# $SO_2$ Modeling Protocol for Alcoa Massena Operations - West Plant

Alcoa Massena Operations - West Plant
Massena, New York

Project number: 60610495.1

June 10, 2022



Delivering a better world

## Quality information

| Prepared by | Checked by | Verified by | Approved by |
|---|---|---|---|
| Adrienne L. Kieldsing<br>Air Quality Scientist | Robert J. Paine<br>Associate Vice President, Air Quality | Jeff Connors,<br>Supervisor, Air Quality Services (East) | Brian L. Stormwind<br>Manager, Air Quality Engineering & Studies |

## Revision History

| Revision | Revision date | Details | Authorized | Name | Position |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

## Distribution List

| # Hard Copies | PDF Required | Association / Company Name |
|---|---|---|
| | | |
| | | |
| | | |
| | | |

## Prepared for:

Alcoa Massena Operations - West Plant
Massena, New York

## Prepared by:

Adrienne L. Kieldsing
Air Quality Scientist
T: 978.905.2271
E: Adrienne.Kieldsing@aecom.com

AECOM
250 Apollo Drive
Chelmsford, MA 01824
aecom.com

# Table of Contents



# Figures

# Tables

# 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 ppb and the form of the standard is the average of the 99$^{th}$ 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 was subject to the DRR provisions is the Alcoa Massena Operations West aluminum smelter (Massena Operations).

Massena Operations has $SO_2$ emissions from three clusters of 12 dry scrubber stacks in the potline area, as well as from an anode bake oven stack (see **Figure 1-1**). It was clear after preliminary modeling with AERMOD with 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 east and Site 2 to the west, was initiated in January 2017. The monitoring sites are shown on a map in **Figure 1-2**.

Through December 2019, the monitored concentrations at Site 1 exceeded the 75 ppb $SO_2$ NAAQS (4$^{th}$ highest peak daily 1-hour maximum) only once over the entire 3-year period. As a result, the 99$^{th}$ 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 exceeded the NAAQS several times each year, so Site 2 had a 3-year design concentration (2017-2019) above the NAAQS[1].

Due to the reported concentrations at Site 2, NYSDEC worked with Alcoa to plan for the next steps, which include:

- Providing a modeling approach to attempt to replicate the monitored concentrations at the two monitoring sites, and
- Using the model to characterize the $SO_2$ concentrations in the vicinity of Massena Operations.

This modeling protocol presents a proposed dispersion modeling approach for documenting AERMOD-predicted 1-hour $SO_2$ concentration patterns resulting from Massena Operations potline and anode bake oven operations with current stacks as well as with a planned future stack arrangement. Other aluminum smelters in the United States have adopted site-specific modeling approaches due to the unique characteristics of the emission source. The modeling approach described in this protocol document involves site-specific source characterization approaches such as rural characterization, consideration of 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.

[1] 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 very steady $SO_2$ emissions from aluminum smelting operations.

**Figure 1-1: Location of the Modeled Alcoa Massena West Sources**

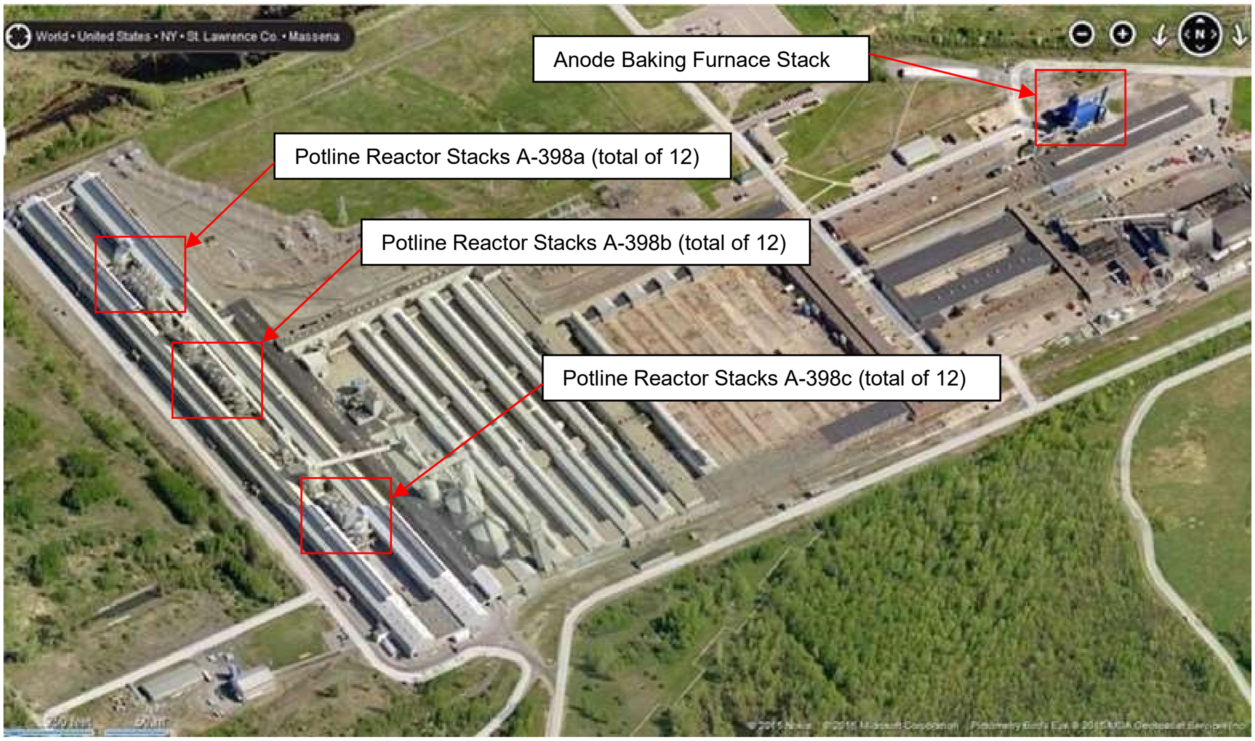
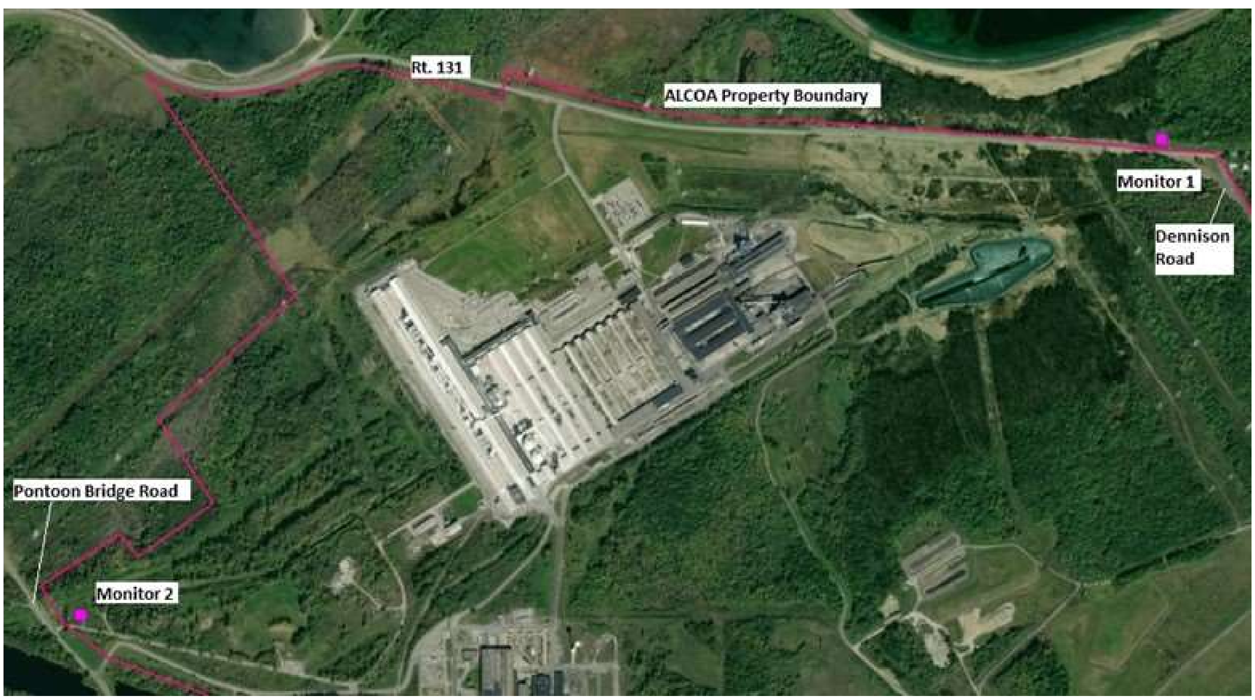



**Figure 1-2: Location of the Two $SO_2$ Monitoring Sites**



## 1.2 Document Organization

Section 2 provides a discussion of the $SO_2$ emission sources at Massena Operations, as well as a review of emission sources within 20 km of Massena Operations, using data from the 2017 National Emissions Inventory (NEI). The proposed approach for modeling Massena Operations as a result of extensive consultation with NYSDEC and EPA is provided in Section 3. Section 4 describes the results of a model evaluation approach to determine the better performing model (between the default AERMOD approach and the proposed site-specific approach that accounts for fugitive heat releases and EPA-provided downwash building pre-processor improvements). Section 5 describes how the better performing model will be used to demonstrate that the proposed stack changes will resolve the current nonattainment situation.

# 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. The facility has one potline consisting of two long rooms with three clusters of dry scrubber stacks between them. These dry scrubber stacks comprise most of the $SO_2$ emissions at the smelter, with one bake oven stack also emitting $SO_2$, as shown in **Figure 1-1**.

The $SO_2$ emission rates that will be used in the modeling of the facility to characterize the concentrations observed at the monitors were derived from the reported monthly averaged $SO_2$ emission rates for the potline systems and the annual averaged $SO_2$ emission rates for the bake oven dry scrubbing systems. Stack exhaust parameters and typical $SO_2$ emission rates are shown in **Table 2-1**.

$SO_2$ 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 directly proportional to the total aluminum produced at a smelter. The total aluminum produced is directly proportional to the operating current of the potline, how effective that current is at reducing alumina to aluminum, which is measured as percent current efficiency and the total number of pots in operation. Higher amperage and higher current efficiency along with total number of pots operating results in more aluminum being produced.

Since the %S in the anodes is very stable and the carbon consumption is consistent, the $SO_2$ emissions are proportional to the amount of aluminum produced. The smelter operates its potline in a very stable manner with very little variability in order to avoid process upsets and damage to the pots.

Due to the very stable operation of the smelter, the calculation of $SO_2$ emissions is done on a monthly basis using a calculation based upon the %S in the coke and the aluminum production. The variability of the aluminum production and the sulfur content in the coke during a month is very low, typically varying less than 5% from the monthly average and well within the permitted 10% relative accuracy specification for Continuous Emission Monitoring systems.

The $SO_2$ emissions from the roof vents on the potlines is very low, comprising no more than 5% of the total potline emissions (typically, 1-3%). Therefore, the roof vent emissions were conservatively added to the emissions for the dry scrubber stacks for modeling purposes. This approach is conservative in that it concentrates the roof vent emissions into the stacks with the vast majority of the emissions rather than dispersing the emissions over the long extent of the roof vents.

**Figure 2-1** shows the location of the modeled property boundary (red outlined area) around the facility within which Massena Operations controls public access. The area enclosed in purple is owned / controlled by Alcoa and the property owner, Arconic Massena LLC. Access to the blue-outlined area is also controlled by Alcoa. As a result, the final modeling for the resolution of the nonattainment area will cover the extent of the nonattainment area while excluding both the purple and blue-outlined areas from ambient air.

### Table 2-1: Typical Exhaust Parameters for $SO_2$ Point Emission Sources

| ID | Number of Stacks | Base Elevation (m) | Release Height (m) | Stack Diameter (m) | Exit Velocity (m/s) | Exit Temperature (K) | Typical $SO_2$ Emission Rate (g/s)(1,2) |
|---|---|---|---|---|---|---|---|
| RS_A | 12 | 65.5 | 23.3 | 1.22 | 10.37 | 369.2 | Monthly Varying |
| RS_B | 12 | 65.5 | 23.3 | 1.22 | 10.37 | 369.2 | Monthly Varying |
| RS_C | 12 | 65.5 | 23.3 | 1.22 | 10.37 | 369.2 | Monthly Varying |
| BAKEOVEN | 1 | 66.3 | 32.0 | 2.13 | 15.99 | 357.4 | Monthly Varying |

(1) See **Table 2-2** for the potline monthly-varying emissions. The modeling will use actual monthly emissions for the period of January 2017 through December 2019.
(2) See **Table 2-3** for the bake oven monthly-varying emissions. The modeling will use actual monthly emissions for the period of January 2017 through December 2019.

### Table 2-2: Monthly-Varying $SO_2$ Emissions (g/s) Per Stack for 2017-2019 for the Dry Potline Scrubber Stacks

| Month | 2017 | 2018 | 2019 |
|---|---|---|---|
| January | 1.841 | 1.850 | 1.882 |
| February | 1.938 | 1.915 | 1.918 |
| March | 1.841 | 1.882 | 1.858 |
| April | 1.924 | 1.808 | 1.934 |
| May | 1.932 | 1.845 | 1.746 |
| June | 1.861 | 1.767 | 1.887 |
| July | 1.881 | 1.871 | 1.969 |
| August | 1.787 | 1.740 | 1.962 |
| September | 1.819 | 1.929 | 1.972 |
| October | 1.849 | 1.900 | 1.852 |
| November | 1.759 | 1.850 | 1.799 |
| December | 1.793 | 1.904 | 1.857 |

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

**Table 2-3: Monthly -Varying $SO_2$ Emissions (g/s) for 2017-2019 for the Anode Bake Furnace Stack**

| Month | 2017 | 2018 | 2019 |
|---|---|---|---|
| January | 2.503 | 2.275 | 2.674 |
| February | 2.363 | 2.607 | 2.544 |
| March | 2.121 | 2.628 | 2.554 |
| April | 2.382 | 2.124 | 2.516 |
| May | 2.598 | 2.273 | 2.485 |
| June | 2.391 | 2.581 | 2.604 |
| July | 2.497 | 2.507 | 2.429 |
| August | 2.311 | 2.597 | 2.422 |
| September | 2.327 | 2.588 | 2.434 |
| October | 2.536 | 2.375 | 2.575 |
| November | 2.627 | 2.331 | 2.301 |
| December | 2.273 | 2.378 | 2.430 |

## 2.2 Nearby $SO_2$ Emission Sources

A review of nearby $SO_2$ sources in the 2017 National Emissions Inventory (NEI), with one update, shows that there are no sources within 50 km that had 2017 $SO_2$ emissions above 1 ton per year. Although the 2017 NEI indicates that Reynolds Metals St. Lawrence (about 11.5 km to the east of Massena Operations) had $SO_2$ emissions of 33.6 tons in 2017, that source was permanently shut down in 2014. Therefore, no sources other than Alcoa Operations will be modeled. The effects of non-modeled sources will be accommodated in regional background.

