GEORGIA DOT RESEARCH PROJECT 20-24

FINAL REPORT

AERMOD, RLINE, AND RLINEXT CASE STUDY ANALYSES IN ATLANTA, GEORGIA

OFFICE OF PERFORMANCE-BASED MANAGEMENT AND RESEARCH

600 WEST PEACHTREE ST. NW
ATLANTA, GA 30308

1. Report No.: FHWA-GA-21-2024	2. Government Accession No.: N/A	3. Recipient's Catalog No.: N/A
4. Title and Subtitle: AERMOD, RLINE, and RLINEXT Case Study Analyses in Atlanta, Georgia		5. Report Date: November 5, 2021 6. Performing Organization Code: N/A
7. Author(s): Randall Guensler, Ph.D., Hongyu Lu, William Reichard, Ziyi Dai, Tian Xia, Angshuman Guin, Ph.D., and Michael O. Rodgers, Ph.D.		8. Performing Organization Report No.: 20-24
9. Performing Organization Name and Address: Georgia Institute of Technology/CEE 790 Atlantic Drive Atlanta, GA 30332 Phone: 404.894.0405 Email: randall.guensler@ce.gatech.edu		10. Work Unit No.: N/A 11. Contract or Grant No.: P.I. NO.: 0017725
12. Sponsoring Agency Name and Address: Georgia Department of Transportation Office of Performance-based Management and Research 600 West Peachtree St. NW Atlanta, GA 30308		13. Type of Report and Period Covered: Final; Nov. 2020 – Nov. 2021 14. Sponsoring Agency Code: N/A
15. Supplementary Notes: Prepared in cooperation with the U.S. Department of Transportation, Federal Highway Administration.		

16. Abstract:

This research assessed the impact of USEPA's AERMOD dispersion model (version of 19191) source types on predicted pollutant concentrations via a case study for the I-75/I-575 Northwest Corridor (NWC) in Atlanta, GA. Using MOVES-Matrix for MOVES 2014b, carbon monoxide (CO) emissions rates for every hour of a one-year study period were generated using traffic volumes and speed from Atlanta Regional Commission's Activity-Based Model (ABM) 2020 and AERMET meteorological profiles provided by the state environmental agency (EPD). To develop concentration profiles to assess prediction differences across source types, consistent input datasets for hourly emissions by ABM link, hourly AERMET data, and gridded receptor placement (standard 20-meter grids and variable grids after link-screening) were run with the AERMOD source types: AREAPOLY (manually created and automatically generated), LINE, VOLUME, RLINE, and RLINEXT (with and without noise barriers). The team processed more than one trillion source-receptor pairs the study area using the cyberinfrastructure resources provided by the Partnership for an Advanced Computing Environment (PACE) at Georgia Tech. The results indicate that predictions from AREAPOLY and LINE are essentially identical. Predictions from RLINE and RLINEXT are almost the same, but these predictions are higher in most cases than any other source type. The VOLUME source type always yields the lowest concentrations and is less sensitive to wind directions and speed, due to the embedded wind meander dispersion parameters implemented only for VOLUME sources. Machine learning results indicate that wind speed, receptor ID (which accounts for adjacent roads and their and their mass flux emission rates in grams/meter^2/second), and wind direction influence the results much more than source type selection. Introducing noise barriers to RLINEXT lowered concentration as expected, but modeling barrier effects was challenging due to the restrictive assumptions in AERMOD. Sensitivity analysis for RLINEXT suggests that barrier height, distance to the roadway, wind speed, and wind direction all affect morel predictions. Modelers need to exercise care in matching barriers to roadway link segments (i.e., barrier edge effects were observed).

17. Keywords: AERMOD, Dispersion Modeling, Source Type, RLINE, RLINEXT	18. Distribution Statement: No Restriction

19. Security Classification (of this report): Unclassified	20. Security Classification (of this page): Unclassified	21. No. of Pages: 139	22. Price: Free

Form DOT 1700.7 (8-69)

GDOT Research Project No. 20-24 / T.O. 2014-109

Final Report

AERMOD, RLINE, AND RLINEXT CASE STUDY ANALYSES IN ATLANTA, GEORGIA

By

Randall Guensler, Ph.D.
Hongyu Lu, William Reichard, Ziyi Dai, Tian Xia,
Angshuman Guin, Ph.D., and Michael O. Rodgers, Ph.D.

School of Civil and Environmental Engineering

Georgia Institute of Technology

Georgia Tech Research Corporation Contract with

Georgia Department of Transportation

In cooperation with

U.S. Department of Transportation Federal Highway Administration

November 5, 2021

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	$candela/m^2$	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	$candela/m^2$	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

The objective of this study was to perform AERMOD microscale dispersion modeling for an entire Atlanta I-75/I-575 Northwest Corridor (NWC) subarea, including freeway corridors, managed lanes, connecting arterials, and intersections serving the NWC system and compare the model results across all AERMOD source types (AREAPOLY, LINE, VOLUME, RLINE, and RLINEXT). The project was designed to assess the relative impacts of various AERMOD input parameters by processing model runs on the PACE supercomputing cluster and using machine learning to assess the relative impacts of model input parameters. Then, the team would undertake specific sensitivity analyses to assess the impacts for a controlled roadway section. A 20-meter standard grid of receptors in combination with a variable grid provided sufficient detail while maintaining reasonable model performance. For all AERMOD source types, the team found that wind speed, receptor ID (which accounts for adjacent roads and their mass flux emission rates in grams/meter^2/second), and wind direction influence the results much more than source type selection. For AREAPOLY sources, polygons generated automatically produce concentration results sufficiently similar to manually generated polygons, demonstrating the feasibility of implementing automatic link generation with no negative impact on AERMOD outputs.

Comparative results across AERMOD source types indicate only trivial prediction differences between AREAPOLY and LINE source types and minor differences between RLINE and RLINEXT source types. The RLINE and RLINEXT source types do predict higher concentrations than the AREAPOLY and LINE source types in most cases, because the RLINE and RLINEXT source types employ different dispersion algorithms within AERMOD (R-LINE dispersion algorithms). Hence, AERMOD with RLINE and RLINEXT source types constitutes a different dispersion model than AERMOD with AREAPOLY and LINE source types. However, using

AERMOD with the VOLUME source type also triggers a third set of dispersion algorithms (additional routines to incorporate wind meander) that are not reflected in predictions for the AREAPOLY and LINE source types. The VOLUME source type predicts lower downwind pollutant concentrations and higher upwind pollutant concentrations than are predicted by AREAPOLY or LINE sources, whenever wind speeds are less than approximately 2.5 meters/second. This holds true for any transportation scenario. The AERMOD uniform and traditional plume algorithms that are invoked by selecting the VOLUME source type is the equivalent of assuming a very large standard deviation in wind direction at low speeds. Hence, AERMOD effectively contains three dispersion models: one for AREAPOLY and LINE source types, a second for RLINE and RLINEXT source types, and a third for VOLUME source types.

The research team has not assessed the validity of the three different dispersion modeling approaches currently employed within AERMOD (which was beyond the scope of this project). However, given the lower predicted concentrations associated with the use of the AERMOD VOLUME source type, the research team recommends a thorough review of the associated dispersion algorithms. It may be that the dispersion algorithms triggered by the VOLUME source type are reasonable and can be supported by the literature, but the research team did not find relevant model shoot-out results to confirm the basis for using the enhanced wind meander approach employed with VOLUME source type. Once an advisory panel reviews the AERMOD dispersion algorithms, and scientific consensus is reached as to the validity of the algorithms under specific conditions, the research team recommends that AERMOD be updated to either: 1) employ one set of dispersion algorithms across all source types, or 2) publish guidance that specifies which algorithms should be applied under specific circumstances. Supplemental guidance will help ensure that model applications yield consistent results for any given transportation scenario. Until

scientific consensus is reached, the research team recommends continued use of AERMOD with AREAPOLY and LINE source types in regulatory analysis, or RLINE and RLINEXT source types, given that the results are more conservative at low wind speeds and less likely to miss any potential exceedance of the NAAQS.

Modeling barrier effects was quite challenging, due to the restrictive assumptions in AERMOD. As expected, noise barriers that are a greater distance from the roadway lead to smaller impacts on pollutant concentrations at downwind receptors. Likewise, noise barriers with lower heights lead to smaller impacts on downwind receptors. Barriers that are only partially aligned with a roadway can cause significant impacts on downwind receptors. Properly splitting links to align with noise barriers is important; however, the act of splitting the links, even while maintaining the exact same geometry, does affect downwind predictions. In many cases, the height, offset, and alignment of a noise barrier segment varied significantly along the length of the barrier; yet, AERMOD requires users to model barriers with a uniform height, offset, and alignment. The research team recommends that additional guidance on integration of noise barriers (coding for model runs) be developed and published by the agencies that require model use in regulatory applications. The research team also recommends that AERMOD barrier modeling be made more robust, such that variable heights, offsets, and alignments may be accounted for.

CHAPTER 1. INTRODUCTION

The Clean Air Act (CAA) of 1970 provides the legal basis for air pollution control in the United States and tasks the US Environmental Protection Agency (USEPA) with developing and implementing National Ambient Air Quality Standards (NAAQS) for specific air pollutants. Section 176(c) of the CAA (42 U.S.C. 7506(c)) and subsequent transportation conformity regulations (40 CFR Parts 51 and 93) stipulate that federal funding can only be provided to highway and transit projects that conform with transportation and air quality plans and that do not cause or contribute to a violation of the NAAQS. When performing an air quality impact assessment for transportation plans and conformity, and when performing comparative alternatives assessment under the National Environmental Policy Act (NEPA), it is important to use accurate pollutant dispersion models. AERMOD is the microscale dispersion modeling software preferred and recommended by the USEPA. AERMOD allows users to model the dispersion of pollutants through a study area, accounting for transportation network geometry, vehicle emissions, meteorological factors, and the presence of influential obstructions such as highway noise barriers. AERMOD allows any transportation facility and scenario to be formatted as a set of polygon, line, or volume sources and modeled as one five different input source types: AREAPOLY, LINE, VOLUME, RLINE, and RLINEXT.

The objective of this study was to perform microscale dispersion modeling for an entire Atlanta I-75/I-575 Northwest Corridor (NWC) subarea and to compare AERMOD outputs for a high-density receptor grid across all five AERMOD source types. The NWC subarea encompasses a variety of transportation projects of potential policy concern, including major intersections, managed lanes, and direct access ramps. These specific transportation facilities

have already been the subject of extensive dispersion modeling activity by the Georgia Institute of Technology. The subarea is modeled by coding the entire transportation link and node network into each of the AERMOD source types, carrying relevant transportation link and noise barrier attributes into the model. Traffic volumes and speed profiles were extracted from output of ARC's ABM2020-TIPA1-2020 (same model used to develop the 2020 TIP). Applicable MOVES2014b emission rates for each link were queried from MOVES-Matrix. Billions of source-receptor pair analyses were performed for this project by running AERMOD on the Partnership for an Advanced Computing Environment (PACE) supercomputing cluster. The large-scale, pairwise, model vs. model analysis reveal the impact of different source type and transportation input parameters on AERMOD outputs.

CHAPTER 2. MODELING OVERVIEW

Chapter two of this report details relevant background information on the project approach, including descriptions of AERMOD link creation, generation of AERMOD emission rates, AERMOD receptor placement, the Northwest Corridor study area, the PACE supercomputing cluster, and scenario development for AERMOD runs.

AERMOD Link Creation

The AERMOD dispersion model can use a variety of source types as model inputs (LINE, AREA, AREAPOLY, VOLUME, RLINE, and RLINEXT). Figure 1 and Figure 2 illustrate some of these source types. The LINE source type allows users to specify the roadway start position, end position, and width. The AREA source type allows users to specify a rectangular source with an arbitrary rotation. The AREAPOLY source allows users to define an irregularly shaped polygon with up to 20 vertices by entering the coordinates of each vertex. The VOLUME source type allows users to define a height and dimension of spherical sources. The RLINE source type integrates R-Line dispersion parameters into the LINE source type. RLINEXT adds additional functionality to the RLINE source type, such as the ability to model the impact of noise barriers and depressed roadway sections.

AERMOD sources are generally split into homogeneous sections, such that the characteristics of each source link do not vary within each source. This requires splitting roadways into multiple individual sources, where each individual source has a uniform emission rate. In this study, the I-75 and I-575 Northwest Corridor were split into several hundred individual AERMOD sources to account for differences in traffic volumes on

different roadway segments, traffic volume differences across lanes, varying roadway geometry, and the presence or absence of noise barriers and depressed sections of roadway.

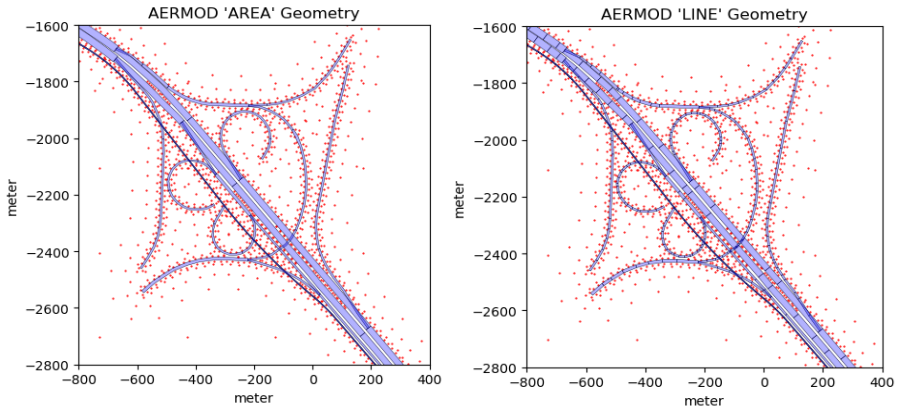


Figure 1 – Geometry of AERMOD Source Types: AREAPOLY and LINE

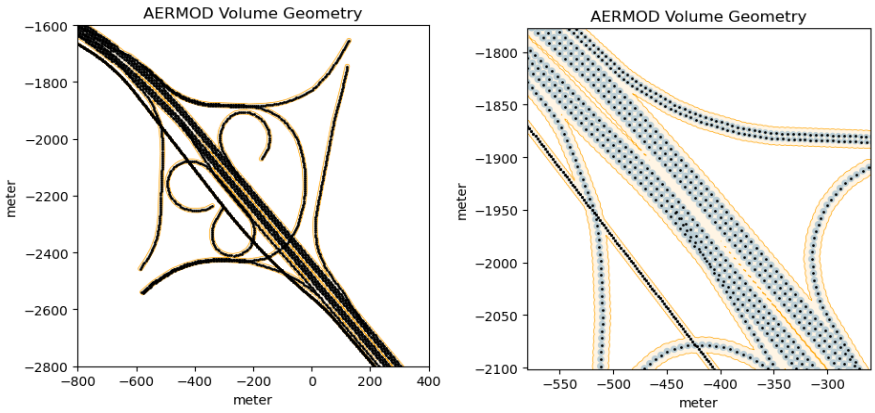


Figure 2 – Geometry of AERMOD Source Types: VOLUME

Generation and Assignment of AERMOD Emission Rates

Once links are generated using one of the AERMOD source types, each link is assigned hourly traffic volume and average speed inputs from the outputs of ARC's ABM2020 travel demand model. Link-specific emission rates for the County fleet composition, vehicle speed, and environmental conditions are pulled from MOVES-Matrix for MOVES 2014b, which generates exactly the same results as running the USEPA MOVES 2014b model for the analysis (Kim, et al., 2020; Liu, et al., 2019; Liu, et al., 2017). The pre-processed lookup

matrix allows emission rates to be queried 200 times faster than performing individual MOVES model runs and without the need to develop a MOVES input file. The fleet composition (source type distribution) for each link was derived from each County's (Cobb County, Cherokee County, and Bartow County) on-road vehicle mix used in the ARC's regional conformity planning (ARC, 2015) and in previous regional research by the team (Xu, et al. 2018). The hourly AERMET meteorological profiles were provided by the GA EPD (24 hours × 365 days of the year for 2019); used to configure the meteorology input to MOVES-Matrix for MOVES 2014b to obtain applicable emission rates. The resultant emissions rates from MOVES-Matrix were assigned to each ABM link, which were converted to the input emissions rates of AERMOD sources. A detailed description of the methodology and data with respect to generation and assignment of these emissions rates are provided later in the report.

Receptor Placement

AERMOD allows users to specify receptor locations. The receptors define the physical locations in x, y, z space for which pollutant concentrations will be predicted for every hour in the simulated year. Receptors allow users to assess pollutant concentration levels relative to nearby locations of concern (e.g., near schools or residential areas where individuals are likely to be exposed to pollutant concentrations for extended periods) and to identify localized areas of high concentration. Assessment of receptor concentrations allows modelers to identify regions that may exceed NAAQS. The computing resources available for this project allowed the research team to assess as many receptors as desired, so a variety of receptor patterns were used in this study, including standard receptor grids and variable receptor grids.

Standard Grids

Standard grids with 200-meter spacing between receptors (Figure 3), 20-meter spacing between receptors, and 5-meter spacing between receptors were used in this study. Receptor grids provide a simple approach that requires minimal forethought, but is computationally inefficient because many of the receptors are placed so far from the roadway that pollutant concentrations are not significant and contribute little information. Likewise, increased receptor density is desired near the roadway, where variation in pollutant concentration is likely to be the largest. It is difficult to strike a balance between sufficient receptor density near roadways and computational efficiency while using a standard grid. Hence, a variable grid approach is also useful as will be discussed in the next section.

Given the large size of the study area, receptors far from roadway emission sources yield very low receptor concentrations (e.g., less than 0.1 μg/m3 across all input emissions rates and meteorology by hour). Removing such receptors from the analyses does not affect research outcomes. A link-screening tool is implemented to examine all link-receptor pairs and filter "non-sensitive" receptors to reduce required computational resources.

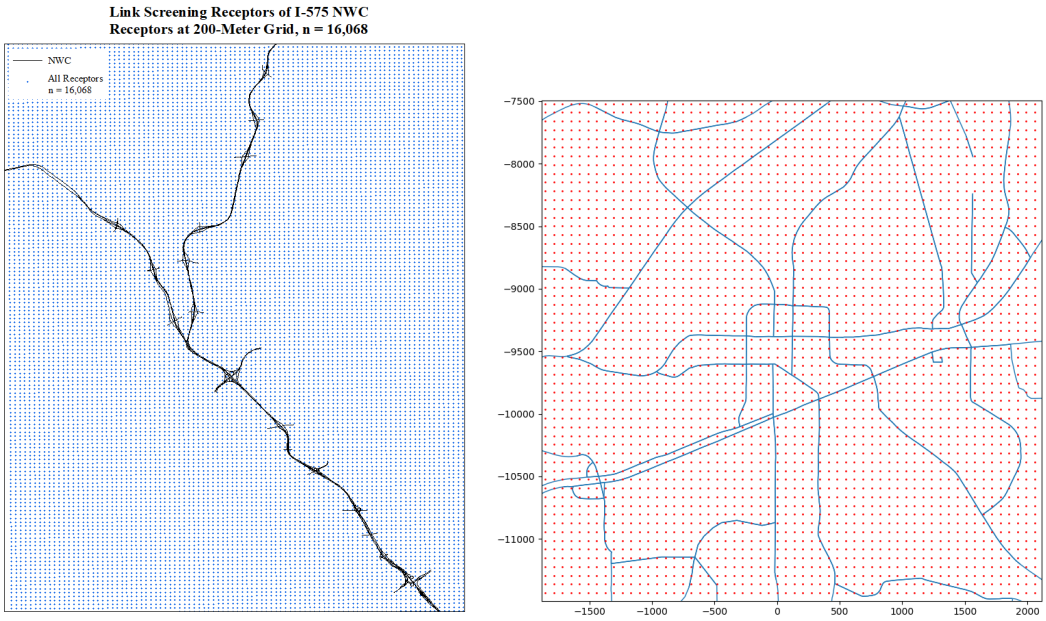


Figure 3 – Placement of Receptors of 200-Meter Standard Grid

Variable Grids

A more advanced approach for placing the receptors is the dynamic grid method (Kim 2020). A link screening process is applied before setting the variable grid, based upon source-receptor centration statistical significance assessed via machine learning. Specifically, a stepwise process is conducted to identify optimal receptor locations, with a forward search used to find the receptor that best fits the $PM_{2.5}$ concentration profile, and a backward search used to remove the receptor that worst fits the $PM_{2.5}$ concentration profile. The iteration process continues, until the marginal change in mean squared error (MSE) is less than some pre-set critical threshold. The dynamic grid-receptor model identifies optimal receptor locations, removing receptors from the grid that contribute no significant additional information to the spatial concentration distributions. The variable grids that evolve from the process tend to have higher receptor density near roadway sources (where small differences

in distance between receptors yield high differences in concentrations) and low receptor density further from roadways sources (where additional receptors yield very similar low concentrations). The more refined receptor grid minimizes the number of receptors used in modeling and speeds up the distributed modeling process.

Northwest Corridor Subarea Description

The research team has performed AERMOD microscale dispersion modeling for an entire Atlanta I-75/I-575 Northwest Corridor (NWC) subarea, including freeway corridors, managed lanes, connecting arterials, and intersections serving the NWC system (Figure 4). This corridor and subarea have been the subject of extensive emissions and dispersion modeling efforts by the Georgia Tech research team (Kim, et al. 2020) and encompasses a variety of projects of potential policy concern, including major intersections, managed lanes, and direct access ramps. The NWC also contains roadway sections with noise barriers, which can be modeled with the AERMOD RLINEXT source type.

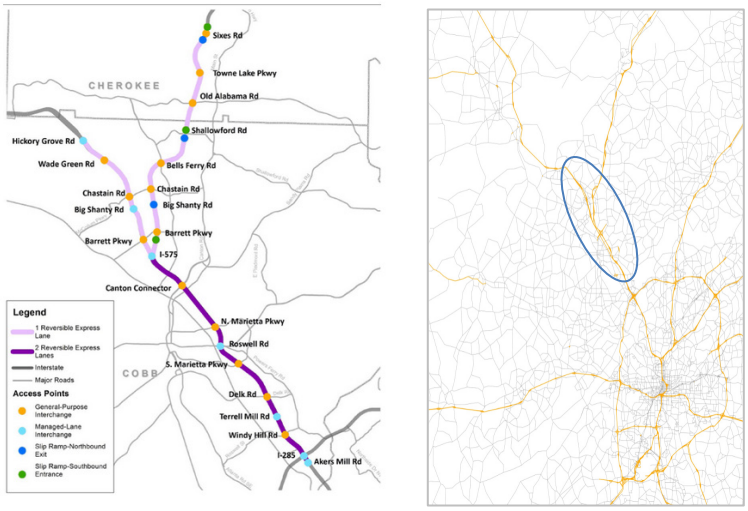


Figure 4 – I-75/I-575 Managed Lanes Corridor and the Modeled Sub Area

Dispersion Modeling on the PACE Supercomputing Cluster

The traditional AERMOD dispersion modeling in this research was supported by the cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing Environment (PACE) at Georgia Tech. PACE is a cluster-based supercomputing platform where users upload data/programs and submit jobs through scheduling systems to access PACE CPU/GPU units and RAM.

Given the huge number of source-receptor pairs that need to be processed, the team split the NWC study area into I-575 and I-75 corridors so that they could be processed separately on PACE. The team then categorized receptors based on grid resolution (5-meter grid, 20-meter grid, 200-meter grid, and variable grid) and receptor distance to the closest roadway establishing a buffer around the roadways for separate near-road analyses (0 to 200 meters from the nearest roadway and 0-500 meters from the nearest roadway).

The research team was given prioritized access to 6,000 CPU cores on the PACE Phoenix Cluster for 500 jobs, with each job using an average of 12 cores. The test runs for the prioritized scenarios indicated that each CPU core on the Phoenix Cluster could process an average of 2.5 AERMOD runs without negatively affecting computational performance. That is, for a full utilization of all 6,000 cores, the team launched 15,000 AERMOD runs at its peak (which varied due to job queueing, wall clock restrictions, and input/output file management).

Model run scenarios were developed based on a combination of receptor grid structure and buffer distance for assigning runs on the supercomputing clusters. Table 1 describes these

scenarios and indicates how many individual source-receptor pairs included in each scenario. Scenarios with fewer source-receptor pairs were prioritized for PACE runs to optimize the use of PACE resources (i.e., low grid resolution runs were run first). Source-receptor pairs across scenarios were coded to run only once to avoid duplicate runs. That is, receptors already processed in the 200-meter grid analyses do not need to be run again in the 20-meter grid analysis because the receptor appears in both grids.

As indicated in Table 1, placing receptors in a 5-meter grid over the entire region would take more than 3,000 days to process, and would still take more than 200 days to process even after link screening was performed to remove inconsequential receptors from the analyses. Fortunately, the vast majority of 5-meter receptors (those located more than 500 meters from the roadway) provide only a trivial improvement in concentration predictions compared to using a 20-meter grid (see Appendix A). Therefore, after finishing the 5-meter receptor grid within 200 meters of the roadway, the team pivoted to the 20-meter receptor grid for the remainder of the project. A variable receptor grid is the most efficient, as unneeded receptors are eliminated from the analyses.

Table 1 – Network and Receptor Scenarios Developed for Supercomputing Runs

Scenario ID	Distance Limit (Meters)	Grid Resolution (Meters)	# of Source-Receptor Pairs	Total CPU Hours Required (Hours)	Estimated Processing Time (Days)
I575only500mMaxDistance5mGrid	500	5	115,133,583,264	9,189,879.8	29.5
I575only500mMaxDistance20mGrid	500	20	7,194,664,764	574,273.0	1.8
I575only200mMaxDistance5mGrid	200	5	50,945,748,504	4,066,453.0	13.0
I575onlyLinkScreened5mGrid	Link Screening	5	802,782,295,392	64,077,505.1	205.4
I575onlyLinkScreened20mGrid	Link Screening	20	50,193,195,759	4,006,384.8	12.8
nwc5mGrid	NWC	5	12,957,538,008,567	1,034,261,359.6	3,314.9
nwc20mGrid	NWC	20	809,406,571,896	64,606,250.1	207.1
nwc200mGrid	NWC	200	7,512,063,156	599,607.5	1.9
nwcVariableGrid	NWC	Variable	12,086,249,484	964,715.7	3.1

For each scenario in Table 1, the team ran every source-receptor pair for each of six AERMOD source types (AREAPOLY, LINE, VOLUME, RLINE, RLINEXT with barriers, and RLINEXT without barriers). Modeling the same transportation network and vehicle activity while varying only the coding of AERMOD source type allowed the team to perform pairwise model vs. model comparisons for hourly pollutant concentration results across the huge array of spatiotemporal receptors. Using enhanced regression tree and machine learning tools, the team was able to identify factors that contribute to modeled differences across receptors and source types, akin to the methods that the team applied to develop the link screening modeling approach (Kim, et al. 2020; Wolf, et al. 1998). The analytical results helped the team identify and isolate the influence of the AERMOD source types, compared to the traditional variables that affect concentration predictions (e.g., wind speed, sensible heat flux, surface roughness, link mass flux emission rate (grams/meter^2/second), source receptor distances, etc.), which remained constant for each source-receptor-hour run. The team then used the evidence-based modeling results to identify potential modeling limitations in applying the various AERMOD source types to various project types within the larger network.

CHAPTER 3. METHODOLOGY

This chapter details the methodology used to complete the project, including transportation network coding, creation of AERMOD polygons, coding of link attributes, AERMOD meteorological configuration, traffic volume model, fleet composition, MOVES-Matrix, noise barriers, and the PACE supercomputing cluster.

Transportation Network Coding

The team has developed the network files for each model run by source type, including the physical layout of each network link, accompanying link characteristics, as well as the required spatial and temporal traffic volumes, on-road operating conditions, and fleet composition. The team has already prepared AERMOD polygon input files for the entire Atlanta metropolitan area (74,000+ network links) and transportation roadway links for RLINE (Kim, et al. 2020). These structures have been adapted to work with the other AERMOD source types and relevant physical network structural data have been assigned to those sources.

Manual Creation of AREAPOLY Polygons

The AERMOD AREAPOLY source type requires users to specify up to 20 vertices that bound a source. The section of I-575 within the study area required the creation of 476 unique polygons to cover 359 unique roadway links. These polygons were created within QGIS using mouse clicks to identify polygon vertices. The analyst would click on a map of the study area to specify the location of the first polygon vertex, click to specify the location of the next vertex, and repeat this process in a clockwise or counterclockwise direction until the polygon was complete. Each polygon could cover no more than one roadway link, and some particularly long and/or curved links required multiple polygons so that no individual polygon would exceed 20 vertices.

Additionally, some Express Lane links within the NWC are reversible lanes, meaning that two unique traffic links exist in the same physical locations. Thus, one polygon could be used to account for both Express Lane links. The process of manually creating and validating polygons for the entire NWC consumed approximately 280 hours of labor.

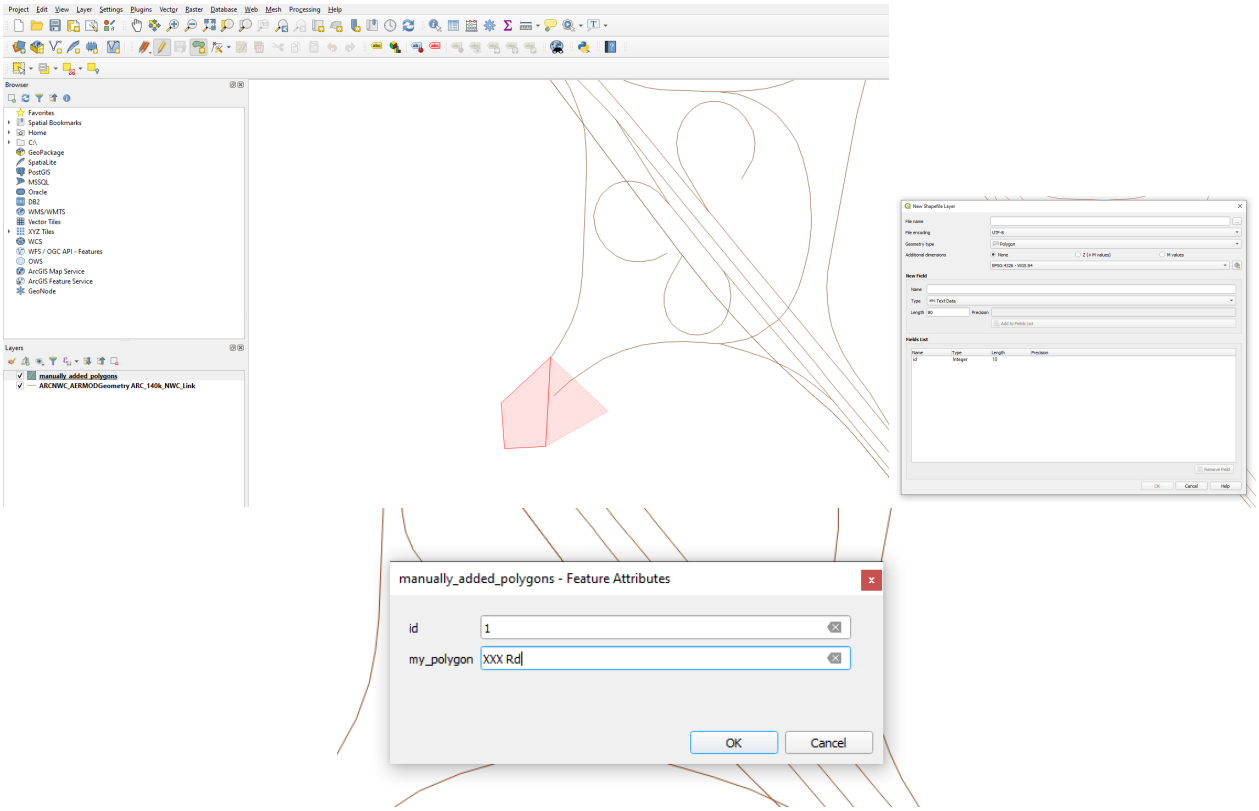


Figure 5 – Manual Polygon Creation Using QGIS

Automatic Polygon Creation

As an alternative to manual creation of AREAPOLY polygons, an automatic polygon creation method was also adopted in the study, using the Python script toolkit developed by Dr. Haobing Liu (Liu, 2021), which enables the automatic conversion from roadway GIS polyline layers to AERMOD sources, including AREA, LINE, VOLUME, RLINE, and RLINEXT defined sources. The tool uses the shape points contained in any travel demand model network file to first break

links into straight segments, and then creates polygons based upon roadway width. The team applied the tool to the ARC ABM2020-TIPA1-2020 model roadway network for this project.

Manual vs Automatic Polygon Creation

Figure 6 compares the polygons generated manually to the polygons generated automatically. There are minimal differences in polygon geometry on the highway itself, with larger differences noted for the arterial roadways that intersect the highway. However, as will be demonstrated in this report, given that the outputs from both methods are so similar, the automated method of polygon generation is preferable to manually generating polygons for the sake of efficiency and consistency.

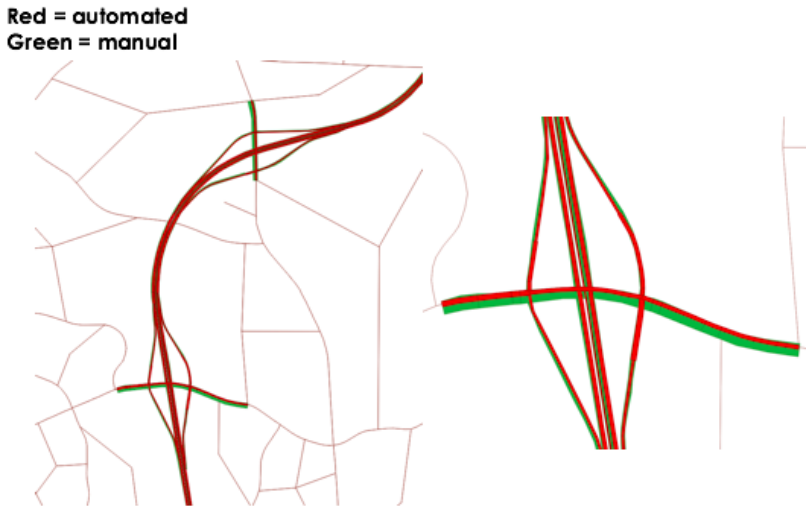


Figure 6 – Differences between Manually and Automatically Generated Polygons

Verification of Links Generated Automatically

Pollutant concentrations are generally the highest when winds run parallel to roadway links, carrying pollutants along the roadway with concentrations building over distance. Hence, proper alignment of links in the spatial modeling domain is critical. The positions of links generated

automatically were verified by visually comparing the links to ABM link geometry. Figure 7 through Figure 10 compare the alignment of the automated links and ABM links, showing that it is acceptable to use the automatically generated links.

The predicted results from AERMOD (source type of AREAPOLY) based on manually created polygons vs. auto generated ones indicated that auto generation does not lead to different concentration results (presented later in this report).

Figure 7 – AERMOD Links vs. ABM Links, LINE Geometry

Figure 8 – AERMOD Links vs. ABM Links, AREAPOLY Geometry

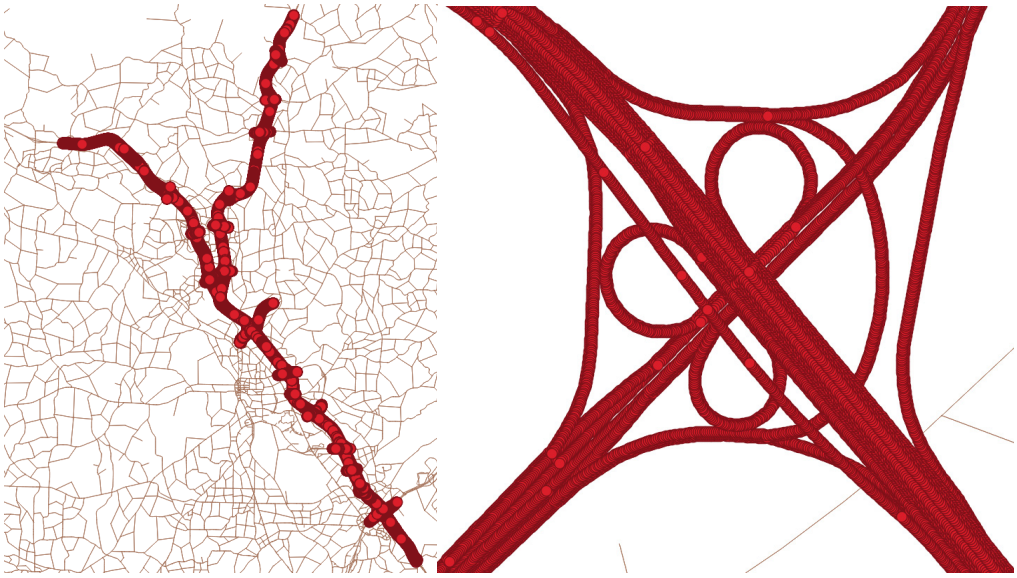


Figure 9 – AERMOD Links vs. ABM Links, VOLUME by Lane Geometry

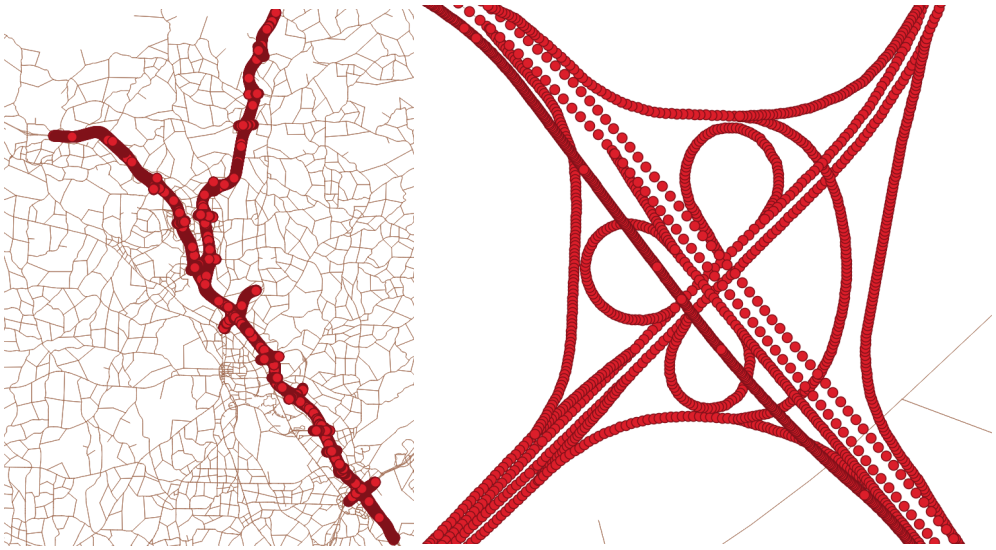


Figure 10 – AERMOD Links vs. ABM Links, VOLUME by Road Geometry

Coding of Link Attributes

Link attributes, such as the presence of noise barriers, were accounted for in coding the RLINEXT source type input. The presence of barriers was identified using Bing Maps Flythrough imagery, which allowed users to note the presence and location of noise barriers and depressed roadway sections. The flagged noise barrier locations were then verified using the GDOT Noise Barrier Inventory, which contains the location of each noise barrier within the study location.

