UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION 2
NEW YORK, N.Y. 10007

August 19, 2024

MEMORANDUM

SUBJECT: Concurrence Request to Use Site-Specific Alternative Approach to Demonstrate Modeled Attainment of the 2010 SO_2 NAAQS at the Alcoa West Aluminum Smelter in Massena, New York

FROM: Brian Marmo, Regional Air Quality Modeler
Permitting Section, Air Programs Branch, Air and Radiation Division
EPA Region 2, New York, New York

BRIAN MARMO Digitally signed by BRIAN MARMO Date: 2024.08.19 10:24:40 -04'00'

THRU: Richard Ruvo, Director
Air and Radiation Division
EPA Region 2, New York, New York

Richard Ruvo Digitally signed by RICHARD RUVO Date: 2024.08.19 11:28:34 -04'00'

TO: George Bridgers, Model Clearinghouse Director
Air Quality Modeling Group, Air Quality Assessment Division
Office of Air Quality Planning and Standards

The U.S. Environmental Protection Agency (EPA) Region 2 seeks concurrence from the Model Clearinghouse regarding the prospective EPA Region 2 approval of an alternative modeling approach as part of New York State Department of Environmental Conservation's (NYSDEC) attainment demonstration of the 2010 SO_2 National Ambient Air Quality Standard (NAAQS). As part of EPA's Round 4 SO_2 NAAQS designation process, a portion of St. Lawrence County, New York surrounding the Alcoa Massena West facility (hereafter, Alcoa Massena) was designated as non-attainment with respect to the 1-hour SO_2 NAAQS. NYSDEC is requesting approval to use a site-specific alternative modeling approach. The proposed approach involves categorizing each of the two roof vents as a series of four-point sources, instead of categorizing each roof vent as a buoyant line (BUOYLINE) source. This method would be modeled in conjunction with the two other site-specific non-guideline modeling approaches that were approved by EPA Region 2 on September 20, 2022 and concurred by the Model Clearinghouse on October 14, 2022 (Record No: 22-II-03). Those two alternative modeling approaches, which modify inputs to the EPA preferred guideline model, AERMOD (**A**merican Meteorological Society/**E**nvironmental Protection Agency **R**egulatory **Mod**el), include using a neutral lapse rate in the lower 100 meters of the atmosphere and incorporating the 2019 draft version of BPIPPRM (**B**uilding **P**rofile **I**nput **P**rogram for PRIME, **P**lume **Ri**se **M**odel **E**nhancements). As with the October 14, 2022, approval, NYSDEC has requested to allow the use of an alternate modeling approach for their air quality modeling analysis, under 40 CFR Part 51, Appendix W §3.2.2(b), Condition (2), for their attainment demonstration. Under Condition (2), an alternative model may be used if the Regional Office finds the conditions specified in Appendix W §3.2.2(d) are satisfied.

NYSDEC submitted their alternative model request on June 18, 2024 (Attachment 1), along with a July 17, 2024 Modeling Protocol (Attachment 2) that included the technical analyses prepared by Alcoa Massena. The request provided evidence and justifications supporting approvability of the alternative modeling approach under Appendix W §3.2.2(b), Condition (2). EPA Region 2 has conducted a thorough review of the request and intends to approve the use of the alternate modeling approach for the Alcoa Massena attainment demonstration of the 2010 SO_2 NAAQS. Region 2 found the proposed application of the model is satisfactory under the requirements of §3.2.2(d). A technical analysis summarizing our review of the submittal is below. Please feel free to contact Brian Marmo at (212) 637-4352, if you have any questions regarding the request.

EPA Region 2's Technical Review of New York State Department of Environmental Conservation's Request to Use Site-Specific Alternative Modeling Approach in AERMOD

1. Background and Project Overview

NYSDEC has requested to use an alternate modeling approach, as provided in §3.2 of the *Guideline* on Air Quality Models (40 CFR Part 51, Appendix W, hereafter referred to as the *Guideline*), to conduct its modeled attainment demonstration of the 1-hour SO_2 National Ambient Air Quality Standard (NAAQS) for Alcoa Massena. Alcoa Massena is an aluminum production facility located in St. Lawrence County on the northern outskirts of the town of Massena, NY. Alcoa Massena has a capacity to produce 136,000 metric tons of primary aluminum per calendar year. The area surrounding the facility is rural with simple terrain within several kilometers and is located along the St. Lawrence Seaway. The facility has one potline building consisting of two long rooms with 36 dry scrubber stacks between them. Emissions from the dry scrubber stacks comprise most of the SO_2 emissions at the smelter, with one bake oven stack also emitting SO_2. Roof vents emit a small fraction of the stack emissions (approximately 5%). Emission rates were found to be approximately uniform across a monthly basis.

There are two site-specific SO_2 ambient monitors near the facility that were sited by NYSDEC using the BLP option in AERMOD in 2017 (see Figure 1). The area was designated nonattainment with respect to the 1-hour SO_2 NAAQS on April 30, 2021, since the two site specific ambient monitors measured concentrations above the 1-hour SO_2 NAAQS between 2017 and 2019.

NYSDEC submitted an alternative model request on June 28, 2022. In this request, NYSDEC and Alcoa concluded that the EPA regulatory default AERMOD model overpredicted concentrations when compared to measurements at the two site-specific SO_2 ambient monitors. After review of the AERMOD debug files, it was found that the plume rise was underestimated. This was likely due to the excess fugitive heat loss emanating from the building, which is not accounted for in AERMOD. To address this issue, the meteorological data were modified such that the lowest 100 meters of the atmosphere featured a neutral lapse rate. In addition, since the downwash effects are overestimated for long narrow buildings, the default BPIPPRM, version 04274, was replaced by the 2019 draft version of BPIPPRM (v19191_DRFT) which reduces the building footprint for elongated rectangular buildings, compared to default BPIPPRM, when the wind is oblique to the building face. On September 20, 2022, EPA Region 2 approved these two alternative modeling approaches with concurrence from the Model Clearinghouse on October 14, 2022. However, even with these modifications to AERMOD, the model continued to overpredict concentrations.

In December 2022, the Alcoa informed NYSDEC that their future stack configuration may feature three 63.2-meter potline reactor stacks. In their July 17, 2024 protocol, Alcoa's proposed to physically merge the 36 dry scrubber stacks into six future potline reactor stacks (two stacks per each of the three courtyards), with a height of 50 meters each. The two stacks per courtyard would be positioned within one diameter of each other, and for modeling purposes, would be merged and treated as a single stack. These future potline reactor stacks would be much taller than the existing stack heights of 23.3 meters but would remain less than the 65 meter de minimis Good Engineering Practice (GEP) stack height. The merging and raising of the stacks are creditable under the Clean Air Act since the total tons of SO_2 emissions are less than 5000 tons per year and the stack heights will be under 65 meters.

ln the October 14, 2022, approval, the small percentage of SO_2 emissions from the two roof vents were added to the emissions from the 36 potline reactor stacks because they were emitted at roughly the same height. However, since the SO_2 emissions from the future potline reactor stacks would potentially be released 27 meters above the SO_2 emissions from the shorter roof vents, the roof vent emissions will now need to be modeled as a separate source.

A modeled attainment demonstration is required for the approval of this strategy into its State Implementation Plan (SIP). As discussed below, categorizing the two roof vents as a series of four-point sources further improves the model performance when compared to measured values at the two SO_2 ambient monitors. EPA Region 2 has reviewed NYSDEC's request to implement this alternative model approach and determined that the use of the proposed method is acceptable.

Figure 1: Alcoa Massena facility and location of the two SO_2 monitoring sites

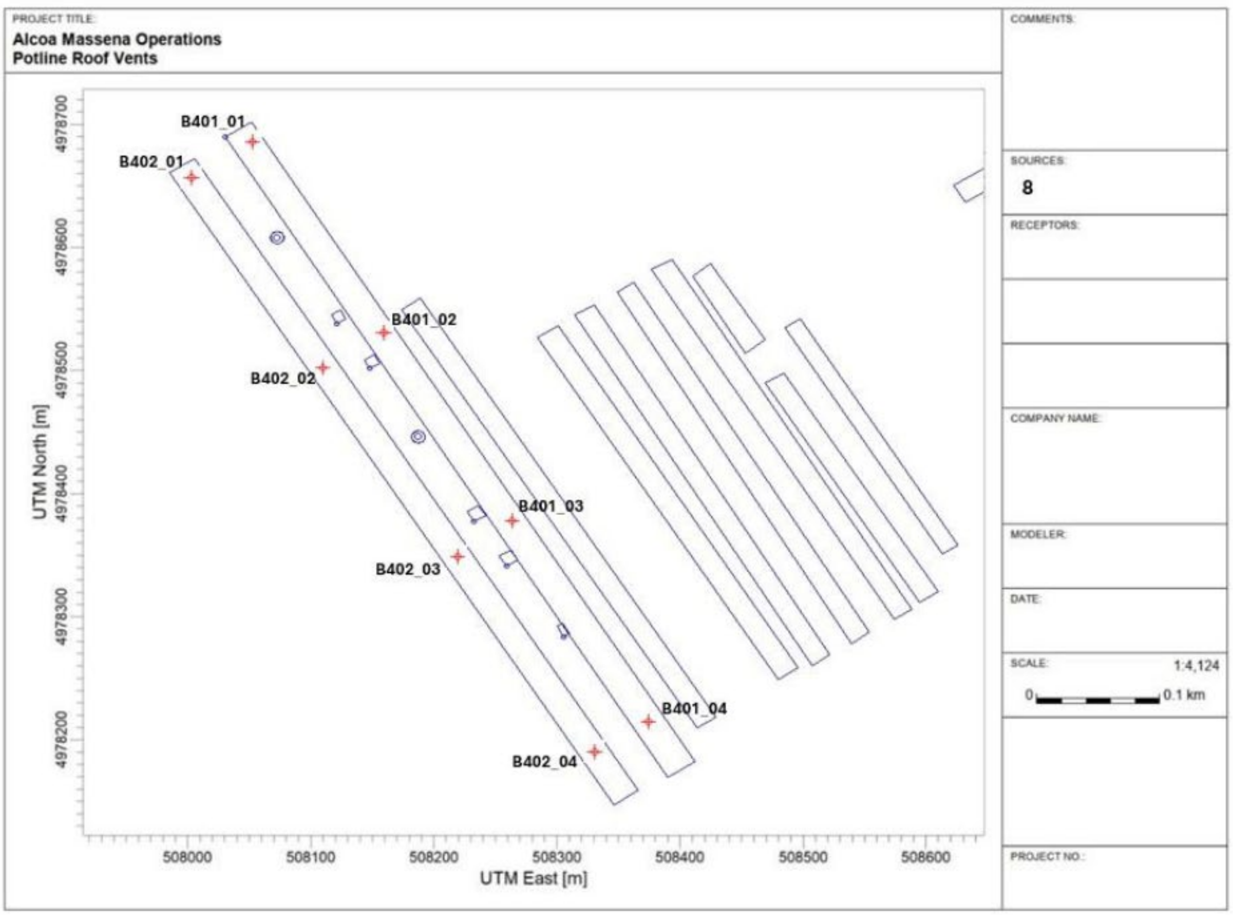
2. Proposed Modeling Approach

The buoyant line source type, BUOYLINE, in AERMOD, equivalent to the Buoyant Line and Point Source (BLP) model operating inside of AERMOD, was developed for aluminum reduction facilities, and is considered the most appropriate approach for modeling roof vents. However, the BUOYLINE source type ignores the user-defined vertical temperature profile data file for its meteorological input. Instead of utilizing the neutral lapse rate in the lower 100 meters of the atmosphere, AERMOD automatically assigns a stable lapse rate for the BUOYLINE source type during nighttime hours for 'rural' source types such as Alcoa Massena. Due to the excess fugitive heat loss emanating from Alcoa Massena, characterizing the two roof vents using the BUOYLINE source type presents a challenge. In contrast, for ‘urban’ source types, the BUOYLINE will utilize the neutral lapse rate for the nighttime hours. While the neutral lapse rate is more representative of the facility in the bottom 100 meters, there is no vertical upper limit to the plume for ‘urban’ sources when modeling with as a BUOYLINE source. Additionally, based on earlier test cases for Alcoa Massena where it was modeled as an urban source, the modeled concentrations were too low and not acceptable. Thus, both NYSDEC and Region 2 agree that the facility should be modeled as a rural source.

To address the issue in which these ‘rural’ buoyant line sources overpredict SO_2 concentrations at the two ambient monitors, Alcoa Massena’s consultant has proposed modeling each of the two roof vents as a series of four-point sources (see Figure 3-5). Overall, there would be eight points sources each with a stack height equal to the two actual roof vents. Furthermore, to ensure the roof vent area, exit velocity, and flow rate are conserved, an effective stack diameter was used for the four-point sources.

To determine the most representative number of stacks, a sensitivity study was performed. In this study, the cases where the roof vents were represented as stacks were compared to the maximum design concentration when the roof vents are modeled as buoyant line sources. The maximum design concentration for the four stacks per roof vent case (9.5 µg/m^3) was the closest to the BUOYLINE source type's maximum design concentration (9.4 µg/m^3). It was also selected because it is slightly higher for conservatism.

Figure 3-5: Potline Roof Vents Modeled as 4 Stacks Along Each Roof Vent

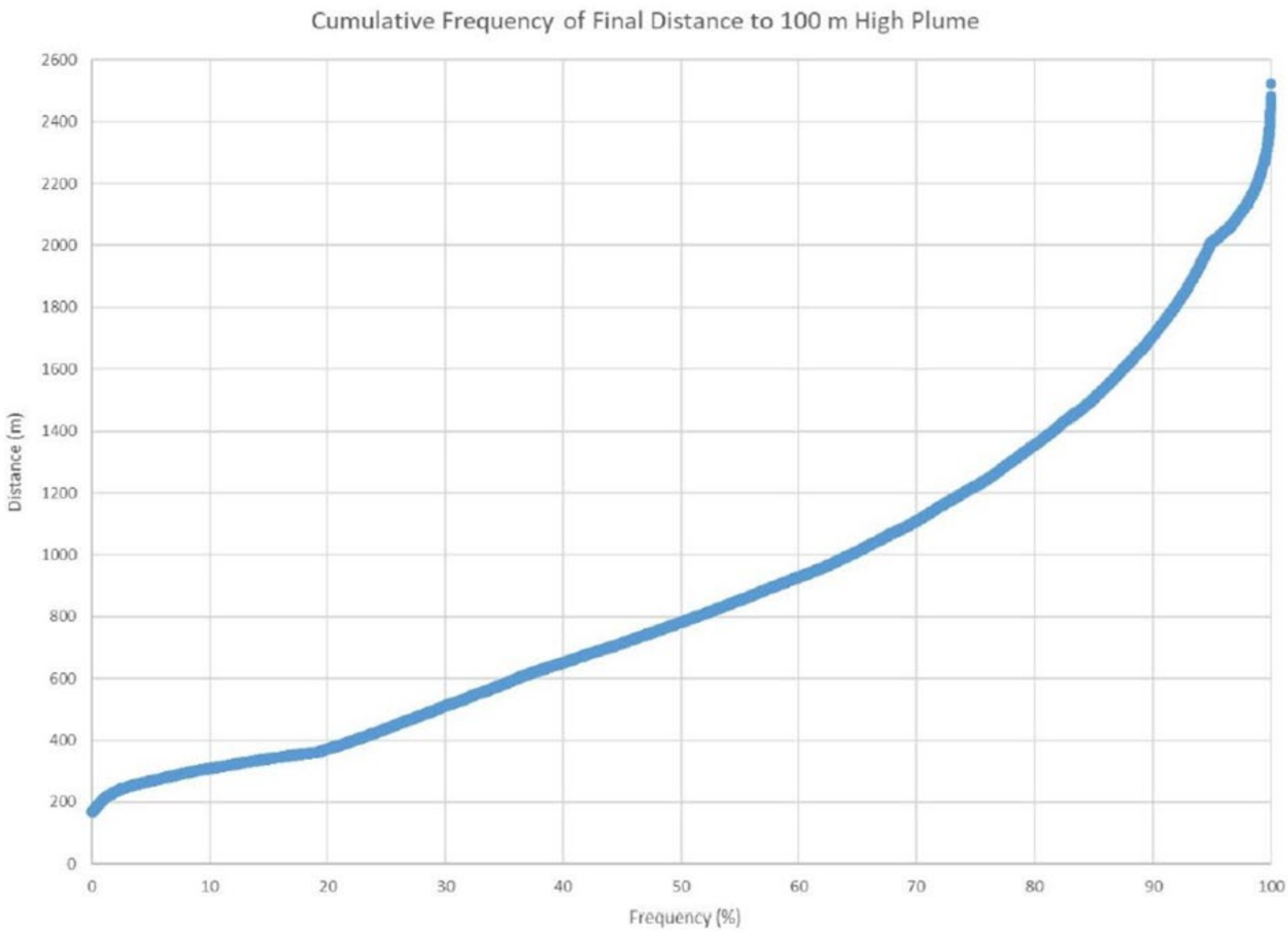


3. Region 2's Review of the Alternative Model Proposal

Though using BUOYLINE source type is the preferred *Guideline* approach for modeling roof vents at an aluminum reduction facility, Region 2 acknowledges that BUOYLINE not utilizing the user-defined vertical temperature profile data file presents challenges. Region 2 also agrees that ignoring the fugitive heat loss from the facility is likely to result in an underestimation in plume rise and an overprediction of SO_2 concentrations.

During a call on January 24, 2023, EPA voiced its concerns regarding the BUOYLINE urban approach and the lack of a vertical upper limit for the lapse rate. To assess the BUOYLINE urban approach for the roof vents, Alcoa used the DEBUG option to evaluate the plume height as a function of distance. The debug option listed seven plume height values at seven corresponding downwind distance, with the seventh value representing the final plume rise. Based on these results, greater than 90% of the nighttime hours (the hours impacted by the urban dispersion option) had final plume rise heights above 100 meters. Next, the facility determined the distance for the buoyant line sources to attain a median plume height of 100 meters. Varying as a function of meteorological conditions, the median distance was around 800 meters (see Figure 3-6). The 800 meters is approximately the closest distance to the facility's ambient air boundary (in the north-northeast direction).

Figure 3-6: Cumulative Frequency Distribution of Final Plume Height Distances



As previously mentioned, NYSDEC and Region 2 agree that the facility should be modeled as a rural source. And since BUOYLINE does not utilize a user-defined vertical temperature profile data file and automatically assigns a stable lapse rate to the nighttime hours for rural sources, Region 2 believes that it is reasonable to categorize each roof vent source as a series of point sources.

To determine the most appropriate number of stacks to properly represent each roof vent, several stack case scenarios were evaluated by Alcoa. This testing methodology utilized the results from the BUOYLINE urban approach for when the plume reached the ambient air boundary, and the plume height was at or below the 100 meters. For the scenario when the roof vents were modeled as buoyant line sources, the 5-year (2017-2021) average emission rate of 1.675 g/s was chosen for each roof vent.

In addition, the roof vent height was used for the release height, the roof vent opening width was used for the line width, and the temperature and flow rate were based on test data from 2017-2020 (see Table 3-1). For the scenarios where the roof vents were modeled as stacks, the roof vent release height was used for the stack height, an effective stack diameter was used to ensure the exit velocity was conserved, and the temperature was based on test data from 2017-2020. For the roof vent emission rate, the values used in BUOYLINE urban case were divided by the number of stacks modeled (see Table 3-2). The maximum design concentration for the four stacks per roof vent case (9.5 µg/m3) was found to be the closest to the BUOYLINE's maximum design concentration (9.4 µg/m3) (see Table 3-3). Additionally, it was selected because it is slightly higher for conservatism.

Table 3-1: Source Parameters and Emissions for the Roof Vents Modeled as BUOYLINE Sources

Source Parameter	Value
Modeled Emission Rate Per Roof Vent (g/s)(1)	1.675
Base Elevation (m)	65.5
Release Height (m)(2)	20.31
Source Length (m)	565.7
Building Length (m)	630.0
Building Height (m)	17.39
Building Width (m)	25.6
Line Width (m)(3)	3.9
Building Separation (m)	25.6
Average Ambient Temperature (K)(4)	280.8
Median Potroom Temperature (K) (5)	305.8
Average Potroom Flow Rate (m^3/s) (6)	2066.04
Buoyancy Parameter, F' (m^4/s^3)	1656.27

(1) 5-year average emissions based on the actual monthly emissions from 2017-2021.
(2) Roof vent height is used for the release height.
(3) Roof vent opening width is used for the line width.
(4) Ambient temperature is based on the average temperature from the meteorological data from 2017-2021.
(5) Temperature data are based on the median potroom test data from 2017-2020.
(6) Flow rate is an average flow rate based on potroom test data from 2017-2019.

Table 3-2: Source Parameters and Emissions for the Roof Vents Modeled as 4 Stacks per Roof Vent

Source Parameter	Value for Each of the 4 Stacks per Roof Vent
Modeled Emission Rate Per Stack (g/s)(1)	0.419
Base Elevation (m)	65.5
Stack Height (m)(2)	20.31
Stack Temperature (K)(3)	305.8
Stack Exit Velocity (m/s)(4)	0.94
Equivalent Stack Diameter (m)(5)	26.5

(1) Total emissions per roof vent are divided by the number of stacks modeled.
(2) Roof vent release height is used for the stack height.
(3) Temperature data are based on the median potroom test data from 2017-2020.
(4) Stack exit velocity is based on the total vent area and vent flow rate in potroom test data from 2017-2019.
(5) Equivalent stack diameter is based on the resulting individual stack area when the roof vent opening area is divided by the 4 modeled stacks.

Table 3-3: Stack Sensitivity Analysis Results

Roof Vents Modeled As...	Modeled Maximum 99th Percentile Design Concentration (µg/m^3)
2 buoyant line sources run with urban dispersion and 250 K population	9.4
3 stacks per roof vent	8.4
4 stacks per roof vent	9.5
5 stacks per roof vent	10.5

4. Regulatory Analysis and Background

Section 3.1.2(a) of the *Guideline* states if a model is required for a particular application, the user must select a model that is specified in the *Guideline*. However, Section 3.1.2(a) also provides that the user may follow the procedures in section 3.2.2 for use of an alternative model or technique. In addition, Section 3.1.2(c) states that the use of the modified model must then be justified as an alternative model on a case-by-case basis to the appropriate reviewing authority and approved by the Regional Administrator.

The alternative model approval process and conditions are outlined in Section 3.2 of the *Guideline*. Section 3.2.2(a) specifies that the determination of acceptability of an alternative model is a Regional Office responsibility in consultation with EPA's Model Clearinghouse. An alternative model may be used subject to Regional Office approval if found to satisfy the requirements listed in Section 3.2.2. Section 3.2.2(b) states the alternative model shall be evaluated from both a theoretical and performance perspective before regulatory use and outlines the three separate conditions where an alternative model may be approved. Condition 2 under Section 3.2.2(b), where a statistical performance evaluation using measured air quality data and the results of that evaluation indicate that the alternative model performs better for the given application than a comparable model in Appendix A to the *Guideline*, applies to this case.

a. Evaluation of Approach under Section 3.2.2(d)

An alternative model is evaluated from both a theoretical and a performance perspective before it is selected for use. The scientific justification provided above addresses the theoretical perspective. For this specific application, NYSDEC and Alcoa selected the model performance procedures for the second of three possible alternative model approaches (Appendix W section 3.2.2(b)(2)): "If a statistical performance evaluation has been conducted using measured air quality data and the results of that evaluation indicate the alternative model performs better for the given application than a comparable model in Appendix A…"

Alcoa conducted three sets of statistical evaluation tests to assess the modified model performance: a) quantile-quantile (Q-Q) plots for each monitor, b) comparison of the modeled and observed 5-year average 1-hour average design concentration for each monitor, and c) the use of the Robust Highest Concentration (RHC) as part of EPA's Cox-Tikvart procedure. Each of the statistical tests showed improvement in performance with the modified model.

The two Q-Q Plots below from Figure 4-5 and 4-6 of the July 17, 2024 Modeling Protocol demonstrate the improved model performance when comparing the default AERMOD ("AERMOD Default") and the modified model ("Massena_MOD") with the alternate model approaches. In addition to modeling the two roof vents as a series of four-point sources, the modified model includes the two approaches that were approved by the Model Clearinghouse on October 14, 2022 (neutral lapse rate in the lower 100 meters of the atmosphere and incorporating the 2019 draft version of BPIPPRM). For Q-Q Plots, a concentration closer to the 1:1 line indicates a better correlation with measured data. In this case, the modified model's maximum concentrations are closer to the 1:1 line with smaller overpredictions, which indicates that there is still a conservative bias.

Figure 4-5: Q-Q Plot for Site 1

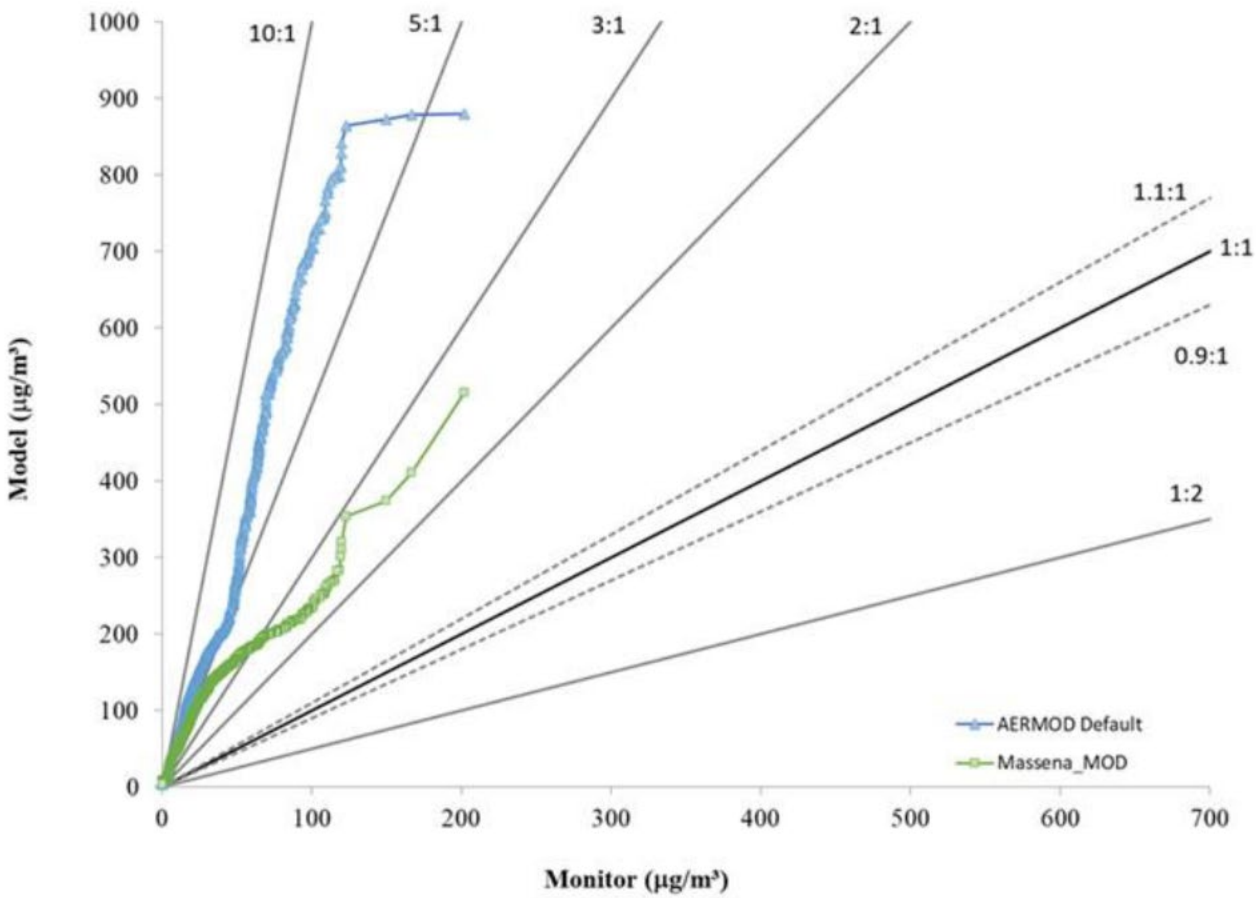
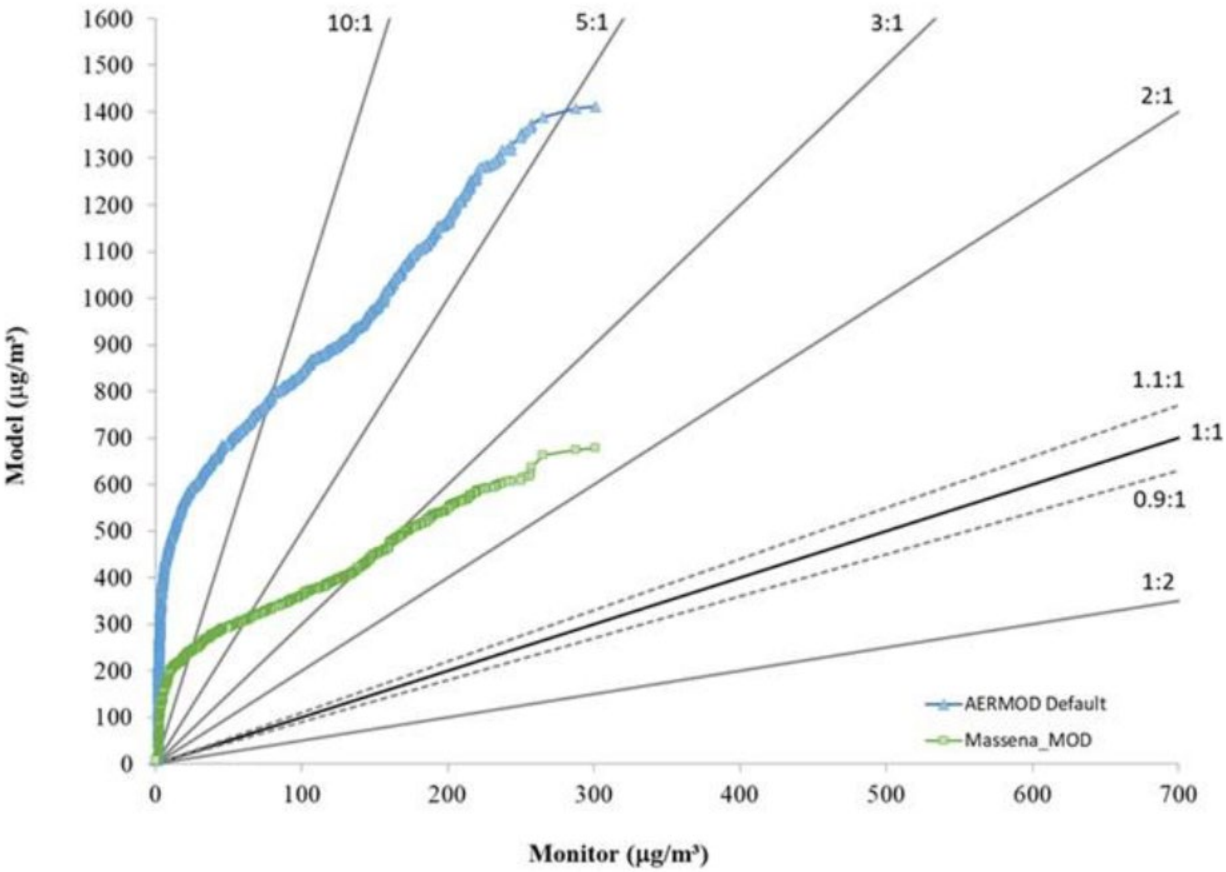


Figure 4-6: Q-Q Plot for Site 2



A separate evaluation to observe the 99th percentile (corresponding to the 4th highest daily 1-hour maximum concentration) design concentration was provided as well. As with the Q-Q Plots, a predicted/observed ratio closer to 1 indicates better performance when compared to measured values. These ratios were provided for both monitoring sites in Table 4-7 and 4-8 of the modeling protocol. In each case, the correlation was improved and positive which demonstrates better correlation with the 99th percentile design values with a conservative bias.

Table 4-7: Modeled-to-Observed Design Concentrations at Site 1

	4th Highest Design Concentration (ug/m³)						
Model Approach	**2017**	**2018**	**2019**	**2020**	**2021**	**5-year Average**	
Observed	118.7	106.6	110.2	110.6	101.1	109.4	**MOR**
AERMOD Default Model	777.5	751.6	612.4	811.5	744.1	739.4	6.8
Massena_MOD Model	240.2	242.8	196.4	270.0	320.7	254.0	2.3

Table 4-8: Modeled-to-Observed Design Concentrations at Site 2

Model Approach	4th Highest Design Concentration (ug/m3)						
	2017	2018	2019	2020	2021	5-year Average	
Observed	237.1	226.7	214.9	232.6	243.6	231.0	**MOR**
AERMOD Default Model	1282.5	1285.3	1031.0	1318.9	1345.0	1252.5	5.4
Massena_MOD Model	588.2	605.1	531.8	616.9	605.2	589.5	2.6

Lastly, the third statistic was provided using the Robust High Concentration. The results are displayed in Tables 4-10 and 4-11 (see below). Once again, the predicted to observed concentrations were improved at both monitoring sites with the modified model with a conservative bias.

Table 4-10: 5-Year Averaged Robust High Concentrations (µg/m3) for Site 1

Model Approach	RHC	Pre/Obs Ratio
Observed	145.6	--
AERMOD Default Model	1028.2	7.1
Massena_MOD Model	309.3	2.1

Table 4-11: 5-Year Averaged Robust High Concentrations (µg/m3) for Site 2

Model Approach	RHC	Pre/Obs Ratio
Observed	282.0	--
AERMOD Default Model	1425.5	5.1
Massena_MOD Model	688.1	2.4

5. Conclusions and Conditions for Use

The statistical evaluations above were done for the current stack configuration. The proposed attainment strategy in this case is to physically merge and raise the stack heights within the GEP stack heights. Emissions will remain below 5000 tons per year and be federally enforceable. Region 2 believes that the alternative model approaches will continue to apply to the Alcoa Massena facility under the proposed attainment strategy for the same reasons as above and may be used for the modeled attainment demonstration.

EPA Region 2 has reviewed the alternative model request submittal provided by NYSDEC and has determined that the proposed modeling approach is acceptable for the attainment demonstration showing how the Alcoa Massena West facility plans to achieve attainment with the 2010 SO_2 NAAQS. Based on our review, we find that the proposed approach addresses the elements contained in Section 3.2.2(d) of the *Guideline*. As such, pursuant to Sections 3.0(b) and 3.2.2(a), Region 2 currently intends to approve the use of this additional site-specific alternative modeling approach in AERMOD. We seek the concurrence from the Model Clearinghouse.

Attachment 1 – NYSDEC Request letter sent on June 18, 2024

NEW YORK STATE DEPARTMENT OF ENVIRONMENTAL CONSERVATION

Division of Air Resources, Bureau of Air Quality Analysis and Research
625 Broadway, Albany, New York 12233-3259
P: (518) 402-8402 | F: (518) 402-9035
www.dec.ny.gov

June 18, 2024

Mr. Richard Ruvo
U.S. EPA Region II
290 Broadway, 25th Floor
New York, N.Y. 10007-1866

Dear Mr. Ruvo:

The New York State Department of Environmental Conservation (NYSDEC) is seeking concurrence with EPA Region 2 on approval of the use of a site-specific modeling approach for sulfur dioxide (SO_2) emissions at Alcoa, an aluminum smelter located in Massena, NY. As part of EPA's Round 4 SO_2 National Ambient Air Quality Standard (NAAQS) designation process, a portion of St. Lawrence County surrounding the Alcoa Massena facility was designated as a non-attainment area. The proposed non-guideline modeling method will be used as part of an attainment demonstration showing how the Alcoa Massena facility plans to achieve attainment with the 2010 SO_2 NAAQS.

The proposed site-specific modeling approach involves categorizing each of the two roof vents as a series of four point sources, instead of categorizing each roof vent as a buoyant line (BUOYLINE) source. These roof vents account for approximately 5% of the total SO_2 emissions from the facility. The roof vent point sources would be modeled in conjunction with two other site-specific non-guideline modeling approaches that had been approved by the EPA Model Clearinghouse on October 14, 2022 (Record No: 22-11-03). These include the use of a neutral lapse rate to simulate the fugitive heat loss continuously emanating from Alcoa Massena's potline buildings, as well as incorporating the 2019 draft version of the Building Profile Input Program for PRIME (BPIPPRM 19191_DRFT).

When NYSDEC presented the alternative modeling approach to EPA Region 2 on June 28, 2022, the future stack configurations at the Alcoa Massena facility were unknown. In December 2022, the facility informed NYSDEC that Alcoa's future stack configuration may include three 63.2-meter potline reactor stacks. In the most recent November 2023 modeling protocol, Section 5 indicated that there may be six future potline reactor stacks, each with an approximate height of 50 meters.

These future potline reactor stacks would be much taller than the existing stack heights of 23.3 meters. In the June 2022 modeling analysis, the small percentage of SO_2 emissions from the two roof vents had been added to the emissions from the 36 potline reactor stacks because they were emitted at roughly the same height. Since the SO_2 emissions from the future potline reactor stacks would potentially be released 27 meters above the shorter roof vents, the SO_2 emissions from each roof vent will now need to be modeled as a separate source.

Section 3.2.2 of Appendix W to 40 CFR Part 51 defines three conditions under which an alternative model may be considered for use. One of these conditions, found in Section 3.2.2(b)(2), states that an alternative model may be considered "if a statistical performance evaluation has been conducted using measured air quality data and the results of that evaluation indicate the alternative model performs better for the given application than a comparable model in Appendix A". A detailed case-specific monitor-to-model comparison was performed by AECOM, and details regarding this analysis can be found in Section 4.2 of "SO_2 Modeling Protocol for Alcoa Massena Operations - West Plant – Draft, Section 4 Only" submitted to NYSDEC on June 14, 2024, by AECOM. This statistical evaluation demonstrated improved model performance using the three non-guideline methods relative to the regulatory default modeling methods.

The 1-hour SO_2 impacts from this facility were originally modeled in 2016 as part of Round 3 of the 2010 SO_2 NAAQS designations. This modeling analysis was used to inform the siting of two ambient SO_2 monitors adjacent to the Alcoa Massena fence line, Alcoa East (Site 1) and Alcoa West (Site 2). The modeling approach used to site these two monitors was approved by EPA, and incorporated the regulatory guideline methods found in Appendix W. The Buoyant Line and Point (BLP) source option was used to model the two roof vents. The 36 potline stacks and the anode bake furnace were modeled as point sources.

After three years (2017-2019) of hourly SO_2 measurements from the two monitors, it was apparent that the regulatory guideline modeling methods used in the 2016 modeling analysis had greatly overpredicted the 1-hour SO_2 impacts from the Alcoa Massena facility. By evaluating output from the DEBUG option in AERMOD, it appeared that these overpredictions were due in part to an underestimation of the plume rise from the long buildings with fugitive heat release located at the Alcoa Massena facility. AERMOD's plume rise calculations do not account for the additional buoyancy due to the fugitive heat released from the top of the long, narrow potline buildings at this location.

The fugitive heat released by the facility to the atmosphere during the nighttime hours was estimated to be on the same order of magnitude as the daytime sensible heat flux due to incoming solar radiation. Based on the year-round steady-state operation of the aluminum smelter, both the daytime and the nighttime stability of the atmosphere above the Alcoa Massena facility would best be categorized as neutral or unstable.

This unique thermal profile above the Alcoa Massena facility presents a challenge when characterizing Alcoa's two roof vents as BUOYLINE source types. The BUOYLINE source type is essentially the BLP algorithms running inside AERMOD, and they do not utilize a user-defined vertical temperature profile data file for its meteorological input. Instead, BUOYLINE automatically assigns a stable lapse rate to the nighttime hours if the model's source type is set to 'rural', as is the case at Alcoa Massena. If the source type is set to 'urban', BUOYLINE assumes a neutral lapse rate during the nighttime hours, which is similar to the actual vertical temperature profile above the facility. However, there is no vertical upper limit to the plume when modeling the BUOYLINE as 'urban', which presents an issue.

NYSDEC and EPA Region 2 both agree that the emission sources at the facility should be modeled as 'rural' and that previous modeling analyses using the BUOYLINE source type had greatly overpredicted the 1-hour SO_2 concentrations at the two monitors. Instead of using the BUOYLINE rural source type to characterize the roof vents, the consultant proposed modeling each of the two roof vents as a series of four point sources. These eight modeled point sources would have the same stack height as the two actual roof vents. The exit velocity across the entire length of each roof vent opening would be conserved and effective stack diameters would be used in the modeling analysis. Using the DEBUG option, the consultant first evaluated the spatial distribution of 100-meter plume heights resulting from the use of the urban BUOYLINE approach. Section 3.3 of the November 2023 AECOM modeling report provides details on the testing methodology that was used to determine the appropriate number of point sources to adequately represent the two roof vents. The evaluation indicated that the modeled maximum SO_2 design concentration at the facility's ambient air boundary was similar using both of these modeling approaches. Based on this comparative modeling analysis, NYSDEC is seeking concurrence with EPA Region 2 to approve the categorization of each roof vent source as a series of four point sources.