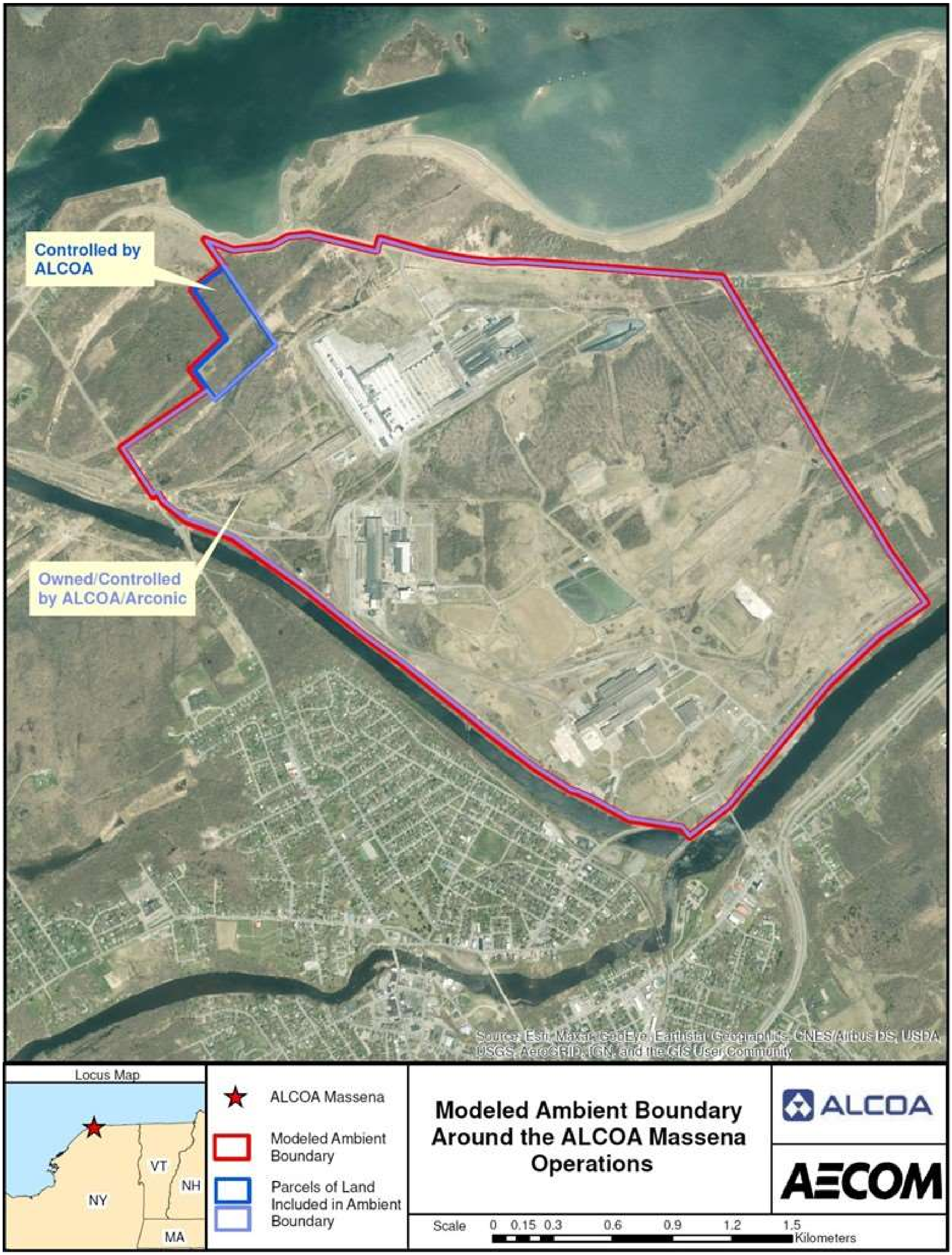
## 2.3 Regional Background

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.

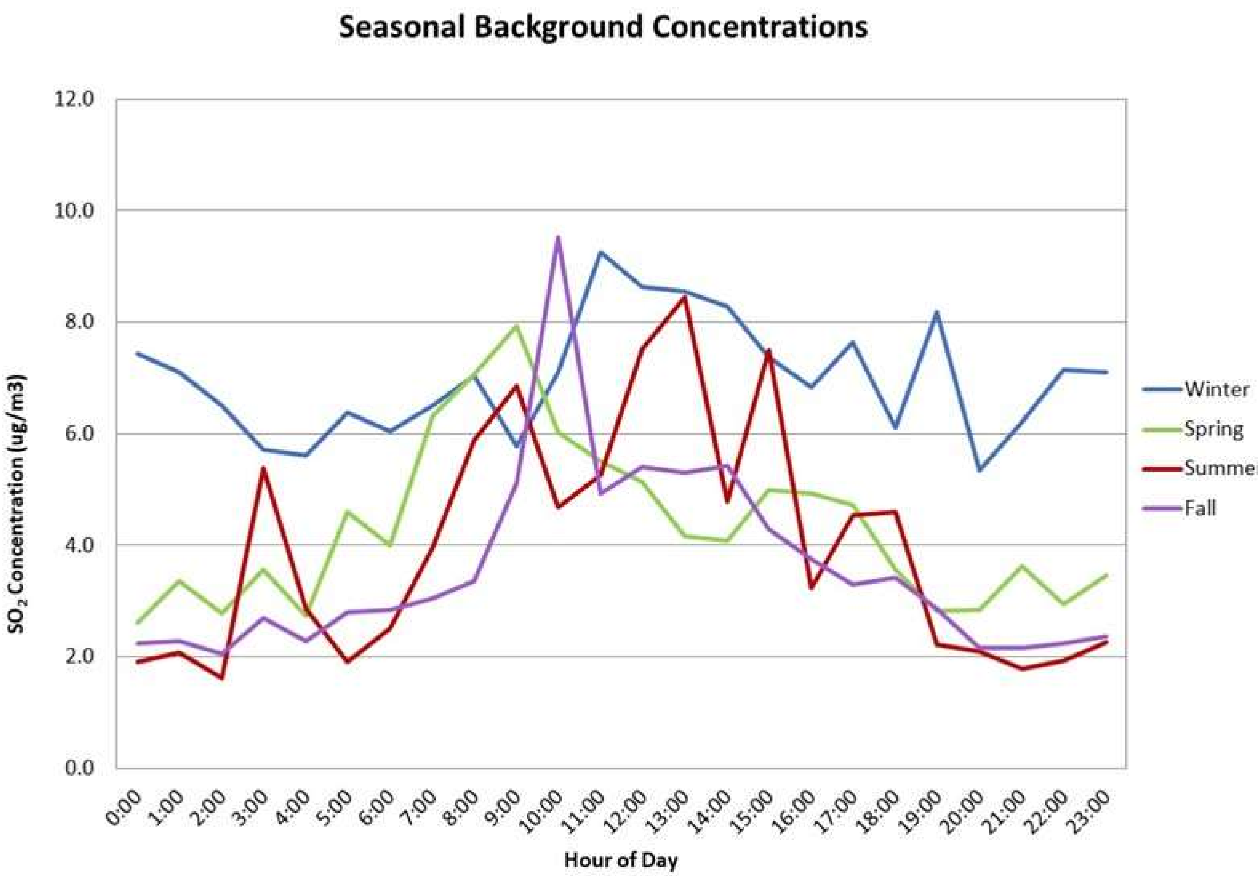
As shown in **Figure 1-2**, there are two monitors at Massena Operations for recording hourly $SO_2$ concentrations. To create regional background values for the modeling, NYSDEC used the hourly $SO_2$ data from the two monitors for the December 2016-November 2019 time period to create a synthetic background 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 the plume from Alcoa Massena Operations. It was determined that Monitor 1 was impacted by the plume when the wind direction was between 230° and 270°, and Monitor 2 was impacted by the plume 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. If the wind was not blowing from the facility towards the monitor, 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[2]) were computed. This approach was approved by EPA as documented in their $SO_2$ Round 4 Technical Support Document for New York[3]. **Figure 2-2** shows the season/hour-of-day values to be used in the modeling.

[2] https://www3.epa.gov/ttn/scram/guidance/clarification/Additional_Clarifications_AppendixW_Hourly-NO2-NAAQS_FINAL_03-01-2011.pdf.
[3] 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. Recently, two aluminum smelters (Alcoa Warrick Operations in Indiana and Intalco Works in Washington) have obtained approval of modeling approaches using site-specific source characterization. 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.

## 3.1 Consideration of Fugitive Heat Releases at the Smelter

The model to be used in this application is the latest AERMOD modeling system (currently, version 21112). The choice of rural or urban for dispersion conditions generally depends upon the land use characteristics within 3 kilometers of the facilities as described in Appendix W to 40 CFR Part 51[4]. 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 US 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, which are 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 K at Massena. This would result in an effective urban population of only 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 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[5] 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 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 would amount to about 50 MW.

**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.

In a recently released white paper[6], 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.

[4] EPA’s Guideline on Air Quality Models, available at https://www3.epa.gov/ttn/scram/guidance/guide/appw_17.pdf.
[5] See discussions at http://peter-entner.com/E/Theory/EBal/EBal.aspx and http://www.tms.org/pubs/journals/JOM/9905/Welch-9905.html.
[6] Hanna, S.R., 2022. Review of LIFTOFF Model as AECOM has Implemented into AERMOD. Report P210. February 6, 2022.

- 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/m2) 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 application was determined to be too small for using an urban dispersion option, the equivalent population of 250,000 supports a quasi-neutral neutral layer above the source of at least 200 m, according to the AERMOD Model Formulation and Evaluation document[7] (Equation 110).  Also, neutral conditions were observed in the lowest 100 m above the US Steel Clairton Works during a 1996 field study[8] 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 deg C per m (or 0.88 deg 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 improved the model performance by reducing overpredictions at the two nearby monitors, but additional changes were required to further improve the model performance, as noted in Section 3.2.

## 3.2 Proposed Enhancement for Treatment of Building Downwash

The residual overpredictions after application of the neutral lapse rate as discussed above are likely to be 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[9] 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")[10] 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, this improvement to the BPIP processing is a promising option, which did result in improved model performance, as documented in Section 4.

For purposes of ease of discussion and labeling, the site-specific application of AERMOD proposed for use at the Alcoa Massena smelter is referenced below as "Massena_MOD".

---

[7] Available at aermod_mfed.pdf (epa.gov).
[8] David Sullivan, 1996.  Review of Meteorology at the Clairton Area: Strengthening Dispersion Modeling for State Implementation Plans.  Submitted to the Allegheny County Health Department.
[9] See, for example, a 2018 presentation at ftp://newftp.epa.gov/Air/aqmg/SCRAM/workshops/2018_RSL_Modelers_Workshop/Presentations/1-13_2018_RSL-EPA_PRIME_Updates.pdf, slide 16.
[10] As documented at BPIPPRM_19191_DRFT-Trans_Memo.pdf (epa.gov).

## 3.3 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. 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 to December 31, 2019), the potline stacks were modeled with all stacks active as shown in **Figure 3-6**. When the western stacks were capped, the subsequent emissions from the western stacks were 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 as shown in **Figure 3-7**.

## 3.4 Meteorological Data Processing

Three full years (2017-2019) of hourly surface observations and one-minute wind speed and direction data from nearby Massena International Airport in Massena, New York will be used in conjunction with the twice-daily soundings upper air data from Albany, New York in AERMET (version 19191), 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 1-minute data to the Alcoa Massena facility. Likewise, Albany, NY is the closest upper air station. 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 3-8** shows the meteorological stations with respect to the Alcoa Massena smelter. The New York State Department of Environmental Conservation (NYSDEC) has provided the meteorological data for 2017-2019.

The surface meteorological data at the Massena International Airport is recorded by an Automated Surface Observing System (ASOS) that records 1-minute measurements of wind direction and wind speed (anemometer height of 10 meters), 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) will be run with the Massena 1-minute data. Five-minute data can 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-meter 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 ($\sigma_\theta$) and sigma-w ($\sigma_w$) 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 2019 revised AERMOD Implementation Guide (AIG).[11]

[11] U.S. EPA 2019. AERMOD Implementation Guide (Revised). U.S. Environmental Protection Agency, Research Triangle Park, NC. August, 2019.

The 2019 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 seasonal category. AERSURFACE will be 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 will be used in AERSURFACE for this modeling application.

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 will be divided into 12 sectors as shown in **Figure 3-9** and **Figure 3-10**.

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 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 will be designated as "airport" with the exception of the 240-270 degree sector which will be 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 3-year 2017-2019 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 will be 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 kilometer from the tower. The “ZOEFF” method is a research-grade method which does not limit the upwind fetch to 1 kilometer 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[12]. The default method is the “ZORAD” method, and this is the method that will be used for this modeling application.

## 3.5 Receptor Processing

Receptor input to AERMOD will be generated for areas of ambient air within the nonattainment area for the compliance modeling. For the model evaluation of the AERMOD guideline (default) model vs. the Massena_MOD approach, the two receptors corresponding to the two monitors will be 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.”[13] **Figure 2-1** (see previous section) illustrates the areas near the Massena Operations where general public access is controlled by Alcoa, and will serve as 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.

- Every 25 meters along the ambient air boundary
- Every 70 meters out to a distance of 2.5 km
- Every 100 meters between 2.5 and 5 km
- Every 500 meters between 5 and 10 km
- Every 1000 meters between 10 and 20 km

**Figures 3-11 and 3-12** provide the receptor grid as viewed in the near-field and far-field. If necessary, additional 100-meter spaced receptors will be placed at the location of the maximum impact should it not already be in an area covered by 100-meter (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.