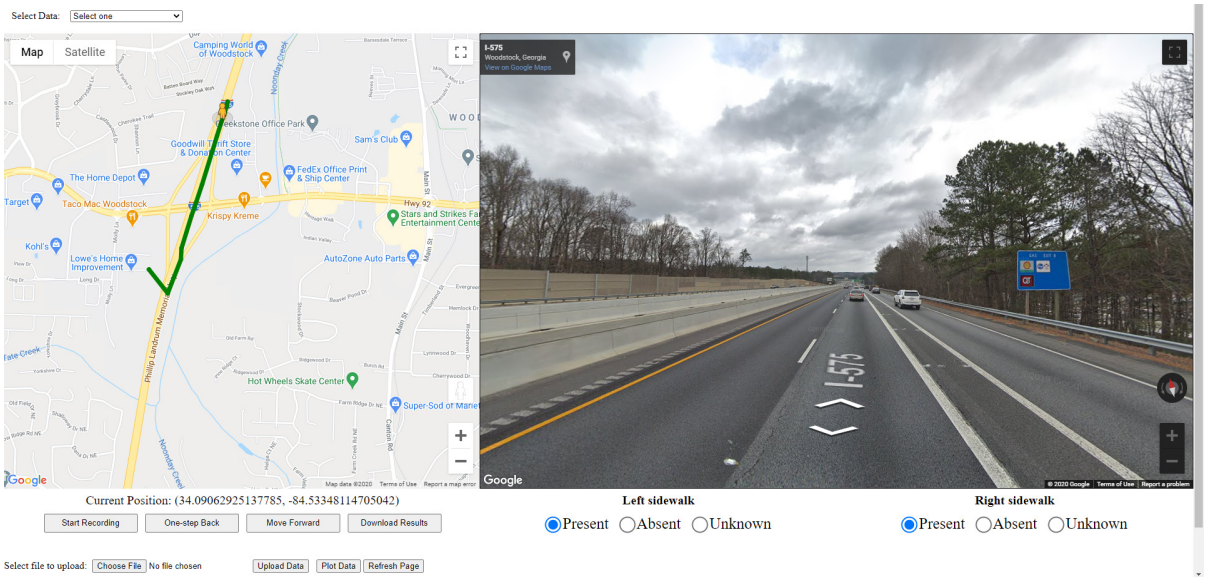


Figure 11 – Bing Maps Flythrough Imagery

AERMOD Meteorological Configuration

A total of 8,760 hours of AERMET profiles (24 hours × 365 days) were obtained from Georgia EPD station #KATL/KFFC, which covers Cobb County, Cherokee County, and Bartow County (the counties in which the study area roads are located), for each hour of 2019, to provide the meteorology input (temperature, relative humidity, wind direction, wind speed, etc.) to both emissions and dispersion modeling of the studied area of NWC. The AERMET files (PFL file and SFC file) serve as an input to AERMOD and the hour-by-hour temperature and relative humidity profiles were used to extract and assign corresponding emission rates from the MOVES-Matrix (for MOVES 2014b) database to each individual hourly model run.

The wind rose for the used AERMET profiles is shown in Figure 12 (prepared using OriginPro™ software). The temperature and humidity distributions are shown in Figure 13 and Figure 14, respectively.

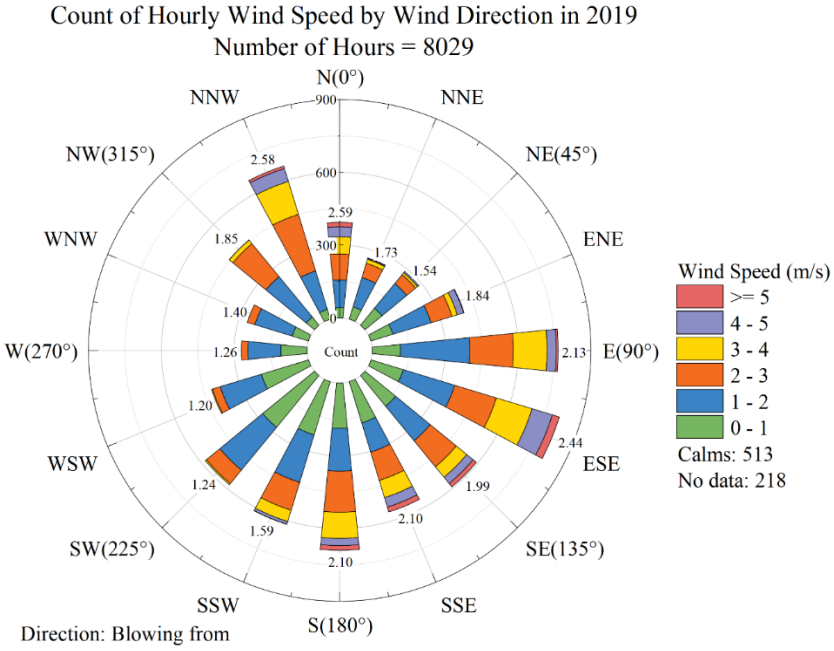


Figure 12 – Wind Direction and Wind Speed Distributions

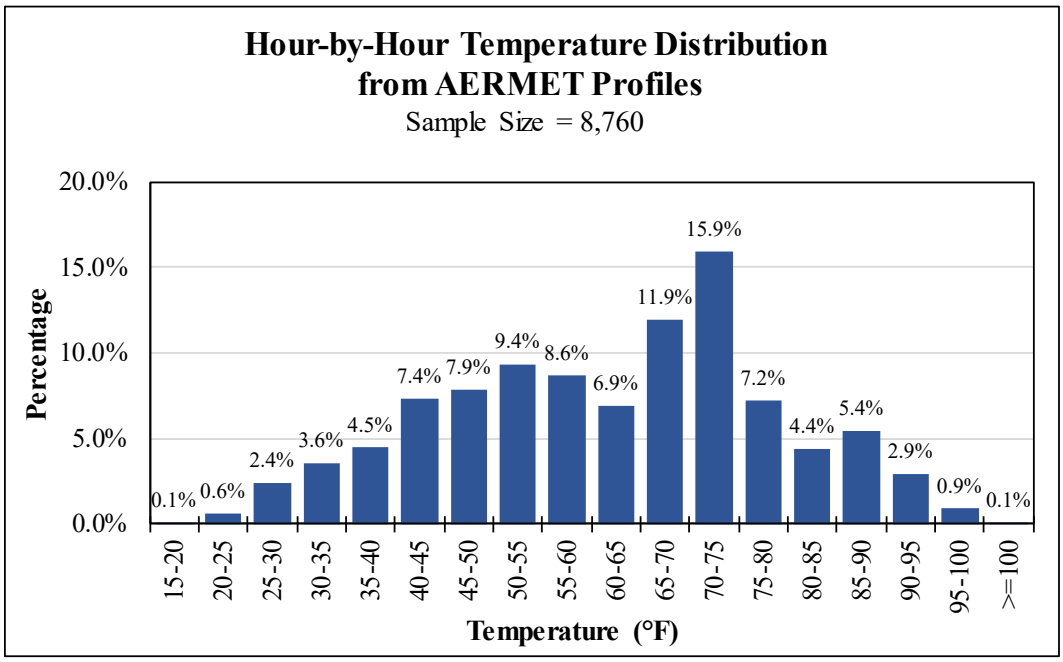


Figure 13 – Temperature Distribution from AERMET Profiles

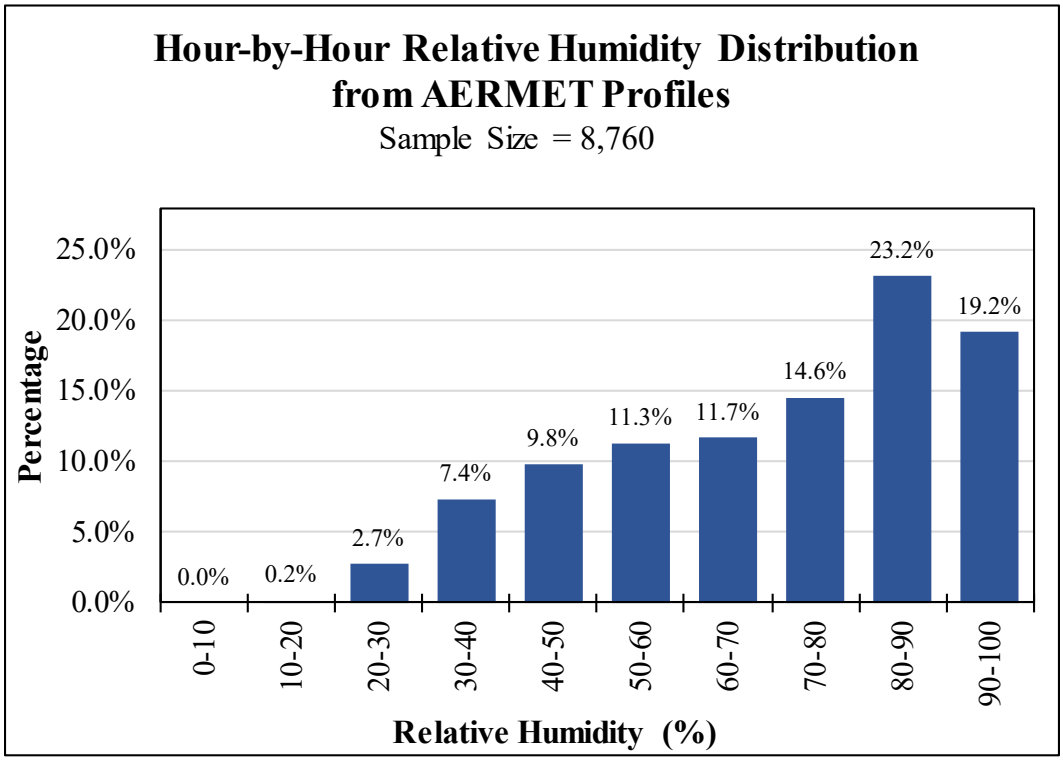


Figure 14 – Relative Humidity Distribution from AERMET Profiles

Traffic Volume and Speed Data

The roadway network adopted in this study is the Activity-Based Travel Demand Model (ABM) by Atlanta Regional Commission (ARC). The specific version employed is ABM2020-TIPA1-2020, which is same version run by ARC staff for the Transportation Improvement Program (TIP), for calendar year 2020, with the path retention method applied and modeled with the ARC planning assumptions for the transportation network, land use, and household demographics for calendar year 2020 (Zhao 2020). The path retention method allows modelers to retain the origin and destination paths used using the Franke-Wolfe algorithm. With path retention, model-predicted link-by-link vehicle traverses through the road network are available for analysis. Therefore, both a full trip record with the departure time and arrival time and origin and destination information and its corresponding specific traces at each time and location are available for use.

The ABM2020-TIPA1-2020 model data is available in geodatabase format and includes a links layer and a nodes layer available. In total, the model network includes 149,967 links and 66,418 unique nodes. The NWC freeway network (and connecting arterial segments) includes 1,570 of the regional roadway links.

For vehicle activity, the team has used the Atlanta Regional Commission's ABM2020 model outputs with path retention, which provides modeled travel demand traffic volumes, link speeds, and individual predicted paths through the network (second-by-second positions can be derived from departure time and average speed on each link traversed by each vehicle). Figure 15 is a screenshot of the 11,399 model-predicted trips that pass through a screen line at Roswell Road and I-75 for one hour of the morning peak period. Because the ABM retains household demographic

data and person assignment for each trip, analysts can apportion modeling results across any demographic cluster derived from ABM model cut points (although demographic analyses not included in this AERMOD project). The team has the ability to generate fleet composition (make, model, and age distributions) for those trips touching freeway links using 2019 freeway license plate observations. The team has also generated fleet composition for arterial trips using random pulls from trip origin TAZ vehicle registration mix employed in regional conformity modeling efforts. However, in this AERMOD project, it was important that the analyses across time and space employ the same fleet composition and model year distributions in all analyses for control purposes. To that end, the team is using fleet composition data employed by the MPO in preparing their regional conformity analyses (discussed later in the report).

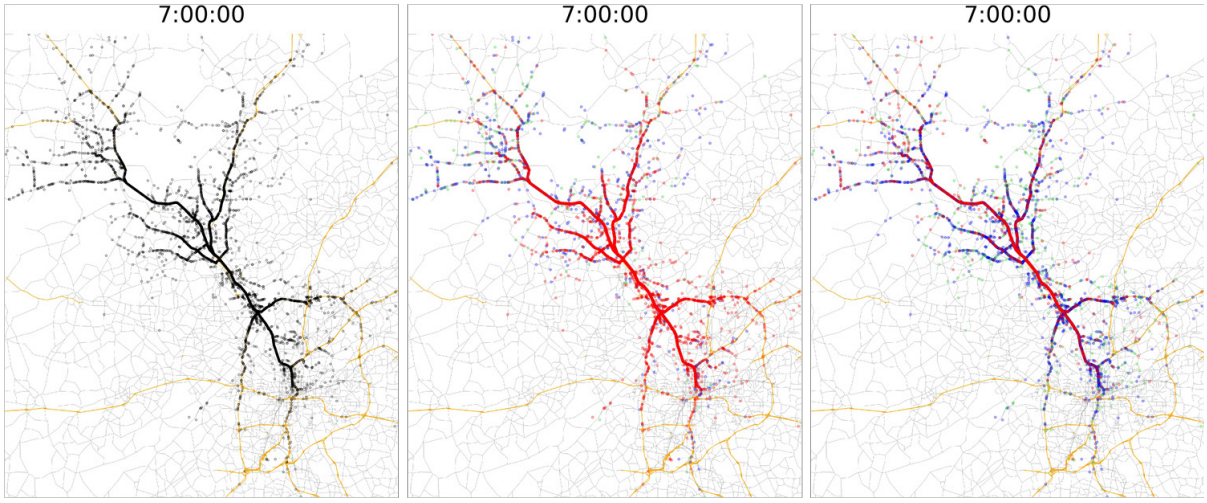


Figure 15 – Path Retention Animation Screenshot

Python scripts have been used to generate AERMOD input files for subsets of sources (link representations) and receptors, run AERMOD on the PACE cluster, and retain and compile results for each receptor in space and time. All source type model runs use the same vehicle fleet composition, traffic volumes, and on-road operating condition inputs.

Hourly Link Volumes and Speeds

The ABM2020-TIPA1-2020 model generates hourly traffic volume and speed, the values remain static within each of the five time periods: EA (early AM), AM (AM peak), MD (midday), PM (PM peak), EV (late evening/night). The EA period ranges from 3:00 AM to 6:00 AM, the AM period ranges from 6:00 AM to 10:00 AM, the MD period ranges from 10:00 AM to 3:00 PM, the PM period ranges from 3:00 PM to 7:00 PM and the EV period ranges from 7:00 PM to 3:00 AM. Specifically, within each of the five time-periods, the ABM processes trips in half-hour increments. More details regarding the ABM2020 model could be found in the model documentation online (ARC 2019).

Fleet Composition

The input fleet composition to emissions modeling were extracted from the previous research by the team (X. Xu et al. 2018) to provide source type distributions and age distributions to MOVES-Matrix. In support of the regional conformity plan, the 20-county nonattainment area was divided into 13-county vs. 7-county areas, where separate fleet compositions were applied for each area in MOVES modeling (ARC 2015). The distribution of counties is shown in Table 2. Cobb County and Cherokee County belong to the 13-county area, and Bartow County belongs to the 7-county area. Four sets of fleet compositions were used in support of the regional conformity plan for the year of 2017, 2024, 2030, and 2040, respectively. The fleet composition calendar year 2017 was used in this research (source type distribution shown in Table 3), and the corresponding fleet composition for each modeled ABM link was allocated based upon county group membership.

Table 2 – County Division of Regional Conformity Plan (ARC 2015)

County Name	Area
Fulton	13-county
DeKalb	13-county
Cobb	13-county
Gwinnett	13-county
Rockdale	13-county
Henry	13-county
Clayton	13-county
Fayette	13-county
Douglas	13-county
Cherokee	13-county
Coweta	13-county
Forsyth	13-county
Paulding	13-county
Bartow	7-county
Carroll	7-county
Spalding	7-county
Newton	7-county
Walton	7-county
Barrow	7-county
Hall	7-county

Table 3 – Input Source Type Distributions

Source Type #	Distribution in 13-County Area	Distribution in 7-County Area
11	2.11%	2.84%
21	53.91%	47.22%
31	31.00%	35.32%
32	10.12%	11.55%
41	0.03%	0.01%
42	0.02%	0.01%
43	0.33%	0.32%
51	0.04%	0.03%
52	1.25%	1.23%
53	0.09%	0.09%
54	0.13%	0.15%
61	0.62%	0.53%
62	0.35%	0.70%

MOVES-Matrix

MOVES-Matrix for MOVES 2014b is a lookup-based energy-use emission rate modeling system developed by the team by pre-running MOVES 2014b in an iterative fashion on the PACE supercomputing cluster across combinations of all model inputs that affect emission rate outputs. To develop the set of emission rates for the Atlanta metro area, the team ran MOVES 146,853 times, iterating across 21 calendar years × 3 fuel specifications (summer, winter and transition) x 111 temperature bins × 21 humidity bins x 21 calendar years (2010-2025, 2030, 2035, 2040, 2045, and 2050). Each iteration produces emission rates for energy use and all pollutants (including pollutant composition types), for every vehicle source type, model year, and on-road operating conditions (by MOVES VSP bin and by average speed and facility type) for each calendar year, fuel specification month, temperature bin, and relative humidity bin. Because MOVES model outputs are also dependent on the specifications of the regional inspection and maintenance (I/M) program strategy (by calendar year), the model runs described above apply to Atlanta and any other region that employs the same fuel specification and I/M programs. Hence, the team is in the process of preparing emissions rate matrices for each of the 117 unique combinations of MOVES 2014b fuel and I/M programs across the United States. As of July 2021, the research team had generated full MOVES-Matrix outputs for 30 of these 117 modeling regions, covering 2,902 out of 3,228 counties as shown in Figure 16 (California does not use the MOVES model).

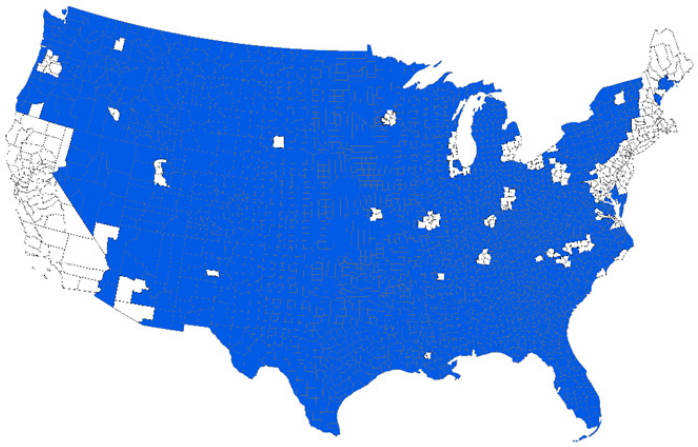


Figure 16 – MOVES-Matrix Coverage Area

By pre-processing all possible combination of input variables, users can query the emission rates directly from MOVES-Matrix, rather than performing new MOVES modeling runs. Users need only identify the subset of the MOVES-Matrix needed for an analysis (calendar year, fuel month, and meteorology data) and run a set of queries to pull required emission rates (by vehicle type and model year and on-road activity) and the queries reassemble the fleet emission rate using the exact same weighting process that is used in MOVES. The queries run 200x faster than running MOVES and users never have to generate MOVES scenario input files, which improves analytical run time efficiency (Liu et al. 2019). The conceptual flow of MOVES-Matrix processing is illustrated in Figure 17.

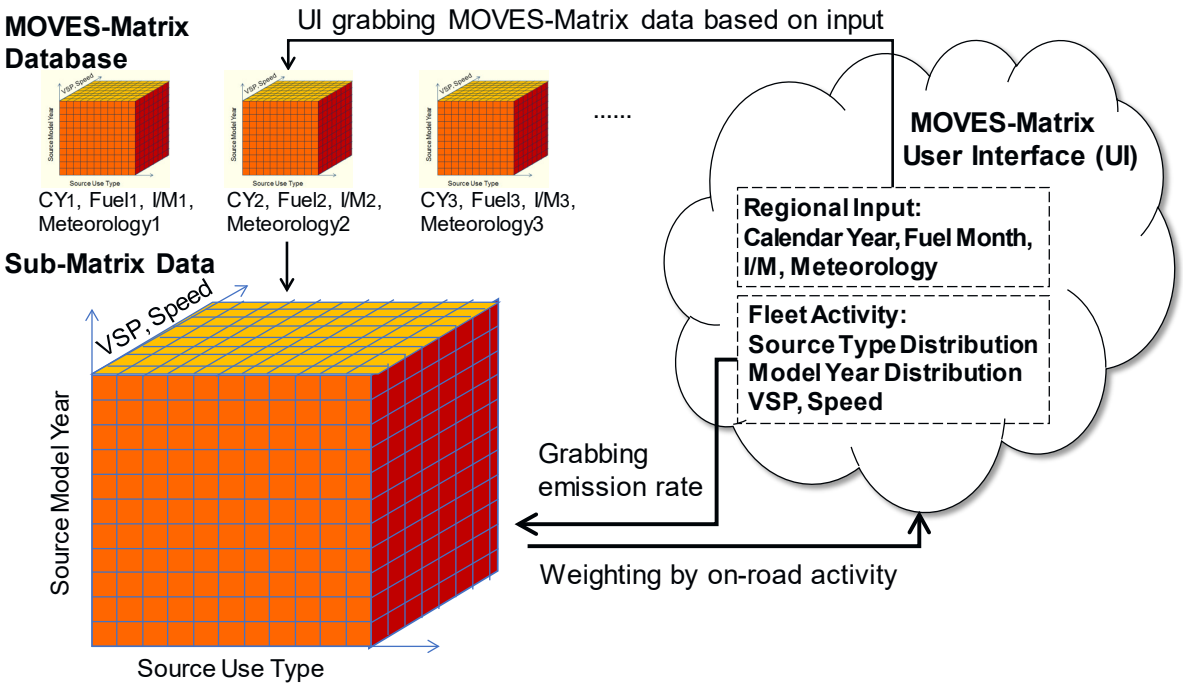


Figure 17 – MOVES-Matrix Conceptual Flow (Liu, et al. 2019)

MOVES-Matrix also enables rapid analyses of engine starts, truck hoteling, evaporative sources, brake/tire wear (X. Xu et al. 2017), and can be used to model the emissions from individual vehicles (Guensler et al. 2017). MOVES-Matrix can be easily coupled with vehicle activity analysis (Li et al. 2017, 2016; Y. Xu et al. 2017) by importing second-by-second vehicle operations. MOVES-Matrix can also be applied to different transportation models, such as travel demand models (X. Xu et al. 2018), and microscopic traffic simulation models (X. Xu et al. 2016).

MOVES-Matrix is highly-desirable for regional-scale dispersion analysis (Kim, et al. 2020), with high-performance to deal with links from large-scale networks, variations in meteorology, and traffic operation input, and with its user-friendly nature to minimize potential human error in running MOVES (especially when it comes increased number of input links).

MOVES-Matrix Emission Rates

As discussed in the previous section, MOVES-Matrix provides a gigantic look-up table for each modeling region. For the Atlanta metro area, MOVES-Matrix contains sub-matrices based on combinations of calendar year, season (Spring/Fall, Summer, Winter fuel season), temperature (0º-110º F with 1º F-bin interval, 111 bins in total for Atlanta), and relative humidity (0%-100% with 5%-bin interval, 21 bins in total for Atlanta). Meteorological data from AERMET are rounded to the appropriate temperature and humidity to link with appropriate sub-matrices for each MOVES-Matrix run.

For each hour in the year of 2019, corresponding sub-matrices within MOVES-Matrix were extracted using Python-based scripts for use in project-level emission rate calculations. Even though the hourly traffic volumes and average link speeds from ABM2020 are consistent across analysis days, and the distributions of vehicle source types and model year are uniform through the year and across analyses, the resulting emissions rates for each hour are different as a function of hourly environmental conditions.

Model Inputs and Outputs

Various sources of input data were used to pull hourly emission rate data from MOVES-Matrix for each analysis, as described in a previous section:

- Speed and volumes are derived from the ARC's ABM2020
- Driving cycles are embedded in MOVES for average speed and facility type
- Source type and vehicle age distributions by source type from regional conformity analyses provided by Georgia EPD
- AERMET meteorology data from regional conformity analyses

Because the sub-matrices needed for the analyses are large (and require a lot of time to upload to PACE), and because MOVES-Matrix is so efficient, the emission rate modeling process was performed on local computers and results were used in PACE. It took less than 12 hours to model all ABM links (1,191 links) for NWC, for the 8,670 hours. The emission rate outputs were compiled and uploaded to PACE to provide the input emissions to dispersion modeling.

Noise Barriers

Roadside noise barriers can have significant effects on the dispersion of traffic-generated air pollutants, especially in the near-road environment (Heist et al. 2014). Hence, accounting for the influence of noise barriers in AERMOD dispersion algorithms is a laudable goal. In AERMOD, barrier impacts are currently modeled using the RLINEXT source type, which allows users to code a noise barrier at a set height and offset from a link. The following subsections detail the process of integrating existing noise barriers into the NWC case study.

GDOT Barrier Inventory

The locations and heights of noise barriers were transcribed from the GDOT noise barrier inventory, which lists the locations and heights of each barrier post. The barrier inventory is kept in PDF format but was converted to an excel spreadsheet with online optical character recognition software and data were manually compared to the original PDF before use. The conversion process was prone to error, particularly involving the numbers 6, 8, and 9, as well as the numbers 4 and 7. After a quick manual scan of the data, the converted locations of each barrier post were plotted in QGIS to identify any unusual post locations. This allowed the team to spot most conversion errors, although there may still have been errors that caused the post location to shift by less than one meter.

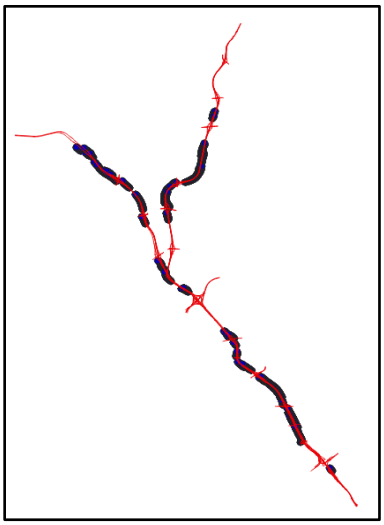


Figure 18 – Noise Barrier Locations

Height Offset and Variability

In total, there are 51 unique barrier segments within the project study area, composed of a combined 5,462 barrier posts and covering a combined length of 33 kilometers. The heights of the barriers range from 2.1 meters to 12.5 meters, with an average height of 6.0 meters. Barrier height even within a segment varies quite a bit. Within each barrier segment, the average height difference between the highest and lowest points is 3.5 meters. Detailed height data were unavailable for about 19% of the barrier posts in the study area (1,048 total).

The variability in noise barrier height complicates the AERMOD modeling process. Barriers modeled within AERMOD must have one consistent height along the entire length of the barrier. However, the barriers within the project study area exhibit noticeably variable heights, as shown in Figure 21 and Figure 22. In these figures, the blue line represents the actual height of the barrier at each barrier post location, and the orange line represents the average height of the entire barrier segment, which might normally be used to model a barrier in AERMOD. Modeling the average of the entire barrier segment fails to account for many local maxima and minima that could alter the

dispersion of emissions nearby. As such, a sensitivity analysis was performed to assess the impact of barrier height on pollutant dispersion. The results of the sensitivity analysis were used to help split AERMOD links into smaller, more homogeneous segments.

Figure 19 shows the distribution of heights among the 4,309 noise barrier posts within the study area. Figure 20 shows the distribution of variances among the 48 noise barriers in the study area. These two charts show that there is a substantial degree of variability in the height of noise barriers within a study area, and within the same barrier.

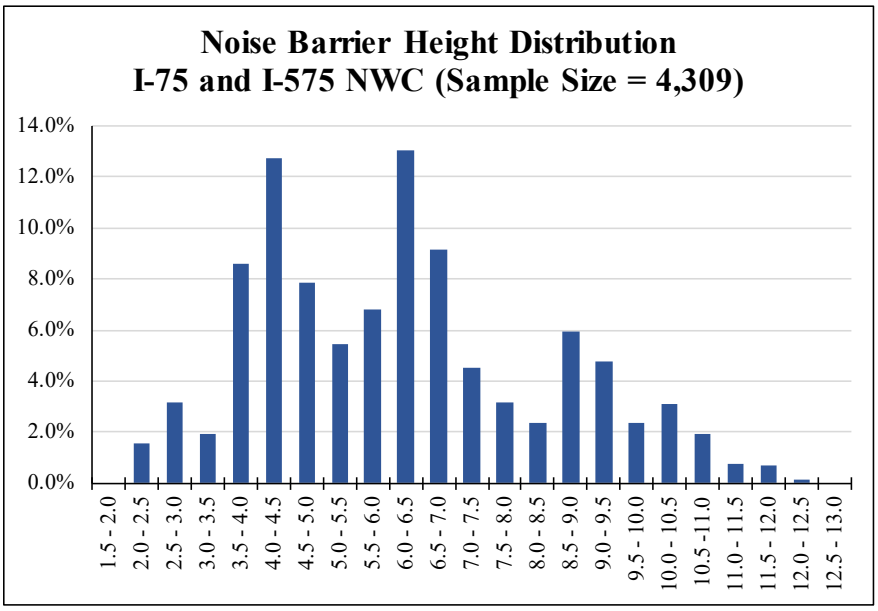


Figure 19 – Noise Barrier Height Distribution

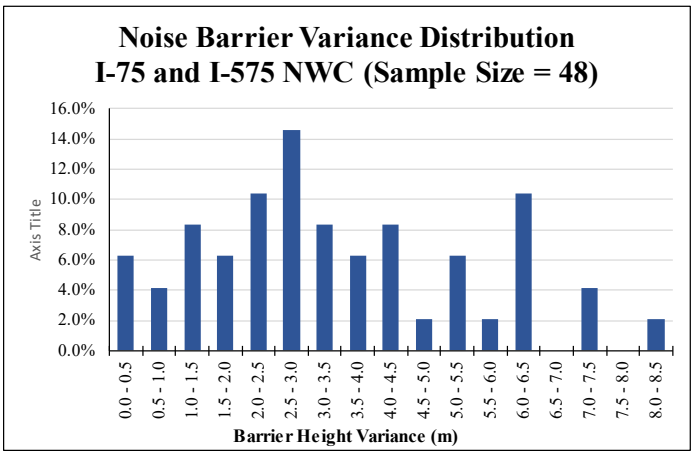


Figure 20 – Noise Barrier Variance Distribution

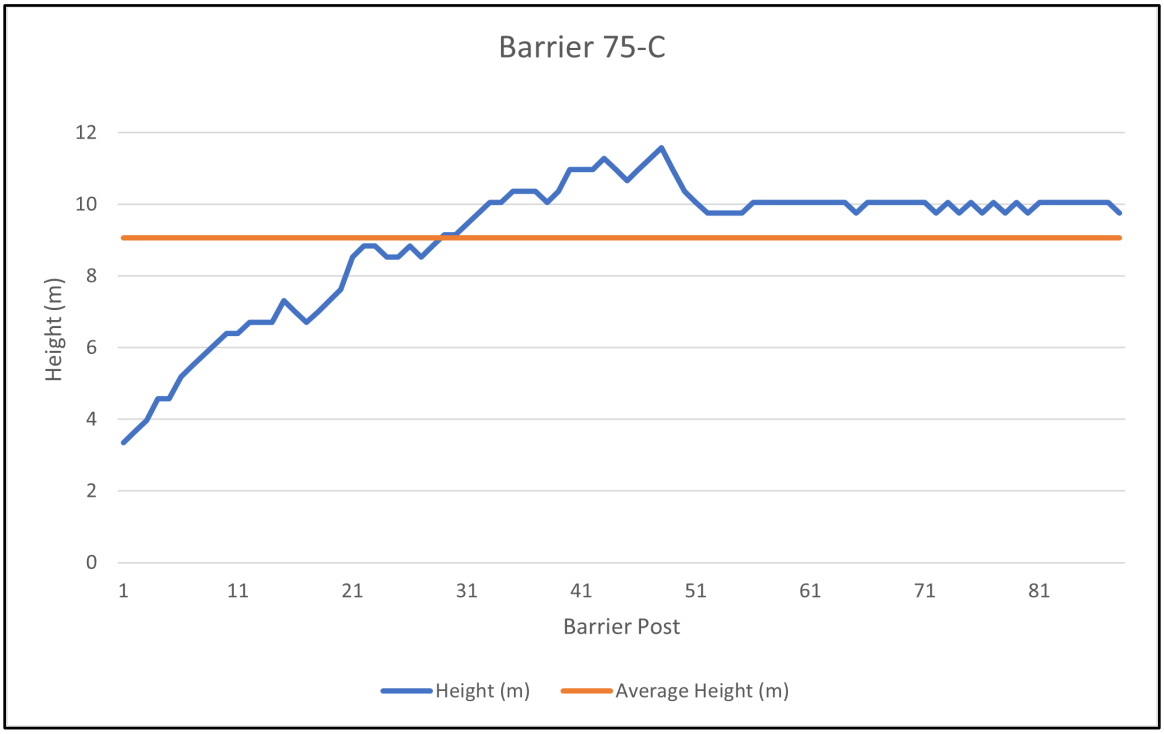


Figure 21 – Barrier 75-C Height Variability Chart

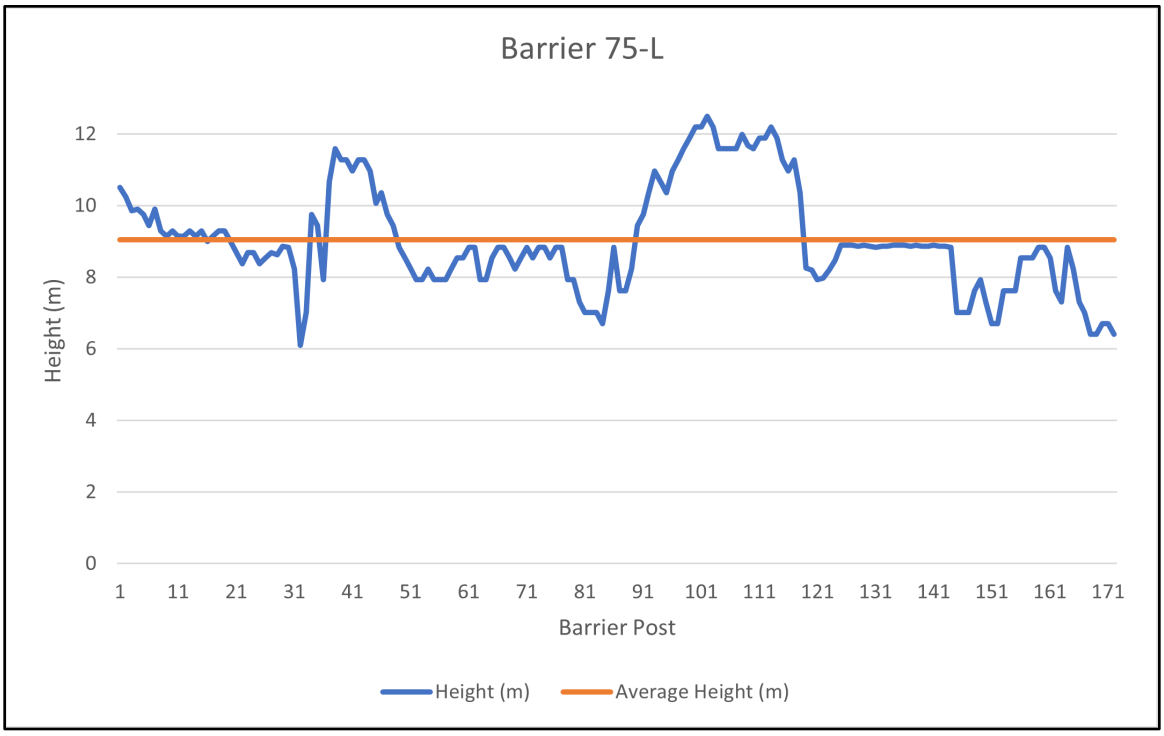


Figure 22 – Barrier 75-L Height Variability Chart

Likewise, there are instances where a barrier is not parallel with the roadway, resulting in a variable barrier offset, as shown in Figure 23. However, AERMOD requires users to define a single offset distance for each barrier. A sensitivity analysis was performed to understand the impact of variation in barrier offset and inform decisions related to splitting larger links into smaller link segments.

