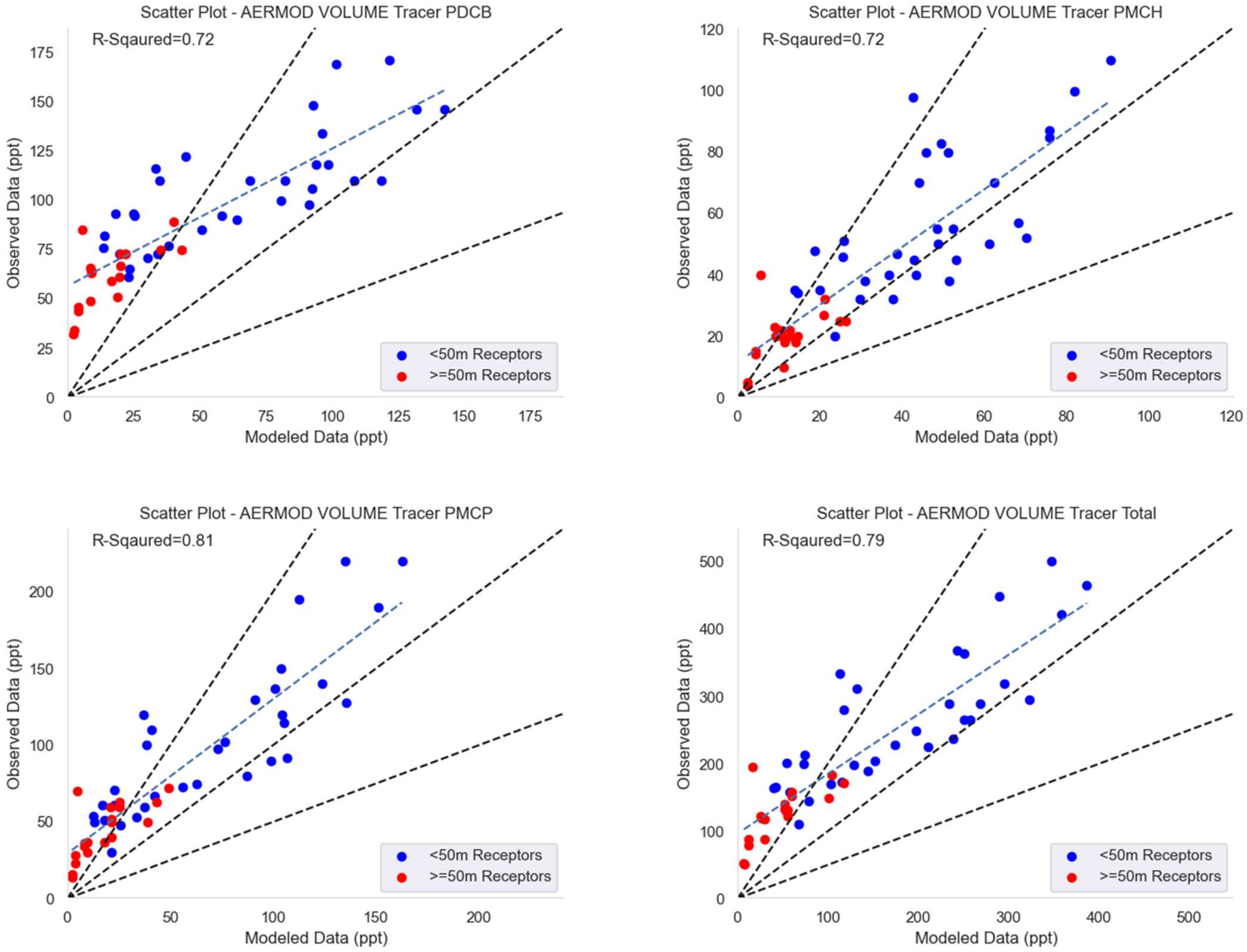
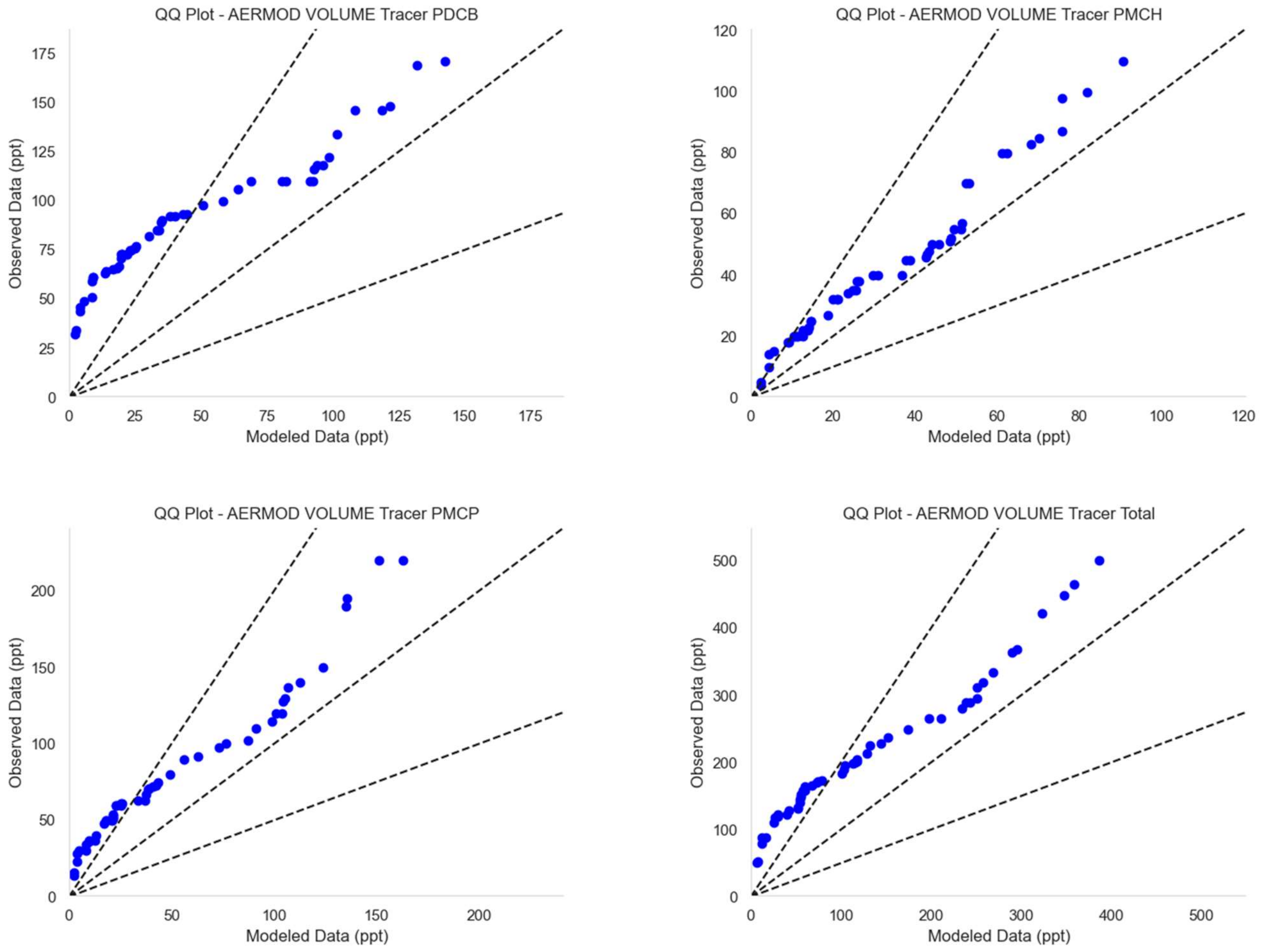


Figure 5-9. Q-Q Plots of Observed vs. AERMOD-Volume Modeled Concentrations for Tracer Evaluation (n=51)



5.1.2.2. Pollutant Evaluation

Table 5-5 provides evaluation metrics for AERMOD-Volume when compared against observed under the pollutant-based evaluation of CO, $PM_{2.5}$, and BC. As shown, similar to other source configurations, AERMOD-Volume also tended to underestimate concentrations of CO and BC while slightly overestimating $PM_{2.5}$ concentrations. This is evident from the FB values for CO and BC being positive, while being negative for $PM_{2.5}$. While for $PM_{2.5}$ the FB value is between -0.67 and 0.67, for the other two pollutants (CO and BC), the FB values are greater than 0.67 illustrating that modeled concentration for these pollutants are underestimated by more than factor of two.

Table 5-5. AERMOD-Volume Quantitative Evaluation Metrics for Pollutant Evaluation

Metric	CO (ppm)	$PM_{2.5}$ (μg/m^3)	BC (μg/m^3)
Average Residual	0.19	-0.28	0.45
Absolute Average Residual	0.19	1.9	0.45
Standard Dev. of Residual	0.083	2.8	0.24
Root Mean Square Error	0.21	2.8	0.51
R-squared (unitless)	0.39	0.003	0.12
Fractional Bias (unitless)	0.963	-0.152	1.121
Robust Highest Concentration – Observed	0.55	4.9	1.1
Robust Highest Concentration – Modeled	0.17	3.1	0.29

Figure 5-10 and Figure 5-11 are the scatter and Q-Q plots when AERMOD-Volume modeled concentrations are compared against observed data. As illustrated, the estimated concentrations were significantly lower than observed data for CO, and BC, while for $PM_{2.5}$ the modeled concentrations were within a factor of two (except for two data points). As noted in earlier discussion, these comparisons suggest that the relatively poor model performance are due to other parts of the modeling chain, including emission modeling emission factors since the tracer evaluation showed relatively good performance measures for both the r squared and RHC.

Figure 5-10. Scatter Plots of Observed vs. AERMOD-Volume Modeled Concentrations for Pollutant Evaluation (n=9)

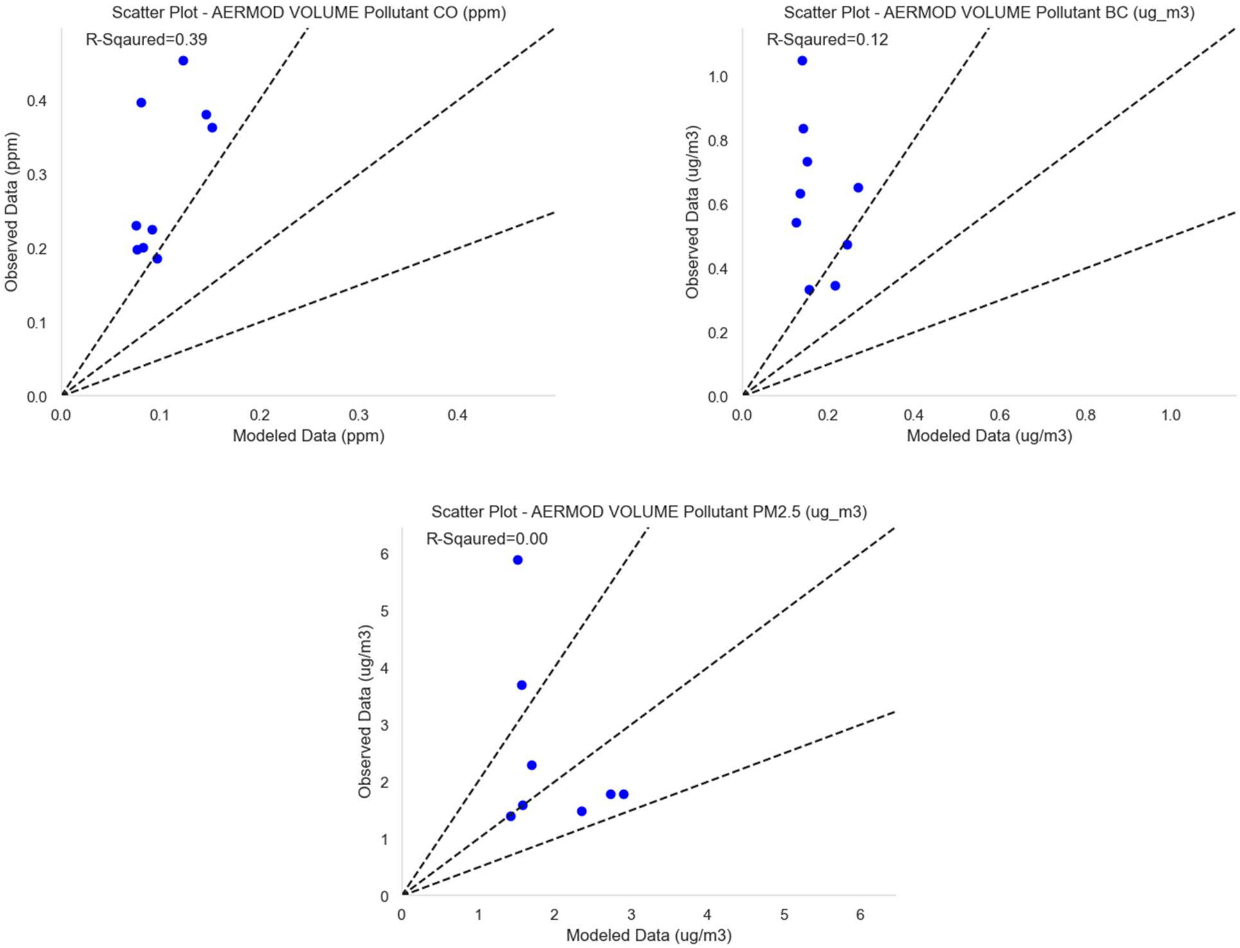
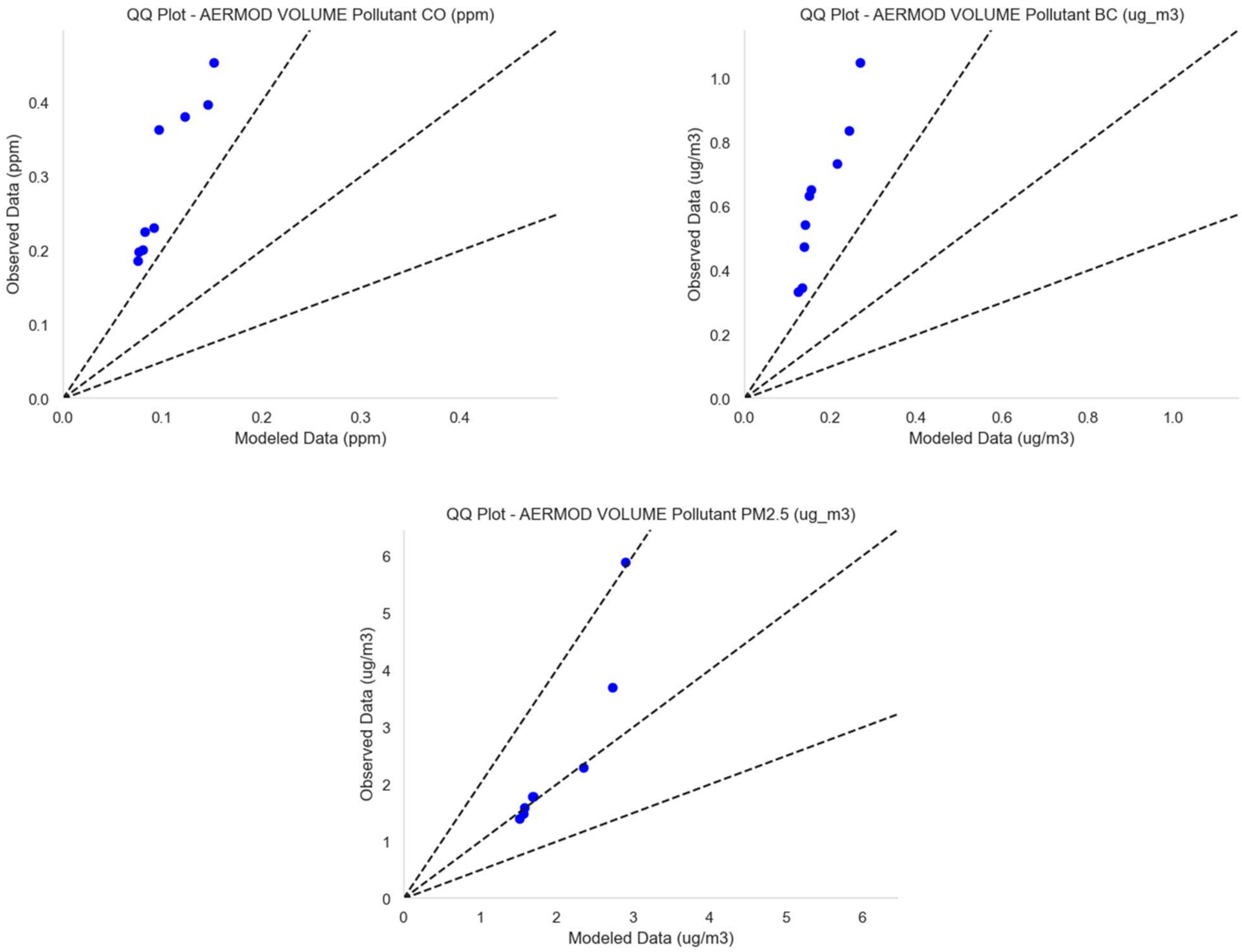


Figure 5-11. Q-Q Plots of Observed vs. AERMOD-Volume Modeled Concentrations for Pollutant Evaluation (n=9)



5.1.2.3. Regulatory Evaluation

Table 5-6 provides a comparison of the DVs for CO and $PM_{2.5}$ concentrations against the DVs calculated based on the near-road monitoring station. As described earlier, two sets of comparisons are provided: a) DVs with outliers, and b) DVs without outliers. As shown in Table 5-6, when outliers were removed, DVs calculated through AERMOD-Volume are very consistent with DVs calculated from observed data. For 1-hour CO and 24-hr PM, DVs were the same, for 8-hour average CO, DVs were different by 0.1 ppm; and for annual $PM_{2.5}$ there was 0.2 μg/m^3 difference.