[12] U.S. EPA 2020. User’s Guide for AERSURFACE Tool. U.S. Environmental Protection Agency, Research Triangle Park, NC. February, 2020..(https://gaftp.epa.gov/Air/aqmg/SCRAM/models/related/aersurface/aersurface_ug_v20060.pdf)
[13] https://www3.epa.gov/scram001/guidance/mch/ama4.txt

**Figure 3-1: Heat Generated by the Potlines**

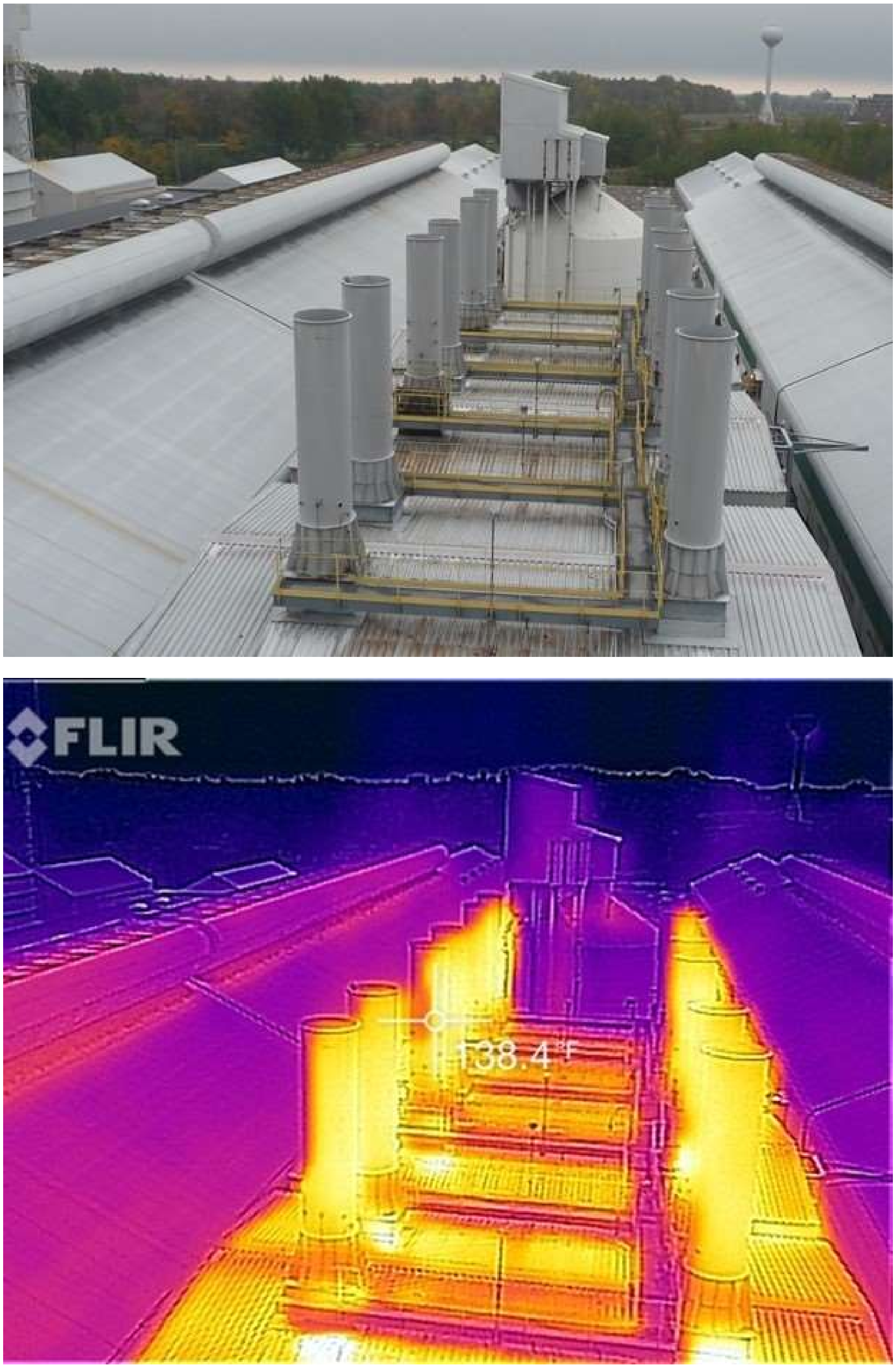


**Figure 3-2: Land Use Around Massena Facility**

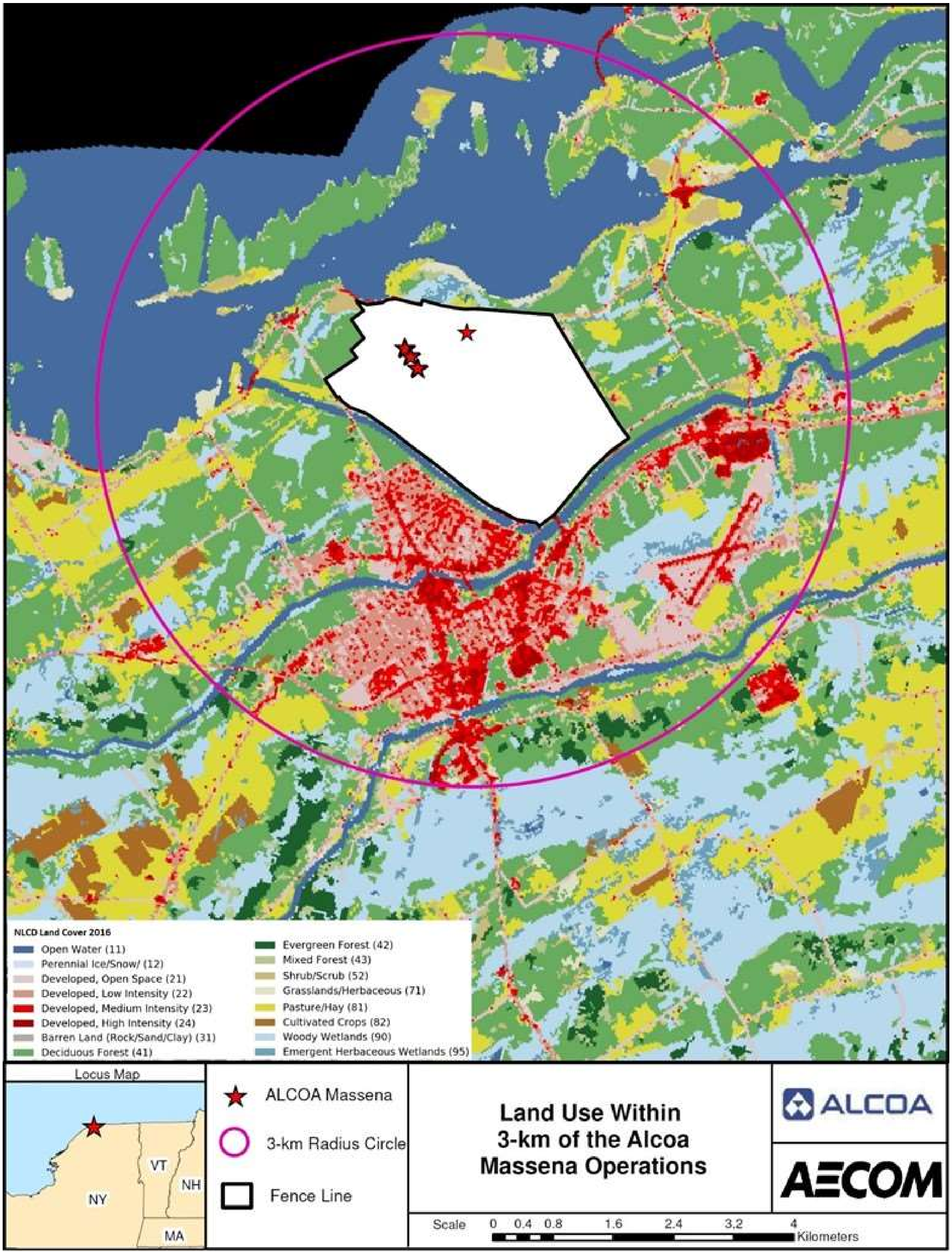


**Figure 3-3: Surface Temperature Data at Alcoa Massena Facility**

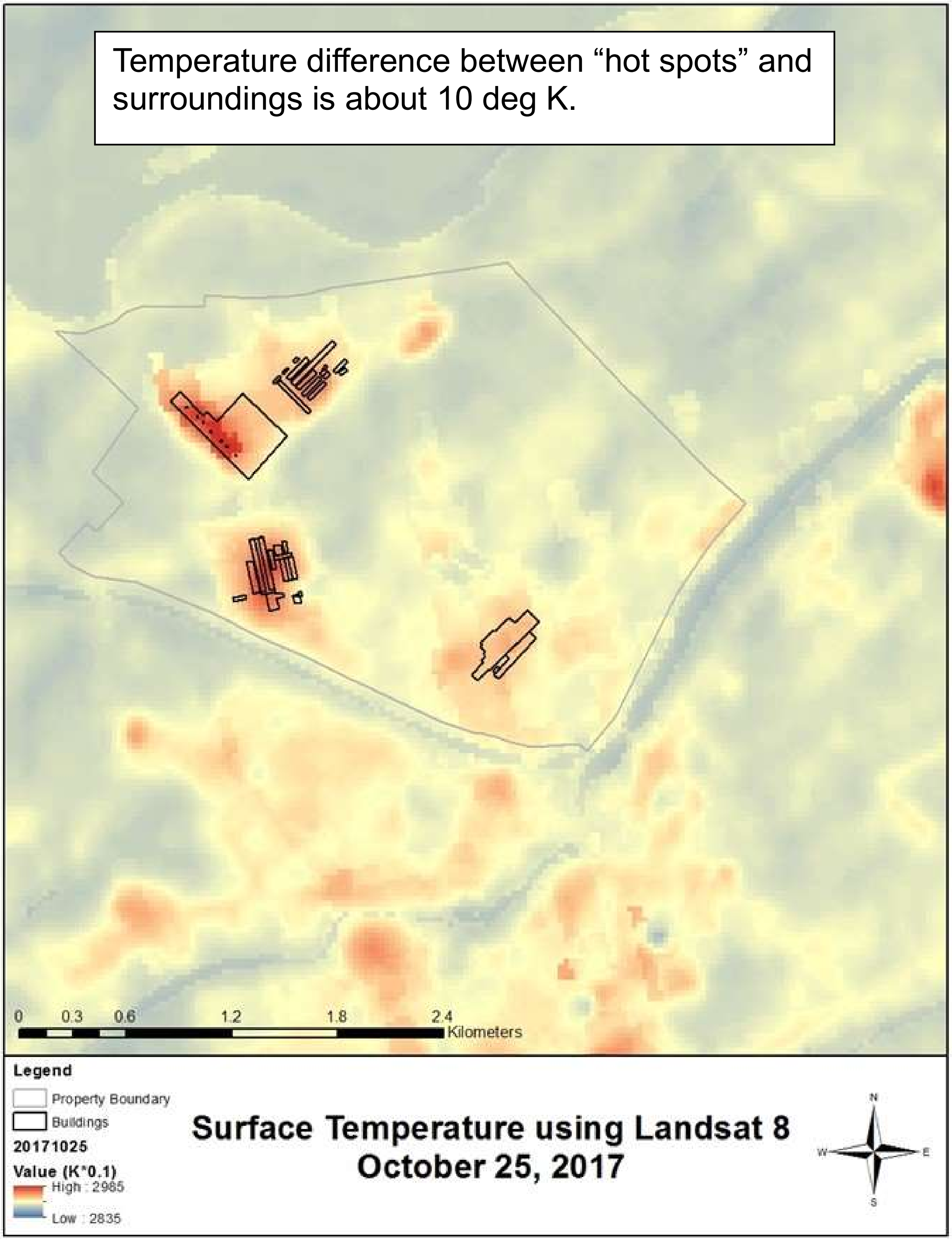
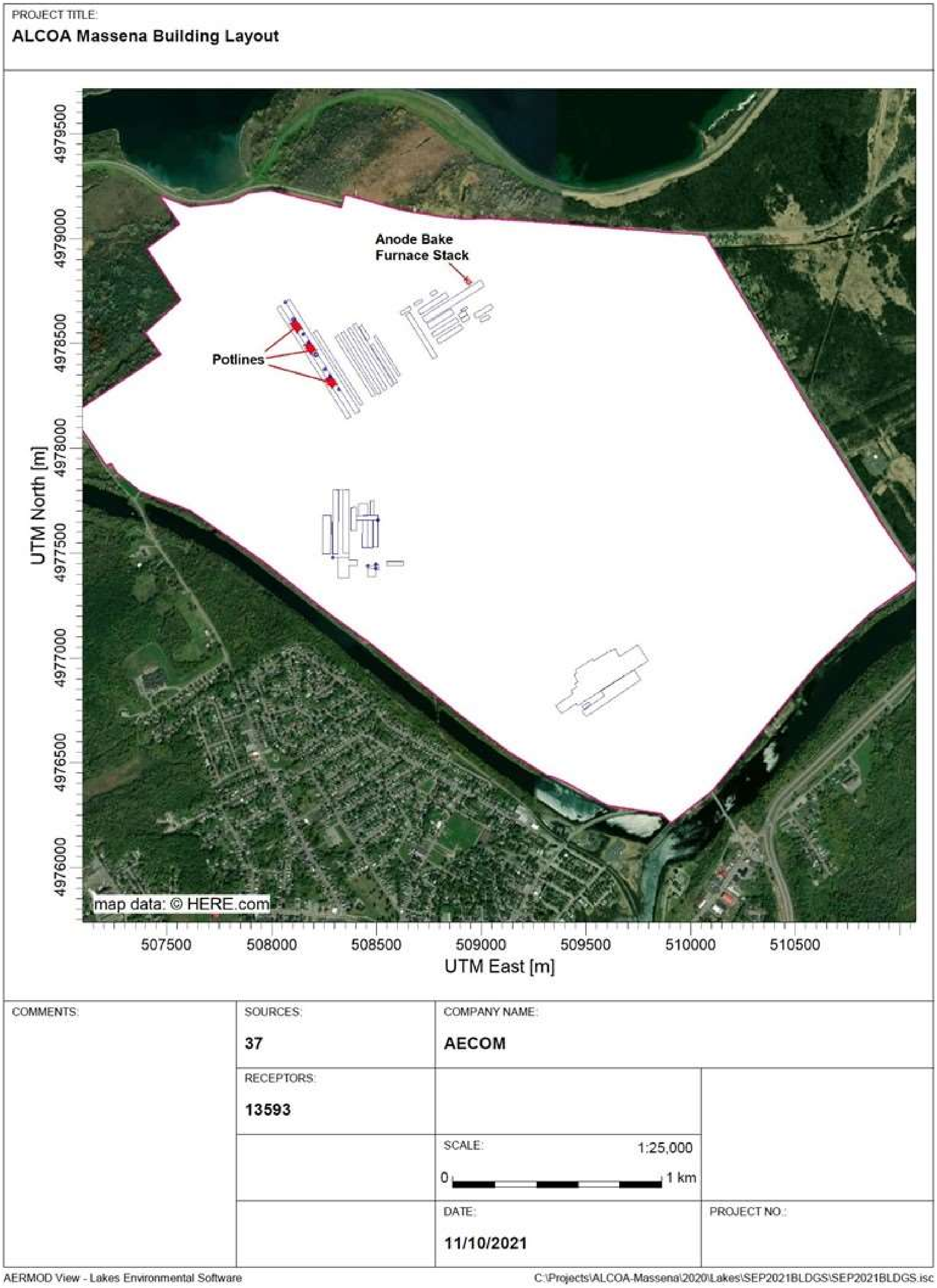