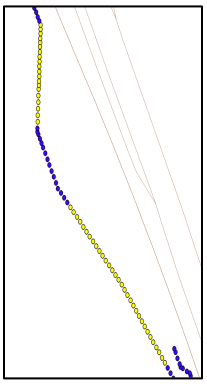


Figure 23 – Noise Barrier with Inconsistent Offset

Noise Barrier Overlap

There are some instances where multiple barriers overlap with the same highway link, as shown in Figure 24. Overlapping barrier configurations are not compatible with AERMOD, which only allows one barrier to be matched with each link. Some NWC roadways may have as many as three barriers along both sides. Analyses traditionally use the barrier nearest to the roadway and ignore the barrier further from the roadway. However, the downwind barrier closest to a given receptor for each link-receptor pair should have the largest impact on dispersion output. As such, links were cloned to create multiple link-barrier pairs for each link. The downwind barrier closest to each receptor is used in modeling emission concentrations at that receptor point. The same receptor may use multiple barrier locations, depending on the direction of the wind at the given moment.

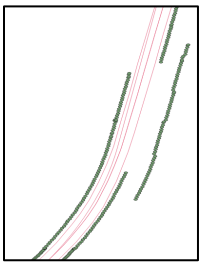


Figure 24 – Overlapping Noise Barriers along a Roadway

Barrier Sensitivity Analysis

As noted earlier, the real-world barriers present in the study area are quite complicated (variability in heights and distances to aligned roadway, offset from parallel, overlapping barrier presence, etc.). However, AERMOD requires that the user specifies a uniform barrier height and roadway offset and then attach the barrier to a single roadway source. Because AERMOD models each barrier as having an impact only on the dispersion from the roadway link to which it is attached, model outputs are dependent on the setup of the roadways and barriers. It is important to avoid improperly associating a barrier with a source that does not have a barrier, and to avoid improperly

failing to associate barriers with their proper links. The AERMOD version used in this research (version 19191) allows attachment of only one barrier to each source. However, splitting source links into smaller link sections (so that barriers can be assigned to link segments) affects model predictions. A sensitivity analysis presented later in this report helps to illustrate the impact of variability in barrier characteristics (heights and distances to roadway), and the impact of splitting of link sources (and association of barrier segments with each source), for various wind speeds and directions.

The first part of the barrier sensitivity analysis investigates the model response to barrier heights and distances to roadway (RLINEXT source) based on an I-575 northbound link placed near the origin of (0, 0), and a reference receptor that was placed at (50, 50) downwind of the source link, as shown in Figure 25. With a fixed wind direction from west to east, barrier setups were compared for upwind vs. downwind. The output at the reference receptor is compared across various barrier heights (2 meters to 10 meters, at an interval of 0.1 meters), and across various barrier distances to roadway (20 meters to 55 meters, at an interval of 5 meters). The other inputs were constant across these scenarios as control variables (cold February morning) to demonstrate the impact of RLINEXT barrier heights and distances to source (including on which side of the roadway) on model output:

- Wind speed: 0.9 m/s
- Wind direction: 270° (west to east)
- Temperature: 36.05 °F
- Humidity: 75%

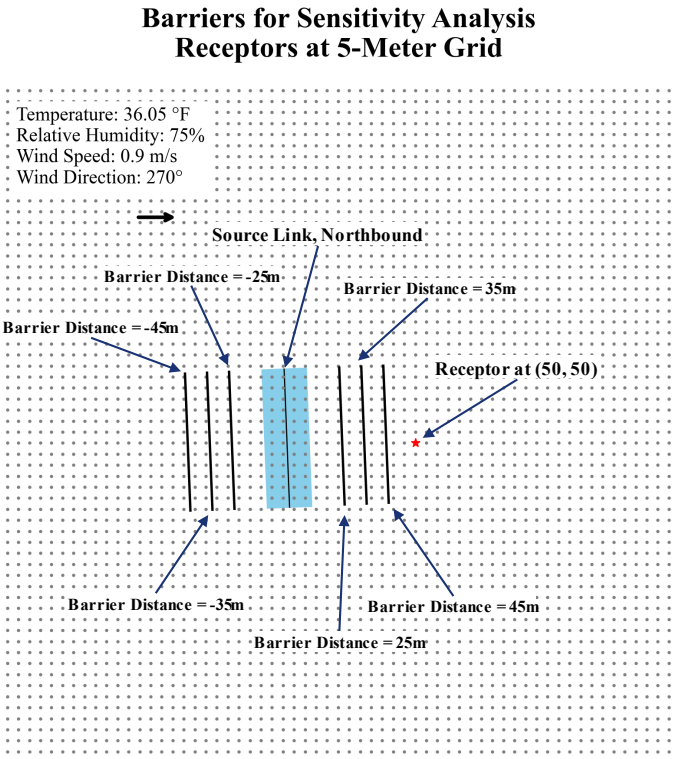


Figure 25 – Source and Receptor Placement for Sensitivity Analysis of Barrier Heights and Distances to Roadway

The second part of the analysis investigates the impact of link-barrier configuration in response to wind characteristics and barrier distance. An imaginary southbound link of 800 meters was split in three ways (no split, even split, uneven split), with various barrier setups (all on the western side of the roadway) to provide ten source-barrier scenarios, as shown in Figure 26 and Table 4. Receptors were placed at 1.5 meter height on a 5-meter grid (excluding over-the-roadway receptors) to compare the predicted concentration profiles across wind direction (0° to 315° at an interval of 45°), wind speed (1.0 m/s to 6.0 m/s at an interval of 0.5 m/s), and barrier distances (20 meters to 50 meters at an interval of 5 meters), with constant temperature (60°F), relative humidity (50%), and input emissions.

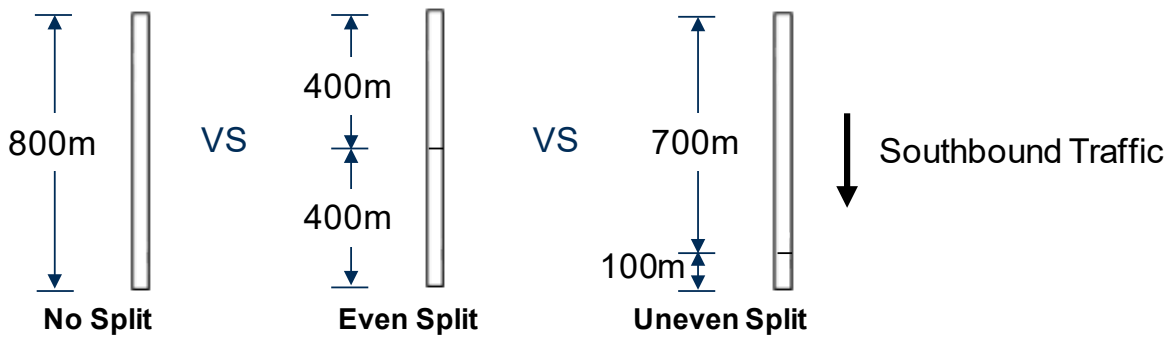


Figure 26 – Scenarios for Sensitivity Analysis by Split Type and Receptor Alignment

Table 4 – Source-Barrier Setup for Sensitivity Analysis

ID	Source Split	Number of Source(s)	Barrier Setup	Number of Barrier(s)
1	No split	1	Without barrier	0
2	No split	1	With barrier	1
3	Even split (5:5)	2	Both without barriers	0
4	Even split (5:5)	2	Upstream source with barrier	1
5	Even split (5:5)	2	Downstream source with barrier	1
6	Even split (5:5)	2	Both with barriers	2
7	Uneven split (7:1)	2	Both without barriers	0
8	Uneven split (7:1)	2	Long source with barrier	1
9	Uneven split (7:1)	2	Short source with barrier	1
10	Uneven split (7:1)	2	Both with barriers	2

The results of the sensitivity analyses indicated that barriers closer to downwind receptors have larger impacts on predicted concentrations, which are also dependent on wind direction (which varies in the AERMET inputs for every hour). Each noise barrier is therefore coded as dynamically aligned with roadway links as a function of the wind direction to minimize the impact of the one-source-one-barrier restriction.

Modeling RLINEXT Sources with Barriers

Based on the sensitivity analyses, the noise barriers were coded following the principles listed below, to compensate for the RLINEXT restrictions mentioned in the previous section.

- No additional line or polygon source splits are implemented that would further divide the sources to potentially better match links with barrier segments (to keep inputs consistent

across LINE, RLINE and RLINEXT source types and to avoid any additional impact from edge effects associated with splitting links that will be discussed later).

- A barrier is attached to a source only when it aligns with at least 50% of the roadway source (with respect to source length).
- A barrier can be attached to multiple sources, if the barrier is more than 50% aligned with more than one link source.
- When two barriers are attached to the same source (e.g., noise barriers on both sides of the roadway), the barrier closer to receptors (downwind of the source) is attached.
- Link-barrier association is therefore modeled dynamically, by hour, to account for changes in hourly wind direction and to meet the other conditions outlined above.

Figure 27 presents a sample of the theoretical design of the methodology to model noise barriers under various wind directions near Chastain Road at I-575. If barriers exist on both sides of the roadway, barriers on the eastern/southern side are modeled when wind blows from northwest (315°), and those on the western/northern side are modeled when wind blows from southeast (135°). The team coded each barrier segment separately with its relative angle to the wind (and consequently downwind receptors), and the actual barriers being modeled vary across hours. Each source aligned with multiple barriers is essentially modeled more than once in the input network (each with one barrier), and the pair with proper barrier attachment (closest to downwind receptors depending on wind direction) is retained in the modeling process, while the remaining barrier(s) are excluded for the modeled hour.

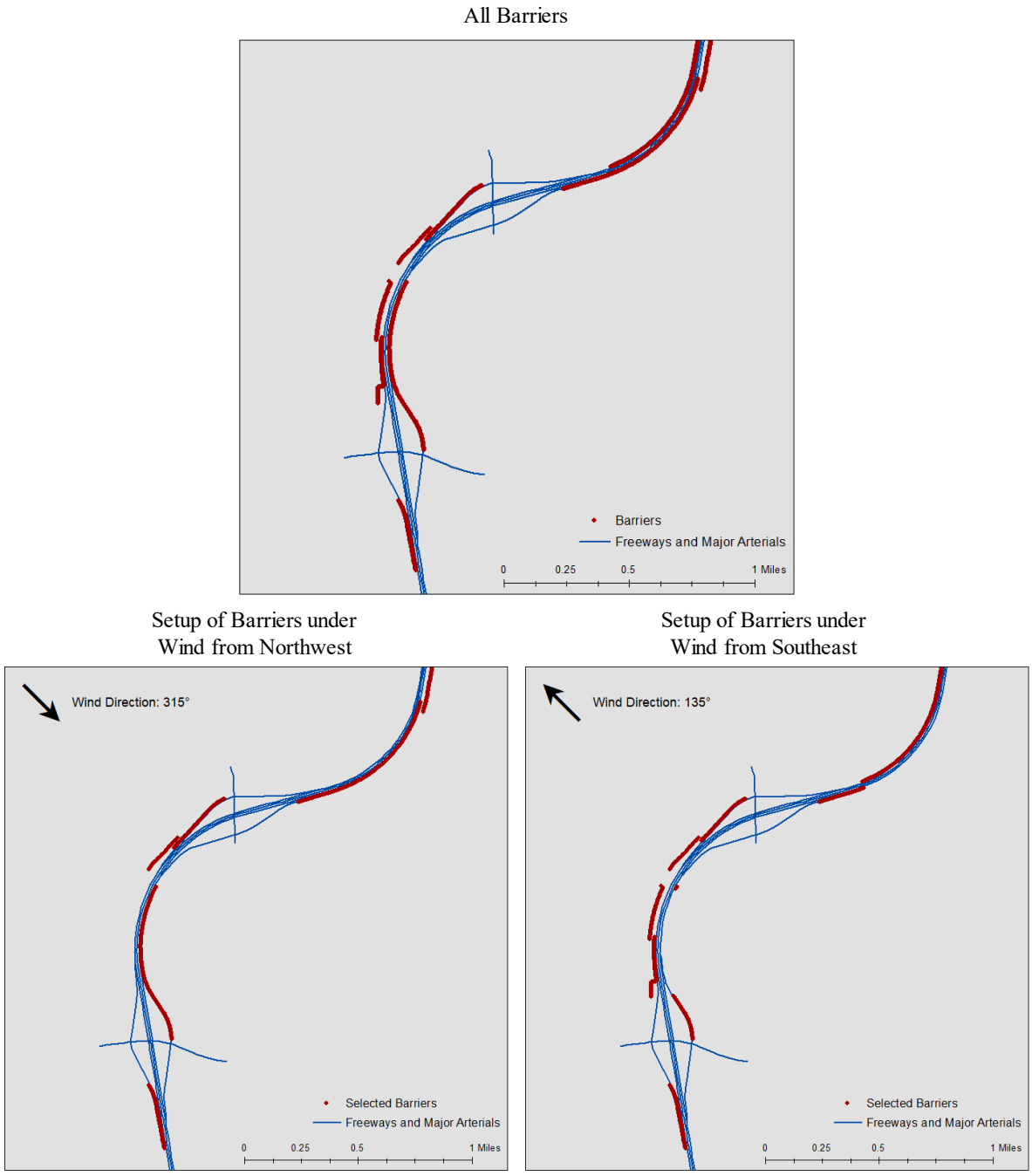


Figure 27 – Sample Barrier Setup at Chastain Road and Bells Ferry Road at I-575

These are theoretical designs in the sense that the dynamically modeled barriers still do not reflect real-world barrier placement, because AERMOD assumes that each barrier is parallel to the transportation link to which it is attached. Because roadway links can have various angles at the

same location (e.g., northbound GP lanes vs. southbound GP lanes vs. Express Lanes), the same barrier can have various impacts (with respect to its spatial setup) across the sources to which it is attached.

AERMOD Runs on the PACE Supercomputing Cluster

As described earlier in the report, the team prioritized finishing receptors of variable grid and of 20-meter standard grid for I-57, with a full utilization of PACE clusters. The emissions modeling based on MOVES-Matrix was conducted locally (running time of less than 12 hours), and the results were uploaded to PACE for Python scripts to pull the emission rates for AERMOD modeling (input generation, parallel launching, and output decoding and compression).

Assignment of AERMOD Emissions Rates

The MOVES-Matrix output provided hour-by-hour (24 hours × 365 days) emissions rates in (grams/link/hour) for each ABM link and the same emissions rates by ABM link were used as the calculation starting point for emissions rate inputs to AERMOD across all source types. These emissions rates were assigned to the corresponding AERMOD sources and were converted to the correct units that correspond to AERMOD source types, as shown in Table 5. For LINE, AREAPOLY, and RLINE sources, if a link was divided into multiple sources, emissions per link were proportionally assigned to each source based on segment length.

Table 5 – AERMOD Unit Conversion from MOVES-Matrix Output

Source Type	Required Unit	Calculation from grams/link/hour
LINE	grams/meter^2/second	Divided by source area and 3,600
AREAPOLY	grams/meter^2/second	Divided by source area and 3,600
VOLUME	grams/second	Divided by number of spheres and 3,600
RLINE	grams/meter^2/second	Divided by source area and 3,600
RLINEXT	grams/meter/second	Divided by source length and 3,600

Link Screening

The link screening tool previously developed by the team (Kim, et al. 2020) was implemented to eliminate roadway source-receptor pairs that do not contribute significantly to predicted downwind concentrations at a receptor. When no link-receptor pairs were significant for a receptor, the receptor was eliminated from all scenario analyses (reducing the number of receptors required to represent downwind concentrations). The machine learning-based screening tool was developed with the support of PACE (Kim, et al. 2020) to examine all source-receptor pairs across all input combinations (emissions rates, wind direction, wind speed, temperature, relative humidity, etc.) to filter receptors with summed predicted concentration (summed from all sources) of less than 0.1 $\mu g/m^3$ under all circumstances. These receptors can then be safely excluded from the analyses without any significant impact on the results.

An iterative filtering process is implemented to reduce the number of source-receptor pairs that need to be screened, further improving the efficiency of the link-screening tool. In the first iteration, receptors on a 200-meter grid (coarse grid to reduce computation burden) enter the screen, and "sensitive" receptors (with results larger than 0.1 $\mu g/m^3$ under any circumstance) are selected (i.e., these receptors are used in the dispersion modeling). The significant receptors naturally form a buffer area in which all receptors on a finer grid (20-meter and 5-meter grids) are selected. In the second iteration, the 20-meter grid receptors immediately outside the border of this area (that is, between a selected receptor and the adjacent screened receptor 200 meters further out) are examined. The 20-meter grid receptors are integrated in the next iteration to expand the selected area (as a result, 5-meter receptors within the expansion are also selected). In the third iteration, the 20-meter grid receptors immediately outside the expanded border are screened, and the final selection is provided as a union of the selected receptors from all three iterations (Figure

28).

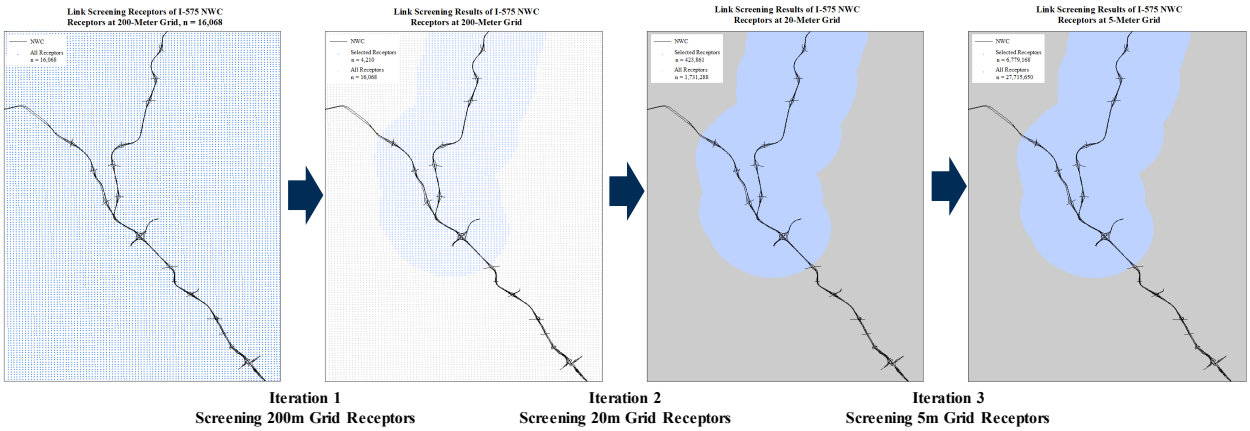


Figure 28 – Iterative Implementation of Link-Screening Tool for I-575 NWC

The iterative implementation of the link-screening tool filtered approximately 75% of the receptors as being non-sensitive and having little influence on the model output, and the total run time for screening was less than two hours (supercomputer not required), as shown in Table 6. The team recommends the use of the link-screening tool in future research projects that involve large-scale networks and fine-grid receptors for dispersion modeling.

Table 6 – Performance of Link-Screening Tool

Network	Grid Resolution (Meter)	Input Number of Receptors	Selected Number of Receptors	Percentage of Receptors Excluded (Workload Reduction)	Screening Running Time (Minute)
I-575 NWC	5	27,715,651	6,779,168	75.5%	80
I-575 NWC	20	1,731,288	423,861	75.5%	5

CHAPTER 4. RESULTS

This chapter provides the results of this analytical work for this study, including machine learning analytics for AERMOD results, predicted concentration across source types, and a sensitivity analysis for RLINEXT with noise barriers.

Machine Learning

In this study, the goal of machine learning model applications was to assess the relative importance of different source types used in the AERMOD model, and to assess how different variables appear to impact AERMOD outputs at different levels and how they interact with each other. Identifying and quantifying the relative feature importance of multiple input variables are the main objectives of these analyses. Whether different AERMOD source types influence predicted concentration results is explored through the feature importance assessment.

Machine Learning Introduction and Methodology

Machine learning (ML) is the study of computer algorithms that automate analytical model building and enable predictive analytics through experience and by the use of data (Mitchell 1997). Machine learning is recognized as part of Artificial Intelligence (AI). Machine learning applications can be used to identify patterns in data and can be applied to a wide variety of applications, such as medicine, email filtering, speech recognition, and computer vision. Machine learning is also commonly used in the transportation field and energy/emission-related analyses.

A variety of algorithms have been developed and refined in the field of machine learning. From a learning approach perspective, machine learning can be divided into three broad categories: supervised learning (with features and labels), unsupervised learning (without labels), and

reinforcement learning. From a use perspective, machine learning can be divided into three broad categories: regression use (continuous numeric prediction), classification use (categorical prediction), and clustering use. Algorithms range from relatively simple designs, including linear regression and naïve Bayes, to more complex designs such as logistic regression, gradient boosted decision trees, and neural-network-based deep learning algorithms (Mohammed 2016).

The specific algorithm selected for analytics in this study is the LightGBM Regressor. The LightGBM algorithm, developed by Microsoft Research Asia in 2017 (Ke, at al., 2017), is a gradient-boosted tree-based (GBDT) algorithm that was refined from traditional GBDT algorithm with Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB). LightGBM has the advantage of supporting multi-scale input variables, which eliminates the need to rescale input features, and the tool supports the use of categorical variables without pre-processing. Moreover, the LightGBM algorithm is good for exploring variable importance and has very fast training speed and high accuracy, which makes it suitable for analyzing millions of data points. If desired, it could also create classifiers with specific thresholds or levels (i.e., LightGBM Classifier).

The goal of feature importance identification is to identify which features are most relevant to the model output (i.e., the dependent variable), and to describe how each feature improves the model. The algorithm selected for exploring feature importance in this section is the SHAP (Shapely Additive Explanations) value, which was developed based on game theory. The approach introduces each feature to the model, one at a time, into the conditional expectation function of the model output, attributing the change produced at each step to the feature that was introduced, and then averaging this process over all possible feature orderings (Lundberg, et al. 2020).

The SHAP feature value summary graphs (e.g., Figure 29) can be interpreted as follows:

- Each feature identified in the plot contributes to the model (each to different degree).
- The importance of input features is ranked from top to bottom, with the top row containing the most important feature and the bottom containing the least important feature.
- Each dot per row represents model output by introducing the specific feature, with the x-axis providing the calculated SHAP value for that output, with records with similar SHAP value clustered together.
- The dot color represents the variable value, where red represents records with large numeric values, blue represents records with small numeric values, and grey represents categorical variables.

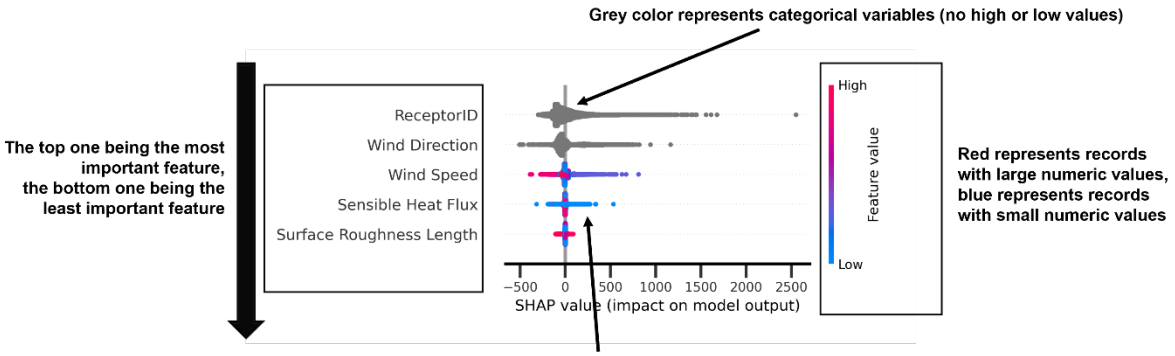


Figure 29 – Example SHAP Summary Plot

The SHAP average value bar plots (e.g., Figure 30) can be interpreted as follows:

- The absolute mean SHAP values for each input feature is calculated by averaging the SHAP values for all model output rows and is presented as a bar plot, regardless of how each individual sample contributes to model outputs.
- All input features are included, and the magnitudes of average SHAP value are presented from top to bottom.

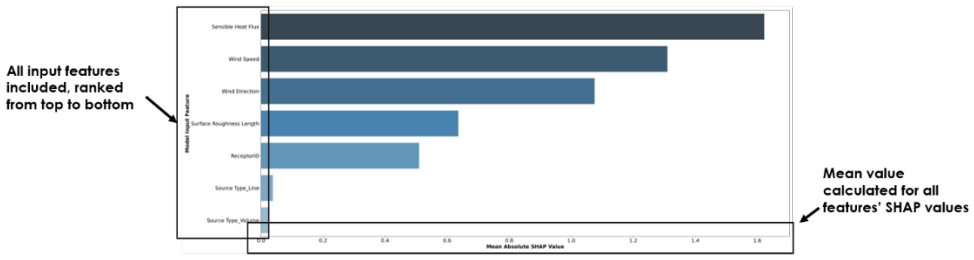


Figure 30 – Example Average SHAP Value Bar Plot

Slight differences can be observed in the ranking of features displayed in the feature value summary graphs vs. the average value bar plots, because the feature value summary graphs accounts for the influence level of different outputs, while the average value bar plots do not.

Data and Setting for Analytical Work

The data used for machine learning analytics in this section are the concentration results for I-575 within 500 meters of the roadway (over-the-road concentrations excluded), with a 20-meter receptor grid, for the 24 hour time period on February 13, 2019 (which represents the worst-case concentration results). The study data include concentration predictions for 60,757 receptors every hour for 24 hours, which yields 1,458,168 AERMOD prediction data rows for analysis. The dependent variable for the LightGBM regression is the prediction concentration in $\mu g/m^3$, the initial set of potential independent variables to be included in the model were identified in previous

work done by Dr. Kim, and excludes variables that showed no statistical significance in his dissertation work (Kim 2020). The variables of interest included source type (LINE/AREA/VOLUME), receptor ID (to control for adjacent roads and their and their mass flux emission rates in grams/meter^2/second), AERMET data that vary by hour, including wind direction (from 0 to 360 degrees, incremented by 1), wind speed (m/s, incremented by 0.01), sensible heat flux (W/m^2, incremented by 0.1) which captures the combined effect on multiple variables of interest, and surface roughness length (meter, incremented by 0.001).

Wind speed, sensible heat flux, and surface roughness length were treated as continuous numeric variables in the model. Receptor ID and wind direction (5° bins) were treated as categorical variables. The AERMOD source types for each run were recoded as dummy variables to allow assessment of differences across model outputs where all other variables are equal for each analysis hour and only source type is changed.

Flow of the Analytical Work in this Report

Figure 31 presents the analytical flow of this section. Analyses presented start with single source type analytics, followed by multiple source type analytics, and then difference analytics. Inside the single source type, clustered modeling was conducted for receptors within different distance ranges, as identified in the previous work that elevated concentration exists near and over the roads and diminishes over distances of 150 to 200 meters (Kim 2020). Over-the-road receptors are not used in regulatory analyses. However, given the charge of identifying differences in predicted downwind concentrations with respect to source type selection, the following comparative analyses include both the over-the-road receptors and away-from-the-road receptors so that the team could explore and assess how the model is treating the "mixing zone" over the roadway by

source type prior to model application of the dispersion algorithms. Multiple source type analytics included all source types in the AERMOD model, with LINE and VOLUME source type coded as dummy variables and distance coded as categorical variables. Difference analytics identified factors leading to the difference in AREAPOLY and LINE as well as AREAPOLY and VOLUME, and visualized receptors with the largest difference in different source types.

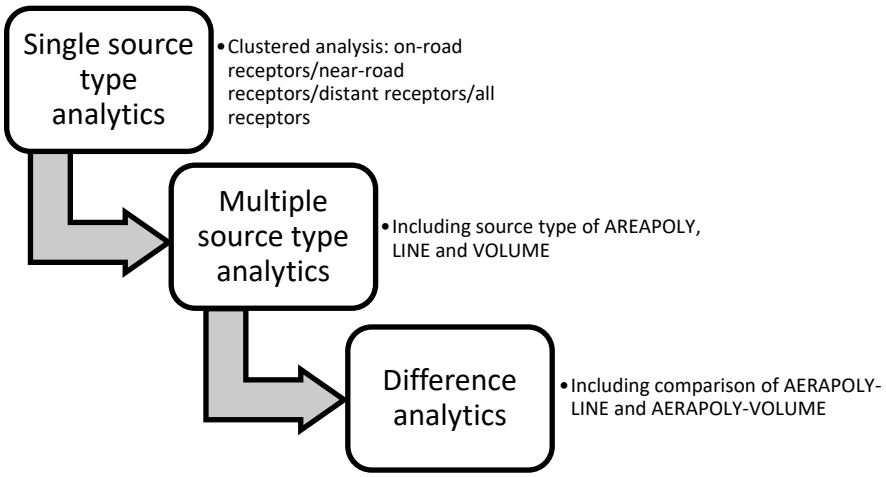


Figure 31 – Machine Learning Analytical Flow

Single Source Type Analytics

Single source type analytics included on-road receptors, near-road receptors (<= 50m) and distance-receptors (> 200m), with their specific receptor number and analysis row number presented in Table 7.

Table 7 – Data Summary for Single Source Type Analytics

Receptor Type	Number of receptors	Number of analysis rows
On-road receptor	2,737	65,677
Near-road receptor	7,282	174,768
Distant receptor	33,864	812,736

Figure 32 through Figure 37 present the SHAP summary plot and average plot for on-road receptors, near-road receptors, and distant receptors. Wind direction and receptor ID are identified as the most influential features, which indicates that higher concentrations arise when the wind

tends to be along primary roadways that have high mass flux emission rates (pollutants accumulate within the plume when the wind runs along roadways with high-emissions rates). On the other hand, surface roughness length was identified as the least influential features. The trend remains the same for all receptors regardless of their distance to roads, with only wind speed and sensible heat flux reflecting slightly different level of feature importance.

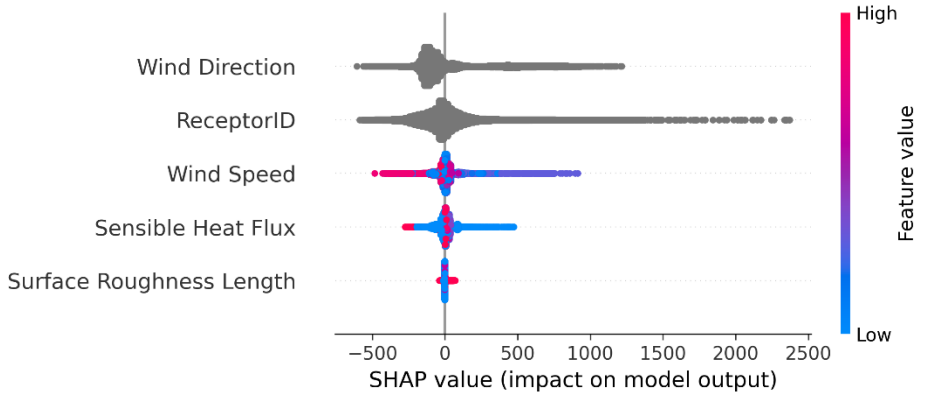


Figure 32 – SHAP Summary Plot for On-Road Receptors

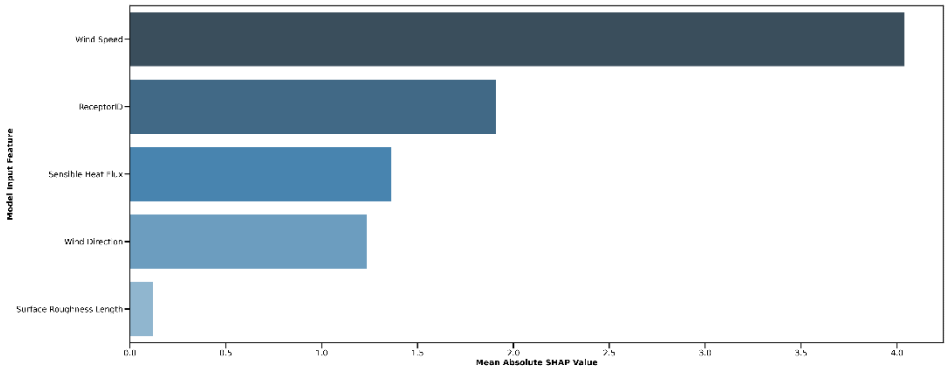


Figure 33 – SHAP Average Plot for On-Road Receptors

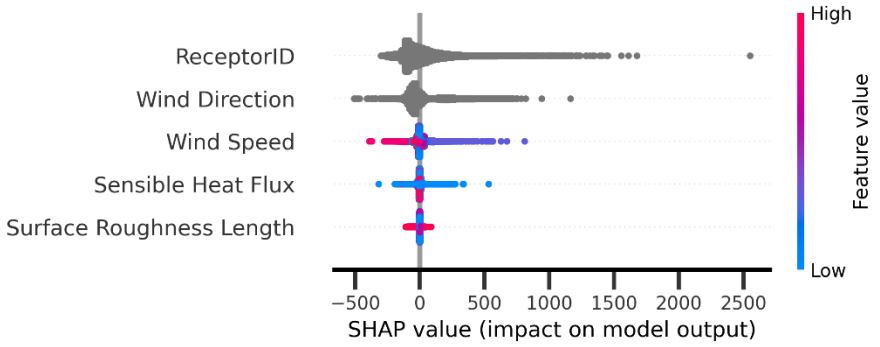


Figure 34 – SHAP Summary Plot for Near Road receptors

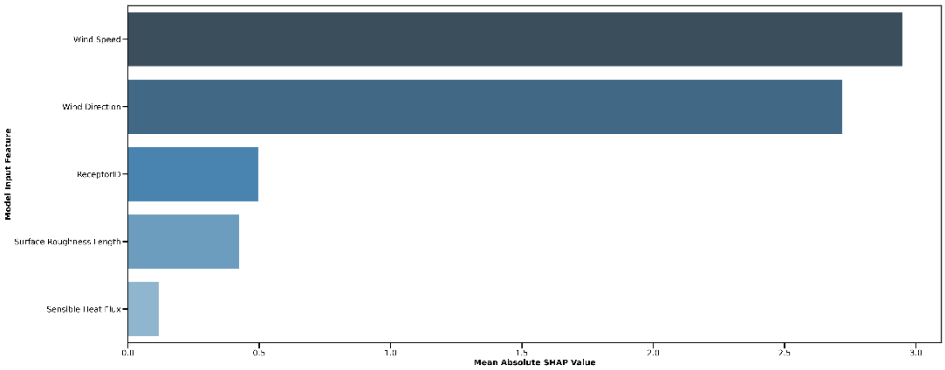


Figure 35 – SHAP Average Plot for Near-Road Receptors

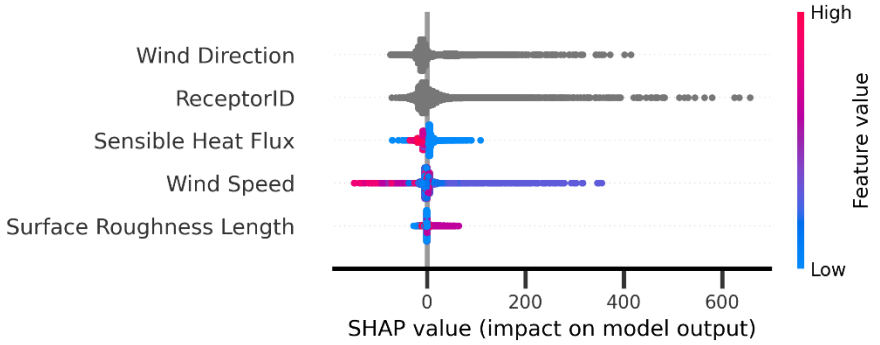


Figure 36 – SHAP Summary Plot for Distant receptors

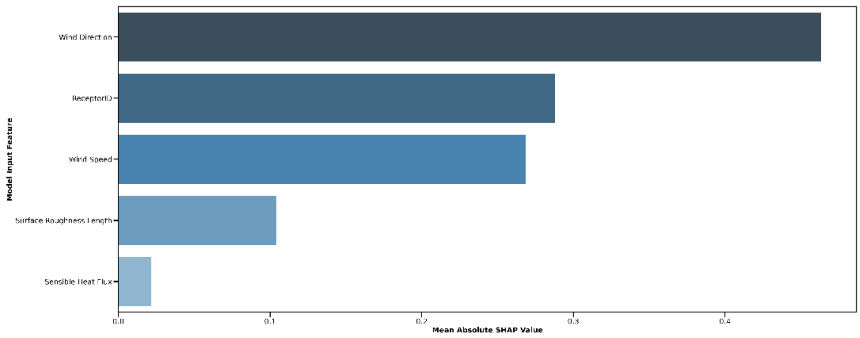


Figure 37 – SHAP Average Plot for Distant Receptors

Multiple Source Type Analytics

Two sets of analyses were conducted for all source types, including an analysis of pollutant concentrations for receptors over-the-road (over-the-road) and traditional modeling receptors located adjacent to the roadways (i.e., not over the road). As discussed earlier, source types were coded as dummy variables to reflect the relative importance of each source type with source type AREAPOLY serving as the baseline, and distance to the nearest roadway was inserted as categorical variable with categories of smaller than 50 meters, 50 to 200 meters, and larger than 200 meters. The specific number of receptors and analysis rows are presented in Table 8.

Table 8 – Receptor Analytics

Receptor Type	Number of Receptors	Number of Analysis Rows
Over-the-Road Receptors	2,737	197,064
Receptors	58,019	4,177,368

Figure 38 through Figure 41 present the SHAP summary plot for on-road receptors and non-road receptors. For on-road receptors (Figure 38 and Figure 39), wind direction, and receptor ID remain the most influential features, followed by wind direction and sensible heat flux. The different source types exert very little influence on the on-road receptor concentration results. For traditional downwind receptors (Figure 40 and Figure 41), distance is the most influential variable (highest feature importance), even larger than wind direction and receptor ID. Similarly, the different source types still show very little influence on the results.