The results from the statistical model performance tests conducted by AECOM are found in the attached document, "SO_2 Modeling Protocol for Alcoa Massena Operations - West Plant – Draft, Section 4 Only".

The default AERMOD analysis incorporated the original BPIPPRM and the original AERMET Profile file with a single observation level at the Massena Airport (KMSS). The AERMOD analysis (Massena_MOD) included the draft version of BPIPPRM_19191 and an AERMET Profile file for the Massena Airport, which included a manually generated second observation level at 100 meters to simulate a neutral lapse rate throughout the lowest 100 meters of the atmosphere. The two roof vents were each categorized as four point sources. These performance evaluations include quantile-quantile (Q-Q) plots for each of the two monitors, design value comparisons between the modeled and observed 1-hour SO_2 impacts and utilizing the Model Evaluation Methodology outlined in EPA's 1992 Protocol for Determining the Best Performing Model.

The results from the Q-Q plots showed that the default model greatly overpredicted the SO_2 impacts at both monitors by at least a factor of 5. The Massena_MOD approach also overpredicted the impacts, however the overprediction was generally between 2 and 3 times the observed concentrations which indicates better performance than the default model. The design value comparison between the default model and Massena_MOD approach showed that both models overpredicted the SO_2 impacts. The modeled-to-observed design value ratios for the default model were between 5.4 and 6.8. The Massena_MOD approach showed better performance with modeled-to-observed design value ratios between 2.3 and 2.6.

EPA's Model Evaluation Methodology was used to analyze the Robust High Concentrations (RHC) for both monitors. Ideally, the ratio of model-predicted RHC to observed RHC should be around 1.0. The default model run showed ratios for the two monitors between 5.1 and 7.1. The Massena_MOD approach showed ratios between 2.1 and 2.4. When using a 90% confidence interval for the Composite Performance Measure, the analysis also indicated that there was a statistically significant difference in the performance between the two models. Finally, the Model Comparison Measure results for both monitors were evaluated and determined to all be positive, indicating that the Massena_MOD modeling approach performs better than the default modeling approach.

Based on the challenges presented when modeling the facility's roof vents as rural BUOYLINE source types, NYSDEC seeks concurrence from EPA Region 2 to approve the categorization of each roof vent source as a series of four point sources. The modeling files and statistical evaluation results are available for download on the NYSDEC File Transfer Service website (https://fts.dec.state.ny.us/fts/). A link and password to access these files will be sent separately via email. If EPA would also like to receive a USB drive containing the modeling files, please let me know.

Sincerely,

Margaret LaFarr
Assistant Director
Division of Air Resources

Ec: Joshua Cook (NYSDEC)
Julia Stuart (NYSDEC)
Bob Bielawa (NYSDEC)
Khai Gibbs (NYSDEC)
Marissa Logan (NYSDEC)
Jennifer Dougherty (NYSDEC)
Nathan Ruta (Alcoa)
Kathy Martin (Alcoa)
Bob Paine (AECOM)
Laura Warren (AECOM)
Neha Sareen (EPA)
Brian Marmo (EPA)
Allie Donohue (EPA)
Kirk Wieber (EPA)

Attachment 2 – Alcoa's 1-hr SO_2 Modeling Protocol

AECOM

SO_2 Modeling Protocol for Alcoa Massena Operations - West Plant

Alcoa Massena Operations - West Plant Massena, New York

Project number: 60726734

July 17, 2024

Delivering a better world

Quality information

Prepared by	Checked by	Approved by
Laura L. Warren Air Quality Scientist	Caryn Sutorus Air Quality Scientist	Robert J. Paine Associate Vice President, Air Quality

Prepared for:

Alcoa Massena Operations - West Plant Massena, New York

Prepared by:

Laura Warren
Air Quality Scientist
E: laura.warren2@aecom.com

AECOM
250 Apollo Drive
Chelmsford, MA 01824
aecom.com

Table of Contents

Figures

Tables

Executive Summary

The United States Environmental Protection Agency (EPA) established the 2010 1-hour National Ambient Air Quality Standard (NAAQS) for sulfur dioxide (SO_2) as 75 parts per billion (ppb). The form of the standard is the average of the 99th percentile of the daily maximum 1-hour average concentrations realized in each of three consecutive calendar years (the "design value," or DV). The EPA implemented the 2010 1-hour SO_2 NAAQS in an approach that involved either a dispersion modeling or monitoring approach to characterize local SO_2 concentrations near isolated emission sources. EPA's Data Requirements Rule (DRR) was finalized on August 21, 2015, and one of the sources in New York that is subject to the DRR provisions is the Alcoa Massena Operations West aluminum smelter (Massena Operations).

Massena Operations emits most of its SO_2 emissions from three clusters ("courtyards") of 12 dry scrubber stacks in the potline area, as well as from an anode bake furnace stack. After preliminary dispersion modeling with AERMOD using default modeling assumptions, it was clear that the monitoring path for the DRR implementation would be the better option due to the unique aspects of an aluminum smelter and heat releases not normally accounted for by the AERMOD model. A two-site monitoring network was developed in consultation with New York State Department of Environmental Conservation (NYSDEC) and approved by EPA. This network began monitoring in 2017 with one monitor to the northeast of the smelter (Site 1), and the other to the west (Site 2).

Through December 2019, the monitored concentrations at Site 1 exceeded the 75 ppb SO_2 NAAQS hourly average for only one hour over the entire 3-year period. As a result, its design value is well below the NAAQS. At Site 2,[1] which is closer to the potline dry scrubber stacks, the monitored concentrations exceeded the NAAQS several times each year, leading to a 3-year design concentration (2017-2019) above the NAAQS (86 ppb, about 15% above the NAAQS)[2]. It is important to note that monitored concentrations show a low peak-to-mean ratio. Over the initial 3-year monitoring period and the follow-on nearly 4 years of additional monitoring, there have been zero 5-minute average SO_2 concentrations above the 200-ppb benchmark upon which the 2010 SO_2 NAAQS is based, and infrequent, modest exceedances of the 1-hour 75 ppb NAAQS. Peak 5-minute concentrations are significant because the 1-hour SO_2 NAAQS was set to 75 ppb as a surrogate for protecting a peak 5-minute concentrations exposure threshold as low as 200 ppb for health effects, as noted in the 2010 SO_2 NAAQS rule (75 FR 35520). In fact, for Massena, the peak 5-minute concentrations at both monitors have never exceeded 150 ppb during the nearly 7 years of continuous monitoring, due in large part to the relatively steady operations and throughput from Massena Operations. The smelter operation leads to low peak-to-mean ratios for the SO_2 5-minute to 1-hour concentrations near the smelter and therefore results in peak 5-minute concentrations that are well below the 200-ppb benchmark level of concern for sensitive populations.

Due to the reported concentrations at Site 2 and following the designation of the nonattainment area (NAA), NYSDEC worked with Alcoa to plan for the next steps, which include:

- Providing a site-specific modeling approach to attempt to replicate the monitored concentrations at the two monitoring sites, and
- Using the model with proposed future source configuration to demonstrate attainment of the SO_2 NAAQS.

It is important for purposes of resolving the NAA that an unbiased and reliable dispersion model is used because the modeling results are used to determine the compliance status of the future source configuration and emissions. If the model is biased high, then the source may be forced to overcontrol or over-engineer a solution to the NAA issue that is unnecessary and goes beyond what is required to achieve the protections to health intended by the NAAQS. The monitoring data for the facility confirm that the AERMOD default model overpredicts by a least a factor of 5 at both monitors using actual emissions and concurrent meteorology for a 5-year period compared to monitored concentrations. Therefore, Alcoa has been working on a series of site-specific modeling approaches with enhancements applied to AERMOD that reduce the overprediction of AERMOD at this facility.

[1] Site 2 is located on Alcoa controlled property, near the property boundary.
[2] Note that Site 2 has had exceedances of the 75 ppb NAAQS on only 3 days in 2023 year-to-date (as of 11/12/2023). As such, the 4th highest peak daily 1-hour maximum for Site 2, year-to-date, is below 75 ppb.

This modeling protocol, updated from previous versions submitted in June 2022, June 2023, and November 2023, presents a site-specific dispersion modeling approach for documenting AERMOD-predicted 1-hour SO_2 concentration patterns resulting from the Massena Operations potline, anode bake furnace operations, and other SO_2 emitting sources with current stacks as well as with a potential change in the future stack configuration for the potline dry scrubber stacks that will be designed to attain the NAAQS.

The June 2022 modeling protocol was approved by EPA Region 2 with concurrence from EPA Model Clearinghouse on October 14, 2022. In the 2022 modeling protocol, which utilized a 3-year modeling period (2017-2019), the approved modeling approach involved site-specific source characterization approaches such as rural characterization, a modification of the thermal temperature profile above the stacks due to the fugitive heat losses from the plant, and building downwash pre-processor enhancements available for the AERMOD modeling system ("Massena_MOD"). In anticipation of taller potline dry scrubber stacks in the future, and after the October 2022 EPA approval of the June 2022 modeling protocol, NYSDEC and EPA requested an update to the protocol in order to revise the manner in which the potline roof vent emissions are characterized in the modeling. Given that a future stack arrangement is planned to include raising the potline dry scrubber stack heights, NYSDEC and EPA no longer considered the June 2022 proposed approach, where potline roof vent emissions were modeled as being emitted out of the potline dry scrubber stacks, to be appropriate. Instead, the reviewing agencies required these sources to be modeled separately from the dry scrubber stacks as proposed in this modeling protocol.

While Alcoa was evaluating the requested adjustments to the June 2022 protocol to address the roof vent emissions, it became clear after additional engineering review that the considered future 65-m stacks for each potline dry scrubber stack courtyard would impart risks to potline operation and trade-offs to reactor performance metrics. The consideration of feasible engineering solutions then brought up the issue of the residual overprediction tendency of the site-specific model proposed in the June 2022 modeling protocol, which did not include the "LIFTOFF" source characterization technique proposed in 2021 that showed significantly improved model performance, but which was not accepted by EPA as of early 2022. In June 2023, Alcoa proposed another approach using an "ENTRAIN" model option that also showed improved model performance, but this approach was not accepted for this modeling application by EPA and NYSDEC as communicated by NYSDEC in August 2023. However, the recent "ENTRAIN" approach was mentioned in the November 2023 protocol and this protocol to document Alcoa's view regarding its scientific merit and superior model performance to accurately and reliably predict SO_2 concentrations in ambient air near the facility. Additionally, discussion of the "Massena_MOD+ENTRAIN" and "Massena_MOD+ENTRAIN 0.2" approaches, presented in the June 2023 protocol, are included herein in **Appendix B** as these approaches have shown improved model evaluation performance relative to Massena_MOD. It is also important to realize that AERMOD's building downwash algorithms do not account for plume meander effects in low winds. This could be a factor in why the default model and Massena_MOD show overprediction tendencies for low wind conditions, and improved model performance could be realized with implementation of a downwash meander formulation.

This protocol focuses upon the Massena_MOD modeling approach for Alcoa Massena presented in the June and November 2023 protocols with an updated roof vent characterization. This updated protocol provides model evaluation results for a 5-year period for several aspects of the model performance at the two monitoring sites: 1) quantile-quantile plots of the modeled and observed concentrations, 2) meteorological conditions associated with the peak 25 modeled and observed concentrations, and 3) a full evaluation using EPA's model evaluation protocol procedures as implemented using the Model Evaluation Methodology software. For the Site 2 monitor, the model evaluation results indicate that the ratio of the model to observed design concentration with a 5-year dataset improves from 5.4 with AERMOD default to 2.6 with Massena_MOD. While the model performance for Massena_MOD is an improvement upon AERMOD default, Alcoa notes that it still exhibits substantial overprediction. Despite this overprediction tendency, the reviewing agencies are requiring Alcoa Massena to utilize Massena_MOD. As a result, Alcoa currently is proceeding with an engineering approach to work with Massena_MOD, even though it results in an over-engineered stack design solution to resolve the NAAQS issue. The future attainment strategy modeling will demonstrate that the proposed stack design, using Massena_MOD, as outlined in Section 5 of this protocol, will result in future monitored SO_2 concentrations, including those at Site 2, that are below the NAAQS, and provide a resolution for the nonattainment area.

1. Introduction

1.1 Background

The United States Environmental Protection Agency (EPA) promulgated a 1-hour National Ambient Air Quality Standard (NAAQS) for sulfur dioxide (SO_2) in 2010. The 1-hour SO_2 NAAQS is set to 75 parts per billion (ppb) and the form of the standard is the average of the 99th percentile of the daily maximum 1-hour average concentrations realized in each of three consecutive calendar years (the "design value," or DV).

The EPA implemented the 2010 1-hour SO_2 NAAQS in an approach that involved either a dispersion modeling or monitoring approach to characterize local SO_2 concentrations near isolated emission sources. EPA's Data Requirements Rule (DRR) was finalized on August 21, 2015, and one of the sources in New York that is subject to the DRR provisions is the Alcoa Massena Operations West aluminum smelter (Massena Operations).

Massena Operations emits most of its SO_2 emissions from three clusters ("courtyards") of 12 dry scrubber stacks in the potline area, as well as from an anode bake furnace stack (see **Figure 1-1**; see **Appendix A** for more detail on SO_2 source locations). It was clear after preliminary dispersion modeling with AERMOD using default modeling assumptions that the monitoring path for the DRR implementation would be the better option due to the unique aspects of an aluminum smelter and heat releases not normally accounted for by the AERMOD model.

Based upon an examination of predominant winds and available sites for monitoring, the New York State Department of Environmental Conservation (NYSDEC) elected to use a 2-monitor network to characterize the SO_2 concentrations in the vicinity of the smelter. This monitoring network, with Site 1 to the northeast and Site 2 to the west, was initiated in January 2017. The monitoring sites are shown on a map in **Figure 1-2** with an indication of the contiguous property boundary (which differs from the ambient air boundary; see **Section 2**).

Through December 2019, the monitored concentrations at Site 1 exceeded the 75 ppb SO_2 NAAQS hourly average for only one hour over the entire 3-year period. As a result, the 99th percentile peak daily 1-hour maximum concentrations are well below the NAAQS at that site. At Site 2, which is closer to the potline dry scrubber stacks, the monitored concentrations modestly exceeded the NAAQS several times each year, leading to a 3-year design concentration (2017-2019) above the NAAQS (86 ppb, about 15% above the NAAQS).[3] As noted in the footnote, the relatively low peak 5-minute concentrations detected at the monitor have resulted in only a modest excursion of the NAAQS at the most-impacted monitor, Site 2.

Due to the reported concentrations at Site 2 and following the determination of the nonattainment area (NAA), NYSDEC worked with Alcoa to plan for the next steps, which include:

- Providing a site-specific modeling approach to attempt to replicate the monitored concentrations at the two monitoring sites, and
- Using the model with the future smelter configuration to demonstrate attainment of the SO_2 NAAQS.

It is important for purposes of resolving the NAA that a relatively unbiased dispersion model is used because the modeling results are used to determine the compliance status of the future source configuration and emissions. If the model is biased high, then the source may be forced to overcontrol or over-engineer a solution to the NAA issue that is unnecessary and overly burdensome. Due to the remaining high bias for the Massena_MOD model that is being submitted for approval herein, Alcoa is planning a stack merging and extension project for the potline dry scrubber stacks in order to demonstrate NAAQS attainment with this model.

[3] However, peak 5-minute concentrations at both monitors are always below EPA's 200 ppb threshold for health effects, as noted in the 2010 SO_2 NAAQS rule (75 FR 35520). In fact, for Massena, the peak 5-minute concentrations at both monitors never exceeded 150 ppb, due in large part to the relatively steady operations and throughput from Massena Operations. The smelter operation leads to low peak-to-mean ratios for the SO_2 5-minute to 1-hour concentrations near the smelter and therefore results in peak 5-minute concentrations that are well below the health benchmark upon which the NAAQS is based.

This modeling protocol presents a proposed site-specific dispersion modeling approach for documenting AERMOD-predicted 1-hour SO_2 concentration patterns resulting from Massena Operations potlines, anode bake furnace operations, and other SO_2 emitting sources with current stacks as well as with a change in the future stack arrangement for the potline dry scrubber stacks (also referred to as reactor stacks) that will be designed to attain the NAAQS. Other aluminum smelters in the United States (e.g., Alcoa's Warrick and former Intalco smelters) have also adopted, with EPA approval, site-specific modeling approaches due to the unique characteristics of the emission source.

This modeling protocol builds off of the modeling protocol (dated June 10, 2022) approved by EPA Region 2 with concurrence from EPA Model Clearinghouse on October 14, 2022.[4] In the 2022 modeling protocol, the approved modeling approach involved site-specific source characterization approaches such as rural characterization, a modification of the thermal temperature profile above the stacks due to the fugitive heat losses from the plant, and building downwash pre-processor enhancements available for the AERMOD modeling system. An update to the protocol is required by EPA in order to revise the manner in which the potline roof vent emissions are characterized in the modeling. Given that a future stack arrangement is planned to include raising the potline dry scrubber stack heights, NYSDEC and EPA no longer consider the original approach, where potline roof vent emissions were modeled as being emitted out of the potline dry scrubber stacks, to be appropriate. Instead, these sources will be modeled separately from the dry scrubber stacks. In the June 2023 and November 2023 protocols, the potline roof vents were characterized as a series of stacks along each vent. Based on the reviewing agencies' feedback, this protocol revises the methodology to calculate the roof vent stack parameters such that roof vent exit velocity and the roof vent opening area are conserved, as required by the agencies. This protocol's model evaluation results were provided to NYSDEC in June 2024 in the form of modeling files and an early release of Section 4 text.

1.1.1 Summary of the Previously Proposed ENTRAIN Modeling Approach

After application of the neutral lapse rate and the site-specific BPIP processing that EPA approved for "Massena_MOD", model performance continued to indicate an overprediction ratio generally between 2 and 3 for the controlling monitor. This level of overprediction is more than the factor-of-2 bias that is referenced by the screening criteria stated in EPA's model evaluation protocol guidance.[5] In an effort to improve the performance of the site-specific model, Alcoa proposed an adjustment to the plume rise characterization through the use of a lower entrainment coefficient (β) to represent the slower rate at which ambient air is ingested into the plumes in close proximity, especially when surrounded by fugitive heat releases, which affects final plume rise. The adjustment is accomplished by applying AERMOD's AWMADWNW ENTRAIN model option and considering a lower plume rise entrainment coefficient best suited for the Massena Operations smelter. In 2021, ENTRAIN was adopted by EPA in AERMOD version 21112 as an alpha option. Since that time, a peer-reviewed journal article[6] was published in the Journal of A&WMA (December 2022) that further supports this option, advancing it closer to beta status. By including the ENTRAIN option for Massena Operations, results indicated improved model performance, especially at Site 2 where measured SO_2 concentrations are the highest.

As communicated by NYSDEC, NYSDEC and EPA did not accept the use of the ENTRAIN option for Alcoa Massena due in part to the fact that this option is still an alpha option, even though Appendix W's Section 3.2.2(b)(2) – the applicable option for approval of an alternative model in this case – makes no mention of the "alpha" or "beta" option issue. Alcoa believes that this option may in the future be further considered for improved plume rise estimates, especially in cases with fugitive heat releases. Accordingly, further discussion of the ENTRAIN option and why it can result in improved model performance for Alcoa Massena is provided in **Appendix B** to provide supporting documentation for this modeling approach. It is also important to realize that AERMOD's building downwash algorithms do not account for plume meander effects in low winds. This could be a factor in why the default model

[4] EPA, 2022. Model Clearinghouse Information Storage and Retrieval System: Case-specific Alternative Approaches to Demonstrate Modeled Attainment of the 2010 SO_2 NAAQS at the Alcoa West Aluminum Smelter in Massena, New York. https://cfpub.epa.gov/oarweb/MCHISRS/index.cfm?fuseaction=main.resultdetails&recnum=22-II-03

[5] EPA, 1992. Protocol for Determining the Best Performing Model. Publication No. EPA–454/R–92–025. Office of Air Quality Planning and Standards, Research Triangle Park, NC. (NTIS No. PB 93–226082). https://www.epa.gov/sites/default/files/2020-10/documents/model_eval_protocol.pdf.

[6] Petersen R. L., J. O. Paumier, and S. A. Guerra, 2022. Development, evaluation, and implementation of building downwash and plume rise enhancements in AERMOD. *J. Air Waste Manag. Assoc*.,72:12,1423-1441. https://doi.org/10.1080/10962247.2022.2120563.

and Massena_MOD show overprediction tendencies for low wind conditions, and improved model performance could also be realized with implementation of a downwash meander formulation.

1.3 Document Organization

Section 2 provides a discussion of the SO_2 emission sources at Massena Operations, as well as a review of nearby SO_2 emission sources. The proposed approaches for modeling Massena Operations as a result of consultation with NYSDEC and EPA is provided in Section 3, including a brief history of other peer-reviewed modeling approaches that were proposed by Alcoa and not accepted by the reviewing agencies. Section 4 describes the results of a model evaluation to determine the better performing model (between the default AERMOD approach and the "Massena_MOD" approach). Section 5 describes how the best performing model will be used to demonstrate that the proposed stack changes will resolve the current nonattainment area. Appendix A includes facility maps identifying the locations and names of modeled sources and buildings. Appendix B includes the use of the "Massena_MOD+ENTRAIN" and "Massena_MOD+ENTRAIN 0.2" approaches as presented in the June 2023 protocol, as these approaches have shown improved model evaluation performance relative to Massena_MOD.

Figure 1-1: Location of the Key SO_2 Sources at Massena Operations

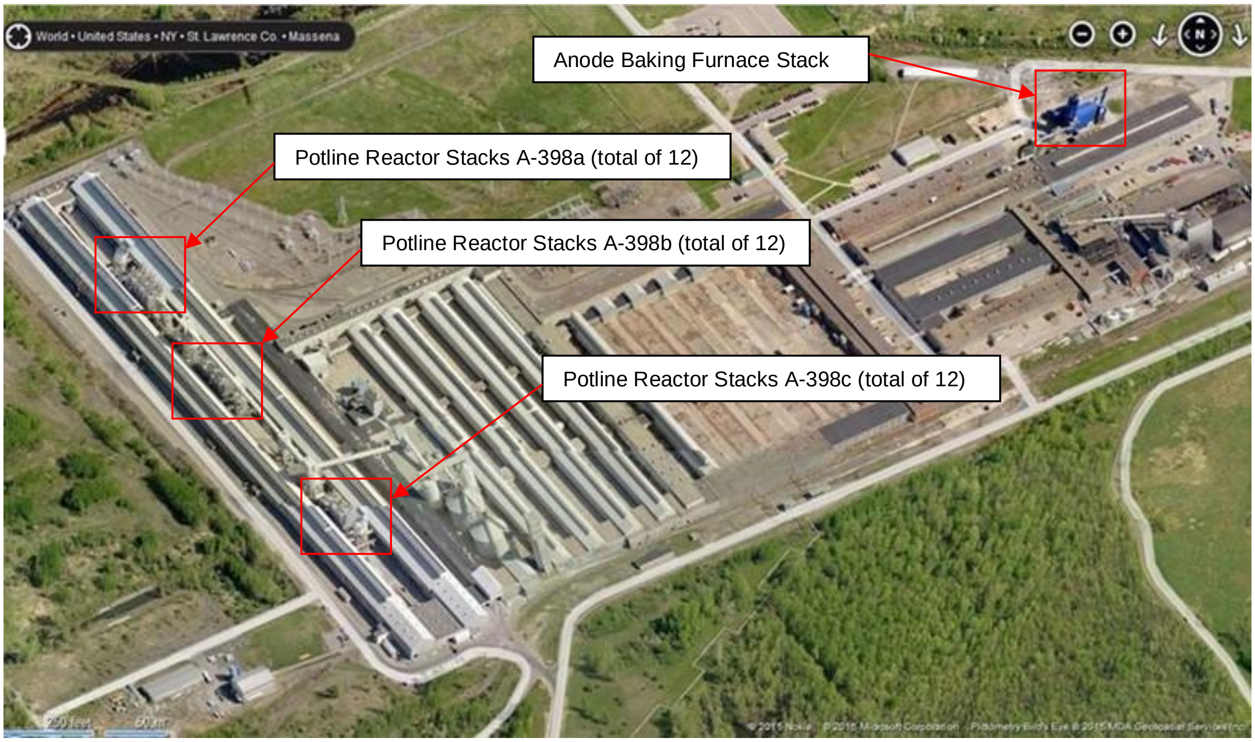
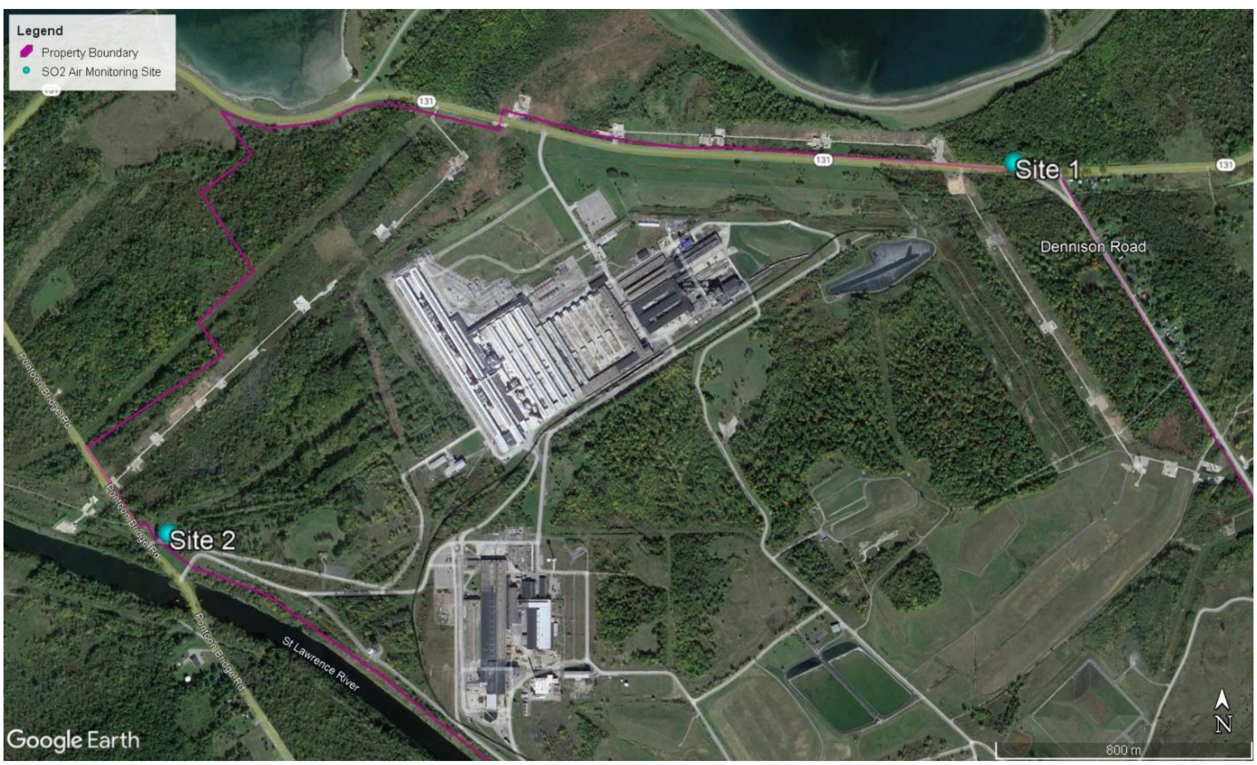


Figure 1-2: Location of the Two SO_2 Monitoring Sites



2. Description of SO_2 Emissions Sources

2.1 Massena West Smelter Operations

Massena Operations is an aluminum production facility with a capacity to produce 136,000 metric tons of primary aluminum per calendar year (full capacity). It is located in St. Lawrence County along the St. Lawrence Seaway, on the northern outskirts of the town of Massena, NY. The area surrounding the smelter is rural with simple terrain within several kilometers (km). The facility has one potline consisting of two long rooms with three clusters of dry scrubber stacks between them in the three “courtyard” areas. These potline dry scrubber stacks comprise most of the SO_2 emissions at the smelter, with one anode bake furnace stack also emitting SO_2, as shown in **Figure 1-1**. There are also several natural gas-fired furnaces, melters/holders, package boilers, and the potline roof vents (see **Figure 3-4**). Facility maps identifying the locations and names of modeled sources and buildings are provided in **Appendix A**.

The SO_2 emission rates that will be used in the modeling of the facility to characterize the concentrations observed at the monitors for the evaluation period of 2017-2021 were derived from the reported monthly averaged SO_2 emission rates for the potline systems (dry scrubber stacks and potline roof vents) and the monthly averaged SO_2 emission rates for the bake furnace dry scrubbing systems. All other SO_2-emitting sources use emissions set to their potential-to-emit emission rates (e.g., melters/holders, boilers). Stack exhaust parameters and typical SO_2 emission rates are shown in **Table 2-1** (Metric) and **Table 2-2** (English), where the dry scrubber/reactor stacks are labeled as “RS”, anode bake furnace as “BF2”, and potline roof vents as “B401” and “B402”. Emission sources where 2017-2021 monthly average emission rates will be used are provided in **Table 2-3, Table 2-4, and Table 2-5**.

SO_2 emissions in the Hall-Heroult aluminum smelting process at Massena Operations are primarily a product of sulfur contained in the carbon anodes being released during electrolysis. Carbon anodes are comprised mainly of coke as a raw material with a lessor amount of pitch as a binding agent. It is in the coke where the majority of sulfur lies (regulated as %S in coke). Carbon consumption is proportional to the total aluminum produced at a smelter. The production of aluminum at the facility is stable and the facility is considered to be operating at full capacity.

The SO_2 emissions from the roof vents on the potlines are very low, comprising no more than 5% of the total potline SO_2 emissions (in the range of 4 to 5% according to tests during the model evaluation period of 2017-2021). In the June 2022 modeling protocol, the roof vent emissions were added to the emissions for the dry scrubber stacks for modeling purposes. Both source types have current release heights that are similar to one another, and this approach was conservative in that it concentrates the roof vent emissions into the stacks with the vast majority of the emissions rather than dispersing the roof vent emissions over the long extent of the roof vents. However, since the planned future potline dry scrubber stack configuration will include higher stack heights, it would no longer be accurate to model the potline roof vent emissions through the potline dry scrubber stacks. This protocol includes an approach for modeling potline roof vent emissions. One possible approach for modeling the roof vents separately is with the BUOYLINE source type. However, the characterization of the lowest 100 meters (m) of the atmosphere with a neutral temperature lapse rate is an important issue for a successful modeling approach for this application. Unfortunately, the BUOYLINE source type, which is basically the Buoyant Line and Point (BLP) model inside AERMOD, has a key limitation in that it does not use the vertical profile of temperature data provided to AERMOD. Instead, with a rural source type, BUOYLINE sources would be modeled with a stable Pasquill-Gifford stability class at night, which uses a stable temperature lapse rate that is inconsistent with the intent of the site-specific modeling approach that was approved by EPA.[4] A more appropriate modeling approach for the roof vent emissions is to characterize the vents as a row of several (i.e., 4) equally-spaced point sources (stacks), with an exit velocity set equal to the actual roof vent velocity and an equivalent diameter that conserves the area of the total roof vent opening among the stacks by dividing the total roof vent opening area by the representative number of stacks, then calculating the equivalent diameter per stack. Therefore, the roof vent stacks will be modeled with rural dispersion using this source characterization.

Total potline emissions used in the modeling of the facility to characterize the concentrations observed at the monitors will be distributed by assigning 95% of total potline emissions to the dry scrubber stacks with the remaining 5% distributed to the potline roof vents, based on analyses conducted during the model evaluation period. **Table 2-1** and

Table 2-2 list the modeled stack parameters for the roof vent stacks, for which the roof vent exit velocity is retained at each stack while equivalent stack diameters are calculated from the total roof vent opening area divided by the number of stacks (the stack diameters were calculated to conserve the area of the total roof vent opening). The number of stacks per roof vent was determined using a modeling analysis presented in Section 3.3, where 4 stacks per roof vent (emissions in **Table 2-4**) will be used for Massena_MOD.

Figure 2-1 shows the location of the modeled ambient air boundary (red outlined area) around the facility within which Massena Operations controls public access and which is not intersected by a public road. It includes a blue outlined area to which access is controlled by Alcoa and/or Arconic. Public roadways will not be excluded from ambient air and will therefore be modeled. A robust system is in place to control public access to the property with security check points, signage, security cameras, and roadways patrols in accordance with the EPA ambient air guidance.

Table 2-1: Current Exhaust Parameters for SO_2 Emission Sources (Metric Units)

ID	Base Elevation (m)	Release Height (m)	Stack Diameter (m)	Exit Velocity (m/s)	Exit Temperature (K)	SO_2 Emission Rate (g/s)
RS_A1[1]	65.5	23.3	1.22	10.37	378.3	See **Table 2-3**
RS_A2[1]	65.5	23.3	1.22	10.37	378.3	
RS_A3[1]	65.5	23.3	1.22	10.37	378.3	
RS_A4[1]	65.5	23.3	1.22	10.37	378.3	
RS_A5[1]	65.5	23.3	1.22	10.37	378.3	
RS_A6[1]	65.5	23.3	1.22	10.37	378.3	
RS_A7[1]	65.5	23.3	1.22	10.37	378.3	
RS_A8[1]	65.5	23.3	1.22	10.37	378.3	
RS_A9[1]	65.5	23.3	1.22	10.37	378.3	
RS_A10[1]	65.5	23.3	1.22	10.37	378.3	
RS_A11[1]	65.5	23.3	1.22	10.37	378.3	
RS_A12[1]	65.5	23.3	1.22	10.37	378.3	
RS_B1[1]	65.5	23.3	1.22	10.37	378.3	
RS_B2[1]	65.5	23.3	1.22	10.37	378.3	
RS_B3[1]	65.5	23.3	1.22	10.37	378.3	
RS_B4[1]	65.5	23.3	1.22	10.37	378.3	
RS_B5[1]	65.5	23.3	1.22	10.37	378.3	
RS_B6[1]	65.5	23.3	1.22	10.37	378.3	
RS_B7[1]	65.5	23.3	1.22	10.37	378.3	
RS_B8[1]	65.5	23.3	1.22	10.37	378.3	
RS_B9[1]	65.5	23.3	1.22	10.37	378.3	
RS_B10[1]	65.5	23.3	1.22	10.37	378.3	
RS_B11[1]	65.5	23.3	1.22	10.37	378.3	
RS_B12[1]	65.5	23.3	1.22	10.37	378.3	
RS_C1[1]	65.5	23.3	1.22	10.37	378.3	
RS_C2[1]	65.5	23.3	1.22	10.37	378.3	
RS_C3[1]	65.5	23.3	1.22	10.37	378.3	
RS_C4[1]	65.5	23.3	1.22	10.37	378.3	
RS_C5[1]	65.5	23.3	1.22	10.37	378.3	
RS_C6[1]	65.5	23.3	1.22	10.37	378.3	
RS_C7[1]	65.5	23.3	1.22	10.37	378.3	
RS_C8[1]	65.5	23.3	1.22	10.37	378.3	
RS_C9[1]	65.5	23.3	1.22	10.37	378.3	
RS_C10[1]	65.5	23.3	1.22	10.37	378.3	
RS_C11[1]	65.5	23.3	1.22	10.37	378.3	
RS_C12[1]	65.5	23.3	1.22	10.37	378.3	
BF2[1]	66.3	32.0	2.13	15.99	357.4	See **Table 2-5**
B401_01[1,2]	65.5	20.3	26.5	0.94	305.8	See **Table 2-4**
B401_02[1,2]	65.5	20.3	26.5	0.94	305.8	

ID	Base Elevation (m)	Release Height (m)	Stack Diameter (m)	Exit Velocity (m/s)	Exit Temperature (K)	SO_2 Emission Rate (g/s)
B401_03[1,2]	65.5	20.3	26.5	0.94	305.8	See **Table 2-4**
B401_04[1,2]	65.5	20.3	26.5	0.94	305.8	
B402_01[1,2]	65.5	20.3	26.5	0.94	305.8	See **Table 2-4**
B402_02[1,2]	65.5	20.3	26.5	0.94	305.8	
B402_03[1,2]	65.5	20.3	26.5	0.94	305.8	
B402_04[1,2]	65.5	20.3	26.5	0.94	305.8	
AI043[3]	78.3	15.9	1.37	3.19	444.8	2.55E-03
AI046[3]	78.3	18.9	1.37	3.19	444.8	2.72E-03
M0031[3]	78.3	22.6	1.27	17.66	487.6	3.18E-03
M024D[3]	78.3	22.6	1.27	10.30	487.6	3.18E-03
M024F[3]	78.3	25.0	2.01	4.14	487.6	2.42E-03
M0034[3]	78.3	22.6	0.81	4.94	396.5	7.62E-04
M0035[3]	78.3	22.6	0.81	4.94	396.5	7.62E-04
M024E[3]	78.3	14.0	0.20	25.07	338.7	8.63E-05
AI026[3]	78.3	18.9	1.37	3.19	444.8	2.27E-03
B000A_B[3]	66.3	13.7	0.72	15.30	383.2	1.50E-03

Notes:

(1) The model evaluation will use calculated monthly emissions for the period of January 2017 through December 2021 based on mass balance assumptions.
(2) Potline roof vents will be modeled as a series of 4 stacks per roof vent.
(3) This source will use its potential-to-emit emission rate for the model evaluation.

Definitions:

RS = Potline dry scrubber stacks/reactor stacks
BF2 = Anode bake furnace
B401/B402 = Potline roof vents, modeled as a series of stacks along each roof vent
AI = Homogenizing heat treat furnaces
M0 = Melters/holders; furnaces (M024E)
B000A_B = Package boilers

Table 2-2: Current Exhaust Parameters for SO_2 Emission Sources (English Units)

ID	Base Elevation (ft)	Release Height (ft)	Stack Diameter (ft)	Exit Velocity (ft/s)	Exit Temperature (F)	SO_2 Emission Rate (lb/hr)
RS_A1[1]	214.9	76.4	4.00	34.02	221.3	See **Table 2-3**
RS_A2[1]	214.9	76.4	4.00	34.02	221.3	
RS_A3[1]	214.9	76.4	4.00	34.02	221.3	
RS_A4[1]	214.9	76.4	4.00	34.02	221.3	
RS_A5[1]	214.9	76.4	4.00	34.02	221.3	
RS_A6[1]	214.9	76.4	4.00	34.02	221.3	
RS_A7[1]	214.9	76.4	4.00	34.02	221.3	
RS_A8[1]	214.9	76.4	4.00	34.02	221.3	
RS_A9[1]	214.9	76.4	4.00	34.02	221.3	
RS_A10[1]	214.9	76.4	4.00	34.02	221.3	
RS_A11[1]	214.9	76.4	4.00	34.02	221.3	
RS_A12[1]	214.9	76.4	4.00	34.02	221.3	
RS_B1[1]	214.9	76.4	4.00	34.02	221.3	
RS_B2[1]	214.9	76.4	4.00	34.02	221.3	
RS_B3[1]	214.9	76.4	4.00	34.02	221.3	
RS_B4[1]	214.9	76.4	4.00	34.02	221.3	
RS_B5[1]	214.9	76.4	4.00	34.02	221.3	
RS_B6[1]	214.9	76.4	4.00	34.02	221.3	
RS_B7[1]	214.9	76.4	4.00	34.02	221.3	
RS_B8[1]	214.9	76.4	4.00	34.02	221.3	
RS_B9[1]	214.9	76.4	4.00	34.02	221.3	
RS_B10[1]	214.9	76.4	4.00	34.02	221.3	
RS_B11[1]	214.9	76.4	4.00	34.02	221.3	
RS_B12[1]	214.9	76.4	4.00	34.02	221.3	
RS_C1[1]	214.9	76.4	4.00	34.02	221.3	
RS_C2[1]	214.9	76.4	4.00	34.02	221.3	
RS_C3[1]	214.9	76.4	4.00	34.02	221.3	
RS_C4[1]	214.9	76.4	4.00	34.02	221.3	
RS_C5[1]	214.9	76.4	4.00	34.02	221.3	
RS_C6[1]	214.9	76.4	4.00	34.02	221.3	
RS_C7[1]	214.9	76.4	4.00	34.02	221.3	
RS_C8[1]	214.9	76.4	4.00	34.02	221.3	
RS_C9[1]	214.9	76.4	4.00	34.02	221.3	
RS_C10[1]	214.9	76.4	4.00	34.02	221.3	
RS_C11[1]	214.9	76.4	4.00	34.02	221.3	
RS_C12[1]	214.9	76.4	4.00	34.02	221.3	
BF2[1]	217.5	105.0	6.99	52.46	183.7	See **Table 2-5**
B401_01[1,2]	214.9	66.6	86.94	3.08	90.8	See **Table 2-4**
B401_02[1,2]	214.9	66.6	86.94	3.08	90.8	

ID	Base Elevation (ft)	Release Height (ft)	Stack Diameter (ft)	Exit Velocity (ft/s)	Exit Temperature (F)	SO_2 Emission Rate (lb/hr)
B401_03[1,2]	214.9	66.6	86.94	3.08	90.8	See **Table 2-4**
B401_04[1,2]	214.9	66.6	86.94	3.08	90.8	
B402_01[1,2]	214.9	66.6	86.94	3.08	90.8	See **Table 2-4**
B402_02[1,2]	214.9	66.6	86.94	3.08	90.8	
B402_03[1,2]	214.9	66.6	86.94	3.08	90.8	
B402_04[1,2]	214.9	66.6	86.94	3.08	90.8	
AI043[3]	256.9	52.2	4.49	10.47	341.0	2.02E-02
AI046[3]	256.9	62.0	4.49	10.47	341.0	2.16E-02
M0031[3]	256.9	74.1	4.17	57.94	418.0	2.52E-02
M024D[3]	256.9	74.1	4.17	33.79	418.0	2.52E-02
M024F[3]	256.9	82.0	6.59	13.58	418.0	1.92E-02
M0034[3]	256.9	74.1	2.66	16.21	254.0	6.05E-03
M0035[3]	256.9	74.1	2.66	16.21	254.0	6.05E-03
M024E[3]	256.9	45.9	0.66	82.25	150.0	6.85E-04
AI026[3]	256.9	62.0	4.49	10.47	341.0	1.80E-02
B000A_B[3]	217.5	44.9	2.36	50.20	230.1	1.19E-02

Notes:

(1) The model evaluation will use calculated monthly emissions for the period of January 2017 through December 2021 based on mass balance assumptions.
(2) Potline roof vents will be modeled as a series of 4 stacks per roof vent.
(3) This source will use its potential-to-emit emission rate for the model evaluation.