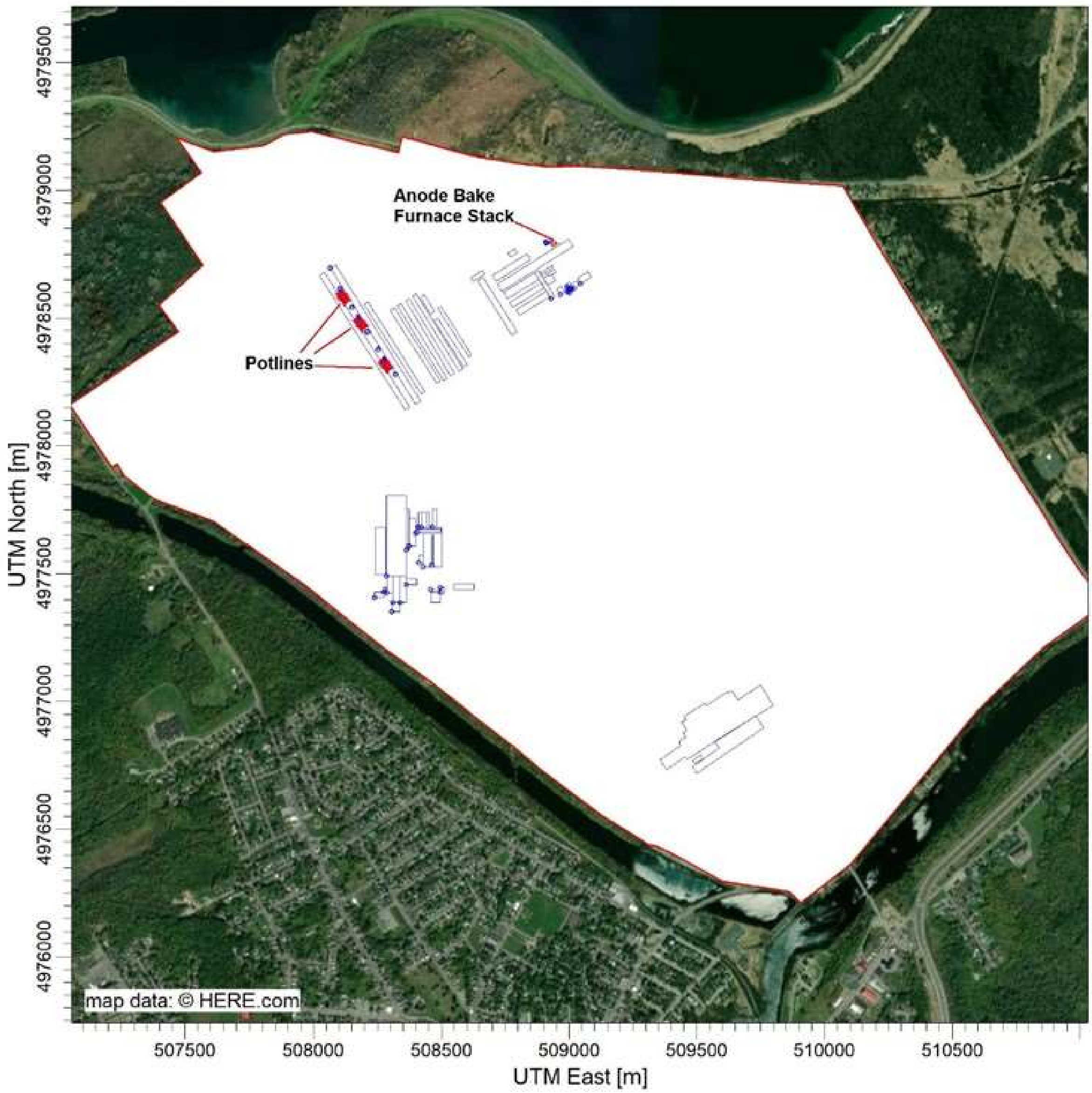
**Figure 3-4: Photo of one of the three potline dry scrubber stack clusters at Massena Operations**

### Figure 3-5: Modeled Building Layout



PROJECT TITLE:

**ALCOA Massena Operations Building Layout**




| COMMENTS: | SOURCES: | COMPANY NAME: | |
|---|---|---|---|
| | 37 | **AECOM** | |
| | RECEPTORS: 13595 | | |
| | | SCALE: 1:25,000 (0 — 1 km) | |
| | | DATE: **5/27/2022** | PROJECT NO.: |