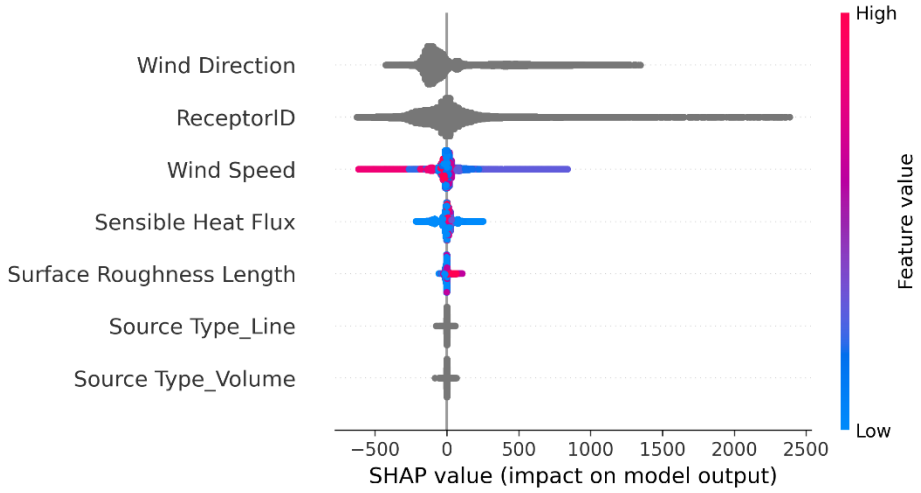


Figure 38 – SHAP Summary Plot for On-Road Receptors

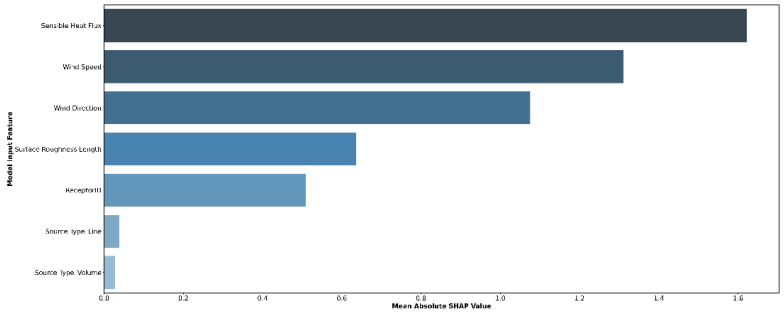


Figure 39 – SHAP Average Plot for On-Road Receptors

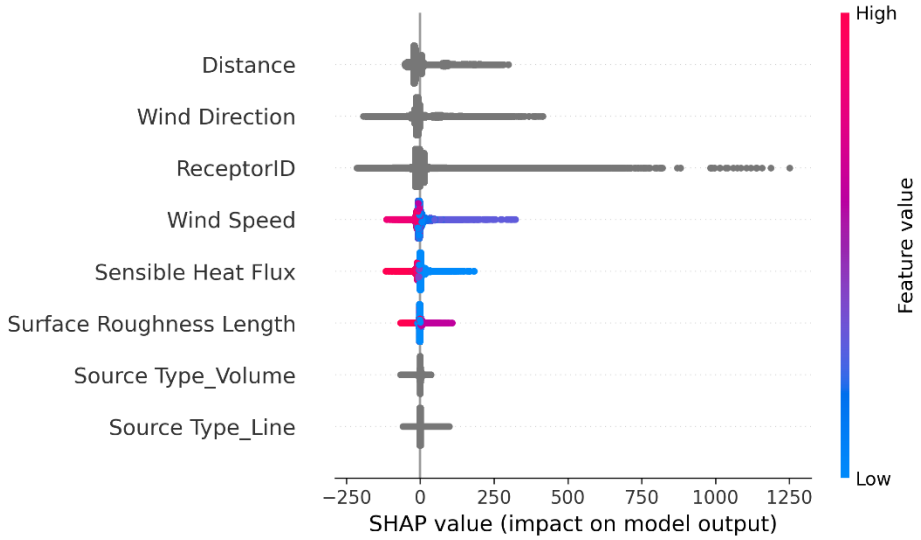


Figure 40 – SHAP Summary Plot for Non-Road Receptors

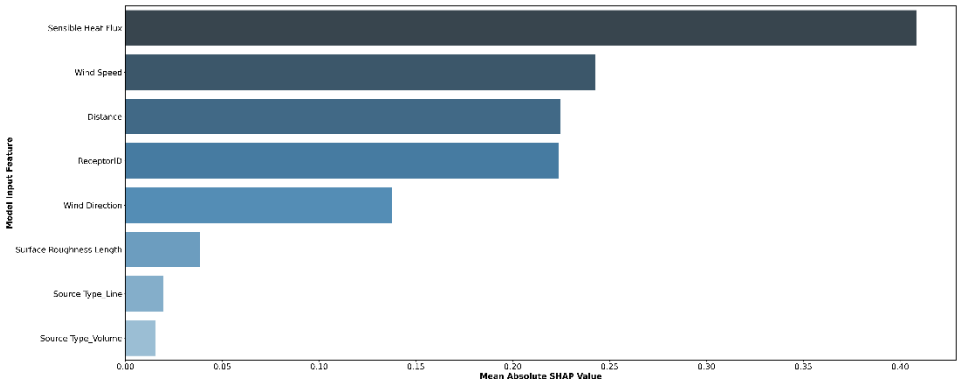


Figure 41 – SHAP Summary Plot for Non-Road Receptors

Difference Analytics

Two rounds of difference analytics were conducted for all 55,341 non-road receptors with 1,328,184 analysis rows, including the AREAPOLY-LINE difference and AREAPOLY-VOLUME difference. Figure 42 through Figure 45 present the SHAP summary plot and average plot for these two models. Wind direction and receptor ID were found to be the most influential features driving concentration differences across source type results, but distance did not appear to exhibit a high feature importance in the difference analytics.

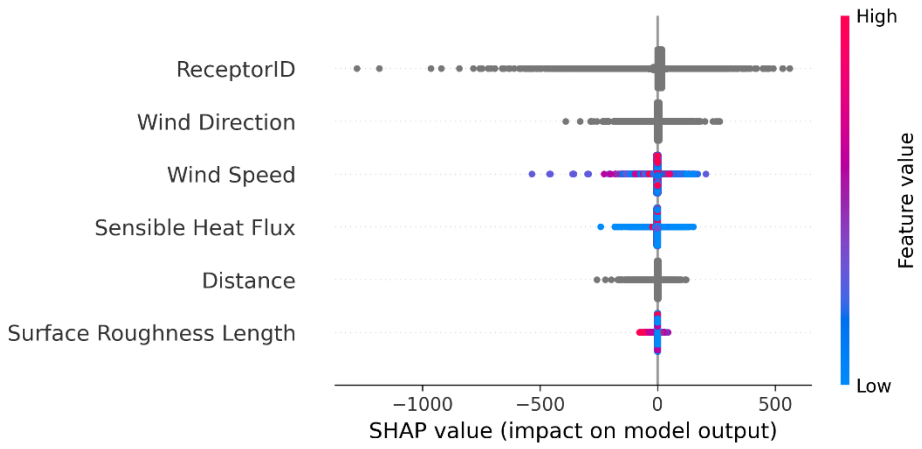


Figure 42 – SHAP Summary Plot for AREAPOLY-LINE Difference

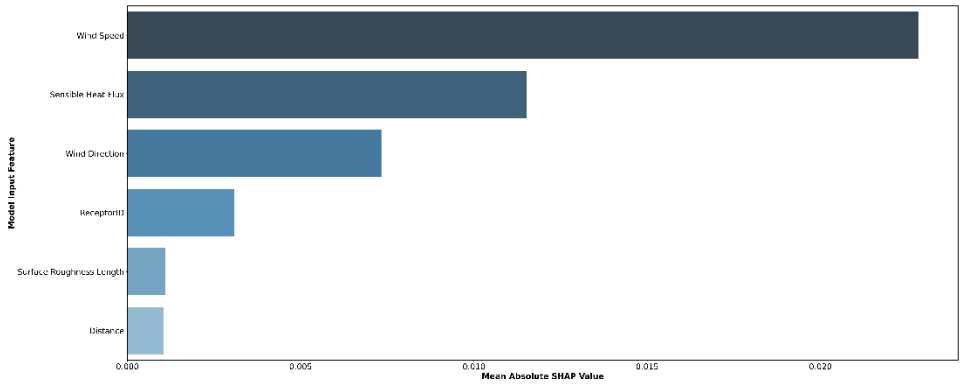


Figure 43 – SHAP Average Plot for AREAPOLY-LINE Difference

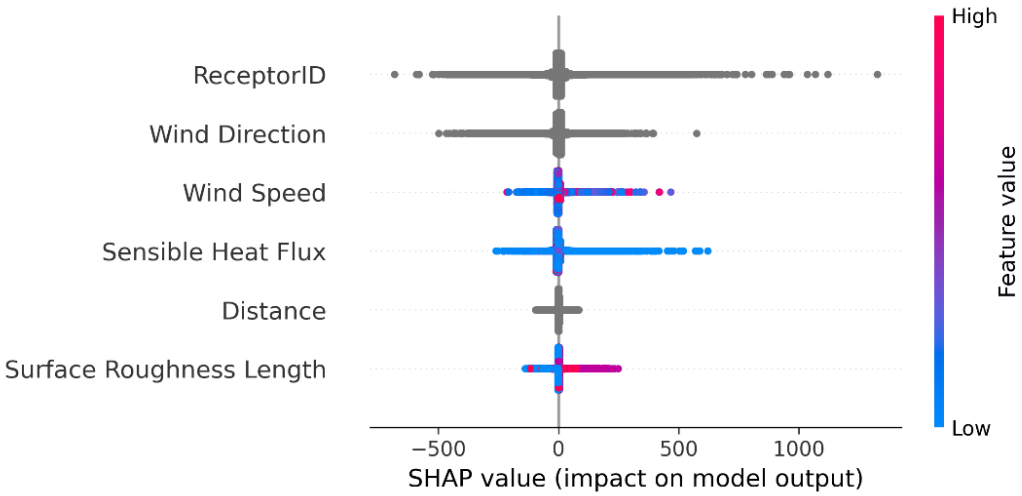


Figure 44 – SHAP Summary Plot for AREAPOLY-VOLUME Difference

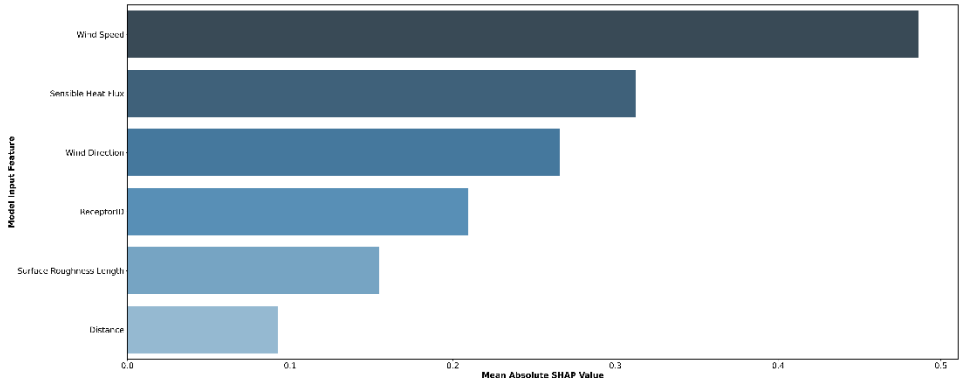


Figure 45 – SHAP Average Plot for AREAPOLY-VOLUME Difference

Receptors with the largest differences in concentrations across the different source types were then checked. Subareas with AREAPOLY-LINE difference and AREAPOLY-VOLUME difference locating outside the 0.01 and 0.99 quantiles were visualized in Figure 46 and Figure 47. Results show that the largest differences appear near roadways, especially near the interchanges with arterials. Hence, the inclusion of arterials' traffic volume and speed appeared to affect the differences in final concentration results. Furthermore, receptors with top 100 largest AREAPOLY-LINE difference and AREAPOLY-VOLUME difference in their concentration results were visualized in Figure 48 and Figure 49. The graphs confirmed that the largest differences between any source type results appear near intersections and along straight-line freeway sections (where downwind concentration differences appear as red bulges).

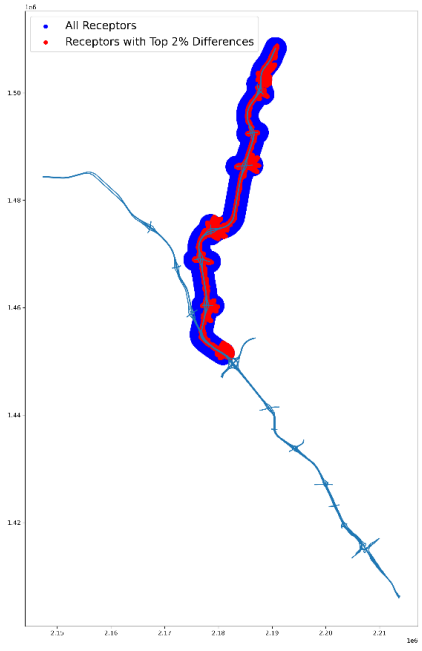


Figure 46 – Receptors with Top 2% Difference in AREAPOLY-LINE

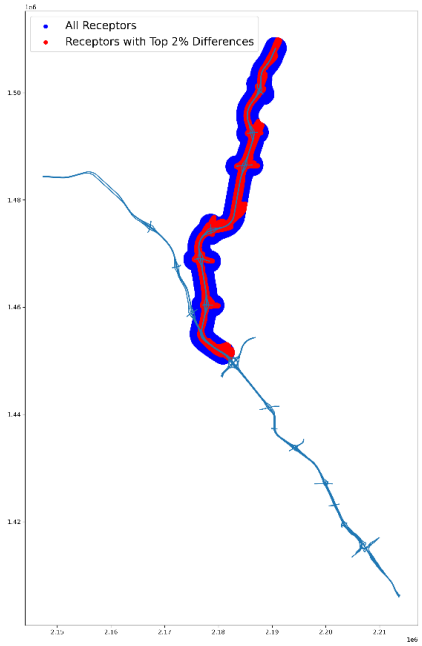


Figure 47 – Receptors with Top 2% Difference in AREAPOLY-VOLUME

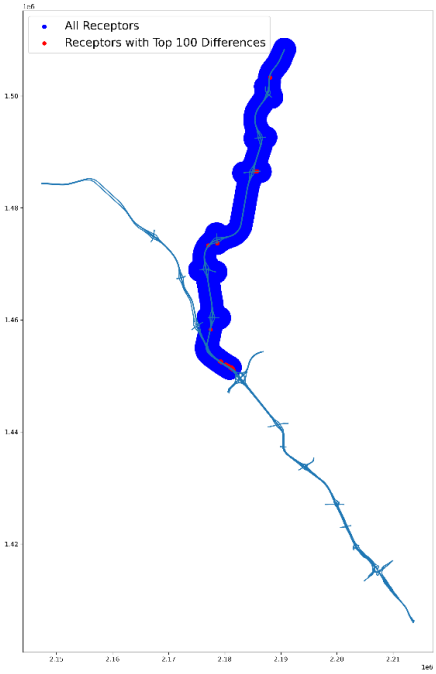


Figure 48 – Receptors with Top 100 Difference in AREAPOLY-LINE

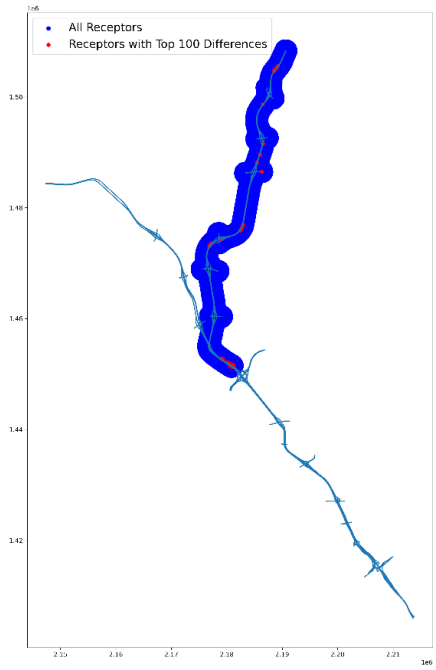


Figure 49 – Receptors with Top 100 Difference in AREAPOLY-VOLUME

Machine Learning Summary

Machine learning analytics were conducted for worst-case single day concentration results along I-575 to identify influential features on results as well as to explore the relative importance of different features. The impacts of different AERMOD source types on results were also studied. For single source type models, it was concluded that using clustered receptor data by distance to roadways helped to improve the model fit and results indicated that near-road receptors and far-from-road receptors are both influenced by mass flux emission rate and wind direction while far-from-road receptors are less influenced by wind speed and more influenced by sensible heat flux. For all source type models, distance to roadway was found to be the most influential variable for non-road receptors, followed by receptor location (and mass flux) and wind direction.

At the sub-regional scale, the source types do have a significant influence on the concentration results, but much less influence than other modeling parameters. Mass flux and wind direction

associated with individual link contributions to downwind concentrations were the most influential factors in prediction differences. Visualizing the results on the receptor grid illustrated that the largest differences across source types occur at near-roadway receptors, with the largest differences located near intersections and downwind of long roadway stretches. Machine learning methods can help to explore feature importance at the aggregate level. However, the team also found that while machine learning results were very useful in previous studies at identifying which parameters are not significant and could be ignored in analyses, machine learning results are less useful at identifying the relative importance of significant factors, given the variability in the AERMET parameters. Nevertheless, the results do indicate that source type significantly influences prediction differences (later confirmed to be related to differences in dispersion algorithms), but that the other factors (wind speed, etc.) are much more influential. Machine learning was not as useful as case study analysis in comparing model outputs across various source types (and other input parameters); hence, routine case study analysis should continue.

Comparison of Sub-Regional Results across Source Types

This section presents the results of the predicted concentrations comparatively across AERMOD source types as heat maps. This research generated the predicted concentration profiles of massive number of runs (8,670 runs per source type), and the output dataset will be uploaded to FHWA's OST Secure File Sharing system. The AERMOD outputs for each modeled hour are presented as heat maps in this research. In accordance with EPA modeling guidelines (USEPA, 2015), receptors over the roadway and within five meters of the roadway (the exclusion zone) have been excluded. The NWC links and pollutant concentrations are illustrated geographically, with predicted concentration (by hour) for each receptor colored in red, with darker color representing higher concentrations (420,753 receptors remain for the 20-meter standard grid after link screening

and removing receptors that fall into over-the-road exclusion zones). A sample of the heat map based on results (RLINE and VOLUME) of 6 PM to 7 PM, April 2, 2019 is presented in Figure 50.

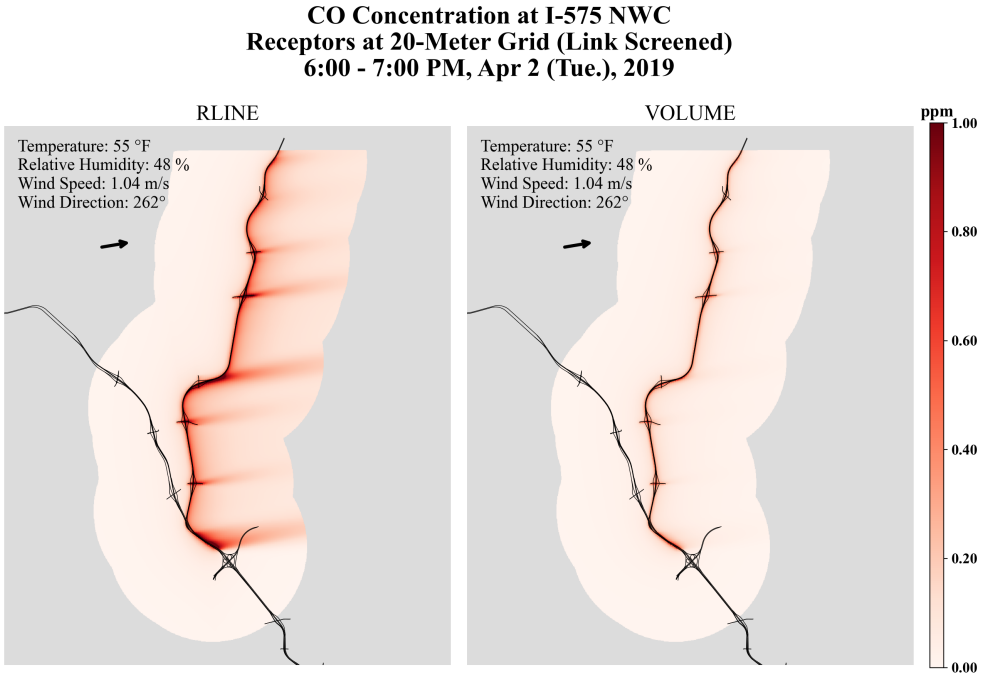


Figure 50 – Heat Map of Concentration Profiles: RLINE vs. VOLUME

The differences across predicted concentrations (by subtracting the results from one source type from the other) are also presented using heat maps, with the color representing the subtraction results (positive as red, and negative as blue). Relative differences are not applicable for most of the cases, because receptors with zero concentration in results from certain source types can have non-zero concentration in results from other source types (to avoid denominators being zero). A sample of the heat map based on differences (VOLUME predictions subtracted from RLINE predictions) for 6 PM to 7 PM, April 2, 2019 is presented in Figure 51.

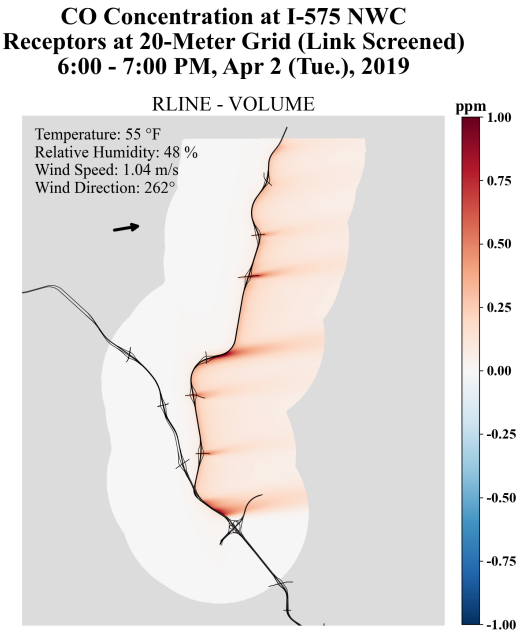


Figure 51 – Heat Map of Differences between Source Types: RLINE vs. VOLUME

The hour-by-hour static heat maps are further converted to animated graphics allowing researchers to visualize the impact of meteorology changes across hours (especially wind direction and speed) on the results, and the animated heat maps for the sample worst cases are uploaded to the YouTube channel created by the team. The video profiles will also be uploaded to FHWA's OST Secure File Sharing system.

It is infeasible to include results for the entire output data set within this report while maintaining a reasonable publication length, so the three worst-cases were developed to present sample output profiles using heat maps of concentration (and differences in concentration) by receptor. These sample cases are shown in Table 9. The first worst case (cold February morning) is presented in the context of this report, while the other two cases (hot summer morning and hot fall evening) are included in Appendix B.

Table 9 – Sample of Worst Cases based on Heat Maps

Date	Weekday	Worst-case Hour	Description
Feb 13, 2019	Wednesday	5 AM to 8 AM	Cold Winter Morning Peak
Jul 29, 2019	Monday	5 AM to 8 AM	Hot Summer Morning Peak
Sep 12, 2019	Thursday	5 PM to 8 PM	Hot Fall Evening Peak

Comparative Results across Receptor Grid Resolution

The results indicate that a 20-meter standard grid combined with a variable grid provide enough detail with respect to comparisons across source types, and is comparable to the 5-meter grid receptor outputs for the LINE worst case results presented in Figure 52. As discussed earlier, the estimated time to run the NWC analyses at 5-meter grid receptors take as long as 200 days (full utilization of PACE resources, without considering job queueing, upload/download time, and system maintenance). The 20-meter standard grid combined with variable grid appears to be sufficient to support the comparative analyses across input source types, and requiring a reasonable amount of computation time. After finishing runs of 5-meter receptors within maximum distance of 200-meters to the roadway (which allowed high-resolution accommodation of near-road results), the team pivoted to running 20-meter grid receptors combined with the variable grid layout.

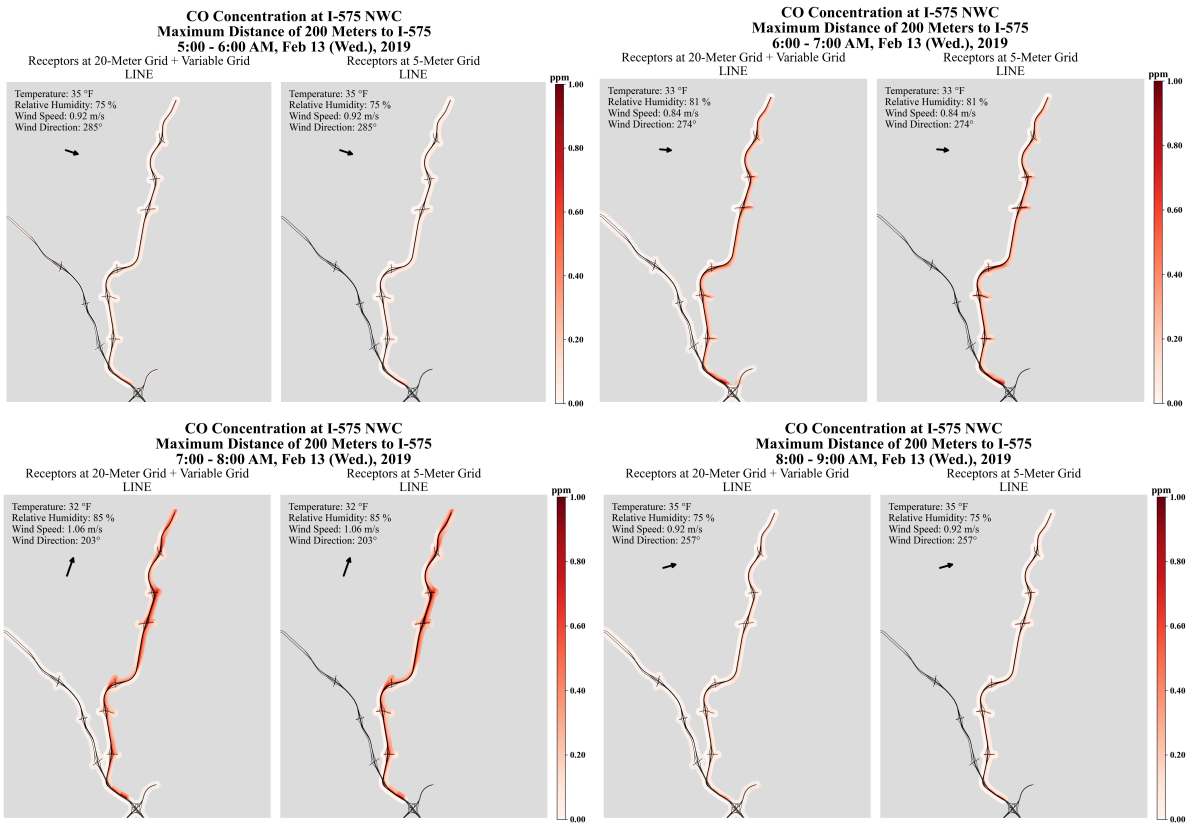


Figure 52 – Predicted CO Concentration of LINE by Receptor Placements (5-Meter Grid vs. 20-Meter Grid Combined with Variable Grid)

Verification of Automatic Generated AREAPOLY Links

The output from AERMOD (source type of AREAPOLY) based on manual creation of polygons vs. automatic generation of polygons are presented in Figure 53, and their differences are shown in Figure 54. The comparison indicates no observable differences. The team concludes that the links generated automatically can be used in AERMOD modeling work without causing any significant impact on the model output (these very small differences can be safely ignored in analyses across source types).

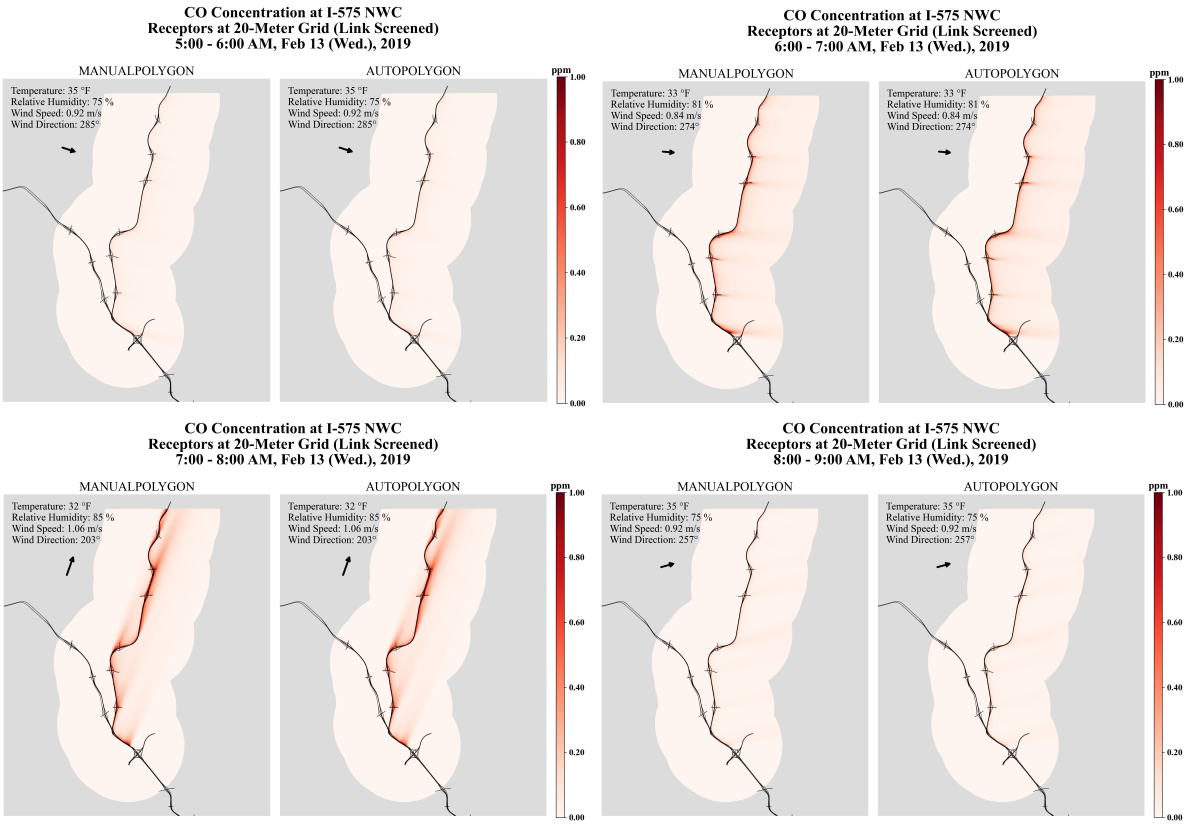


Figure 53 – CO Concentration Comparison at I-575 NWC from AREAPOLY (Manually Created vs. Auto Generated), Cold Winter Morning Peak

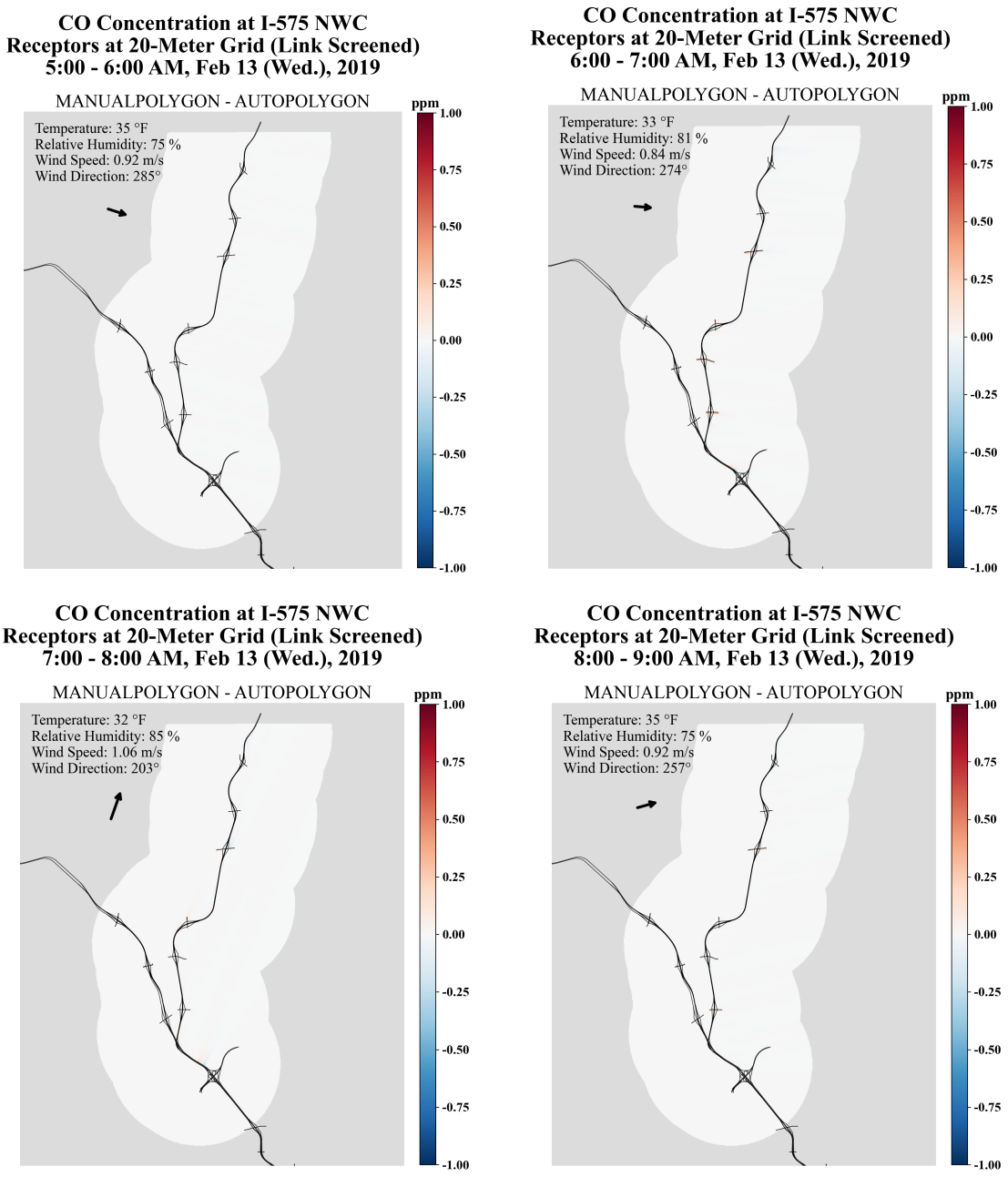


Figure 54 – Differences in CO Concentration at I-575 NWC from AREAPOLY (Manually Created vs. Auto Generated), Cold Winter Morning Peak

Sample Worst Case (Cold Winter Morning) Comparative Results

This section presents the concentration profiles of CO of the sample worst case of cold winter mornings (standard 20-meter receptor grid after link screening), predicted for the various source types, as well as the differences across the source types. As with the previous analyses, receptors over the roadway and within five meters of the roadway have been excluded (USEPA, 2015). The other two sample worst cases (hot summer morning and hot fall evening) are presented in Appendix B. Heat maps across comparable source types for the same hour are presented alongside each other (AREAPOLY vs. LINE vs. VOLUME, LINE vs. RLINE vs. RLINEXT (no barrier), and RLINEXT with barriers vs. RLINEXT without barrier).

Comparative Results for AREAPOLY, LINE, and VOLUME

The CO concentration results of AREAPOLY, LINE, and VOLUME are presented in Figure 55, and their differences are shown in Figure 56. No significant differences between results from LINE and AREAPOLY are observed, which was anticipated because the modeling algorithms are similar. This is even more significant, given that AREAPOLY shapes were automatically generated to match ABM links, and not as rectangles to match LINE sources.

Significant differences were noted between VOLUME and LINE/AREAPOLY with respect to downwind and upwind concentrations. The predicted downwind concentrations from the VOLUME source type are more evenly distributed in x-y space and much lower than the results from AREAPOLY and LINE. Concentrations from VOLUME are also significant and non-zero for receptors located upwind of the source. This indicates that a significant wind meander element has been triggered within AERMOD dispersion routines when the VOLUME source type is used as input (not triggered for the AREAPOLY or LINE source types). The VOLUME source upwind

impacts can clearly be seen in Figure 55 (in the small figure appearing in Row 2, Column 3) and the red tinge representing an upwind difference that can be seen in Figure 56.