Table 5-6. AERMOD-Volume Design Values for Regulatory Evaluation[42]

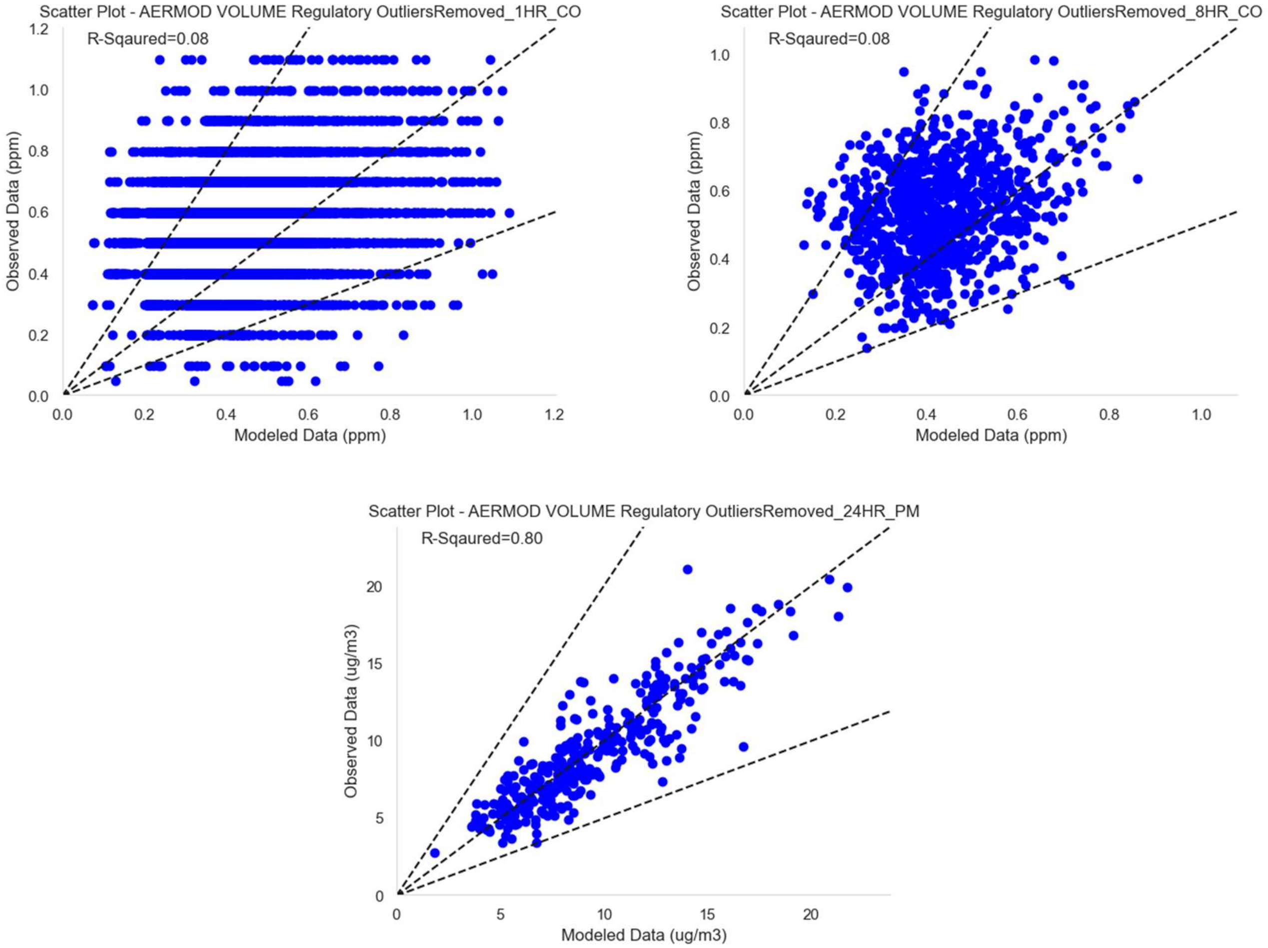
Averaging Period and Pollutant	With Outliers		Without Outliers	
	Modeled	Observed	Modeled	Observed
1-hour CO (ppm)	1.9	5.6	1.1	1.1
8-hour CO (ppm)	1.2	1.2	0.9	1.0
24-hour $PM_{2.5}$ (μg/m^3)	21	20	19	19
Annual $PM_{2.5}$ (μg/m^3)	9.7	9.5	9.5	9.3

Figure 5-12, Figure 5-13, and Figure 5-14 are scatter, Q-Q, and time-series plots, respectively, comparing modeled vs. observed 1-hour and 8-hour averaged CO as well as 24-hour averaged $PM_{2.5}$ concentrations. Similar to the pollutant-based evaluation, the Q-Q plot demonstrates lower modeled CO concentrations as compared to the observed data. As noted earlier, the underestimation is likely due to other parts of the modeling chain, including the emissions modeling since the tracer evaluation showed much better model performance. Note that, as discussed above, longer term averages inherently show lower scatter.

---

[42] Consistent with U.S. EPA reporting of design values, the DVs for 1-hr CO, 8-hr CO, and annual PM2.5 are reported with one decimal, while it is rounded to nearest whole number for 24-hr PM2.5.

Figure 5-12. Scatter Plots of Observed vs. AERMOD-Volume Modeled Concentrations for Regulatory Evaluation (n=8,760 for 1 hour CO, n=1,095 for 8-hour CO, and n=365 for 24-hr $PM_{2.5}$)


Scatter Plot - AERMOD VOLUME Regulatory OutliersRemoved_1HR_CO

Scatter Plot - AERMOD VOLUME Regulatory OutliersRemoved_8HR_CO

Scatter Plot - AERMOD VOLUME Regulatory OutliersRemoved_24HR_PM

Figure 5-13. Q-Q Plots of Observed vs. AERMOD-Volume Modeled Concentrations for Regulatory Evaluation (n=8,760 for 1 hour CO, n=1,095 for 8-hour CO, and n=365 for 24-hr $PM_{2.5}$)

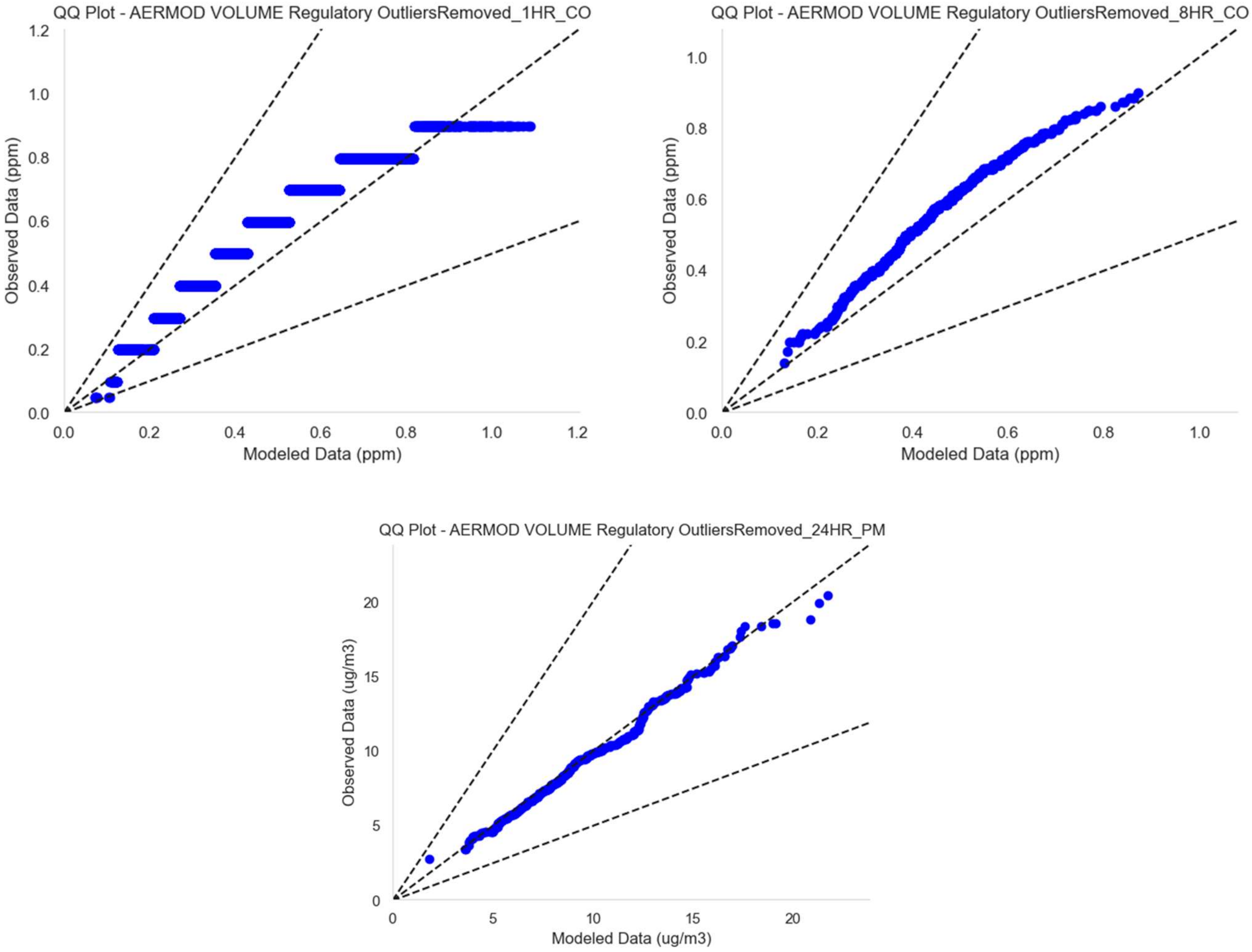
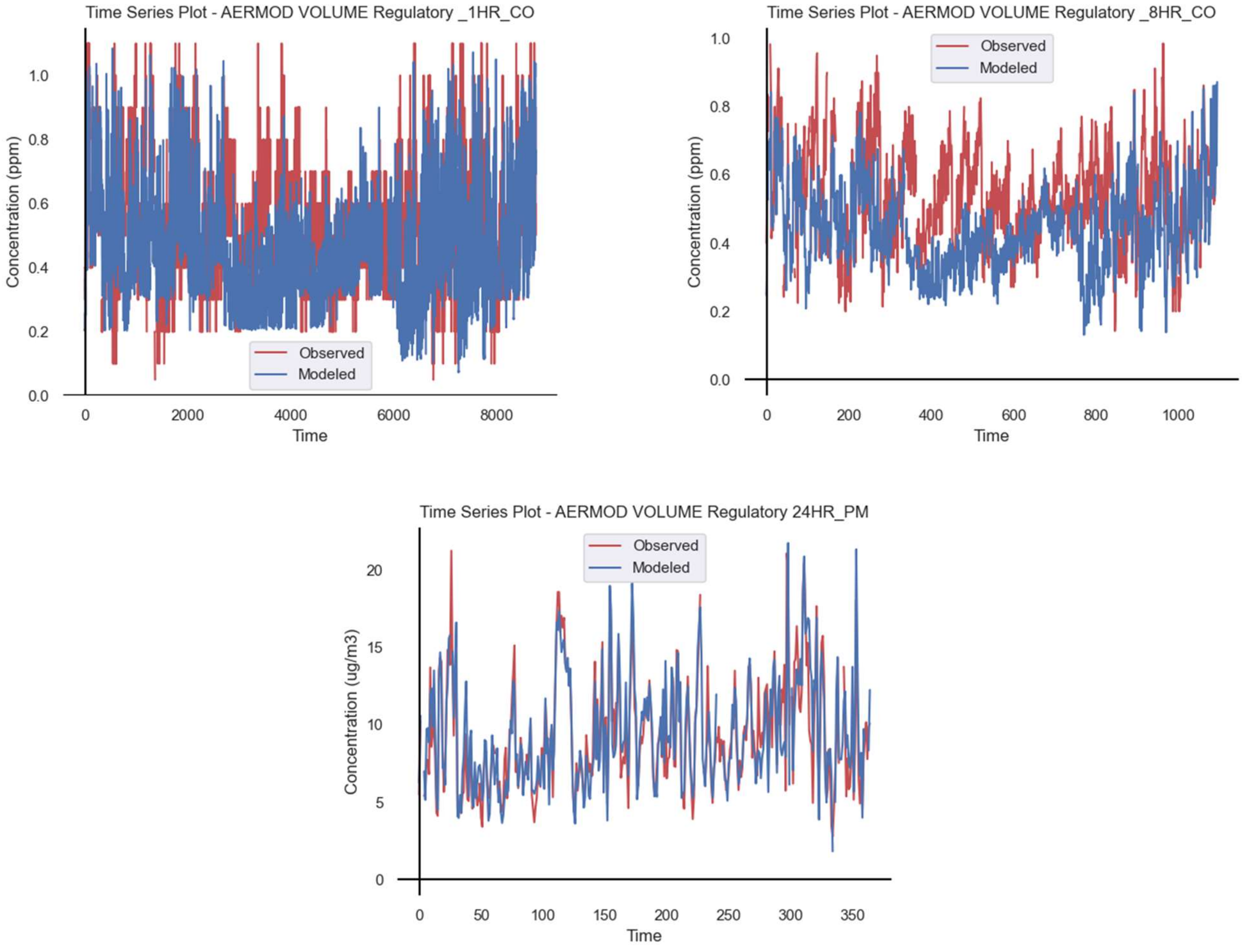


Figure 5-14. Timeseries Plots of Observed vs. AERMOD-Volume Modeled Concentrations for Regulatory Evaluation in 2019. For CO, x-axis represents 1-hr increments for 1-hr averaged CO and 8-hour increments for 8-hr averaged CO, while for $PM_{2.5}$, the x-axis represents 24-hr increments.



5.1.3. AERMOD-RLINE

5.1.3.1. Tracer Evaluation

Table 5-7 provides evaluation metrics for AERMOD with the RLINE source configuration (referred to as AERMOD-RLINE) when compared against observed data under the tracer evaluation. As indicated by the residual values, AERMOD-RLINE tended to underestimate the concentrations of all three tracers. The correlation coefficient (R-squared) ranges between 0.73 to 0.83 across the tracers identical to those seen with AERMOD-AREAPOLY. While for PMCH the fractional bias is less than 0.67, for the other two tracers, the fractional bias of means are greater than 0.67 illustrating that modeled concentration for these tracers are underestimated by more than factor of two.

Table 5-7. AERMOD-RLINE Quantitative Evaluation Metrics under Tracer Evaluation (ppt unless otherwise stated)

Metric	PDCB	PMCH	PMCP	Total
Average Residual	53.8	16.5	42.5	112.7
Absolute Average Residual	53.8	17.4	42.5	112.7
Standard Dev. of Residual	15.0	13.8	24.5	47.7
Root Mean Square Error	55.8	21.5	49.0	122.4
R-squared (unitless)	0.78	0.73	0.83	0.82
Fractional Bias (unitless)	0.864	0.477	0.712	0.720
Robust Highest Concentration – Observed	194.8	114.2	261.1	563.8
Robust Highest Concentration – Modeled	106.9	72.4	124.7	311.6

Figure 5-15 and Figure 5-16 are the scatter and Q-Q plots when AERMOD-RLINE modeled concentrations are compared against observed values. Blue dots represent concentrations at near-field receptors (less than 50 m from the roadway) while red dots represent concentrations at far-field receptors (≥50 m receptors). Similar to AERMOD-AREAPOLY and AERMOD-Volume, the model tended to have better performance for near-field receptors as compared to far-field receptors.

Figure 5-15. Scatter Plots of Observed vs. AERMOD-RLINE Modeled Concentrations for Tracer Study (n=51)