Definitions:

RS = Potline dry scrubber stacks/reactor stacks
BF2 = Anode bake furnace
B401/B402 = Potline roof vents, modeled as a series of stacks along each roof vent
AI = Homogenizing heat treat furnaces
M0 = Melters/holders; furnaces (M024E)
B000A_B = Package boilers

Table 2-3: Monthly Average SO_2 Emissions Per Stack for the Potline Dry Scrubber Stacks

Month	Emissions (g/s)					Emissions (lb/hr)				
	2017	2018	2019	2020	2021	2017	2018	2019	2020	2021
January	1.749	1.758	1.788	1.732	1.732	13.88	13.95	14.19	13.74	13.75
February	1.841	1.820	1.822	1.815	1.963	14.61	14.44	14.46	14.40	15.58
March	1.749	1.788	1.765	1.831	1.856	13.88	14.19	14.01	14.54	14.73
April	1.828	1.718	1.838	1.821	1.850	14.51	13.63	14.59	14.46	14.68
May	1.835	1.752	1.659	1.792	1.764	14.57	13.91	13.17	14.23	14.00
June	1.768	1.679	1.793	1.800	1.707	14.03	13.33	14.23	14.29	13.55
July	1.787	1.777	1.871	1.891	1.731	14.18	14.11	14.85	15.01	13.74
August	1.698	1.653	1.864	1.975	1.767	13.48	13.12	14.79	15.67	14.02
September	1.728	1.832	1.873	1.717	1.660	13.72	14.54	14.86	13.63	13.17
October	1.757	1.805	1.760	1.651	1.619	13.94	14.32	13.97	13.10	12.85
November	1.671	1.757	1.709	1.748	1.609	13.26	13.95	13.57	13.87	12.77
December	1.703	1.809	1.764	1.651	1.609	13.52	14.35	14.00	13.10	12.77

Note: There is a total of 36 dry scrubber stacks at the smelter.

Table 2-4: Monthly Average SO_2 Emissions Per Potline Roof Vent Representative Stack

Month	Emissions (g/s)					Emissions (lb/hr)				
	2017	2018	2019	2020	2021	2017	2018	2019	2020	2021
January	0.414	0.416	0.423	0.410	0.410	3.29	3.30	3.36	3.26	3.26
February	0.436	0.431	0.432	0.430	0.465	3.46	3.42	3.43	3.41	3.69
March	0.414	0.424	0.418	0.434	0.440	3.29	3.36	3.32	3.44	3.49
April	0.433	0.407	0.435	0.431	0.438	3.44	3.23	3.45	3.42	3.48
May	0.435	0.415	0.393	0.425	0.418	3.45	3.29	3.12	3.37	3.32
June	0.419	0.398	0.425	0.426	0.404	3.32	3.16	3.37	3.38	3.21
July	0.423	0.421	0.443	0.448	0.410	3.36	3.34	3.52	3.55	3.25
August	0.402	0.392	0.441	0.468	0.419	3.19	3.11	3.50	3.71	3.32
September	0.409	0.434	0.444	0.407	0.393	3.25	3.44	3.52	3.23	3.12
October	0.416	0.427	0.417	0.391	0.383	3.30	3.39	3.31	3.10	3.04
November	0.396	0.416	0.405	0.414	0.381	3.14	3.30	3.21	3.28	3.02
December	0.403	0.428	0.418	0.391	0.381	3.20	3.40	3.32	3.10	3.02

Table 2-5: Monthly Average SO_2 Emissions for the Anode Bake Furnace Stack

Month	Emissions (g/s)					Emissions (lb/hr)				
	2017	2018	2019	2020	2021	2017	2018	2019	2020	2021
January	2.503	2.275	2.674	2.499	2.640	19.87	18.06	21.23	19.83	20.96
February	2.363	2.607	2.544	2.502	2.598	18.75	20.69	20.19	19.85	20.62
March	2.121	2.628	2.554	2.550	2.589	16.83	20.85	20.27	20.24	20.55
April	2.382	2.124	2.516	2.446	2.436	18.91	16.86	19.97	19.41	19.33
May	2.598	2.273	2.485	2.426	2.442	20.62	18.04	19.72	19.26	19.38
June	2.391	2.581	2.604	2.402	2.412	18.98	20.49	20.66	19.06	19.15
July	2.497	2.507	2.429	2.339	2.441	19.82	19.90	19.28	18.57	19.37
August	2.311	2.597	2.422	2.408	2.447	18.34	20.61	19.22	19.11	19.42
September	2.327	2.588	2.434	2.421	2.460	18.47	20.54	19.32	19.21	19.52
October	2.536	2.375	2.575	2.504	2.658	20.13	18.85	20.43	19.87	21.09
November	2.627	2.331	2.301	2.588	2.419	20.85	18.50	18.26	20.54	19.20
December	2.273	2.378	2.430	2.640	2.436	18.04	18.87	19.29	20.95	19.34

2.2 Nearby SO_2 Emission Sources

A review of nearby SO_2 sources shows that there are no sources within 50 km of the facility with SO_2 emissions greater than 1 ton per year using the 2020 EPA National Emissions Inventory.[7] Therefore, no sources other than Massena Operations will be modeled. The effects of non-modeled sources are represented in regional background.

2.3 Regional Background

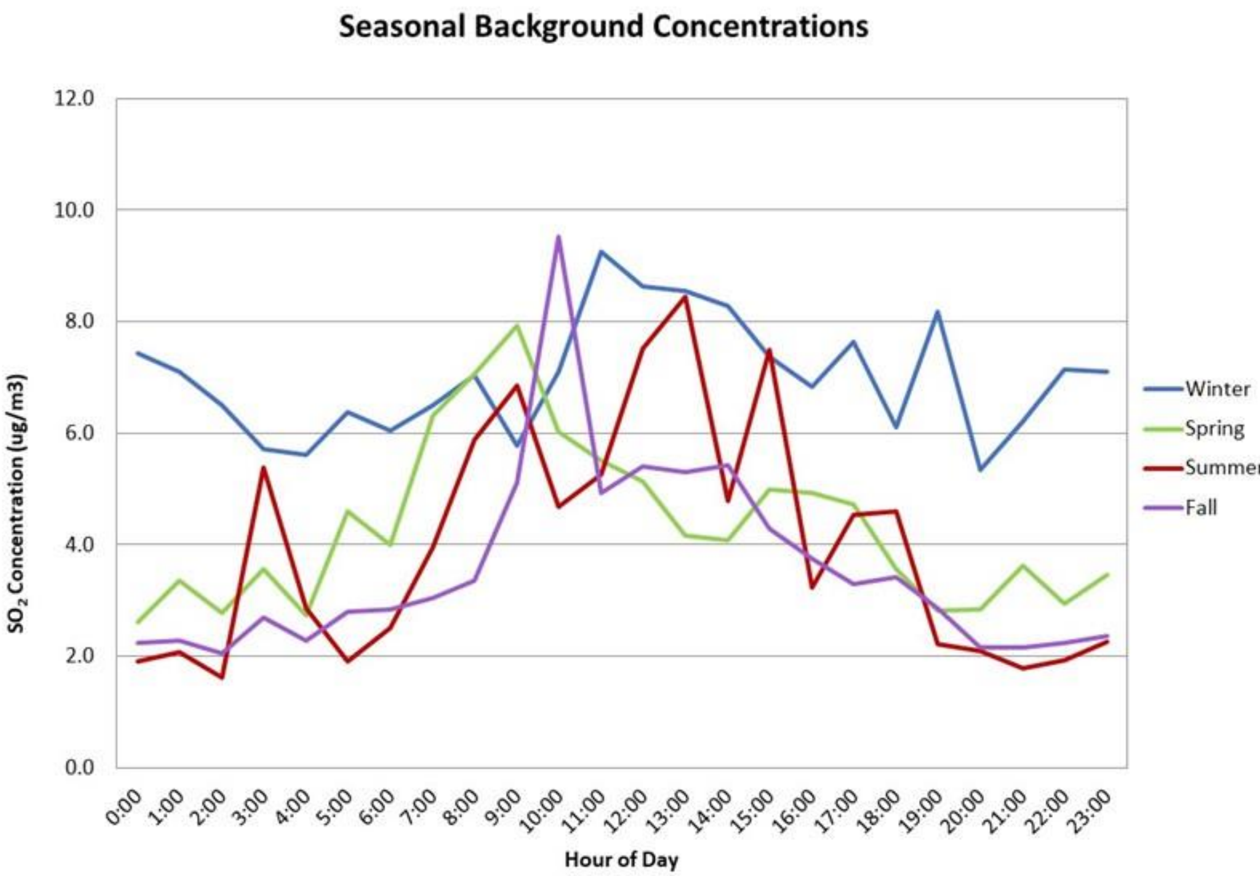
Regional background concentrations are used in modeling to represent emission sources that are not directly modeled, as well as naturally occurring levels of the pollutant of interest. Once regional background levels have been identified, they are added to the modeled results at each receptor for a cumulative modeling result.

As shown in **Figure 1-2**, there are two SO_2 monitors in the vicinity of Massena Operations. To create regional background values for the modeling, NYSDEC used the hourly SO_2 data from the two monitors for the December 2016-November 2019 time period to create a synthetic background monitor time series of concentrations. Beginning the dataset with December 2016 ensured that the "winter" season for each of the three years included contiguous monthly data from December through February by using the December data from the previous year. NYSDEC took the lower concentration of the two monitors for each hour when data was present for both monitors. If only one monitor recorded valid data for a particular hour, NYSDEC looked at the wind direction for that hour to determine if that monitor was impacted by emissions from Alcoa Massena Operations. It was determined that Site 1 was impacted by the smelter emissions when the wind direction was between 230° and 270°, and Site 2 was impacted by the smelter emissions when the wind direction was between 30° and 90°. If the wind was blowing from the facility towards either monitor with data for that hour, then the data for that monitor was not used, and the hour was considered missing for that monitor. Otherwise, the concentration was considered valid (as a background value) and was used in the determination of the hourly background value. Once a dataset of sequential hourly background values was created, values for the season / hour-of-day background approach (as discussed in EPA's March 1, 2011 guidance[8]) were computed. This approach was approved by EPA as documented in their SO_2 Round 4 Technical Support Document for New York.[9] **Figure 2-2** shows the season/hour-of-day values to be used in the modeling.

[7] EPA, 2020. National Emissions Inventory (NEI) Data: Online 2020 NEI Data Retrieval Tool. https://www.epa.gov/air-emissions-inventories/2020-national-emissions-inventory-nei-data.
[8] EPA, 2011. Additional Clarifications Regarding Application of Appendix W Modeling Guidance for the 1-hour NO_2 NAAQS. https://www.epa.gov/sites/default/files/2015-07/documents/appwno2_2.pdf.
[9] EPA, 2020. Technical Support Document: Chapter 6. Final Round 4 Area Designations for the 2010 1-Hour SO_2 Primary National Ambient Air Quality Standard for New York. Available at 06-ny-rd4_final_so2_designations_tsd.pdf (epa.gov). Pages 19-20.

Figure 2-1: Ambient Air Boundary for Massena Operations Modeling

Figure 2-2: Background Concentrations by Season/Hour-of-Day



3. Dispersion Modeling Approach

Modeling aluminum smelters is challenging due to issues such as the presence of a localized heat signature that can be generated from the facility itself as shown in **Figure 3-1**. As such, aluminum smelters cannot be accurately characterized using the guideline model, AERMOD, without consideration of site-specific features associated with this type of industrial area. This has been established in other recent modeling analyses for SO_2 NAAQS at aluminum smelters in Indiana and Washington (Alcoa Warrick Operations and former Intalco Works, respectively) where accepted modeling approaches used site-specific source characterizations. These cases, approved by the respective state agencies and EPA, are examples of the use of site-specific source characterization for modeling aluminum smelters, especially with the availability of nearby monitors to test the model performance.

Two enhancements and one source characterization change are proposed in this updated protocol for a site-specific modeling approach that more accurately characterize SO_2 concentrations around the Alcoa Massena Operations in the NAA. The two enhancements were already approved in October 2022 for use for Massena Operations by EPA[4] (100-m deep neutral temperature profile, and enhancement for treatment for building downwash). The new source characterization change proposes a method to represent the potline roof vent emission sources so that they are explicitly modeled at their source locations and release heights instead of adding their emissions to the dry scrubber stack source locations as was the method used in the 2022 modeling protocol. The next sections describe these proposed enhancements as well as the source characterization change in greater detail.

3.1 Consideration of Fugitive Heat Releases at the Smelter

The proposed model to be used in this application is the latest AERMOD modeling system (version 23132 at the time of this protocol). The choice of rural or urban for dispersion conditions generally depends upon the land use characteristics within 3 km of the facilities as described in Appendix W to 40 CFR Part 51.[10] Factors that affect the rural/urban choice, and thus the dispersion, include the extent of vegetated surface area, the water surface area, types of industry and commerce, and building types and heights within this area. According to EPA modeling guidelines, if more than 50% of an area within a 3-km radius of the facility is classified as rural, then rural dispersion coefficients are to be used in the dispersion modeling analysis. Conversely, if more than 50% of the area is urban, urban dispersion coefficients are used. As shown in **Figure 3-2**, this analysis would indicate that the land use around the Massena smelter is rural.

Emission sources such as aluminum smelters are unique in that they are associated with large fugitive heat releases that result in a local urban-like dispersion environment. Both the Warrick and Intalco facilities, at least twice the capacity of Massena Operations, were modeled using an urban source characterization. However, the Massena smelter has a smaller heat signature (temperature-wise and in aerial coverage) from the Warrick and Intalco smelters as shown in the satellite image in **Figure 3-3**. The urban-rural temperature difference between the "hot spots" and cooler air is about 10 Kelvin (K) at Massena. This would result in an effective urban population of about 250,000. Based on discussions with NYSDEC, this effective population and the smaller thermal footprint relative to the other Alcoa smelters are not considered large enough to be able to consider the Massena Operations smelter as an urbanized "highly industrialized area". Therefore, Massena Operations sources will be modeled using rural dispersion characteristics.

Highly industrialized areas, especially those associated with facilities with substantial fugitive heat losses that have not been factored into dispersion model development and evaluation by EPA, present unique challenges for routine model applications. Aluminum smelters, especially those put into place several decades ago, are among this type of facility. According to heat balance studies[11] of aluminum smelters such as Alcoa's Massena Operations, the waste heat from the electrical demand used in the Hall-Heroult process constitutes 45-50% of the total energy input. For

[10] EPA, 2017. Guideline on Air Quality Models. Available at https://www.epa.gov/scram/2017-appendix-w-final-rule.
[11] See discussions at http://peter-entner.com/E/Theory/EBal/EBal.aspx and http://www.tms.org/pubs/journals/JOM/9905/Welch-9905.html.

Massena Operations, the electrical input amounts to 240 megawatts (MW). If only 50% of the waste heat eventually escapes to the atmosphere (a conservatively low fraction), the resulting heat loss to the atmosphere from the potline area would amount to about 50 MW. Other areas of heat releases occur in connection with SO_2 emission sources associated with anode baking operations as well as holder/melting operations.

Figure 3-1 provides an infrared photo (Flir camera) as well as a normal (visible spectrum) image of the hot process area associated with the potline area of the smelter. **Figure 3-3** shows a thermal satellite image showing the temperature excesses in the smelter area as well as other areas in the facility. These figures document that there are significant fugitive heat losses from the potline area of the smelter. The thermal satellite image further shows that the bake furnace, package boilers, and melters/holders are also located in areas with significant fugitive heat losses due to their industrial processes.

In a recently released white paper,[12] Dr. Stephen Hanna commented in support of a neutral temperature lapse rate above the smelter stacks. Applicable text from his white paper is provided below in support of this concept.

- “I conclude that is important to account for the fact that the fugitive heat emissions are causing the local stability to remain neutral during the night. This correction will help reduce the significant overpredictions by the default AERMOD model.”
- “The modification to account for fugitive heat releases holds promise. The current AERMOD model and PRIME downwash model neglect the effects of fugitive heat releases around the building, which will lead to underpredictions of plume rise and hence overprediction biases in concentration. The fugitive heat releases are spread across the broad building roof and are about the same magnitude (a few hundred watts/m^2) as the natural daytime sensible heat flux due to solar warming of the ground surface. Therefore, stable boundary conditions do not occur at night over and around the building, and the effective stability is neutral or slightly unstable.”
- “In order to accommodate the effects of fugitive heat releases at night, a neutral temperature lapse rate could be assumed in the vicinity of the sources affected by the fugitive heat releases.”

Quasi-neutral conditions have been used in modeling of large industrial areas, such as for the urban option in AERMOD. Although the equivalent urban population for the Massena Operations application was determined by NYSDEC to be too small for using an urban dispersion option, the equivalent population of 250,000 supports a quasi-neutral layer above the source of at least 200 m, according to the AERMOD Model Formulation document[13] (Equation 110). Also, neutral conditions were observed in the lowest 100 m above the US Steel Clairton Works during a 1996 field study[14] due to fugitive heat releases from that facility, even though the land use was rural.

To accommodate this effect in the modeling, AECOM simulated a 100-m deep neutral layer by adding a 100-m temperature “observation” to the PROFILE input to the modeling. This temperature implemented the neutral lapse rate for all hours by applying a dry adiabatic lapse rate of 0.0098 degrees Celsius (C) per m (or 0.88 degree C drop in ambient temperature between 10 and 100 m). During daytime conditions, the temperature lapse rate is not used in AERMOD; this change only affected nighttime conditions.

This change was approved by EPA to be used as part of the site-specific modeling approach for Massena Operations in October 2022.[4] It demonstrated improved model performance by reducing overpredictions at the two nearby monitors, but an additional change was previously proposed to further improve the model performance, as noted in Section 3.2.

[12] Hanna, S.R., 2022. Review of LIFTOFF Model as AECOM has Implemented into AERMOD. Report P210. February 6, 2022.
[13] EPA, 2023. AERMOD Model Formulation. https://gaftp.epa.gov/Air/aqmg/SCRAM/models/preferred/aermod/aermod_mfd.pdf.
[14] David Sullivan, 1996. Review of Meteorology at the Clairton Area: Strengthening Dispersion Modeling for State Implementation Plans. Submitted to the Allegheny County Health Department.

3.2 Proposed Enhancement for Treatment of Building Downwash

The residual overpredictions after application of the neutral lapse rate as discussed above are likely related to building downwash effects. Therefore, we applied an available building pre-processor program that addresses known limitations for long and narrow buildings.

BPIPPRM_19191_DRFT

For several years prior to 2019, EPA modeling workshops[15] had noted a problem with how the Building Profile Input Program (BPIP) depicts building dimensions on angular approaches to long and narrow buildings. For angular approaches, the BPIP program could produce overly large building dimensions corresponding to the product of the diagonal dimensions for some wind direction approaches. To respond to this shortcoming, EPA's Office of Research and Development (ORD) provided an updated BPIP program in 2019 ("BPIPPRM_19191_DRFT")[16] that restricts the building footprint for any approach angle to the actual building footprint area. Due to the presence of very long and narrow buildings for this modeling application as shown in **Figure 3-4**, this improvement to the BPIP processing is a promising option for use with emissions sources with building downwash influenced by long and narrow buildings, which results in improved model performance, as documented in Section 4.

This additional change improved model performance at the two nearby monitors and was also approved by EPA for use with the 100-m deep neutral layer refinement for the site-specific modeling approach for Massena Operations in October 2022.[4]

3.3 Characterization of Potline Roof Vents

The most appropriate approach for modeling roof vents is with the BUOYLINE source type, since the Buoyant Line and Point Source (BLP) model for this source type was developed for aluminum reduction facilities. However, for this application at Massena Operations, the characterization of the lowest 100 m of the atmosphere with a neutral temperature lapse rate is an important issue for a successful modeling approach. Unfortunately, the BUOYLINE source type, which is basically the BLP model inside AERMOD, has a key limitation in its current implementation in AERMOD in that it does not use the vertical profile of temperature data provided to the model through the meteorological profile file. Instead, with a rural source type, BUOYLINE sources would be modeled with a stable Pasquill-Gifford stability class at night, which uses a stable temperature lapse rate that is inconsistent with actual conditions and with the intent of the site-specific modeling approach that was approved by EPA.[2] An alternative modeling approach for roof vent emissions at Massena Operations is to characterize the vents as a row of several point sources (stacks), using the roof vent dimensions and measured flow rate or velocity, equivalent stack parameters were calculated where the flow rate/velocity and roof vent area are conserved. The exit velocity is set to the actual roof vent exit velocity, which is based on the area of the roof vent opening and its flow rate. To accomplish this, an equivalent stack diameter is calculated based on the number of representative stacks based on total area of the roof vent opening divided by the number of stacks. In other words, the combined cross-sectional area of the stacks modeled equal the roof vent opening area, since the exit velocity matches the actual roof vent velocity. The next sections describe the development of this approach.

Methodology to Determine a Recommended Potline Roof Vent Modeling Approach

In order to properly implement a neutral temperature lapse rate for the BUOYLINE sources, a simple approach is to model the BUOYLINE sources only as urban. AECOM conducted modeling with this approach and our findings and areas for further review are noted below:

- BUOYLINE sources modeled as urban essentially ignore the population setting, and simply sets the nighttime stability class to 4 (neutral); results are nearly identical no matter what population is entered. Generally, smaller populations result in somewhat fewer hours being set to stability class 4 than larger populations.

[15] See, for example, the EPA 2018 Regional/State/Local Modeler' Workshop presentation, Downwash Summit Report Out: EPA (slide 16): https://gaftp.epa.gov/Air/aqmg/SCRAM/workshops/2018_RSL_Modelers_Workshop/2018_RSL_Modelers_Workshop-Final_Agenda.pdf.

[16] EPA, 2019. Release of BPIPPRM, Version 19191_DRFT, for Public Review and Comment. Available at BPIPPRM_19191_DRFT-Trans_Memo.pdf (epa.gov).

- EPA registered its concern on a January 24, 2023 call about the BUOYLINE urban approach – not because of the neutral temperature lapse rate, but because there is no vertical upper limit assumed for this lapse rate with the BLP model within AERMOD. Through use of the BLP debug output, AECOM determined the BUOYLINE plume height as a function of distance to understand whether the plume height exceeded 100 m well before the distance to peak smelter ground-level impacts.
- With the possibility that the BUOYLINE plume rise was above 100 m well before the distance to the peak smelter impact, AECOM worked on the alternative approach with a row of point sources. The issue to be resolved was the number of point sources along the two roof vents to use in the modeling instead of a BUOYLINE approach.

Potline Roof Vent Exhaust Parameters for Line and Stack Source Model Testing

Alcoa has conducted periodic testing on the potline roof vent exhaust. The results of the testing were used to determine model inputs for the BUOYLINE input data. For the alternative approach using a row of identical stacks, the flow rate and area for the entire roof vent was conserved while the total roof vent area was evenly divided among number of the stacks modeled to calculate an equivalent stack diameter per stack. **Table 3-1** details the roof vent exhaust parameters when modeled as two BUOYLINE sources. **Table 3-2** details the roof vent exhaust parameters when modeled as two rows of 4 stacks, for a total of 8 stacks. **Figure 3-5** shows how the roof vents are represented in the modeling analysis as both the BUOYLINE source and stacks. The emissions for all scenarios for this specific test were based on actual emissions by using the 5-year average emissions for the modeling evaluation period, as requested by the reviewing agencies. However, the emissions could be any arbitrary value for this test as long as they are consistent between the line and point source depictions. These parameters were used for model input, paired with a receptor grid described in Section 3.6.

Table 3-1: Source Parameters and Emissions for the Roof Vents Modeled as BUOYLINE Sources

Source Parameter	Value
Modeled Emission Rate Per Roof Vent (g/s)(1)	1.675
Base Elevation (m)	65.5
Release Height (m)(2)	20.31
Source Length (m)	565.7
Building Length (m)	630.0
Building Height (m)	17.39
Building Width (m)	25.6
Line Width (m)(3)	3.9
Building Separation (m)	25.6
Average Ambient Temperature (K)(4)	280.8
Median Potroom Temperature (K) (5)	305.8
Average Potroom Flow Rate (m^3/s) (6)	2066.04
Buoyancy Parameter, F' (m^4/s^3)	1656.27

(1) 5-year average emissions based on the actual monthly emissions from 2017-2021.
(2) Roof vent height is used for the release height.
(3) Roof vent opening width is used for the line width.
(4) Ambient temperature is based on the average temperature from the meteorological data from 2017-2021.
(5) Temperature data are based on the median potroom test data from 2017-2020.
(6) Flow rate is an average flow rate based on potroom test data from 2017-2019.

Table 3-2: Source Parameters and Emissions for the Roof Vents Modeled as 4 Stacks per Roof Vent

Source Parameter	Value for Each of the 4 Stacks per Roof Vent
Modeled Emission Rate Per Stack (g/s)(1)	0.419
Base Elevation (m)	65.5
Stack Height (m)(2)	20.31
Stack Temperature (K)(3)	305.8
Stack Exit Velocity (m/s)(4)	0.94
Equivalent Stack Diameter (m)(5)	26.5

(1) Total emissions per roof vent are divided by the number of stacks modeled.
(2) Roof vent release height is used for the stack height.
(3) Temperature data are based on the median potroom test data from 2017-2020.
(4) Stack exit velocity is based on the total vent area and vent flow rate in potroom test data from 2017-2019.
(5) Equivalent stack diameter is based on the resulting individual stack area when the roof vent opening area is divided by the 4 modeled stacks.

Findings with the BUOYLINE Urban Approach

AECOM has reviewed the BLP debug results after running BUOYLINE sources for characterizing the roof vent emissions with urban dispersion enabled. The debug output lists a series of seven plume rise amounts at seven downwind distances and assumes that the seventh distance and plume height establish the final plume height. For more than 90% of the nighttime hours (affected by the urban dispersion option), the resulting final plume height was above 100 m.

The next step was to determine the distance for BUOYLINE sources to attain a median height of 100 m. Then, modeling at that distance for the roof vents depicted as a buoyant line source versus a series of identical stacks would determine the stack arrangement that best matches the BUOYLINE approach. The distance to a BUOYLINE height of 100 m for nighttime hours was found to vary as a function of meteorological conditions (e.g., wind speed). The distance to a 100-m BUOYLINE plume height for half of the nighttime hours was found to be about 800 m, or about the distance to the closest plant's ambient air boundary for a direction toward the north-northeast, aligned with the most frequent wind direction sector. **Figure 3-6** shows a cumulative frequency distribution of the nighttime final plume height distances for a 100-m BUOYLINE plume height.

Determining the Most Representative Number of Stacks for Modeling the Potline Roof Vent Emissions

The objective of this part of the analysis was to find, for the BUOYLINE source distance for which the plume reached ambient air with a median of the cases at or below the 100-m plume height, a stack arrangement that provided a similar result. All ambient air boundary receptors were used for this analysis. Receptors beyond the ambient air boundary were not used because those locations would involve greater frequency of BUOYLINE plume rises beyond 100 m.

Several stack cases were run with a constant emission rate for the roof vent emissions, each with a different number of stacks representing each roof vent. **Table 3-3** describes the cases closest to BUOYLINE result for Massena_MOD and shows the maximum design concentration at the plant's ambient air boundary. For Massena_MOD, the case with 4 stacks representing each roof vent has a maximum design concentration (9.5 µg/m^3) that is slightly higher than the result associated with the roof vents modeled as buoyant line sources (9.4 µg/m^3). The use of fewer stacks than 4 results in an impact less than 8.4 µg/m^3, so the case that was shown to be slightly higher for conservatism was selected to serve as a representative alternate modeling approach in lieu of the buoyant line source approach.

This proposed potline roof vent characterization will be evaluated in a comparison with SO_2 ambient air monitoring data through the updated model evaluation in Section 4.

Table 3-3: Stack Sensitivity Analysis Results

Roof Vents Modeled As...	Modeled Maximum 99th Percentile Design Concentration (μg/m^3)
2 buoyant line sources run with urban dispersion and 250 K population	9.4
3 stacks per roof vent	8.4
4 stacks per roof vent	9.5
5 stacks per roof vent	10.5

3.4 Emissions Processing for Model Evaluation

Alcoa has documented monthly emission rates and stack parameters for the SO_2 sources at the smelter. There is a complication in that the western stacks along the potlines were capped from September 18, 2018 to November 19, 2019. During this time, all western stacks at all three courtyards (A, B, and C) were capped, for a total of 18 stacks capped. To deal with this issue, an hourly emissions file will be used to turn the appropriate stacks on and off for the times when the western stacks were capped vs. uncapped. When the western stacks were not capped (from January 1, 2017 to September 17, 2018 and November 20, 2019 on), the potline stacks will be modeled with all stacks active, example as shown for Courtyard A in **Figure 3-7**. When the western stacks were capped, the subsequent emissions from the western stacks will be modeled to be emitted through the eastern stacks, which handled the merged exhaust. Therefore, the "merged" stacks were represented by modeling all 6 of the eastern stacks - example as shown in for Courtyard A in **Figure 3-8**.

3.5 Meteorological Data Processing

Five full years (2017-2021) of surface observations from nearby Massena International Airport in Massena, NY were used in conjunction with the twice-daily soundings upper air data from Albany, NY in AERMET (version 23132), the meteorological preprocessor for AERMOD, which is consistent with guidance stated in 9.3.1.2 of 40 CFR Part 51, Appendix W (EPA modeling guidelines). Massena International Airport is the closest Automated Surface Observation Station (ASOS) station (within 2 km of the smelter) with shorter average data (e.g., 1-minute) to the Alcoa Massena facility. Likewise, Albany, NY is the closest upper air station in the U.S. The Integrated Surface Hourly (ISH) data was obtained from the National Climatic Data Center's (NCDC) ftp site and the Forecast Systems Laboratory (FSL) formatted upper air data was obtained from NCDC's FSL website. **Figure 3-9** shows the meteorological stations with respect to Massena Operations. The meteorological data was processed for this modeling protocol due to the model evaluation presented herein. NYSDEC provided the meteorological data for 2017-2021 using AERMET version 22112 and AECOM used NYSDEC processing files to update the meteorological data with AERMET version 23132.

The surface meteorological data at the Massena International Airport is recorded by an ASOS that records 1-minute measurements of wind direction and wind speed (anemometer height of 10-m), along with hourly surface observation data including temperature, cloud cover, ceiling height, surface pressure, etc. The 1-minute data is especially desirable because it provides a more accurate depiction of the average winds during the hour as opposed to a snap-shot in time. In addition, it ensures greater temporal resolution wind measurements that results in fewer calm winds (which are excluded from modeling) compared to stations that do not record minute data. Also, EPA specifically prefers that surface National Weather Service stations with 1-minute data should be used for AERMOD modeling. The current EPA-approved version of AERMINUTE (version 15272) was run with the Massena 1-minute data. Five-minute data were also be used as a substitute for any missing 1-minute data.

AERMET creates two output files for input to AERMOD:

- SURFACE: a file with boundary layer parameters such as sensible heat flux, surface friction velocity, convective velocity scale, vertical potential temperature gradient in the 500-m layer above the planetary boundary layer, and convective and mechanical mixing heights. Also provided are values of Monin-Obukhov length, surface roughness, albedo, Bowen ratio, wind speed, wind direction, temperature, and heights at which measurements were taken.

- PROFILE: a file containing multi-level meteorological data with wind speed, wind direction, temperature, sigma-theta (σ_θ) and sigma-w (σ_w) when such data are available. As stated above, the simulation of a 100-m deep neutral layer was done by inserting an assumed temperature "observation" at the 100-m level that was 0.88 deg C lower than the hourly temperature observed at the 10-m level.

Although the Massena airport temperature has been input to AERMET at the 2-m level, AERMET outputs the data in the SURFACE and PROFILE files for temperature at the 10-m level to match the levels of the wind data.

AERMET requires specification of the meteorological station site characteristics including surface roughness (z_o), albedo (r), and Bowen ratio (B_o). These parameters have been developed according to the guidance provided by EPA in the AERMOD Implementation Guide (AIG).[17]

The AIG provides the following recommendations for determining the site characteristics:

1. The determination of the surface roughness length should be based on an inverse distance weighted geometric mean for a default upwind distance of 1 kilometer relative to the measurement site. Surface roughness length may be varied by sector to account for variations in land cover near the measurement site; however, the sector widths should be no smaller than 30 degrees.

2. The determination of the Bowen ratio should be based on a simple unweighted geometric mean (i.e., no direction or distance dependency) for a representative domain, with a default domain defined by a 10-km by 10-km region centered on the measurement site.

3. The determination of the albedo should be based on a simple unweighted arithmetic mean (i.e., no direction or distance dependency) for the same representative domain as defined for Bowen ratio, with a default domain defined by a 10-km by 10-km region centered on the measurement site.

The AIG recommends that the surface characteristics be determined based on digitized land cover data. EPA has developed a tool called AERSURFACE that determines the site characteristics in accordance with the recommendations from the AIG. AERSURFACE incorporates look-up tables of representative surface characteristic values by land cover category (and tree canopy and impervious surface categories where available) and seasonal category. AERSURFACE was applied with the instructions provided in the AERSURFACE User's Guide.

The current version of AERSURFACE (version 20060) was developed to support the use of land cover data from the USGS National Land Cover Data (NLCD) developed after 1992. The post-1992 land cover data, which has been developed for years 2001, 2006, 2011, and 2016, has a different classification scheme than the 1992 land cover data. The USGS provides all of the data at a spatial resolution of 30 meters x 30 meters. While the NLCD 1992 archive uses a 21-category classification scheme applied over the continental U.S., the NLCD data archives from the post-1992 data use a modified 16-category classification scheme applied over the continental U.S., Hawaii, and Puerto Rico, with four additional categories which are only applicable to Alaska. AERSURFACE version 20060 still reads the 1992 NLCD data. Some of the post-1992 NLCD data also include separate percent tree canopy data and percent impervious data files which are used to supplement the land cover data. EPA recommends using all three files if they are all available for a particular year. Land cover, percent tree canopy data, and percent impervious data from 2016 were used in AERSURFACE for this modeling application. To date, there have been no notable changes to the airport landscape after 2016.

The AIG recommends that the surface characteristics be determined based on the land use surrounding the site where the surface meteorological data were collected. As recommended in the AIG for surface roughness, the 1-km radius circular area centered at the meteorological station site was created. For this analysis, the area was divided into 12 sectors as shown in **Figure 3-10** and **Figure 3-11**.

Beginning with AERSURFACE version 20060, each sector can be specified individually as "airport" or "non-airport". In previous versions the "airport" and "non-airport" classifications were applied to all sectors. A sector should be

[17] EPA 2023. AERMOD Implementation Guide (Revised). US Environmental Protection Agency, Research Triangle Park, NC. https://gaftp.epa.gov/Air/aqmg/SCRAM/models/preferred/aermod/aermod_implementation_guide.pdf.

considered "airport" if the sector is dominated by large impervious surfaces represented by runways, roads, parking lots, and other paved surfaces. A sector should be considered "non-airport" if the sector is dominated by buildings or other structures or by vegetation.

For this application, all sectors were designated as "airport" with the exception of the 240-270-degree sector which was designated as "non-airport". In AERSURFACE, the various land cover categories are linked to a set of seasonal surface characteristics. As such, AERSURFACE requires specification of the seasonal category for each month of the year. The following five seasonal categories are supported by AERSURFACE. The applicable seasons associated with the modeling period for this site are listed below.

1. Midsummer with lush vegetation (June, July, August)
2. Autumn with un-harvested cropland (September, October, November)
3. Late autumn after frost and harvest, or winter with no snow (category not used)
4. Winter with continuous snow on ground (January, February, December)
5. Transitional spring with partial green coverage or short annuals (March, April, May)

For Bowen ratio, the land use values are linked to three categories of surface moisture corresponding to average, wet, and dry conditions. For this application all three years were processed using average conditions.

Another modification to AERSURFACE version 20060 is the ability to specify the calculation method for the surface roughness length. This version of AERSURFACE has two different methods referred to as the "ZORAD" method and the "ZOEFF" method. The "ZORAD" method is the method used in previous versions of AERSURFACE where surface roughness length is calculated as an inverse distance weighted geometric mean based on the land cover within the area around the meteorological tower out to 1 km from the tower. The "ZOEFF" method is a research-grade method which does not limit the upwind fetch to 1 km from the tower. The "ZOEFF" method uses the estimated growth of the internal boundary layer (IBL) to determine how far from the tower to go in each sector. These methods are described in more detail in the AERSURFACE User's Guide for version 20060.[18] The default method is the "ZORAD" method, and this is the method that was used for this modeling application.

3.6 Receptor Processing

Model receptors are the locations for which the model will calculate modeled concentrations. For this application, the requirements of the receptor grid are defined by EPA guidance as well as NYSDEC guidance.[19,20] Receptor input to AERMOD will be generated for areas of ambient air for the compliance modeling, extending out to about 20 km from the facility (excluding Canada) as required by NYSDEC. For the model evaluation, the two receptors corresponding to the two monitors were processed using AERMAP. For general modeling involving all ambient air receptors within the nonattainment area, the discussion below applies.

According to a 1986 EPA memo, ambient air is defined as "…that portion of the atmosphere, external to buildings, to which the general public has access."[21] **Figure 2-1** (see previous section) illustrates the areas near the Massena Operations where general public access is controlled by Alcoa, and illustrates the ambient air boundary (red outline) for this analysis. A nested Cartesian (rectangular) receptor grid will be used with the receptor spacing as described below. This receptor grid includes a minor revision where public roadways will be modeled.

- Every 25 m along the ambient air boundary
- Every 70 m out to a distance of 2.5 km
- Every 100 m between 2.5 and 5 km

[18] EPA, 2020. User's Guide for AERSURFACE Tool. US Environmental Protection Agency, Research Triangle Park, NC. February 2020. https://gaftp.epa.gov/Air/aqmg/SCRAM/models/related/aersurface/aersurface_ug_v20060.pdf
[19] EPA, 2014. Guidance for 1-Hour SO_2 NAA State Implementation Plans Submissions. https://www.epa.gov/so2-pollution/guidance-1-hour-sulfur-dioxide-so2-nonattainment-area-state-implementation-plans-sip
[20] NYSDEC, 2020. DAR-10 – NYSDEC Guidelines on Dispersion Modeling Procedures for Air Quality Impact Analysis. https://www.dec.ny.gov/docs/air_pdf/dar10.pdf
[21] EPA, 2019. Revised Policy on Exclusions from "Ambient Air". https://www.epa.gov/nsr/ambient-air-guidance

- Every 500 m between 5 and 10 km
- Every 1000 m between 10 and 20 km (excluding Canada)

Figures 3-12 and 3-13 provide the receptor grid as viewed in the near-field and far-field. If necessary, additional 100-m spaced receptors will be placed at the location of the maximum impact should it not already be in an area covered by 100-m (or denser) grid spacing. Receptors will not be modeled over Canada.

Receptor height scales at each receptor location will be developed by AERMAP (version 18081), the terrain preprocessor for AERMOD, which requires processing of terrain data files. Terrain elevations from USGS National Elevation Dataset (NED) will be used to develop the receptor terrain elevations required by AERMOD.