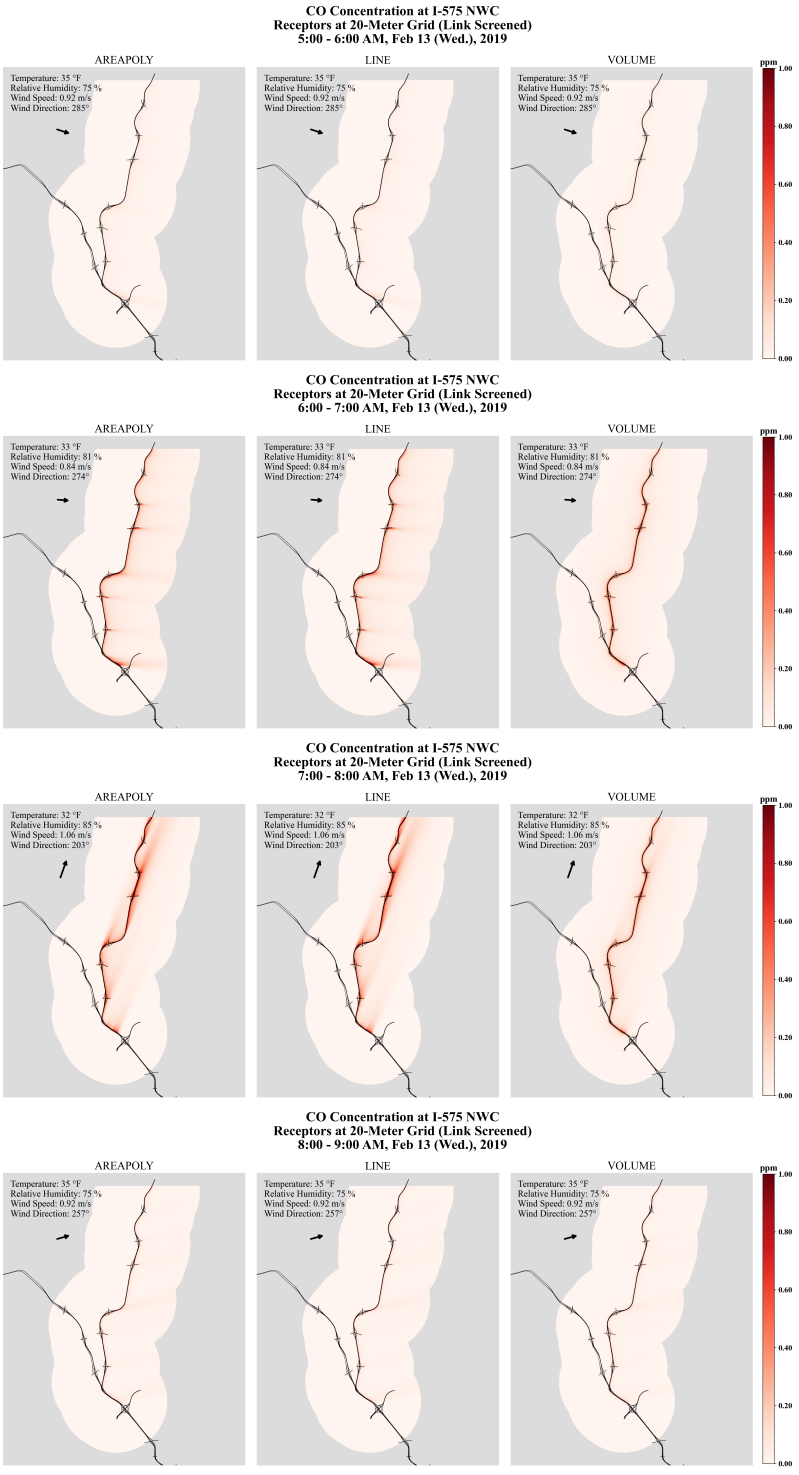


Figure 55 – Predicted CO Concentration at I-575 NWC by Source Type (AREAPOLY, LINE, and VOLUME), Cold Winter Morning Peak

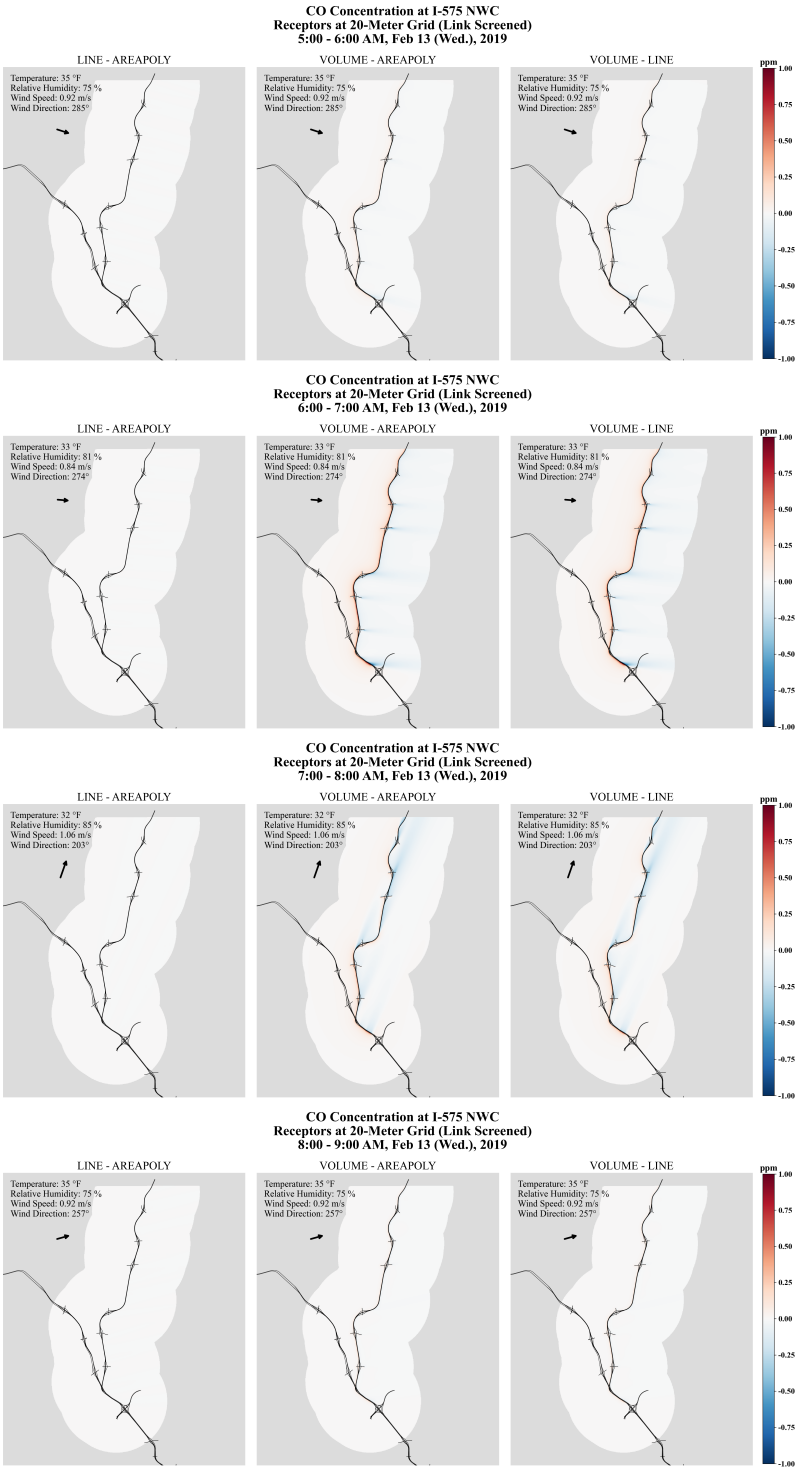


Figure 56 – Differences of Predicted CO Concentration Comparison at I-575 NWC by Source Type (AREAPOLY vs. LINE vs. VOLUME), Cold Winter Morning Peak

Assessing the Impact of the VOLUME Dispersion Algorithms

As discussed earlier, when the VOLUME source type is used, AERMOD predicted higher upwind concentrations and lower downwind concentrations in worst-case analyses than predicted using the AREAPOLY or LINE source types. The differences are the result of wind meander algorithms employed only with the VOLUME source type. To assess the impacts of the wind meander algorithms, the team performed additional VOLUME source type case study analyses with the wind meander code activated (MaxFRAN=1, default) and with the wind meander code deactivated (MaxFRAN=0). Sensitivity analysis was performed as a function of wind speed.

A case study North-South link of 400 meters was created and coded with volume source spheres eight meters in diameter. The wind direction was fixed at 270 degrees (from West to East), and the wind speed was varied from 0.1 m/s to 6.0 m/s in 0.1 m/s intervals. All other meteorology inputs were set to the same real-world AERMET profiles for 5:00 AM to 6:00 AM, Feb 13, 2019 (the worst-case cold winter morning presented in the previous sections). Receptors were placed at 5-meter horizontal resolution and 0.5-meter vertical resolution (a standard 3-D grid), with no receptors placed over the roadway or in the exclusion zone.

Figure 57 shows the resulting heat map for a wind speed of 1.0 m/s and Figure 58 shows the resulting heat map for a wind speed of 3.0 m/s. These results confirm the upwind and downwind impacts noted in the previous worst-case analyses for VOLUME sources (i.e., higher upwind and lower downwind concentrations than obtained from LINE and AREAPOLY sources) when the wind meander option is turned on. The results also infer that VOLUME source results are similar to AREAPOLY and LINE at higher wind speeds (e.g., 3.0 m/s as shown in Figure 58).

The concentration profiles vs. wind speed at a downwind receptor (50 m, 0 m) were extracted from the results and illustrated in Figure 59. Similarly, concentration profiles vs. wind speed at an upwind receptor (-20 m, 0) were extracted from the results and shown in Figure 60 (the upwind receptor is 10 meters away from the exclusion zone). Figure 59 and Figure 60 demonstrate that VOLUME source types with wind meander activated lead to very different results when wind speed is lower than approximately 2.5 m/s. Turning off the wind meander (setting MaxFRAN=0) produces the same concentrations for VOLUME vs. LINE and AREAPOLY. The differences in predictions are large enough to warrant a review of the associated dispersion algorithms employed with VOLUME source type and the associated AERMOD FORTRAN code.

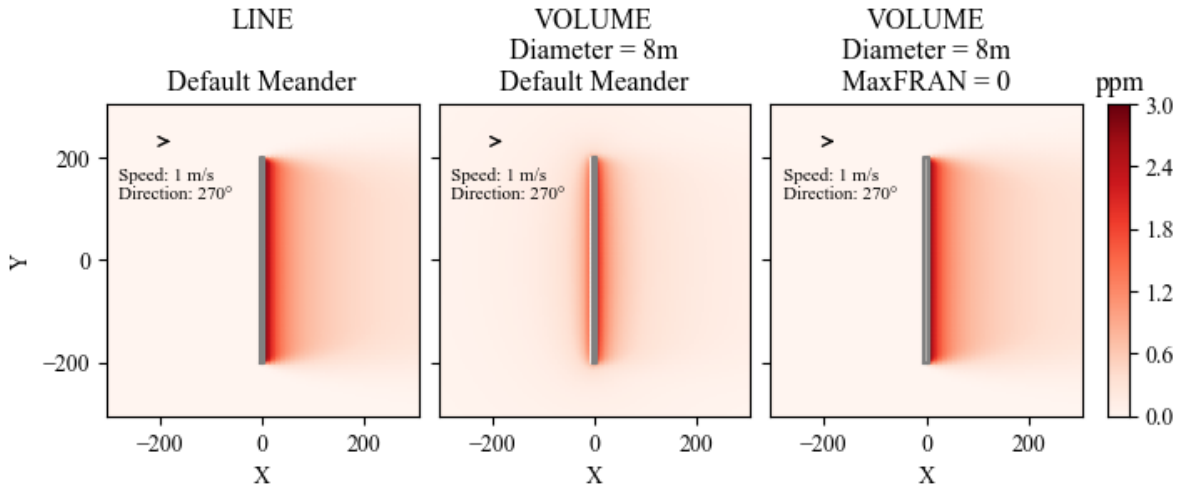


Figure 57 – Predicted CO Concentration for Meander Impact Assessment for a Wind Speed of 1.0 Meter/Second

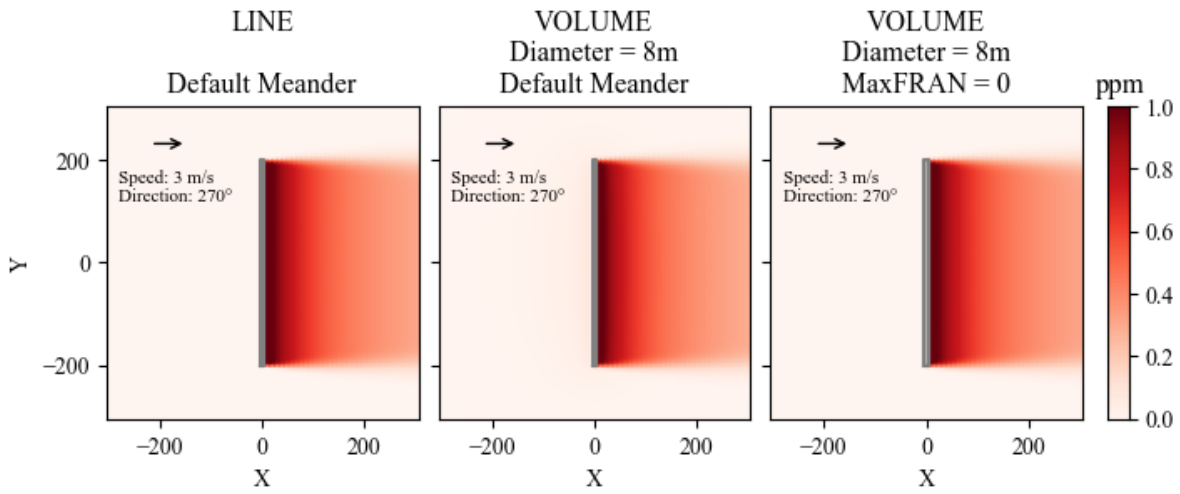


Figure 58 – Predicted CO Concentration for Meander Impact Assessment for a Wind Speed of 3.0 Meters/Second

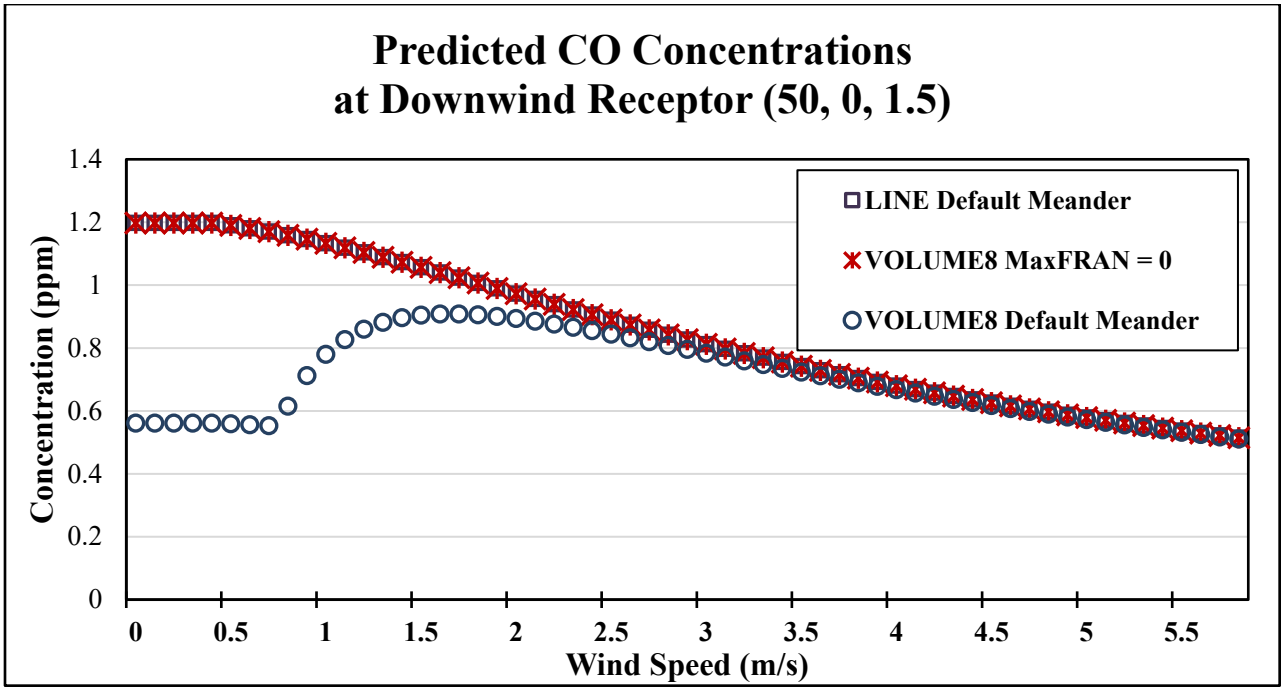


Figure 59 – Predicted CO Concentration vs. Wind Speed at a Downwind Reference Receptor (50, 0, 1.5)

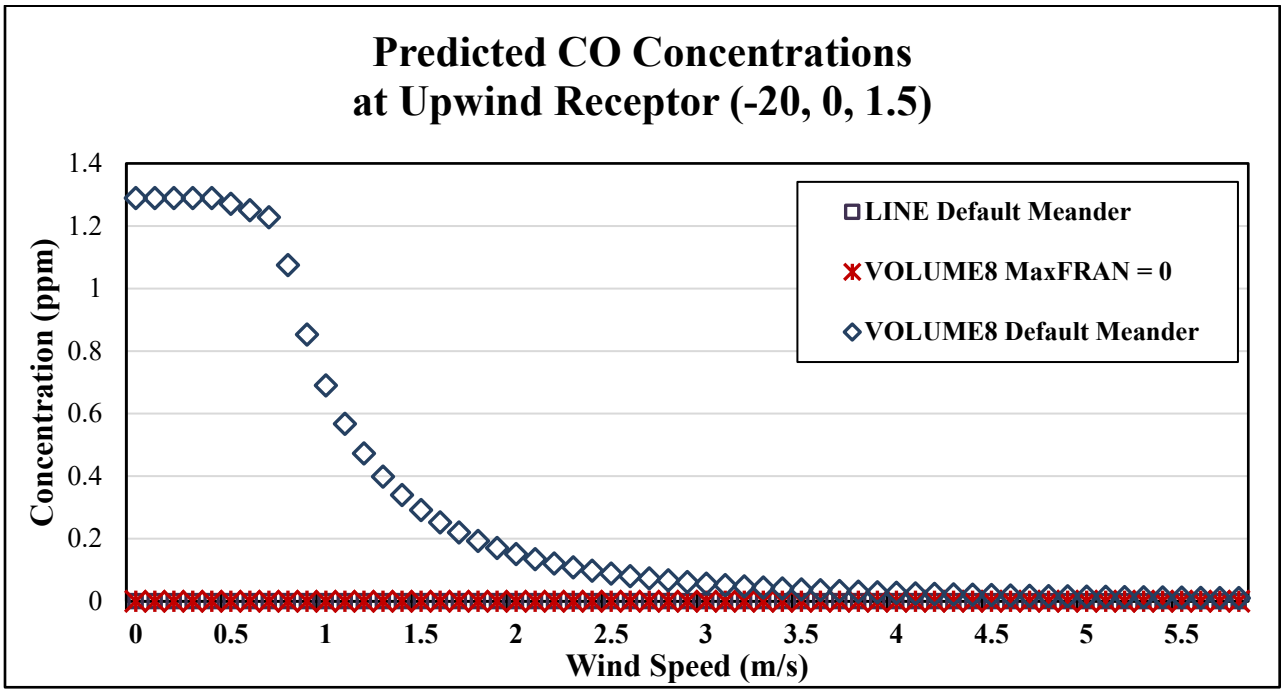


Figure 60 – Predicted CO Concentration vs. Wind Speed at an Upwind Reference Receptor (-20, 0, 1.5)

Comparative Results for LINE, RLINE, and RLINEXT (No Barriers)

The CO concentration results of LINE, RLINE, and RLINEXT are presented in Figure 61, and their differences are shown in Figure 62. Overall, no major differences were noted between RLINE and RLINEXT when no barriers are present, which is not surprising given that their dispersion algorithms are the same. However, differences were still observed at some receptors and, surprisingly, these differences tended to be a bit larger than the differences noted earlier between LINE and AREAPOLY (typically 1%, but as large as around 5% for some receptors). Overall, the concentrations predicted using the RLINE and RLINEXT source types without barriers, and their associated dispersion algorithms, are higher than the concentrations predicted using the AREAPOLY and LINE source types. However, for 8:00 to 9:00 AM, Feb 13 (Figure 61 and Figure 62), RLINE and RLINEXT generate slightly lower results than AREAPOLY/LINE.

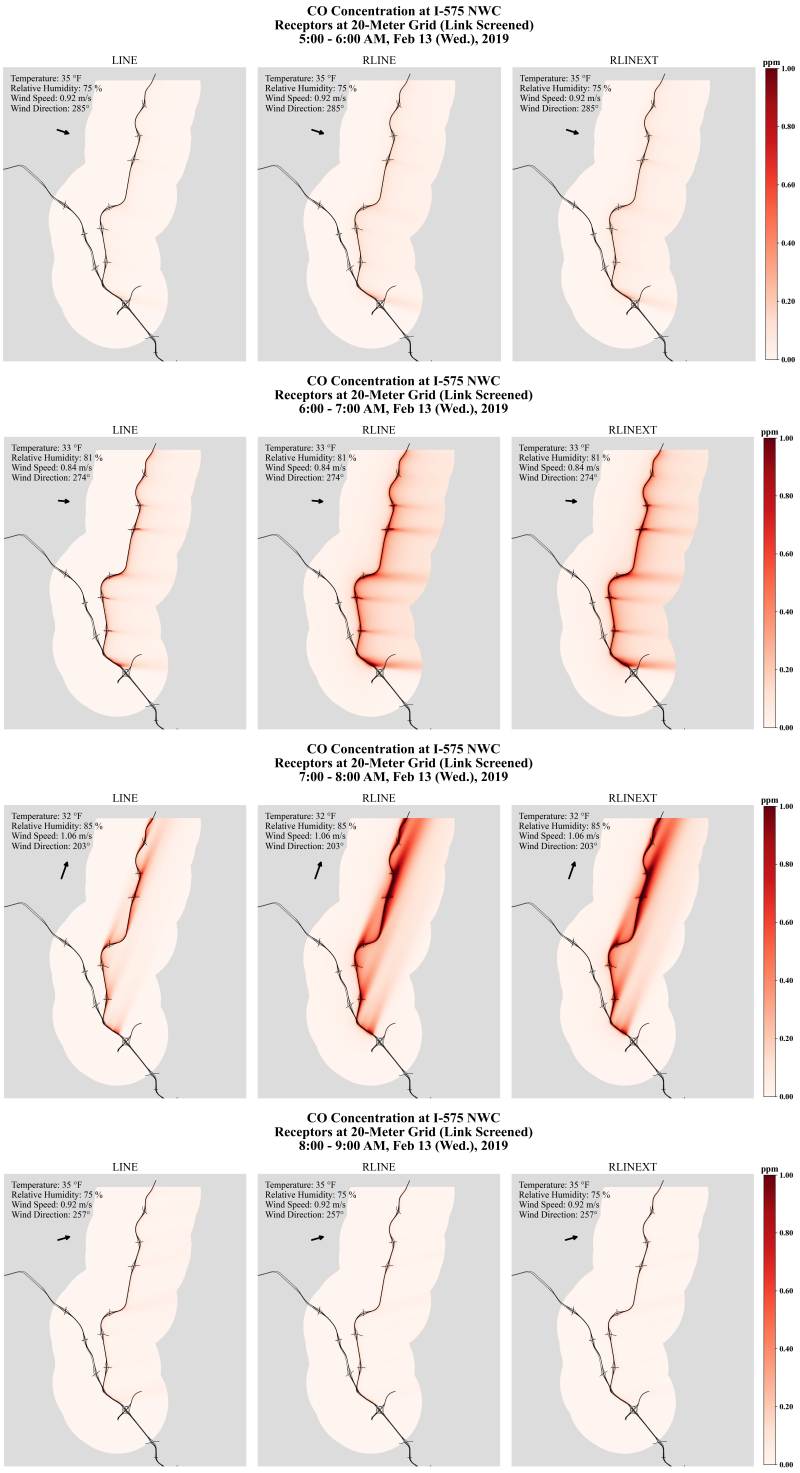


Figure 61 – Predicted CO Concentration at I-575 NWC by Source Type (LINE, RLINE, and RLINEXT no Barrier), Cold Winter Morning Peak

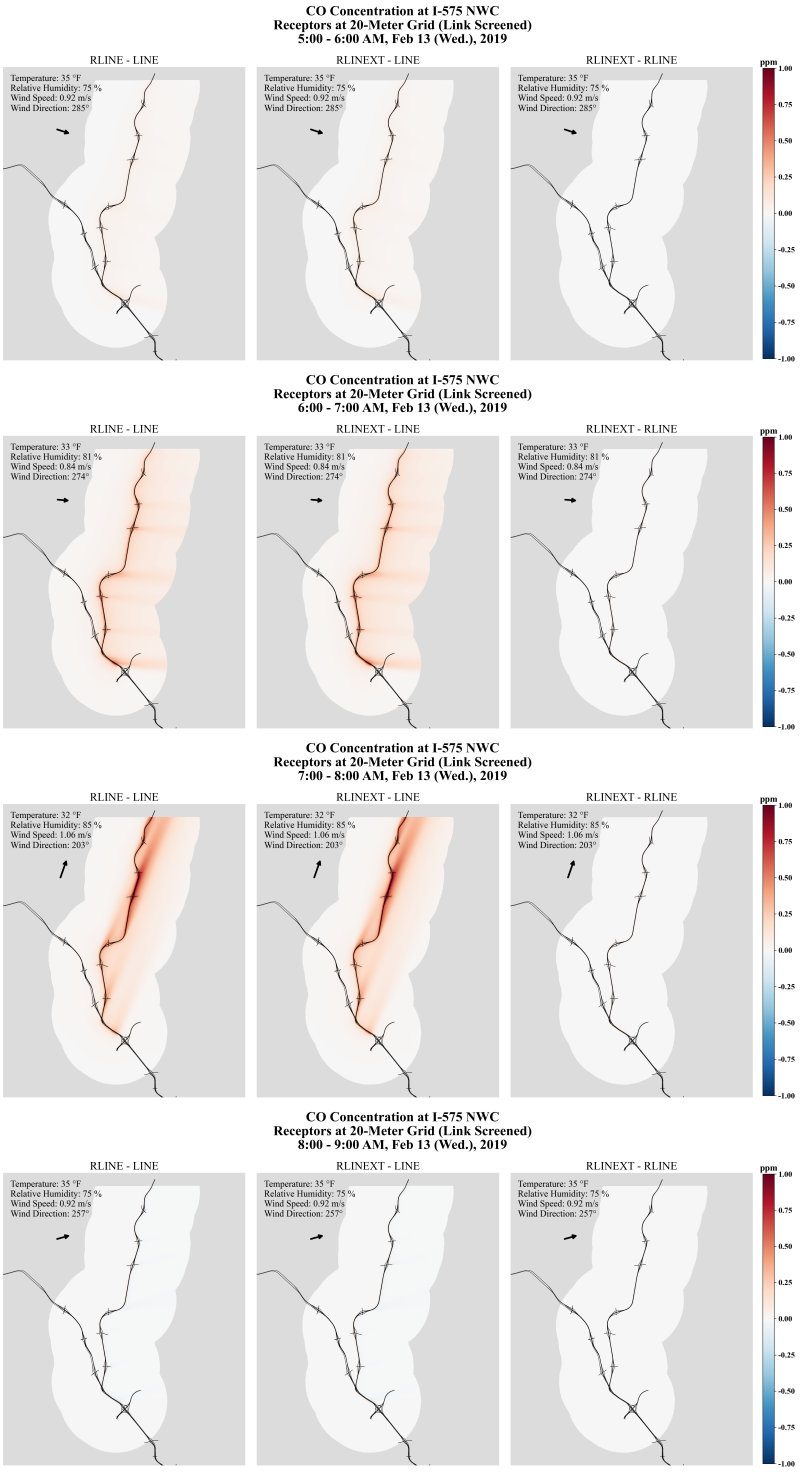


Figure 62 – Differences of Predicted CO Concentration at I-575 NWC by Source Type (LINE, RLINE, and RLINEXT no Barrier), Cold Winter Morning Peak

The team believes that the observed prediction differences between RLINE and RLINEXT are related to differences in the release height parameters used in RLINE vs. RLINEXT model runs. In AERMOD, users can specify the release height of RLINE sources (set to 1.3 meters in this study for all sources). However, users cannot specify release height for RLINEXT sources. A preliminary test found that RLINE and RLINEXT generate the same results when release height is set to 0 for all RLINE sources and DCL is set to 0 for all RLINEXT sources. The team suspects that AERMOD automatically sets the release height for all RLINEXT sources to 0. The user guide could be updated to reflect the differences between RLINE and RLINEXT input parameters.

RLINEXT with and without Noise Barriers

The CO concentration results of RLINEXT with and without noise barriers are presented in Figure 63, and the differences between without vs. with barriers are shown in Figure 64. As expected, differences were observed for RLINEXT with vs. without noise barriers. When the wind direction is perpendicular to the road and barriers, the concentrations at receptors near the barriers from RLINEXT with barriers are much lower than the results from RLINEXT without barriers.

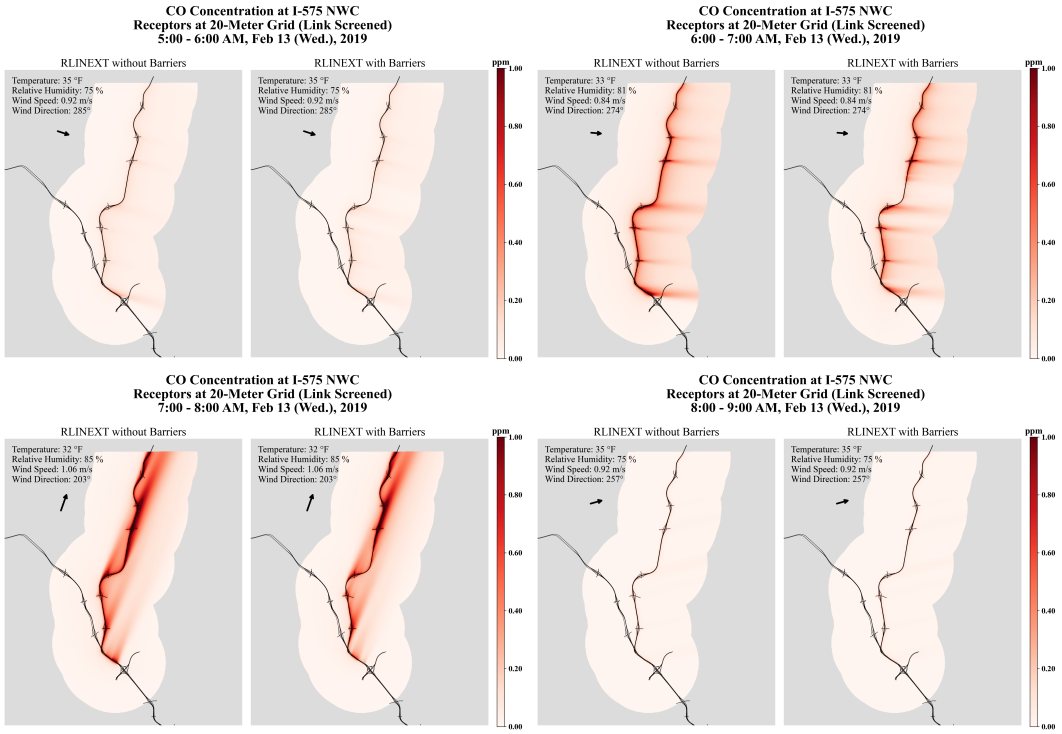


Figure 63 – Predicted CO Concentration at I-575 NWC (RLINEXT with and without Noise Barriers), Cold Winter Morning Peak

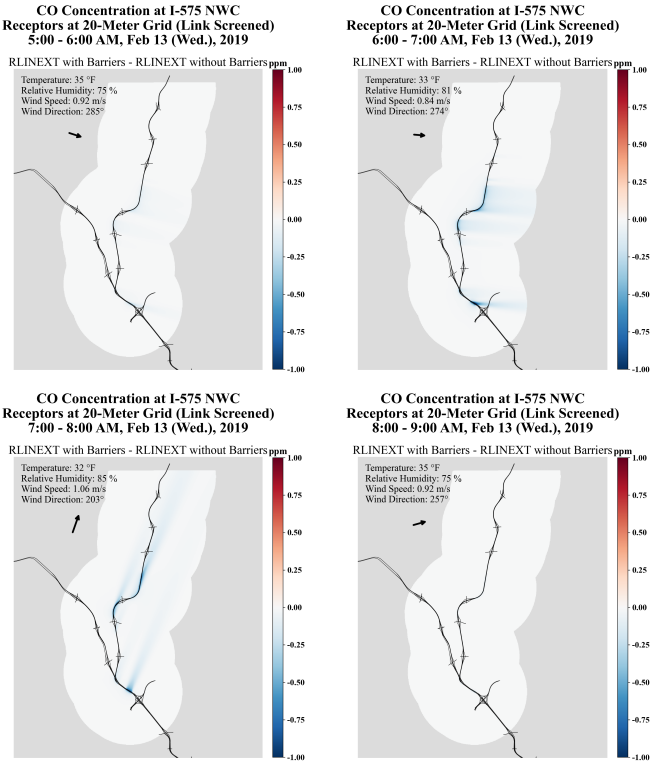
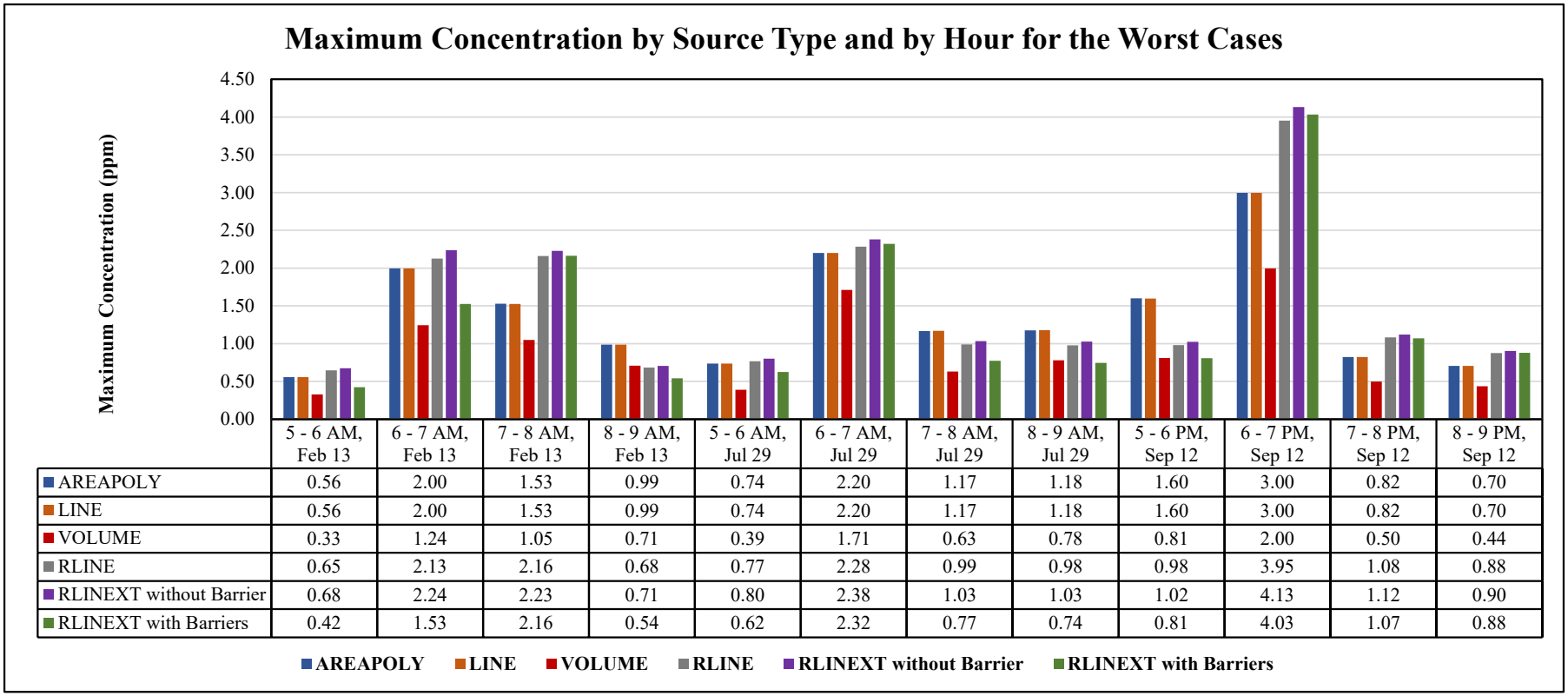


Figure 64 – Differences of Predicted CO Concentration at I-575 NWC (RLINEXT with vs. without Noise Barriers), Cold Winter Morning Peak

Maximum Pollutant Concentrations (AREAPOLY, LINE, VOLUME, RLINE, and RLINEXT)

The maximum concentrations (across link-screened receptors of the I-575 network after removing receptors that fall into the exclusion zones) by hour and by source type are shown in Figure 65. The maximum concentration predicted using the VOLUME source type is always much smaller than predicted using AREAPOLY and LINE (approximately 30% to 50% lower). The RLINE and RLINEXT source types generate larger maximum concentrations than the other source types for most cases. In some cases, AREAPOLY and LINE generate larger maximum concentrations than RLINEXT and RLINE at some of the receptors (although overall results are larger). RLINEXT generates even larger results than RLINE (approximately 5%), most likely due to the release height variable issue discussed earlier. Introducing barriers to RLINEXT leads to much lower maximum concentration at ground level, which is not surprising because the maximum concentration usually occurs near the roadway where the noise barriers are placed.



	5 - 6 AM, Feb 13	6 - 7 AM, Feb 13	7 - 8 AM, Feb 13	8 - 9 AM, Feb 13	5 - 6 AM, Jul 29	6 - 7 AM, Jul 29	7 - 8 AM, Jul 29	8 - 9 AM, Jul 29	5 - 6 PM, Sep 12	6 - 7 PM, Sep 12	7 - 8 PM, Sep 12	8 - 9 PM, Sep 12
AREAPOLY	0.56	2.00	1.53	0.99	0.74	2.20	1.17	1.18	1.60	3.00	0.82	0.70
LINE	0.56	2.00	1.53	0.99	0.74	2.20	1.17	1.18	1.60	3.00	0.82	0.70
VOLUME	0.33	1.24	1.05	0.71	0.39	1.71	0.63	0.78	0.81	2.00	0.50	0.44
RLINE	0.65	2.13	2.16	0.68	0.77	2.28	0.99	0.98	0.98	3.95	1.08	0.88
RLINEXT without Barrier	0.68	2.24	2.23	0.71	0.80	2.38	1.03	1.03	1.02	4.13	1.12	0.90
RLINEXT with Barriers	0.42	1.53	2.16	0.54	0.62	2.32	0.77	0.74	0.81	4.03	1.07	0.88

Figure 65 – Maximum Predicted Concentration by Source Type and by Hour

Dispersion Model Algorithms across Source Types

One set of dispersion modeling routines are used in AERMOD for AREAPOLY and LINE sources and a different set of dispersion modeling routines are used for VOLUME sources. A third set of dispersion routines (based upon R-Line) are employed with RLINE and RLINEXT source types. Hence, AERMOD effectively provides three different dispersion models that can be selected by the modeler for any transportation scenario; one for AREAPOLY and LINE source types, a second for RLINE and RLINEXT, and a third for VOLUME source types.