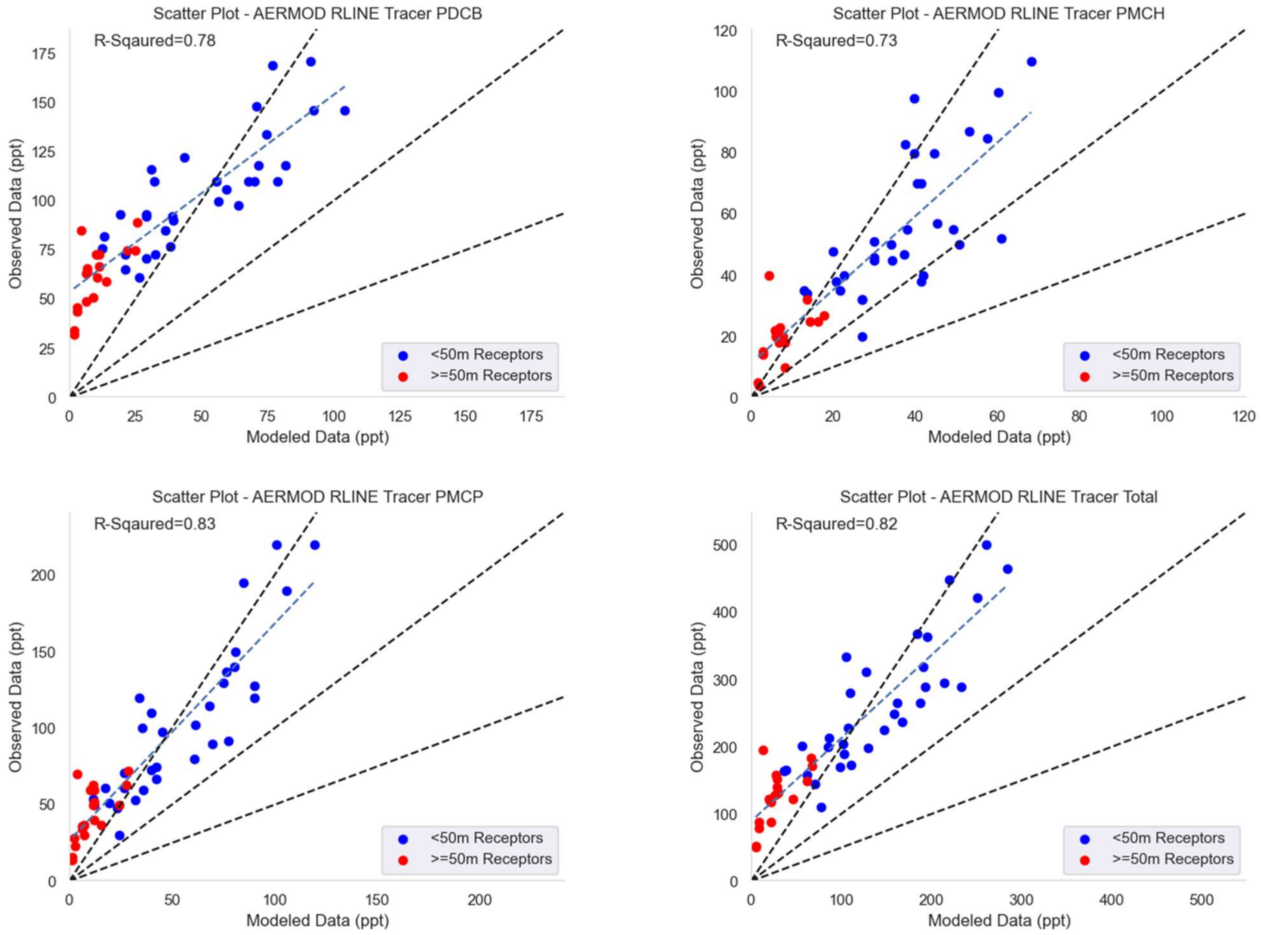
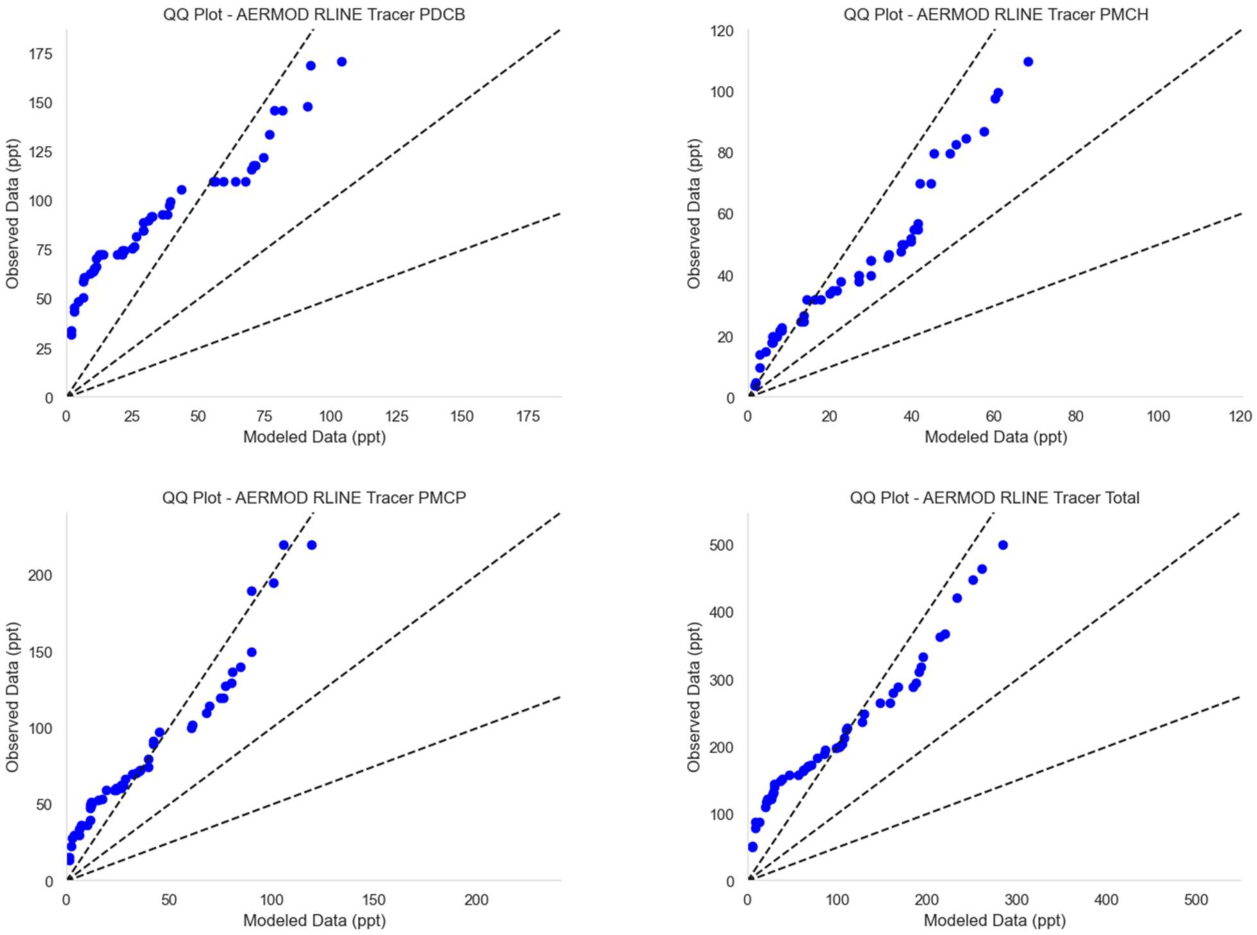


Figure 5-16. Q-Q Plots of Observed vs. AERMOD-RLINE Modeled Concentrations for Tracer Study (n=51)



5.1.3.2. Pollutant Evaluation

Table 5-8 provides evaluation metrics for AERMOD-RLINE when compared against observed data under the pollutant-based evaluation for CO, $PM_{2.5}$, and BC. Similar to AERMOD-AREAPOLY and AERMOD-Volume, AERMOD-RLINE tended to underestimate concentrations for CO and BC, and also it is showing slight underestimation for $PM_{2.5}$. While for $PM_{2.5}$ the Fractional Bias value is less than 0.67, for the other two pollutants (CO and BC), the FB values are greater than 0.67 illustrating that modeled concentration for these pollutants are underestimated by more than factor of two.

Table 5-8. AERMOD-RLINE Quantitative Evaluation Metrics for Pollutant Evaluation

Metric	CO (ppm)	$PM_{2.5}$ (μg/m^3)	BC (μg/m^3)
Average Residual	0.23	0.32	0.50
Absolute Average Residual	0.23	1.8	0.50
Standard Dev. of Residual	0.09	2.9	0.22
Root Mean Square Error	0.24	2.9	0.54
R-squared (unitless)	0.14	0.17	0.00
Fractional Bias (unitless)	1.250	0.216	1.319
Robust Highest Concentration – Observed	0.55	4.9	1.13
Robust Highest Concentration – Modeled	0.12	2.17	0.19

Figure 5-17 and Figure 5-18 are the scatter and Q-Q plots when AERMOD-RLINE modeled concentrations are compared against observed data. As illustrated, the modeled concentrations were significantly lower than observed data for CO and BC, while for $PM_{2.5}$ the modeled concentrations were within factor of two (except for two data points). As discussed for the other models, these comparisons are dependent upon other parts of the modeling chain, including the emission factors and with the relatively good performance when using the tracer emission rates, most of the discrepancy is likely associated with the emission factor model.

Figure 5-17. Scatter Plots of Observed vs. AERMOD-RLINE Modeled Concentrations for Pollutant Evaluation (n=9)

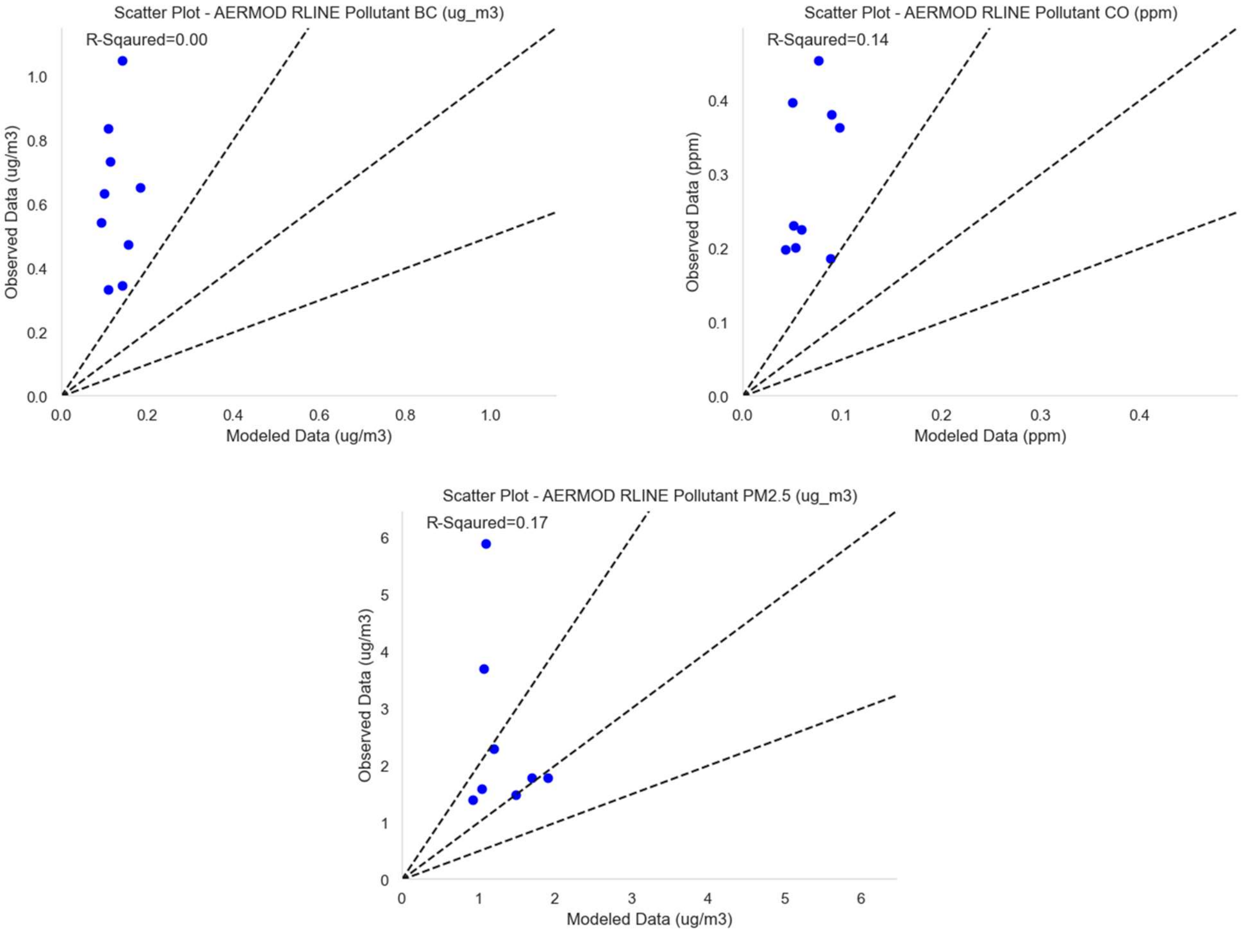
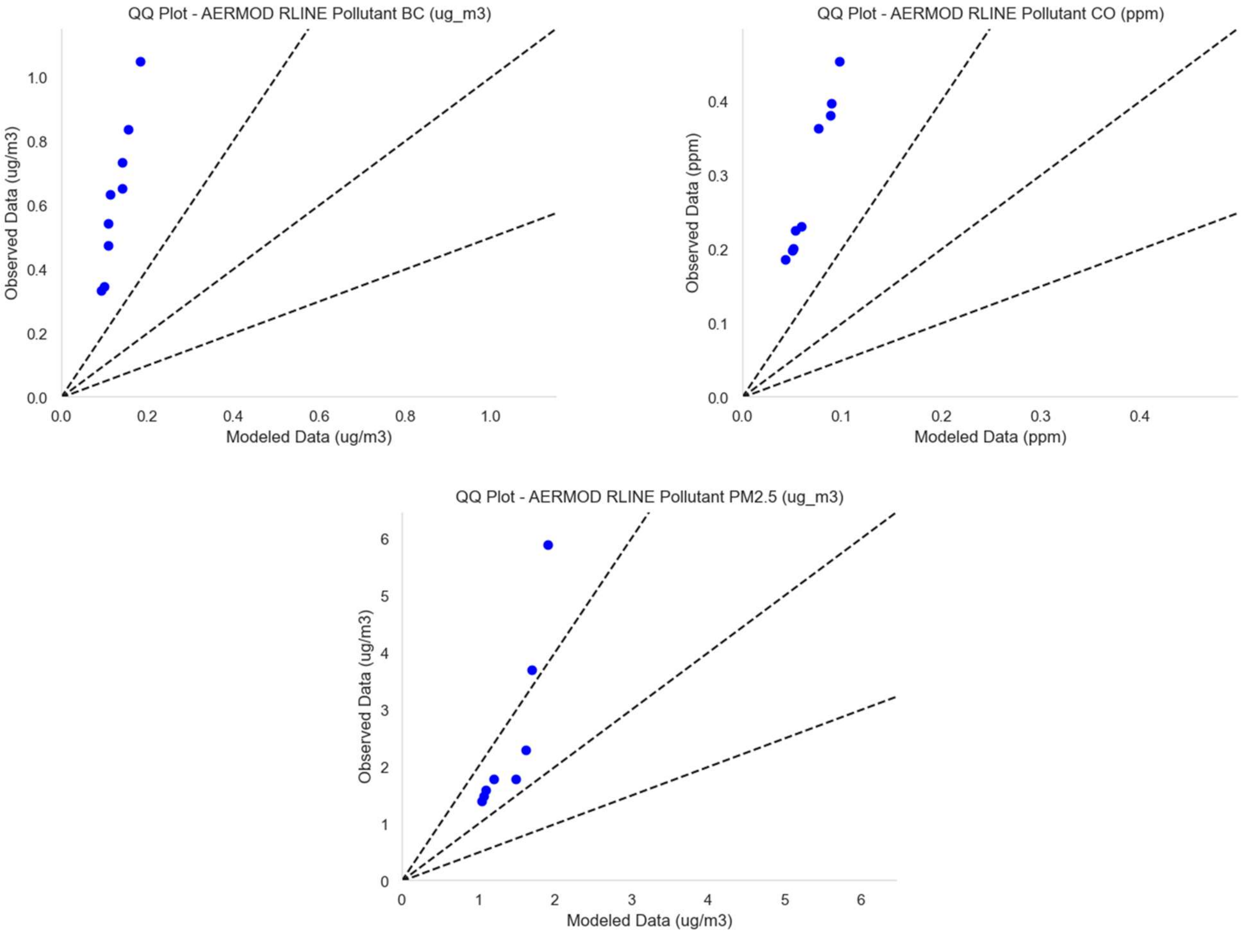


Figure 5-18. Q-Q Plots of Observed vs. AERMOD-RLINE Modeled Concentrations for Pollutant Evaluation (n=9)


QQ Plot - AERMOD RLINE Pollutant BC (ug_m3)

QQ Plot - AERMOD RLINE Pollutant CO (ppm)

QQ Plot - AERMOD RLINE Pollutant PM2.5 (ug_m3)

5.1.3.3. Regulatory Evaluation

Table 5-9 provides a comparison of the DVs for CO and $PM_{2.5}$ concentrations against the DVs calculated based on near-road monitoring station. As described earlier, two sets of comparisons are provided: a) DVs with outliers, and b) DVs without outliers. As shown in Table 5-9, when outliers are removed, DVs calculated through AERMOD-RLINE are relatively similar to design values calculated from observed data. For 1-hour and 8-hour average CO, DVs are different by 0.1 ppm or less; for 24-hour $PM_{2.5}$ the modeled design value is 1 $\mu g/m^3$ higher than observed, and for annual $PM_{2.5}$ there is 0.4 $\mu g/m^3$ difference.

Table 5-9. AERMOD-RLINE Design Values for Regulatory Evaluation[43]

Averaging Period and Pollutant	With Outliers		Without Outliers	
	Modeled	Observed	Modeled	Observed
1-hour CO (ppm)	2.0	5.6	1.2	1.1
8-hour CO (ppm)	1.3	1.2	0.9	1.0
24-hour $PM_{2.5}$ ($\mu g/m^3$)	22	20	20	19
Annual $PM_{2.5}$ ($\mu g/m^3$)	9.9	9.5	9.7	9.3

Figure 5-19, Figure 5-20, and Figure 5-21 are scatter, Q-Q, and time-series plots, respectively, comparing modeled vs. observed 1-hour and 8-hour averaged CO as well as 24-hour averaged $PM_{2.5}$ concentrations. Similar to the pollutant-based evaluation, the Q-Q plot demonstrates lower modeled CO concentrations as compared to the observed data. As noted earlier, model performance here is a combination of the dispersion model, the emissions modeling, and background concentrations. (See Section 5.1.1.3). Note that, as discussed above, longer term averages inherently show lower scatter.

---

[43] Consistent with U.S. EPA reporting of design values, the DVs for 1-hr CO, 8-hr CO, and annual PM2.5 are reported with one decimal, while it is rounded to nearest whole number for 24-hr PM2.5.