Figure 3-1: One of the Three Potline Dry Scrubber Stack Clusters - Heat Generated by the Potlines

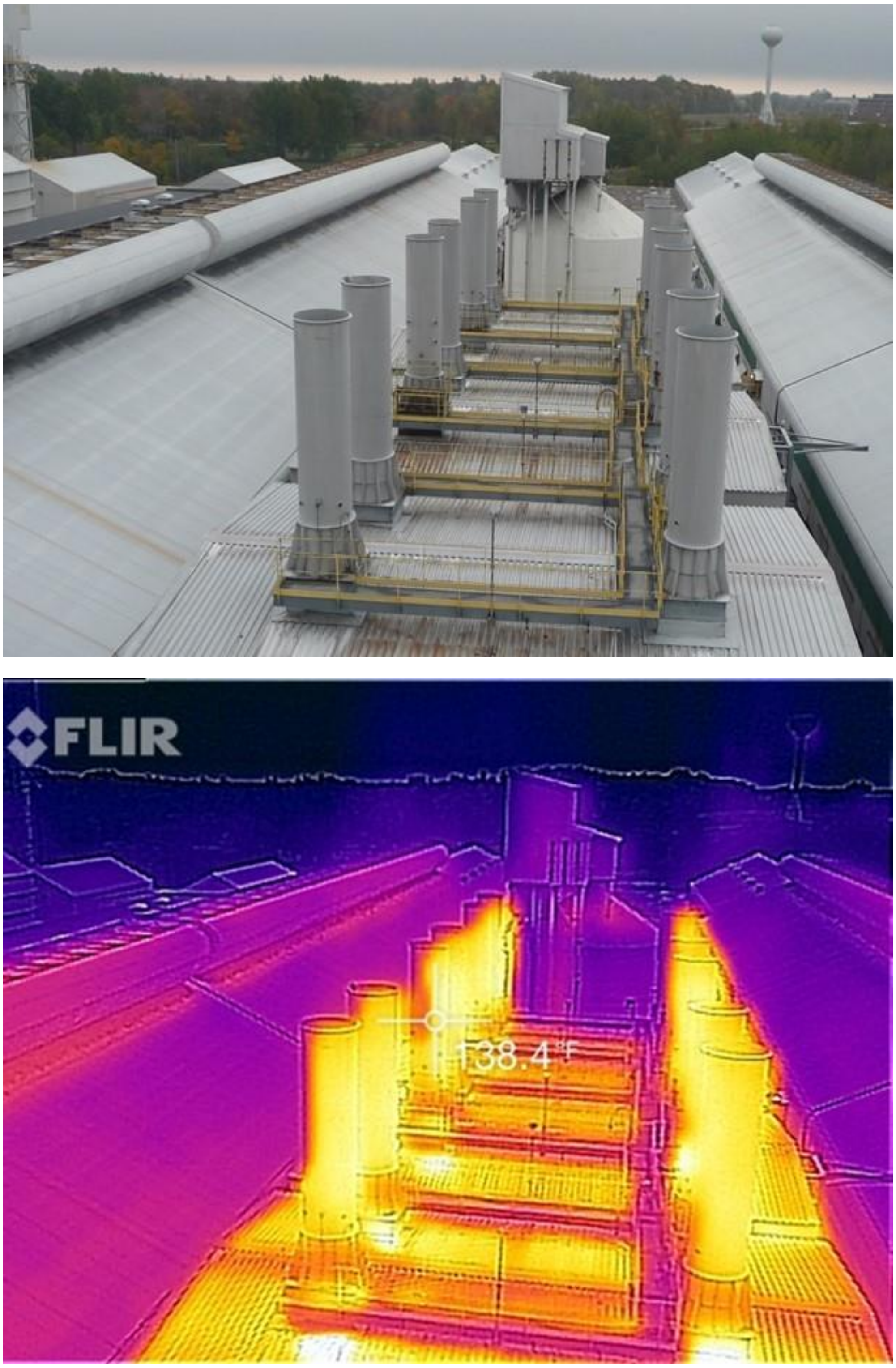


Figure 3-2: Land Use Around Massena Facility

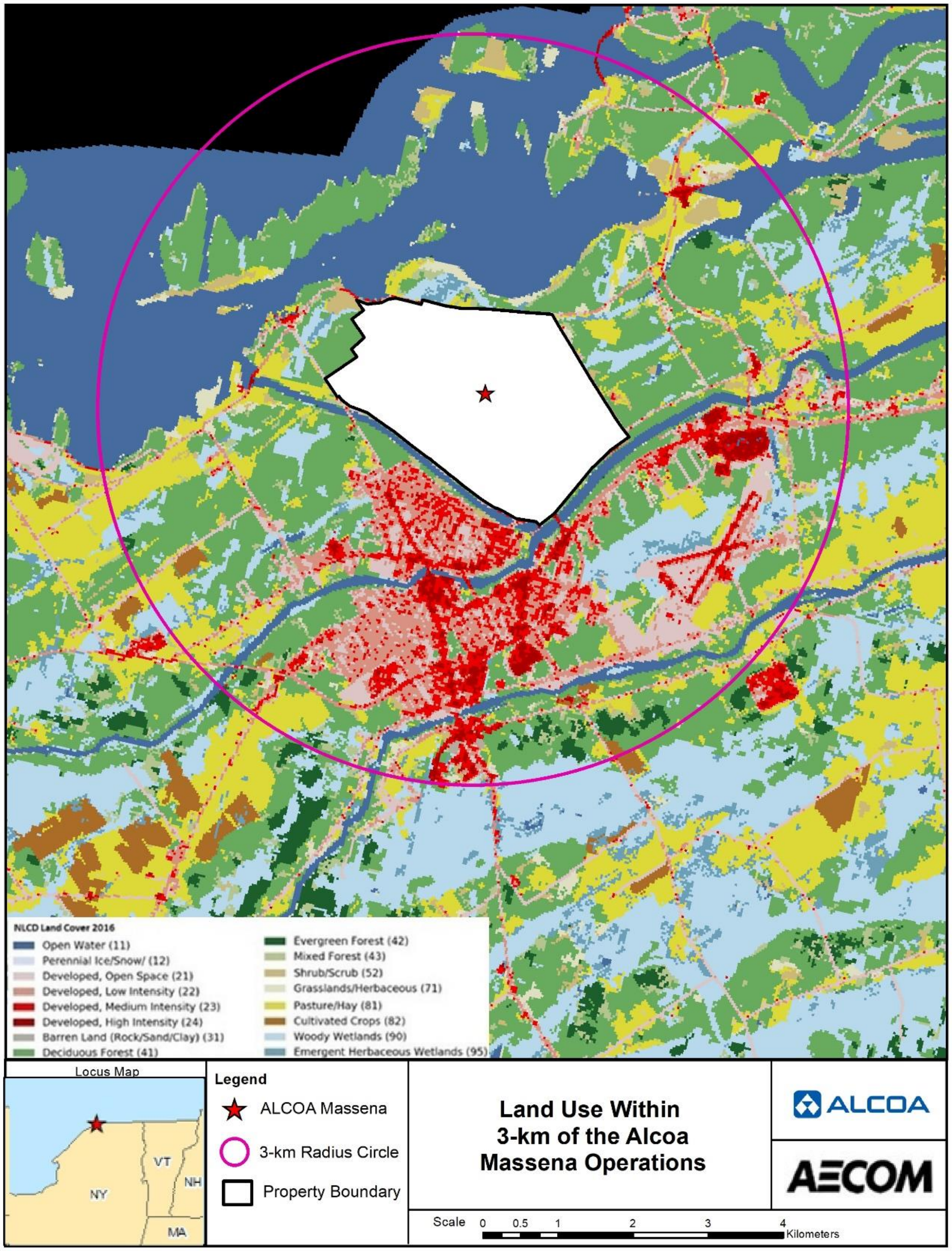


Figure 3-3: Surface Temperature Data at Massena Operations

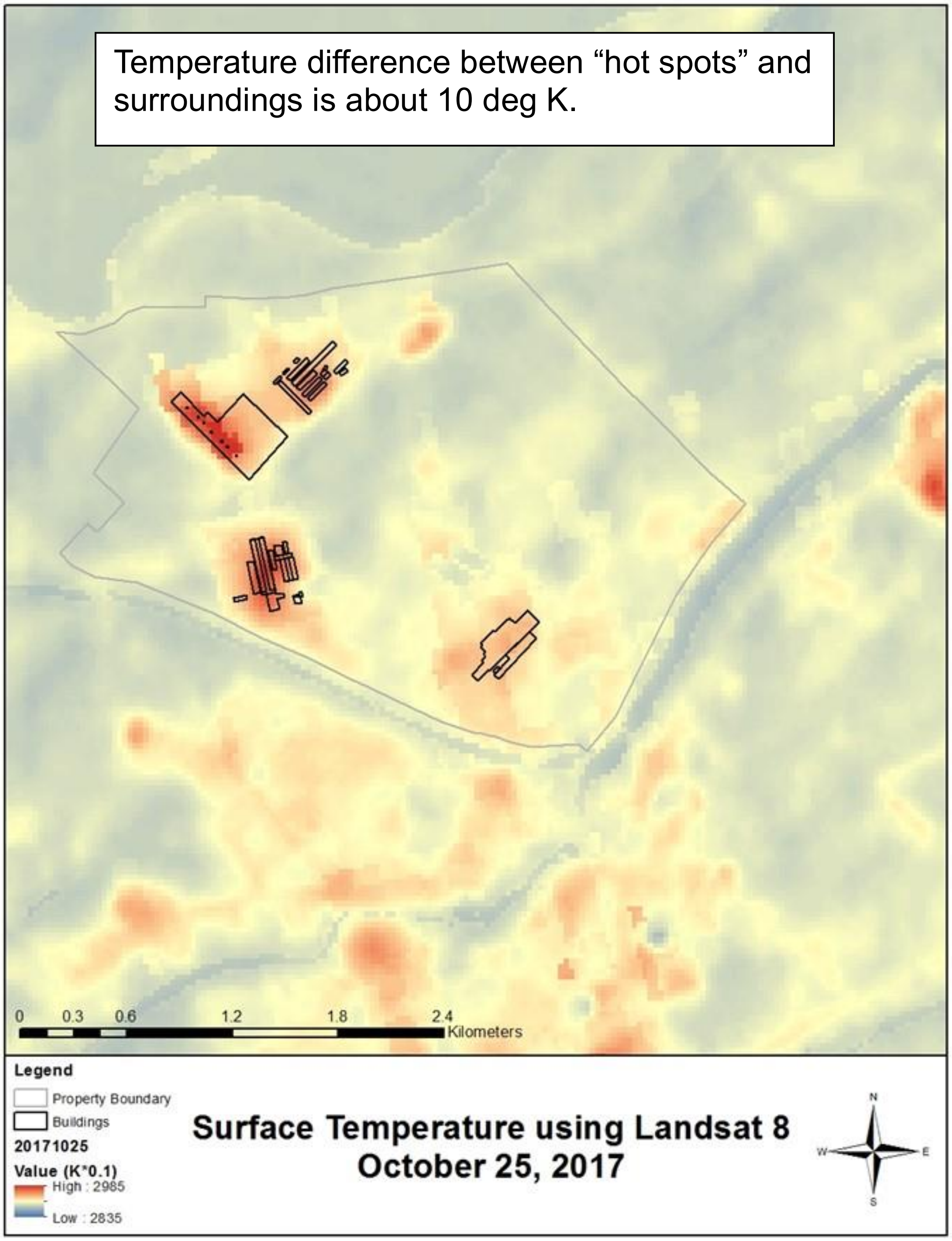


Figure 3-4: Modeled Building Layout

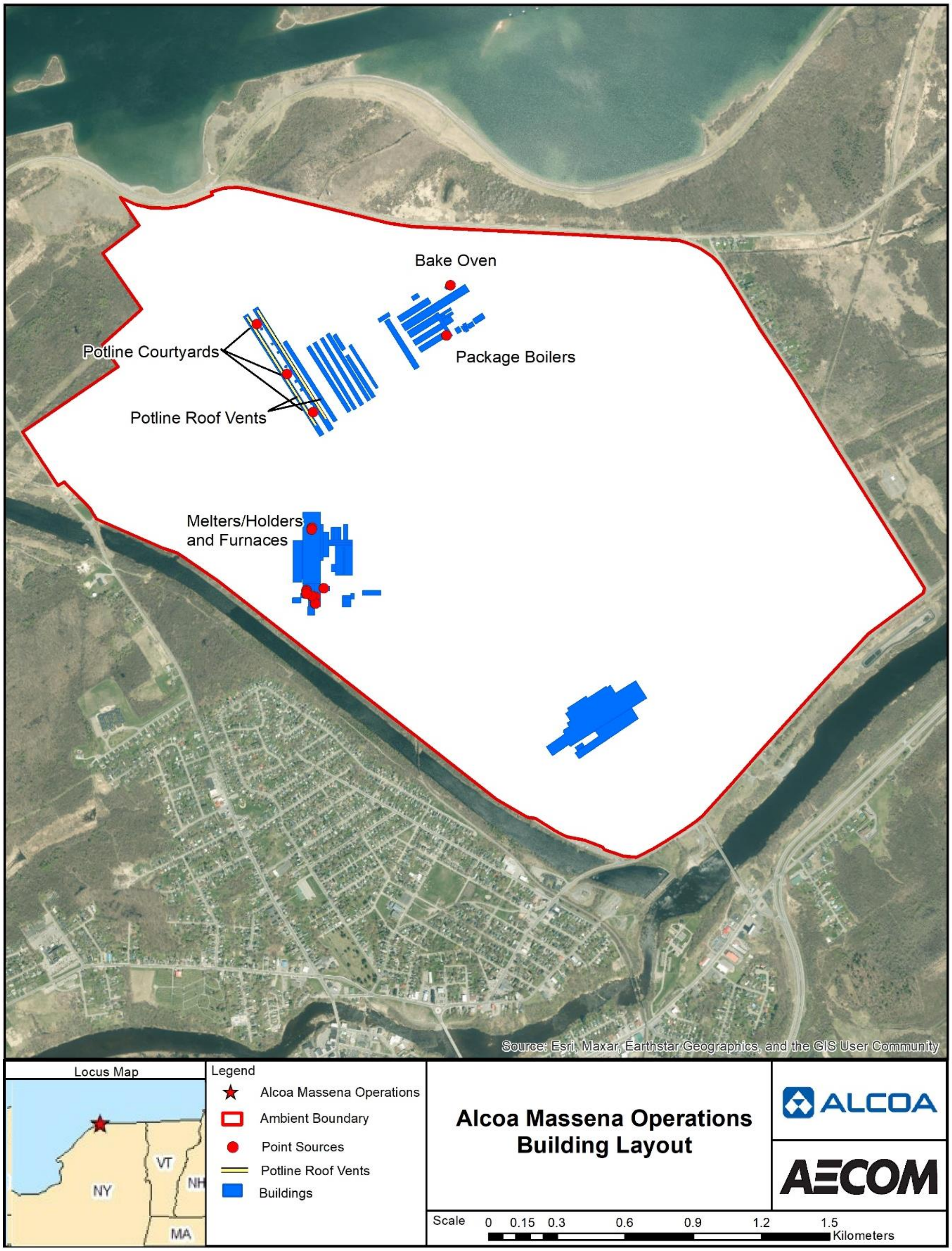


Figure 3-5: Potline Roof Vents Modeled as 4 Stacks Along Each Roof Vent

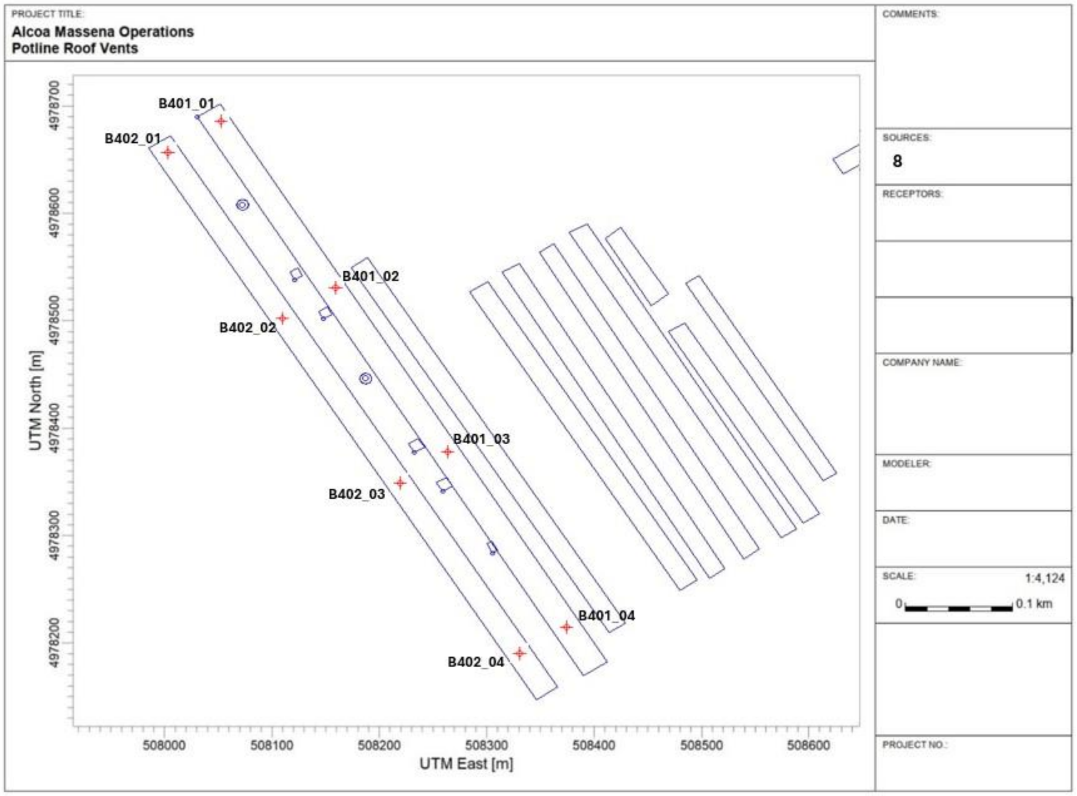


Figure 3-6: Cumulative Frequency Distribution of Final Plume Height Distances

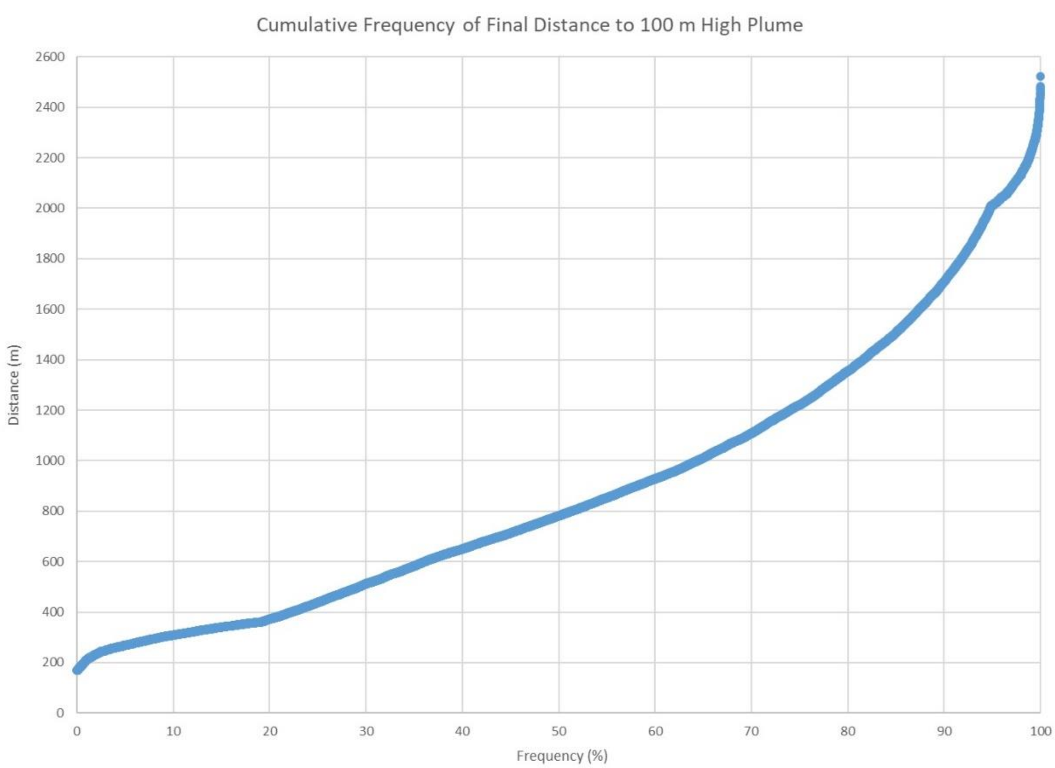


Figure 3-7: Modeled Location of Merged Stacks When No Stacks Are Capped for Courtyard A

Figure 3-8: Modeled Location of Merged Stacks When Western Stacks Are Capped for Courtyard A

Figure 3-9: Locations of Meteorological Stations

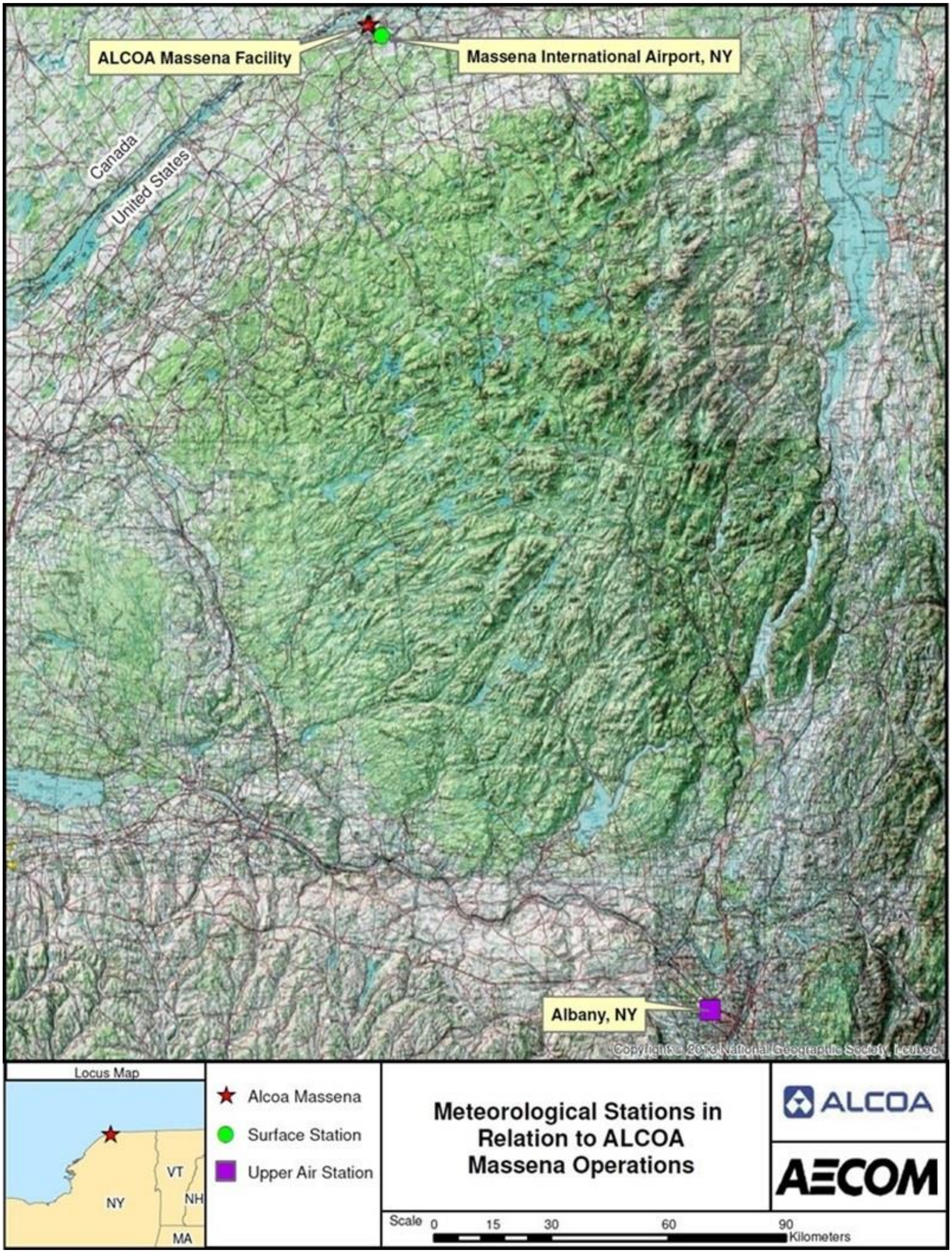


Figure 3-10: Land Use Sectors Overlaid Over NLCD Data

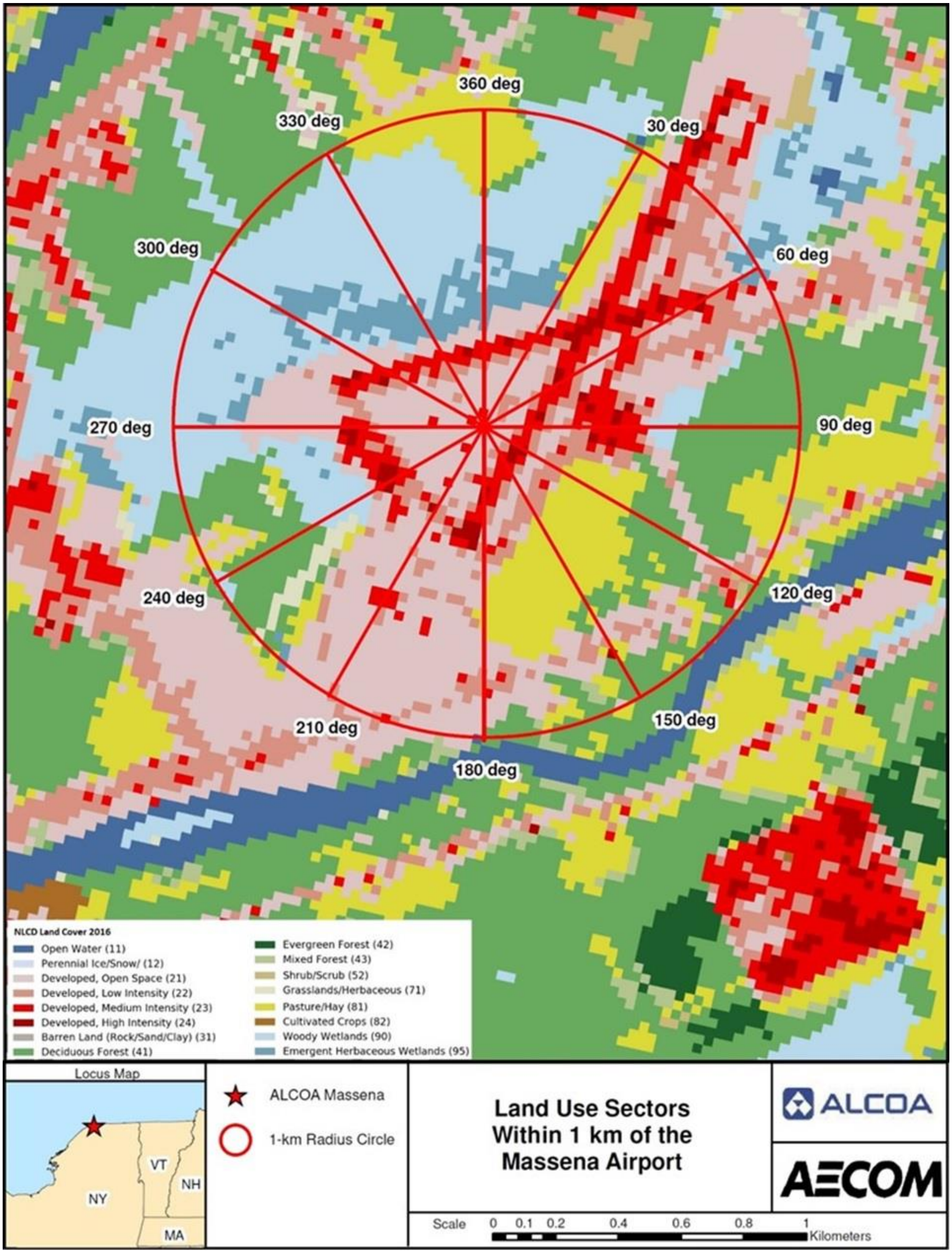


Figure 3-11: Land Use Sectors Overlaid Over Aerial Photo

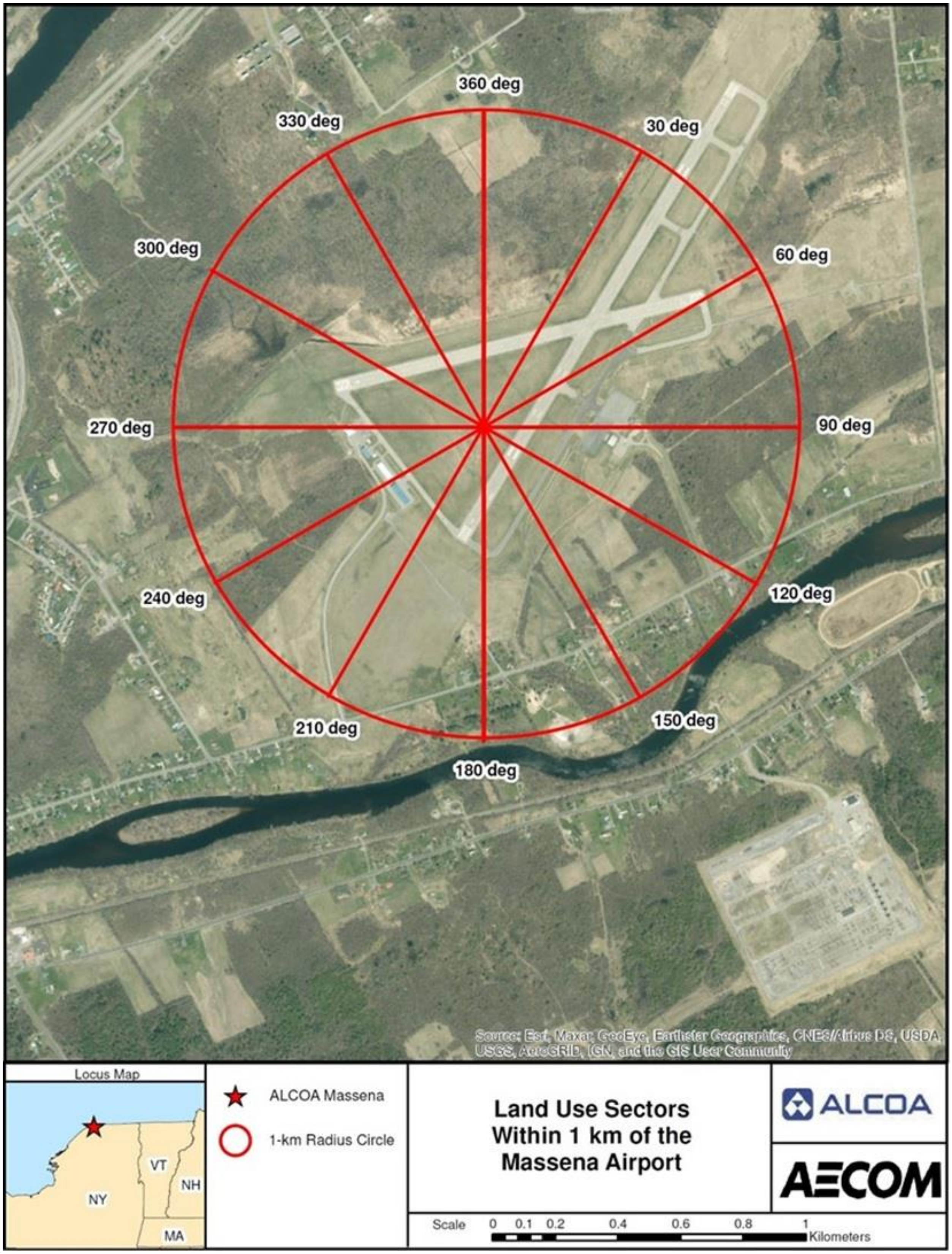


Figure 3-12: Near-Field Receptor Grid

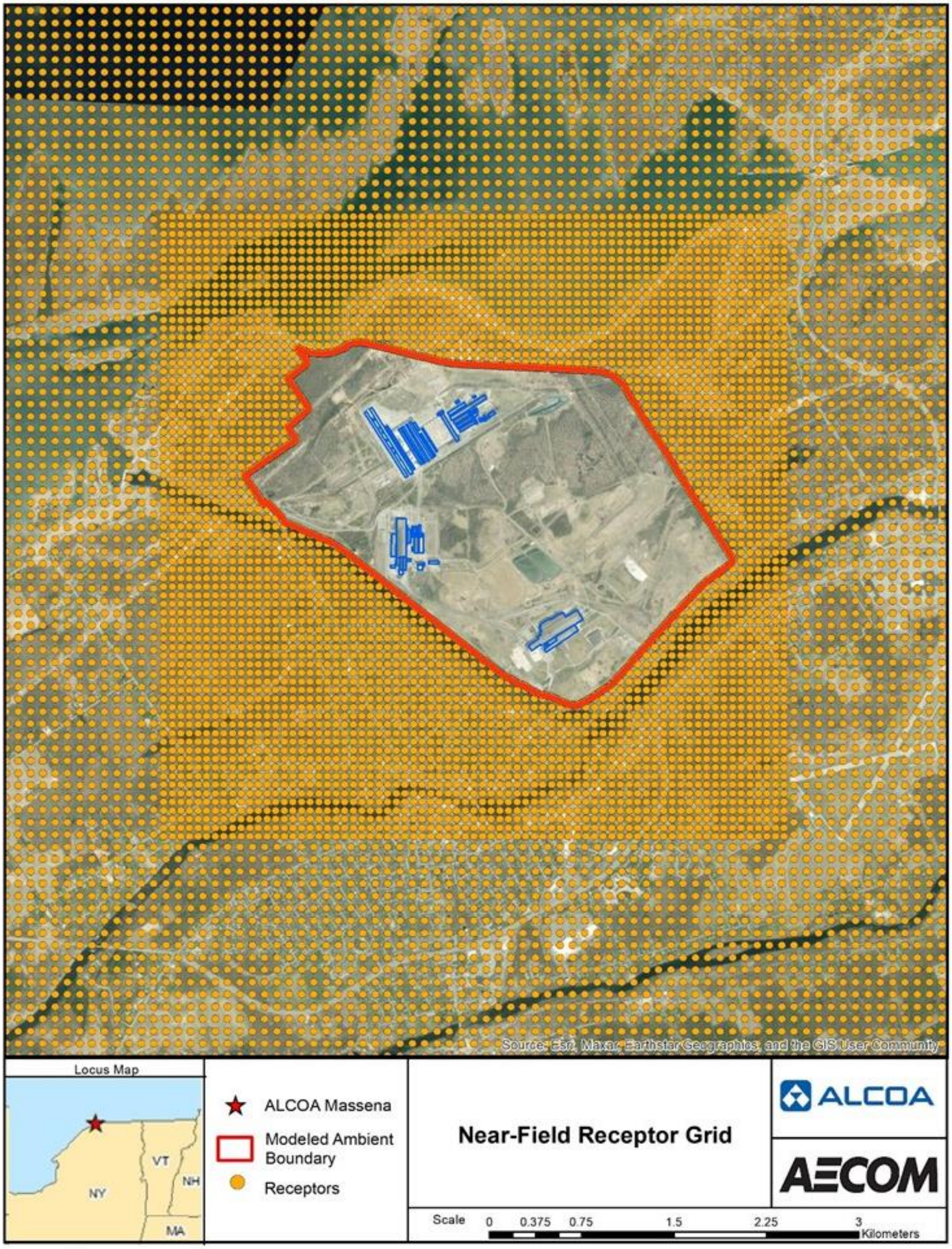
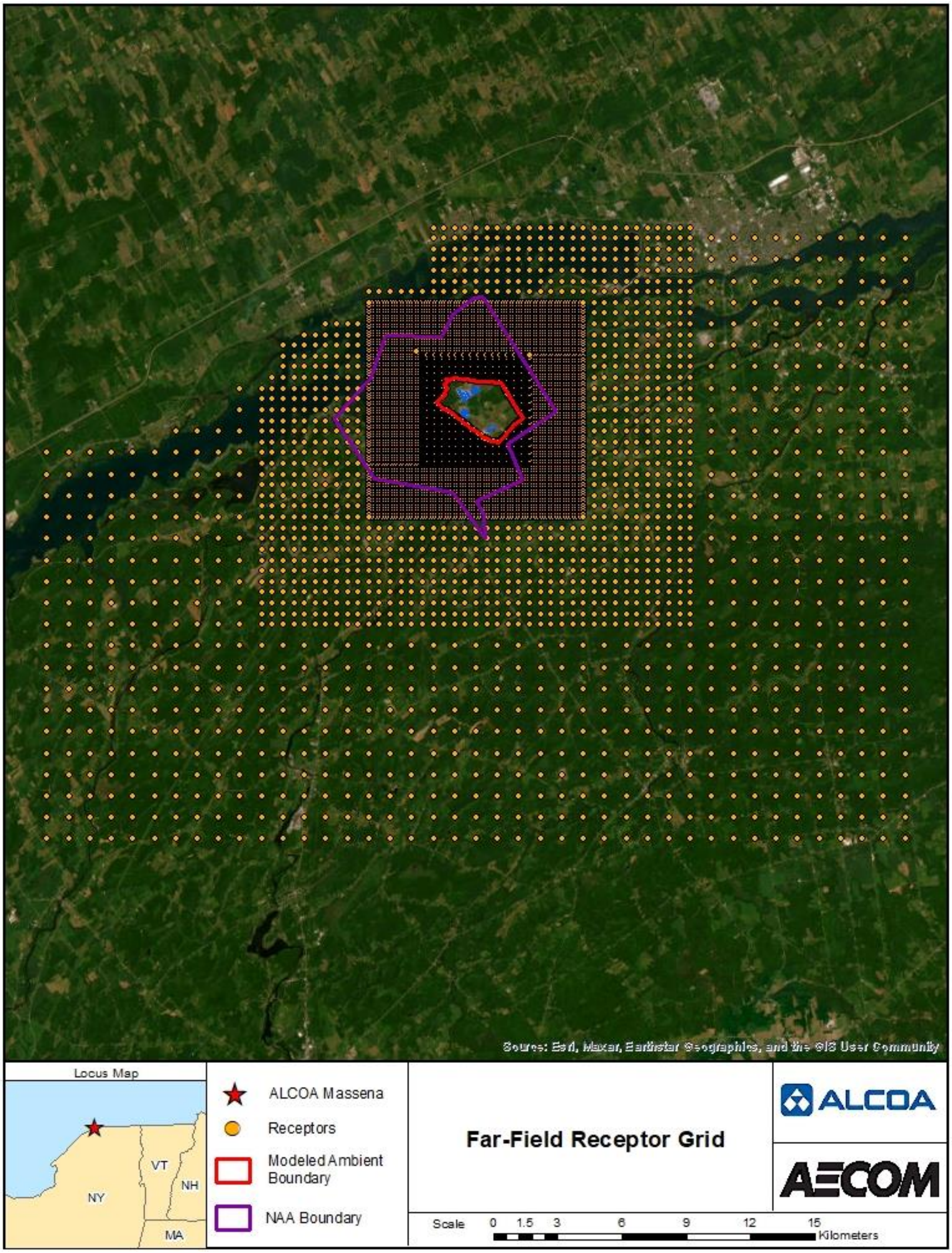


Figure 3-13: Far-Field Receptor Grid



4. Evaluation of Proposed Modeling Approach

The use of AERMOD for this modeling application with the neutral temperature lapse rate in the lowest 100-m and EPA's draft (ORD) BPIPPRM is proposed because the model performance is better matched to monitored conditions than the default model, as outlined below. This proposed model is considered to be a nonguideline technique; however, there is strong scientific justification for this approach. As stated in the 2017 Appendix W to 40 CFR Part 51 (EPA's modeling guidance; published in the Federal Register at 82 FR 5182 – January 17, 2017) in Section 3.2.2:

"Any … modification to a preferred model that would result in a change in the concentration estimates … alters its status so that it is no longer a preferred model. Use of the modified model must then be justified as an alternative model on a case-by-case basis to the appropriate reviewing authority and approved by the Regional Administrator."

"EPA has determined and applied a specific evaluation protocol that provides a statistical technique for evaluating model performance for predicting peak concentration values, as might be observed at individual monitoring locations. This protocol is available to assist in developing a consistent approach when justifying the use of other-than-preferred models recommended in the Guideline (i.e., alternative models). The procedures in this protocol provide a general framework for objective decision-making on the acceptability of an alternative model for a given regulatory application. These objective procedures may be used for conducting both the technical evaluation of the model and the field test or performance evaluation."

An alternative model is evaluated from both a theoretical and a performance perspective before it is selected for use. The scientific justification provided above addresses the theoretical perspective. For this specific application, Alcoa uses the model performance procedures for the second of three possible alternative model approaches (Appendix W section 3.2.2(b)(2)): "If a statistical performance evaluation has been conducted using measured air quality data and the results of that evaluation indicate the alternative model performs better for the given application than a comparable model in Appendix A", then the alternative model may be approved for use for the proposed application.