AERMOD View - Lakes Environmental Software

C:\Projects\ALCOA-Massena\2020\Lakes\UnmgStksRevBldgs\UnmgStksRevBldgs.isc

**Figure 3-6: Modeled Location of Merged Stacks When No Stacks Are Capped**


**Figure 3-7: Modeled Location of Merged Stacks When Western Stacks Are Capped**

**Figure 3-8: Locations of Meteorological Stations**

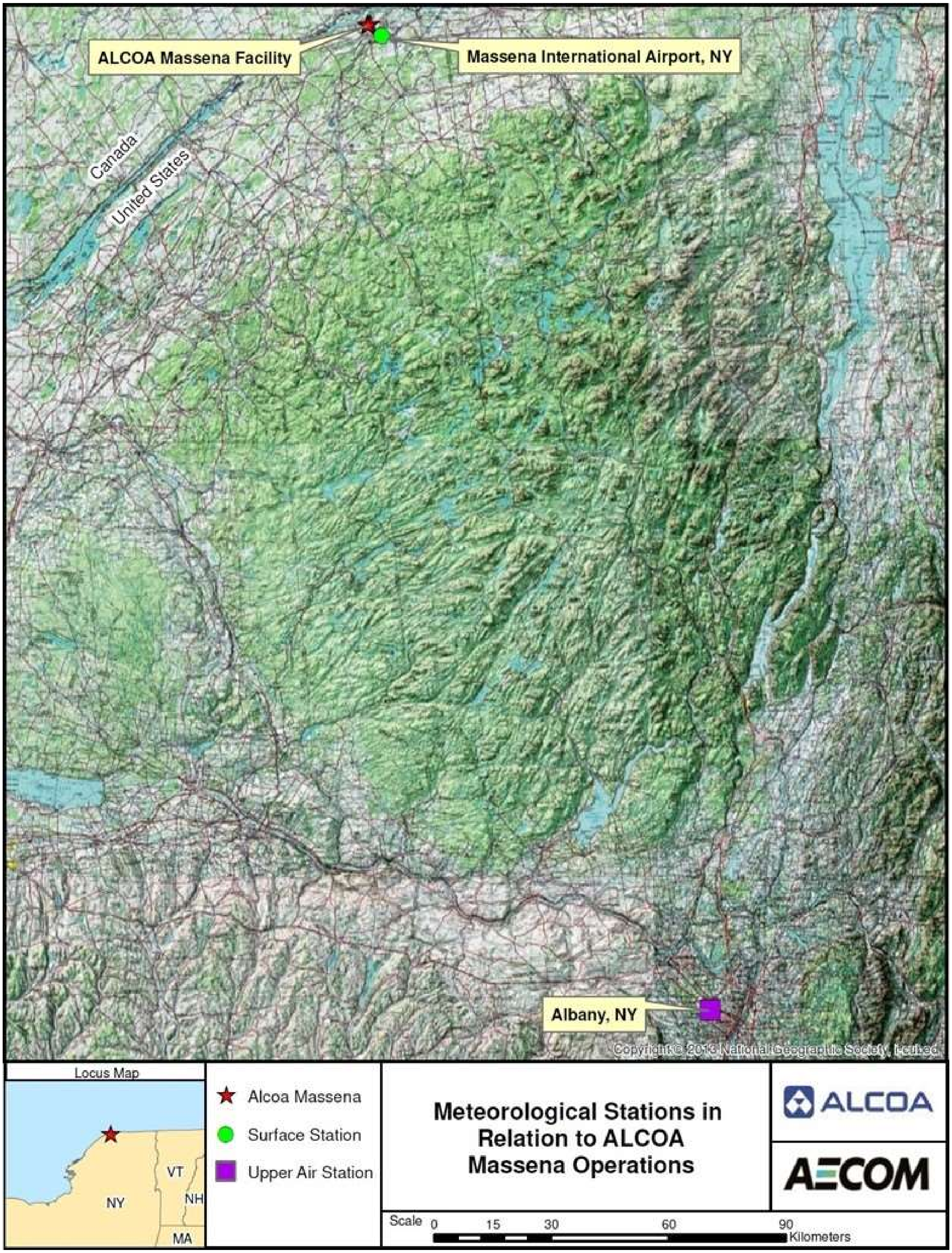


**Figure 3-9: Land Use Sectors Overlaid Over NLCD Data**

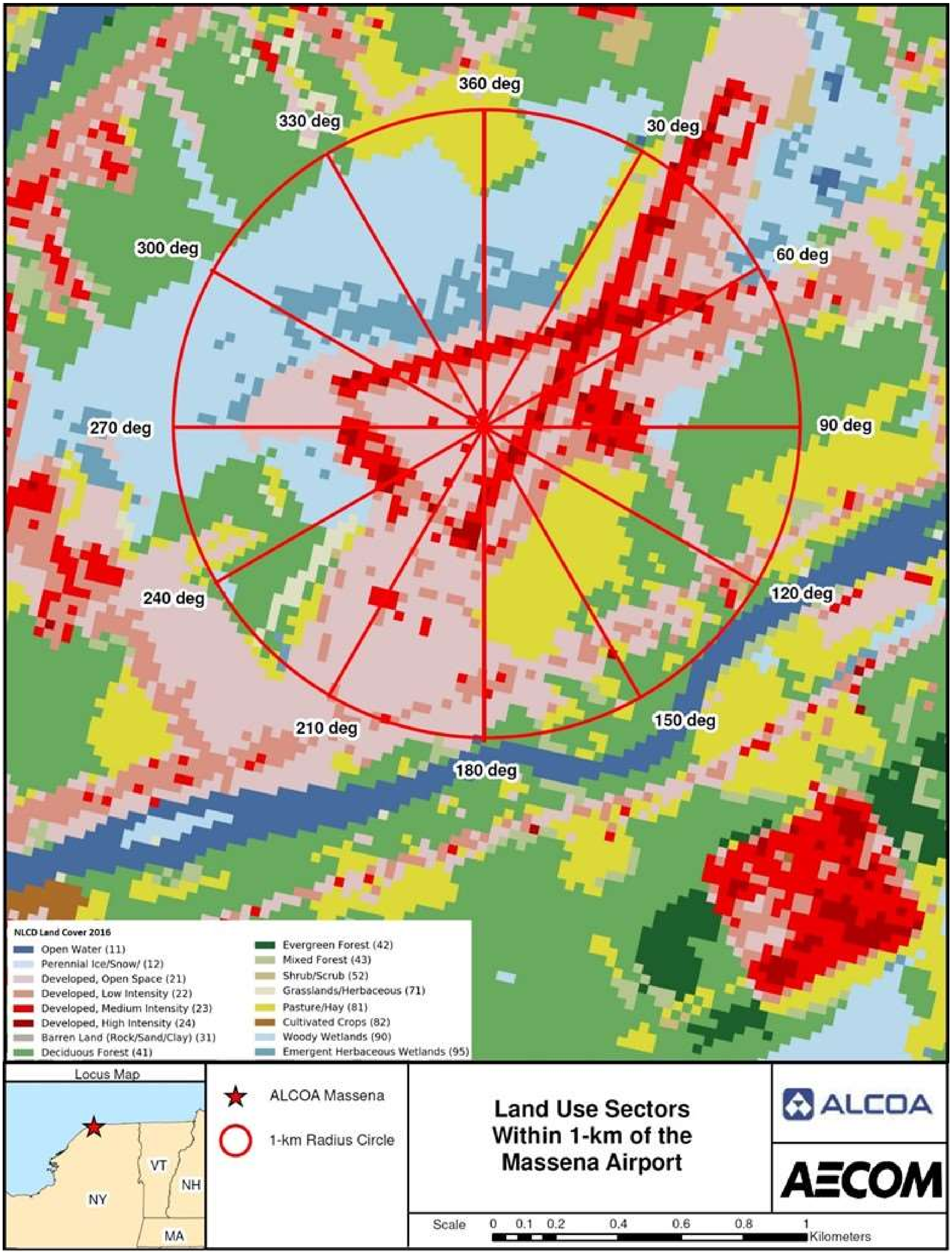


**Figure 3-8: Land Use Sectors Overlaid Over Aerial Photo**

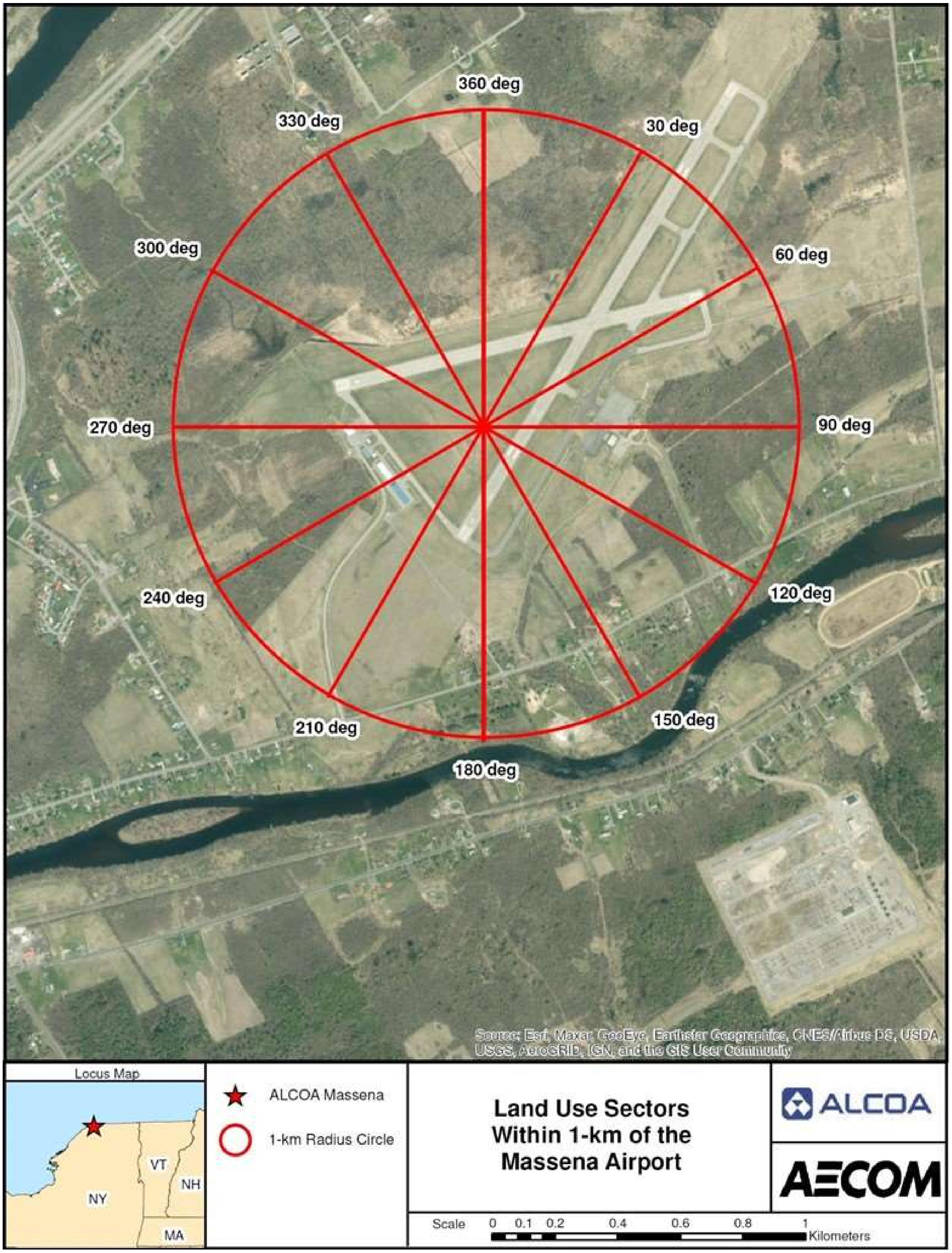
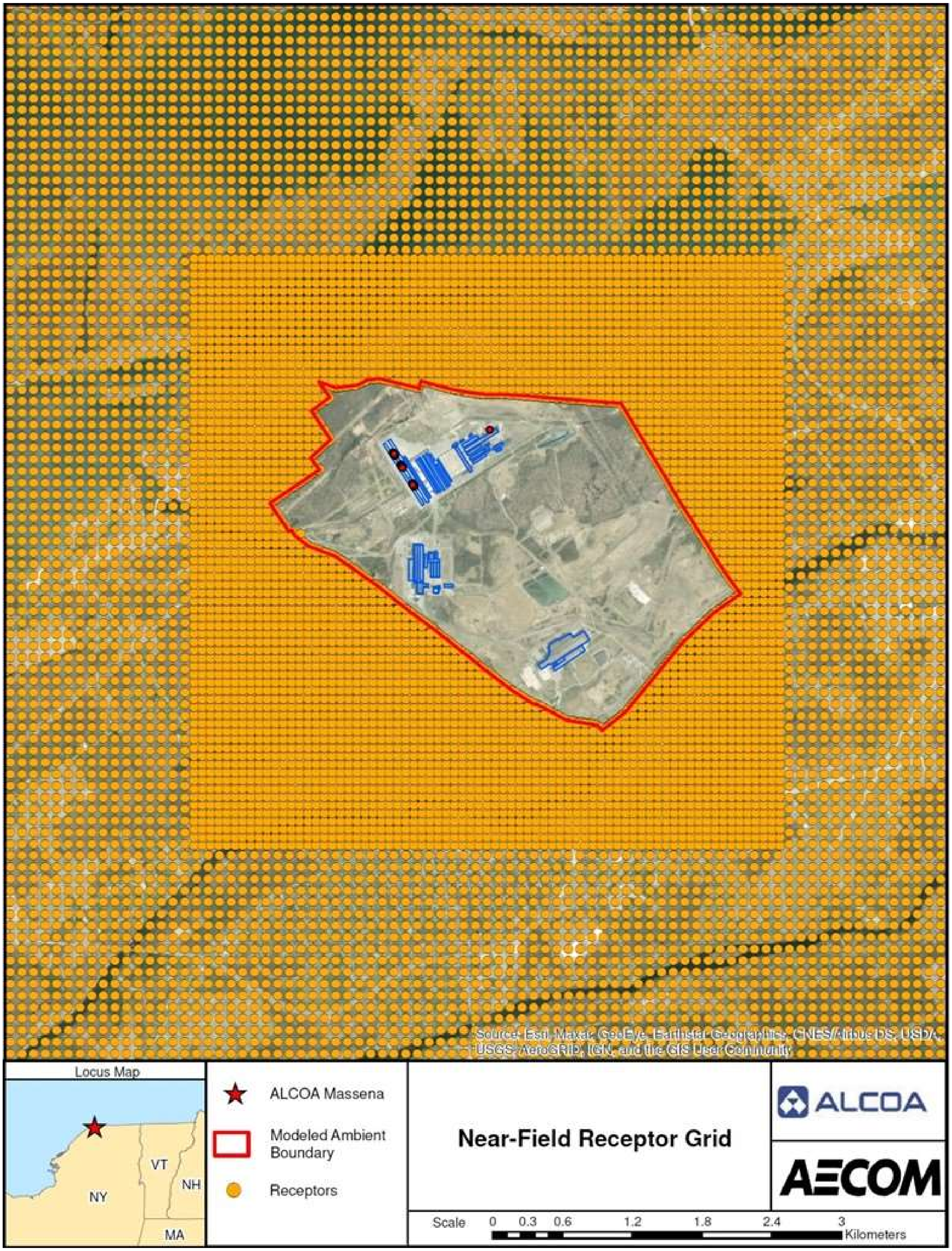
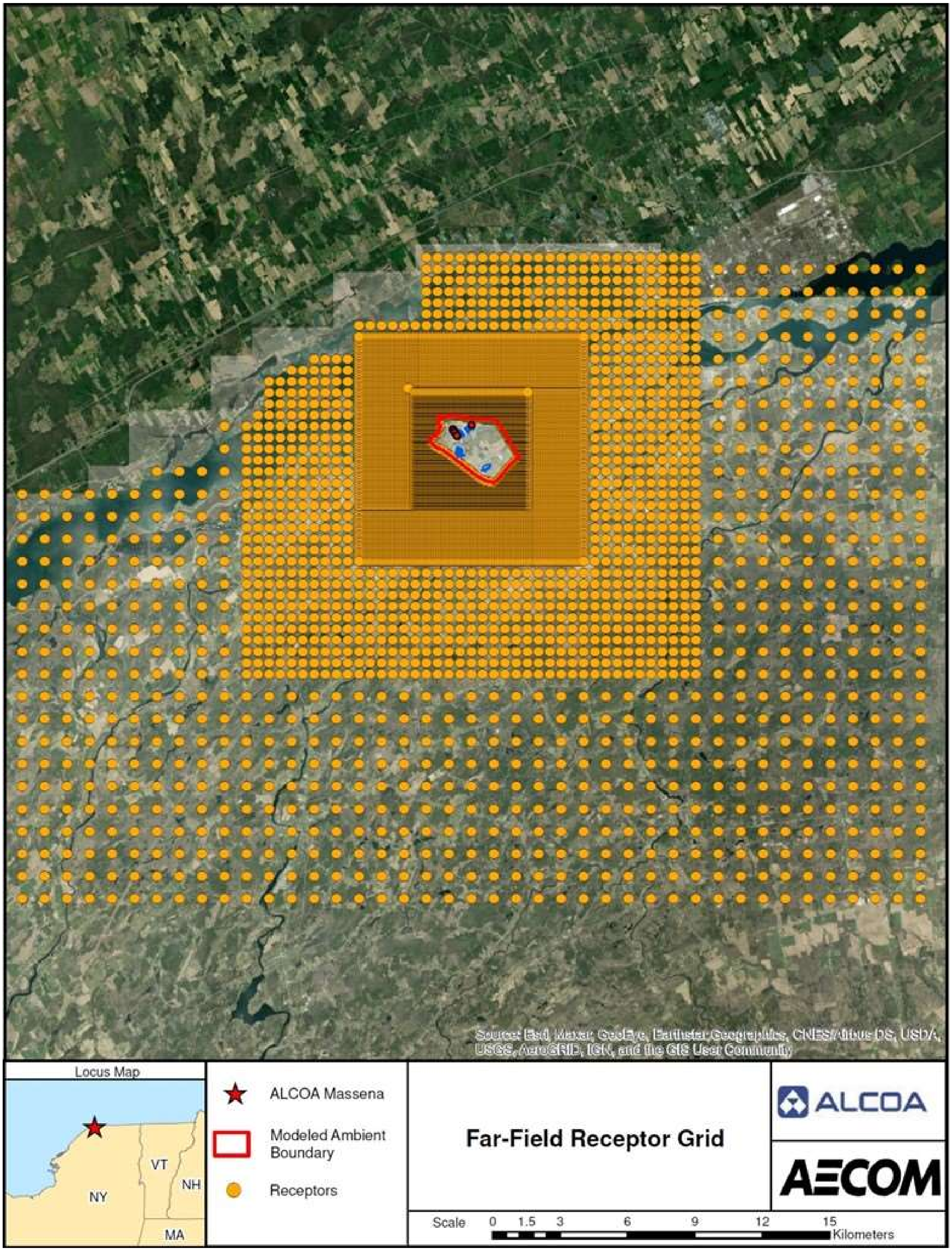


**Figure 3-11: Near-Field Receptor Grid**



## Figure 3-12: 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 ("Massena_MOD") is considered as a nonguideline technique, even though there is a strong scientific justification for this approach. It is also noteworthy that the site-specific changes involve only pre-processor steps, and the AERMOD model to be used has default options applied. 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 selects 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" (e.g., AERMOD run in default mode including downwash effects), then the alternative model may be approved for use for the proposed application.

Section 4.1 describes an analysis of meteorological conditions associated with the top 25 observed and top 25 predicted concentrations for the default (EPA's original BPIPPRM and the original AERMET-produced PROFILE file) and Massena_MOD (ORD BPIPPRM and a neutral lapse rate) models. This preliminary 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). Three full years of data (calendar years 2017 – 2019) will be used in for model evaluation of the site-specific source characterization to be used in this application. Meteorological data was processed using procedures described in Section 3.

## 4.1 Preliminary Analysis for Reviewing Model Performance: AERMOD-Default vs. Massena_MOD

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 preliminary 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, while Tables 4-5 and 4-6 provide the same information for Massena_MOD.

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 2017-2019 had 25 out of 25 events occurring at night, and Site 2 had 21 out of 25 of the hours occurring at night. In comparison, AERMOD-Default's top 25 modeled hours at Site 1 had 25 out of 25 hours at night, and had 25 out of 25 hours at night at 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 22 out of 25 hours at night, and had 25 out of 25 hours at night at Site 2.

Another preliminary 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 4.8 m/s at Site 1 and 5.4 m/s at Site 2. The average wind speed for over the top 25 modeled hours with AERMOD-Default at both sites was significantly lower at 0.9 m/s at Site 1 and 0.8 m/s at Site 2. The peak modeled events using AERMOD-Default have unexpectedly low wind speeds, indicating a significant mismatch in the meteorological conditions between the peak observations and peak modeled hours.

The average wind speed for over the top 25 modeled hours with Massena_MOD at both sites was 2.0 m/s at Site 1 and 2.2 m/s at Site 2. These averages are more consistent with the observations, indicating that Massena_MOD is performing better than AERMOD-Default at matching the wind speed conditions associated with peak concentration hours. Figures 4-1 and 4-2 show quantile-quantile plots of the ranked modeled and observed wind speeds, showing that the Massena_MOD model's peak concentration wind speed distribution is more consistent with the wind speed distribution for the top 25 observations than AERMOD-Default.

Yet another preliminary 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 546.1 m at Site 1 and 540.5 m at Site 2. The average mixing height over the top 25 modeled hours with AERMOD-Default at both sites was much lower at 52.4 m at Site 1 and 37.5 m at Site 2.

The average mixing height for over the top 25 modeled hours with Massena_MOD at both sites was 331.0 m at Site 1 and 112.0m at Site 2. These averages are more consistent with the observations, indicating that Massena_MOD is performing better than AERMOD-Default at matching the mixing heights associated with peak concentration hours. Figures 4-3 and 4-4 show quantile-quantile plots of the ranked modeled and observed mixing heights, showing that the Massena_MOD model's peak concentration mixing height distribution is more consistent with the mixing height distribution for the top 25 observations than AERMOD-Default.

Overall, these preliminary results indicate that Massena_MOD's results for the top 25 hourly concentrations is more consistent with observations than AERMOD-Default.

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

| Date | Hour | Monitored SO2 Concentration (ug/m3) | Wind Direction (degrees) | Wind Speed (m/s) | Mechanical Mixing Height (m) | Convective Mixing Height (m) |
|---|---|---|---|---|---|---|
| 3/22/2018 | 20 | 202.13 | 236 | 3.2 | 263 | N/A |
| 4/8/2017 | 21 | 167.08 | 222 | 3.37 | 285 | N/A |
| 2/27/2017 | 18 | 149.68 | 245 | 3.57 | 325 | N/A |
| 3/22/2018 | 19 | 136.66 | 242 | 3.22 | 356 | N/A |
| 2/14/2019 | 3 | 123.19 | 243 | 5 | 533 | N/A |
| 4/8/2017 | 20 | 120.31 | 238 | 3.1 | 302 | N/A |
| 2/27/2018 | 2 | 119.97 | 230 | 3.24 | 210 | N/A |
| 3/6/2019 | 3 | 119.71 | 244 | 4.13 | 507 | N/A |
| 4/3/2019 | 23 | 119.39 | 249 | 7.37 | 1255 | N/A |
| 12/19/2017 | 24 | 118.97 | 244 | 5.53 | 621 | N/A |
| 2/13/2017 | 2 | 118.69 | 241 | 4.29 | 418 | N/A |
| 2/14/2019 | 2 | 117.22 | 246 | 5.79 | 667 | N/A |
| 3/6/2019 | 1 | 116.12 | 234 | 3.45 | 295 | N/A |
| 2/13/2017 | 1 | 115.44 | 246 | 4.75 | 490 | N/A |
| 3/4/2019 | 23 | 110.20 | 247 | 4.64 | 608 | N/A |
| 1/25/2019 | 19 | 109.18 | 247 | 6.31 | 764 | N/A |
| 2/16/2019 | 6 | 108.99 | 249 | 6.16 | 735 | N/A |
| 4/4/2019 | 1 | 108.86 | 248 | 6.7 | 1089 | N/A |
| 2/7/2018 | 23 | 108.49 | 244 | 3.01 | 239 | N/A |
| 12/20/2017 | 3 | 107.87 | 246 | 5.82 | 672 | N/A |
| 2/14/2019 | 4 | 106.92 | 243 | 5.41 | 599 | N/A |
| 12/9/2018 | 22 | 106.61 | 238 | 4.12 | 306 | N/A |
| 1/25/2019 | 18 | 106.58 | 247 | 6.67 | 844 | N/A |
| 2/14/2019 | 6 | 106.45 | 238 | 5.19 | 438 | N/A |
| 2/26/2017 | 2 | 101.87 | 248 | 6.67 | 832 | N/A |