Figure 5-19. Scatter Plots of Observed vs. AERMOD-RLINE Modeled Concentrations for Regulatory Evaluation (n=8,760 for 1 hour CO, n=1,095 for 8-hour CO, and n=365 for 24-hr PM$_{2.5}$)

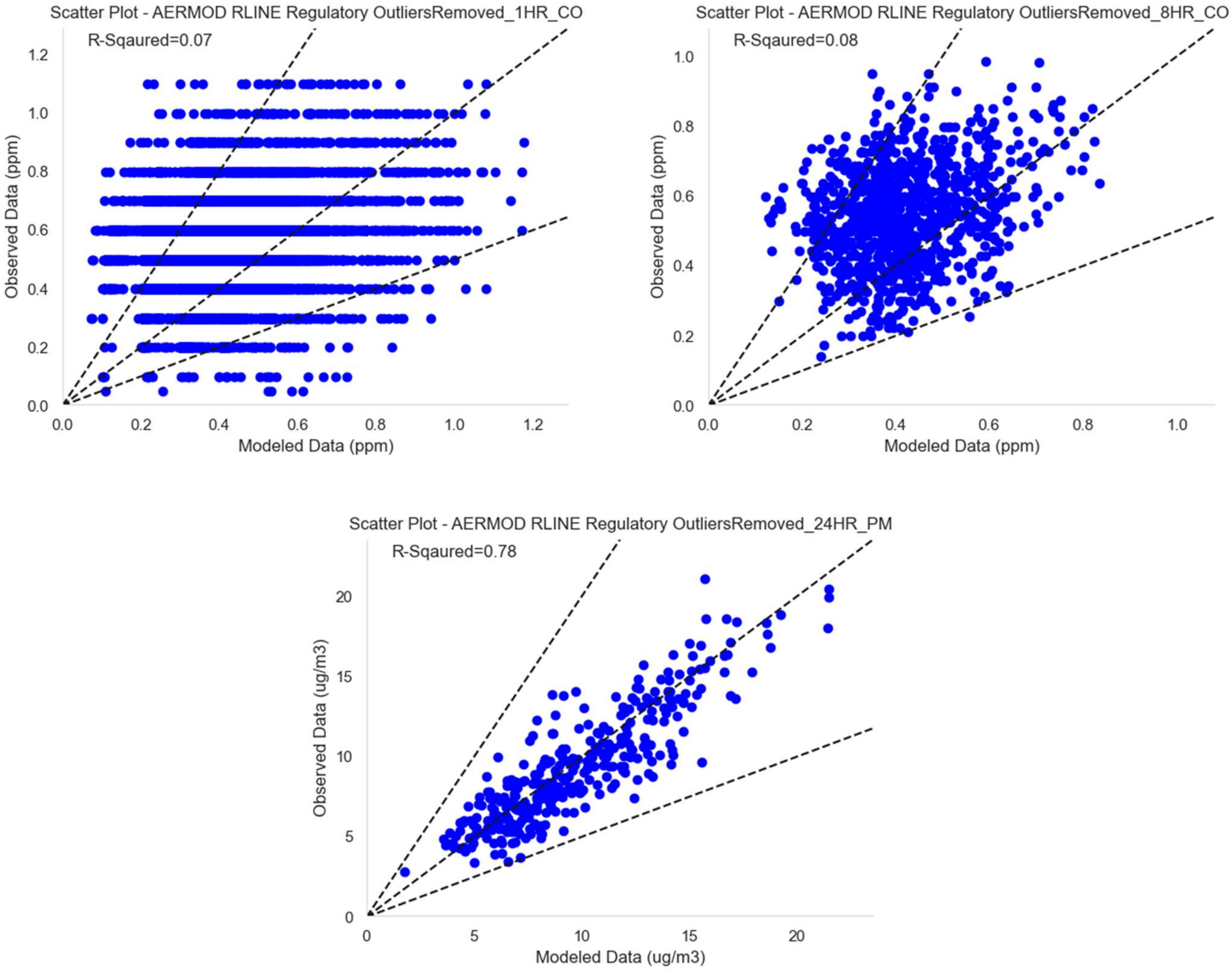


Figure 5-20. Q-Q Plots of Observed vs. AERMOD-RLINE Modeled Concentrations for Regulatory Evaluation (n=8,760 for 1 hour CO, n=1,095 for 8-hour CO, and n=365 for 24-hr $PM_{2.5}$)

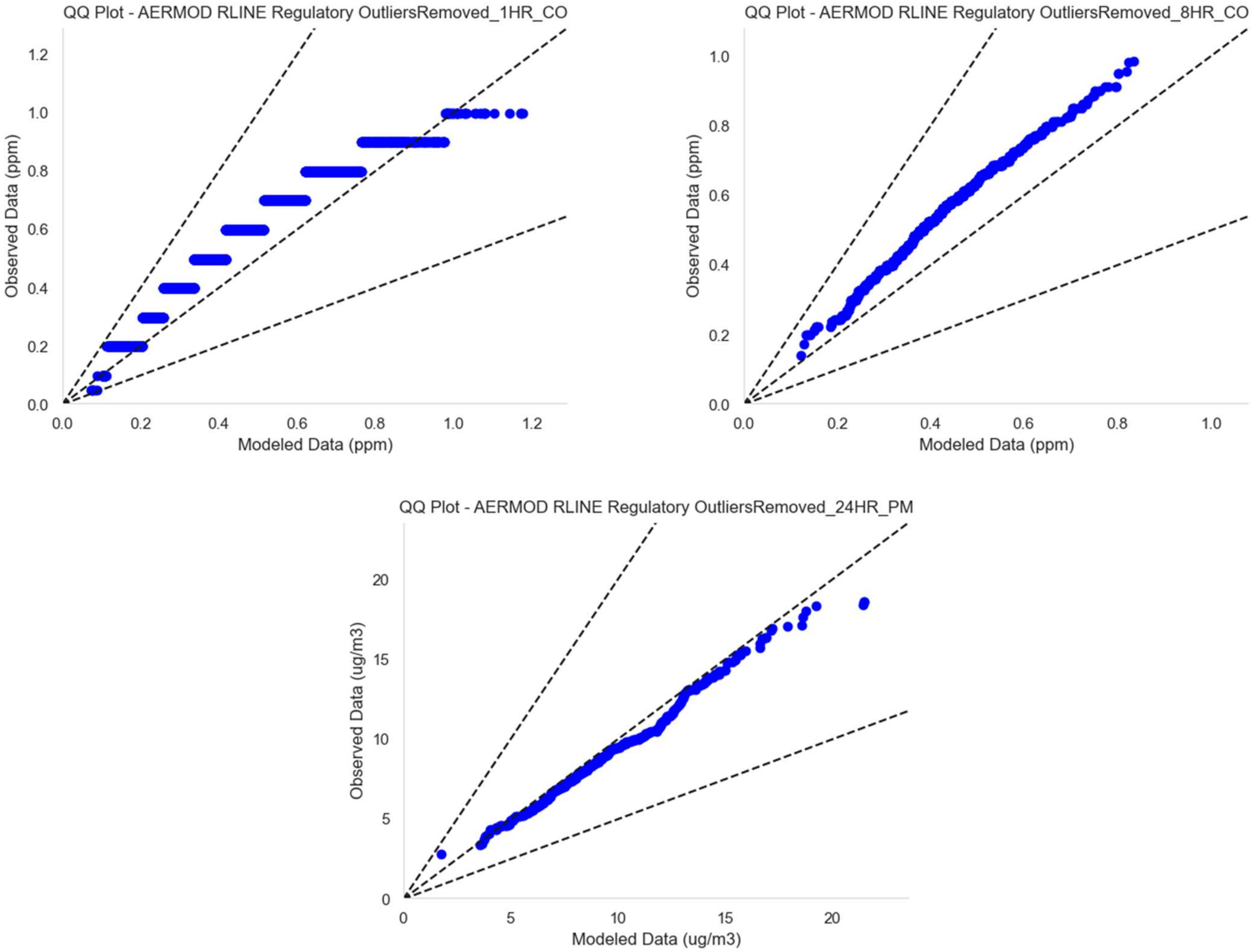
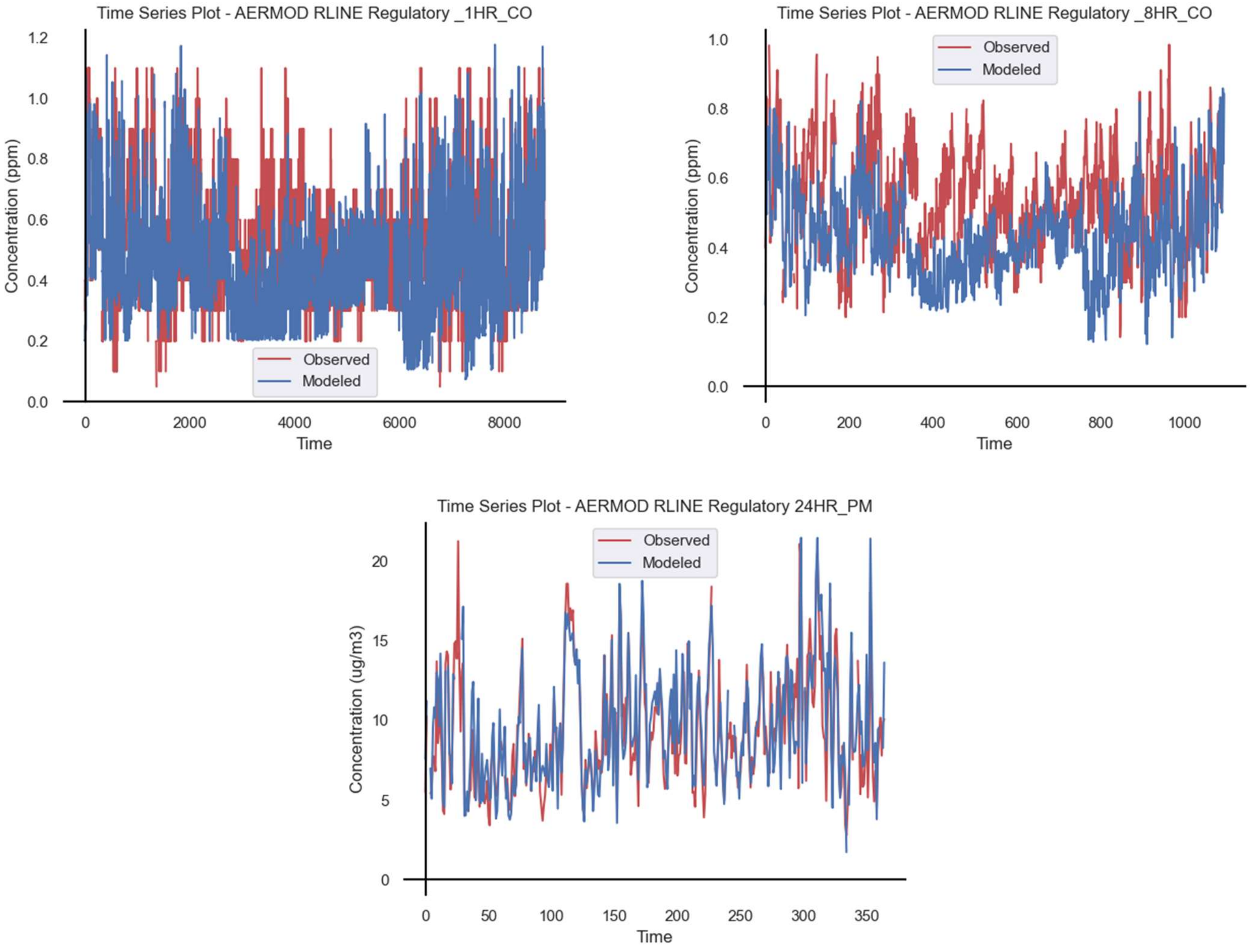


Figure 5-21. Timeseries Plots of Observed vs. AERMOD-RLINE Modeled Concentrations for Regulatory Evaluation for 2019. For CO, x-axis represents 1-hr increments for 1-hr averaged CO and 8-hour increments for 8-hr averaged CO, while for $PM_{2.5}$, the x-axis represents 24-hr increments.



5.1.4. ADMS-Roads

5.1.4.1. Tracer Evaluation

Table 5-10 provides evaluation metrics for ADMS-Roads when compared against observed data under the tracer evaluation. As indicated by the residual values, the ADMS-Roads model underestimates the observed concentrations of all three tracers with PMCH showing the least amount of underestimation. This is also evident from the magnitude of the FBs being lower for PMCH than the other two tracers, indicating better performance of the model in estimating PMCH concentrations. While for PMCH the FB value is less than 0.67, for the other two tracers, the FB values are greater than 0.67 illustrating that modeled concentration for these tracers are underestimated by more than factor of two. The correlation coefficient (R-squared) ranged between 0.70 to 0.82 across the tracers, similar to other models.

Table 5-10. AMDS-Roads Quantitative Evaluation Metrics for Tracer Evaluation (ppt unless otherwise stated)

Metric	PDCB	PMCH	PMCP	Total
Average Residual	54.7	12.9	49.1	116.7
Absolute Average Residual	54.7	14.0	49.1	116.7
Standard Dev. of Residual	17.1	15.1	31.2	58.2
Root Mean Square Error	57.3	19.9	58.2	130.4
R-squared (unitless)	0.74	0.70	0.82	0.80
Fractional Bias (unitless)	0.886	0.355	0.871	0.756
Robust Highest Concentration – Observed	194.8	114.2	261.1	563.8
Robust Highest Concentration – Modeled	99.4	70.8	100.4	261.7

Figure 5-22 and Figure 5-23 are the scatter and Q-Q plots when ADMS-Roads modeled concentrations are compared against observed data. Blue dots represent concentrations at near-field receptors (less than 50 m from the roadway) while red dots represent concentrations at far-field receptors (≥50 m receptors). Unlike AERMOD, the performance of ADMS-Roads did not substantially vary for near- vs. far-field receptors. That is, the concentrations for ADMS, considering the total of all tracers, lies neatly near the 2:1 implying (except for PMCH) that the model consistently underestimates observed values by about a factor of 2. Whereas AERMOD-AREAPOLY, for example, shows differences in performance in the blue (less than 50m) and red (more than 50m) receptors (Figure 5-1).

Figure 5-22. Scatter Plots of Observed vs. ADMS-Roads Modeled Concentrations for Tracer Study (n=51)

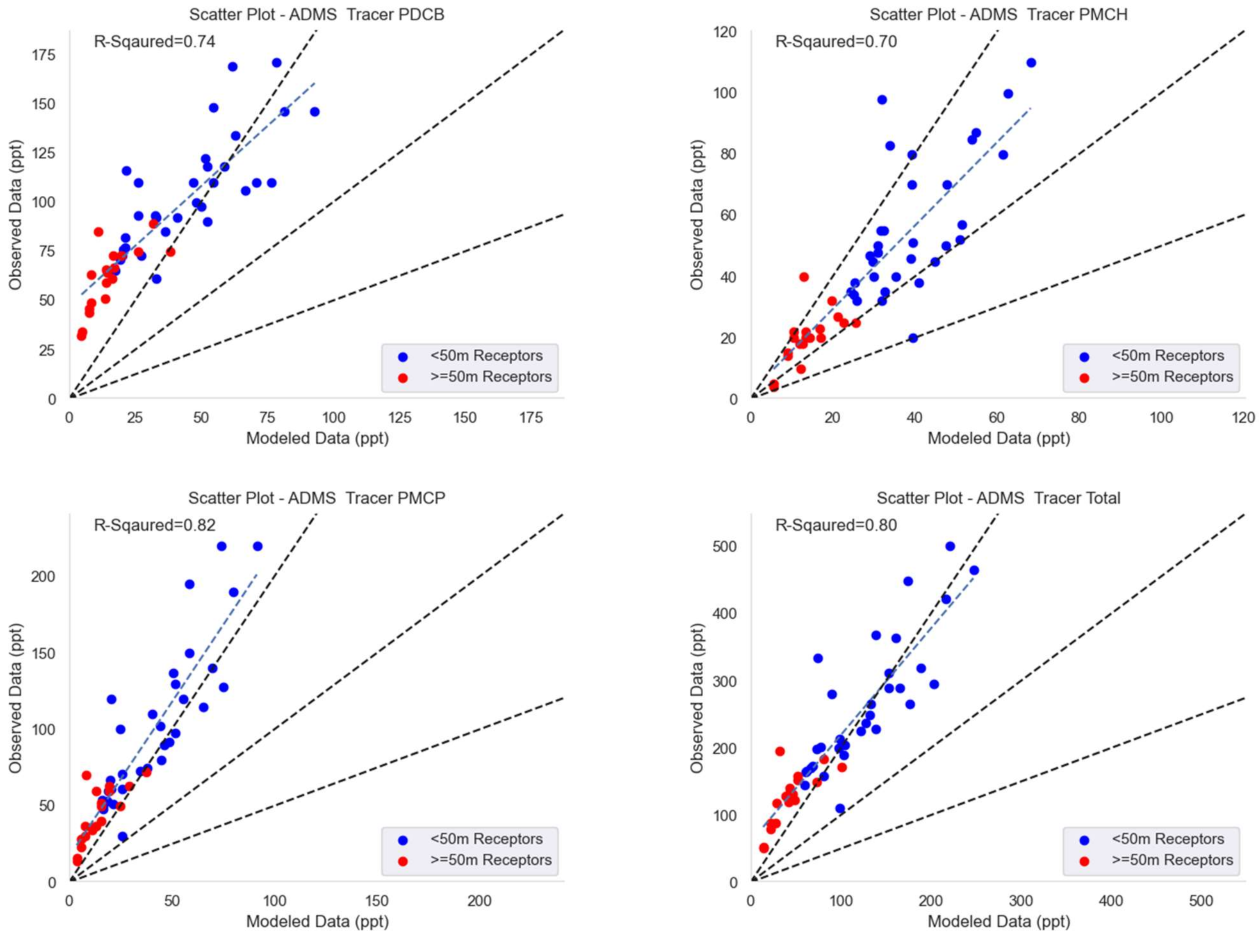
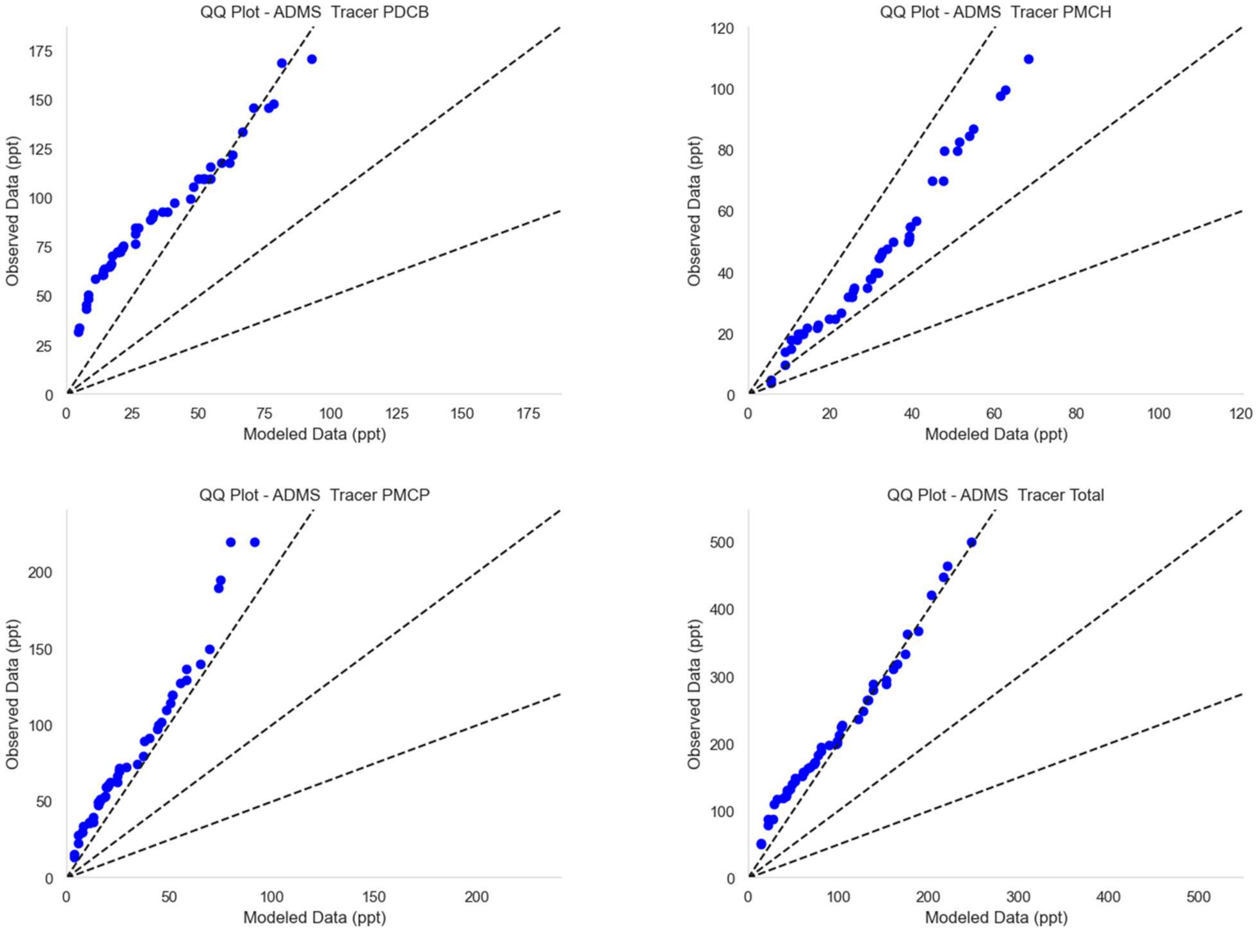