In this model application, the default (EPA's original BPIPPRM and the original AERMET-produced PROFILE file) and the site-specific modeling approach are evaluated. The site-specific approach will utilize the already approved enhancements included in "Massena_MOD" (ORD BPIPPRM and a neutral lapse rate). However, Massena_MOD will now also use the characterization of emissions from the two potline roof vents using a total of 4 identical stacks per roof vent, distributed evenly along each vent.

Section 4.1 describes an analysis of meteorological conditions associated with the top 25 observed and top 25 predicted concentrations for the default model as well as Massena_MOD. This analysis provides a sense as to whether the better performing model to be determined from the statistical tests described in Section 4.2 is performing better for the right conditions. Section 4.2 describes the statistical tests to be used for the performance evaluation. This protocol document is an important element in the process of the acceptance of an alternative modeling approach. An important consideration for the selection of the better modeling approach is the performance of the models being considered for the two local monitors (Site 1 and Site 2). For this modeling protocol, five full years of data (calendar years 2017 – 2021) are used for the model evaluation of the site-specific source characterization to be used in this application. This is an expanded dataset from the three years of data (2017 – 2019) used in the 2022 version of the modeling protocol. Meteorological data was processed using procedures described in Section 3. Modeling evaluation files will be provided for agency review in a modeling archive.

4.1 Analysis for Reviewing Model Performance

The reviewing agencies have asked for a review of the meteorological conditions associated with the peak observed and modeled concentrations to determine day vs. night and meteorological conditions (e.g., wind speed and mixing height) associated with these peak concentrations. This analysis looks at the top 25 1-hour observed and modeled concentrations.

Tables 4-1 and 4-2 provide a tabulation of the dates and hours of top 25 ranked observed concentrations at Sites 1 and 2, respectively. The tables include the wind speed and mixing height information as well. Similarly, **Tables 4-3 and 4-4** provide a tabulation of the top 25 modeled concentrations at Sites 1 and 2 for AERMOD-Default. The same information is provided for Massena_MOD (**Table 4-5 and 4-6**).

For the comparison of daytime vs. night hours, a daytime hour is defined as one with a defined convective mixing height (negative Monin-Obukhov length). The top 25 observed hourly SO_2 concentrations at Site 1 during 2017-2021 had all 25 events occurring at night, and Site 2 had 20 hours occurring at night. In comparison, AERMOD-Default's top 25 modeled hours had all 25 hours at night for both Site 1 and Site 2 while Massena_MOD's top 25 modeled hours at Site 1 had 16 hours at night and had all 25 hours at night at Site 2.

One of the most important points of comparison is the distribution of wind speeds for the top 25 observed and modeled hours. Over the top 25 observed hours, the average wind speed was about 4.5 meters per second (m/s) at Site 1 and 4.8 m/s at Site 2. For AERMOD-Default, the averages for both sites were significantly lower at 0.8 m/s, unexpectedly low wind speeds that indicate a significant mismatch in the meteorological conditions between the peak observations and peak modeled hours. The average wind speed for the top 25 modeled hours with Massena_MOD is 1.4 m/s at Site 1 and 2.3 m/s at Site 2. The Massena_MOD averages were closer to the observations, indicating that the site-specific modeling approach performs better than AERMOD-Default at matching the wind speed conditions associated with peak concentration hours. **Figures 4-1 and 4-2** show the top 25 ranked modeled and observed concentrations vs. wind speeds in plot form, showing Massena_MOD's improvement upon AERMOD-Default with more consistent wind speed distribution in comparison to observations.

Yet another point of comparison is the distribution of mixing heights for the top 25 observed and modeled hours. For each hour, the higher of the mechanical and convective mixing heights was selected, with only the mechanical mixing height for stable hours. Over the top 25 observed hours, the average mixing height was 498 m at Site 1 and 438 m at Site 2. The average mixing height over the top 25 modeled hours with AERMOD-Default at both sites was much lower at 49 m at Site 1 and 37 m at Site 2 whereas Massena_MOD's were 115 m at Site 1 and 131 m at Site 2. These averages were closer to the observations, indicating that the site-specific modeling approach performs better than AERMOD-Default at matching the mixing heights associated with peak concentration hours. **Figures 4-3 and 4-4** show this data in plot form.

Overall, these results indicate that the best performing model that matches the top 25 hourly concentrations more closely with observations is Massena_MOD.

Table 4-1: Top 25 Ranked Observed Concentrations at Site 1

Rank	Date	Hour	Monitored SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	3/22/2018	20	202.13	236	3.2	263	--
2	4/8/2017	21	167.08	222	3.4	285	--
3	2/27/2017	18	149.68	245	3.6	325	--
4	3/22/2018	19	136.66	242	3.2	356	--
5	2/14/2019	3	123.19	243	5.0	533	--
6	4/8/2017	20	120.31	238	3.1	302	--
7	2/27/2018	2	119.97	230	3.2	210	--
8	3/6/2019	3	119.71	244	4.1	507	--
9	4/3/2019	23	119.39	249	7.4	1255	--
10	12/19/2017	24	118.97	244	5.5	621	--
11	2/13/2017	2	118.69	241	4.3	418	--
12	2/14/2019	2	117.22	246	5.8	667	--
13	3/14/2021	3	116.77	240	3.6	411	--
14	3/6/2019	1	116.12	234	3.5	295	--
15	2/12/2020	4	115.78	247	5.3	592	--
16	2/13/2017	1	115.44	246	4.8	490	--
17	3/17/2020	24	115.33	239	4.1	392	--
18	3/4/2020	23	113.18	237	3.5	308	--
19	2/16/2021	20	112.63	241	5.7	648	--
20	3/14/2021	1	111.85	240	3.5	389	--
21	2/6/2021	19	110.93	248	5.2	564	--
22	2/11/2020	24	110.59	241	4.8	501	--
23	3/4/2019	23	110.20	247	4.6	608	--
24	1/25/2019	19	109.18	247	6.3	764	--
25	2/16/2019	6	108.99	249	6.2	735	--

Table 4-2: Top 25 Ranked Observed Concentrations at Site 2

Rank	Date	Hour	Monitored SO_2 Concentration ($\mu g/m^3$)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	5/23/2020	20	301.17	48	5.5	506	--
2	2/1/2021	4	287.91	31	2.8	145	--
3	2/1/2021	5	272.24	35	3.1	169	--
4	2/1/2021	2	272.09	35	4.2	269	--
5	2/1/2021	3	269.60	25	3.0	215	--
6	2/1/2021	19	267.37	58	4.7	322	--
7	4/4/2017	2	265.43	58	6.8	691	--
8	4/24/2020	7	257.10	41	4.4	402	200
9	5/25/2017	8	256.71	50	5.7	613	674
10	4/8/2021	18	256.08	44	6.9	747	716
11	2/1/2021	6	255.03	30	2.5	117	--
12	5/18/2018	21	252.99	50	5.0	433	--
13	1/24/2020	6	250.58	46	4.2	269	--
14	1/31/2021	23	250.31	39	2.8	144	--
15	4/16/2018	21	249.76	53	6.0	598	--
16	4/8/2021	19	247.56	49	6.2	615	--
17	5/18/2018	20	246.59	48	5.3	474	--
18	5/25/2017	9	243.90	48	6.3	707	948
19	2/7/2018	14	243.63	49	5.8	484	113
20	2/2/2021	2	243.63	44	4.2	270	--
21	5/25/2017	21	240.62	52	7.2	757	--
22	3/26/2017	5	240.23	68	2.9	221	--
23	4/24/2020	6	240.23	51	3.0	197	--
24	5/12/2017	19	237.14	45	5.2	467	--
25	3/18/2017	18	236.17	50	7.5	826	--

Table 4-3: Top 25 Ranked Modeled Concentrations for AERMOD Default Model at Site 1

Rank	Date	Hour	Modeled SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	4/10/2021	23	879.79	253	0.8	49	--
2	2/20/2017	19	878.71	251	0.8	42	--
3	2/2/2020	22	872.66	252	0.7	40	--
4	12/3/2019	7	864.69	252	0.8	42	--
5	2/24/2020	20	841.93	254	1.1	52	--
6	8/21/2020	20	828.47	253	1.0	61	--
7	8/7/2020	22	811.51	250	0.7	47	--
8	4/14/2017	4	809.39	252	0.7	45	--
9	5/7/2018	20	798.84	253	0.9	54	--
10	12/19/2020	1	797.78	253	0.7	42	--
11	7/9/2021	21	797.57	252	0.8	50	--
12	1/13/2021	1	796.00	254	1.0	50	--
13	4/18/2020	5	794.85	254	0.9	52	--
14	4/23/2017	3	789.89	253	1.0	56	--
15	4/23/2017	4	784.90	251	0.8	48	--
16	7/9/2021	20	783.84	253	0.8	56	--
17	4/1/2020	23	783.28	255	0.8	47	--
18	4/22/2017	24	777.55	251	0.9	52	--
19	9/16/2018	24	776.23	250	0.7	45	--
20	3/18/2018	4	767.06	253	0.8	52	--
21	3/30/2018	24	751.55	254	1.0	56	--
22	7/15/2018	21	750.75	250	0.8	48	--
23	6/6/2018	21	746.35	252	0.7	45	--
24	6/10/2018	22	744.45	252	0.7	46	--
25	9/13/2021	3	744.07	251	0.8	48	--

Table 4-4: Top 25 Ranked Modeled Concentrations for AERMOD Default Model at Site 2

Rank	Date	Hour	Modeled SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	10/24/2017	17	1411.35	55	0.7	35	--
2	7/10/2020	4	1408.26	52	0.7	34	--
3	9/13/2018	4	1387.71	55	0.7	35	--
4	5/23/2017	24	1373.33	54	0.8	36	--
5	8/10/2021	2	1367.93	54	0.8	37	--
6	2/26/2021	22	1365.01	54	1.0	36	--
7	8/25/2021	6	1357.62	54	0.7	35	--
8	8/16/2020	6	1354.40	54	1.0	41	--
9	8/27/2020	6	1350.81	51	0.9	39	--
10	4/14/2021	22	1344.95	53	0.7	35	--
11	7/15/2018	20	1328.66	52	0.6	33	--
12	3/26/2020	7	1318.88	55	0.9	35	--
13	8/14/2020	24	1318.41	51	0.6	34	--
14	8/15/2020	23	1317.35	51	1.0	47	--
15	7/12/2021	3	1303.96	55	0.8	38	--
16	2/10/2018	20	1299.62	53	0.9	35	--
17	10/7/2017	1	1295.92	55	0.7	34	--
18	5/28/2021	24	1295.74	55	0.8	36	--
19	4/17/2020	24	1291.78	53	0.8	38	--
20	8/22/2020	21	1286.81	57	0.8	37	--
21	4/20/2020	23	1285.34	53	0.8	38	--
22	5/28/2018	4	1285.27	54	0.9	39	--
23	5/13/2018	5	1284.57	55	0.8	36	--
24	2/22/2017	5	1282.55	55	1.0	36	--
25	2/14/2018	1	1281.04	55	0.9	36	--

Table 4-5: Top 25 Ranked Modeled Concentrations for Massena_MOD Model at Site 1

Rank	Date	Hour	Modeled SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	3/22/2021	8	515.00	258	0.7	65	57
2	6/24/2021	6	410.64	244	0.7	60	72
3	2/7/2021	12	374.48	248	1.2	81	94
4	9/21/2020	8	353.80	250	0.7	74	83
5	3/23/2021	8	320.69	263	0.7	62	96
6	9/3/2018	22	311.33	254	1.2	73	--
7	3/21/2021	8	301.32	253	1.1	108	63
8	1/21/2017	20	283.79	256	1.4	73	--
9	7/30/2020	21	282.98	254	1.2	75	--
10	9/22/2020	8	282.04	253	1.1	121	106
11	7/4/2020	24	269.98	254	1.1	73	--
12	5/19/2021	22	269.15	255	1.2	72	--
13	4/7/2019	10	268.52	257	0.9	95	106
14	2/1/2020	3	266.78	254	1.4	69	--
15	8/21/2020	20	263.96	253	1.0	61	--
16	8/11/2020	22	261.76	252	3.4	402	--
17	8/17/2020	8	259.94	247	1.1	110	127
18	2/21/2021	5	259.69	251	1.8	106	--
19	12/26/2017	20	254.62	255	1.7	99	--
20	7/11/2018	22	253.08	254	1.1	69	--
21	1/16/2021	22	252.69	253	2.1	132	--
22	5/23/2017	20	251.46	254	1.6	110	--
23	7/12/2020	5	251.24	253	3.1	365	--
24	12/23/2021	16	249.72	253	1.8	105	--
25	1/15/2020	1	246.96	255	2.0	127	--

Table 4-6: Top 25 Ranked Modeled Concentrations for Massena_MOD Model at Site 2

Rank	Date	Hour	Modeled SO_2 Concentration ($\mu g/m^3$)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	8/14/2020	19	678.08	53	2.3	204	--
2	7/15/2020	23	674.78	54	2.1	113	--
3	4/10/2021	19	662.89	52	2.1	112	--
4	7/15/2020	20	656.27	51	2.3	149	--
5	7/15/2020	22	651.18	51	2.2	128	--
6	8/4/2020	5	636.71	51	1.8	87	--
7	6/30/2017	2	622.88	54	2.1	117	--
8	6/22/2020	20	616.93	57	1.9	145	--
9	9/10/2018	3	614.78	53	2.4	139	--
10	2/18/2021	18	614.49	53	2.9	152	--
11	9/16/2018	20	608.30	56	1.6	76	--
12	2/20/2018	19	607.39	56	2.5	120	--
13	8/29/2021	6	607.04	52	2.5	149	--
14	8/28/2021	22	605.24	51	2.3	133	--
15	8/13/2018	22	605.13	53	1.9	102	--
16	2/20/2018	5	604.18	56	2.4	107	--
17	5/18/2020	2	602.69	53	2.6	156	--
18	9/22/2021	18	602.64	53	2.2	124	--
19	6/10/2020	3	600.84	51	2.2	126	--
20	5/30/2021	21	597.97	53	2.5	147	--
21	6/4/2017	22	595.74	53	2.5	154	--
22	5/21/2017	2	594.99	56	1.9	95	--
23	6/4/2017	23	593.25	55	2.1	114	--
24	2/19/2021	8	592.94	52	3.1	164	--
25	8/27/2021	21	592.50	56	2.5	153	--

Figure 4-1: Top 25 Ranked Observed and Modeled Concentrations vs. Wind Speeds for Site 1

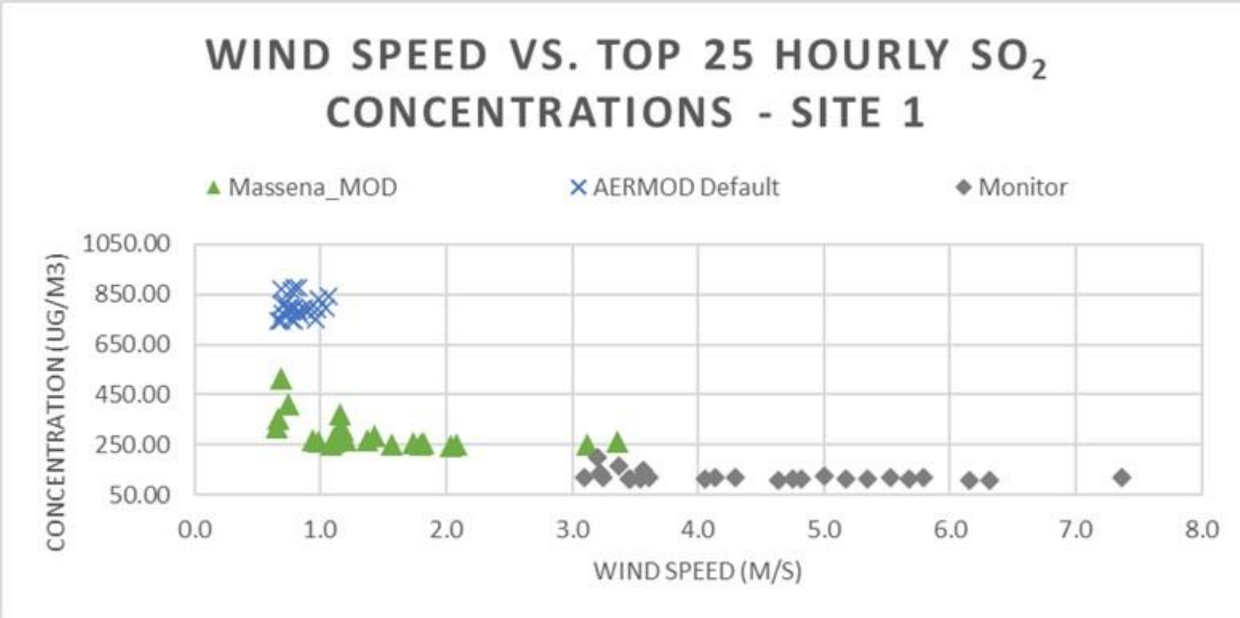


Figure 4-2: Top 25 Ranked Observed and Modeled Concentrations vs. Wind Speeds for Site 2

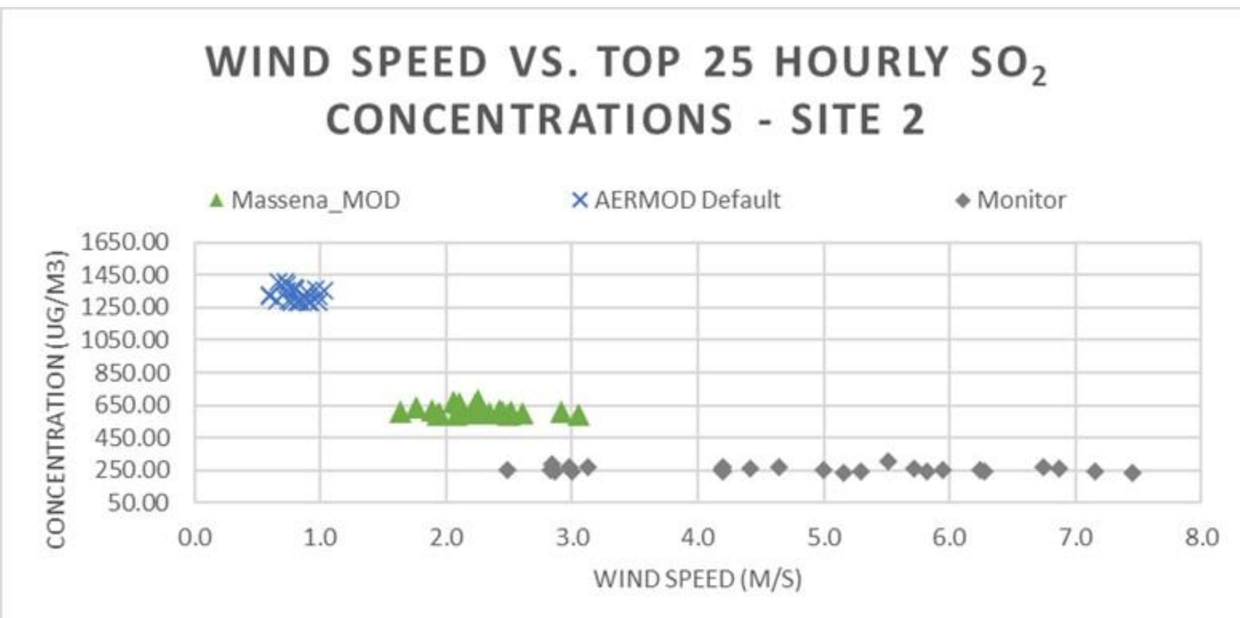


Figure 4-3: Top 25 Ranked Observed and Modeled Concentrations vs. Mixing Heights for Site 1

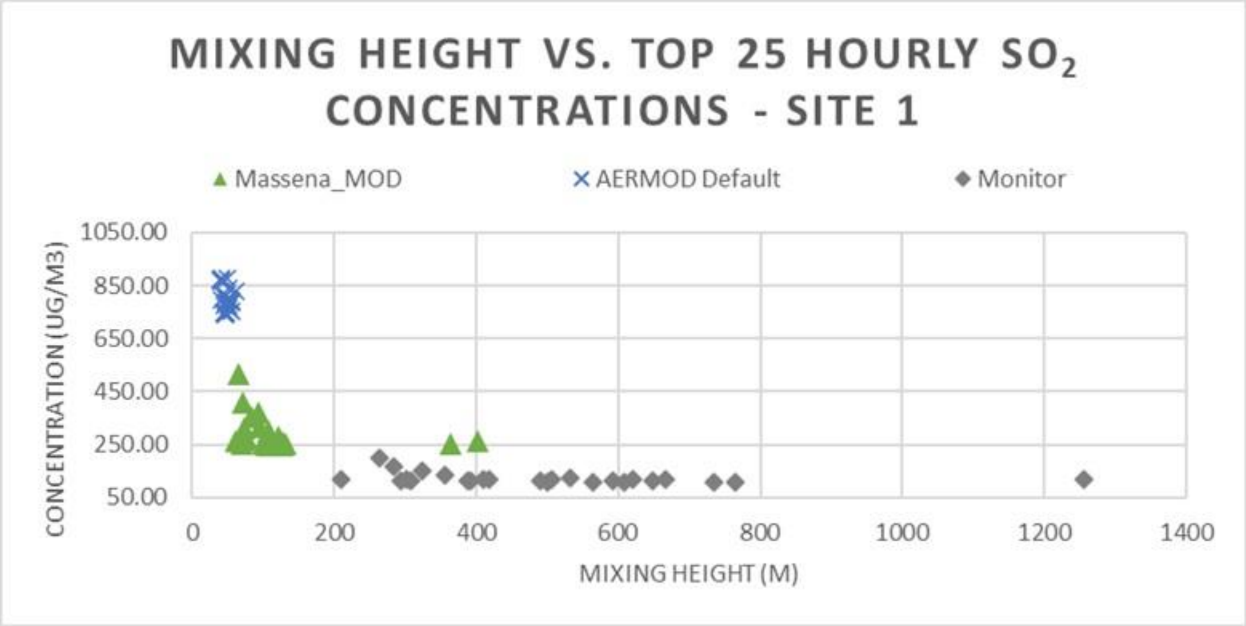
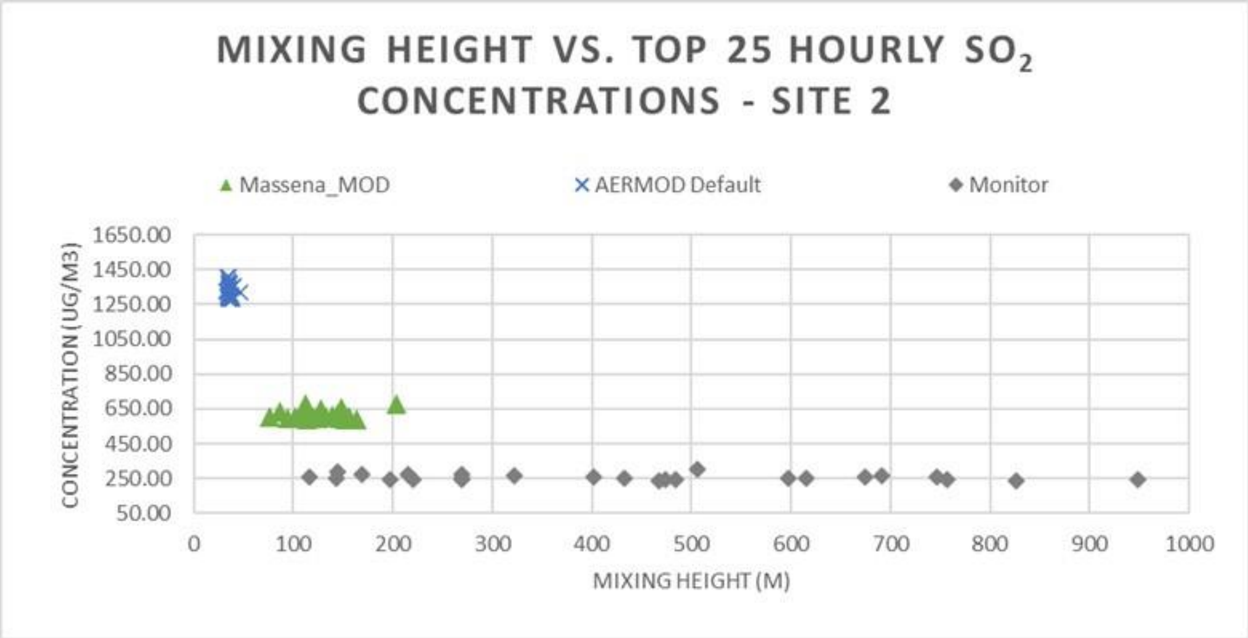


Figure 4-4: Top 25 Ranked Observed and Modeled Concentrations vs. Mixing Heights for Site 2



4.2 Results of Statistical Performance Tests

Three sets of statistical evaluation tests were conducted: a) quantile-quantile (Q-Q) plots for each monitor, b) comparison of the modeled and observed 5-year average 1-hour average design concentration[22] for each monitor, and c) the use of the Robust Highest Concentration (RHC) as part of EPA's Cox-Tikvart[23] procedure as described in EPA's 1992 model evaluation procedures[6], as implemented in EPA's Model Evaluation Methodology (MEM) software.[24] Further discussion of each of these tests and the results of the testing are provided below.

Quantile-Quantile Plots

Operational performance of models for predicting compliance with air quality regulations, especially those involving a peak or near peak value at some unspecified time and location, can be assessed with Q-Q plots (Chambers et al., 1983).[25] Q-Q plots are created by sorting by rank the predicted and the observed concentrations from a set of predictions initially paired in time. The sorting is generally done for individual monitors, such that the comparison is still paired in space, but not in time. The sorted sets of predicted concentrations are then plotted by rank against the observed concentrations, which are also sorted by rank. While these concentration sets are no longer paired in time, the plot is useful for answering the question, "Over a period of time, does the distribution of the model predictions match those of observations?" Scatterplots, which use data paired in time and space, provide a stricter test, answering the question: "At a given time and place, does the magnitude of the model prediction match the observation?" It is the experience of model developers (e.g., Weil, et al., 1992[26] and Liu and Moore, 1984[27]) that wind direction uncertainties can and do cause disappointing scatterplot results from what are otherwise well-performing dispersion models. Therefore, the Q-Q plot instead of the scatterplot is a more pragmatic procedure for demonstrating model performance of applied models. Venkatram (2001)[28] further discusses the attributes for the use of Q-Q plots for evaluating regulatory models.

The Q-Q plot for Site 1, which compares the default model with the site-specific approach, is presented in **Figure 4-5** and the Q-Q plot for Site 2 is presented in **Figure 4-6**. In general, since SO_2 monitors have a +/- 10% tolerance for calibration accuracy,[29] ranked modeled concentrations that are within 10% of a "perfect model" result are considered to be unbiased. EPA's 1992 model evaluation procedures[30] indicate that an acceptable model should have peak predictions within a factor of 2 of observations. For both sites, it is clear that the default model grossly overpredicts by about a factor of 5 or more (aside from the highest concentration). The Massena_MOD model at Site 1 and Site 2 is generally showing ranked modeled-to-observed pairs with overpredictions between a factor of 2 and a factor of 3. Based on these Q-Q plot results, the performance of Massena_MOD is clearly better than that of the default model, though substantial overprediction remains.

---

[22] The "design concentration" is the 99th percentile of the peak daily 1-hour maximum concentration computed for each calendar year, and averaged over the five years included in the evaluation. For any given year, assuming that there are at least 301 days with valid peak daily observations, the fourth highest daily 1-hour maximum would constitute the 99th percentile value.
[23] William M. Cox and Joseph A. Tikvart, 1990. A statistical procedure for determining the best performing air quality simulation model, *Atmospheric Environment*. Part A. General Topics, Volume 24, Issue 9, Pages 2387-2395. ISSN 0960-1686, https://doi.org/10.1016/0960-1686(90)90331-G.
[24] Strimaitis, D., E. Insley, M. Korc, and F. Lurmann, 1993. User's Guide for the Model Evaluation Methodology (MEM System for Comparing Model Performance Version 1.0. STI-93261-1392-FR.
[25] Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A., 1983. Chapter 3: Comparing Data Distributions. *Graphical Methods for Data Analysis. (Bell Laboratories).* Wadsworth International Group and Duxbury Press.
[26] Weil J.C, Sykes and Venkatram A., 1992. Evaluating air-quality models: Review and outlook. *J. Appl. Met*., 31, p 1121-1144.
[27] Liu, M. K., and G. E. Moore, 1984. Diagnostic validation of plume models at a plains site. EPRI Report No. EA-3077, Research Project 1616-9, Electric Power Research Institute, Palo Alto, CA.
[28] Venkatram, A., R. W. Brode, A. J. Cimorelli, J. T. Lee, R. J. Paine, S. G. Perry, W. D. Peters, J. C. Weil, and R. B. Wilson, 2001. A complex terrain dispersion model for regulatory applications. *Atm.Env*., 35, 4211-4221.
[29] Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II, Ambient Air Quality Monitoring Program, 2013, available at http://www.epa.gov/ttnamti1/files/ambient/pm25/qa/QA-Handbook-Vol-II.pdf. (Table 10-3 and Appendix D, page 13).
[30] EPA, 1992. Protocol for Determining the Best Performing Model. Publication No. EPA–454/R–92–025. Office of Air Quality Planning and Standards, Research Triangle Park, NC. (NTIS No. PB 93–226082). https://www.epa.gov/sites/default/files/2020-10/documents/model_eval_protocol.pdf.

Figure 4-5: Q-Q Plot for Site 1

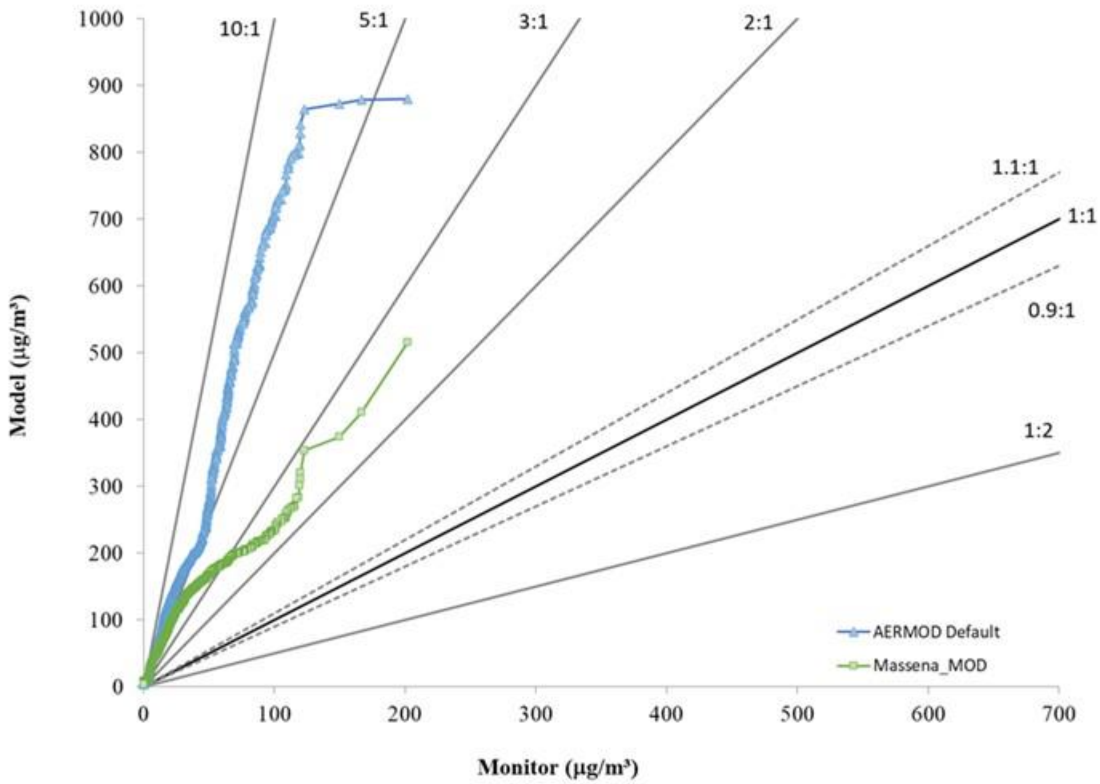
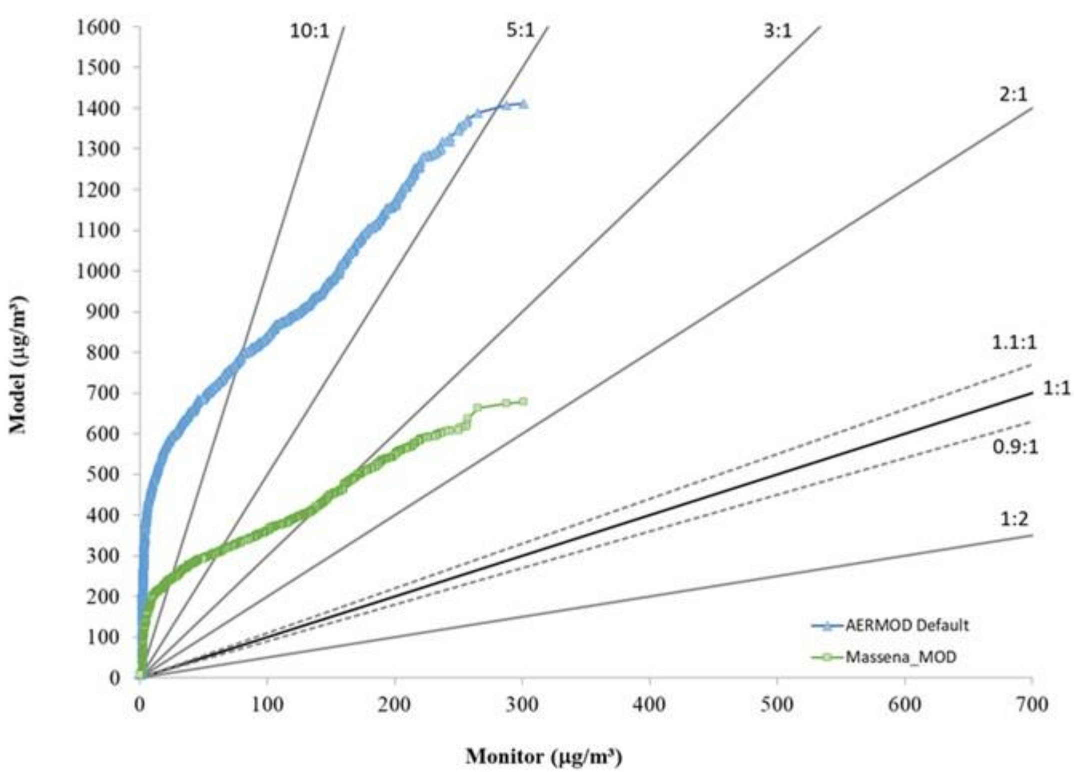


Figure 4-6: Q-Q Plot for Site 2



Comparison of Modeled and Predicted Design Concentrations

A key metric that should be included in any model evaluation involves the modeled and observed design concentration, which corresponds to the form of the ambient standard that is being evaluated. For SO_2, this is the 99th percentile peak daily 1-hour maximum concentration averaged over the 5-year modeling period. A separate calculation is made for each monitor. Similar to the discussion in the Q-Q plot metric noted above, a modeled-to-observed ratio ("MOR") design concentration within 10% indicates an unbiased model.

Tabulated values for this comparison are provided in **Table 4-7** for Site 1 and **Table 4-8** for Site 2. As can be seen in both tables, both models overpredict, but the default model has a MOR value of nearly 7 at Site 1 and over 5 at Site 2. In contrast, the Massena_MOD has MOR values of about 2.3 at Site 1 and 2.6 for Site 2. As a result, Massena_MOD improves model-to-monitor performance, although still overpredicting.

Table 4-7: Modeled-to-Observed Design Concentrations at Site 1

Model Approach	4th Highest Design Concentration (ug/m^3)						
	2017	**2018**	**2019**	**2020**	**2021**	**5-year Average**	
Observed	118.7	106.6	110.2	110.6	101.1	109.4	**MOR**
AERMOD Default Model	777.5	751.6	612.4	811.5	744.1	739.4	6.8
Massena_MOD Model	240.2	242.8	196.4	270.0	320.7	254.0	2.3

Table 4-8: Modeled-to-Observed Design Concentrations at Site 2

Model Approach	4th Highest Design Concentration (ug/m^3)						
	2017	**2018**	**2019**	**2020**	**2021**	**5-year Average**	
Observed	237.1	226.7	214.9	232.6	243.6	231.0	**MOR**
AERMOD Default Model	1282.5	1285.3	1031.0	1318.9	1345.0	1252.5	5.4
Massena_MOD Model	588.2	605.1	531.8	616.9	605.2	589.5	2.6

Testing with the Robust Highest Concentration

The Model Evaluation Methodology (MEM) software was designed to evaluate model performance by implementing the statistical analysis procedures contained in EPA's 1992 Protocol for Determining the Best Performing Model (EPA-454/R-92-025). MEM evaluates model performance through two stages. The first step is a screening test to flag models that fail to perform at a minimum operation level. The fractional bias (= 2 * (observed – predicted) / (observed + predicted)) of the mean and the fractional bias of the standard deviation are used to qualify performance. The fractional bias has been selected as the basic measure of performance in the MEM. Values for the fractional bias range between -2 and +2 (over prediction, under prediction). Also, the fractional bias is a good proxy for result comparisons. Fractional biases that are equal to -0.67 are equivalent to overpredictions by a factor of 2, while a fractional bias of +0.67 is equivalent to an underprediction by a factor of 2. The absolute fractional bias (AFB) statistic, which is just the absolute value of the fractional bias (FB), is computed for each of the individual models.

The second stage is a resampling technique (bootstrapping) which generates a probability distribution of possible data outcomes. Five years of data can be arranged into seasonal blocks (DFJ, MAM, JJA and SON); the MEM software can be recompiled to accept multiple years of data. Within each season, the pieces are sampled with replacement until a total season is created. This process is repeated using each of the four seasons to construct a complete bootstrap year. Sampling within seasons guarantees that each season will be represented by only days chosen from that season. Since sampling is done with replacement, some days are represented more than once, while other days are not represented at all. Next, the data generated for the bootstrap year are used to calculate the composite performance measures for each model. This process is repeated until sufficient samples are available to calculate a meaningful standard error, which is the standard deviation of the measure over all of the bootstrap-generated outcomes (the sample size of which has been set to 1,000).

The method of bootstrapping is used to estimate the standard error of the composite performance measure of each model. Using this estimation, the statistical significance of the difference between models is then assessed. A test statistic, the Robust Estimate of the Highest Concentration (RHC), is then conducted within MEM using a subset, N, of the highest concentrations (Equation 1).

$$RHC = X(N) + [\bar{X} - X(N)] * \ln\left[\frac{3N-1}{2}\right] \quad \text{(Equation 1)}$$

where:

N = number of values;

$\bar{X}$ = average of the N-1 largest values; and

$X(N)$ = Nth largest value.

The assigned number of values, N, typically ranges between 11 and 26; 26 is suggested for this application.

After the RHC calculations, the model comparison statistics are then conducted. The first comparison measure that is calculated is the Composite Performance Measure (CPM). The CPM is a weighted linear combination of the individual fractional biased components. A CPM is calculated for each model (Equation 2).

$$CPM = \frac{1}{3} * \overline{AFB_{r,s}} + \frac{2}{3}\left[\frac{(AFB)_3 + (AFB)_{24}}{2}\right] \quad \text{(Equation 2)}$$

where:

$(AFB)_{r,s}$ = Absolute Fractional Bias for diagnostic conditions r at station s;

$(AFB)_3$ = Absolute Fractional Bias for 3-hour averages; and

$(AFB)_{24}$ = Absolute Fractional Bias for 24-hour averages.

The final performance measure calculated is the Model Comparison Measure (MCM) – mean and confidence interval. The MCM is the difference between the CPM for two models (Equation 3).

$$MCM_{i,j} = CPM_i - CPM_j \quad \text{(Equation 3)}$$

where:

CPM_i = Composite Performance Measure for model i; and

CPM_j = Composite Performance Measure for model j.

The magnitude and sign of the MCM are indicative of the relative performance of each pair of models. The smaller the composite performance measure, the better the overall performance of a model. This means that for two arbitrary models, Model A and Model B, a negative MCM (i.e., the difference between the CPM for Model A and Model B) implies that Model A is performing better (Model A has the smaller CPM), while a positive value indicates that Model B is performing better. For each pair of model comparisons, the significance of the MCM depends upon whether or not its confidence interval, e.g., 90th percentile, overlaps zero. If the confidence interval overlaps zero, the two models may not be performing at a level which is statistically different, although one model may still exhibit a notable tendency to have a lower bias, and therefore can be judged to have superior performance. If the confidence interval does not overlap zero, (upper and lower limits are both negative or both positive), then there exists a statistically significant difference between the two models at the stated level of confidence. In previous work, EPA has used a 90th percentile level of confidence.[31] The MEM software also computes a Combined Model Comparison Measure (CMCM) to provide a model performance assessment over all monitoring sites.