The research team has not assessed the validity of the three different dispersion modeling approaches currently employed within AERMOD (which was beyond the scope of this project). However, given the lower predicted concentrations associated with the use of the AERMOD VOLUME source type, the research team recommends convening an advisory panel to review all of the dispersion algorithms. It may be that the dispersion algorithms triggered by the VOLUME source type are reasonable and are supported by the scientific literature, but the research team did not find relevant model shoot-out results to confirm the basis for using the enhanced wind meander approach employed with VOLUME source type. Until scientific consensus is reached, the research team recommends continued use of AERMOD with AREAPOLY and LINE source types (or RLINE and RLINEXT source types) in regulatory analysis, because the results are more conservative than the ones obtained from VOLUME sources. It is also worth noting that running AERMOD with the VOLUME source type takes 10 to 20 times longer than model runs for other source types, because the number of VOLUME sources (spheres) used to represent a facility are much larger (one link has to be broken into dozens of spheres, rather than a single polygon or line). Hence, avoiding the VOLUME source type provides a run-time advantage in regional applications.

Sensitivity Analysis for RLINEXT with Respect to Noise Barriers

The sensitivity analyses for barrier height and distance indicate that barriers placed downwind (Figure 66) and upwind (Figure 67) of a source both significantly affect predicted concentrations at downwind receptors. Downwind barriers have a larger impact on these receptors than upwind barriers. The team also suspects that mass flux becomes stacked behind the barrier and then lofts over the barrier as a function of wind speed, reducing predicted concentrations at all downwind ground-level receptors (1.5-meter height). Developing vertical concentration profiles was beyond the scope of this report; however, the research team plans to expand these sensitivity analyses to quantify the impact of noise barrier placement and configuration on vertical concentration profiles as part of an independent research effort for journal article publication.

Figure 66 – Predicted CO Concentration by Barrier Height and Distance to Roadway (Barrier Downwind of the Source)

Figure 67 – Predicted CO Concentration by Barrier Height and Distance to Roadway (Barrier Upwind of the Source)

The barrier sensitivity analyses also reveal that predicted concentrations at downwind receptors decrease as barrier height increases, and that the barrier influence diminishes as the distance between barrier and source increases (see Figure 68), which was anticipated by the research team based upon theory. Distant barriers with low height (e.g., upwind barriers that are 50 meters away and 4 meters high) have no significant impact on predictions at the reference receptor. However, given the real-world variability in barrier height distributions (e.g., 84.3% of the barriers in this research were between 4 meters and 10 meters in height), and the fact that most barriers are

immediately adjacent to the road (e.g., within 50 meters), barriers do influence model output. The team recommends that once the dual-barrier algorithms in the latest version of AERMOD are verified, analyses include both barriers for each link, rather than choosing one barrier per link.

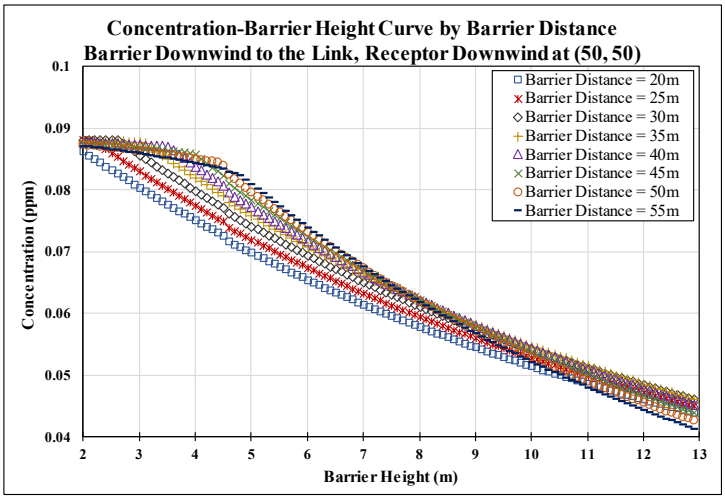


Figure 68 – Sensitivity of Predicted CO Concentration to Barrier Height and Distance to Roadways

Various link splitting strategies do not yield significantly different predicted concentrations when barriers are not present (Figure 69) and when continuous barriers (identified in figures as "fully-aligned" barriers) are present along all roadway links (Figure 70). However, the impact of link splitting can be significant when coupled with partial barrier alignments, where barriers are tied to only some of the roadway sections, as shown in Figure 71 (no source split), Figure 72 (uneven source split), and Figure 73 (even source split). Even when a small portion of the source is aligned with a noise barrier (e.g., one eighth, or 12.5%, as shown in Figure 72), the impact can be large for receptors near the source, and these impacts have to be considered jointly with variability in wind direction. The impact of the relative position between barrier and link varies across downwind receptors because the wind direction varies by hour.

The minimum predicted concentration at any receptor is 0.000001 $\mu g/m^3$ (non-zero). Relative differences can be obtained and are presented in Figure 74 (no barrier) and Figure 75 (fully aligned barriers). The sensitivity analysis for source-barrier setups indicates that the impact of splitting the sources is not significant with respect to concentration profiles, but these impacts can be as large as approximately 5% in relative difference (seen as striated patterns of positive and negative difference). It is worth noting that the differences near the roadway in Figure 75 occur between the barrier and the source link. Although these differences are significant, they are smaller compared with the differences caused by mismatched barrier alignment. However, if a receptor of interest is near the modeled network and between the roadway and the noise barriers, the users might want to evaluate the impact of link splitting when preparing RLINEXT source type input files.

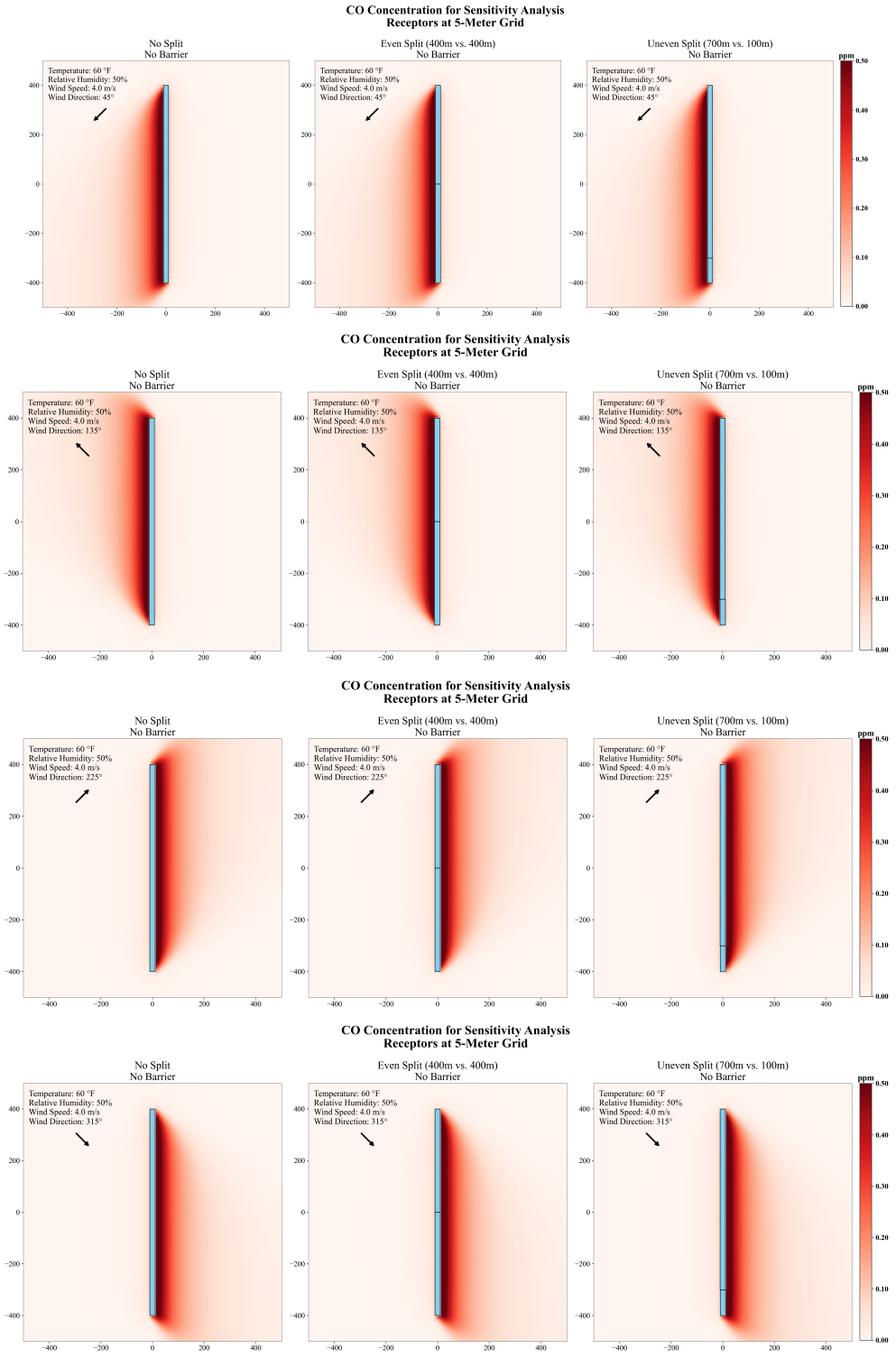


Figure 69 – Predicted RLINEXT CO Concentrations without Barrier by Split Type (No Split, Even Split, and Uneven Split)

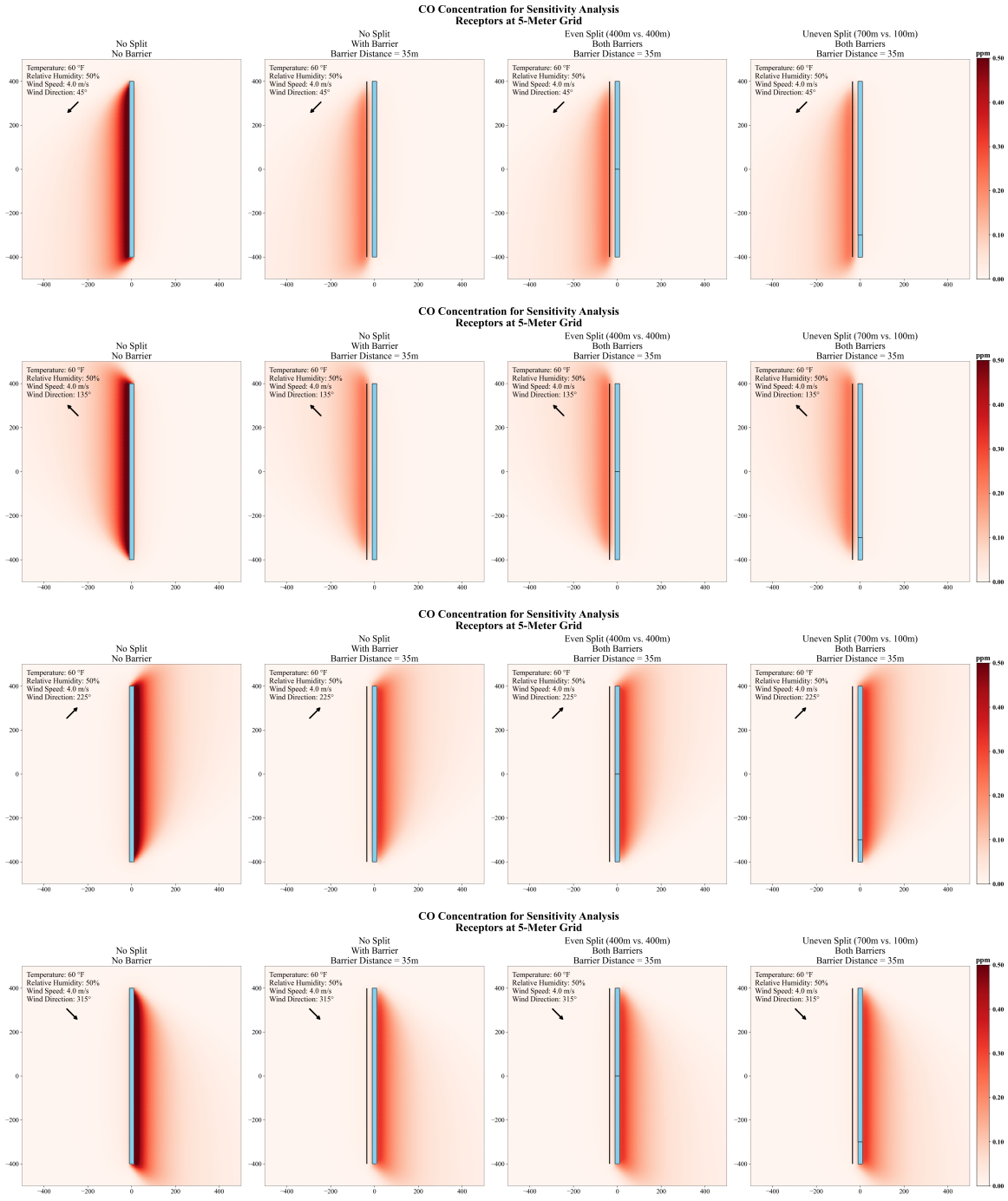


Figure 70 – Predicted RLINEXT CO Concentrations with Fully Aligned Barriers by Split Type (No Split, Even Split, and Uneven Split)

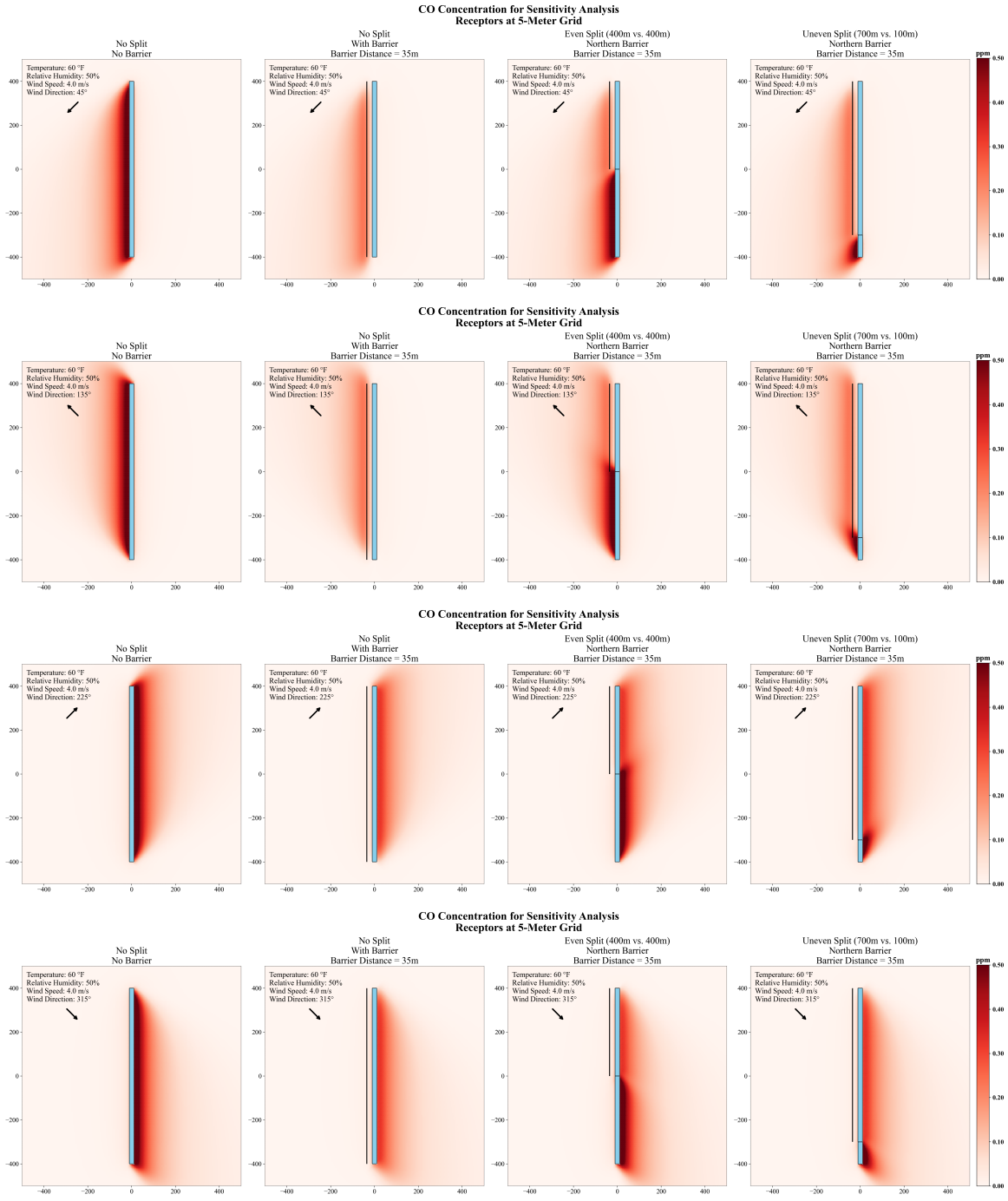


Figure 71 – Predicted RLINEXT CO Concentrations with Partially Aligned Barriers by Split Type (No Split, Even Split, and Uneven Split)

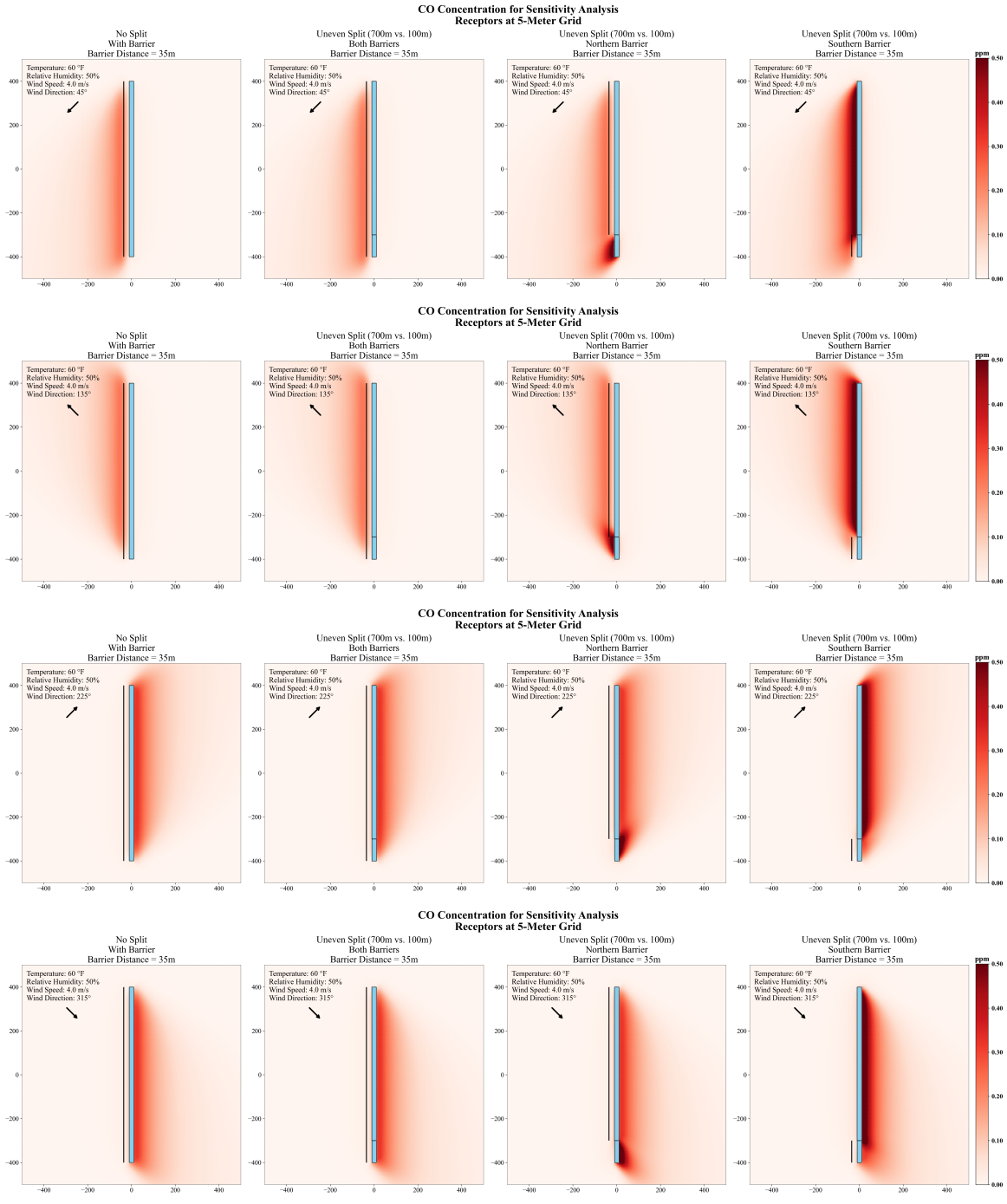


Figure 72 – Predicted RLINEXT CO Concentrations for Unevenly Split Links by Barrier Alignment (Fully Aligned, 87.5% Aligned, and 12.5% Aligned)

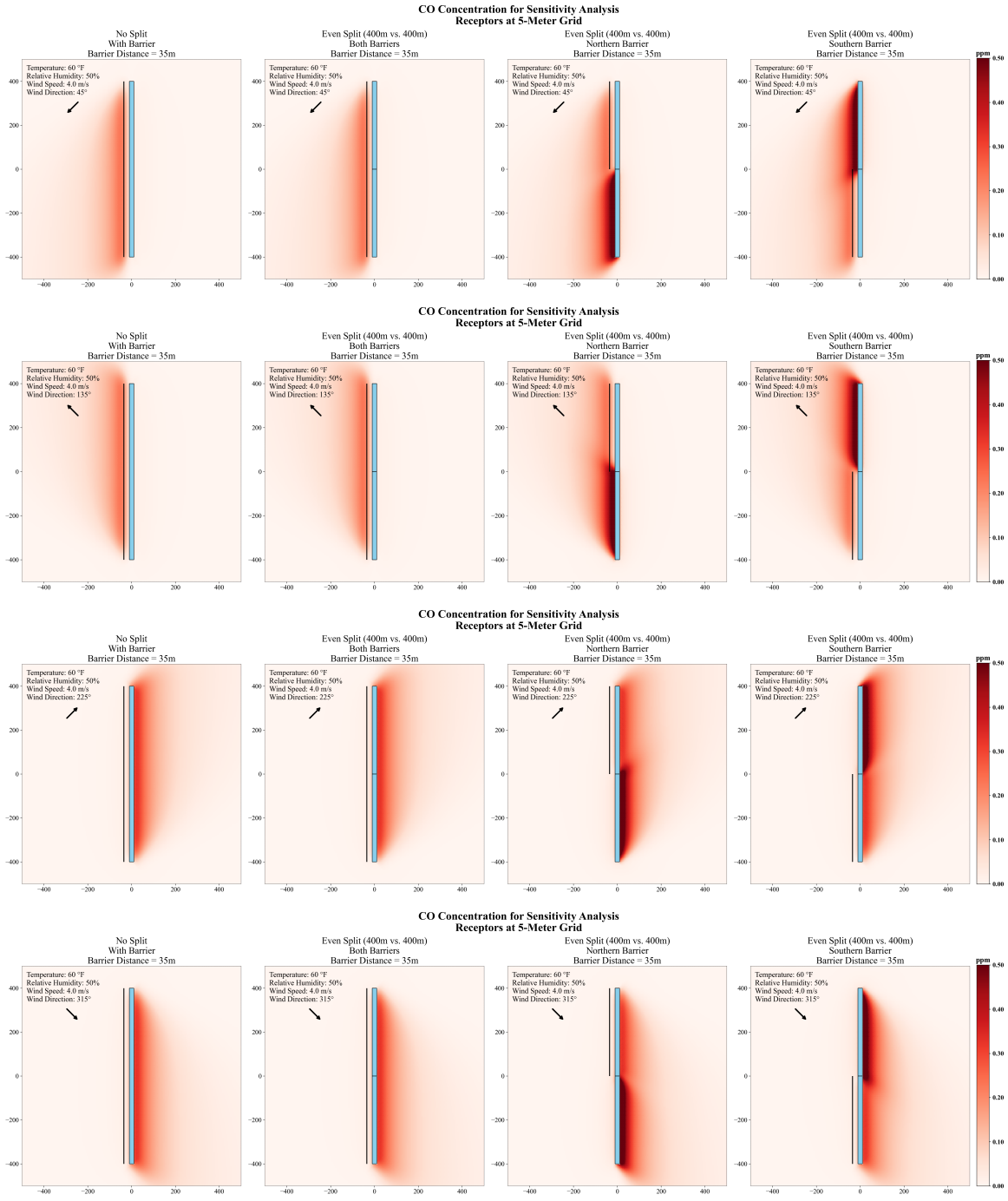


Figure 73 – Predicted RLINEXT CO Concentrations for Evenly Split Links by Barrier Alignment (Fully Aligned, Northern Half Aligned, Southern Half Aligned)

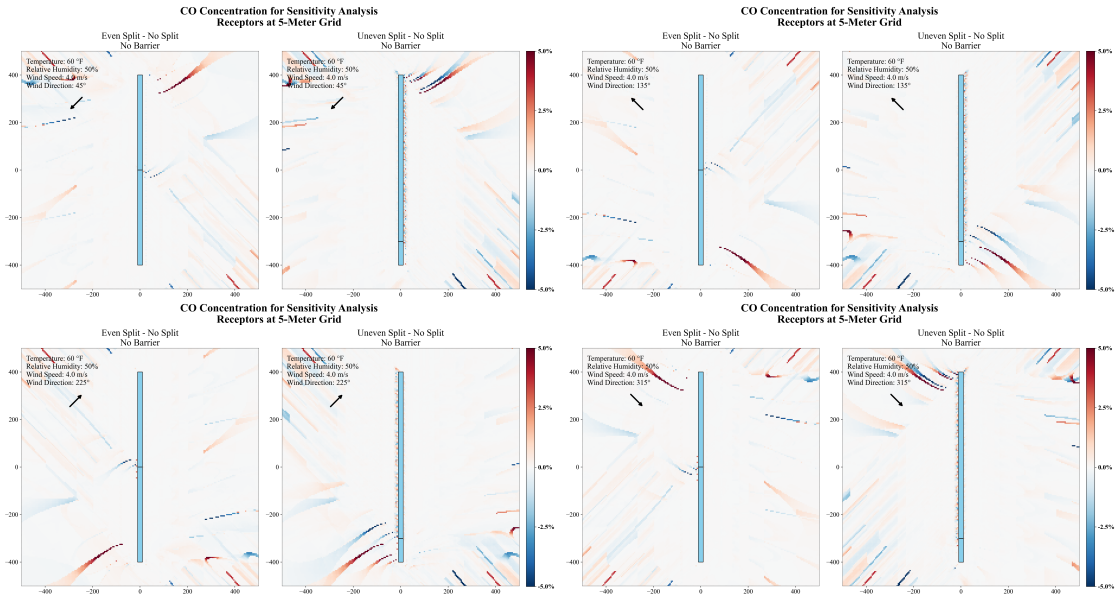


Figure 74 – Relative Differences in RLINEXT Predicted CO Concentrations, without Barrier by Split Type (No Split vs. Even Split vs. Uneven Split)

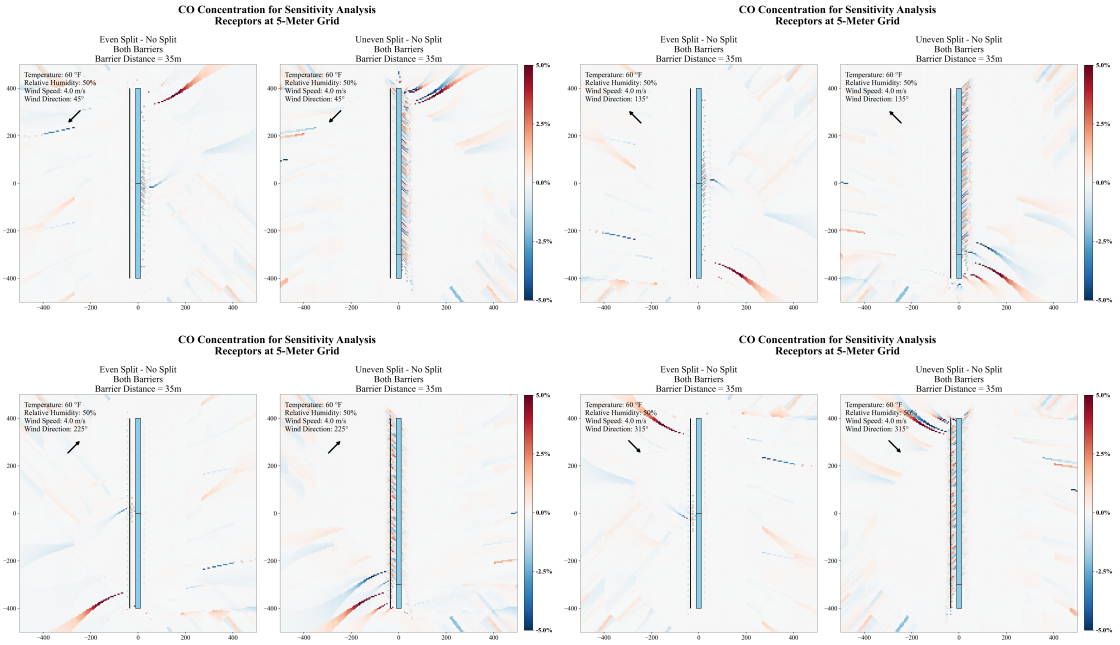


Figure 75 – Relative Differences in RLINEXT Predicted CO Concentrations with Fully Aligned Barrier(s) by Split Type (No Split vs. Even Split vs. Uneven Split)

The results for link split type coupled with various barrier distances are shown in Figure 76 (uneven split) and Figure 77 (even split). The figures reveal that besides barrier alignment, variability in barrier distance from the roadway (which occurs frequently in the real world) also affects relative model output, and these impacts vary across wind directions.

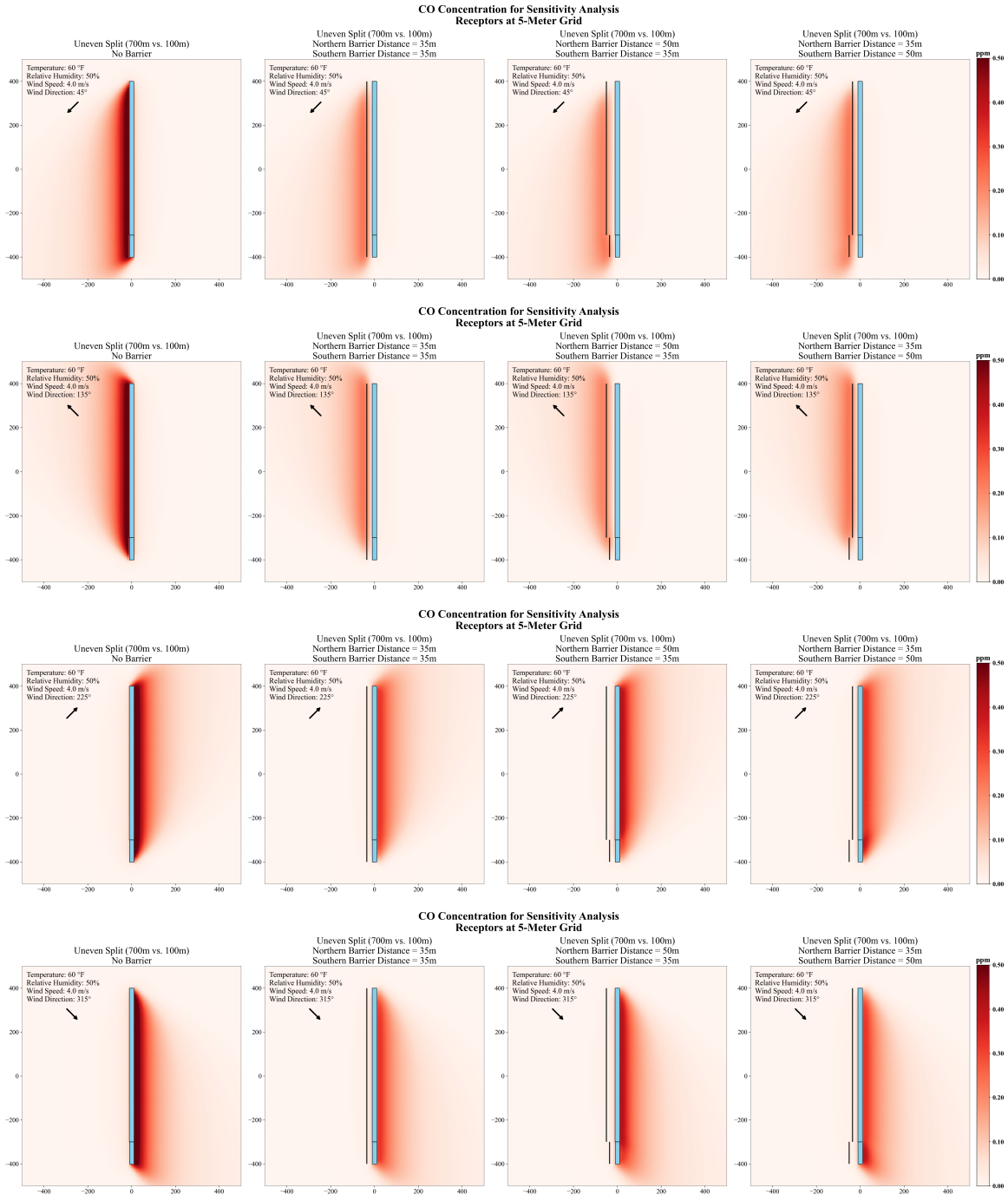


Figure 76 – Predicted RLINEXT CO Concentrations for Unevenly Split Links by Barrier Alignment (Fully Aligned, Northern Half Aligned, and Southern Half Aligned)

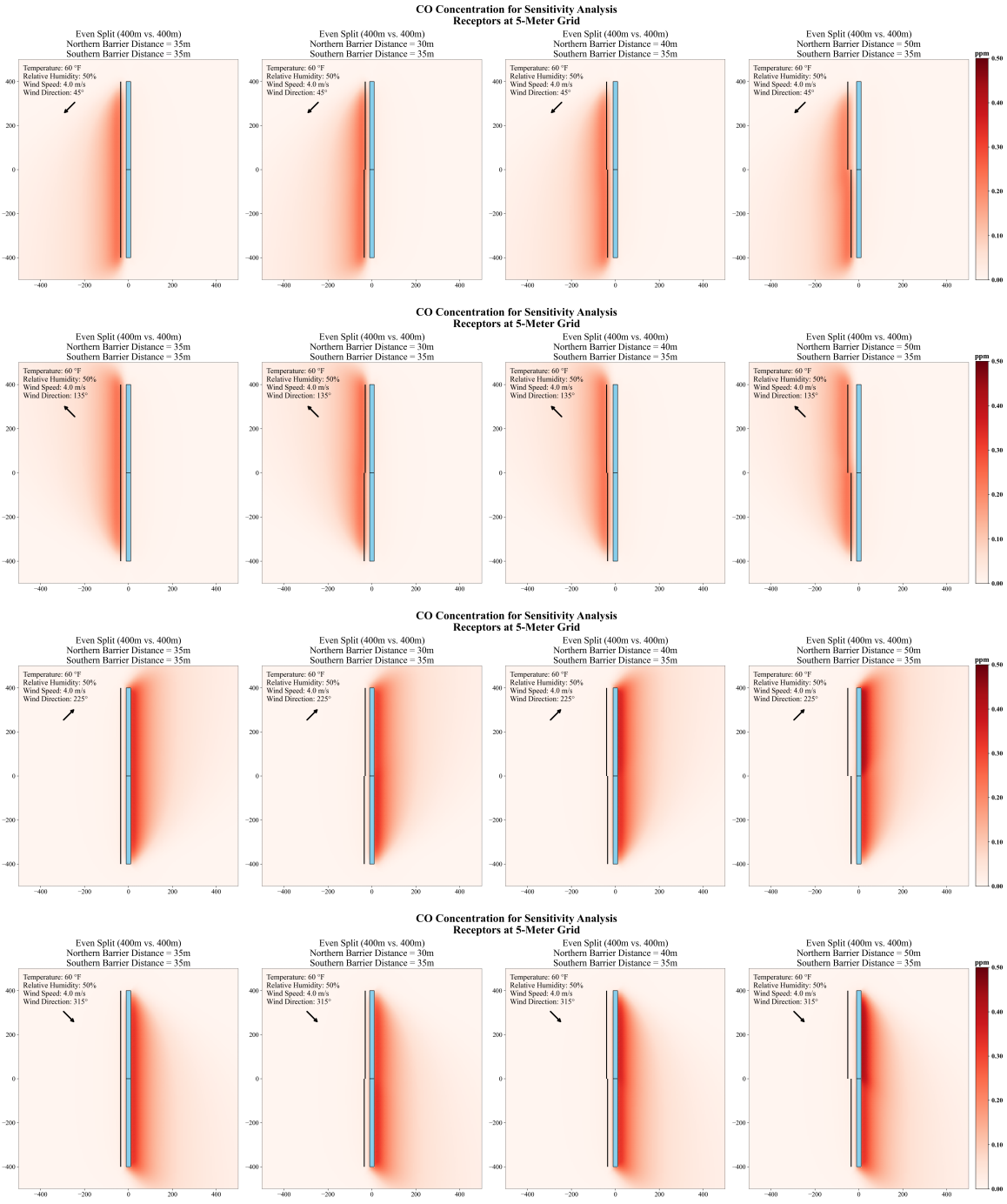


Figure 77 – Predicted RLINEXT CO Concentrations for Evenly Split Links by Barrier Alignment (Fully Aligned, Northern Half Aligned, Southern Half Aligned)

Barrier Impacts for Receptors Located Over the Roadway

The team recognizes that receptors are not placed over the roadway for regulatory analysis. However, assessing what happens within the model over the roadway remained of interest, given that different dispersion algorithms are employed to model barrier effects. The team conducted a sensitivity analysis for receptors placed over the roadway in a 1-meter grid, and the results are shown in Figure 78 and Figure 79. The relative differences across link splits are presented in Figure 80 and Figure 81. The RLINEXT results are not evenly distributed over the roadway, and the differences near the link border can be as large as 5% (depending on wind direction). A division of differences (a parallel line near the median of the link) is observed when wind blows perpendicular to the source, and the patterns of these differences indicate the possibility of resultant bias that could be further examined if a future review panel is convened to assess AERMOD dispersion algorithms.