Figure 5-23. Q-Q Plots of Observed vs. ADMS-Roads Modeled Concentrations for Tracer Study (n=51)



5.1.4.2. Pollutant Evaluation

Table 5-11 provides evaluation metrics for ADMS-Roads when compared to observations under the pollutant-based evaluation of CO, $PM_{2.5}$, and BC. As with the other model/source configurations, ADMS-Roads also tended to underestimate concentrations of all three pollutants. This is evident from the FB values being underpredict for all three pollutants. Similar to the other source configurations, while for $PM_{2.5}$ the Fractional Bias value is less than 0.67, for the other two pollutants (CO and BC), the FB values are greater than 0.67 illustrating that modeled concentration for these pollutants are underestimated by more than factor of two. Also, comparing the Robust Highest Concentration (RHC) values demonstrated the model's poor performance in capturing the maximum observed concentrations.

Table 5-11. ADMS-Roads Quantitative Evaluation Metrics for Pollutant Evaluation

Metric	CO (ppm)	$PM_{2.5}$ (µg/m³)	BC (µg/m³)
Average Residual	0.22	0.28	0.51
Absolute Average Residual	0.22	1.87	0.51
Standard Dev. of Residual	0.09	2.98	0.22
Root Mean Square Error	0.23	2.99	0.55
R-squared (unitless)	0.10	0.34	0.01
Fractional Bias (unitless)	1.158	0.186	1.376
Robust Highest Concentration – Observed	0.55	4.92	1.13
Robust Highest Concentration – Modeled	0.14	2.44	0.20

Figure 5-24 and Figure 5-25 are the scatter and Q-Q plots when ADMS-Roads modeled concentrations are compared against observed data. As illustrated, the modeled concentrations are significantly lower than observed data for CO and BC, while for $PM_{2.5}$ the modeled concentrations were within factor of two (except for two data points). As noted for the other models, these comparisons include performance of both the dispersion model and other parts of the modeling chain, including emissions modeling.

Figure 5-24. Scatter Plots of Observed vs. ADMS-Roads Modeled Concentrations for Pollutant Evaluation (n=9)

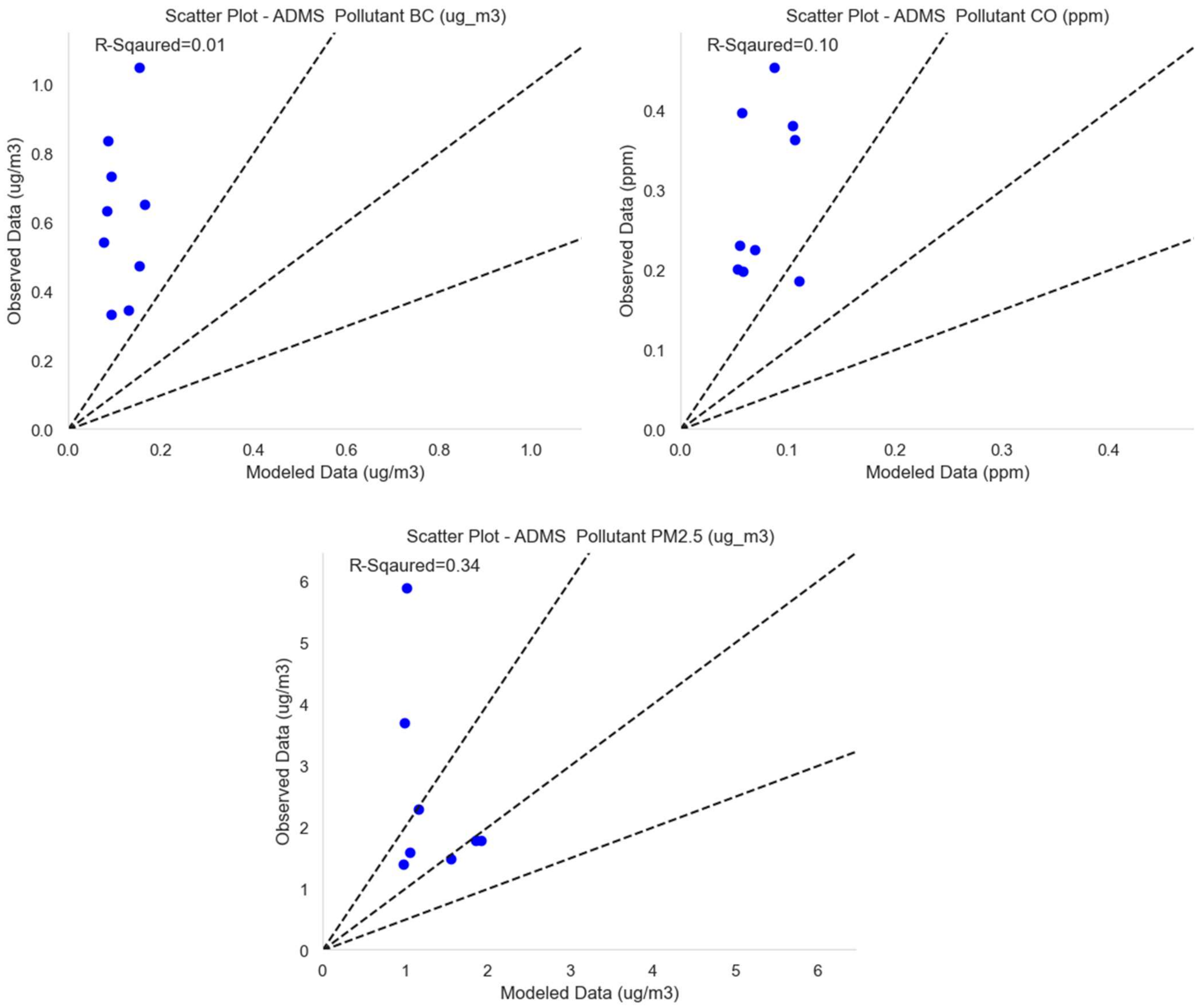
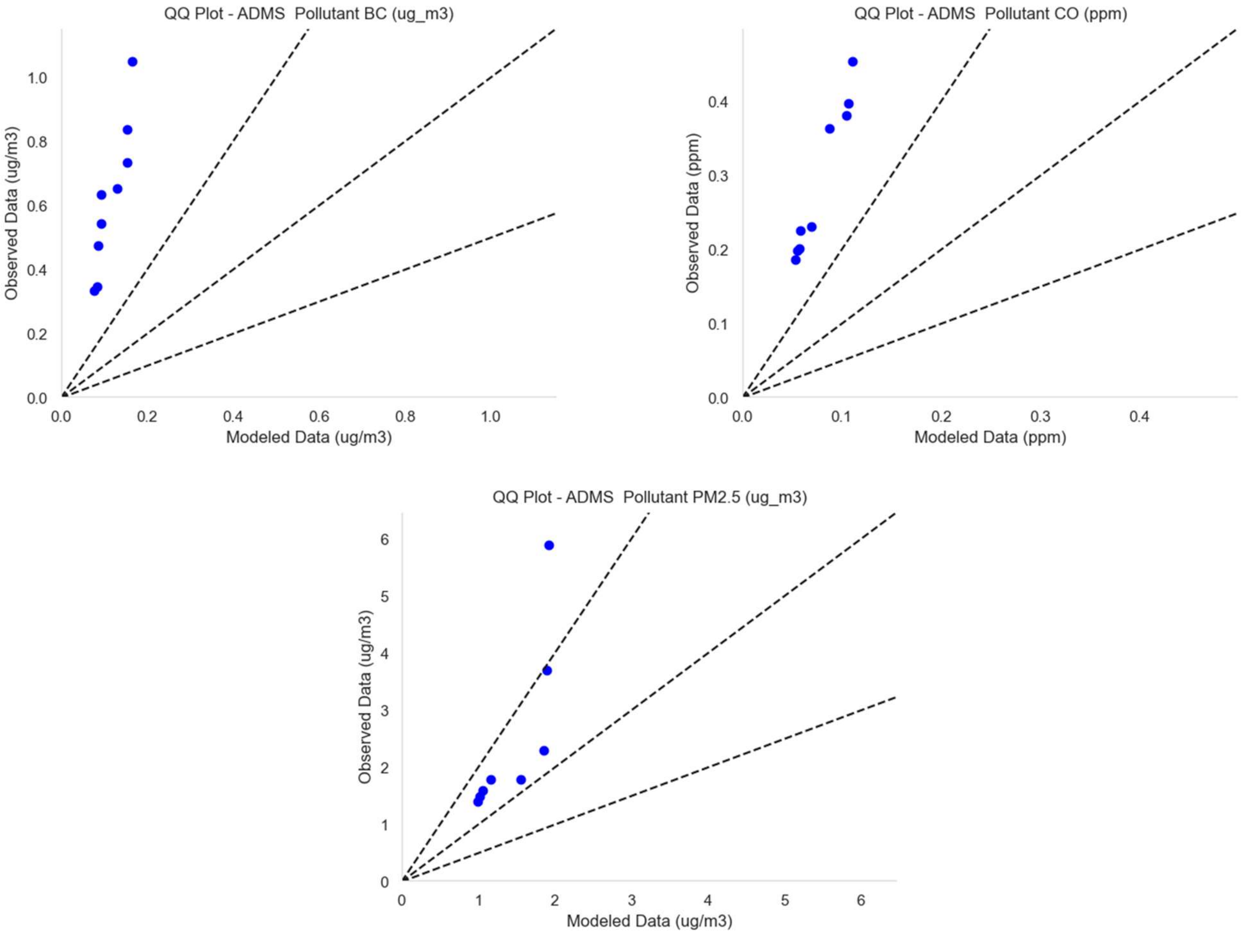


Figure 5-25. Q-Q Plots of Observed vs. ADMS-Roads Modeled Concentrations for Pollutant Evaluation (n=9)



5.1.4.3. Regulatory Evaluation

Table 5-12 provides a comparison of the DVs for CO and $PM_{2.5}$ concentrations against the DVs calculated based on the near-road monitoring station. As described earlier, two sets of comparisons are provided: a) DVs with outliers, and b) DVs without outliers. As shown in Table 5-12, when outliers were removed, DVs calculated through ADMS-Roads were fairly consistent with DVs calculated from observed data. For 1-hour averaged CO, DVs were different by 0.5 ppm; for 8-hour averaged CO, DVs were different by 0.1 ppm; for 24-hour $PM_{2.5}$, the modeled DV was 1 μg/m^3 higher than observed; and for annual $PM_{2.5}$, the modeled design value was 0.9 μg/m^3 higher than the observed one.

Table 5-12. ADMS-Roads Design Values for Regulatory Evaluation[44]

Averaging Period and Pollutant	With Outliers		Without Outliers	
	Modeled	Observed	Modeled	Observed
1-hour CO (ppm)	1.6	5.6	1.6	1.1
8-hour CO (ppm)	1.1	1.2	1.1	1.0
24-hour $PM_{2.5}$ (μg/m^3)	20	20	20	19
Annual $PM_{2.5}$ (μg/m^3)	10.2	9.5	10.2	9.3

Figure 5-26, Figure 5-27, and Figure 5-28 are scatter, Q-Q, and time-series plots, respectively, comparing modeled vs. observed 1-hour and 8-hour average CO as well as 24-hour average $PM_{2.5}$ concentrations. Similar to the pollutant-based evaluation, the Q-Q plot demonstrates lower modeled CO concentrations as compared to the observed data. As described earlier, performance for the regulatory applications is a function of the dispersion model, the emissions modeling, and the background concentrations. (See Section 5.1.1.3). Note that, as discussed above, longer term averages inherently show lower scatter.

---

[44] Consistent with U.S. EPA reporting of design values, the DVs for 1-hr CO, 8-hr CO, and annual PM2.5 are reported with one decimal, while it is rounded to nearest whole number for 24-hr PM2.5.

Figure 5-26. Scatter Plots of Observed vs. ADMS-Roads Modeled Concentrations for Regulatory Evaluation (n=8,760 for 1 hour CO, n=1,095 for 8-hour CO, and n=365 for 24-hr $PM_{2.5}$)

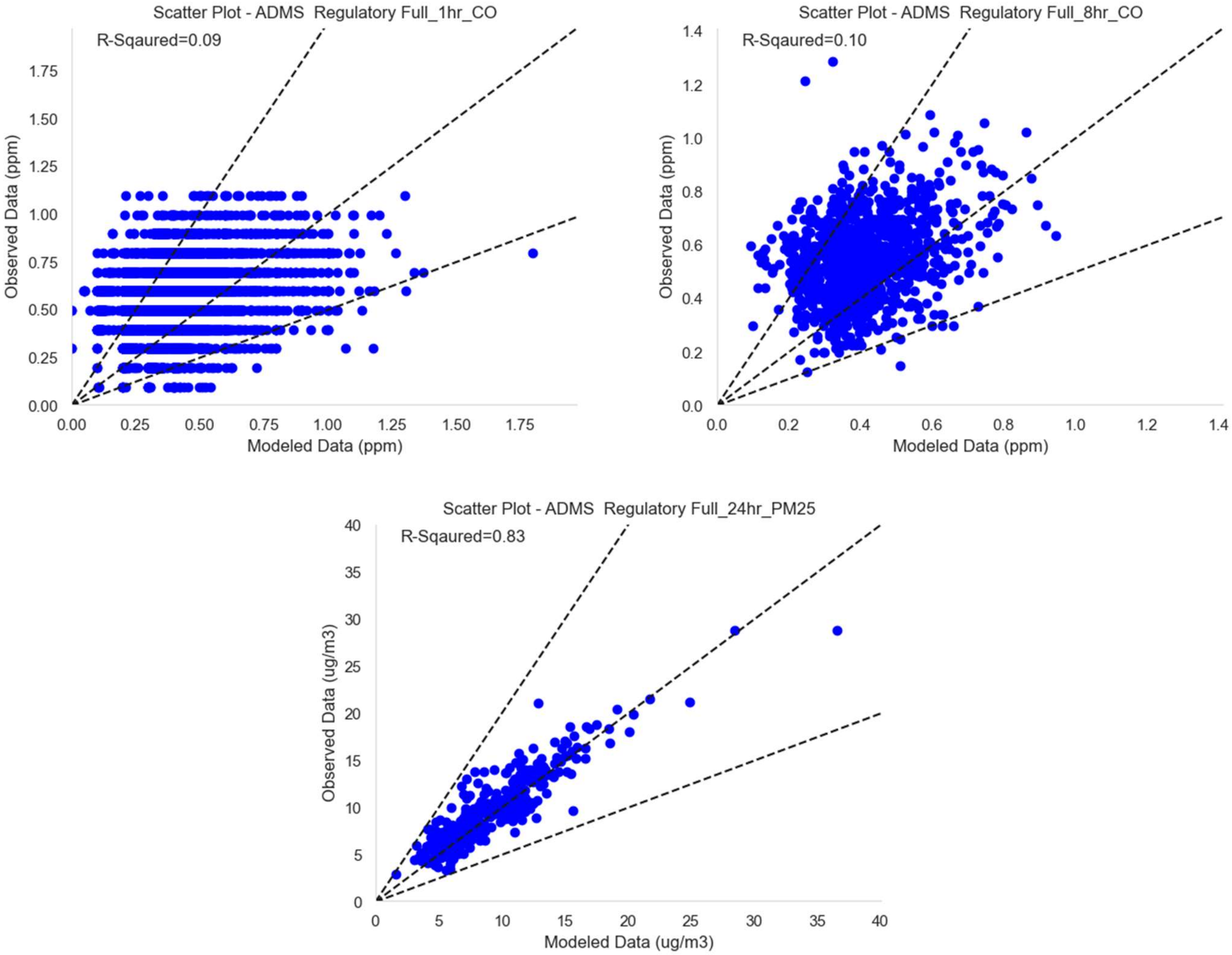


Figure 5-27. Q-Q Plots of Observed vs. ADMS-Roads Modeled Concentrations for Regulatory Evaluation (n=8,760 for 1 hour CO, n=1,095 for 8-hour CO, and n=365 for 24-hr $PM_{2.5}$)