Table 4-9 shows the average fractional biases for the modeling approaches at both monitors for the RHC estimate, where a value of zero would indicate a perfect model. A negative fractional bias indicates that the model is overpredicting. A model overprediction by a factor of 2 results in a fractional bias value of -0.67, while an underprediction by a factor of 2 results in a fractional bias of +0.67. The resulting overprediction ratio for the default model is much higher for both sites with the default model than with Massena_MOD, where Massena_MOD has fractional bias values closer to (though still greater than) a factor of 2 overprediction.

Table 4-9: Average Fractional Biases for Site 1 and Site 2 for RHC Estimate

Site	Model Approach	FB_{Avg}
Site 1	AERMOD Default Model	-1.47
	Massena_MOD Model	-0.81
Site 2	AERMOD Default Model	-1.35
	Massena_MOD Model	-0.83

Tables 4-10 and 4-11 show the RHC for Site 1 and Site 2, respectively. A ratio of model predicted RHC to the observed RHC at or slightly above 1.0 is considered ideal. As seen in tables, the default model runs have RHC predicted-to-observed ratios between 5 and 7, while Massena_MOD has RHC ratios range between 2.1 and 2.4.

Figures 4-7 and 4-8 show the CPM values for Sites 1 and 2, respectively, and **Figure 4-9** shows the MCM values for both monitors separately and the combined (CMCM) for Massena_MOD vs. the default model. The CPM values for Masssena_MOD are smaller than the default model approach. Also, as expected, the 90% confidence intervals for the CPMs for Massena_MOD do not overlap (by a wide margin) with the default model, meaning that a statistically significant difference exists between the performance of Massena_MOD vs. the default model. This is reflected in the MCM values, which are all positive, meaning that “model J” (Massena_MOD) performs much better than “model I” (default model).

[31] See, for example, the EPA 2015 presentation available at: ftp://newftp.epa.gov/Air/aqmg/SCRAM/conferences/2015_11th_Conference_On_Air_Quality_Modeling/Presentations/1-5_Proposed_Updates_AERMOD_System.pdf.

Table 4-10: 5-Year Averaged Robust High Concentrations (µg/m³) for Site 1

Model Approach	RHC	Pre/Obs Ratio
Observed	145.6	--
AERMOD Default Model	1028.2	7.1
Massena_MOD Model	309.3	2.1

Table 4-11: 5-Year Averaged Robust High Concentrations (µg/m³) for Site 2

Model Approach	RHC	Pre/Obs Ratio
Observed	282.0	--
AERMOD Default Model	1425.5	5.1
Massena_MOD Model	688.1	2.4

Figure 4-7: Plot of CPM for Site 1

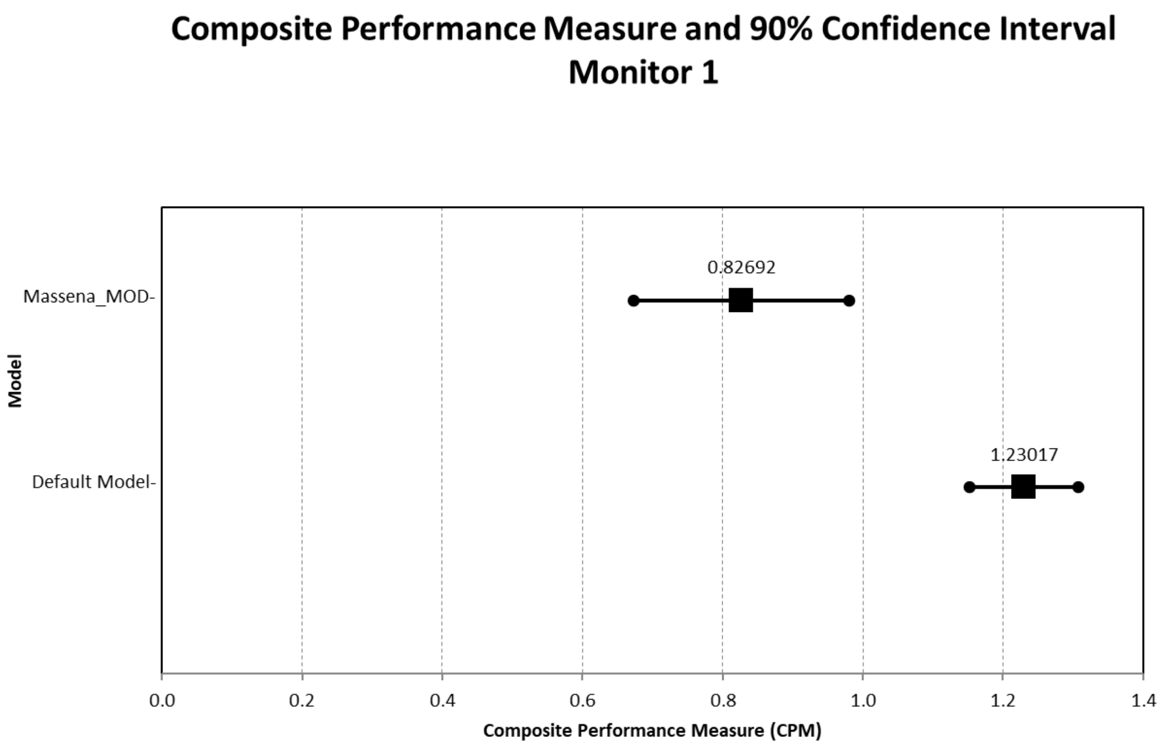


Figure 4-8: Plot of CPM for Site 2

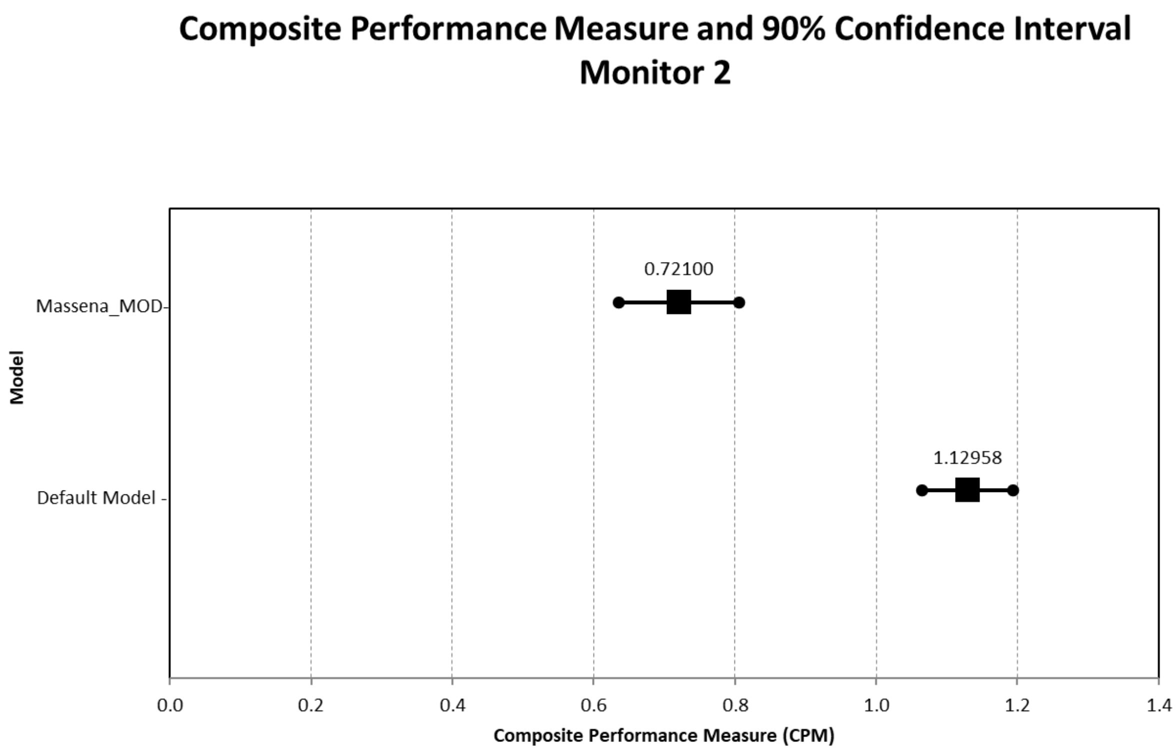
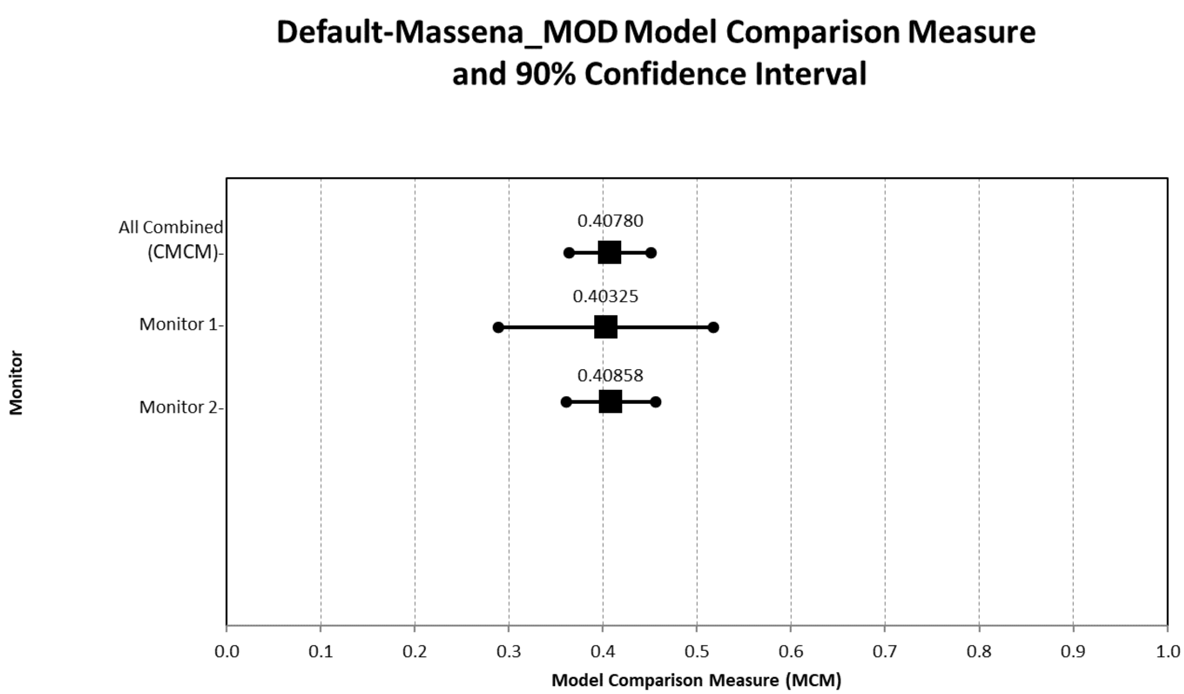


Figure 4-9: AERMOD Default vs. Massena_MOD MCM, CMCM, and 90% Confidence Interval for Sites 1 and 2



5. Use of Site-Specific Model for NAAQS Compliance Modeling

This protocol proposes the use of a site-specific modeling approach for Massena Operations based upon an updated model evaluation with five-year of SO_2 and meteorological measurements at two ambient air monitoring sites. Scientific justification is presented for the proposed alternative model for Alcoa's Massena Operations (Massena_MOD) in comparison to AERMOD-Default. The results of the model evaluation elements described above indicate that AERMOD-Default grossly overpredicts at the two monitors while the Massena_MOD approach overpredicts less and it performs better on all of the evaluation tests. Therefore, we conclude that Massena_MOD outperforms AERMOD-Default by a statistically significant margin.

Alcoa will not be changing its method of operation with the facility changes. A solution has been identified in which groups of dry scrubber stacks along the potlines will likely be merged and/or raised. The merging is creditable because the plantwide SO_2 emission limit will be less than the 5,000 tons per year referenced in the 1985 Stack Height Regulations for this modeling credit. The final stack heights will be below the default Good Engineering Practice (GEP) height of 65 m.

Alcoa has evaluated several stack design alternatives to bring the area into SO_2 NAAQS attainment. The most feasible that has been identified at this time is a dual 50-m stack option that combines three of the six reactor baghouses into one stack, and the other three into a second stack, and positioning the stacks within one diameter of each other. Specifically, for each courtyard, there would be 2 stacks planned to be located within 1 stack diameter's distance from one another. Each set of 2 stacks would utilize a common support structure of scaffolding and each individual stack would serve 3 of the 6 reactors (i.e., compartments of the common air pollution control device, or dry alumina scrubber) in a given courtyard (see **Figure 5-1**).

With the stacks configured in this manner, the two stacks of equal height would qualify them as a "continuous source". This means that these 2 stacks per courtyard will have distinct emission points close enough together such that it will result in a merged plume.[32] Therefore, each pair of stacks would be merged and treated as a single stack for modeling purposes, following established EPA policy. As support for this approach, the EPA Model Clearinghouse concluded that, for a model application in New Jersey, since the stacks were located within 1 stack diameter of one another *"the 'gap' can be considered closed and the 3 stacks a continuous source".*[33] The Model Clearinghouse also referenced this 1990 case as an example of when plume merging is appropriate in a 1996 communication with Ohio EPA regarding the "Interpretation of Modeling Guidance Plume Rise".[34] The 1996 communication indicated *"The logic for this decision (to allow plume merging) is based in Section 3.3.2 of the GEP Stack height guideline, buildings that are sufficiently close together should be treated as a single building for purposes of determining L in the stack height formula. This logic was extended to closely separated stacks, with the general result that if the stacks are separated by less than their width (diameter), they could be treated as one (C/H Record 91-II-01)."*

To account for this effect in the model, each pair of stacks would be modeled as one equivalent stack with equivalent stack parameters, consistent with modeling conducted for other approved model applications such as Plant Scherer[35] in Monroe County, GA and Stuart Generating Station[36] in Adams County, OH, as well as others[37,38]. Equivalent stack parameters would represent the combined flow for both stacks along with an equivalent stack diameter. The

[32] Hanna, 1982. Handbook on Atmospheric Dispersion, "[t]he reduced entrainment and increased buoyancy of the merged plumes may increase plume rise significantly."). https://www.osti.gov/biblio/5591108
[33] EPA, 1990. https://cfpub.epa.gov/oarweb/MCHISRS/index.cfm?fuseaction=main.resultdetails&recnum=91-II%20%20-01.
[34] EPA, 1992. https://cfpub.epa.gov/oarweb/MCHISRS/index.cfm?fuseaction=main.resultdetails&recnum=96-V%20%20%20-10
[35] Georgia Environmental Protection Division, 2015. SO_2 Designations Round 2 State Recommendations, Attachment 2, see Source Data. https://www.epa.gov/sulfur-dioxide-designations/so2-designations-round-2-georgia-state-recommendation-and-epa-response
[36] Ohio EPA, 2016. SO_2 Designations Round 3 State Recommendations, Appendix Y, see Emission Sources. https://epa.ohio.gov/divisions-and-offices/air-pollution-control/state-implementation-plans/division-of-air-pollution-control-sip-so2-so2-app8
[37] Alaska Department of Environmental Conservation, 2018. Modeling Review Procedures Manual, see Stack Modifications. https://dec.alaska.gov/air/air-permit/dispersion-modeling/
[38] Pima County Department of Environmental Quality, 2022. Air Quality Permit #1052 Technical Support Document, see Emission and Stack Data. https://content.civicplus.com/api/assets/85a50a58-8848-43a2-a91a-ec348b890917

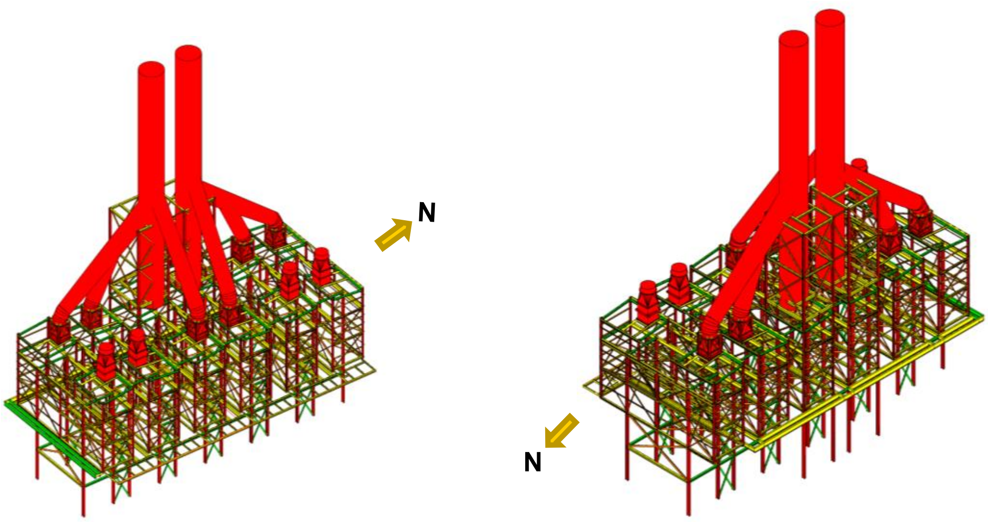
equivalent diameter would be calculated such that it results in the combined area for the individual stacks (which, for a 2-stack arrangement with the same stack diameter, is equal to √2 times the diameter of a single stack). The equivalent stack diameter would then be used along with the combined flow rate to calculate a representative equivalent stack exit velocity. The 2 future stacks in each courtyard would be built identical to one another (e.g., stack height, diameter, serving 3 reactors each for the common air pollution control device); therefore, flows and velocities would also be expected to be the same under normal operating conditions. Alcoa will consider conditions for which reactor outages could result in alternative stack parameters. Cases that occur very infrequently will not be considered in accordance with the EPA's March 1, 2011 guidance[8] involving situations that occur so infrequently that they are not expected to affect the probabilistic 1-hour SO_2 NAAQS.

When the proposed modeling is conducted to demonstrate future SO_2 NAAQS attainment, the 1-hour emission rates determined to demonstrate compliance (i.e., critical emission values, or CEVs) for the potline reactor stacks will be presented as 1 CEV per courtyard. Presentation of CEVs in this manner will be consistent with the proposed modeling of 1 merged stack per courtyard.

Multiple combinations of emission rates for the three merged courtyard reactor stacks may be modeled to provide Alcoa with a range of compliant SO_2 emission rates. This will be done due to the possibility that for any given hour, SO_2 emissions at one of the courtyards could be higher than those at the other two courtyards.

A separate modeling report will be submitted using the modeling procedures that are approved by EPA following review of this modeling protocol to demonstrate future SO_2 NAAQS compliance with five years of Massena airport data (2017-2021). Pending approval of this modeling protocol, Alcoa will proceed with design of the dual 50-m stack alternative including representing these stacks as 1 equivalent stack per courtyard for modeling purposes and developing CEVs for these sources as 1 CEV per courtyard.

Figure 5-1: Planned Future Dual 50-m Stack Configuration



6. Conclusions

EPA approved the use of Massena_MOD in October 2022 with site-specific use of a 100-m neutral lapse rate and a more appropriate version of the Building Profile Input Program (BPIP) for this site. However, changes in the site-specific model are needed due to the need to separately model the roof vent emissions if the future potline dry scrubber stacks are to be raised.

This protocol describes in the detail the scientific justification for the proposed improvements. In addition, the protocol provides the results of model evaluation results for several aspects of the model performance at the two monitoring sites: 1) quantile-quantile plots of the modeled and observed concentrations, 2) meteorological conditions associated with the peak 25 modeled and observed concentrations, and 3) a full evaluation using EPA's model evaluation protocol procedures as implemented using the Model Evaluation Methodology software. For the key Site 2 monitor, the model evaluation results indicate that the ratio of the model to observed design concentration with a 5-year data set improves from 5.4 with AERMOD default and 2.6 with Massena_MOD. Therefore, Alcoa proposes the use of the revised Massena_MOD site-specific modeling approach.

Alcoa wishes to work collaboratively with NYSDEC to resolve the SO_2 nonattainment issue and thus, will move forward at this time with Massena MOD notwithstanding its material overprediction tendency. Alcoa maintains that the remaining overprediction tendency for Massena_MOD may result in an over-engineered solution to the SO_2 nonattainment designation. As with the November 2023 protocol, discussion of previously proposed site-specific approaches, “Massena_MOD+ENTRAIN” and “Massena_MOD+ENTRAIN 0.2” fromthe June 2023 protocol are included in **Appendix B** as these approaches have shown improved model evaluation performance relative to Massena_MOD.

Appendix A Facility Maps of Buildings and SO_2 Emission Sources

Facility maps are provided herein to identify the locations and names of each of the modeled buildings and SO_2 sources. Each SO_2 source is labeled as the modeled ID. Definitions of the modeled source IDs are provided below.

Modeled Source ID Definitions:
RS = Potline dry scrubber stacks/reactor stacks
BF2 = Anode bake furnace
B401/B402 = Potline roof vents, modeled as a series of stacks along each roof vent
AI = Homogenizing heat treat furnaces
M0 = Melters/holders; furnaces (M024E)
B000A_B = Package boilers

Figure A-1: Facility Layout of Modeled Buildings and SO_2 Sources – Northern Section

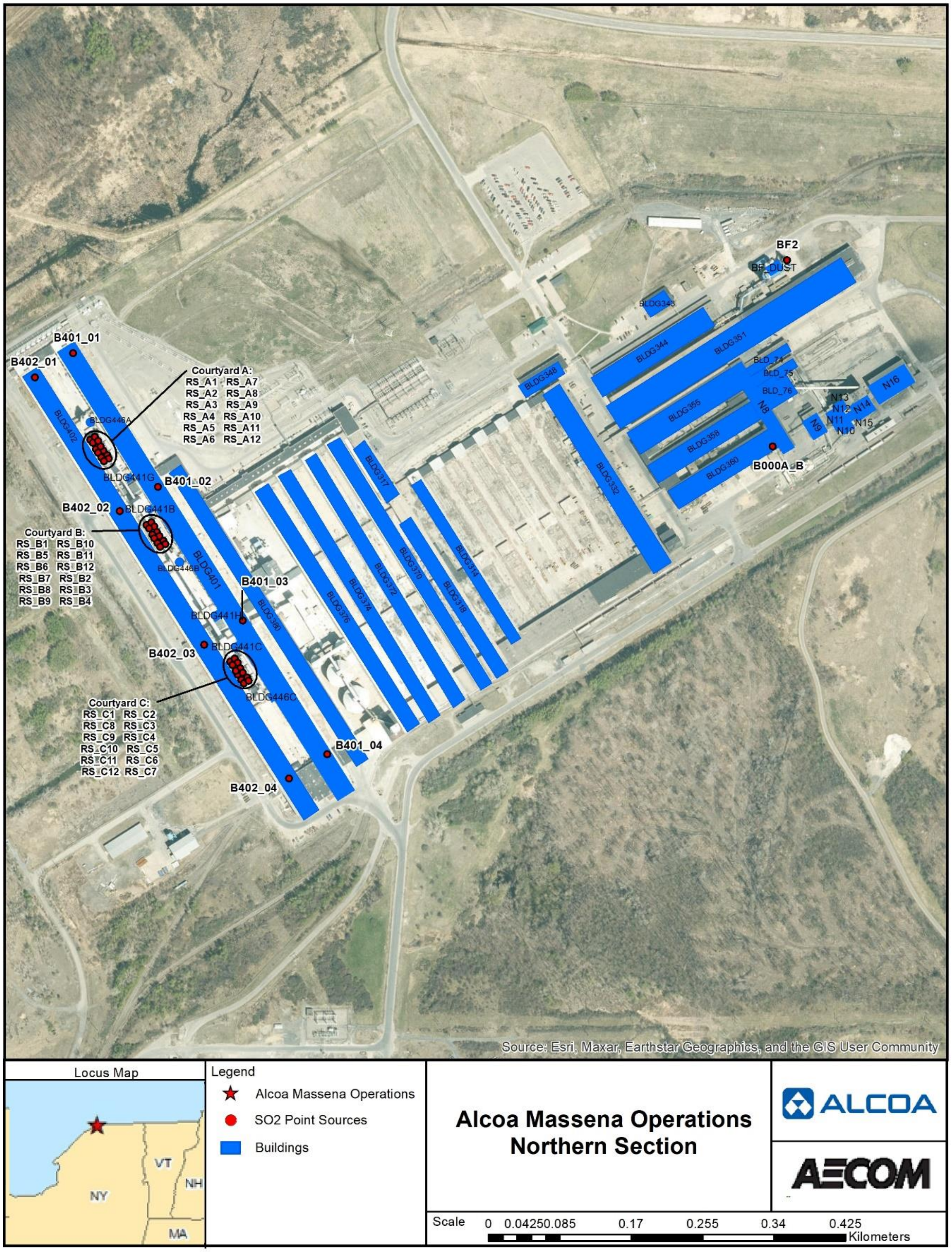
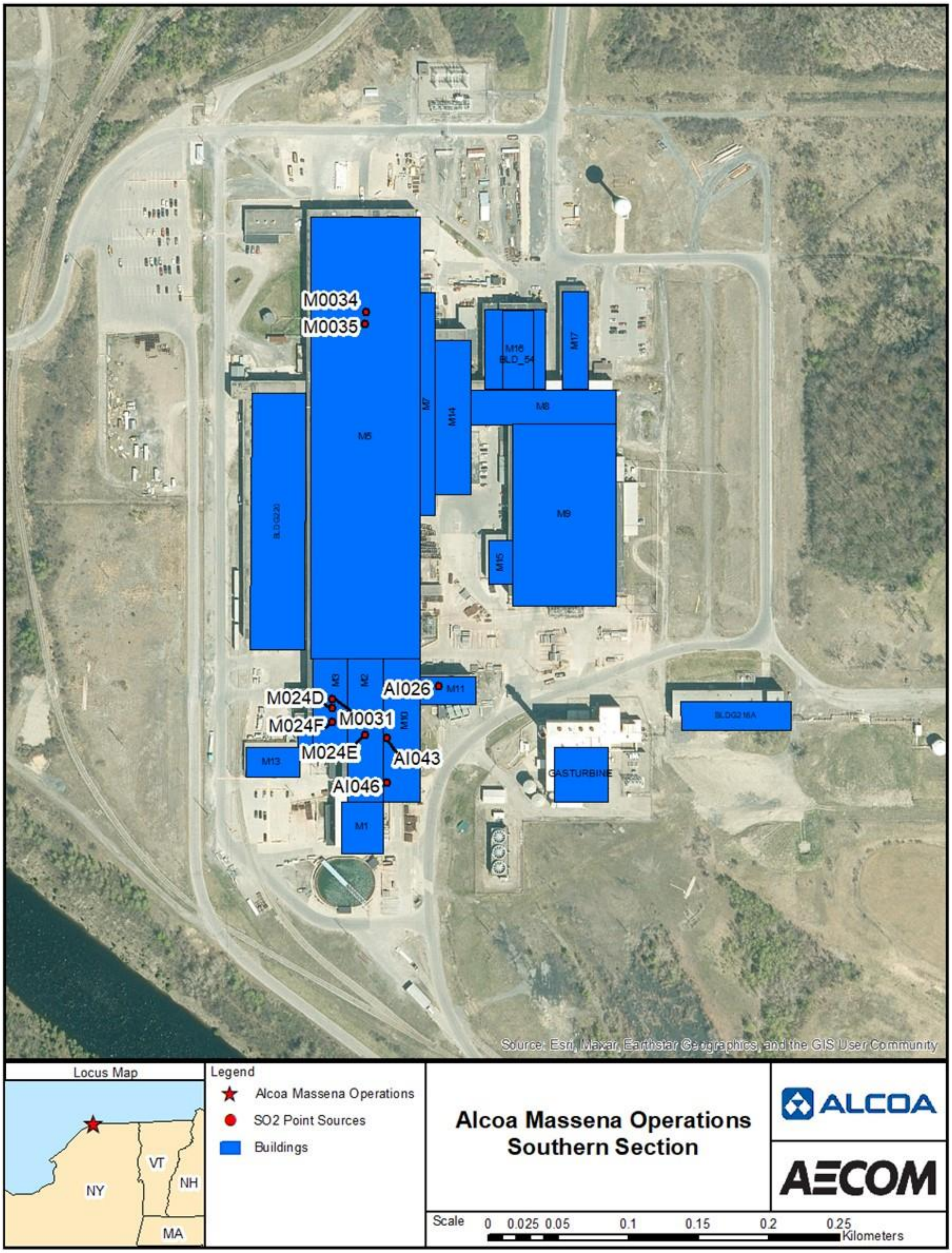


Figure A-2: Facility Layout of Modeled Buildings and SO_2 Sources – Southern Section



Appendix B Formulation and Performance of the AERMOD ENTRAIN Option as Applied to the Alcoa Massena Modeling Application

B.1 Introduction

It is important for purposes of resolving the nonattainment area in the vicinity of the Alcoa Massena facility that a relatively unbiased dispersion model is used because the modeling results are used to determine the compliance status of the future source configuration and emissions. If the model is biased high, then the source may be forced to overcontrol or over-engineer a solution to the nonattainment issue that is unnecessary and overly burdensome. Alcoa has evaluated an additional site-specific modeling approach where a model option called ENTRAIN is used. Use of this option was proposed in the June 2023 version of the modeling protocol and serves as Alcoa's preferred site-specific modeling approach because it reduces, to the extent possible, the overprediction tendency of AERMOD at this facility and has been shown to outperform Massena_MOD.

The following information presented in this appendix provides background on the ENTRAIN model option and presents the model evaluation from the June 2023 modeling protocol. The SO_2 source characterization in this current protocol vs. the June 2023 protocol varied slightly in that the potline roof vent sources were characterized with a different number of stacks and the exit temperatures varied for the potline dry scrubber stacks and potline roof vents and a minor revision was made to the modeled ambient air boundary and receptor grid. For the roof vents, 10 stacks/vent was used to characterize the roof vents for all evaluated site-specific modeling approaches presented in this appendix except "Massena_MOD+ENTRAIN 0.2" used 14 stacks/vent. Stack parameters for the roof vent stacks were calculated where stack diameter was set to the roof vent opening width and exit velocity was calculated to conserve the total roof vent flow rate divided by the number of stacks modeled. For exit temperatures, the dry scrubber stacks utilized an 89.7 K above ambient temperature while the roof vents utilized 25 K above ambient temperature. Lastly, the June 2023 modeling protocol model evaluation used AERMOD version 22112 whereas the current protocol uses AERMOD 23132 (released October 25, 2023). Overall, these changes are not expected to have much effect on the outcome of the model evaluation presented herein.

B.2 Background

ENTRAIN is currently an alpha option developed by the Air and Waste Management Association's (A&WMA) PRIME2 committee to refine downwash effects handled by PRIME algorithms, specifically adjusting the entrainment coefficient used for downwash plume rise. AERMOD uses a 0.6 coefficient by default while the ENTRAIN option sets it to 0.35. In 2021, ENTRAIN was adopted by EPA in AERMOD version 21112 as an alpha option. Since that time, a peer-reviewed journal article[39] was published in the Journal of A&WMA (December 2022) that further supports this option, providing more support to advance this option to beta status, especially for modeling applications with field data to support its use. By including the ENTRAIN option for Massena Operations, results indicate improved model performance, especially at Site 2 where measured SO_2 concentrations are the highest.

The development of plume rise formulas for air quality dispersion models was described in landmark chapters authored by Dr. Gary Briggs[40,41] in 1975 and 1984. Briggs notes that a key aspect of plume rise is the increase of the

[39] Petersen R. L., J. O. Paumier, and S. A. Guerra, 2022. Development, evaluation, and implementation of building downwash and plume rise enhancements in AERMOD. *J. Air Waste Manag. Assoc*.,72:12,1423-1441. https://doi.org/10.1080/10962247.2022.2120563.
[40] Briggs, G. A., 1975. Plume Rise Predictions. Chapter 3 in Haugen, D. A. Lectures on air pollution and environmental impact analyses pp, 59-111, Boston, MA, 29 September 1975-3 October 1975. American Meteorological Society, Boston, MA.
[41] Briggs, G. A., 1984. Plume Rise and Buoyancy Effects. Chapter 8 in Randerson, D. 1984. Atmospheric Science and Power Production. Prepared for United States Department of Energy. DOI:10.2172/6503687.

plume volume flux with height due to entrainment of ambient air into the plume as it rises. The ultimate height of the final plume rise will depend upon the rate of ambient air entrainment, which will act to cool the plume to near ambient temperatures, such that the remaining plume upward velocity is comparable to typical atmospheric vertical turbulent velocities. Plume rise formulations such as those noted above refer to the “entrainment coefficient” or “entrainment constant” as being proportional to the rate of increase of the effective plume radius with height. This constant is not specified by plume rise theory because of the “closure problem” in which there is one more variable than equations available for the plume rise formulation. Therefore, the closure for the plume rise formulation is typically completed by developing an entrainment coefficient (β) based upon a comparison of modeled plume behavior (or ground-level concentrations) to field studies.

The value for β used in AERMOD and many other models for point source plume rise is based upon experiments done in neutrally stable ambient fluids by Hoult and Weil[42] as well as other field studies using tall-stack plumes.[43] However, these plumes were characterized by single, isolated stacks well above buildings with no fugitive heat sources in the vicinity of the plumes that would complicate the plume rise assessment. Groups of closely-spaced plumes, such as those resulting from adjacent stacks along with fugitive heat releases that form a “heat envelope” would result in a larger effective plume radius, and this plume radius would grow more slowly with height than isolated plumes would. This means the plume’s size would not be altered as rapidly by entrainment of cooler, ambient air as isolated plumes would be.

Briggs established a “best fit” plume rise entrainment coefficient as 0.6 based upon isolated tall stack studies, which is the value used by AERMOD. For larger plumes in terms of building wakes affected by building downwash, Petersen et al.[39] noted from wind tunnel observations of plume rise affected by building downwash that an entrainment coefficient of 0.35 resulted in a better match of plume rise with visible plumes in their wind tunnel. However, due to site-specific considerations for plume rise for cases of fugitive heat releases that can affect plume rise that were not considered in past studies, it is best to allow the user to test a range of values for the entrainment coefficient.

AECOM has generalized the ENTRAIN option by adding a capability for the user to specify the coefficient value as an input. The revised AERMOD executable, programming code, and AERMOD equivalency test were provided to the reviewing agencies in the June 30, 2023 modeling archive for the model evaluation demonstration.

B.2.1 Field Data Use for Estimating Optimal Entrainment Coefficient for Massena Operations

For large industrial area sites where the plumes are not easily visible, the best way to determine the optimal entrainment coefficient is to test AERMOD’s ability to match its ground-level predictions to observations. The optimal entrainment coefficient is expected to be lower than the default value of 0.6 used in AERMOD that is applicable to isolated stacks in an environment without fugitive heat releases.

Figure B-1 shows a visible and infrared photo of the potline area, indicating significant heat releases from this area, consistent with the discussion of heat losses of roughly 50 MW from the potline area alone. Thermal satellite imagery in **Figure B-2** indicates other areas of the smelter with SO_2 sources (the bake oven area and the melters / holders area) also have significant heat losses. In these areas, stack releases of SO_2 emissions would have buoyancy enhancements due to the fugitive heat releases, the plumes from adjacent stacks would tend to partially merge, and fugitive heat surrounding the stacks would tend to “insulate” the plumes from intrusion of ambient air in the rising plume. These effects all lead to a lower effective plume rise entrainment coefficient relative to the default value used for isolated stacks without fugitive heat releases of 0.6. The testing of the model versus observed concentrations, especially at Site 2, is a key tool to optimize the value of the entrainment coefficient, factoring in the comparison of concentration magnitudes as well as meteorological conditions for peak observed and modeled concentrations. Model evaluation results that were submitted with the June 30, 2023 modeling protocol indicate that using entrainment coefficients lower than 0.6 (such as 0.35 or even as low as 0.2) reduces the overprediction bias in the

[42] Hoult, D. B. and J. C. Weil, 1972. Turbulent plume in a laminar cross flow, *Atmospheric Environment*, Volume 6, Issue 8, 513-531, ISSN 0004-6981, https://doi.org/10.1016/0004-6981(72)90069-8.
[43] Hoult, D. B, J. A. Fay, and L. J. Forney, 1969. A Theory of Plume Rise Compared with Field Observations. *Journal of the Air Pollution Control Association*, 19:8, 585-590, DOI: 10.1080/00022470.1969.10466526.

site-specific model as well as improves upon the relationship between the top 25 modeled vs. monitored SO_2 concentrations' meteorological conditions.

For purposes of ease of discussion and labeling, the site-specific modeling approaches of AERMOD evaluated for proposed use at the Massena Operations smelter will refer to the 100-m deep neutral layer and enhanced treatment of building downwash as "Massena_MOD", the use of these two enhancements with the addition of the ENTRAIN model option (entrainment coefficient of 0.35 by default) will be referred to as "Massena_MOD+ENTRAIN", and the use of these three enhancements with the ENTRAIN model option using a lower entrainment coefficient of 0.2 will be referred to as "Massena_MOD+ENTRAIN 0.2". Note that model results shown herein for AERMOD Default Massena_MOD were based on a preliminary potline roof vent characterization where they were modeled as 10 stacks per roof vent rather than the final 4 stacks per potline roof vent. Therefore, AERMOD Default and Massena_MOD results shown herein will vary slightly from those results shown in the main text of the modeling protocol.

Figure B-1: One of the Three Potline Dry Scrubber Stack Clusters - Heat Generated by the Potlines

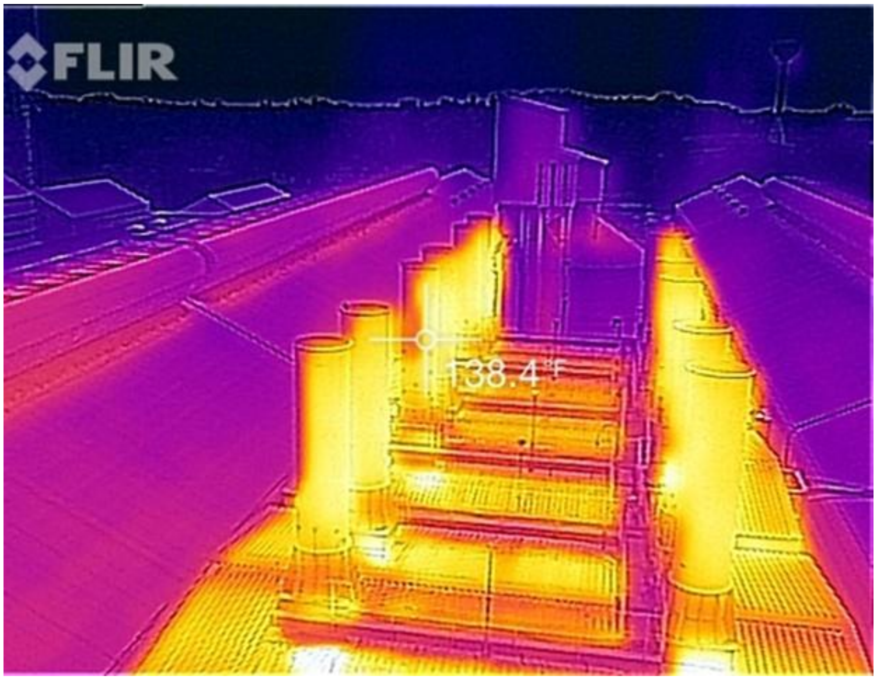
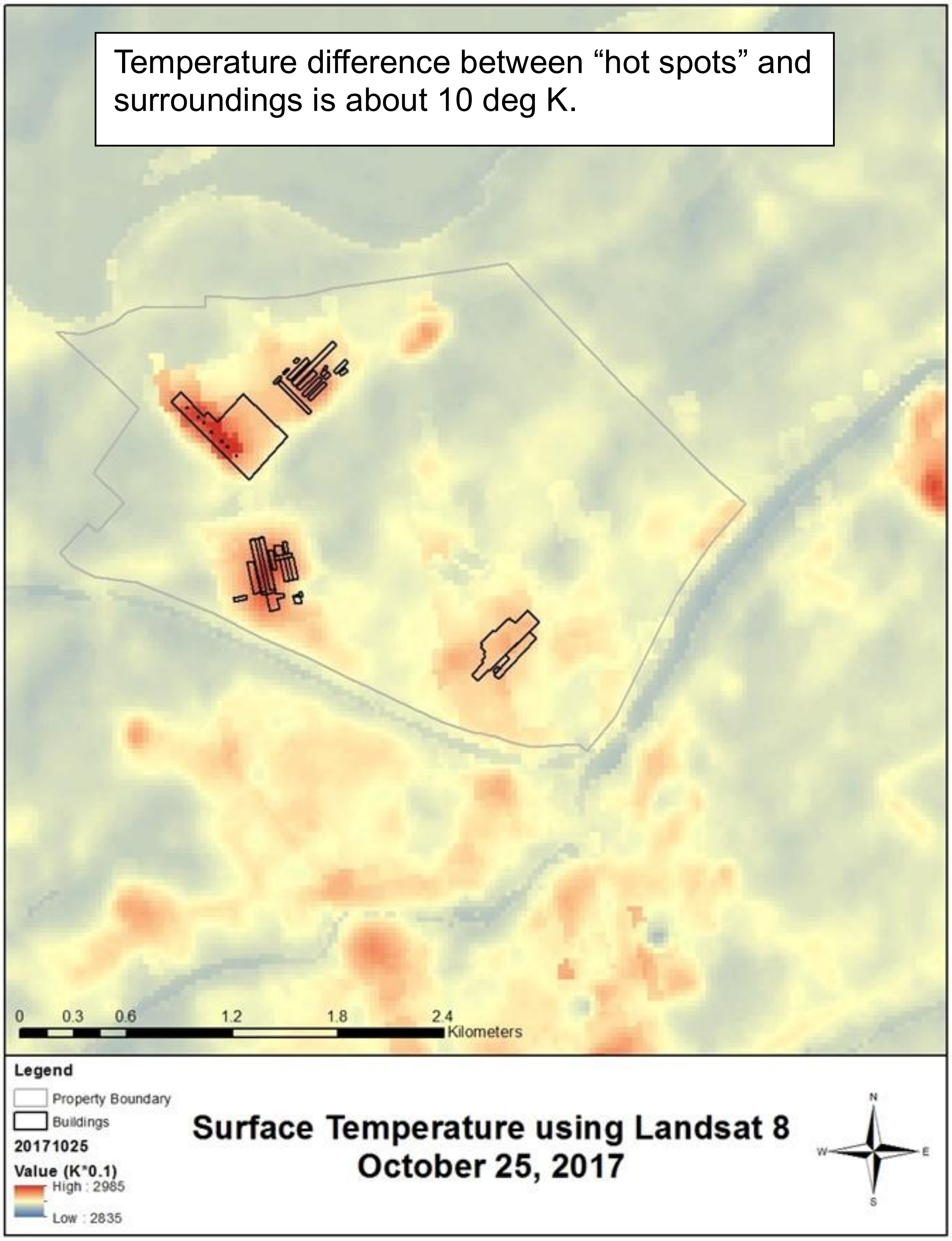

FLIR
138.4°F

Figure B-2: Surface Temperature Data at Massena Operations



B.3 Evaluation of the ENTRAIN Option

A model evaluation has been conducted with the default modeling approach (EPA's original BPIPPRM and the original AERMET-produced PROFILE file) and three site-specific modeling approaches. All three approaches utilize the already approved enhancements included in "Massena_MOD" (ORD BPIPPRM and a neutral lapse rate). However, Massena_MOD is further refined to include the ENTRAIN model option. One uses the ENTRAIN option with its default entrainment coefficient of 0.35 ("Massena_MOD+ENTRAIN") while the other, the preferred model, uses the ENTRAIN option with a user-provided entrainment coefficient of 0.2 ("Massena_MOD+ENTRAIN 0.2").