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

| Date/Hour | Date | Hour | Monitored SO2 Concentration (ug/m3) | Wind Direction (degrees) | Wind Speed (m/s) | Mechanical Mixing Height (m) | Convective Mixing Height (m) |
|---|---|---|---|---|---|---|---|
| 17040402 | 4/4/2017 | 2 | 265.43 | 58 | 6.75 | 691 | N/A |
| 17052508 | 5/25/2017 | 8 | 256.71 | 50 | 5.72 | 613 | 674 |
| 18051821 | 5/18/2018 | 21 | 252.99 | 50 | 5 | 433 | N/A |
| 18041621 | 4/16/2018 | 21 | 249.76 | 53 | 5.95 | 598 | N/A |
| 18051820 | 5/18/2018 | 20 | 246.59 | 48 | 5.29 | 474 | N/A |
| 17052509 | 5/25/2017 | 9 | 243.90 | 48 | 6.28 | 707 | 948 |
| 18020714 | 2/7/2018 | 14 | 243.63 | 49 | 5.82 | 484 | 109 |
| 17052521 | 5/25/2017 | 21 | 240.62 | 52 | 7.15 | 757 | N/A |
| 17032605 | 3/26/2017 | 5 | 240.23 | 68 | 2.86 | 221 | N/A |
| 17051219 | 5/12/2017 | 19 | 237.14 | 45 | 5.16 | 467 | N/A |
| 17031818 | 3/18/2017 | 18 | 236.17 | 50 | 7.45 | 826 | N/A |
| 17042421 | 4/24/2017 | 21 | 235.93 | 54 | 5.44 | 493 | N/A |
| 17040411 | 4/4/2017 | 11 | 234.28 | 57 | 4.59 | 433 | 310 |
| 17042420 | 4/24/2017 | 20 | 234.15 | 48 | 4.88 | 416 | N/A |
| 17060521 | 6/5/2017 | 21 | 233.86 | 43 | 5.12 | 476 | N/A |
| 17040404 | 4/4/2017 | 4 | 233.84 | 57 | 6.72 | 686 | N/A |
| 17031819 | 3/18/2017 | 19 | 231.19 | 54 | 7.48 | 812 | N/A |
| 18041620 | 4/16/2018 | 20 | 230.87 | 50 | 8.54 | 1016 | N/A |
| 17040318 | 4/3/2017 | 18 | 230.38 | 52 | 6.96 | 754 | N/A |
| 18051804 | 5/18/2018 | 4 | 228.49 | 45 | 4.66 | 387 | N/A |
| 17031823 | 3/18/2017 | 23 | 227.47 | 66 | 4.97 | 528 | N/A |
| 19101120 | 10/11/2019 | 20 | 226.73 | 46 | 3.36 | 246 | N/A |
| 18032106 | 3/21/2018 | 6 | 226.71 | 58 | 1.69 | 77 | N/A |
| 18060304 | 6/3/2018 | 4 | 226.42 | 54 | 3.46 | 258 | N/A |
| 17071221 | 7/12/2017 | 21 | 225.01 | 52 | 4.26 | 358 | N/A |

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

| Date | Hour | Default Modeled SO2 Concentration (ug/m3) | Wind Direction (degrees) | Wind Speed (m/s) | Mechanical Mixing Height (m) | Convective Mixing Height (m) |
|---|---|---|---|---|---|---|
| 2/20/2017 | 19 | 941.81 | 251 | 0.79 | 42 | N/A |
| 12/3/2019 | 7 | 926.41 | 252 | 0.77 | 42 | N/A |
| 4/14/2017 | 4 | 869.34 | 252 | 0.7 | 45 | N/A |
| 5/7/2018 | 20 | 849.27 | 253 | 0.92 | 54 | N/A |
| 4/23/2017 | 3 | 841.34 | 253 | 0.97 | 56 | N/A |
| 4/23/2017 | 4 | 840.87 | 251 | 0.77 | 48 | N/A |
| 4/22/2017 | 24 | 829.74 | 251 | 0.89 | 52 | N/A |
| 9/16/2018 | 24 | 824.26 | 250 | 0.7 | 45 | N/A |
| 3/18/2018 | 4 | 819.60 | 253 | 0.84 | 52 | N/A |
| 6/6/2018 | 21 | 800.66 | 252 | 0.67 | 45 | N/A |
| 3/30/2018 | 24 | 800.34 | 254 | 0.96 | 56 | N/A |
| 6/10/2018 | 22 | 798.43 | 252 | 0.69 | 46 | N/A |
| 7/15/2018 | 21 | 791.69 | 250 | 0.77 | 48 | N/A |
| 10/2/2017 | 4 | 786.63 | 253 | 0.6 | 43 | N/A |
| 11/4/2017 | 7 | 778.83 | 254 | 0.84 | 51 | N/A |
| 7/29/2017 | 23 | 762.32 | 250 | 0.8 | 50 | N/A |
| 9/14/2018 | 22 | 746.27 | 249 | 0.79 | 48 | N/A |
| 8/16/2017 | 20 | 743.94 | 253 | 0.91 | 65 | N/A |
| 7/11/2018 | 21 | 735.48 | 250 | 0.87 | 57 | N/A |
| 3/24/2017 | 20 | 720.05 | 251 | 1.04 | 61 | N/A |
| 6/15/2018 | 21 | 716.14 | 250 | 0.82 | 51 | N/A |
| 9/3/2018 | 22 | 714.34 | 254 | 1.18 | 73 | N/A |
| 8/27/2017 | 2 | 705.09 | 251 | 0.97 | 60 | N/A |
| 2/28/2019 | 8 | 704.09 | 254 | 1.11 | 51 | N/A |
| 7/11/2018 | 22 | 703.27 | 254 | 1.09 | 69 | N/A |

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

| Date | Hour | Default Modeled SO2 Concentration (ug/m3) | Wind Direction (degrees) | Wind Speed (m/s) | Mechanical Mixing Height (m) | Convective Mixing Height (m) |
|---|---|---|---|---|---|---|
| 10/24/2017 | 17 | 1513.38 | 55 | 0.72 | 35 | N/A |
| 9/13/2018 | 4 | 1498.25 | 55 | 0.74 | 35 | N/A |
| 5/23/2017 | 24 | 1475.49 | 54 | 0.79 | 36 | N/A |
| 7/15/2018 | 20 | 1436.94 | 52 | 0.6 | 33 | N/A |
| 10/7/2017 | 1 | 1421.76 | 55 | 0.65 | 34 | N/A |
| 2/10/2018 | 20 | 1389.75 | 53 | 0.89 | 35 | N/A |
| 5/13/2018 | 5 | 1388.18 | 55 | 0.77 | 36 | N/A |
| 5/9/2018 | 21 | 1381.81 | 52 | 0.63 | 33 | N/A |
| 8/1/2017 | 21 | 1372.59 | 54 | 0.57 | 33 | N/A |
| 2/14/2018 | 1 | 1367.43 | 55 | 0.91 | 36 | N/A |
| 12/24/2019 | 20 | 1367.27 | 53 | 0.8 | 33 | N/A |
| 2/22/2017 | 5 | 1364.33 | 55 | 0.99 | 36 | N/A |
| 5/28/2018 | 4 | 1360.91 | 54 | 0.92 | 39 | N/A |
| 8/3/2018 | 21 | 1351.67 | 53 | 0.79 | 36 | N/A |
| 5/9/2018 | 20 | 1349.90 | 52 | 0.89 | 49 | N/A |
| 10/18/2017 | 23 | 1343.02 | 52 | 0.79 | 36 | N/A |
| 4/24/2018 | 4 | 1327.60 | 53 | 0.72 | 35 | N/A |
| 2/21/2017 | 2 | 1325.42 | 51 | 0.86 | 34 | N/A |
| 9/20/2017 | 20 | 1317.38 | 53 | 0.96 | 40 | N/A |
| 6/9/2018 | 22 | 1315.94 | 55 | 0.82 | 65 | N/A |
| 10/6/2017 | 21 | 1314.89 | 51 | 0.8 | 36 | N/A |
| 9/7/2017 | 3 | 1305.75 | 54 | 0.96 | 41 | N/A |
| 9/19/2017 | 6 | 1305.48 | 53 | 0.96 | 40 | N/A |
| 5/24/2017 | 1 | 1304.80 | 50 | 0.79 | 36 | N/A |
| 8/14/2017 | 23 | 1299.47 | 51 | 0.72 | 35 | N/A |

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

| Date | Hour | Massena_MOD Modeled SO2 Concentration (ug/m3) | Wind Direction (degrees) | Wind Speed (m/s) | Mechanical Mixing Height (m) | Convective Mixing Height (m) |
|---|---|---|---|---|---|---|
| 9/3/2018 | 22 | 362.12 | 254 | 1.18 | 73 | N/A |
| 1/21/2017 | 20 | 322.65 | 256 | 1.43 | 73 | N/A |
| 7/11/2018 | 22 | 294.40 | 254 | 1.09 | 69 | N/A |
| 5/23/2017 | 20 | 285.12 | 254 | 1.57 | 110 | N/A |
| 12/26/2017 | 20 | 283.34 | 255 | 1.74 | 99 | N/A |
| 4/7/2019 | 10 | 281.35 | 257 | 0.94 | 95 | 106 |
| 5/24/2018 | 1 | 281.06 | 255 | 1.37 | 88 | N/A |
| 2/12/2018 | 20 | 266.58 | 253 | 2.13 | 137 | N/A |
| 6/8/2017 | 22 | 264.73 | 250 | 1.21 | 79 | N/A |
| 10/21/2017 | 22 | 263.61 | 253 | 1.45 | 100 | N/A |
| 6/23/2017 | 20 | 261.93 | 253 | 2.81 | 315 | N/A |
| 5/17/2018 | 6 | 252.16 | 254 | 2.78 | 291 | N/A |
| 2/1/2017 | 12 | 250.57 | 256 | 0.79 | 70 | 136 |
| 8/20/2017 | 4 | 246.49 | 251 | 3.17 | 370 | N/A |
| 6/15/2018 | 5 | 246.33 | 252 | 1.35 | 94 | N/A |
| 2/20/2019 | 10 | 246.09 | 243 | 0.67 | 54 | 112 |
| 8/20/2017 | 5 | 244.88 | 252 | 3.1 | 355 | N/A |
| 1/2/2018 | 2 | 243.55 | 251 | 1.55 | 84 | N/A |
| 5/28/2018 | 20 | 242.40 | 257 | 1.45 | 96 | N/A |
| 6/3/2017 | 5 | 240.86 | 252 | 2.86 | 312 | N/A |
| 9/3/2017 | 24 | 240.14 | 251 | 3.41 | 396 | N/A |
| 6/1/2017 | 21 | 240.00 | 251 | 3.42 | 414 | N/A |
| 1/13/2018 | 24 | 238.69 | 251 | 1.54 | 82 | N/A |
| 2/24/2018 | 6 | 236.69 | 250 | 4.17 | 400 | N/A |
| 8/13/2017 | 6 | 236.57 | 249 | 2.79 | 304 | N/A |

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

| Date | Hour | Massena_MOD Modeled SO2 Concentration (ug/m3) | Wind Direction (degrees) | Wind Speed (m/s) | Mechanical Mixing Height (m) | Convective Mixing Height (m) |
|---|---|---|---|---|---|---|
| 9/16/2018 | 20 | 697.62 | 56 | 1.64 | 76 | N/A |
| 6/30/2017 | 2 | 686.23 | 54 | 2.11 | 117 | N/A |
| 8/13/2018 | 22 | 672.41 | 53 | 1.94 | 102 | N/A |
| 9/10/2018 | 3 | 669.98 | 53 | 2.42 | 139 | N/A |
| 5/21/2017 | 2 | 664.32 | 56 | 1.93 | 95 | N/A |
| 2/20/2018 | 5 | 660.54 | 56 | 2.35 | 107 | N/A |
| 2/20/2018 | 19 | 659.22 | 56 | 2.52 | 120 | N/A |
| 6/4/2017 | 23 | 654.55 | 55 | 2.08 | 114 | N/A |
| 9/2/2017 | 22 | 653.06 | 56 | 1.99 | 108 | N/A |
| 6/23/2018 | 23 | 651.94 | 53 | 1.98 | 105 | N/A |
| 2/28/2018 | 20 | 648.73 | 53 | 2.32 | 105 | N/A |
| 6/4/2017 | 22 | 645.94 | 53 | 2.5 | 154 | N/A |
| 5/17/2018 | 22 | 644.65 | 52 | 2.37 | 147 | N/A |
| 7/13/2017 | 24 | 644.60 | 55 | 1.77 | 88 | N/A |
| 2/15/2017 | 8 | 639.88 | 53 | 2.61 | 127 | N/A |
| 2/15/2017 | 6 | 639.50 | 53 | 2.66 | 131 | N/A |
| 12/9/2019 | 16 | 635.19 | 53 | 2.36 | 117 | N/A |
| 5/17/2017 | 5 | 629.99 | 50 | 1.99 | 100 | N/A |
| 4/19/2017 | 2 | 629.90 | 54 | 2.66 | 161 | N/A |
| 5/1/2017 | 24 | 627.79 | 57 | 1.86 | 92 | N/A |
| 2/15/2017 | 4 | 627.42 | 52 | 2.37 | 108 | N/A |
| 9/13/2018 | 19 | 627.13 | 51 | 1.62 | 78 | N/A |
| 9/18/2017 | 19 | 625.66 | 54 | 1.6 | 72 | N/A |
| 12/30/2019 | 23 | 625.43 | 56 | 2.66 | 131 | N/A |
| 5/13/2017 | 4 | 624.26 | 50 | 2.05 | 105 | N/A |