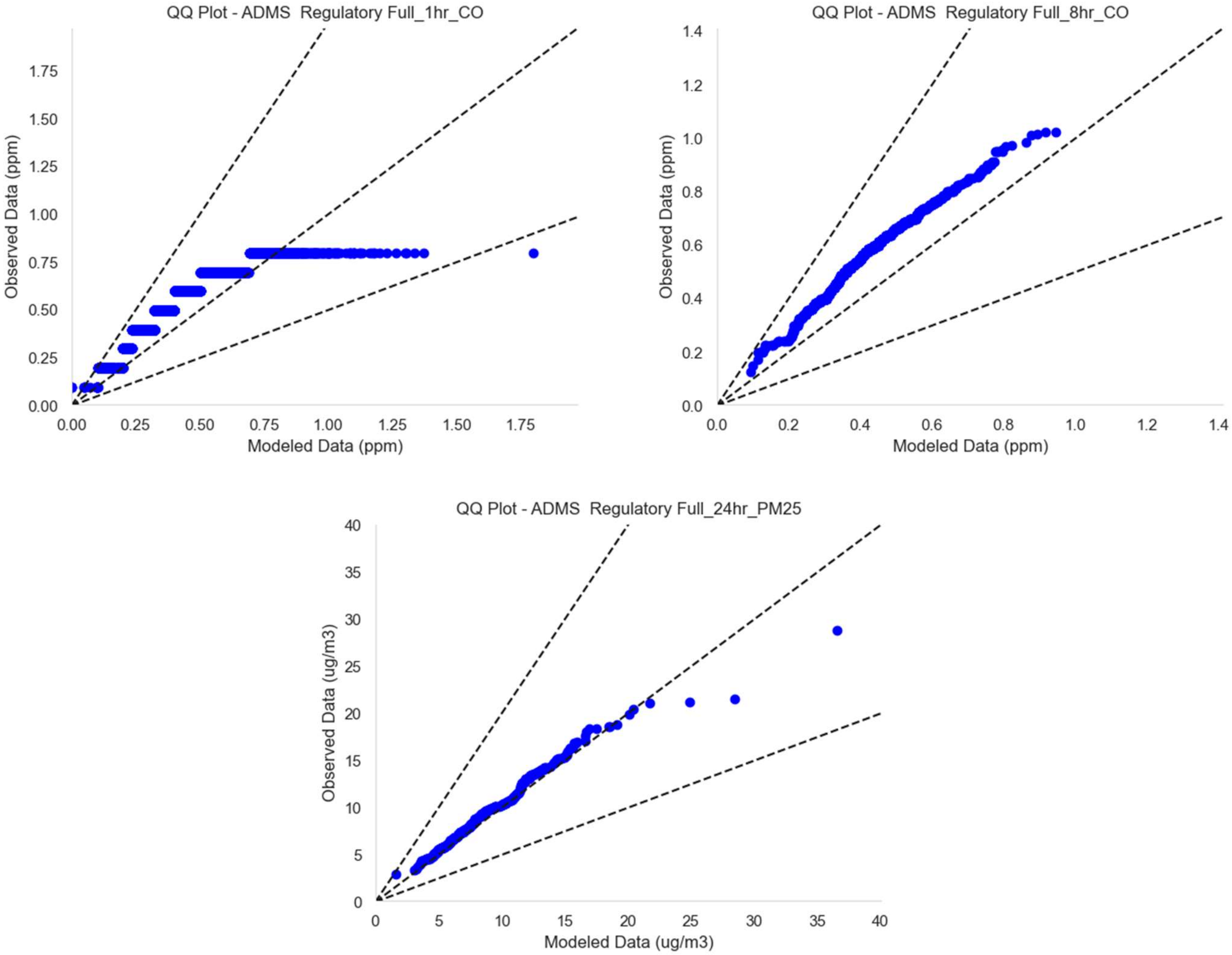
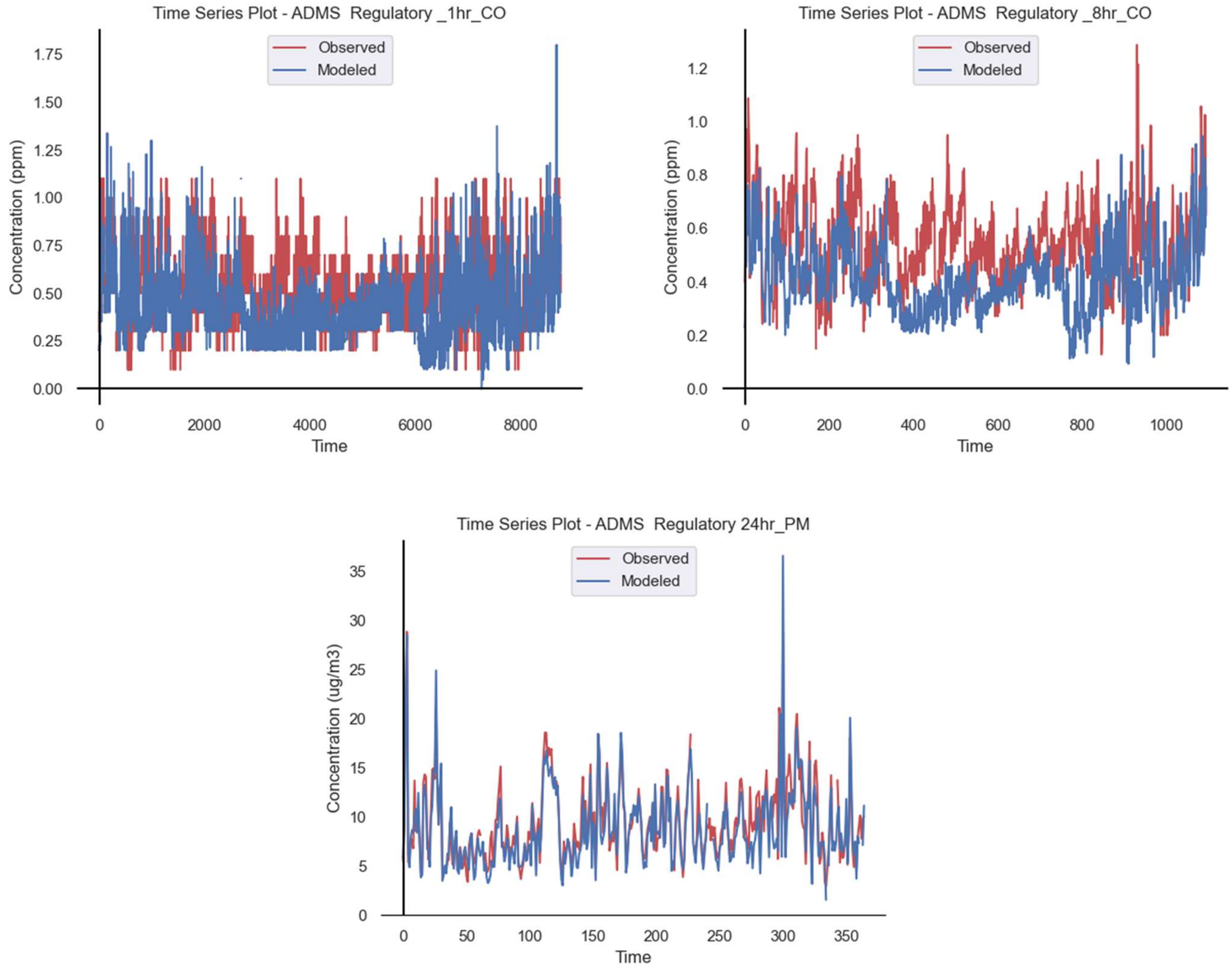


Figure 5-28. Timeseries Plots of Observed vs. ADMS-Roads Modeled Concentrations for Regulatory Evaluation in 2019. For CO, x-axis represents 1-hr increments for 1-hr averaged CO and 8-hour increments for 8-hr averaged CO, while for $PM_{2.5}$, the x-axis represents 24-hr increments.



5.1.5. AERSCREEN

Table 5-13 provides a comparison of AERSCREEN maximum modeled concentrations versus maximum observed concentrations for tracer, pollutant, and regulatory evaluations[45].

Note that for observed regulatory data, we show only data without outliers. (See Section 5.1.1.3.) As expected, AERSCREEN overpredicted maximum concentration between a factor of 3.5-8.8 times higher. Specifically on average, ratios of maximum modeled to maximum observed concentrations ranged between 3.5 for 24-hr $PM_{2.5}$ concentrations under regulatory evaluation to 8.8 for hourly $PM_{2.5}$ under pollutant evaluation and hourly PMCH under tracer evaluation. For the regulatory evaluation, note that there are no recommended time-average scaling factors for area sources in AERSCREEN, and therefore the model-to-observed comparisons for 8-hour CO and 24-hour $PM_{2.5}$ involved the maximum 1-hour AERSCREEN results versus the 8-hour CO and 24-hour $PM_{2.5}$ observed values.

Table 5-13. Maximum Concentration Estimated by AERSCREEN vs. Observed for Tracer, Pollutant, and Regulatory Evaluations. There are no values available for cells colored in dark gray.

Maximum Concentration Value	Pollutant-based		Regulatory		Tracer	
	AERSCREEN	Observed	AERSCREEN	Observed	AERSCREEN	Observed
Hourly CO (ppm)	3.2	0.45	4.6	1.10		
Hourly $PM_{2.5}$ ($\mu g/m^3$)	51.9	5.9				
Hourly BC ($\mu g/m^3$)	4.2	1.1				
8-hour CO (ppm)			4.2	1.03		
24-hr $PM_{2.5}$ ($\mu g/m^3$)			99.7	28.8		
Annual $PM_{2.5}$ ($\mu g/m^3$)						
Hourly PDCB (ppt)					949	171
Hourly PMCP (ppt)					869	220
Hourly PMCH (ppt)					971	110
Hourly Total Tracer (ppt)					2,788	501

---

[45] Note that for regulatory based evaluation, maximum of the background concentration is added to the maximum modeled concentrations to compare against maximum observed concentration from the near-road monitor. For Pollutant based evaluation, the maximum of modeled concentrations is compared against maximum of background corrected observed concentrations (i.e., near-road monitor – background concentration).

5.1.6. CAL3QHC

Table 5-14 provides a comparison of CAL3QHC maximum modeled concentrations versus maximum observed concentrations for tracer, pollutant, and regulatory evaluations[46].

As expected, CAL3QHC overpredicted maximum concentration by a factor from 1.3 to 2.6 times higher. On average, ratios of maximum modeled to maximum observed concentrations ranged between 1.3 for 8-hour CO concentrations under regulatory evaluation and 1-hour CO concentrations under pollutant evaluation to 2.6 for hourly PMCH and PDCB under tracer evaluation. **Compared to AERSCREEN, CAL3QHC resulted in significantly lower maximum modeled-to-observed concentration ratios while still overpredicting maximum concentrations under all examined scenarios**. Thus, CAL3QHC should continue to remain available as a regulatory CO screening model capability as it provides a continued usefulness to successfully screen potential projects and minimize the burden for refined modeling analysis, whereas AERSCREEN screening may more frequently trigger refined modeling**.** For the regulatory evaluation, note that we used the recommended default persistence factor (0.7) for CAL3QHC results to obtain 8-hour CO averages, though it is possible that more current CO background conditions may be less time variable leading to a higher persistence factor. CAL3QHC and AERSCREEN also used different meteorological values as dictated by the model input requirements. (See Sections 2.2.6.3 and 2.2.6.5.)

Table 5-14. Maximum Concentration Estimated by CAL3QHC vs. Observed for Tracer, Pollutant, and Regulatory Evaluations. There are no values available for cells colored in dark gray.

Maximum Concentration Value	Pollutant-based		Regulatory		Tracer	
	CAL3QHC	Observed	CAL3QHC	Observed	CAL3QHC	Observed
Hourly CO (ppm)	0.6	0.45	1.90	1.10		
Hourly $PM_{2.5}$ ($\mu g/m^3$)						
Hourly BC ($\mu g/m^3$)						
8-hour CO (ppm)			1.31	1.03		
24-hr $PM_{2.5}$ ($\mu g/m^3$)						
Annual $PM_{2.5}$ ($\mu g/m^3$)						
Hourly PDCB (ppt)					438.2	171
Hourly PMCP (ppt)					478.8	220
Hourly PMCH (ppt)					283.9	110
Hourly Total Tracer (ppt)					1146	501

---

[46] Note that for regulatory based evaluation, maximum of the background concentration is added to the maximum modeled concentrations to compare against maximum observed concentration from the near-road monitor. For Pollutant based evaluation, the maximum of modeled concentrations is compared against maximum of background corrected observed concentrations (i.e., near-road monitor – background concentration).

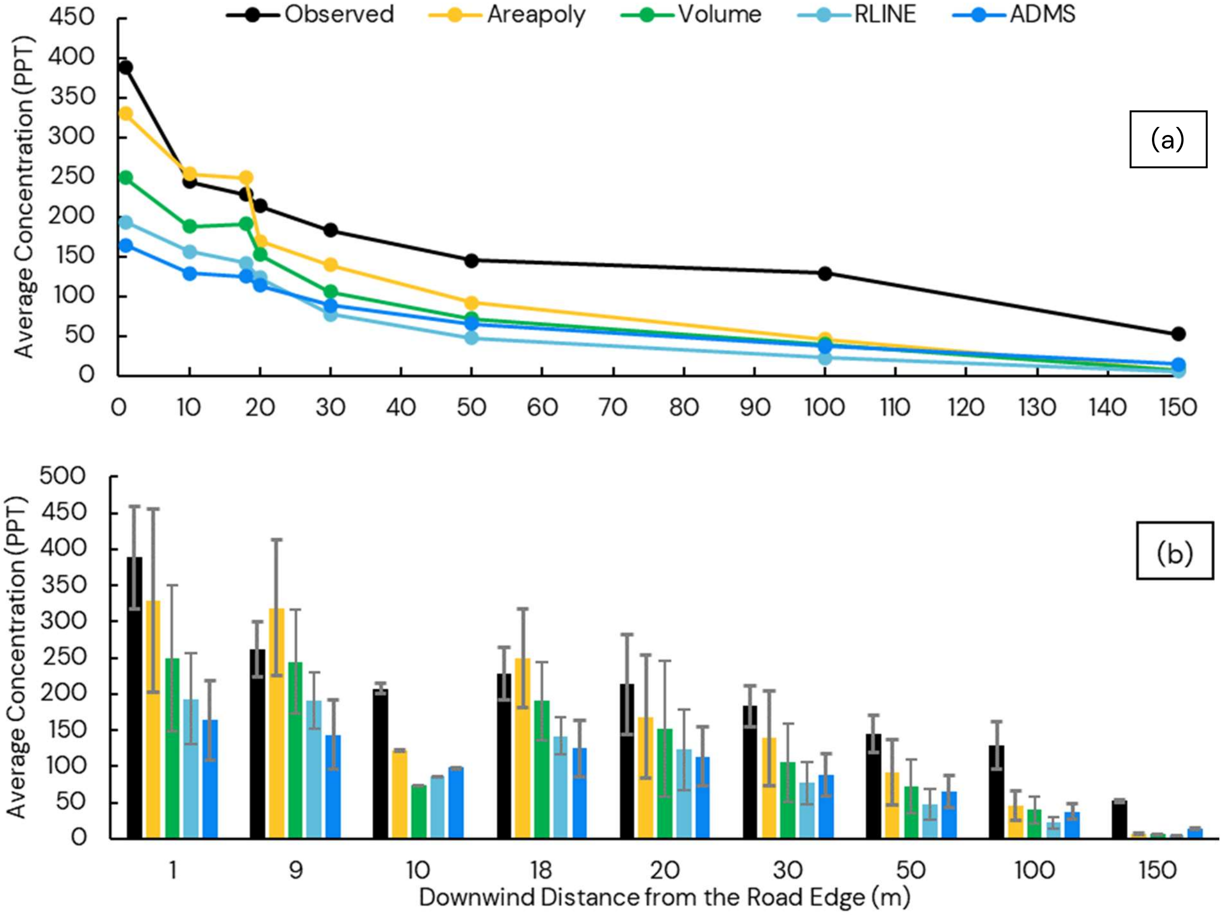
5.2. Quantitative Results Across Models

Table 5-15, Table 5-16, Table 5-17, and Table 5-18 provide a side-by-side comparison of model performance metrics under the tracer evaluation for PDCB, PMCH, PMCP, and the total of all tracers, respectively.

Interestingly, while the performance of models varied across different tracers, the relative performance of models against each other did not vary substantially. For example, if AERMOD-AREAPOLY demonstrated lower absolute average residual than other models for PDCB, it also showed similar relative performance for other tracers as well. This is consistent with the idea that each vehicle's release had relatively steady average concentrations over the experimental hour. Table 5-15 compares the quantitative metrics across models for the tracer evaluation for the sum of all three tracers. Of the four models examined for the tracer study, AERMOD with source configuration of AREAPOLY demonstrated relatively better performance when compared to the other three models/source configurations tested. AERMOD-AREAPOLY had relatively lower absolute average residual, fractional bias, RMSE and importantly the best pairing of the RHC observed to modeled. On the other hand, the ADMS-Roads model showed relatively poorer performance in estimating the observed tracer concentration as compared to the other three models. The ADMS-Roads model had the highest average residual, RMSE, and fractional bias. All four model/source configurations showed similar R-squared values. It should be noted that aside from differences embedded in the methodologies, these models often take different set of inputs which could also lead to differences in their performance. For inputs that were common across different model, the project team used consistent assumptions and inputs.

All four models also showed a small tendency for underpredictions, as demonstrated by fractional bias values. However, this underprediction tends to occur most at lower concentrations and further distances downwind from the roadway. This is also seen in Figure 5-29, which shows observed and modeled values for the average concentration of the sum of all three tracers versus the distance from the roadway at which the measurements were made. This combines both the north and south experiment locations and averages over all experiments, meaning that two different locations are shown aggregated on this chart. As an average over all experiments there is substantial variation in concentrations at each location. We have shown these results in two parts. Figure 5-29 (a) shows the average concentrations as a function of distance, without any measure of the variance in the values. This demonstrates the general shape of the dispersion curves from each of the models. Figure 5-29 (b) shows the average and one standard deviation of the concentration at each measurement location. This emphasizes how well the models perform at each distance when considering the variation across experiments. Note that receptor locations were not the same in all experiments (on the southern end of the project, while the monitors were deployed in the same general area along a stretch of side road, the exact locations differed from one experiment day to the next, and the May 3 experiment utilized an additional row of air monitors).