In Section B.3.1, this evaluation describes an analysis of meteorological conditions associated with the top 25 observed and top 25 predicted concentrations for the default, Massena_MOD, Massena_MOD+ENTRAIN, and Massena_MOD+ENTRAIN 0.2 models. This analysis provides a sense as to whether the better performing model to be determined from the statistical tests described in subsequent section is performing better for the right conditions. Section B.3.2 describes the statistical tests to be used for the performance evaluation. These tests are an important element in the process of the demonstrating applicability of an alternative modeling approach. An important consideration for the selection of the better modeling approach is the performance of the models being considered for the two local monitors (Site 1 and Site 2). Five full years of data (calendar years 2017 – 2021) were used for the model evaluation of the site-specific source characterization for Alcoa Massena. Modeling evaluation files were provided to the reviewing agencies in the June 30, 2023 modeling archive for the model evaluation demonstration.

B.3.1 Analysis for Reviewing Model Performance

A review was conducted of the peak observed and modeled concentrations to determine their occurrence during the day vs. night and meteorological conditions (e.g., wind speed and mixing height) associated with these peak concentrations. This analysis looks at the top 25 1-hour observed and modeled concentrations.

Tables B-1 and B-2 provide a tabulation of the dates and hours of top 25 ranked observed concentrations at Sites 1 and 2, respectively. The tables include the wind speed and mixing height information as well. Similarly, **Tables B-3 and B-4** provide a tabulation of the top 25 modeled concentrations at Sites 1 and 2 for AERMOD-Default. The same information is provided for Massena_MOD (**Table B-5 and B-6**), Massena_MOD+ENTRAIN (**Tables B-7 and B-8**), and Massena_MOD+ENTRAIN 0.2 (**Tables B-9 and B-10**).

For the comparison of daytime vs. night hours, a daytime hour is defined as one with a defined convective mixing height (negative Monin-Obukhov length). The top 25 observed hourly SO_2 concentrations at Site 1 during 2017-2021 had all 25 events occurring at night, and Site 2 had 20 hours occurring at night. In comparison, AERMOD-Default's top 25 modeled hours had all 25 hours at night for both Site 1 and Site 2. While AERMOD-Default had the same number of nighttime hours out of the top 25 concentrations as the observations at Site 1, it had more nighttime hours out of the top 25 concentrations than the observations at the critical Site 2. Massena_MOD's top 25 modeled hours at Site 1 had 17 hours at night and had all 25 hours at night at Site 2. Massena_MOD+ENTRAIN had 13 hours at night for Site 1 and 24 hours at night for Site 2. Massena_MOD+ENTRAIN 0.2 had 13 hours at night for Site 1 and 20 hours at night for Site 2 (matching the number of nighttime hours out of the top 25 observed for Site 2).

Another point of comparison is the distribution of wind speeds for the top 25 observed and modeled hours. Over the top 25 observed hours, the average wind speed was about 4.5 meters per second (m/s) at Site 1 and 4.8 m/s at Site 2. For AERMOD-Default, this average for both sites was significantly lower at 0.8 m/s, unexpectedly low wind speeds that indicate a significant mismatch in the meteorological conditions between the peak observations and peak modeled hours.

The average wind speed for the top 25 modeled hours with Massena_MOD is 1.2 m/s at Site 1 and 2.3 m/s at Site 2. For Massena_MOD+ENTRAIN, the average is 1.9 m/s at Site 1 and 2.7 m/s at Site 2 while Massena_MOD+ENTRAIN 0.2 is 2.3 m/s at Site 1 and 3.0 m/s at Site 2. These averages are more consistent with the observations, indicating that the site-specific modeling approaches perform better than AERMOD-Default at matching the wind speed conditions associated with peak concentration hours. With each refinement to Massena_MOD (ENTRAIN then ENTRAIN 0.2), the average wind speed matches better with the observed average wind speed. **Figures B-3 and B-4** show the top 25 ranked modeled and observed concentrations vs. wind speeds in plot form, showing improvement upon AERMOD-Default with more consistent wind speed distribution in comparison to observations with Massena_MOD, then further improvement with Massena_MOD+ENTRAIN, and finally additional improvement with Massena_MOD+ENTRAIN 0.2.

Yet another point of comparison is the distribution of mixing heights for the top 25 observed and modeled hours. For each hour, the higher of the mechanical and convective mixing heights was selected, with only the mechanical mixing height for stable hours. Over the top 25 observed hours, the average mixing height was 498 m at Site 1 and 438 m at Site 2. The average mixing height over the top 25 modeled hours with AERMOD-Default at both sites was much lower at 49 m at Site 1 and 35 m at Site 2.

The average mixing height for the top 25 modeled hours with Massena_MOD is 88 m at Site 1 and 124 m at Site 2. For Massena_MOD+ENTRAIN, this value is 192 m at Site 1 and 152 m at Site 2 while Massena_MOD+ENTRAIN 0.2 is 240 m at Site 1 and 200 m at Site 2. These averages are more consistent with the observations, indicating that the site-specific modeling approaches perform better than AERMOD-Default at matching the mixing heights associated with peak concentration hours. As with the wind speed comparison, with each refinement to Massena_MOD (ENTRAIN then ENTRAIN 0.2), the average mixing height matches better with the observed average mixing height. **Figures B-5 and B-6** show this data in plot form.

Overall, these results indicate that the best performing model that matches the top 25 hourly concentrations more consistently with observations is Massena_MOD+ENTRAIN 0.2, followed by Massena_MOD+ENTRAIN, then Massena_MOD, and lastly AERMOD-Default.

Table B-1: Top 25 Ranked Observed Concentrations at Site 1

Rank	Date	Hour	Monitored SO_2 Concentration ($\mu g/m^3$)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	3/22/2018	20	202.13	236	3.2	263	--
2	4/8/2017	21	167.08	222	3.4	285	--
3	2/27/2017	18	149.68	245	3.6	325	--
4	3/22/2018	19	136.66	242	3.2	356	--
5	2/14/2019	3	123.19	243	5.0	533	--
6	4/8/2017	20	120.31	238	3.1	302	--
7	2/27/2018	2	119.97	230	3.2	210	--
8	3/6/2019	3	119.71	244	4.1	507	--
9	4/3/2019	23	119.39	249	7.4	1255	--
10	12/19/2017	24	118.97	244	5.5	621	--
11	2/13/2017	2	118.69	241	4.3	418	--
12	2/14/2019	2	117.22	246	5.8	667	--
13	3/14/2021	3	116.77	240	3.6	411	--
14	3/6/2019	1	116.12	234	3.5	295	--
15	2/12/2020	4	115.78	247	5.3	592	--
16	2/13/2017	1	115.44	246	4.8	490	--
17	3/17/2020	24	115.33	239	4.1	392	--
18	3/4/2020	23	113.18	237	3.5	308	--
19	2/16/2021	20	112.63	241	5.7	648	--
20	3/14/2021	1	111.85	240	3.5	389	--
21	2/6/2021	19	110.93	248	5.2	564	--
22	2/11/2020	24	110.59	241	4.8	501	--
23	3/4/2019	23	110.20	247	4.6	608	--
24	1/25/2019	19	109.18	247	6.3	764	--
25	2/16/2019	6	108.99	249	6.2	735	--

Table B-2: Top 25 Ranked Observed Concentrations at Site 2

Rank	Date	Hour	Monitored SO_2 Concentration ($\mu g/m^3$)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	5/23/2020	20	301.17	48	5.5	506	--
2	2/1/2021	4	287.91	31	2.8	145	--
3	2/1/2021	5	272.24	35	3.1	169	--
4	2/1/2021	2	272.09	35	4.2	269	--
5	2/1/2021	3	269.60	25	3.0	215	--
6	2/1/2021	19	267.37	58	4.7	322	--
7	4/4/2017	2	265.43	58	6.8	691	--
8	4/24/2020	7	257.10	41	4.4	402	203
9	5/25/2017	8	256.71	50	5.7	613	674
10	4/8/2021	18	256.08	44	6.9	747	704
11	2/1/2021	6	255.03	30	2.5	117	--
12	5/18/2018	21	252.99	50	5.0	433	--
13	1/24/2020	6	250.58	46	4.2	269	--
14	1/31/2021	23	250.31	39	2.8	144	--
15	4/16/2018	21	249.76	53	6.0	598	--
16	4/8/2021	19	247.56	49	6.2	615	--
17	5/18/2018	20	246.59	48	5.3	474	--
18	5/25/2017	9	243.90	48	6.3	707	948
19	2/7/2018	14	243.63	49	5.8	484	113
20	2/2/2021	2	243.63	44	4.2	270	--
21	5/25/2017	21	240.62	52	7.2	757	--
22	3/26/2017	5	240.23	68	2.9	221	--
23	4/24/2020	6	240.23	51	3.0	197	--
24	5/12/2017	19	237.14	45	5.2	467	--
25	3/18/2017	18	236.17	50	7.5	826	--

Table B-3: Top 25 Ranked Modeled Concentrations for AERMOD Default Model at Site 1

Rank	Date	Hour	Modeled SO_2 Concentration (μg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	2/2/2020	22	939.32	252	0.7	40	--
2	2/20/2017	19	932.33	251	0.8	42	--
3	4/10/2021	23	901.59	253	0.8	49	--
4	12/3/2019	7	900.02	252	0.8	42	--
5	2/24/2020	20	878.02	254	1.1	52	--
6	12/19/2020	1	870.52	253	0.7	42	--
7	4/14/2017	4	857.24	252	0.7	45	--
8	1/13/2021	1	846.07	254	1.0	50	--
9	4/18/2020	5	844.01	254	0.9	52	--
10	4/23/2017	4	829.58	251	0.8	48	--
11	8/21/2020	20	829.16	253	1.0	61	--
12	4/23/2017	3	827.15	253	1.0	56	--
13	5/7/2018	20	826.07	253	0.9	54	--
14	4/1/2020	23	818.08	255	0.8	47	--
15	8/7/2020	22	814.41	250	0.7	47	--
16	4/22/2017	24	812.92	251	0.9	52	--
17	7/9/2021	21	803.49	252	0.8	50	--
18	9/16/2018	24	799.84	250	0.7	45	--
19	3/18/2018	4	797.64	253	0.8	52	--
20	7/9/2021	20	786.13	253	0.8	56	--
21	3/30/2018	24	779.69	254	1.0	56	--
22	6/6/2018	21	776.27	252	0.7	45	--
23	6/10/2018	22	774.53	252	0.7	46	--
24	7/15/2018	21	769.07	250	0.8	48	--
25	11/4/2017	7	768.78	254	0.8	51	--

Table B-4: Top 25 Ranked Modeled Concentrations for AERMOD Default Model at Site 2

Rank	Date	Hour	Modeled SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	2/26/2021	22	1452.50	54	1.0	36	--
2	9/13/2018	4	1451.24	55	0.7	35	--
3	5/23/2017	24	1407.78	54	0.8	36	--
4	4/14/2021	22	1397.41	53	0.7	35	--
5	10/24/2017	17	1396.64	55	0.7	35	--
6	7/15/2018	20	1388.05	52	0.6	33	--
7	7/10/2020	4	1376.89	52	0.7	34	--
8	3/26/2020	7	1376.84	55	0.9	35	--
9	8/27/2020	6	1374.37	51	0.9	39	--
10	5/28/2021	24	1370.13	55	0.8	36	--
11	2/15/2021	8	1369.53	56	0.8	31	--
12	2/8/2021	22	1369.25	56	0.8	33	--
13	4/20/2020	23	1362.40	53	0.8	38	--
14	4/17/2020	24	1362.33	53	0.8	38	--
15	8/16/2020	6	1357.52	54	1.0	41	--
16	8/25/2021	6	1351.39	54	0.7	35	--
17	10/7/2017	1	1351.24	55	0.7	34	--
18	2/10/2018	20	1349.97	53	0.9	35	--
19	8/10/2021	2	1347.05	54	0.8	37	--
20	5/13/2018	5	1345.42	55	0.8	36	--
21	8/14/2020	24	1341.73	51	0.6	34	--
22	5/9/2018	21	1335.85	52	0.6	33	--
23	2/22/2017	5	1331.66	55	1.0	36	--
24	2/21/2017	2	1331.12	51	0.9	34	--
25	2/14/2018	1	1328.94	55	0.9	36	--

Table B-5: Top 25 Ranked Modeled Concentrations for Massena_MOD Model at Site 1

Rank	Date	Hour	Modeled SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	3/22/2021	8	540.05	258	0.7	65	56
2	6/24/2021	6	417.86	244	0.7	60	72
3	2/7/2021	12	411.21	248	1.2	81	91
4	9/21/2020	8	369.89	250	0.7	74	82
5	9/3/2018	22	347.33	254	1.2	73	--
6	2/1/2020	3	333.21	254	1.4	69	--
7	3/23/2021	8	330.85	263	0.7	62	95
8	12/26/2017	20	322.14	255	1.7	99	--
9	2/21/2021	5	320.47	251	1.8	106	--
10	1/21/2017	20	319.80	256	1.4	73	--
11	3/21/2021	8	312.20	253	1.1	108	61
12	12/23/2021	16	297.69	253	1.8	105	--
13	9/22/2020	8	289.13	253	1.1	121	109
14	7/11/2018	22	282.69	254	1.1	69	--
15	7/30/2020	21	281.42	254	1.2	75	--
16	1/16/2021	22	279.20	253	2.1	132	--
17	12/27/2020	15	277.98	253	1.4	85	--
18	1/15/2020	1	277.98	255	2.0	127	--
19	4/7/2019	10	277.59	257	0.9	95	106
20	5/2/2020	1	274.85	252	1.3	77	--
21	5/19/2021	22	271.25	255	1.2	72	--
22	7/4/2020	24	269.90	254	1.1	73	--
23	5/24/2018	1	269.59	255	1.4	88	--
24	5/2/2020	4	267.06	253	1.1	68	--
25	8/21/2020	20	263.10	253	1.0	61	--

Table B-6: Top 25 Ranked Modeled Concentrations for Massena_MOD Model at Site 2

Rank	Date	Hour	Modeled SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	7/15/2020	23	670.26	54	2.1	113	--
2	9/16/2018	20	669.59	56	1.6	76	--
3	8/14/2020	19	667.15	53	2.3	204	--
4	2/18/2021	18	664.50	53	2.9	152	--
5	4/10/2021	19	652.94	52	2.1	112	--
6	7/15/2020	22	648.85	51	2.2	128	--
7	7/15/2020	20	646.65	51	2.3	149	--
8	8/13/2018	22	645.90	53	1.9	102	--
9	9/10/2018	3	644.58	53	2.4	139	--
10	2/19/2021	8	642.66	52	3.1	164	--
11	2/15/2017	8	637.55	53	2.6	127	--
12	2/15/2017	6	636.54	53	2.7	131	--
13	5/21/2017	2	636.02	56	1.9	95	--
14	2/20/2018	5	635.08	56	2.4	107	--
15	2/20/2018	19	634.37	56	2.5	120	--
16	2/5/2021	4	633.36	57	2.3	101	--
17	8/4/2020	5	630.65	51	1.8	87	--
18	2/15/2017	4	629.72	52	2.4	108	--
19	6/30/2017	2	626.12	54	2.1	117	--
20	6/23/2018	23	625.79	53	2.0	105	--
21	2/28/2018	20	623.84	53	2.3	105	--
22	5/17/2018	22	620.03	52	2.4	147	--
23	3/23/2020	21	619.90	54	2.6	151	--
24	3/19/2020	20	616.87	51	2.2	116	--
25	5/3/2021	2	615.98	52	2.4	132	--

Table B-7: Top 25 Ranked Modeled Concentrations for Massena_MOD+ENTRAIN Model at Site 1

Rank	Date	Hour	Modeled SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	3/22/2021	8	540.05	258	0.7	65	56
2	6/24/2021	6	417.86	244	0.7	60	72
3	2/7/2021	12	411.21	248	1.2	81	91
4	9/21/2020	8	369.89	250	0.7	74	82
5	3/23/2021	8	330.85	263	0.7	62	95
6	3/21/2021	8	312.20	253	1.1	108	61
7	9/22/2020	8	289.13	253	1.1	121	109
8	4/7/2019	10	277.59	257	0.9	95	106
9	2/21/2021	5	270.27	251	1.8	106	--
10	8/17/2020	8	257.06	247	1.1	110	129
11	8/11/2020	22	255.88	252	3.4	402	--
12	2/1/2017	12	252.21	256	0.8	70	136
13	2/7/2021	18	246.82	251	3.8	347	--
14	7/12/2020	5	245.75	253	3.1	365	--
15	5/12/2021	6	244.25	253	2.7	266	--
16	2/20/2019	10	243.54	243	0.7	54	111
17	1/16/2021	22	243.32	253	2.1	132	--
18	5/17/2018	6	241.78	254	2.8	291	--
19	12/23/2021	16	238.78	253	1.8	105	--
20	12/26/2017	20	238.55	255	1.7	99	--
21	3/9/2021	8	238.42	250	2.6	269	--
22	12/19/2020	13	237.51	245	1.4	115	101
23	8/11/2020	21	233.94	255	2.9	313	--
24	2/2/2017	17	233.87	252	4.4	434	--
25	2/7/2021	19	232.20	249	4.5	448	--

Table B-8: Top 25 Ranked Modeled Concentrations for Massena_MOD+ENTRAIN Model at Site 2

Rank	Date	Hour	Modeled SO_2 Concentration (µg/m^3)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	2/18/2021	18	534.26	53	2.9	152	--
2	2/20/2018	19	533.17	56	2.5	120	--
3	3/10/2021	9	530.20	53	0.8	57	93
4	2/14/2017	24	528.17	56	3.3	181	--
5	2/1/2021	21	527.44	57	3.0	173	--
6	2/19/2021	8	525.05	52	3.1	164	--
7	2/19/2021	6	524.21	53	3.4	195	--
8	2/13/2021	23	518.95	52	3.3	188	--
9	8/27/2021	21	518.80	56	2.5	153	--
10	12/30/2019	23	518.62	56	2.7	131	--
11	2/10/2018	1	517.41	56	2.9	147	--
12	5/19/2018	5	514.54	56	2.6	156	--
13	8/27/2021	20	513.76	56	2.6	162	--
14	2/14/2017	20	508.93	57	2.8	145	--
15	2/20/2018	5	506.31	56	2.4	107	--
16	2/19/2021	3	497.91	55	3.6	215	--
17	2/18/2020	5	495.35	57	2.8	143	--
18	3/19/2020	24	493.49	57	2.4	134	--
19	7/23/2018	3	490.26	51	2.6	165	--
20	9/10/2018	3	489.88	53	2.4	139	--
21	1/27/2018	1	489.35	56	3.1	165	--
22	4/19/2017	2	489.15	54	2.7	161	--
23	2/22/2017	17	489.12	55	3.1	173	--
24	2/15/2017	6	488.98	53	2.7	131	--
25	2/5/2021	4	487.73	57	2.3	101	--

Table B-9: Top 25 Ranked Modeled Concentrations for Massena_MOD+ENTRAIN 0.2 Model at Site 1

Rank	Date	Hour	Modeled SO_2 Concentration ($\mu g/m^3$)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	3/22/2021	8	542.34	258	0.7	65	56
2	6/24/2021	6	420.26	244	0.7	60	72
3	2/7/2021	12	412.87	248	1.2	81	91
4	9/21/2020	8	371.15	250	0.7	74	82
5	3/23/2021	8	331.88	263	0.7	62	95
6	3/21/2021	8	313.53	253	1.1	108	61
7	9/22/2020	8	289.99	253	1.1	121	109
8	4/7/2019	10	278.71	257	0.9	95	106
9	8/17/2020	8	258.18	247	1.1	110	129
10	8/11/2020	22	255.22	252	3.4	402	--
11	2/1/2017	12	252.95	256	0.8	70	136
12	2/7/2021	18	245.96	251	3.8	347	--
13	2/20/2019	10	244.74	243	0.7	54	111
14	7/12/2020	5	244.44	253	3.1	365	--
15	5/12/2021	6	241.61	253	2.7	266	--
16	5/17/2018	6	240.80	254	2.8	291	--
17	12/19/2020	13	238.21	245	1.4	115	101
18	3/9/2021	8	237.88	250	2.6	269	--
19	2/2/2017	17	233.69	252	4.4	434	--
20	2/7/2021	19	232.33	249	4.5	448	--
21	8/11/2020	21	231.82	255	2.9	313	--
22	4/23/2021	6	229.55	251	3.9	471	--
23	2/17/2021	5	229.28	249	4.1	388	--
24	2/17/2017	3	229.07	253	4.0	374	--
25	2/4/2017	2	228.19	250	4.2	409	--

Table B-10: Top 25 Ranked Modeled Concentrations for Massena_MOD+ENTRAIN 0.2 Model at Site 2

Rank	Date	Hour	Modeled SO_2 Concentration ($\mu g/m^3$)	Wind Direction (degrees)	Wind Speed (m/s)	Mechanical Mixing Height (m)	Convective Mixing Height (m)
1	3/10/2021	9	532.59	53	0.8	57	93
2	1/31/2020	13	456.14	48	0.8	37	76
3	11/25/2021	12	443.39	64	1.0	73	80
4	2/14/2017	24	411.50	56	3.3	181	--
5	12/11/2020	11	406.82	60	1.2	78	88
6	3/27/2021	23	377.85	56	3.3	225	--
7	8/26/2019	7	376.05	51	0.8	49	102
8	2/12/2017	4	375.97	57	3.4	195	--
9	2/1/2021	21	373.57	57	3.0	173	--
10	2/20/2018	6	372.23	57	3.4	196	--
11	2/7/2017	2	371.56	57	3.9	241	--
12	2/19/2021	3	367.14	55	3.6	215	--
13	2/20/2018	8	366.20	56	4.4	287	--
14	2/16/2021	1	366.01	53	4.5	296	--
15	7/23/2018	1	365.19	56	3.4	251	--
16	2/14/2021	1	363.15	54	4.5	296	--
17	2/19/2021	6	358.79	53	3.4	195	--
18	10/11/2017	21	357.98	56	2.9	189	--
19	4/27/2017	5	357.04	56	3.6	261	--
20	1/21/2018	18	356.25	57	3.4	189	--
21	3/29/2018	22	354.87	56	3.5	246	--
22	12/14/2019	5	354.28	57	3.9	243	--
23	3/27/2017	3	352.60	56	3.4	231	--
24	10/12/2017	3	352.28	56	2.9	185	--
25	3/2/2020	5	352.16	56	3.6	259	--

Figure B-3: Top 25 Ranked Observed and Modeled Concentrations vs. Wind Speeds for Site 1

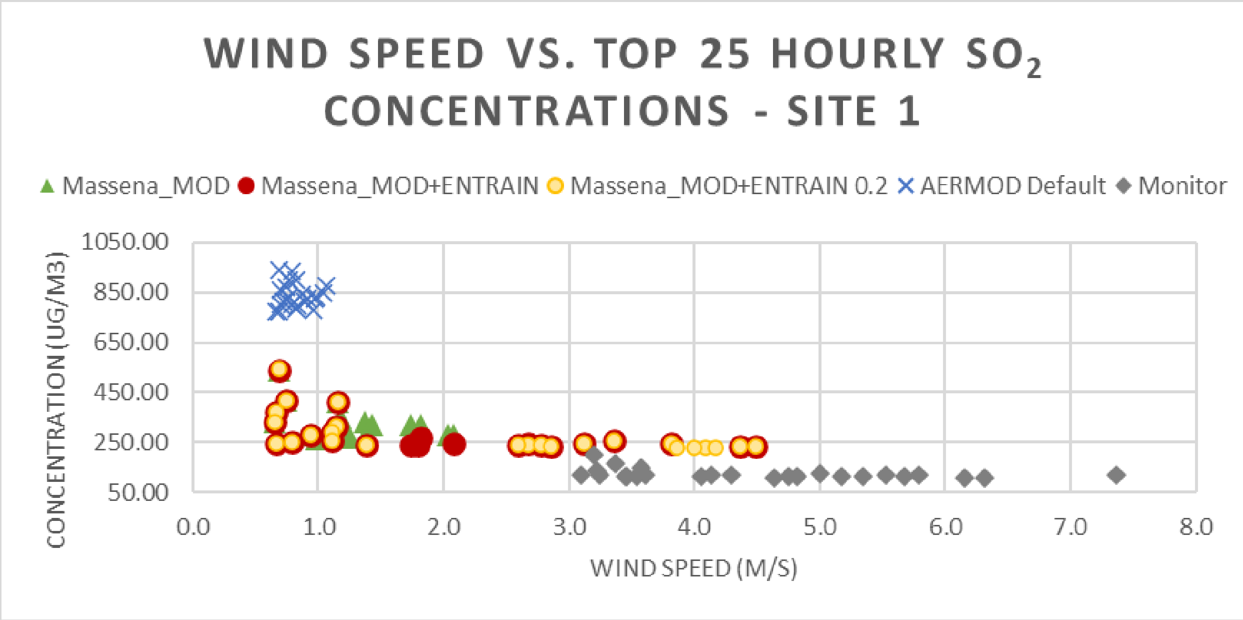


Figure B-4: Top 25 Ranked Observed and Modeled Concentrations vs. Wind Speeds for Site 2

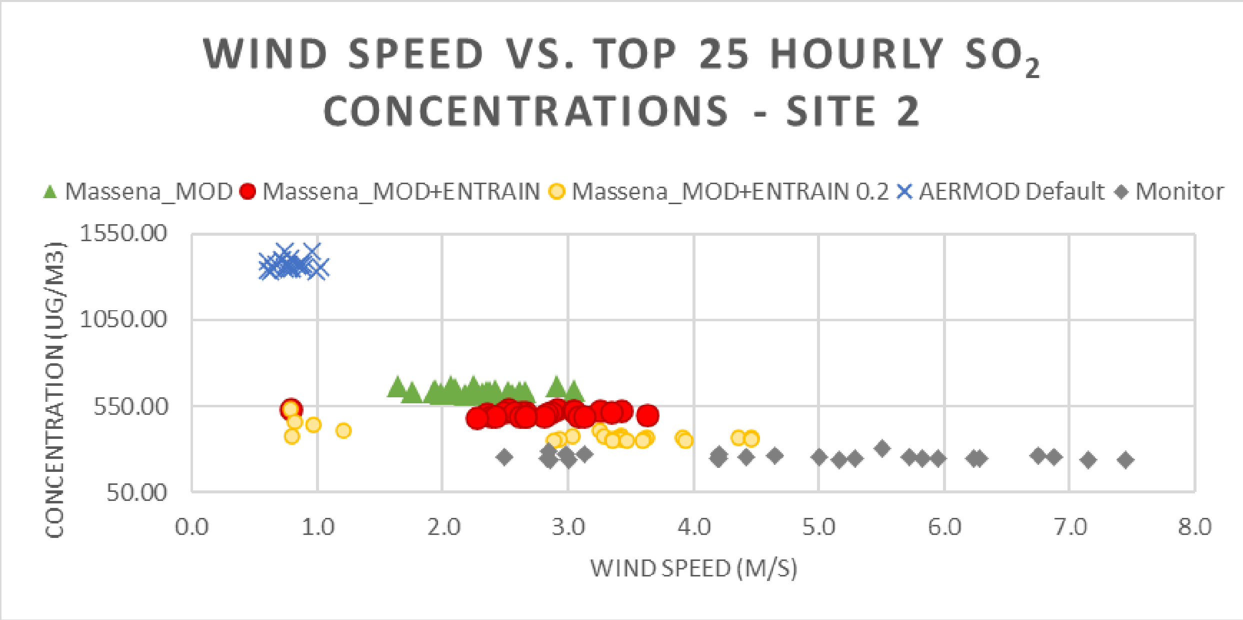


Figure B-5: Top 25 Ranked Observed and Modeled Concentrations vs. Mixing Heights for Site 1

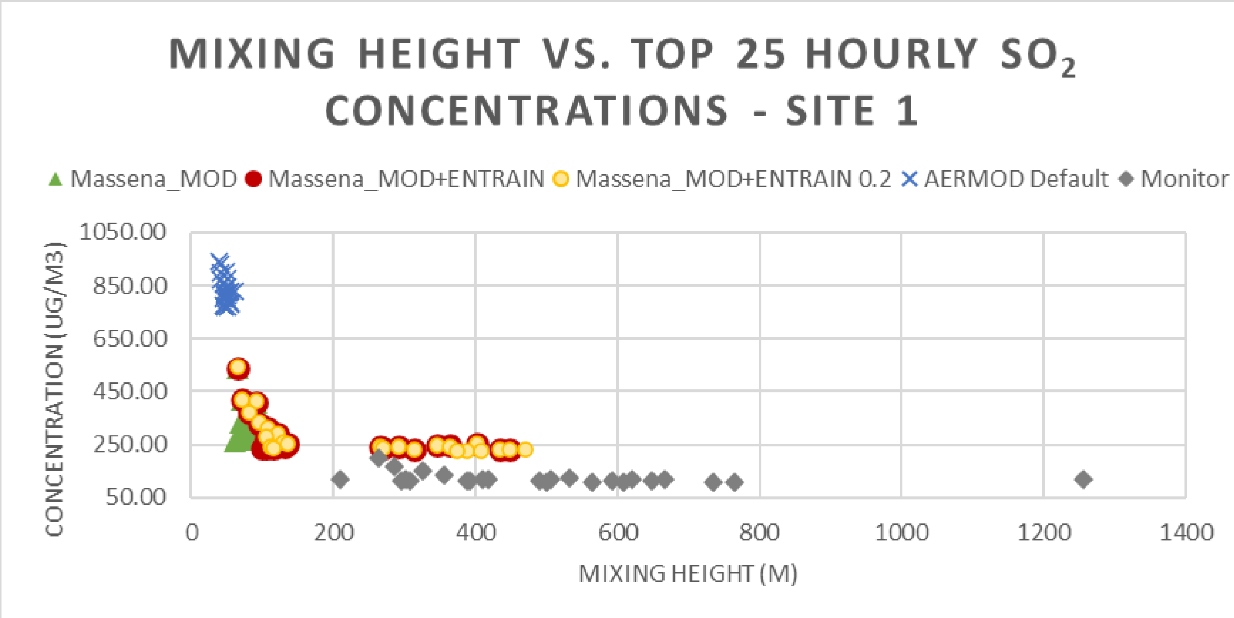
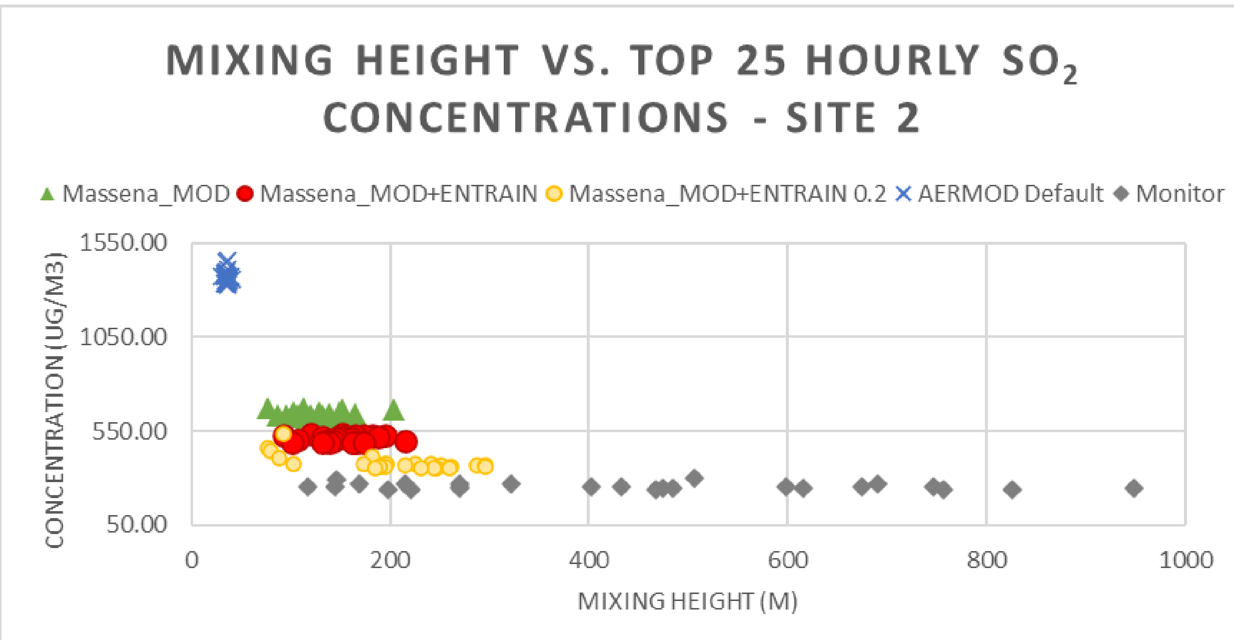


Figure B-6: Top 25 Ranked Observed and Modeled Concentrations vs. Mixing Heights for Site 2



B.3.2 Results of Statistical Performance Tests

Three sets of statistical evaluation tests were conducted: a) quantile-quantile (Q-Q) plots for each monitor, b) comparison of the modeled and observed 5-year average 1-hour average design concentration[44] for each monitor, and c) the use of the Robust Highest Concentration (RHC) as part of EPA's Cox-Tikvart[45] procedure as described in EPA's 1992 model evaluation procedures[6], as implemented in EPA's Model Evaluation Methodology (MEM) software.[46] Further discussion of each of these tests and the results of the testing are provided below.

Quantile-Quantile Plots

Operational performance of models for predicting compliance with air quality regulations, especially those involving a peak or near peak value at some unspecified time and location, can be assessed with Q-Q plots (Chambers et al., 1983).[47] Q-Q plots are created by sorting by rank the predicted and the observed concentrations from a set of predictions initially paired in time. The sorting is generally done for individual monitors, such that the comparison is still paired in space, but not in time. The sorted sets of predicted concentrations are then plotted by rank against the observed concentrations, which are also sorted by rank. While these concentration sets are no longer paired in time, the plot is useful for answering the question, "Over a period of time, does the distribution of the model predictions match those of observations?" Scatterplots, which use data paired in time and space, provide a stricter test, answering the question: "At a given time and place, does the magnitude of the model prediction match the observation?" It is the experience of model developers (e.g., Weil, et al., 1992[48] and Liu and Moore, 1984[49]) that wind direction uncertainties can and do cause disappointing scatterplot results from what are otherwise well-performing dispersion models. Therefore, the Q-Q plot instead of the scatterplot is a more pragmatic procedure for demonstrating model performance of applied models. Venkatram (2001)[50] further discusses the attributes for the use of Q-Q plots for evaluating regulatory models.

The Q-Q plot for Site 1, which compares the default model with the three site-specific approaches, is presented in **Figure B-7** and the Q-Q plot for Site 2 is presented in **Figure B-8**. In general, since SO_2 monitors have a +/- 10% tolerance for calibration accuracy,[51] ranked modeled concentrations that are within 10% of a "perfect model" result are considered to be unbiased. EPA's 1992 model evaluation procedures[52] indicate that an acceptable model should have peak predictions within a factor of 2 of observations. For both sites, it is clear that the default model grossly overpredicts by at least a factor of 5. While the Massena_MOD model at Site 1 is generally showing ranked modeled-to-observed pairs with overpredictions between a factor of 2 and a factor of 3, the peak elevated concentrations are within a factor of 2 of the observations. This is also the case for Massena_MOD+ENTRAIN and Massena_MOD+ENTRAIN 0.2 as this monitor (with no observed NAAQS violations) does not appear to be sensitive to the ENTRAIN enhancements. At Site 2, the Massena_MOD elevated concentrations are between a factor of 2 to 3 of the observations with the peak elevated concentrations within a factor of 2.2 of the observations. Massena_MOD+ENTRAIN and Massena_MOD+ENTRAIN 0.2 both reduce the overprediction tendency further to about 1.8 for the peak elevated concentration. For Massena_MOD+ENTRAIN 0.2, the overprediction tendency drops

[44] The "design concentration" is the 99th percentile of the peak daily 1-hour maximum concentration computed for each calendar year, and averaged over the five years included in the evaluation. For any given year, assuming that there are at least 301 days with valid peak daily observations, the fourth highest daily 1-hour maximum would constitute the 99th percentile value.
[45] William M. Cox and Joseph A. Tikvart, 1990. A statistical procedure for determining the best performing air quality simulation model, *Atmospheric Environment*. Part A. General Topics, Volume 24, Issue 9, Pages 2387-2395. ISSN 0960-1686, https://doi.org/10.1016/0960-1686(90)90331-G.
[46] Strimaitis, D., E. Insley, M. Korc, and F. Lurmann, 1993. User's Guide for the Model Evaluation Methodology (MEM System for Comparing Model Performance Version 1.0. STI-93261-1392-FR.
[47] Chambers, J. M., Cleveland, W. S., Kleiner, B., and Tukey, P. A., 1983. Chapter 3: Comparing Data Distributions. *Graphical Methods for Data Analysis. (Bell Laboratories).* Wadsworth International Group and Duxbury Press.
[48] Weil J.C, Sykes and Venkatram A., 1992. Evaluating air-quality models: Review and outlook. *J. Appl. Met*., 31, p 1121-1144.
[49] Liu, M. K., and G. E. Moore, 1984. Diagnostic validation of plume models at a plains site. EPRI Report No. EA-3077, Research Project 1616-9, Electric Power Research Institute, Palo Alto, CA.
[50] Venkatram, A., R. W. Brode, A. J. Cimorelli, J. T. Lee, R. J. Paine, S. G. Perry, W. D. Peters, J. C. Weil, and R. B. Wilson, 2001. A complex terrain dispersion model for regulatory applications. *Atm.Env*., 35, 4211-4221.
[51] Quality Assurance Handbook for Air Pollution Measurement Systems, Volume II, Ambient Air Quality Monitoring Program, 2013, available at http://www.epa.gov/ttnamti1/files/ambient/pm25/qa/QA-Handbook-Vol-II.pdf. (Table 10-3 and Appendix D, page 13).
[52] EPA, 1992. Protocol for Determining the Best Performing Model. Publication No. EPA–454/R–92–025. Office of Air Quality Planning and Standards, Research Triangle Park, NC. (NTIS No. PB 93–226082). https://www.epa.gov/sites/default/files/2020-10/documents/model_eval_protocol.pdf.

to about 1.5 for the 20th highest concentration (a metric more closely aligned with the form of the NAAQS for a 5-year modeling period) while Massena_MOD+ENTRAIN is 2.1. Based on these Q-Q plot results, the performance of Massena_MOD+ENTRAIN 0.2 is clearly better than that of the default model as well as the other site-specific model approaches.