**Figure 4-1: Q-Q Plot of Observed and Modeled Ranked Wind Speeds for Monitor 1**

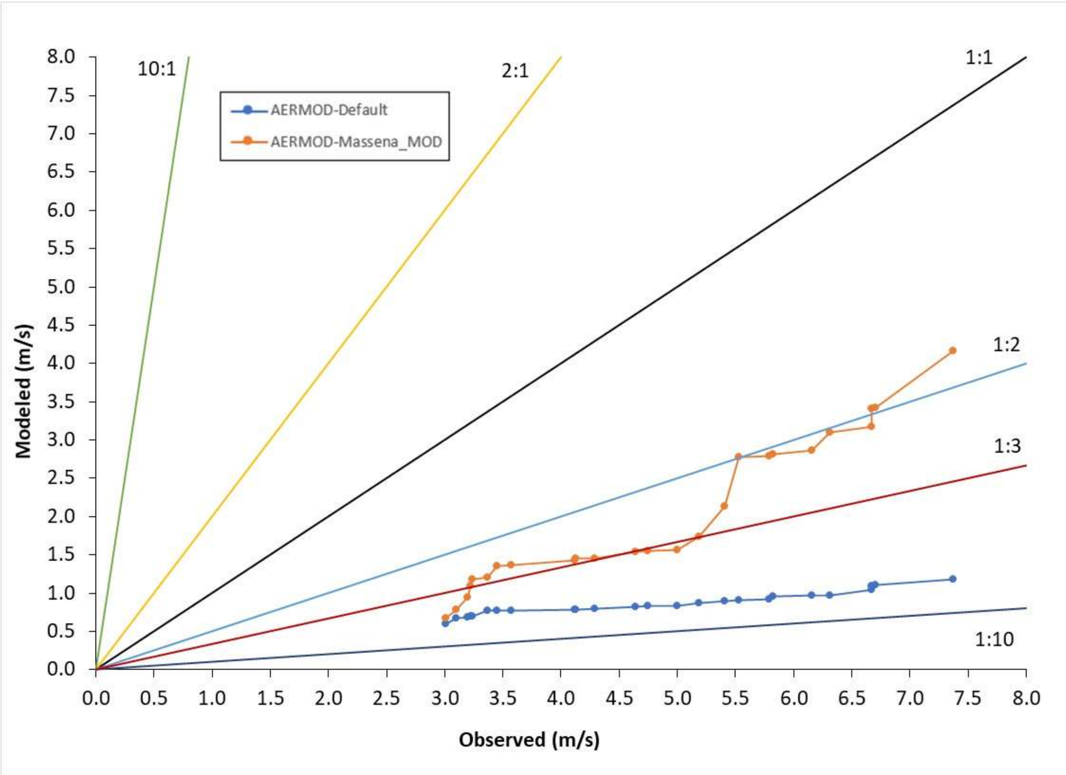



**Figure 4-2: Q-Q Plot of Observed and Modeled Ranked Wind Speeds for Monitor 2**

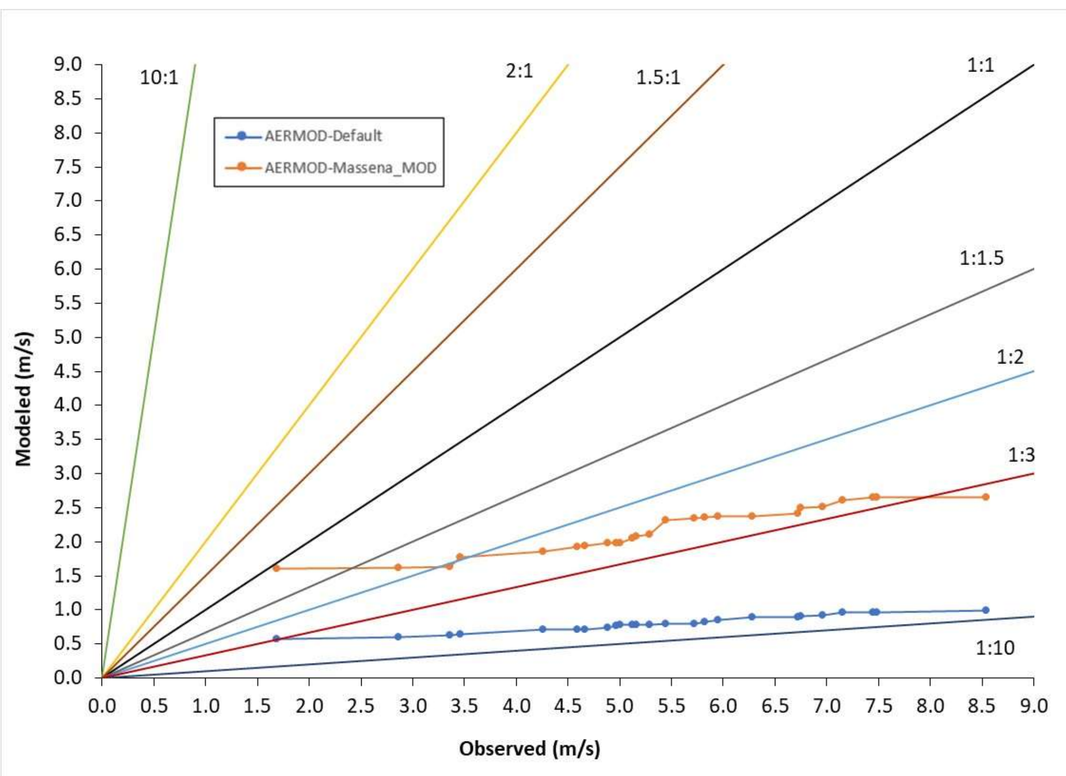


**Figure 4-3: Q-Q Plot of Observed and Modeled Ranked Mixing Heights for Monitor 1**

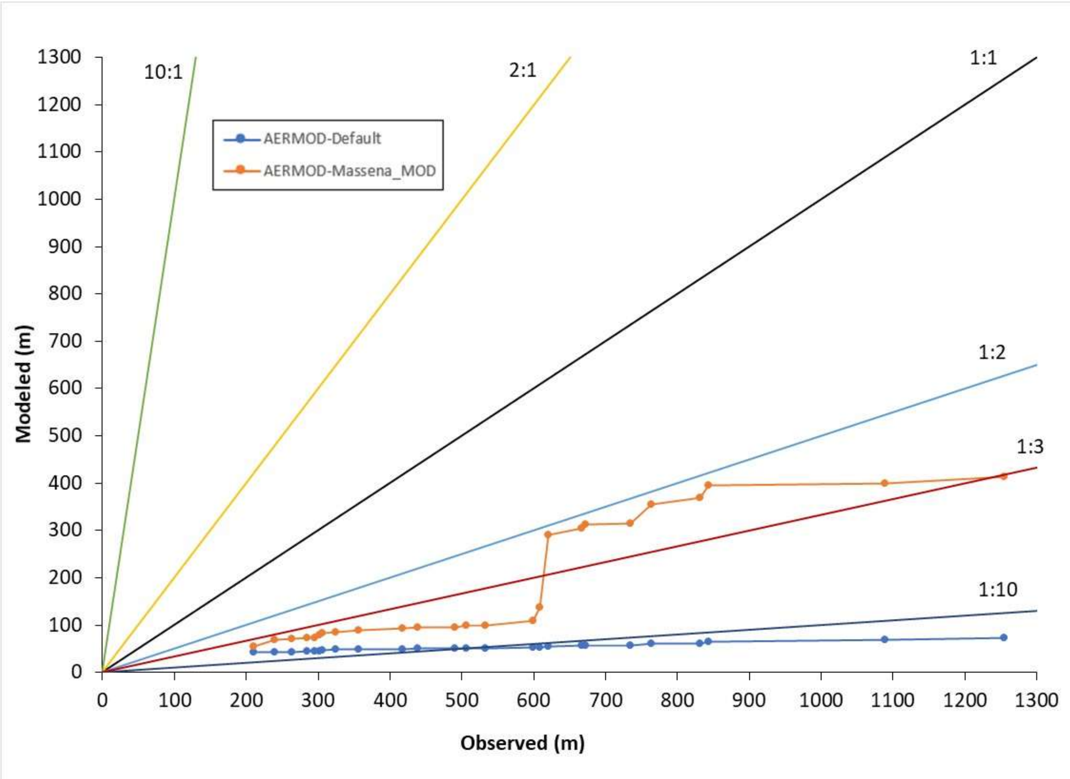
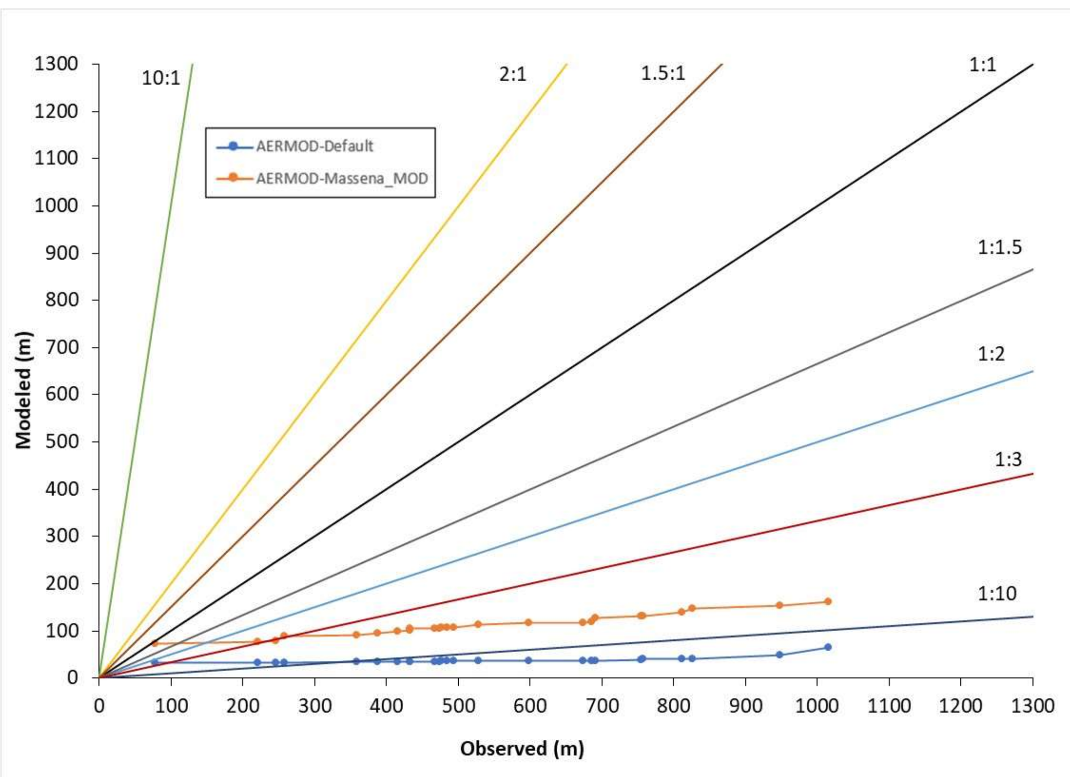



**Figure 4-4: Q-Q Plot of Observed and Modeled Ranked Mixing Heights for Monitor 2**



## 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 3-year average 1-hour average design concentration[14] for each monitor, and c) the use of the Robust Highest Concentration (RHC) as part of EPA's Cox-Tikvart[15] procedure as described in EPA's 1992 model evaluation procedures[6], as implemented in EPA's Model Evaluation Methodology (MEM) software[16]. Further discussion of each of these tests and the results of the testing are provided below.

**Quantile-Quantile Plots**

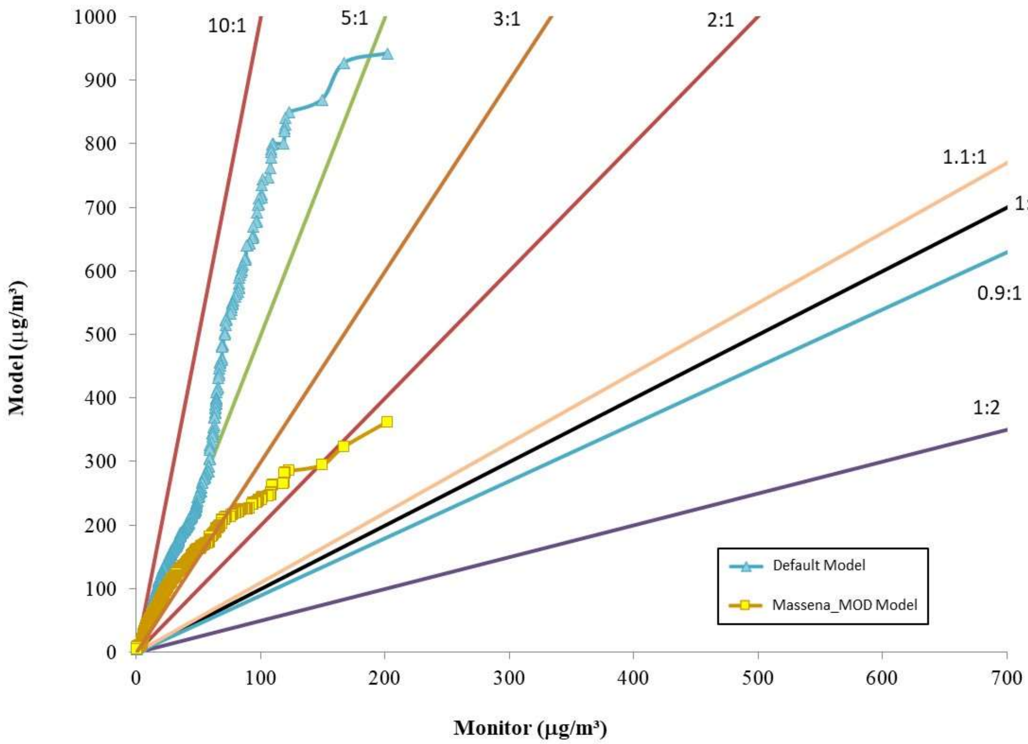
Operational performance of models for predicting compliance with air quality regulations, especially those involving a peak or near peak value at some unspecified time and location, can be assessed with quantile-quantile (Q-Q) plots (Chambers et al., 1983)[17]. Q-Q plots are created by sorting by rank 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)[18] and Liu and Moore, 1984[19]) 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)[20] 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 Massena_MOD 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[21], ranked modeled concentrations that are within 10% of a "perfect model" result are considered to be unbiased. EPA's 1992 model evaluation procedures[22] indicate that an acceptable model should have peak predictions within a factor of 2 of observations. In both Figures 4-5 and 4-6, 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. At Site 2, the Massena_MOD elevated concentrations are within a factor of 3 of the observations with the peak elevated concentrations within a factor of 2.7 of the observations. Even though the elevated concentrations at Site 2 for Massena_MOD are slightly more than a factor of 2 of the observed concentrations, the overprediction tendency is much lower than that of the default model. Based on these Q-Q plot results, the performance of Massena_MOD is clearly better than that of the default model.

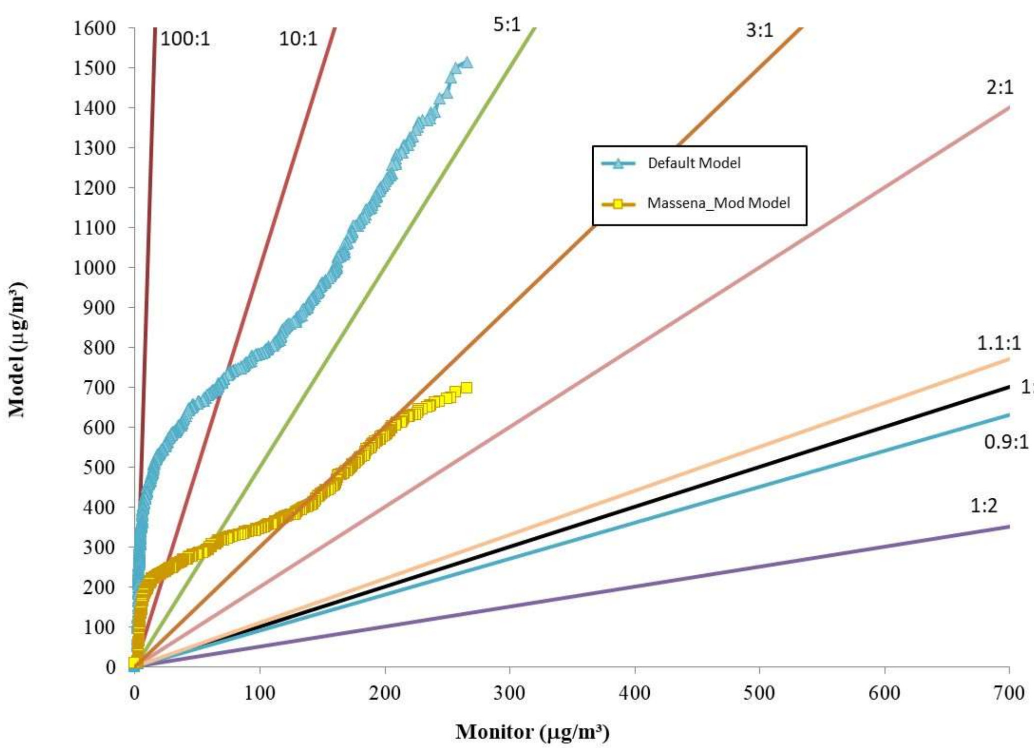
---

[14] The "design concentration" is the 99th percentile of the peak daily 1-hour maximum concentration computed for each calendar year, and averaged over the three 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.
[15] 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.
[16] 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.
[17] 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.
[18] Weil J.C, Sykes and Venkatram A., 1992. Evaluating air-quality models: Review and outlook. *J. Appl. Met*., 31, p 1121-1144.
[19] 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.
[20] 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.
[21] 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).
[22] U.S. Environmental Protection Agency, 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). model_eval_protocol.pdf (epa.gov).