Figure 5-29. Average observed and modeled concentration as a function of downwind distance. All tracers and all experiments combined. (a) average over all experiments, (b) the average and one standard deviation.



In many applications it is the highest predicted concentrations that matter most, such as for demonstrating project level conformity with air quality standards. Performance for peak concentrations can be assessed by the high-concentration end of the points shown in the Q-Q plots for each model, such as Figure 5-2, and the RHC values in Table 5-18. In both cases, AERMOD-AREAPOLY is the best model configuration closely representing the highest concentrations. While it overestimates concentrations at 9 and 18 m receptors, and slightly underestimates concentrations at other distances, the modeled concentrations from AERMOD-AREAPOLY are closely representing the observed concentrations. All other models show a higher underprediction bias, with ADMS performing the worst, at roughly half the observed value. Figure 5-30 illustrates the performance of models for the receptor nearest the road in all experiments, where the highest concentrations are expected to occur. Of the four models, AERMOD-AREAPOLY has the closest mean and maximum concentrations at 1-m receptor to observed data. ADMS-ROADS, by contrast, shows maxima below the minimum observed values at this important location. We suspect that the cause for some of this better model performance for AREAPOLY characterization is that given the 10 lanes of freeway and close proximity of the receptors to the freeway that the emissions are best characterized as an area source rather than a line or set of volume sources. For freeways with fewer lanes and receptors at distances further from the road we would expect better agreement between all of the source characterization types.

Figure 5-30. Box plots illustrating the observed and modeled concentration of all tracers (total) at 1-m receptors across all experiment hours

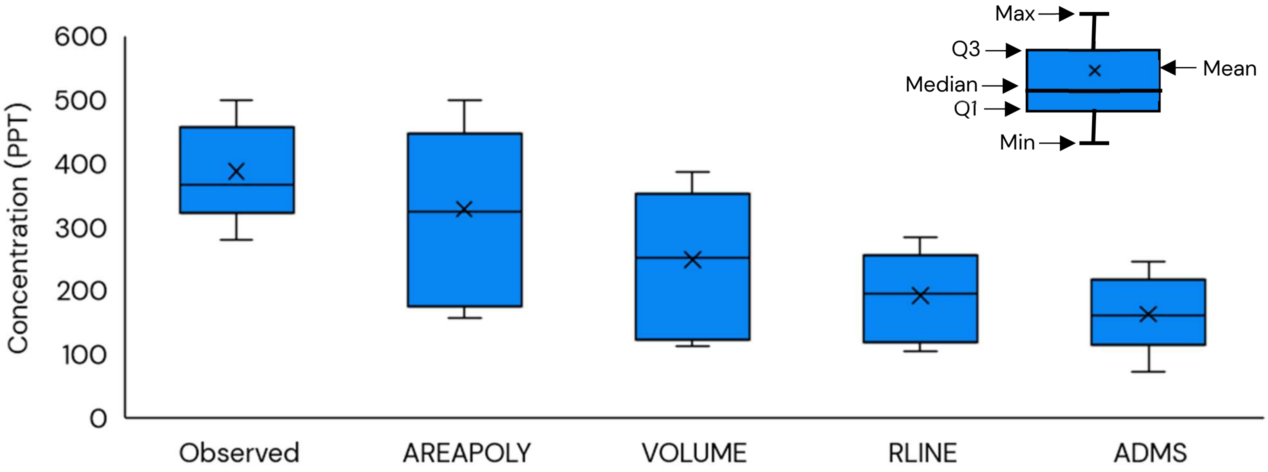


Table 5-15. Tracer Evaluation Metrics Across All Refined Models/Source Configurations – PDCB Tracer (ppt unless otherwise stated)

Metric	AERMOD			ADMS Roads
	AREAPOLY	Volume	RLINE	
Average Residual	29.6	42.3	53.8	54.7
Absolute Average Residual	35.1	42.7	53.8	54.7
Standard Dev. of Residual	26.1	20.7	15.0	17.1
Root Mean Square Error	39.4	47.1	55.8	57.3
R-squared (unitless)	0.77	0.72	0.78	0.74
Fractional Bias (unitless)	0.398	0.623	0.864	0.886
Robust Highest Concentration – Observed	194.8	194.8	194.8	194.8
Robust Highest Concentration – Modeled	204.3	154.4	106.9	99.4

Table 5-16. Tracer Evaluation Metrics Across All Refined Models/Source Configurations – PMCH Tracer (ppt unless otherwise stated)

Metric	AERMOD			ADMS Roads
	AREAPOLY	Volume	RLINE	
Average Residual	-0.38	9.4	16.5	12.9
Absolute Average Residual	11.5	12.6	17.4	14.0
Standard Dev. of Residual	15.3	13.7	13.8	15.1
Root Mean Square Error	15.3	16.6	21.5	19.9
R-squared (unitless)	0.74	0.72	0.73	0.7
Fractional Bias (unitless)	-0.009	0.247	0.477	0.355
Robust Highest Concentration – Observed	114.2	114.2	114.2	114.2
Robust Highest Concentration – Modeled	144.4	97.8	72.4	70.8

Table 5-17. Tracer Evaluation Metrics Across All Refined Models/Source Configurations – PMCP Tracer (ppt unless otherwise stated)

Metric	AERMOD			ADMS Roads
	AREAPOLY	Volume	RLINE	
Average Residual	15.9	29.7	42.5	49.1
Absolute Average Residual	23.5	31.2	42.5	49.1
Standard Dev. of Residual	23.5	21.9	24.5	31.2
Root Mean Square Error	28.3	36.9	49.0	58.2
R-squared (unitless)	0.83	0.81	0.83	0.82
Fractional Bias (unitless)	0.217	0.449	0.712	0.871
Robust Highest Concentration – Observed	261.1	261.1	261.1	261.1
Robust Highest Concentration – Modeled	234.2	181.5	124.7	100.4

Table 5-18. Tracer Evaluation Metrics Across All Refined Models/Source Configurations – Sum of All Tracers (PDCB + PMCH + PMCP) (ppt unless otherwise stated)

Metric	AERMOD			ADMS Roads
	AREAPOLY	Volume	RLINE	
Average Residual	45.0	81.4	112.7	116.7
Absolute Average Residual	62.8	82.5	112.7	116.7
Standard Dev. of Residual	59.1	49.9	47.7	58.2
Root Mean Square Error	74.3	95.4	122.4	130.4
R-squared (unitless)	0.82	0.79	0.82	0.8
Fractional Bias (unitless)	0.236	0.472	0.720	0.756
Robust Highest Concentration – Observed	563.8	563.8	563.8	563.8
Robust Highest Concentration – Modeled	585.3	425.9	311.6	261.7

Table 5-19, Table 5-20, and Table 5-21 provide a side-by-side comparison of model performance metrics under the pollutant evaluation for CO, BC, and $PM_{2.5}$, respectively. As described earlier, unlike the tracer evaluations, these comparisons are highly dependent upon the emission rates used in the modeling.

Generally, all four model/source configurations underestimated concentrations of CO and BC, while their estimates were mostly within factor of two for $PM_{2.5}$. It is noteworthy to mention that while CO and BC emission rates were all based on the EMFAC2021 model, for $PM_{2.5}$, we used a combination of EMFAC2021 vehicle emission rates and road dust emission factors based on EPA's AP-42 approach and California-specific silt loading factors. (See Section 4.7.) All models had low r-squared values in comparison to the tracer experiments suggesting much of the uncertainty is in other parts of the modeling chain, including the emission rates used in the modeling.

Table 5-19. Pollutant-based Evaluation Metrics Across All Refined Models/Source Configurations – CO (ppm unless otherwise stated)

Metric	AERMOD			ADMS Roads
	AREAPOLY	Volume	RLINE	
Average Residual	0.18	0.191	0.23	0.22
Absolute Average Residual	0.18	0.191	0.23	0.22
Standard Dev. of Residual	0.09	0.083	0.09	0.09
Root Mean Square Error	0.20	0.21	0.24	0.23
R-squared (unitless)	0.22	0.39	0.14	0.10
Fractional Bias (unitless)	0.894	0.963	1.250	1.158
Robust Highest Concentration – Observed	0.55	0.55	0.55	0.55
Robust Highest Concentration – Modeled	0.19	0.169	0.12	0.14

Table 5-20. Pollutant-based Evaluation Metrics Across All Refined Models/Source Configurations – BC (μg/m^3 unless otherwise stated)

Metric	AERMOD			ADMS Roads
	AREAPOLY	Volume	RLINE	
Average Residual	0.43	0.45	0.50	0.51
Absolute Average Residual	0.43	0.45	0.50	0.51
Standard Dev. of Residual	0.23	0.24	0.22	0.22
Root Mean Square Error	0.49	0.51	0.54	0.55
R-squared (unitless)	0.02	0.12	0.00	0.01
Fractional Bias (unitless)	1.060	1.121	1.319	1.376
Robust Highest Concentration – Observed	1.13	1.13	1.13	1.13
Robust Highest Concentration – Modeled	0.30	0.29	0.19	0.20

Table 5-21. Pollutant-based Evaluation Metrics Across All Refined Models/Source Configurations – PM$_{2.5}$ (μg/m^3 unless otherwise stated)

Metric	AERMOD			ADMS Roads
	AREAPOLY	Volume	RLINE	
Average Residual	-0.45	-0.28	0.32	0.28
Absolute Average Residual	1.9	1.9	1.8	1.9
Standard Dev. of Residual	2.9	2.8	2.9	3.0
Root Mean Square Error	3.0	2.8	2.9	3.0
R-squared (unitless)	0.08	0.003	0.17	0.34
Fractional Bias (unitless)	-0.236	-0.152	0.216	0.186
Robust Highest Concentration – Observed	4.9	4.9	4.9	4.9
Robust Highest Concentration – Modeled	3.6	3.1	2.2	2.4

Table 5-22 provides a comparison for the regulatory evaluation of calculated DVs using the four model/source configuration versus observed values. Of the four models/configurations examined, AERMOD-AREAPOLY showed modeled DVs that were closest to the observed values, whereas ADMS-Roads had the highest deviation from the observed values. Only ADMS-ROADS overpredicted for 8-hour CO. All models except AERMOD in the area source configuration overpredicted the annual average PM$_{2.5}$ design value. As described earlier, while all four models/configurations showed DVs that were close to observed values, a high fraction (more than 80 percent) of modeled DVs were driven by background concentrations, so only 20 percent of modeled DVs were due to near-receptor, traffic-related emissions. Also, the results are dependent on both the dispersion models and most importantly the emission factor model (i.e., the modeling chain). Therefore, we cannot draw specific conclusion about performance of the near-road dispersion modeling component of this modeling chain in isolation as the emission rates and background concentrations are inputs to the dispersion model as used in this regulatory context. The purpose of the regulatory evaluation is to evaluate the performance of modeling chain (traffic, emission, and dispersion) in the regulatory context whereas the tracer evaluation is a better mean to assess the performance of the dispersion model alone and the pollutants-based evaluation for roadway impacts with a more refined set of observations.

Table 5-22. Design Values for Regulatory Evaluation Across All Refined Models/Source Configurations, Without Outliers

Averaging Period and Pollutant	AERMOD			ADMS-Roads	Observed
	AREAPOLY	Volume	RLINE		
1-hour CO (ppm)	1.2	1.1	1.2	1.62	1.1
8-hour CO (ppm)	1.0	0.94	0.9	1.07	0.98
24-hour $PM_{2.5}$ (µg/m^3)	17.9	19.0	20.1	19.5	18.6
Annual $PM_{2.5}$ (µg/m^3)	8.8	9.5	9.7	10.2	9.3

As discussed earlier, considering the uncertainties within the pollutant and regulatory evaluations related to emission factors and upwind background concentrations for the regulatory evaluation, the tracer evaluation is the most reliable method for assessing the performance of dispersion components of the models examined in this study. As shown in Table 5-18, of the four models and source configuration examined, the AERMOD-AREAPOLY showed the best performance in terms of having lower residual values, lower RMSE, lower absolute fractional bias, and having the closest robust highest concentration to the observed values. AERMOD-AREAPOLY also presented a better performance than the other three model/source configurations under pollutant (except for $PM_{2.5}$) and regulatory evaluations. Therefore, the AERMOD-AREAPOLY is the best-performing model based on evaluation metrics used in this study.

With respect to the differences in performance of AERMOD-AREAPOLY and AERMOD-RLINE, Wayson and Voigt (2022)[47] has recently completed a set of sensitivity tests for the AERMOD-RLINE model when applied to roadways which included:

- release height,
- number of lanes,
- elevation of roadway
- initial vertical dimension, and
- winds perpendicular and parallel to roadway.

Their sensitivity tests, conducted under worst-case meteorological conditions, illustrated that at a distance of only 25-m away from the roadway centerline, doubling the height of the release reduces the concentration by half and this parameter is affected by traffic mix, vehicle speed, traffic volume, and width of roadway. Similarly, the downwind concentrations showed a 30-percent decrease when the initial vertical dimension is doubled from 1 to 2 meters. Parallel winds over the freeway showed increasing concentrations with increasing roadway width along the downwind axis. Urban and rural settings were also shown to affect results. These results show the model's sensitivity to these parameters including targeted worst-case meteorological data. Significant

---

[47] Wayson, Roger and Voigt, Christopher, (2022) AERMOD Source Types RLINE and RLINEXT Testing, FHWA Office of Natural Environment and FHWA Virginia Division Office, March.

sensitivities were also observed in regulatory testing using five-years of meteorological data for northern Virginia for initial vertical dispersion coefficients and urban setting.

ICF reviewed the key paper (Gillies et. al, 2005[48]) used by EPA to recommend plume height and initial vertical dimension from vehicles most notably under stable conditions (we used this scaling factor in all AERMOD modeling). The paper reported that the initial vehicle height is scaled by a factor of 1.7 to determine release height. However, this was based on a small fleet of nine vehicles with no vehicle height being higher than 3.3 meters. We suspect that for taller vehicles this 1.7 factor may be an overestimate. Additionally, as noted by Gillies et. al. (2005), in addition to atmospheric stability, the shape of the vehicle, and the angle of the ambient wind with respect to the direction of vehicle travel, play a critical role in determining this scaling height factor. Further study is needed to better estimate this scaling factor. We believe the current guideline is likely overestimating the release height and initial vertical dispersion, so that even when this parameterization is coupled with the better source characterization line source algorithm in AERMOD-RLINE (numerically integrating point sources) the model tends to underpredict observed concentrations.

We would recommend that a field study program be undertaken to provide a more robust estimate of the initial plume rise height under a variety of atmospheric conditions, wind speeds, and roadway widths that includes a minimum of three FHWA vehicle size classifications (Class 1-2, class 3-5, class 6-12) and includes measurements of the vertical concentration profile.

5.3. Qualitative Ease of Use Evaluation

To provide a qualitative evaluation of the four model/source configurations examined in this study (the focus being on the refined models), ICF created a scoring matrix based on the criteria describe earlier in Section 2.3.3. For each evaluation criteria, each model gets a score of 0 through 10. A score of 10 indicates it is relatively strong in that area and a score of 0 indicates that the model has significant relative weakness. Scoring was based on professional judgement and experience by the ICF team that carried out the modeling for the study.

---

[48] Gillies,J.A., V.Etyemezian, H.Kuhns, D.Nikolic, D.A. Gillette, Effect of vehicle characteristics on unpaved road dust emissions, (2005) Atmospheric Environment 39 (2005) 2341–2347.

Table 5-23. Qualitative Evaluation of Dispersion Models

Evaluation Criteria	AERMOD			ADMS-Roads
	AREAPOLY	Volume	RLINE	
User Interface	0	0	0	10
Ease of Use	3	1	3	8
Model Transparency	7	7	7	4
Runtime	10	5	9	10
Output Resolution	8	8	8	6
Transportation Conformity and Hot-Spot Analysis	5	4	5	0
Suitability of boundary-layer parameterization	8	8	8	8
Initial Plume Dispersion	7	7	7	6
Model Maintenance	10	10	10	10
Cost	10	10	10	0

User Interface

- **AERMOD**: AERMOD has no user interface, as it runs from the command line with text input files. Third-party software is available that provides graphical interfaces.
- **ADMS-Roads**: ADMS-Roads comes with an easy-to-use user interface that allows the user to pick various options from source configuration, meteorology, receptor locations, and type of output data needed. Despite looking outdated, it does provide all the necessary features to modify critical inputs for dispersion modeling. In addition to input options, ADMS-Roads also has easy-to-use visualization tool, called "mapper", that allows the user to visualize the source configuration as well as output concentration data.

Ease of Use

- **AERMOD**: As noted above, the lack of a user interface means that users must develop all inputs and analyze all outputs without any graphical assistance from the software. Due to the Fortran source code, the model can be stringent about formatting of inputs. The model comes with a detailed user's guide, implementation guide, and technical and supporting documentation, which is sufficient in most cases, but some instructions are not as clear as they could be. During runtime, AERMOD's warning messages can be vague and there is not a glossary for all the keywords used in those messages. AERMOD comes with helpful data pre-processors for terrain, meteorology, and surface characteristics, though each has similar issues to those noted here about the parent model. For complex site-specific model scenarios, users almost definitely will need to make use of a graphical user interface (GUI) software. The model and its pre-processors have good flexibility about some sources of data that they can use, including the ability to process some free-form data. Volume sources, when numerous as in this project, are onerous to create and can result in very large input files, particularly when using hourly emissions, which can make it very difficult to create, quality-control, and store/share the model input files.

- **ADMS-Roads**: AMDS is generally easy to use model, although a basic knowledge of dispersion modeling is certainly needed to provide the necessary inputs and prepare the model for the case being studied. The model also comes with a detailed user's guide and technical documentation that provides step-by-step instructions on how to run the model and how to prepare the inputs files. The model also allows both direct inputs through the user interface as well as pre-processed input files that the model can read and process. This feature makes the model easy to use for both simple and sophisticated source and meteorology conditions.