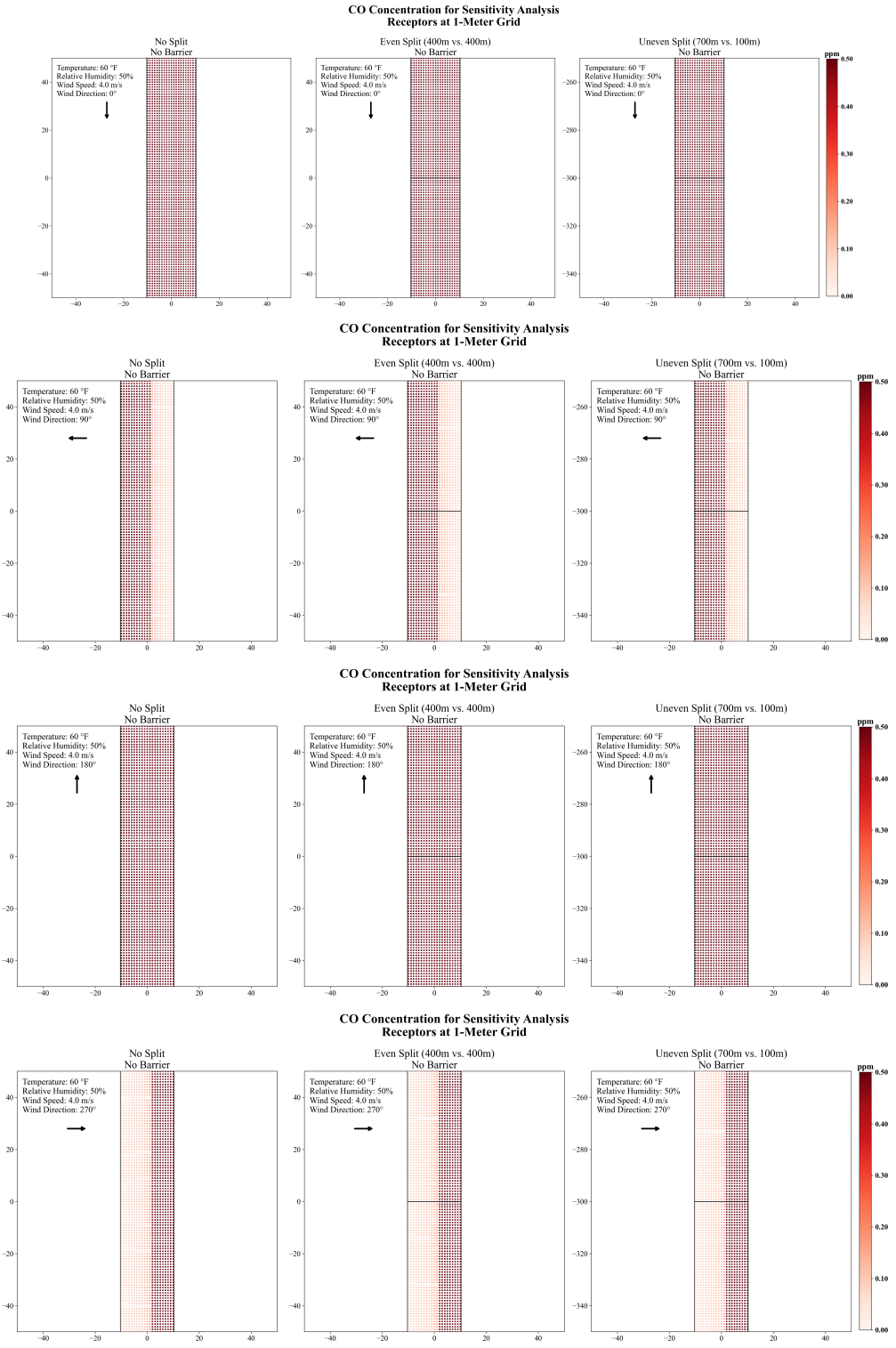


Figure 78 – Predicted RLINEXT CO Concentration Over-the-Roadway without Barrier by Split Type (No Split, Even Split, and Uneven Split)

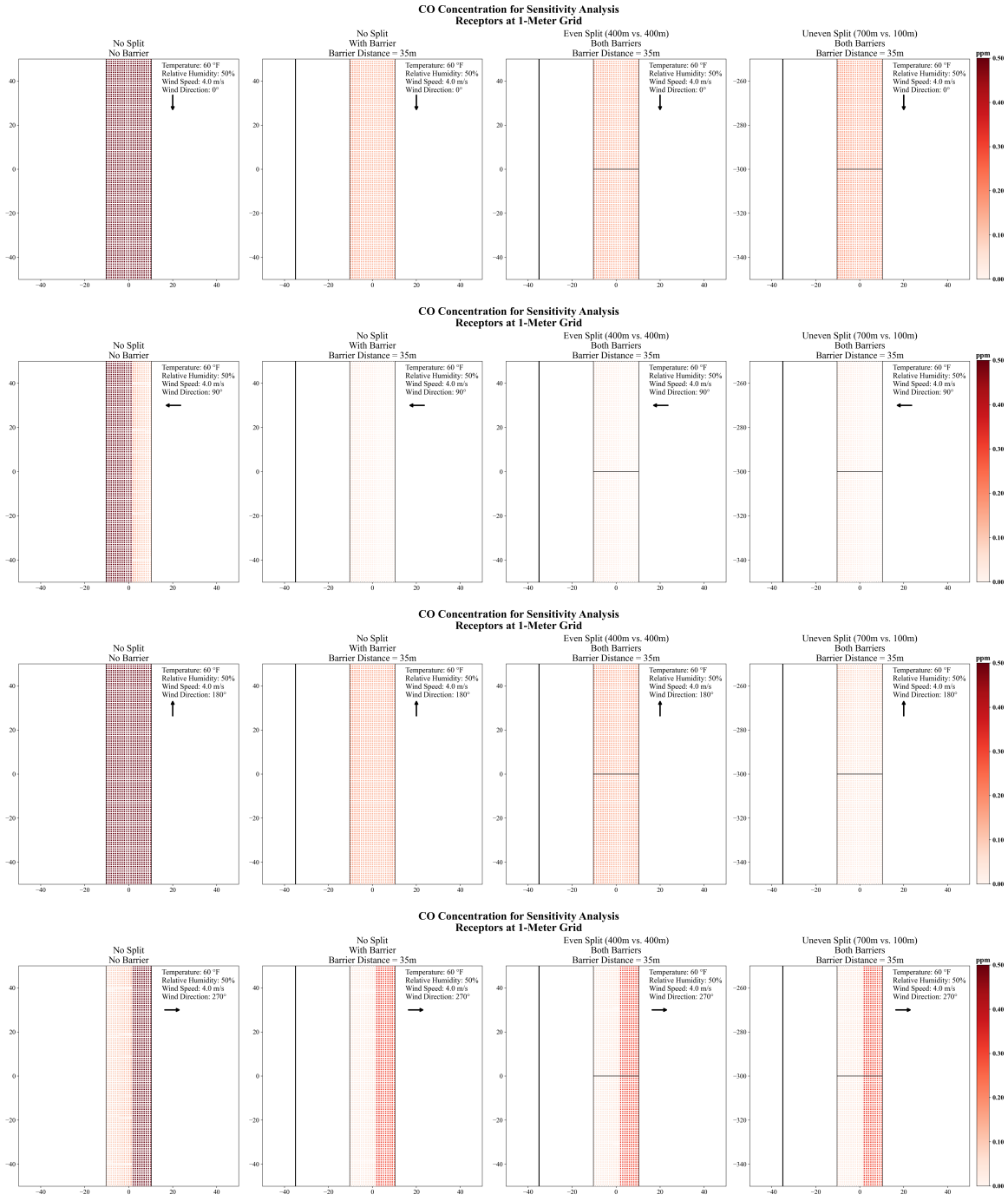


Figure 79 – Predicted RLINEXT CO Concentration Over-the-Roadway with Fully Aligned Barrier by Split Type (No Split, Even Split, and Uneven Split)

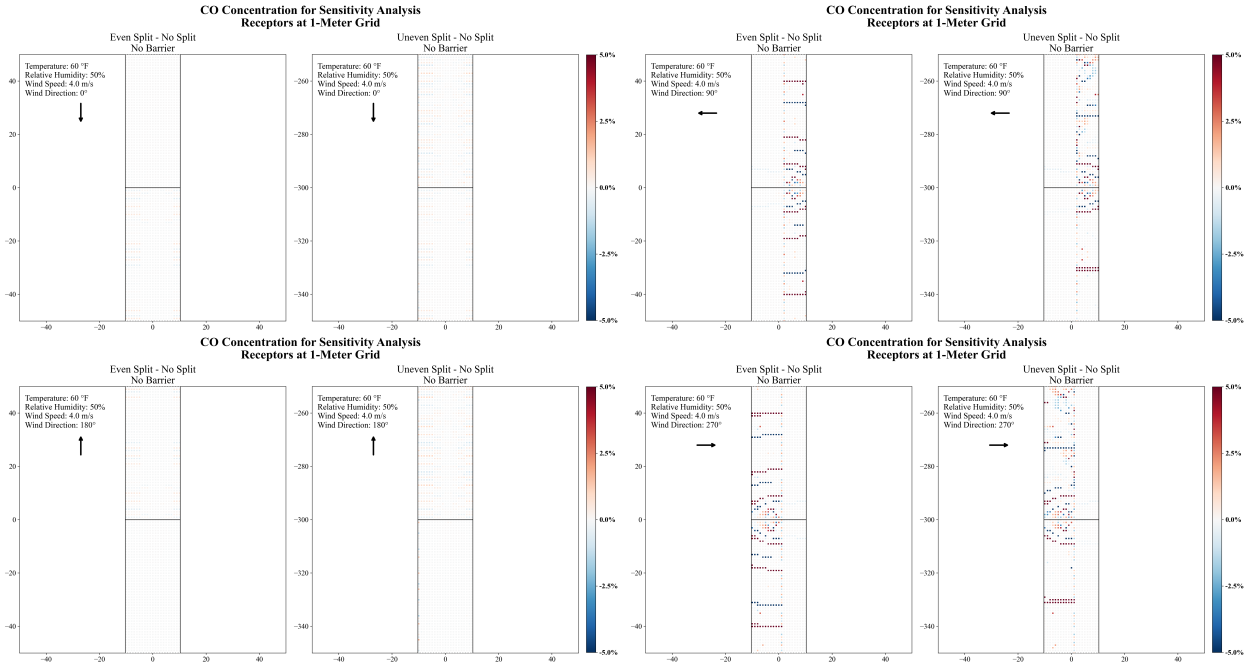


Figure 80 – Relative Differences in Predicted CO Concentration Over the Roadway without Barriers by Split Type (No Split, Even Split, and Uneven Split) for RLINEXT

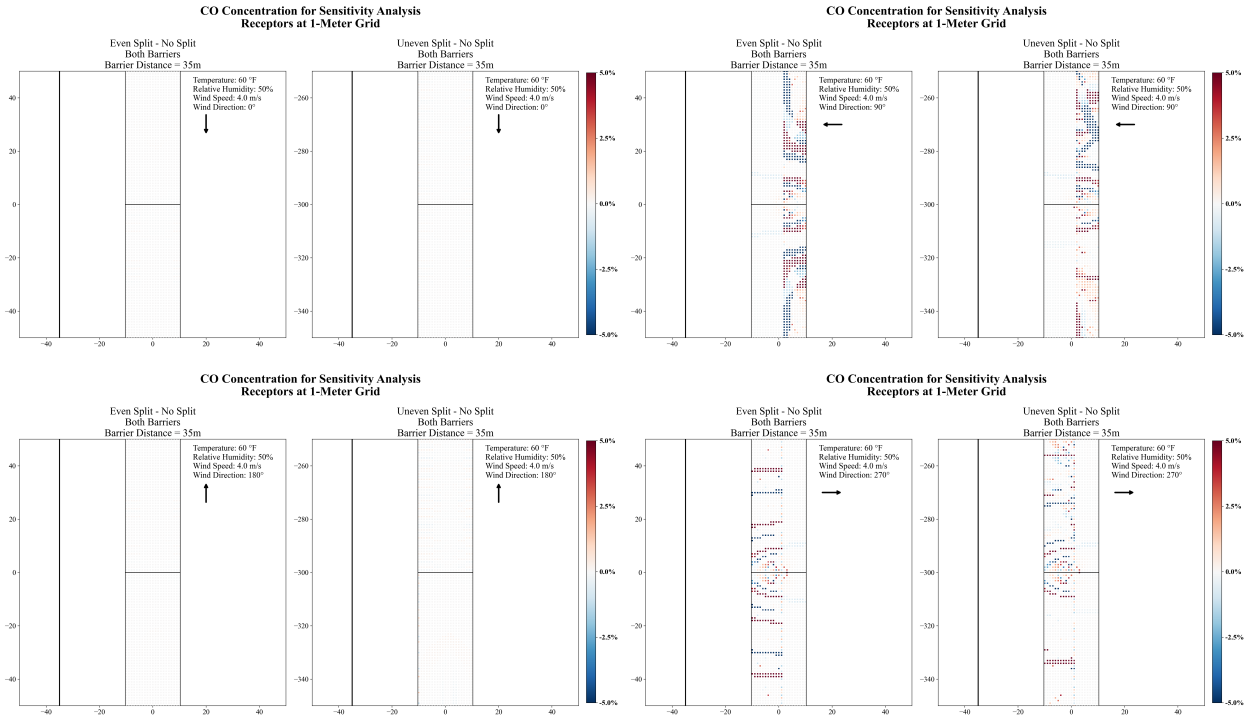


Figure 81 – Relative Differences in Predicted CO Concentration Over the Roadway with Barrier(s) by Split Type (No Split, Even Split, and Uneven Split) for RLINEXT

CHAPTER 5. CONCLUSIONS

This study performed AERMOD microscale dispersion modeling for an entire Atlanta I-75/I-575 Northwest Corridor (NWC) subarea, including freeway corridors, managed lanes, connecting arterials, and intersections serving the NWC system. This sub-regional modeling effort allowed for the investigation of the relative impacts of modeling parameters on model outputs, with special attention given to the impact of noise barriers.

The significance of many of the parameters was tested using machine learning. For single source type models, the team concluded that using clustered receptor data by distance to roadways helps improve the model fit, and results found that near-road receptors and far-from-road receptors are both influenced by wind speed, receptor ID (which accounts for adjacent roads and their and their mass flux emission rates in grams/meter^2/second), and wind direction. The largest differences across source types were noted for near-roadway receptors located near intersections. However, the source types were found to have very little influence on the concentration results compared to the factors that relate to wind parameters and mass flux from links upwind of the receptors.

The comparison of concentration profiles resulting from various receptor grid resolutions (20-meter grid plus variable grid vs. 5-meter grid, within maximum distances of 200 meters to the roadway) indicated that 20-meter standard grid combined with variable grid is sufficient to provide output detail, while maintaining reasonable performance for comparative analyses across AERMOD source types. For AREAPOLY sources, concentration results from automatically generated polygons are essentially the same as those from manually created polygons. The automatic polygon generation process was demonstrated to be efficient and feasible and does not negatively affect AERMOD outputs (in fact, the process reduces human error potential).

The comparative results across source types indicated that concentrations from RLINEXT and RLINE (both use the model of R-LINE) were very similar (the concentration fields are not exactly the same, but the differences were very small), and their results were higher than other source types. The differences between RLINEXT vs. RLINE may relate to the release height input variable in RLINE, which cannot be manually specified in RLINEXT. The results from AREAPOLY sources (both automatically generated and manually created) were nearly identical to those from LINE sources, and both are lower than RLINE and RLINEXT results in most cases. At low wind speed, the VOLUME source type yields significantly lower pollutant concentrations downwind of the source and higher pollutant concentrations upwind of the source than all of the other source types. The change in the concentration field is the result of the implementation of a wind meander effect in dispersion algorithms invoked by AERMOD only for the VOLUME source type.

Given the lower predicted downwind concentrations associated with the use of the VOLUME source type, the research team recommends a thorough review of the embedded dispersion algorithms. It may be that the dispersion algorithms triggered by the VOLUME source type are reasonable and can be supported by the literature, but the research team did not find relevant model shoot-out results to confirm the basis for using the enhanced wind meander approach employed with VOLUME source type. Until a thorough scientific review is conducted for all AERMOD dispersion algorithms and scientific consensus is reached, the research team recommends continued use of AERMOD with AREAPOLY and LINE source types in regulatory analysis, given that the results are more conservative at low wind speeds and less likely to miss any potential exceedance of the NAAQS. Once scientific consensus is reached, the research team

recommends that AERMOD be updated to employ one set of dispersion algorithms across all source types, or that specific guidance be published as to which algorithms should be applied under specific circumstances, so that model applications yield consistent results for any given transportation scenario.

Accounting for noise barriers (one barrier maximum) in RLINEXT leads to lower resultant downwind concentrations, compared with results from no-barrier runs, which was anticipated. Sensitivity analyses with respect to RLINEXT with noise barriers indicate that barriers with greater distance separation from the roadway yield smaller impacts on receptors downwind of the roadway, and lower height barriers yield smaller impacts on downwind receptors (for receptor heights at 1.5 meters). The results also indicate that barriers on both sides of the roadway impact predicted results when tested separately, with downwind barriers yielding a larger impact. Barriers partially aligned with roadways (even if this represents only a small fraction) yield significantly different predictions compared barriers that stretch the entire length of roadways (i.e., barrier edge effects are noted). Splitting the RLINEXT sources to match barrier alignment can affect predicted concentrations by approximately 5% for some receptors.

Modeling noise barriers using the RLINEXT source type is challenging. The current version of AERMOD (version 19191) requires that each barrier be matched with only one link (the next version of AERMOD allows the use of both an upwind and downwind barrier). However, properly splitting the input sources to align with noise barriers is important, because the splits can change model outputs. The assumption that each barrier is perfectly parallel to the attaching source, and the restriction that the attached barrier only impacts the source, makes it difficult to model complicated real-world barrier systems.

All of the analyses presented in this report employed receptor grids with a standard 1.5-meter vertical height for all receptors. Given the observed influence of barriers on horizontal concentration profiles presented in this report, team is currently expanding the barrier sensitivity analyses to bring in vertical profiles using a 3-dimensional grid with receptors located in vertical layers at 1.5, 3.0, 4.5, 6.0, 12.0, and 24.0 meters. The team anticipates completing this follow-up research and submitting a journal manuscript for publication in December 2021.

The considerable library of model run results generated for this project are being archived by the Federal Highway Administration and will be available upon request. The model input files and outputs contain a diverse set of input combinations, including various source types, link network geometry, receptor placement, and noise barrier combinations. Rather than creating entirely new AERMOD runs, researchers can use this archive as a foundation for creating new AERMOD runs, or can use the modeling framework designed for this study (file/code structure) to perform similar analyses. The output dataset also facilitates a variety of further applications, such as modeling chain of environmental impact of the transportation system (i.e., from vehicle activity to human exposure), equity assessment with respect to traffic-related pollution, sensitivity and uncertainty analysis of dispersion modeling, etc.

APPENDIX A: COMPARISON OF RECEPTOR RESOLUTION

This appendix presents the concentration results predicted at receptors of 5-meter standard grid vs. 20-meter standard grid combined with variable grids.

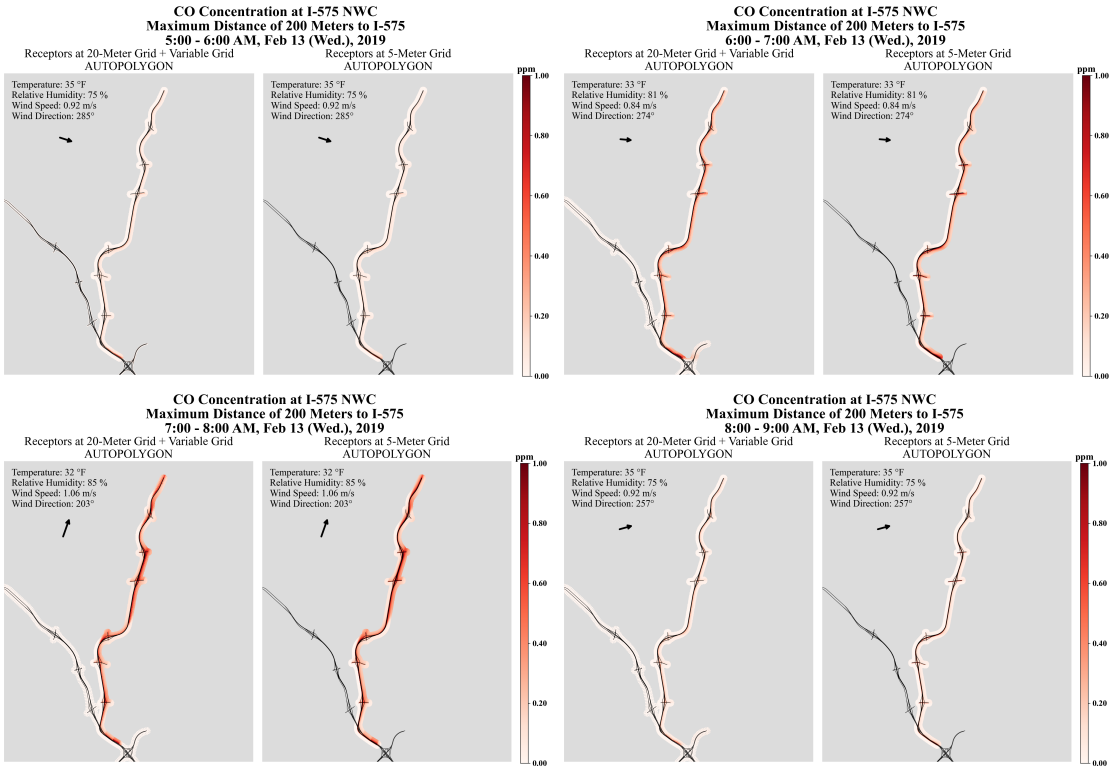


Figure 82 – Predicted AREAPOLY CO Concentrations by Receptor Placement (5-Meter Grid vs. 20-Meter Grid Combined with Variable Grid)

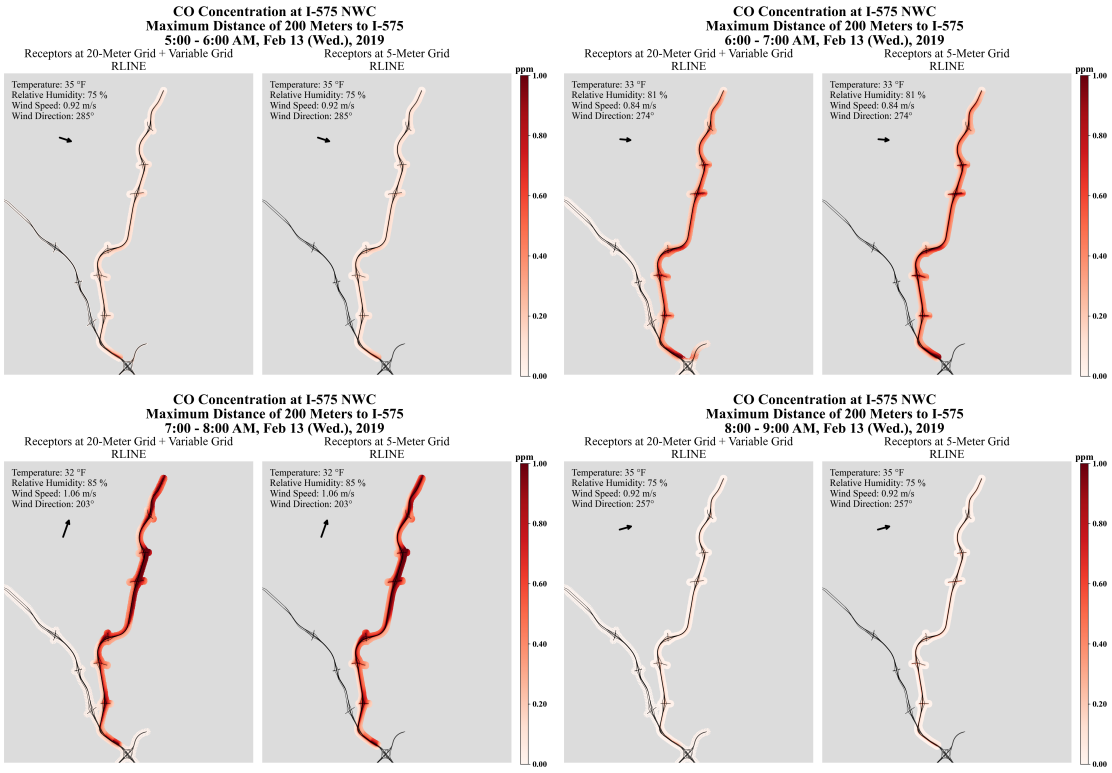


Figure 83 – Predicted LINE CO Concentrations by Receptor Placement (5-Meter Grid vs. 20-Meter Grid Combined with Variable Grid)

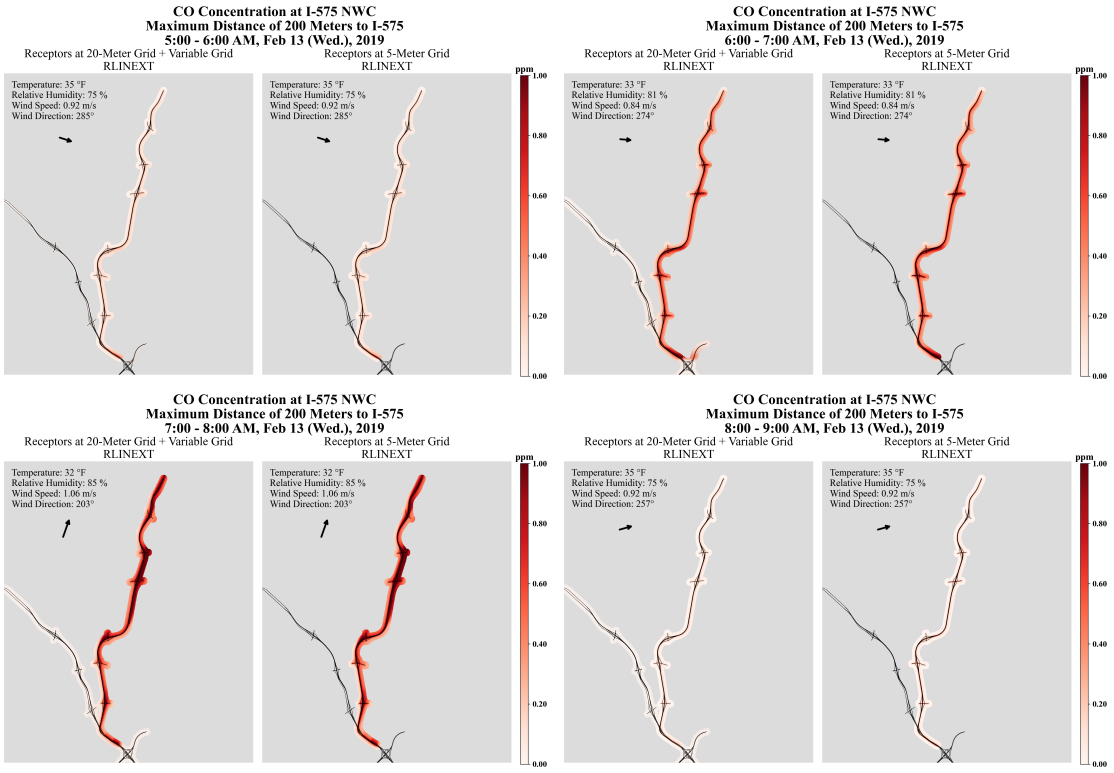


Figure 84 – Predicted LINE CO Concentrations by Receptor Placement (5-Meter Grid vs. 20-Meter Grid Combined with Variable Grid)

APPENDIX B: SAMPLE WORST CASES ACROSS SOURCE TYPES OF COMPARATIVE RESULTS

This appendix presents two sample worst-case heat map visualizations of 20-meter grid receptors for the I-575 network after link screening for predicted concentration profiles across source types. As with the previous analyses, receptors over the roadway and within five meters of the roadway have been excluded (USEPA, 2015). The heat maps for the hot summer morning peak (July 29, 2019) are presented in Figure 85 through Figure 90, and the heat maps of the hot fall evening peak are presented in Figure 91 through Figure 96.

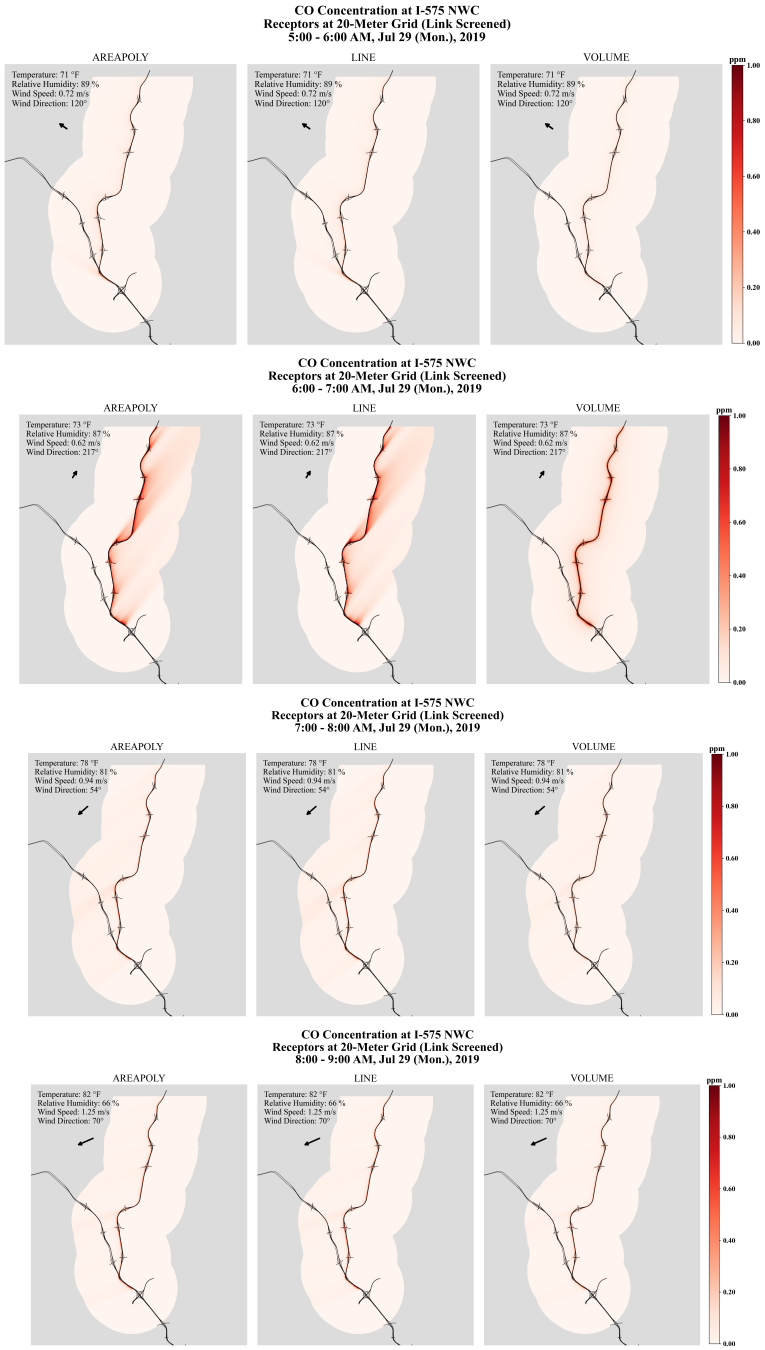


Figure 85 – Predicted CO Concentrations by Source Type (AREAPOLY, LINE, and VOLUME), Hot Summer Morning Peak

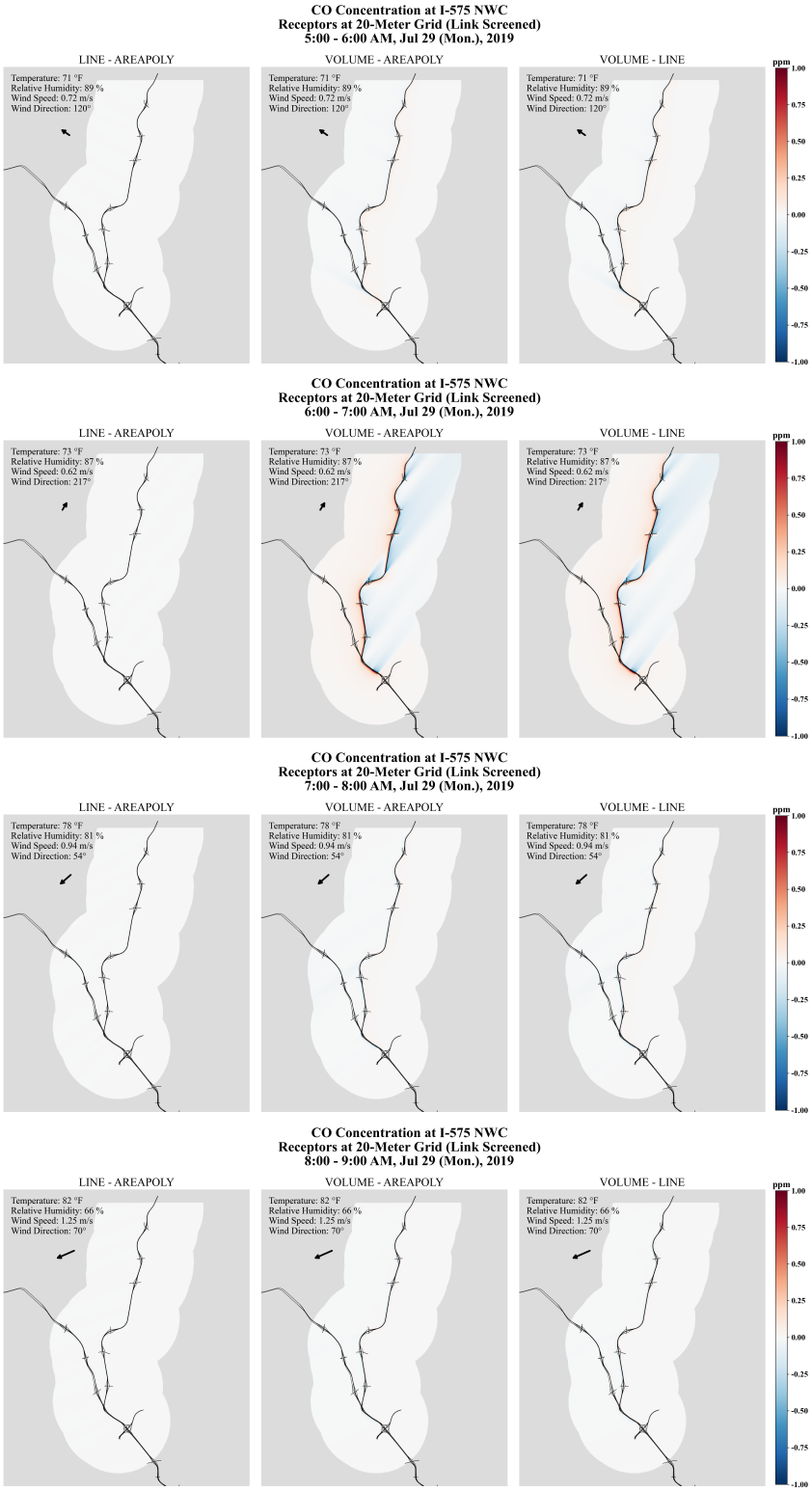


Figure 86 – Differences in Predicted CO Concentrations by Source Type (AREAPOLY vs. LINE vs. VOLUME), Hot Summer Morning Peak