## 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 three-year average of the 99$^{th}$ percentile peak daily 1-hour maximum concentration. 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 over 6 at Site 1 and over 5 at Site 2. In contrast, the Massena_MOD model has a MOR value between 2 and 3 at both monitoring sites. As a result, Massena_MOD is determined to have a better performance, although still overpredicting.

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

| Model Option | 4th Highest Design Concentration µg/m³ | | | | |
|---|---|---|---|---|---|
| | 2017 | 2018 | 2019 | 3-yr Ave | |
| Observed | 118.69 | 106.61 | 110.20 | 111.83 | Pre/obs |
| Default Model | 829.74 | 800.66 | 655.64 | 762.01 | 6.81 |
| Massena_MOD Model | 264.73 | 266.58 | 215.88 | 249.06 | 2.23 |

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

| Model Option | 4th Highest Design Concentration µg/m³ | | | | |
|---|---|---|---|---|---|
| | 2017 | 2018 | 2019 | 3-yr Ave | |
| Observed | 237.14 | 226.71 | 214.89 | 226.25 | Pre/obs |
| Default Model | 1372.59 | 1388.18 | 1103.08 | 1287.95 | 5.69 |
| Massena_MOD Model | 653.06 | 660.54 | 589.46 | 634.36 | 2.80 |

**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 two, while a fractional bias of +0.67 is equivalent to an underprediction by a factor of two. 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. Three 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 (CPM) 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).

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

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

$$MCM_{i,j} = CPM_i - CPM_j$$ (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[23].

The Model Evaluation Methodology 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 both model runs at both monitors for the RHC estimate. The resulting overprediction ratio for the default model is about 7 for Monitor 1 and about 6 for Monitor 2. For Massena_MOD, the overprediction ratio is much lower, between 2 and 3 for both monitors.

[23] See, for example, the EPA presentation 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 4-9: Average Fractional Biases for Monitors 1 and 2 for RHC Estimate**

| Site | Case | $FB_{Avg}$ |
|---|---|---|
| Monitor 1 | Default Model | -1.47 |
| | MASSENA_MOD | -0.74 |
| Monitor 2 | Default Model | -1.41 |
| | MASSENA_MOD | -0.93 |

Tables 4-10 and 4-11 show the Robust High Concentrations for monitor 1 and monitor 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 4-10 and 4-11, the default model runs have RHC predicted-to-observed ratios between 5 and 8, while the Massena_MOD RHC ratios range between 1.9 and 2.8. As expected, the 90% confidence intervals for the CPM between the two models (see Figures 4-7 and 4-8) do not overlap by a wide margin, meaning that there exists a statistically significant difference between the performance of the two models.

**Table 4-10: 3-Year Averaged Robust High Concentrations (μg/m³) for Monitor 1**

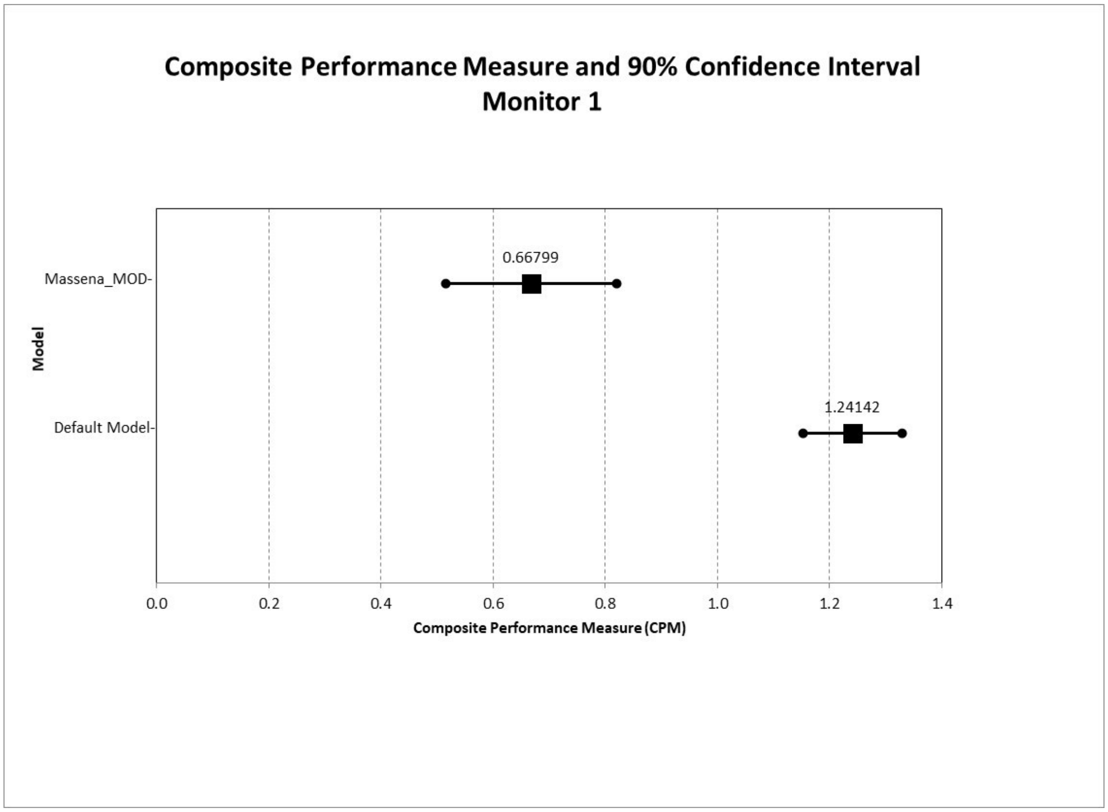
| *Model Option* | *RHC* | Pre/Obs Ratio |
|---|---|---|
| Observed | 149.92 | - |
| Default Model | 1066.46 | 7.11 |
| MASSENA_MOD | 298.44 | 1.99 |

**Table 4-11: 3-Year Averaged Robust High Concentrations (μg/m³) for Monitor 2**

| *Model Option* | *RHC* | Pre/Obs Ratio |
|---|---|---|
| Observed | 268.47 | - |
| Default Model | 1463.09 | 5.45 |
| MASSENA_MOD | 750.09 | 2.79 |

Figures 4-7 and 4-8 show the CPM values for monitors 1 and 2, respectively, and Figure 4-9 shows the MCM values for both monitors separately and the combined (CMCM). As can be seen in Figures 4-7 and 4-8, the CPM values for the Massena_MOD approach are smaller than the CPM values for the default model approach. This is reflected in the MCM values, which are all positive, meaning that model 2 (Massena_MOD) performs much better than model 1 (default model).

### Figure 4-7: Plot of CPM for Monitor 1




### Figure 4-8: Plot of CPM for Monitor 2

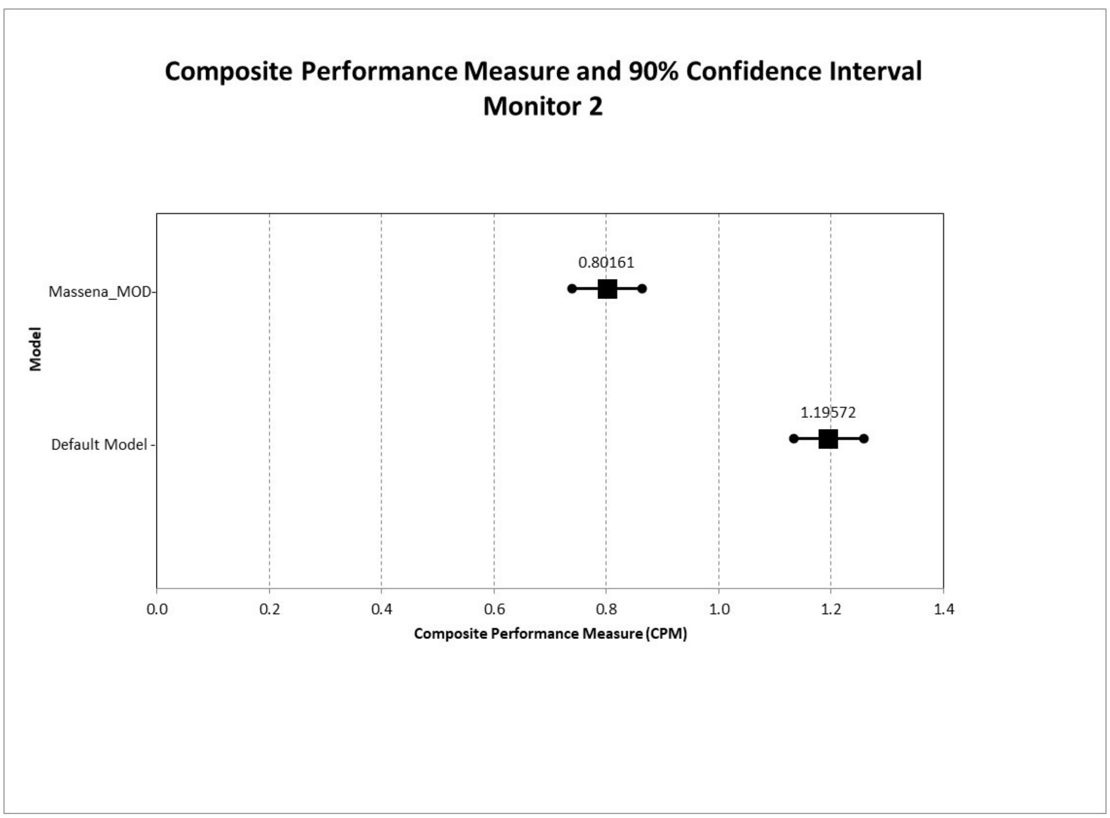
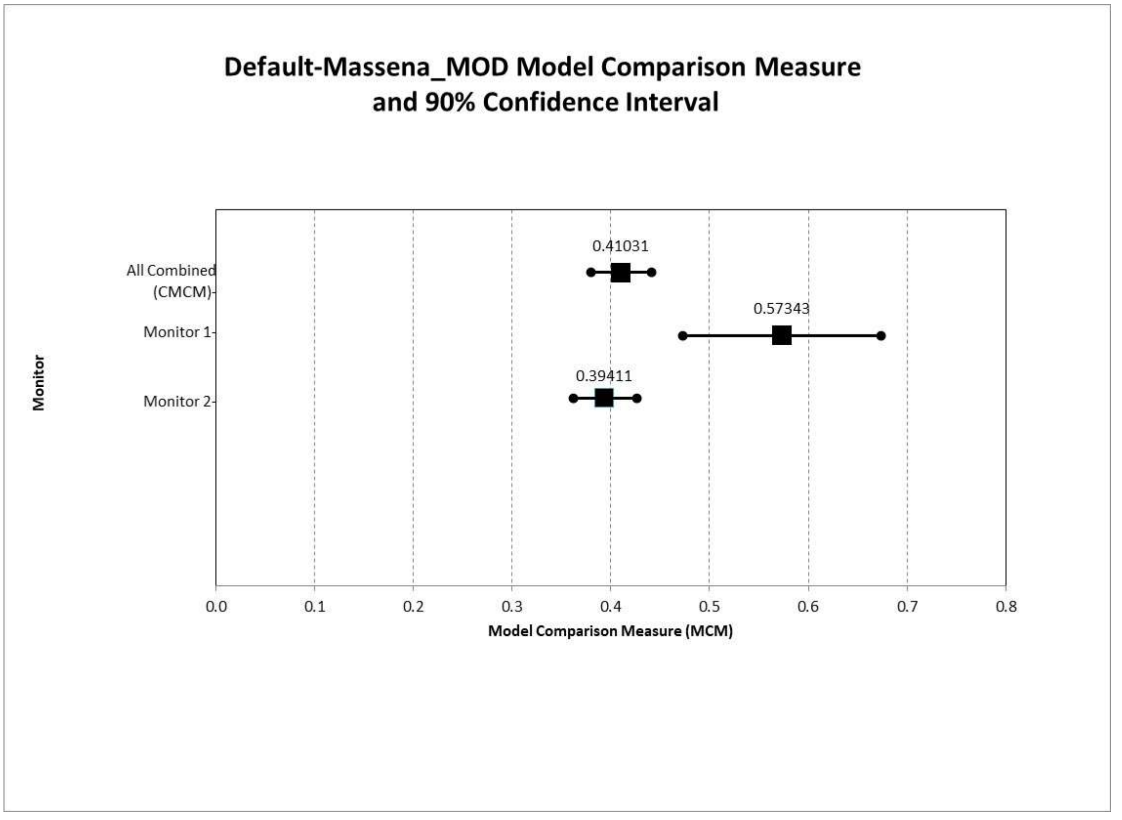


**Figure 4-9: Plot of MCM and CMCM and 90% Confidence Interval for Monitors 1 and 2**



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

This protocol presents the scientific justification for the proposed alternative model for Alcoa's Massena Operations (Massena_MOD). The results of the model evaluation elements described above indicate that AERMOD-Default grossly overpredicts at the two monitors. The Massena_MOD approach overpredicts much 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, it still overpredicts, and it should be considered as an acceptable model for use at the Alcoa Massena smelter.

Alcoa will not be changing its method of operation with the facility changes, but groups of dry scrubber stacks along the potlines will be merged and raised. The merging is creditable because the plantwide $SO_2$ emission limit will be below the 5000 tons per year referenced in the 1985 Stack Height Regulations for this modeling credit. The final stack heights will be well below the default Good Engineering Practice (GEP) height of 65 m. A separate modeling report will be submitted using the Massena_MOD modeling procedures to demonstrate future $SO_2$ NAAQS compliance using the Massena_MOD model with five years of Massena airport data (2017-2021).

aecom.com

aecom.com