Model Transparency

- **AERMOD**: As described earlier, AERMOD comes with several user, technical, and supporting documents that provide the necessary information to utilize the model for numerous scenarios, though sophisticated users will find the documentation more helpful than will new or basic users. The model's source code is freely available to study and manipulate if needed. EPA provides numerous test cases and evaluation databases as well.
- **ADMS-Roads**: As described earlier, ADMS-Roads comes with detailed technical documentation and user's guide that provides both basic and sophisticated users with the necessary information to use the model for their studies. While unlike AERMOD, the underlying codes for the model are not accessible, the model documentation is easy to follow. Considering that AERMOD makes the underlying codes accessible, we gave ADMS-Roads a lower score for this category.

Runtime

- **AERMOD**: All AREAPOLY and RLINE runs were short, at less than 1 minute per run, and while the initial model setup time at the beginning of run execution was longer for RLINE (due to having more than 60 sources to setup, taking less than 30 seconds), the dispersion calculations for RLINE likely were slightly faster. The volume-source runs took significantly longer, at 3 minutes per tracer run (most of which was due to having to set up more than 6,500 sources and read in almost 1 million lines of hourly emissions data) and closer to 90 minutes for the regulatory runs (the $PM_{2.5}$ regulatory run had nearly 9,500 sources and over 7 GB of hourly emissions data). The volume source regulatory run would have taken hours had we used five years of meteorological data, and had a larger number of volume sources been used (e.g., smaller source spacing, a larger project domain, or additional vehicle sources or chemicals being modeled) it would have been increasingly difficult to manage the creation of that large number of sources and in particular the hourly emissions file.
- **ADMS-Roads**: ADMS-Roads has a short runtime. For an annual run of 8,760 hours of meteorological inputs the model took 30 minutes and 46 seconds. That can be translated to about 0.220 second per each meteorology input (i.e., hourly meteorology input). For the pollutant and tracer evaluation the model only took 4 – 5 seconds to finish the runs. The model runtime does not appear to be sensitive to the number of receptors.

Output Resolution

- **AERMOD**: With the right combination of model settings, AERMOD provides many options for time resolution (hourly or longer; including some U.S.-based design values for air quality standards), peak and ranked results, and source stratifications (by every source or by custom groups of sources). All of these are very helpful for regulatory evaluations and conformity determinations.
- **ADMS-Roads**: ADMS-Roads provides detailed level concentration data at second, minute, hourly levels, however, it does not provide 8-hour, daily, or annual averages (at least for the ADMS-Roads that we used in this study). Considering that the model is originally designed for use in the UK, the model does not have the ability to calculate U.S.-based design values for air quality standards. While this may not be necessarily a weakness for U.K. users, we are considering it as a downside for U.S. based users considering the need to post-process the hourly concentration data to calculate design values.

Transportation Conformity and Hot-Spot Analysis

- **AERMOD**: AERMOD has been tested many times for U.S.-based transportation projects, and EPA and other U.S.-based agencies have guidance for such usage of the model. This guidance helps, though as noted earlier the model itself has no built-in features to help in things like determining source geometries for roadways (i.e., there is no map-based software or default values to select), pulling in monitored air quality data and traffic data from known sources, etc. Therefore, substantial "offline" research is required, plus processing of the raw data found, to create the inputs AERMOD needs. Volume sources in particular can be onerous to set up due to needing to parameterize roadways as series of adjacent "circular" sources and the avoidance of receptor exclusion zones where larger volume sources can render receptors very close to the roadway unable to be modeled in AERMOD.
- **ADMS-Roads**: As described earlier, ADMS-Roads is originally designed for use in U.K., and therefore it does not provide post-processed air pollution concentration data for use in hot-spot and project level transportation conformity analysis. One of the benefits of ADMS-Roads is the built-in feature in the model to estimate the roadway emissions. For cases where uses would like to use the model calculated emission factors, the model uses the U.K Emissions Factors Toolkit (EFT) databases along with user inputs on traffic flow (i.e., vehicle counts by hour), traffic speed, and road grade to determine the emission rates. However, it is users' responsibility to define the source geometry, and road configurations for the purpose of dispersion modeling. Also the emission factor feature in ADMS-Roads is certainly not applicable to U.S., due to differences in U.K. vehicle mix, and emission standards versus those in the U.S.

Suitability of boundary-layer parameterization

- **AERMOD**: As detailed in Section 2.2.6, AERMOD incorporates numerous surface and atmospheric parameters both from on-site sources and representative off-site sources. Those parameters include hourly values of heat fluxes, surface roughness, mixing heights, stability representations, cloud cover, temperature, and wind flow, including sub-hourly wind variances (i.e., turbulence indicators), with some capability of multiple vertical measurements, all of which are processed in AERMOD's AERMET meteorological processor and AERSURFACE land-cover processor. While some of these parameters represent conditions on sub-hourly scales, AERMOD does not have the ability to model at sub-hourly timesteps. And, while multiple meteorological datasets can be combined to represent project conditions, AERMOD cannot use different meteorology data for different modeling areas (without dividing those areas up into different model runs).

AERMOD is a steady-state hourly model, so it does not retain any dispersed chemical from one timestep to the next. A series of AERMOD sensitivity tests was conducted by (Long et al., 2008)[49] who varied a number of meteorological and land-use parameters over the possible variation found over the San Francisco Bay Area. He found that changes in maximum model concentrations of up to 50% occurred with surface roughness as derived from land-cover and that cloud cover, solar radiation, and urban population showed smaller sensitives of up to 20% over the possible ranges.

- **ADMS-Roads**: As described in Section 2.2.6.4, the ADMS-Roads model takes in detailed meteorological data similar to AERMOD for modeling the dispersion. The user has the ability to provide surface roughness, wind speed, wind direction, cloud cover, boundary layer height, surface temperature, meandering, relative humidity, and surface heat flux for the model to characterize the boundary layer. According to our assessment, we believe ADMS-Roads' boundary layer parametrization is as sophisticated as AERMOD and other available dispersion models.

Initial Plume Dispersion

- **AERMOD**: AERMOD parameterizes traffic-induced vertical turbulence through input of the initial vertical dimension. It is up to the user to develop this parameter, though EPA provides guidance where, for transportation projects, it is based on vehicle heights. These values can be varied by hour if desired, to reflect hourly variations in traffic fleet mix. Similarly, if a volume source characterization is used then the traffic-induced horizontal turbulence must also be input specified in the initial lateral dimension. EPA provides guidance on the development of this term in Appendix J of EPA's *PM Hot-spot Guidance: Transportation Conformity Guidance for Quantitative Hot-spot Analyses in $PM_{2.5}$ and PM_{10} Nonattainment and Maintenance Areas*, Office of Transportation and Air Quality, EPA-420-B-21-037, October 2021.

- **ADMS-Roads**: ADMS-Roads has the ability to account for the traffic-induced turbulence by introducing an extra lateral spread ($\sigma_{y_{road}}$) when modeling dispersion. For each source, the model takes in the traffic flow as well as vehicle speed for light and heavy-duty vehicles. However, the model only takes one value for traffic flow, and if the user is running the model for multiple hours, the model will assume the same traffic flow in all those hours, unless the user wants to run each hour separately which of course is a limitation of the model.

Model Maintenance

- **AERMOD**: The AERMOD system is fully supported by EPA. New versions are released as needed for bug fixes and enhancements, for example there have been 3 revisions to the parent model in the last four years. New features can be introduced in test phases (e.g., alpha and beta testing) before they may become fully supported features, with or without status as default or non-default regulatory options.

- **ADMS-Roads**: The ADMS-Roads model is maintained and regularly updated by Cambridge Environmental Research Consultants (CERC), United Kingdom. The model is updated for any bug fixes, new methodology

---

[49] Long G.E., Cordova J.F., and Tanrikulu, S. 2004 An Analysis of AERMOD sensitivity to input parameters in the San Francisco Bay Area, 13th Conference on the Applications of Ari Pollution Meteorology with the Air And Waste Management Association, American Meteorological Society, Vancouver, BC.

development, updated emission factors, etc. The last update to the model, ADMS 5.0 was released in April 2020 and is a major update of ADMS-Urban, ADMS-Roads and ADMS-Airport with many powerful new features to aid modelling, data processing and visualization.

Cost

- **AERMOD**: AERMOD is available at no additional cost to all users.

- **ADMS-Roads**: Unlike AERMOD, ADMS-Roads is not free, but can be licensed for a fee to any user. Two types of licenses are available: an annual license is valid for one year and is inclusive of support for the period covered by the license; a permanent license is valid for use indefinitely and support for the first year is covered by the license. A single user license allows use of the software by 1 user at a time; further users can be added at a reduced rate. A research license is available that allows user of the software for non-commercial purposes at research and academic organizations; also, a teaching license is available that allows use of the software for teaching purposes at academic institutions. The licensing of the software provides revenue for CERC to maintain and improve the software.

6. Representative Background Concentration Determination Needed for NAAQS Modeling

The AERMOD and ADMS models can incorporate background concentrations in a variety of ways to assess comparison with NAAQS. Several approaches are available for determining background concentrations depending upon pollutant. The most conservative choice is to use the highest concentration monitored based on the form of the standard from the most recent past 3-years of representative background concentration. We discuss below the current EPA recommended procedures for developing background concentrations from representative background monitors.

6.1. CO Background Determination

EPA has not published guidance on recommendations on how to determine background concentrations for CO other than to recommend that the background concentration should be from a monitoring site not affected by the analysis location. However, EPA has published guidance for determining short-term averaging periods to determine background concentrations with the emphasis though on the 1-hour NO_2 standard[50].

The key principle in the guidance is to determine background concentrations associated with "meteorological conditions accompanying the concentration of concern". Based on this principle the appropriate methodology for incorporating background concentrations in the impact assessment for the 1-hour CO standard would be to use the most recent three years of the high second high (H2H) of the available background concentrations by season and hour-of-day. However, for situations where mobile sources are the principal source, inclusion of a day-of-week component to the temporal variability is appropriate.

The rank associated with the H2H of daily maximum 1-hour values should be generally consistent with the number of "samples" within that time period for each combination based on the temporal resolution but also account for the number of samples "ignored" in specifying the H2H value. EPA recommends that the background values by season and hour-of-day used in determining NAAQS compliance be based on the high first high for each season and hour-of-day combination, whereas the H2H value should be used if values vary by hour-of-day only. A similar approach is used for the 8-hour CO standard although no hour-of-day would be used as only non-overlapping 8-hour periods should be used in determining background values.

6.2. $PM_{2.5}$ Background Determination

On May 20, 2014, EPA released a memo entitled "*Guidance on $PM_{2.5}$ Permit Modeling*". Then on April 17, 2018, EPA released an updated memo entitled "*Guidance on Significant Impact Levels for Ozone and Fine Particulates in the Prevention Significant Deterioration Permitting Program* ". The 2018 memo superseded the 2014 memo for applying the recommended $PM_{2.5}$ significant impact level values to determine compliance with the applicable $PM_{2.5}$ NAAQS and PSD Increments. However, the 2014 memo is still used to calculate the two-tiered approach for determining the 24-hour and annual $PM_{2.5}$ background values, and to combine modeled and monitored $PM_{2.5}$ background values from a representative monitor.

---

[50] EPA, 2011 Memorandum, "Additional Clarification Regarding Application of Appendix W Modeling Guidance for the l-hour NO2, National Ambient Air Quality Standard", Tyler Fox, Leader Air Quality Modeling Group, Research Triangle Park, NC. March.

6.2.1. First Tier Approach

The first-tier approach is for the modeled design value (DV) to be added to the monitored DV from a representative monitor. However, the modeled 24-hour concentration to be added to the monitored DV is based on the multi-year average of the 98th percentile of modeled 24-hour concentrations rather than the multi-year average of the highest (i.e., 100th percentile value) of modeled, annual, 24-hour average concentrations. The monitor should be representative in that it accounts for the contribution of both primary and secondary $PM_{2.5}$ associated with existing sources within the project area. The total impact for comparison to the NAAQS should be based on the sum of the modeled design value for $PM_{2.5}$ impacts (from the dispersion model) and the monitored design value that includes both primary and secondary $PM_{2.5}$.

6.2.2. Second Tier Approach

In the second-tier approach for the 24-hour analyses, it is recommended that the distribution of ambient air quality monitoring data equal to and less than the annual 98th percentile be divided into seasons (or quarters) for each of the three years that are used to develop the monitored DV. This results in a dataset for each year (for three years) which contains one season (quarter) with the 98th percentile value and three seasons (quarters) with the maximum values which are less than or equal to the 98th percentile value. The maximum concentration from each of the seasonal (or quarterly) subsets should then be averaged across these three years of air quality monitoring data. The resulting average maximums should then be included as the four seasonal background values within the AERMOD model for determination of the total concentration for comparison to the NAAQS.

EPA's 2021 PM Hotspot Guidance, Chapter 8,[51] also provides discussion on determining background concentrations from adjacent emission sources.

6.3. Selection of Representative Background Monitor

When trying to identify a representative background monitor the following criteria should be considered:

- Proximity of the background monitor from the project study area
- Similarity of characteristics (e.g., density of emission sources, land-use, terrain) between the monitor location and the project study area
- Wind-flow pattern between the background monitor and project study area with preference of the background monitor upwind of the prevailing upwind direction for the roadway location
- Purpose and geographic scale of monitoring station
- Collection frequency of pollutant of concern
- Data completeness[52]

Further discussion and guidance on the use of ambient monitoring data to estimate background concentration can be found in Section 8.3.1 of EPA's "PM Hot-spot Guidance: Transportation Conformity Guidance for Quantitative Hot-spot Analyses in $PM_{2.5}$ and PM10 Nonattainment and Maintenance Areas", Office of Transportation and Air Quality, EPA-420-B-21-037, October 2021.

---

[51] Transportation Conformity Guidance for Quantitative Hot-spot Analyses in PM2.5 and PM10 Nonattainment and Maintenance Areas, EPA-420-B-21-037, October 2021. Available at https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P1013C6A.pdf.

[52] Data completeness is defined by EPA is 75% of scheduled samples in each calendar quarter of each year.

In this study we collected ambient air quality data from the BAAQMD's Aquatic Park, Berkeley near-road monitor. The BAAQMD's 2019 Network Plan[53] report was reviewed for sites that could potentially be considered representative background monitors. Air quality monitoring stations in proximity (< 10 miles) and located on the Bay side of the Berkeley Hills and San Pablo Ridge are:

- San Pablo
- Richmond
- Point Richmond
- Oakland West
- Laney College
- Oakland East

The Point Richmond and Richmond monitor were eliminated from consideration as they only measure hydrogen sulfide, sulfur dioxide and air toxics. Laney College is a near road monitor, not a background monitor. Oakland East is influenced by large emissions sources in the areas, and high volume of traffic in the city center and several nearby major freeways. Oakland West is located one mile of the prevailing downwind of the Port of Oakland which is a major source of particulate matter and heavy truck and vehicle traffic.

San Pablo site is considered a middle scale for CO monitoring based on traffic counts and the distance between the nearest traffic lane and the air monitor. The closest roadway is Rumrill Road, which has an annual average daily traffic (AADT) of just 15,433 so is mostly measuring regional CO background level – this in comparison with Aquatic Park sites 280,400 AADT. The site is on the west side of I-80 so has frequent exposure to the relatively clean air from San Pablo Bay. For $PM_{2.5}$ it's a bit more problematic being downwind of the Chevron refinery, but still classified as middle scale for $PM_{2.5}$. However, the BAAQMD identified this monitor to be representative of areawide $PM_{2.5}$ concentrations and so reasonable to use as the background monitoring station. This made the San Pablo site the clear choice for the background monitor for both CO and $PM_{2.5}$. Also, both the San Pablo and Aquatic Park use same instrument types to measure CO and $PM_{2.5}$, this is ideal particularly for $PM_{2.5}$ measurement.

6.4. Use of Chemical Transport Models for Background Determination

In some cases, it may be appropriate to use background concentrations that have been determined based on modeled outputs from a chemical transport model (CTM). CTMs are most often gridded photochemical models that are used in both regulatory and environmental impact assessments, as well as in attainment demonstrations for non-attainment and maintenance areas for State Implementation Plans and by EPA in regulatory assessments to support national or regional rulemakings.

Use of CTM output is particularly valuable where no representative ambient monitoring data is available. In these types of analyses, CTM modeling is completed for a base year and then the resulting concentrations are used in to estimate background concentration levels. This method applies in areas that have appropriate photochemical modeling outputs available. Depending on the nature of the modeling, it may be possible to obtain CTM outputs that can be used to derive background concentrations. This may be an option if the standard post-processed data includes only a subset of monitoring sites in the domain or a subset of averaging times (e.g., annual average results are available, but not 24-hour average results). Further discussion on the use of CTM can be found in Section 8.3.2 *of EPA's PM Hot-spot Guidance: Transportation Conformity*

---

[53] 2019 AIR MONITORING NETWORK PLAN, July 1, 2019. Meteorology and Measurement Division. Available at: https://www.baaqmd.gov/~/media/files/technical-services/2019_network_plan-pdf.pdf?la=en.

Guidance for Quantitative Hot-spot Analyses in $PM_{2.5}$ and PM_{10} Nonattainment and Maintenance Areas, Office of Transportation and Air Quality, EPA-420-B-21-037, October 2021.

For this assessment we explored the use of CTM approach using the CTM study done to support California Assembly Bill No. 617 "Community Air Protection Program") performed by the BAAQMD54. In this study BAAQMD applied the EPA's Community Multi-Scale Air Quality (CMAQ) to simulate pollutant concentrations at a 1-km horizontal resolution that encompassed the entire San Francisco Bay Area (including over water) for 2016. While the study focused on $PM_{2.5}$ it also included many other pollutants as well as CO. However, in discussion with the BAAQMD it was found that the hourly concentration data needed from the model output had not been saved due to the large data storage requirements for retaining hourly concentrations, for each grid cell and layer for each pollutant over the one-year period. Although we considered rerunning the CTM and saving only the outputs needed for this assessment, the computational power needed to do so was beyond available resources.

---

[54] Fine Particulate Matter Data Analysis and Regional Modeling in the San Francisco Bay Area to Support Assembly Bill 617, Publication No. 201901-017PM, January 2019.