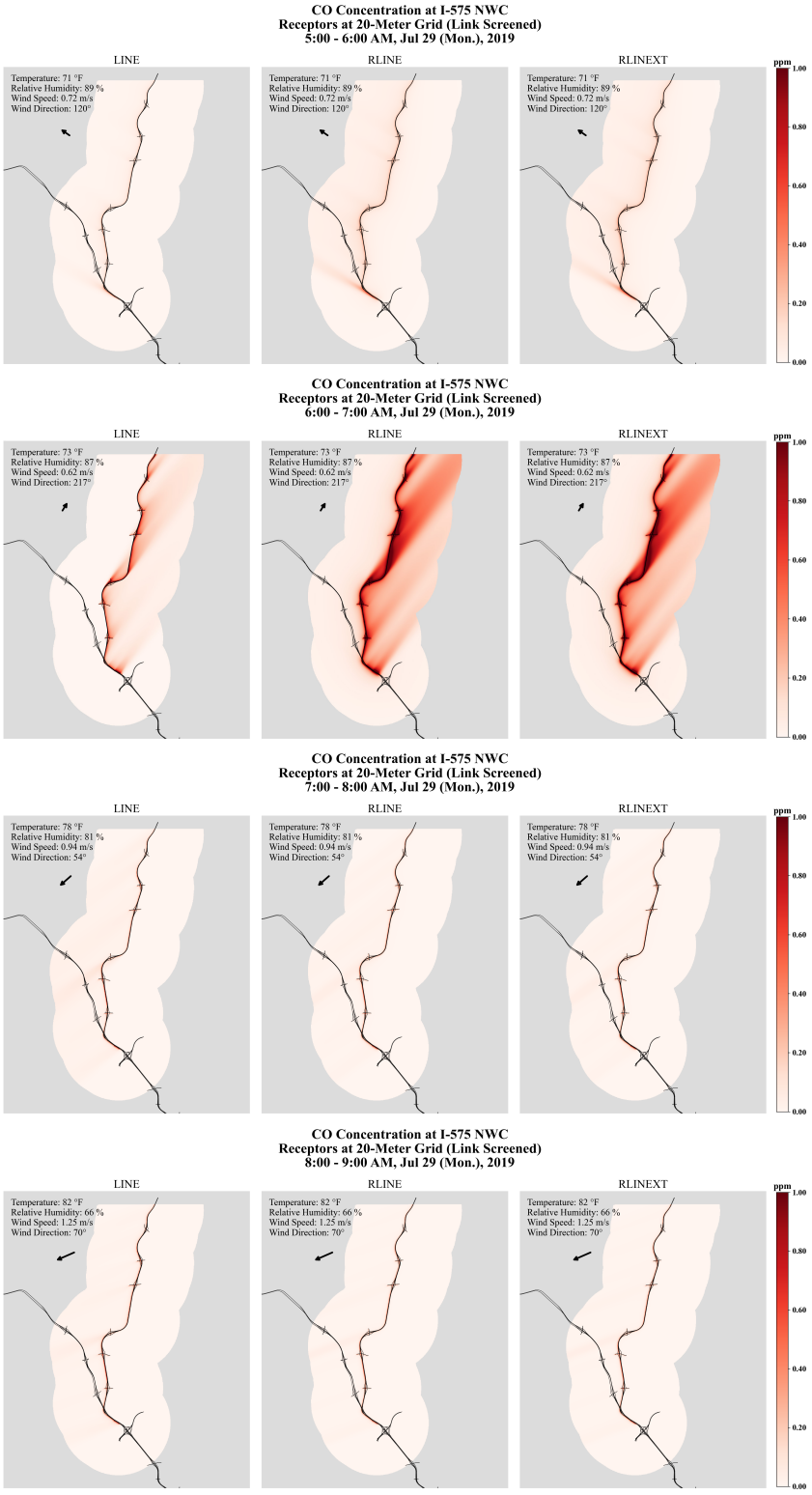


Figure 87 – Predicted CO Concentrations by Source Type (LINE, RLINE, and RLINEXT without Barrier), Hot Summer Morning Peak

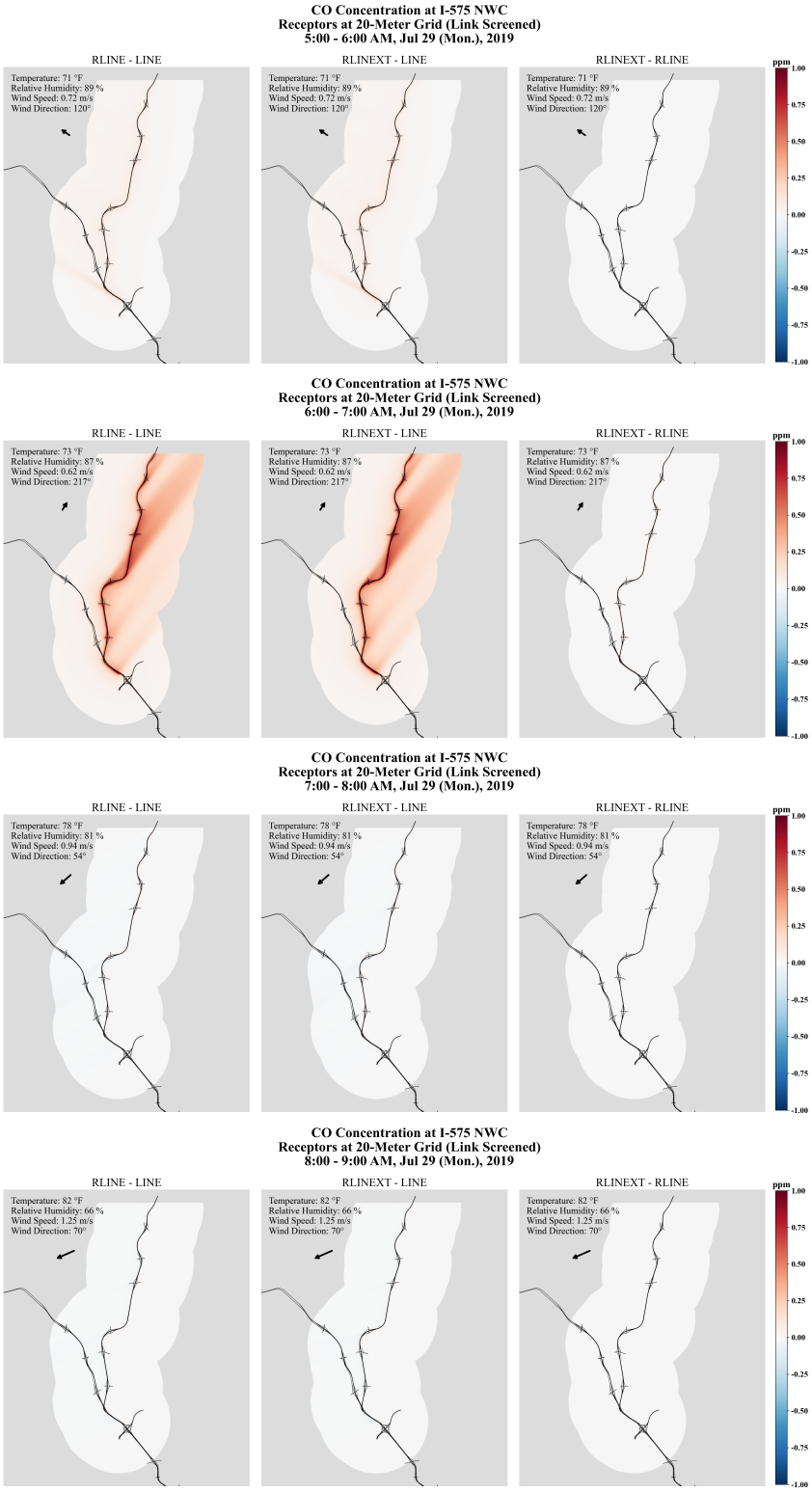


Figure 88 – Differences in Predicted CO Concentrations by Source Type (LINE vs. RLINE vs. RLINEXT without Barrier), Hot Summer Morning Peak

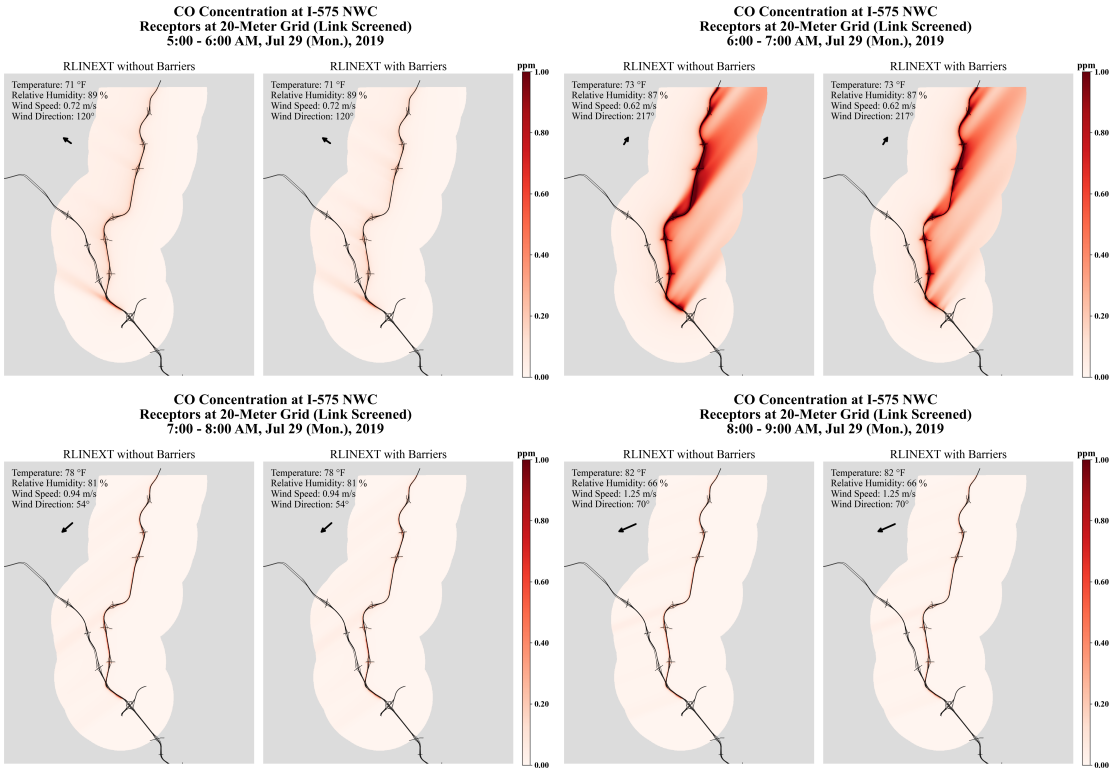


Figure 89 – Predicted RLINEXT CO Concentrations with and without Noise Barriers, Hot Summer Morning Peak

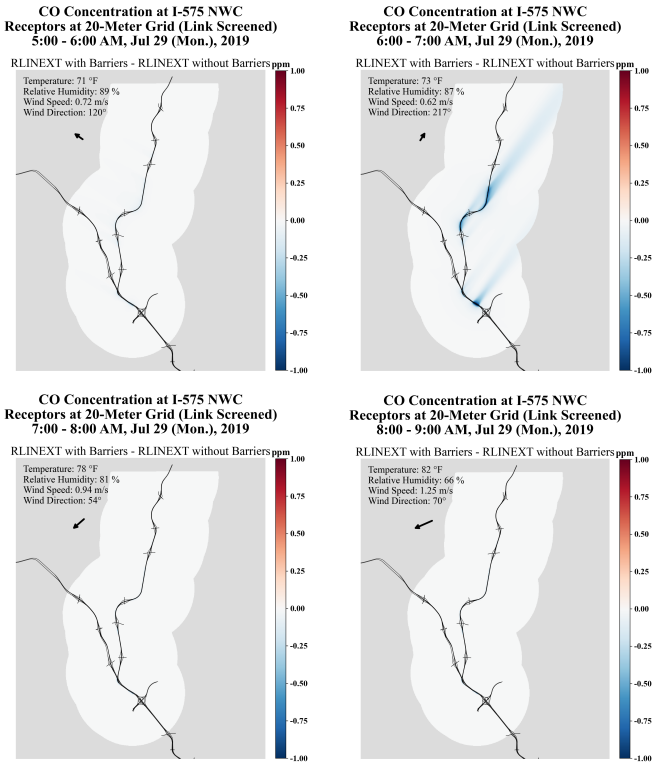


Figure 90 – Differences in Predicted RLINEXT CO Concentrations, with and without Noise Barriers, Hot Summer Morning Peak

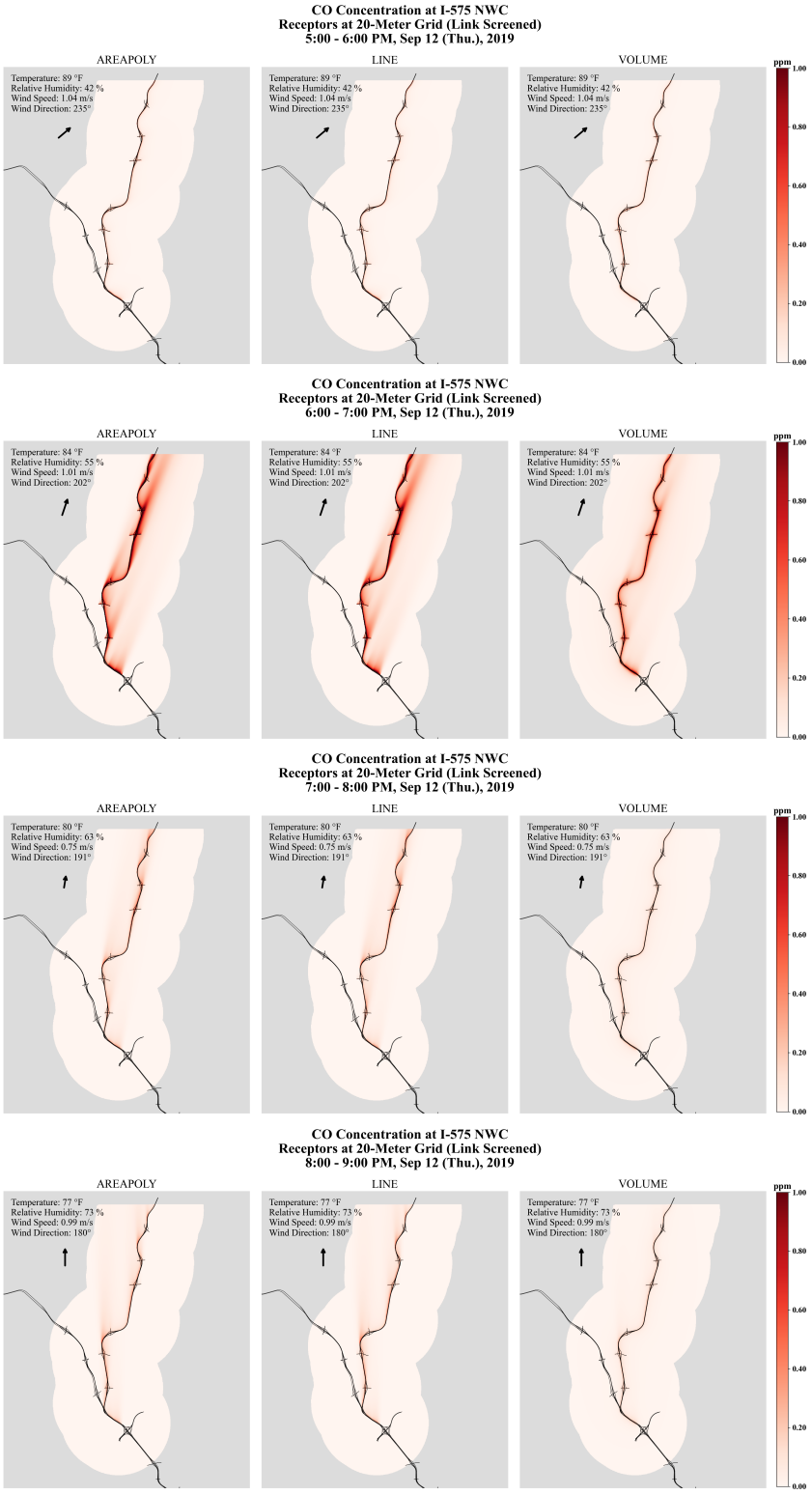


Figure 91 – Predicted CO Concentration by Source Type (AREAPOLY vs. LINE vs. VOLUME), Hot Fall Evening Peak)

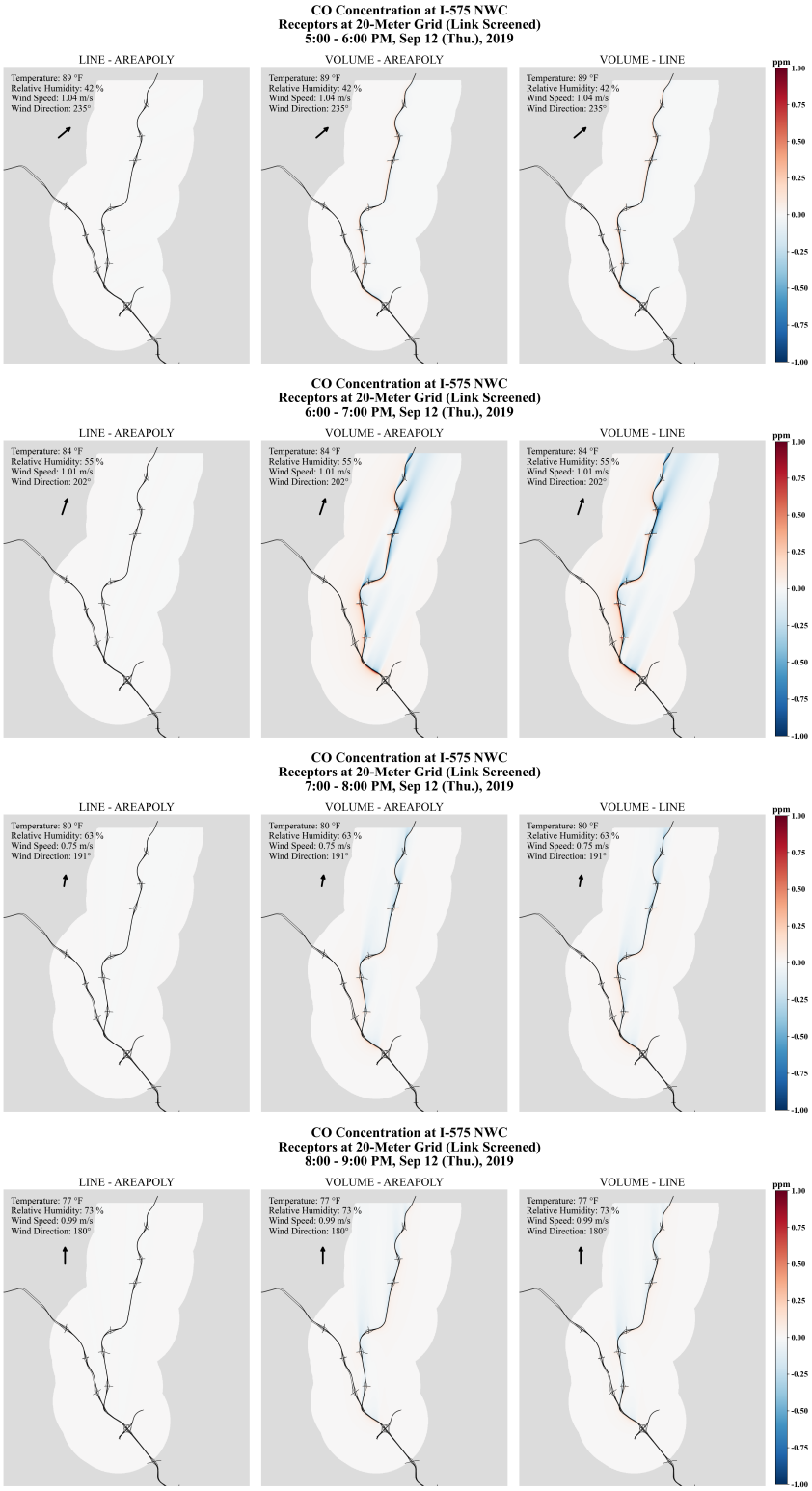


Figure 92 – Differences in Predicted CO Concentrations by Source Type (AERAPOLY vs. LINE vs. VOLUME), Hot Fall Evening Peak

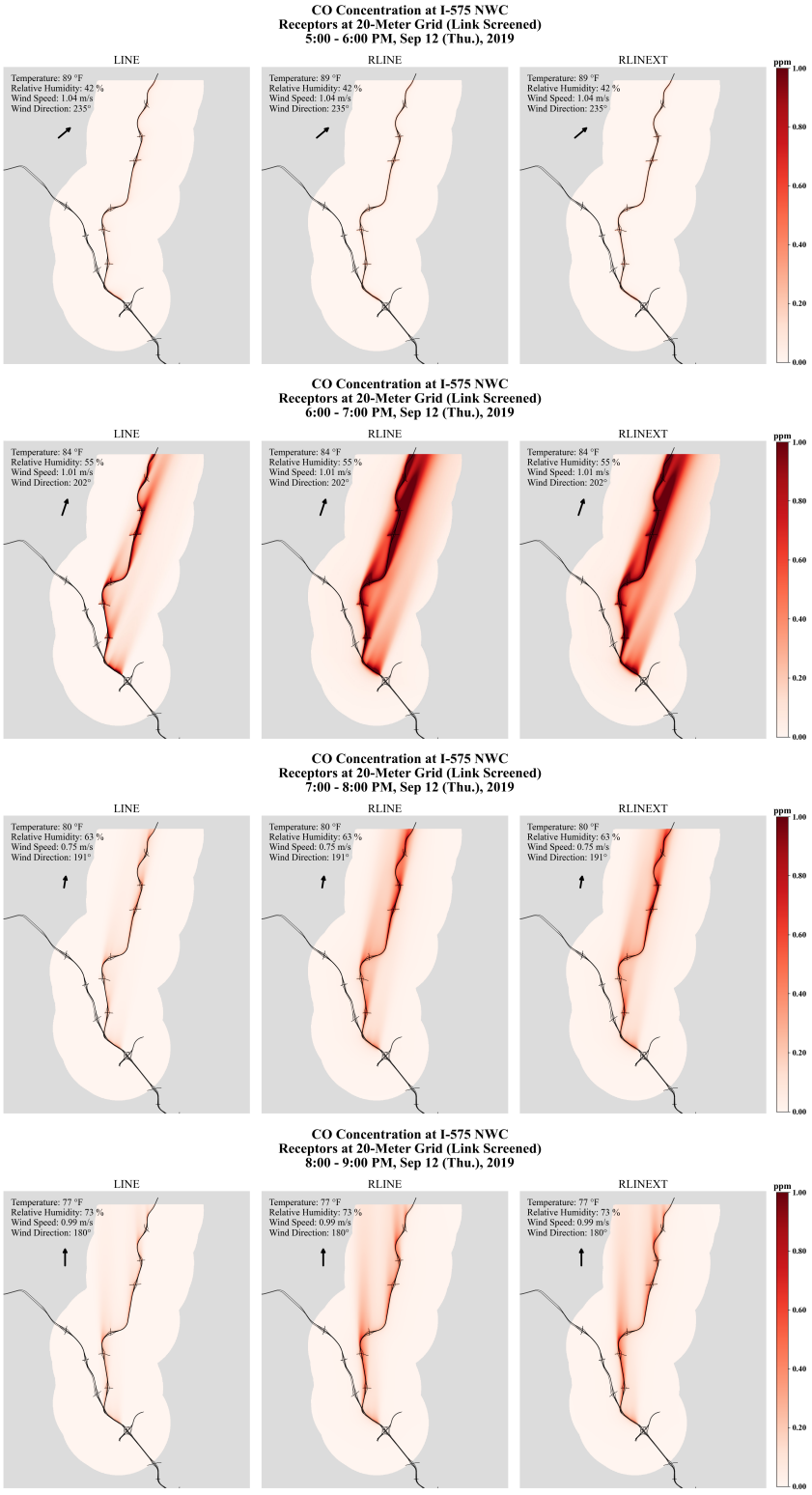


Figure 93 – Predicted CO Concentrations by Source Type (LINE, RLINE, and RLINEXT without Barrier), Hot Fall Evening Peak

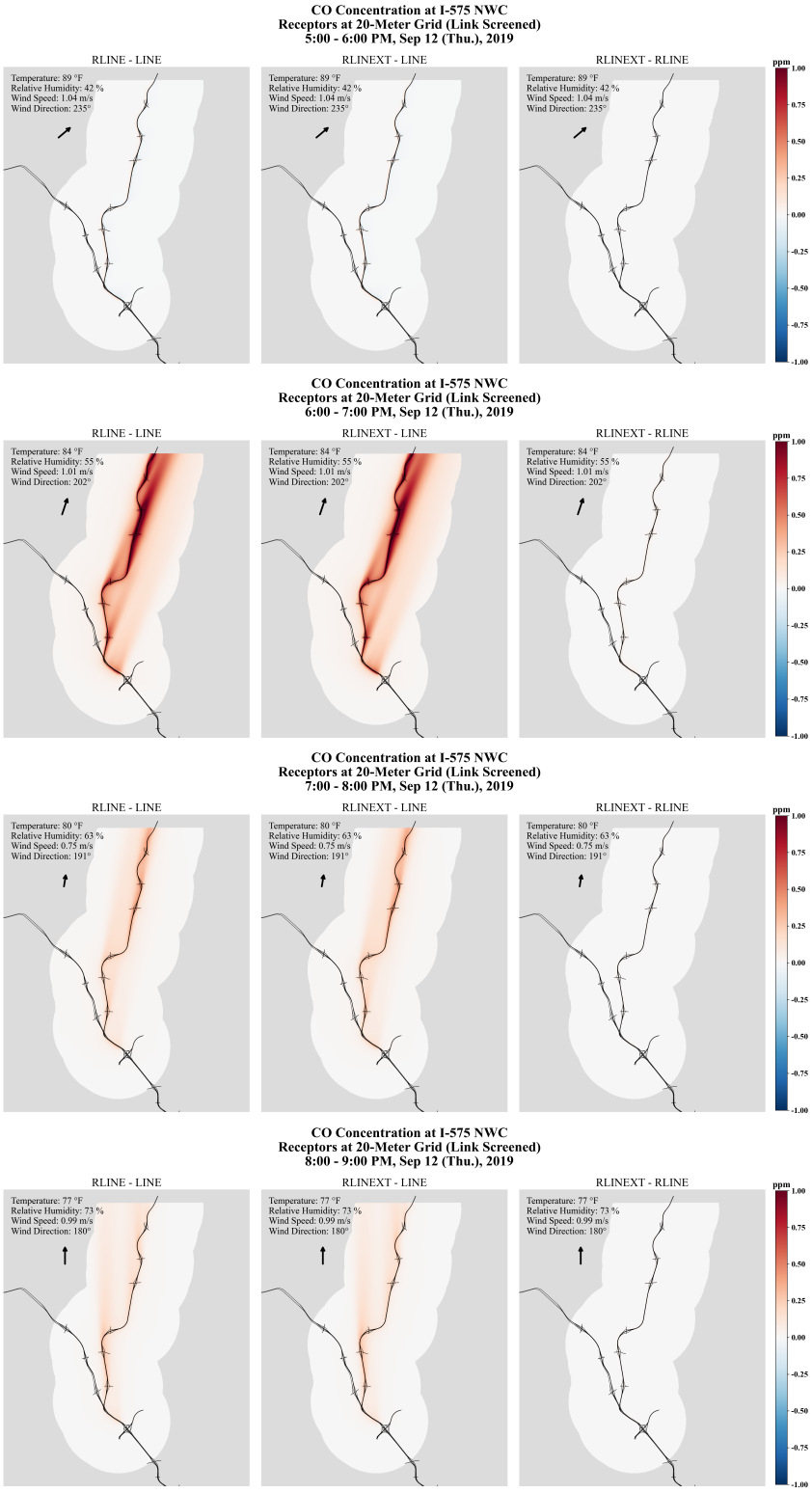


Figure 94 – Differences in Predicted CO Concentration by Source Type (LINE vs RLINE vs. RLINEXT without Barrier), Hot Fall Evening Peak

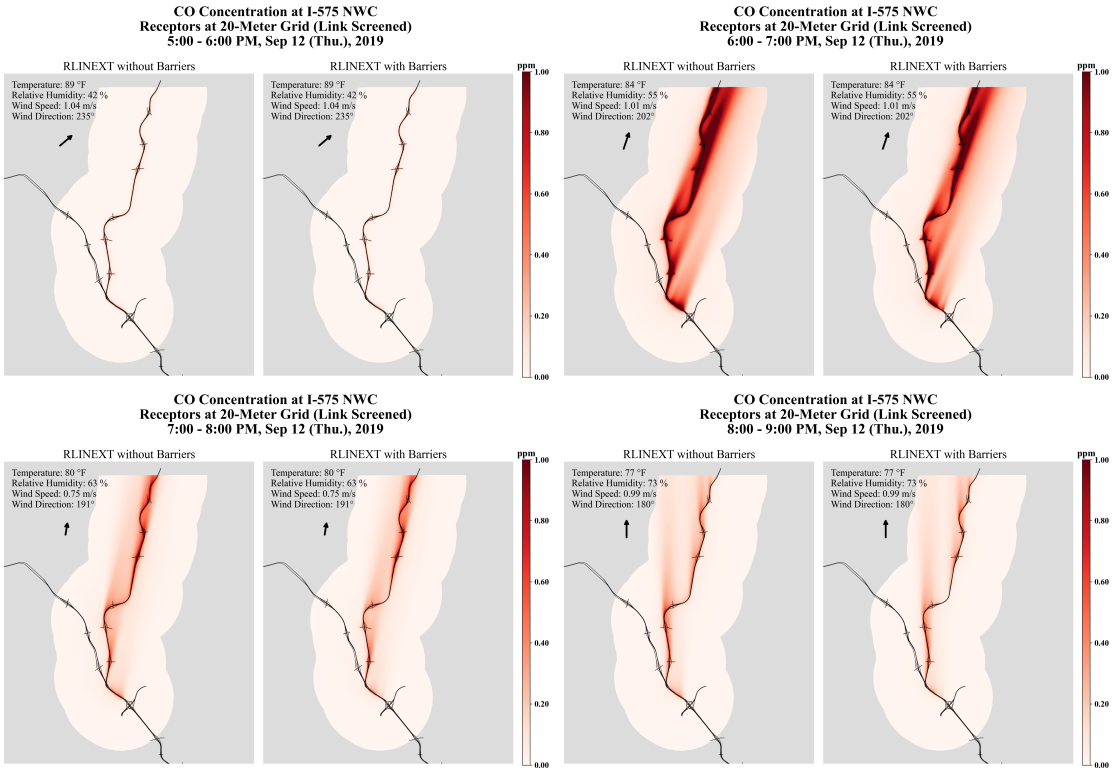


Figure 95 – Predicted RLINEXT CO Concentrations with and without Noise Barriers, Hot Fall Evening Peak

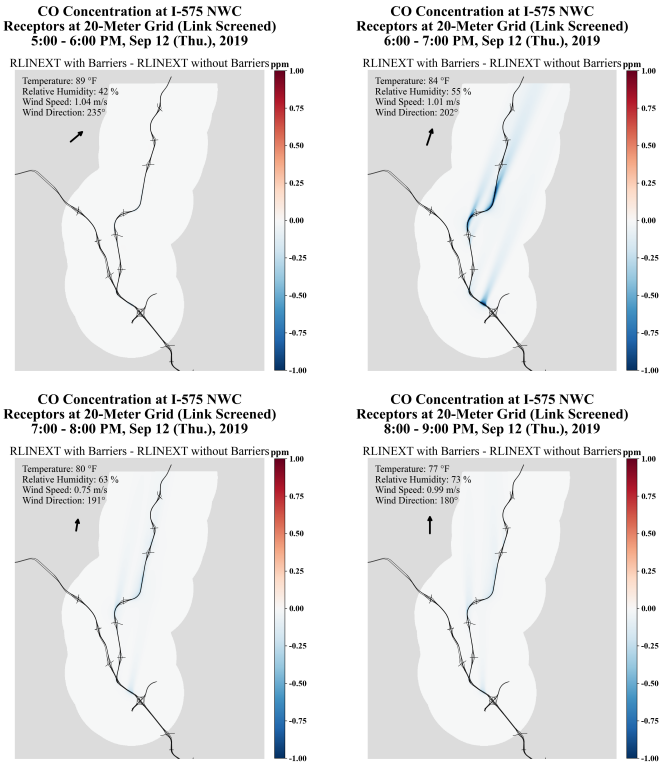


Figure 96 – Differences in Predicted RLINEXT CO Concentrations with and without Noise Barriers, Hot Fall Evening Peak

ACKNOWLEDGMENTS

The authors of this report wish to thank the Georgia Department of Transportation, Georgia State Road and Tollway Authority, and the Atlanta Regional Commission for their data contributions, support, and assistance throughout this research effort.

REFERENCES

ARC, Atlanta Regional Commission (2019). "Model Documentation for the ARC Travel Demand Model." Atlanta, GA. https://atlregional.github.io/ARC_Model/ index.html#calibration, accessed June 2021.ARC. 2015. "The Atlanta Region's Plan – Conformity Determination Report." Atlanta, GA. https://documents.atlantaregional.com/The-Atlanta-Region-s-Plan/rtp/Conformity-Determination-Report.pdf.

Guensler, Randall, Haobing Liu, Yanzhi (Ann) Xu, Alper Akanser, Daejin Kim, Michael P. Hunter, and Michael O. Rodgers. 2017. "Energy Consumption and Emissions Modeling of Individual Vehicles." *Transportation Research Record: Journal of the Transportation Research Board,* 2627: 93–102. https://doi.org/10.3141/2627-11.

Heist, D., C. Hood, A. Venkatram, M. Snyder, V. Isakov, S. Perry, D. Carruthers, J. Stocker, and S. Smith. 2014. "Dispersion Modeling Approaches for Near Road." In *16th International Conference on Harmonisation within Atmospheric Dispersion Modeling for Regulatory Purposes*. Varna, Bulgaria. https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=310695&Lab=NERL.

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye and Tie-Yan Liu. Lightgbm: "A Highly Efficient Gradient Boosting Decision Tree." *Advances in Neural Information Processing Systems* (2017), 30, 3146-3154.

Kim, Daejin. 2020. "Large-Scale, Dynamic, Microscopic Simulation for Region-Wide Line Source Dispersion Modeling." Doctoral dissertation, Georgia Institute of Technology. https://smartech.gatech.edu/handle/1853/64620

Kim, Daejin, Haobing Liu, Michael O. Rodgers, and Randall Guensler. 2020. "Development of Roadway Link Screening Model for Regional-Level Near-Road Air Quality Analysis: A Case Study for Particulate Matter." *Atmospheric Environment*, 237: 117677. https://doi.org/10.1016/j.atmosenv.2020.117677.

Kim, Daejin, Haobing Liu, Xiaodan Xu, Hongyu Lu, Roger Wayson, Michael O. Rodgers, and Randall Guensler. 2020. "Distributed Computing for Region-Wide Line Source Dispersion Modeling." *Computer-Aided Civil and Infrastructure Engineering*, 1–15. https://doi.org/10.1111/mice.12639.

Li, Hanyan, Haobing Liu, Xiaodan Xu, Yanzhi "Ann" Xu, Michael O Rodgers, and Randall Guensler. 2016. "Emissions Benefits from Reducing Local Transit Service Deadheading: An Atlanta Case Study." In *95th Annual Meeting of the Transportation Research Board*. Washington, D.C., U.S.

Li, Hanyan, Yuanbo Wang, Xiaodan Xu, Haobing Liu, Angshuman Guin, Michael O Rodgers, Michael P Hunter, Jorge A Laval, and Randall Guensler. 2017. "Assessing the Time, Monetary, and Energy Costs of Alternative Modes." In *97th Annual Meeting of Transportation Research Board*. Washington, D.C., U.S.

Liu, Haobing (2021). Converting Roadway GIS Layer to AERMOD Geometry for Refined Particulate Matter Project-Level Modeling. University of New Mexico. Project Report Prepared for the Federal Highway Administration. Washington, DC. June 2021.

Liu, Haobing, Randall Guensler, Hongyu Lu, Yanzhi Xu, Xiaodan Xu, and Michael O. Rodgers. 2019. "MOVES-Matrix for High-Performance On-Road Energy and Running Emission Rate Modeling Applications." Journal of the Air & Waste Management Association, July, 10962247.2019.1640806. https://doi.org/10.1080/10962247.2019.1640806.

Liu, H., X. Xu, M. Rodgers, Y. Xu, and R. Guensler (2017). "MOVES-Matrix and Distributed Computing for Microscale Line Source Dispersion Analysis." Journal of the Air and Waste Management Association. DOI: 10.1080/10962247.2017.1287788. Volume 67, Issue 7, pp. 763-775. April 2017.

Lundberg, Scott M, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M. Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal and Su-In Lee. 2020. "From Local Explanations to Global Understanding with Explainable AI for Trees." Nature Machine Intelligence, 2(1), 56-67.

Mitchell, Tom Michael (1997). Machine Learning. New York: McGraw-Hill. 1997.

Mohammed, Mohssen. Machine Learning: Algorithms and Applications. CRC Press, 2016.

USEPA (2015). "Transportation Conformity Guidance for Quantitative Hot-spot Analyses in PM2.5 and PM10 Nonattainment and Maintenance Areas". EPA-420-B-15-084. https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100NMXM.pdf.

Wolf, J., R. Guensler and S. Washington (1998). "High Emitting Vehicle Characterization Using Regression Tree Analysis." Transportation Research Record. Number 1641. pp. 58-65. Transportation Research Board. Washington, DC. 1998.

Xu, Xiaodan, Haobing Liu, James M Anderson, Yanzhi Xu, Michael P. Hunter, and Michael O. Rodgers. 2016. "Estimating Project-Level Vehicle Emissions with VISSIM and MOVES-Matrix." In 95th Annual Meeting of the Transportation Research Board. Washington, D.C., U.S.

Xu, Xiaodan, Haobing Liu, Angshuman Guin, Michael O Rodgers, and Randall Guensler. 2018. "Regional Emission Analysis with Travel Demand Models and MOVES-Matrix." In *97th Annual Meeting of Transportation Research Board*. Washington, D.C., U.S.

Xu, Xiaodan, Haobing Liu, Hanyan Li, Michael O. Rodgers, and Randall Guensler. 2017. "Integrating Engine Start, Soak, Evaporative, and Truck Hoteling Emissions into MOVES-Matrix." In *97th Annual Meeting of Transportation Research Board*. Washington, D.C., U.S. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85039147023&partnerID=40&md5=da9d9ace4e50055e864be9eb63a6bd84.

Xu, Yanzhi, Hanyan Li, Haobing Liu, Michael O. Rodgers, and Randall L. Guensler. 2017. "Eco-Driving for Transit: An Effective Strategy to Conserve Fuel and Emissions." *Applied Energy*, 194: 784–97. https://doi.org/10.1016/j.apenergy.2016.09.101.

Zhao, Yingping. 2020. "Distributional Impact Assessment of Transportation Policy Using Activity-Based Models with Path Retention." Doctoral Dissertation, Georgia Institute of Technology.