Figure B-7: Q-Q Plot for Site 1

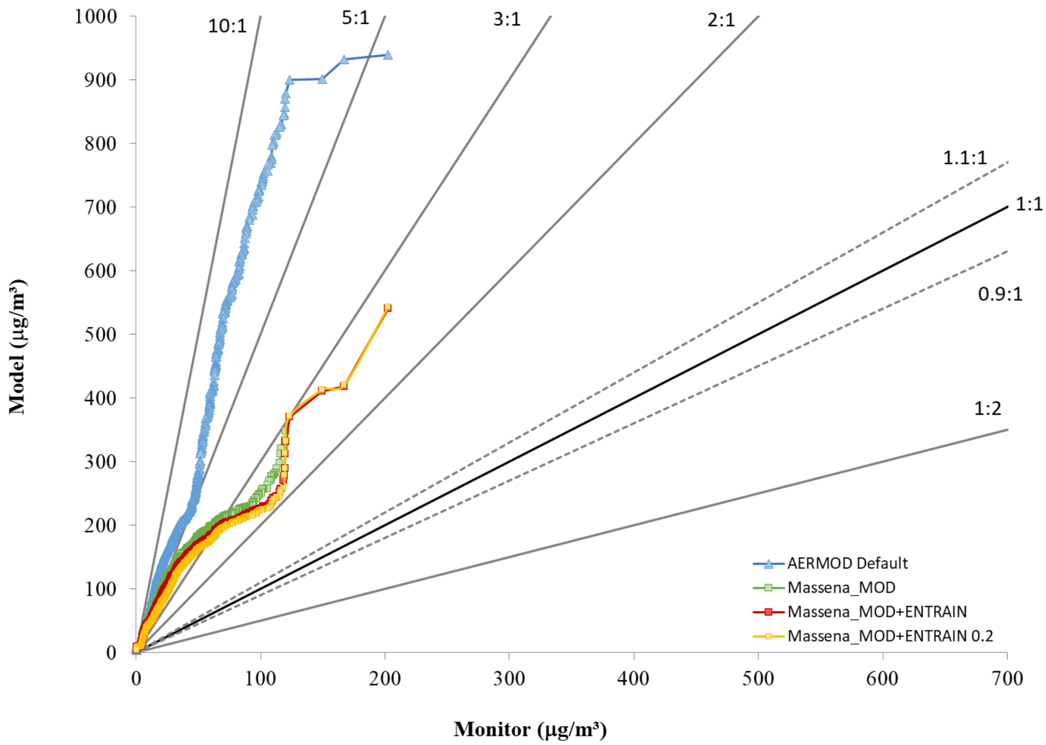
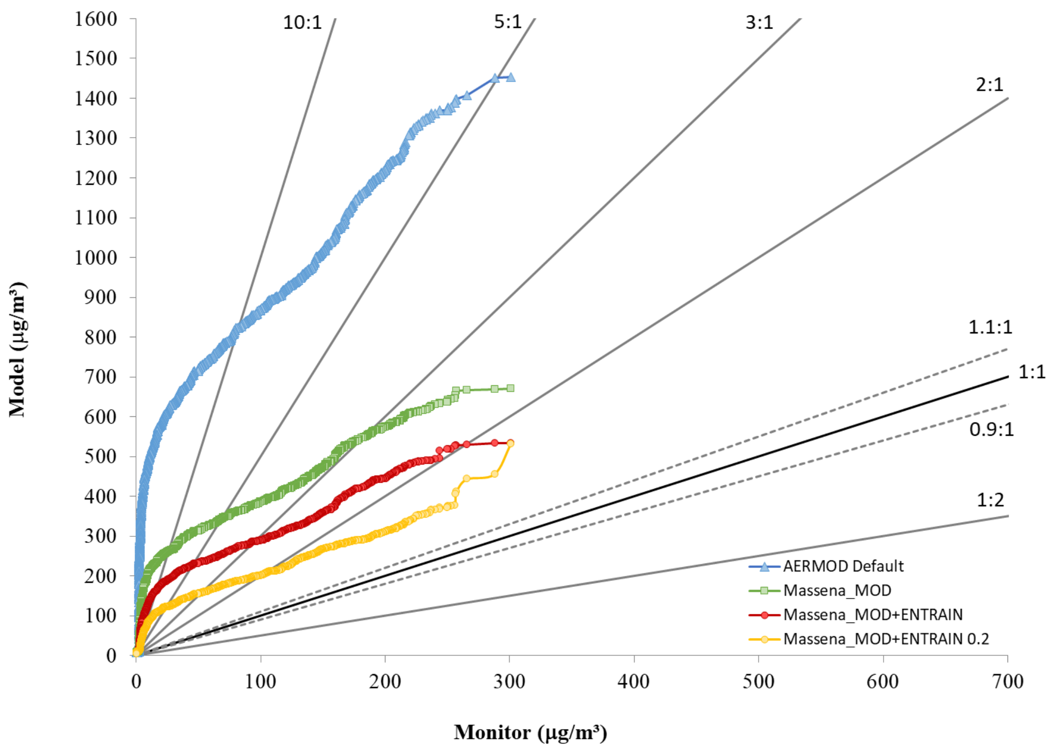


Figure B-8: Q-Q Plot for Site 2



Comparison of Modeled and Predicted Design Concentrations

A key metric that should be included in any model evaluation involves the modeled and observed design concentration, which corresponds to the form of the ambient standard that is being evaluated. For SO_2, this is the 99th percentile peak daily 1-hour maximum concentration averaged over the 5-year modeling period. A separate calculation is made for each monitor. Similar to the discussion in the Q-Q plot metric noted above, a modeled-to-observed ratio ("MOR") design concentration within 10% indicates an unbiased model.

Tabulated values for this comparison are provided in **Table B-11** for Site 1 and **Table B-12** for Site 2. As can be seen in both tables, all models overpredict, but the default model has a MOR value of 7 at Site 1 and over 5 at Site 2. In contrast, the Massena_MOD, Massena_MOD+ENTRAIN, and Massena_MOD+ENTRAIN 0.2 models have MOR values of about 2.3 to 2.4 at Site 1. For Site 2, Massena_MOD and Massena_MOD+ENTRAIN have MOR values between 2 and 3 while the Massena_MOD+ENTRAIN 0.2 model has a MOR value of 1.5. As a result, Massena_MOD+ENTRAIN 0.2 is determined to have a best performance, although still overpredicting, with Massena_MOD+ENTRAIN as the second-best model among the three site-specific modeling approaches.

To understand the spatial pattern of the design concentrations, modeling was conducted with the three site-specific model approaches on a full receptor grid using the 5-year model evaluation dataset. **Figure B-9 to Figure B-11** provide isopleths of modeled design concentrations for Massena_MOD, Massena_MOD+ENTRAIN, and Massena_MOD+ENTRAIN 0.2, respectively. As shown, all site-specific model approaches result in very similar spatial patterns. As expected, the maximum modeled concentration is along the western fenceline very close to Site 2. For Massena_MOD+ENTRAIN 0.2, the maximum modeled concentration is about 500 m north of Site 2 and a concentration closer to Site 2 is also shown for reference. For Massena_MOD+ENTRAIN 0.2, it is apparent that it results in the lowest maximum concentration (closest to matching the monitors) but continues to demonstrate conservatism in that this approach retains overprediction at both monitoring locations by a factor of at least 1.5. At Site 1 in particular, the modeling results indicate design concentrations exceeding the NAAQS; however, the Site 1 monitor's measured design concentrations have always been much less than the NAAQS.

Table B-11: Modeled-to-Observed Design Concentrations at Site 1

Model Approach	4th Highest Design Concentration (ug/m^3)						
	2017	2018	2019	2020	2021	5-year Average	
Observed	118.7	106.6	110.2	110.6	101.1	109.4	**MOR**
AERMOD Default Model	812.9	779.7	642.4	844.0	753.4	766.5	7.0
Massena_MOD Model	252.0	257.0	207.9	281.4	330.8	265.8	2.4
Massena_MOD+ENTRAIN Model	229.8	222.1	207.1	255.9	330.8	249.1	2.3
Massena_MOD+ENTRAIN 0.2 Model	228.2	214.5	206.0	255.2	331.9	247.2	2.3

Table B-12: Modeled-to-Observed Design Concentrations at Site 2

Model Approach	4th Highest Design Concentration (ug/m^3)						
	2017	2018	2019	2020	2021	5-year Average	
Observed	237.1	226.7	214.9	232.6	243.6	231.0	**MOR**
AERMOD Default Model	1331.7	1345.4	1074.0	1362.4	1369.5	1296.6	5.6
Massena_MOD Model	610.7	635.1	566.4	619.9	633.4	613.1	2.7
Massena_MOD+ENTRAIN Model	489.0	490.3	419.6	486.4	525.1	482.1	2.1
Massena_MOD+ENTRAIN 0.2 Model	358.0	354.9	297.8	339.1	373.6	344.7	1.5

Testing with the Robust Highest Concentration

The Model Evaluation Methodology (MEM) software was designed to evaluate model performance by implementing the statistical analysis procedures contained in EPA's 1992 Protocol for Determining the Best Performing Model (EPA-454/R-92-025). MEM evaluates model performance through two stages. The first step is a screening test to flag models that fail to perform at a minimum operation level. The fractional bias (= 2 * (observed – predicted) / (observed + predicted)) of the mean and the fractional bias of the standard deviation are used to qualify performance. The fractional bias has been selected as the basic measure of performance in the MEM. Values for the fractional bias range between -2 and +2 (over prediction, under prediction). Also, the fractional bias is a good proxy for result comparisons. Fractional biases that are equal to -0.67 are equivalent to overpredictions by a factor of 2, while a fractional bias of +0.67 is equivalent to an underprediction by a factor of 2. The absolute fractional bias (AFB) statistic, which is just the absolute value of the fractional bias (FB), is computed for each of the individual models.

The second stage is a resampling technique (bootstrapping) which generates a probability distribution of possible data outcomes. Five years of data can be arranged into seasonal blocks (DFJ, MAM, JJA and SON); the MEM software can be recompiled to accept multiple years of data. Within each season, the pieces are sampled with replacement until a total season is created. This process is repeated using each of the four seasons to construct a complete bootstrap year. Sampling within seasons guarantees that each season will be represented by only days chosen from that season. Since sampling is done with replacement, some days are represented more than once, while other days are not represented at all. Next, the data generated for the bootstrap year are used to calculate the composite performance measures for each model. This process is repeated until sufficient samples are available to calculate a meaningful standard error, which is the standard deviation of the measure over all of the bootstrap-generated outcomes (the sample size of which has been set to 1,000).

The method of bootstrapping is used to estimate the standard error of the composite performance measure of each model. Using this estimation, the statistical significance of the difference between models is then assessed. A test statistic, the Robust Estimate of the Highest Concentration (RHC), is then conducted within MEM using a subset, N, of the highest concentrations (Equation 1).

$$RHC = X(N) + [\bar{X} - X(N)] * \ln\left[\frac{3N-1}{2}\right] \qquad \text{(Equation 1)}$$

where:

N = number of values;

$\bar{X}$ = average of the N-1 largest values; and

$X(N)$ = Nth largest value.

The assigned number of values, N, typically ranges between 11 and 26; 26 is suggested for this application.

After the RHC calculations, the model comparison statistics are then conducted. The first comparison measure that is calculated is the Composite Performance Measure (CPM). The CPM is a weighted linear combination of the individual fractional biased components. A CPM is calculated for each model (Equation 2).

$$CPM = \frac{1}{3} * \overline{AFB_{r,s}} + \frac{2}{3}\left[\frac{(AFB)_3 + (AFB)_{24}}{2}\right] \qquad \text{(Equation 2)}$$

where:

$(AFB)_{r,s}$ = Absolute Fractional Bias for diagnostic conditions r at station s;

$(AFB)_3$ = Absolute Fractional Bias for 3-hour averages; and

$(AFB)_{24}$ = Absolute Fractional Bias for 24-hour averages.

The final performance measure calculated is the Model Comparison Measure (MCM) – mean and confidence interval. The MCM is the difference between the CPM for two models (Equation 3).

$$MCM_{i,j} = CPM_i - CPM_j \qquad \text{(Equation 3)}$$

where:

CPM_i = Composite Performance Measure for model i; and

CPM_j = Composite Performance Measure for model j.

The magnitude and sign of the MCM are indicative of the relative performance of each pair of models. The smaller the composite performance measure, the better the overall performance of a model. This means that for two arbitrary models, Model A and Model B, a negative MCM (i.e., the difference between the CPM for Model A and Model B) implies that Model A is performing better (Model A has the smaller CPM), while a positive value indicates that Model B is performing better. For each pair of model comparisons, the significance of the MCM depends upon whether or not its confidence interval, e.g., 90th percentile, overlaps zero. If the confidence interval overlaps zero, the two models may not be performing at a level which is statistically different, although one model may still exhibit a notable tendency to have a lower bias, and therefore can be judged to have superior performance. If the confidence interval does not overlap zero, (upper and lower limits are both negative or both positive), then there exists a statistically significant difference between the two models at the stated level of confidence. In previous work, EPA has used a 90th percentile level of confidence.53 The MEM software also computes a Combined Model Comparison Measure (CMCM) to provide a model performance assessment over all monitoring sites.

Table B-13 shows the average fractional biases for the modeling approaches at both monitors for the RHC estimate, where a value of zero would indicate a perfect model. A negative fractional bias indicates that the model is overpredicting. A model overprediction by a factor of 2 results in a fractional bias value of -0.67, while an underprediction by a factor of 2 results in a fractional bias of +0.67. The resulting overprediction ratio for the default model is about 7 for Site 1 and about 5 for Site 2. The overprediction ratios are much lower for the site-specific models, where the Massena_MOD+ENTRAIN 0.2 has the lowest (best) ratios for both sites, followed by Massena_MOD+ENTRAIN, and lastly Massena_MOD. For Site 2 specifically, the absolute value of the fractional biases for AERMOD default as well as Massena_MOD are greater than 0.67 while Massena_MOD+ENTRAIN and Massena_MOD+ENTRAIN 0.2 have absolute values less than 0.67. According to EPA's "Protocol for Determining the Best Performing Model", if the absolute value of the fractional bias exceeds 0.67 (i.e., the model overpredicts or underpredicts by more than a factor of 2), "consideration may be given to excluding that model from further evaluation due to its limited credibility for refined regulatory analysis."[30] In this evaluation, we have retained the AERMOD default model and Massena_MOD for the full suite of evaluation results in order to determine how much better the more refined candidate models using the ENTRAIN option perform. However, it is clear from the EPA guidance that the EPA's default model and Massena_MOD can be disqualified from consideration for the NAA model due to their poor evaluation performance based on the fractional bias results.

Table B-13: Average Fractional Biases for Site 1 and Site 2 for RHC Estimate

Site	Model Approach	FB_{Avg}
Site 1	AERMOD Default Model	-1.48
	Massena_MOD Model	-0.88
	Massena_MOD+ENTRAIN Model	-0.79
	Massena_MOD+ENTRAIN 0.2 Model	-0.78
Site 2	AERMOD Default Model	-1.37
	Massena_MOD Model	-0.86
	Massena_MOD+ENTRAIN Model	-0.66
	Massena_MOD+ENTRAIN 0.2 Model	-0.39

Tables B-14 and B-15 show the RHC for Site 1 and Site 2, respectively. A ratio of model predicted RHC to the observed RHC at or slightly above 1.0 is considered ideal. As seen in tables, the default model runs have RHC predicted-to-observed ratios between 5 and 7, while the site-specific models have RHC ratios range between 1.5 and 2.5, where Massena_MOD+ENTRAIN 0.2 achieves a 1.5 RHC at Site 2.

[53] See, for example, the EPA 2015 presentation available at: ftp://newftp.epa.gov/Air/aqmg/SCRAM/conferences/2015_11th_Conference_On_Air_Quality_Modeling/Presentations/1-5_Proposed_Updates_AERMOD_System.pdf.

Figures B-12 and B-13 show the CPM values for Sites 1 and 2, respectively, and **Figure B-14 through Figure B-16** show the MCM values for both monitors separately and the combined (CMCM) for each model approach. The CPM values for the site-specific model approaches are smaller than the CPM values for the default model approach. As with the fractional bias comparison, CPM results for the site-specific model approaches indicate that Massena_MOD+ENTRAIN 0.2 performs best, followed by Massena_MOD+ENTRAIN, and lastly Massena_MOD. Also, as expected, the 90% confidence intervals for the CPMs for the site-specific model approaches do not overlap by a wide margin with the default model, meaning that there exists a statistically significant difference between the performance of each of the site-specific model vs. the default model. This is reflected in the MCM values, which are all positive, meaning that "model J" (Massena_MOD, Massena_MOD+ENTRAIN, or Massena_MOD+ENTRAIN 0.2) performs much better than "model I" (default model).

Table B-14: 5-Year Averaged Robust High Concentrations (µg/m^3) for Site 1

Model Approach	RHC	Pre/Obs Ratio
Observed	145.6	--
AERMOD Default Model	1046.9	7.2
Massena_MOD Model	329.2	2.3
Massena_MOD+ENTRAIN Model	296.2	2.0
Massena_MOD+ENTRAIN 0.2 Model	291.3	2.0

Table B-15: 5-Year Averaged Robust High Concentrations (µg/m^3) for Site 2

Model Approach	RHC	Pre/Obs Ratio
Observed	282.0	--
AERMOD Default Model	1435.7	5.1
Massena_MOD Model	706.0	2.5
Massena_MOD+ENTRAIN Model	555.0	2.0
Massena_MOD+ENTRAIN 0.2 Model	413.2	1.5

Figure B-9: Spatial Pattern of Design Concentrations for Massena_MOD

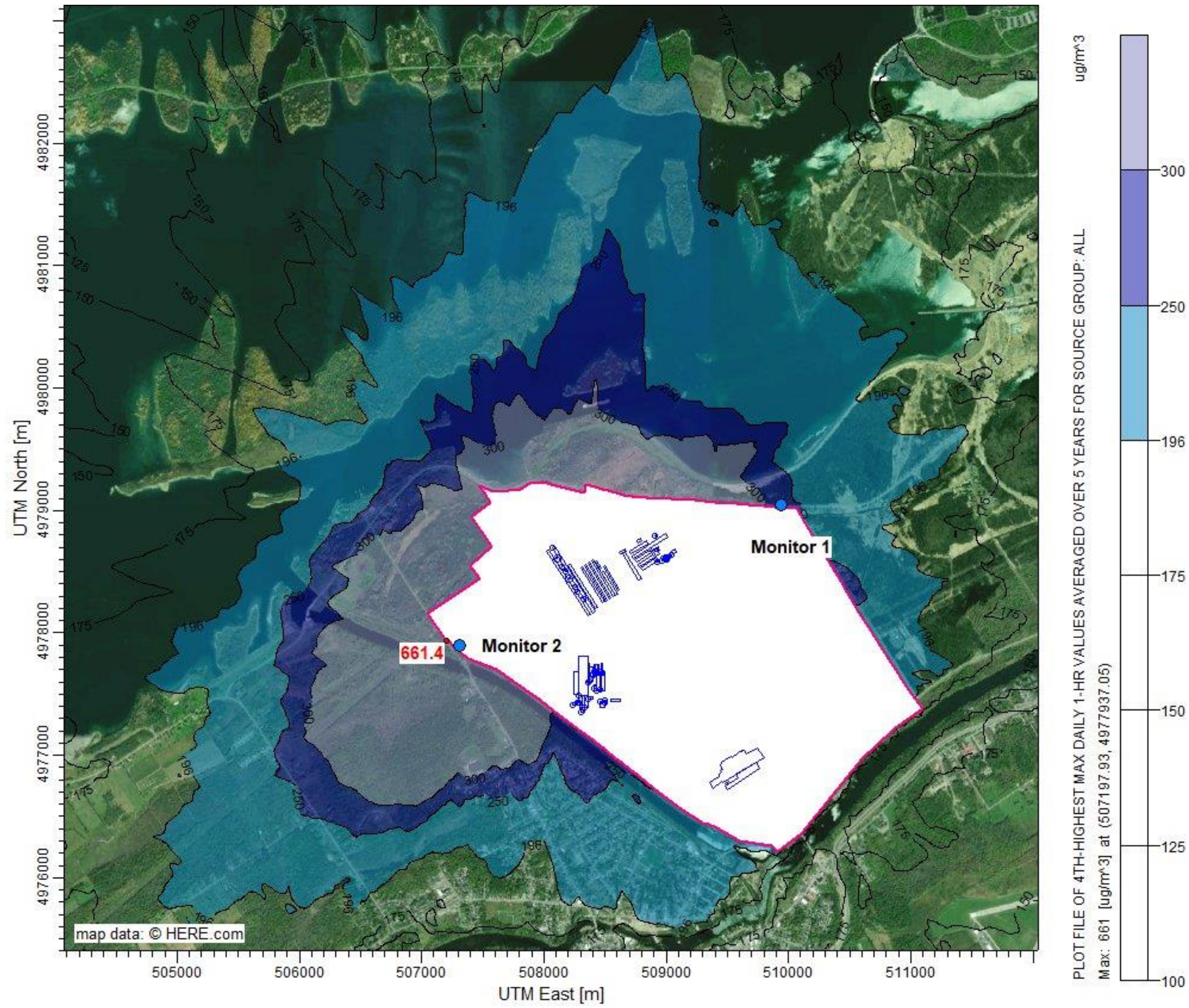


Figure B-10: Spatial Pattern of Design Concentrations for Massena_MOD+ENTRAIN

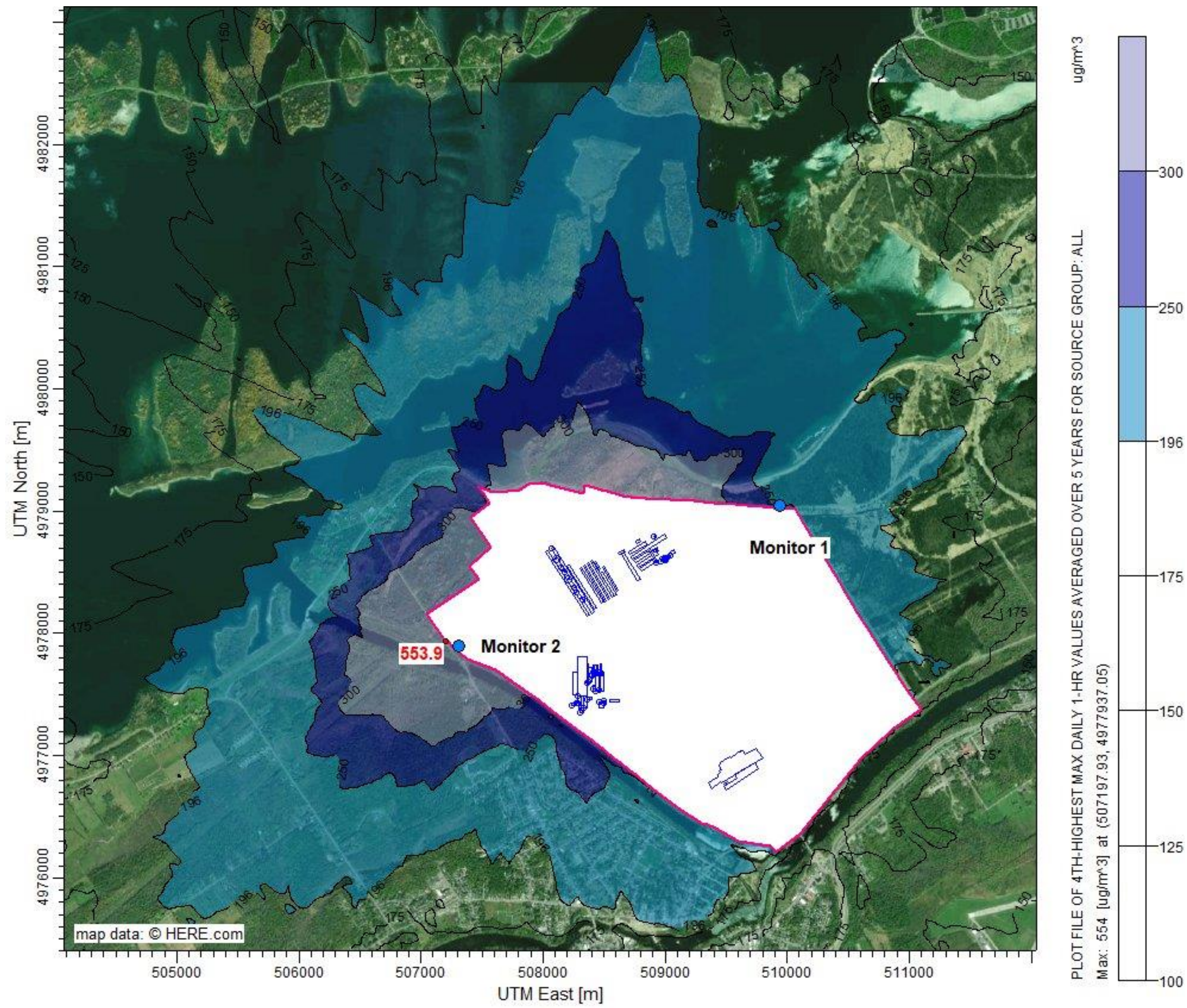


Figure B-11: Spatial Pattern of Design Concentrations for Massena_MOD+ENTRAIN 0.2

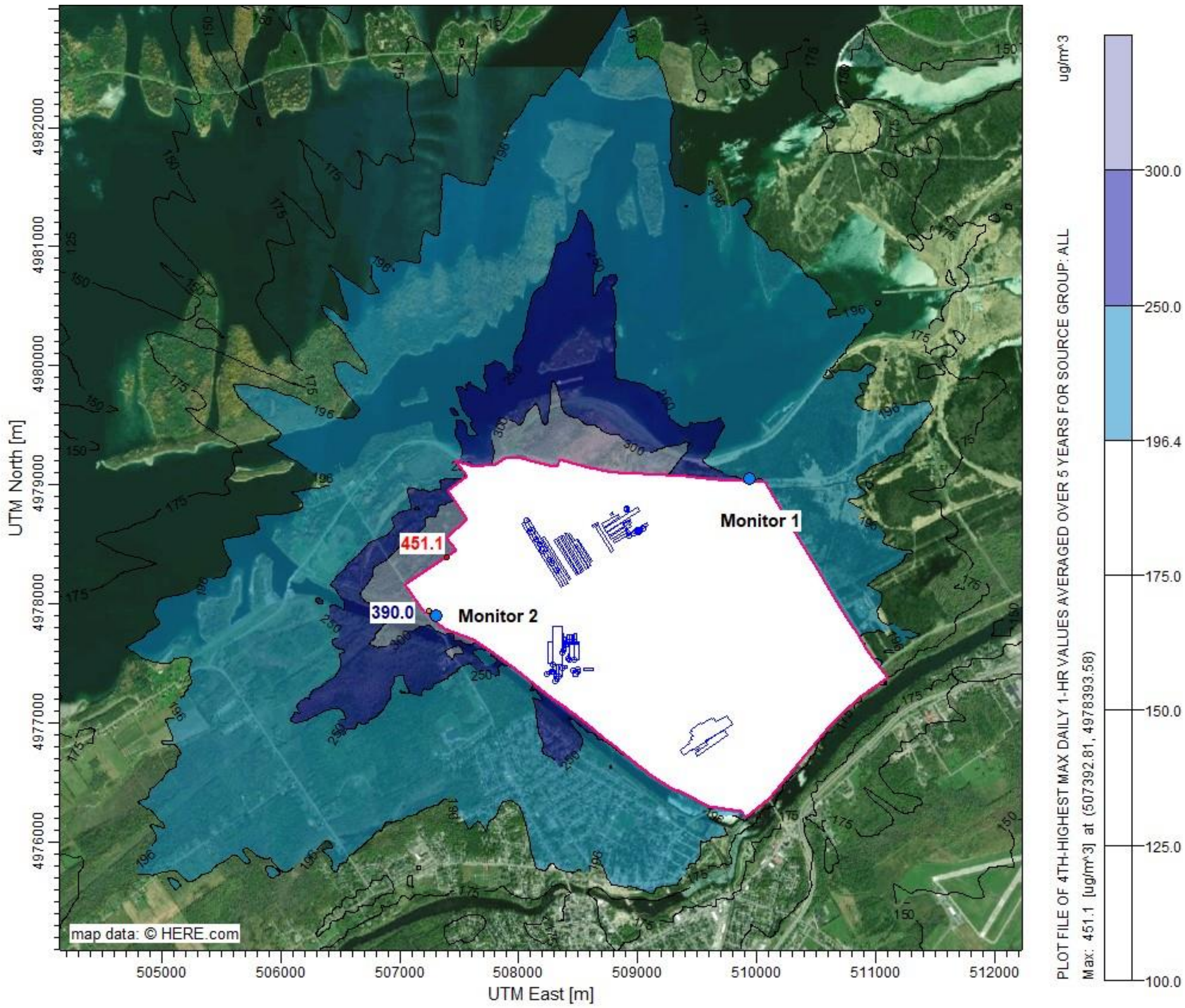


Figure B-12: Plot of CPM for Site 1

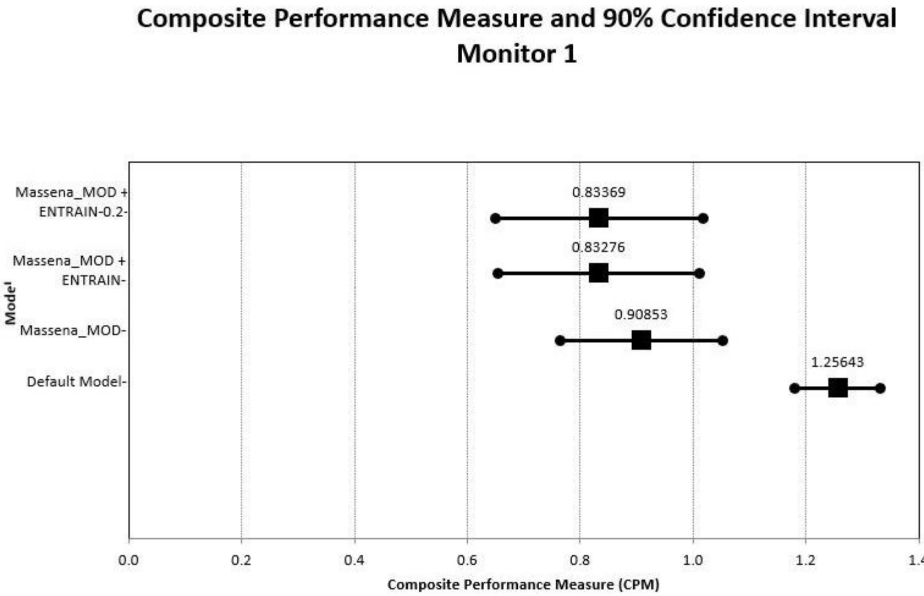


Figure B-13: Plot of CPM for Site 2

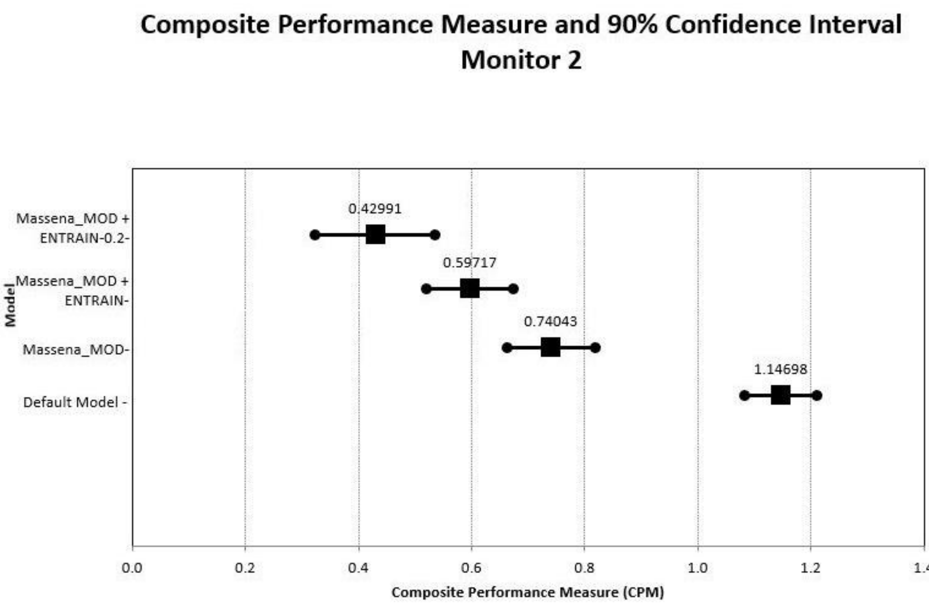


Figure B-14: AERMOD Default vs. Massena_MOD MCM, CMCM, and 90% Confidence Interval for Sites 1 and 2

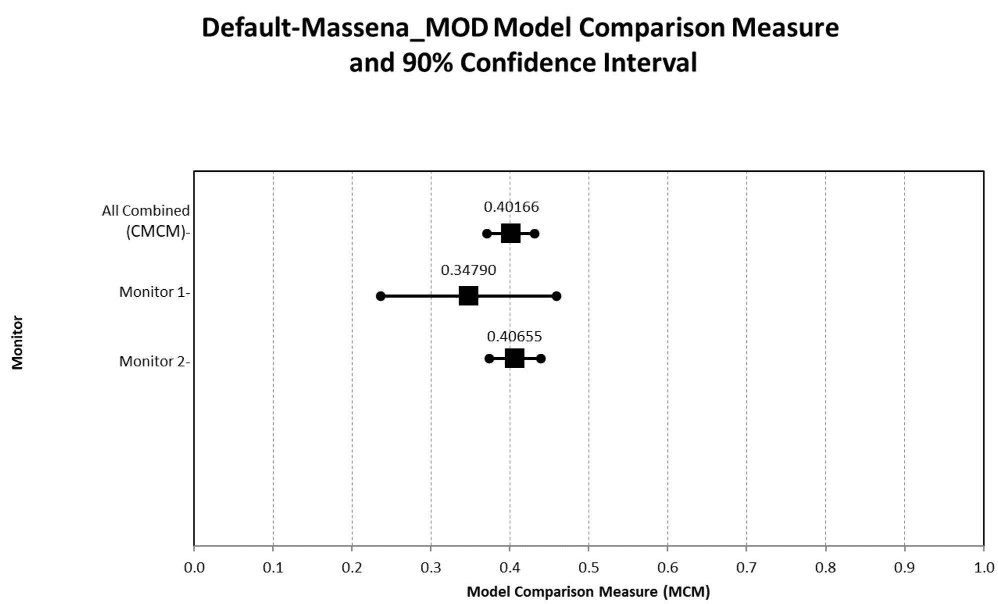


Figure B-15: AERMOD Default vs. Massena_MOD+ENTRAIN MCM, CMCM, and 90% Confidence Interval for Sites 1 and 2

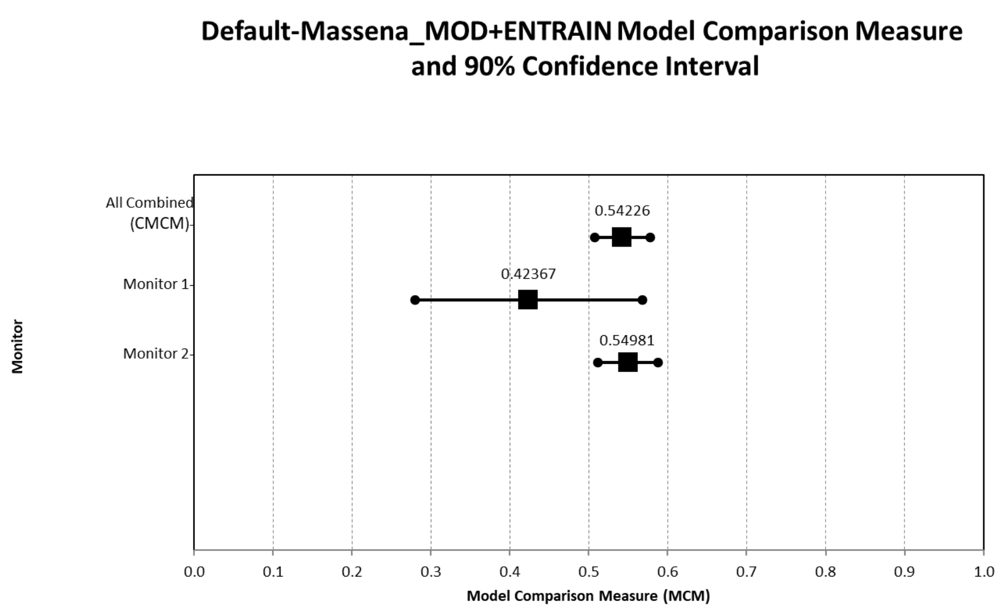
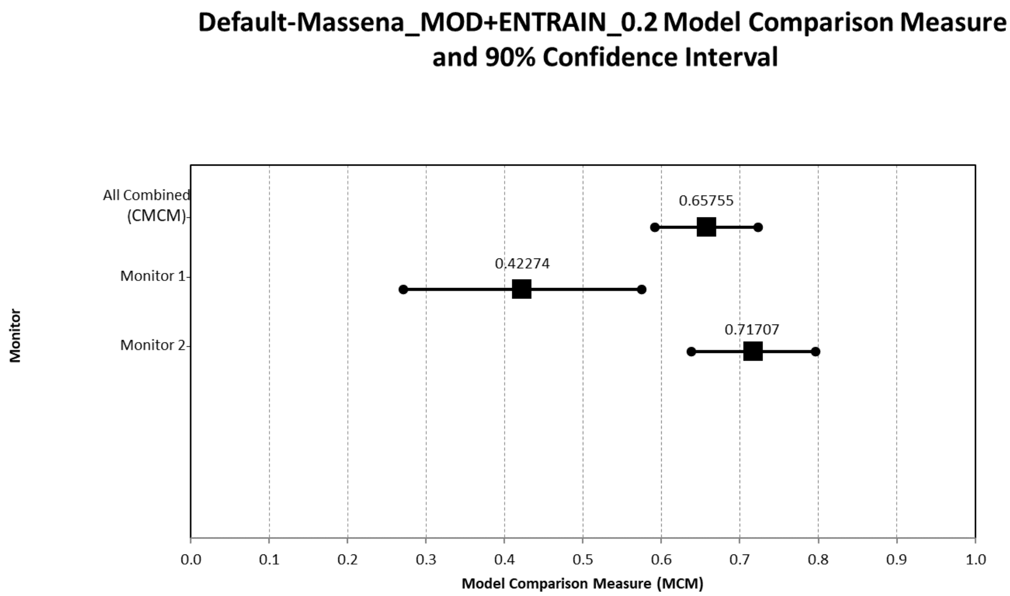


Figure B-16: AERMOD Default vs. Massena_MOD+ENTRAIN 0.2 MCM, CMCM, and 90% Confidence Interval for Sites 1 and 2



B.4 Conclusions

It is important for purposes of resolving the nonattainment area in the vicinity of the Alcoa Massena facility that a relatively unbiased dispersion model is used because the modeling results are used to determine the compliance status of the future source configuration and emissions. If the model is biased high, then the source may be forced to overcontrol or over-engineer a solution to the nonattainment issue that is unnecessary and overly burdensome. Therefore, Alcoa sought to use a site-specific model that reduces, to the extent possible, the overprediction tendency of AERMOD at this facility in the June 2023 modeling protocol.

EPA approved the use of Massena_MOD in October 2022 with site-specific use of a 100-m neutral lapse rate and a more appropriate version of the Building Profile Input Program (BPIP) for this site. However, following that approval, changes in the site-specific model were needed due to the need to separately model the roof vent emissions if the future potline reactor stacks are to be raised. In addition to this change, an improvement to the plume rise estimates was identified that incorporates use of two candidate site-specific entrainment coefficients (values of 0.35 and 0.2), both of which result in better model performance in two key areas: 1) the ratio of modeled to observed peak concentrations, and 2) a comparison of modeled and observed meteorological conditions for peak concentration events.

This appendix describes in detail the scientific justification for the improvements as well as provides the model evaluation results for several aspects of the model performance at the two monitoring sites: 1) quantile-quantile plots of the modeled and observed concentrations, 2) meteorological conditions associated with the peak 25 modeled and observed concentrations, and 3) a full evaluation using EPA's model evaluation protocol procedures as implemented using the Model Evaluation Methodology software. For the key Site 2 monitor, the model evaluation results indicate that the ratio of the model to observed design concentration with a 5-year data set improves from 5.6 with AERMOD default, 2.7 with Massena_MOD, 2.1 with Massena_MOD+ENTRAIN, and 1.5 with Massena_MOD+ENTRAIN 0.2. Therefore, the use of the Massena_MOD+ENTRAIN 0.2 site-specific modeling approach will be considered in the future as it was the best-performing model for this site.

aecom.com

aecom.com