United States Environmental Protection Agency

Evaluation of Prognostic Meteorological Data in AERMOD Applications

EPA-454/R-16-004
December 2016

Evaluation of Prognostic Meteorological Data in AERMOD Applications

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Air Quality Assessment Division
Air Quality Modeling Group
Research Triangle Park, North Carolina

Preface

This document provides an evaluation of prognostic meteorological data in AERMOD. Included in this document are descriptions of the inputs, evaluation of the meteorological data, and evaluation of AERMOD results using observed meteorological data and prognostic meteorological data. This document is an update to the July 2015 document "Evaluation of Prognostic Meteorological data in AERMOD Applications" (EPA-454/R-15-004). This document includes corrections to typographical errors and includes a new Appendix C, which are further evaluations of MMIF output with emphasis on comparisons of grid resolution, mixing height methodology, and number of output vertical levels. EPA will respond to specific requests for subsets of the data or for specific additional inputs and outputs of the modeling process, depending on the availability of the data. Requests for electronic copies of the air quality modeling data used for this rule should be sent to James Thurman (Thurman.james@epa.gov).

Acknowledgements

This report was developed as part of the 2015 proposal of The Guideline on Air Quality Models, Appendix W with input from the meteorological data workgroup comprised of staff from EPA's Office of Air Quality Planning and Standards and Regions 5, 7, and 8. WRF and MMIF processing for the Gibson, IN evaluation were processed by Computer Science Corporation. WRF and MMIF processing for Martins Creek, PA and Herculaneum, MO were performed by Andy Hawkins of EPA's Region 7. Evaluations discussed in Appendix B were performed by Rebecca Matichuk of EPA's Region 8. The workgroup acknowledges Kali Frost of the Indiana Department of Environmental Management for providing AERMOD inputs for Gibson, IN.

Contents

Preface ii
Acknowledgements iii
Figures vi
Tables xi
1. Introduction 1
2. Methodology 2
2.1 Study areas 2
2.1.1 Gibson 3
2.1.2 Martins Creek 7
2.1.3 Herculaneum 10
2.2 Meteorological data evaluation 14
2.3 Model evaluation methodology 15
3. Results 17
3.1 Gibson 17
3.1.1 Meteorological data comparisons 17
3.1.2 AERMOD results 25
3.2 Martins Creek 33
3.2.1 Meteorological data comparisons 33
3.2.2 AERMOD results 40
3.3 Herculaneum 48
3.3.1 Meteorological data comparisons 48
3.3.2 AERMOD results 58
4. Summary and Conclusions 59
5. References 61
Appendix A. Meteorological data comparisons A-1
A.1 Gibson A-1
A.2 Martins Creek A-17
A.3 Herculaneum A-33
Appendix B. NWS/AERMET and WRF/MMIF analyses for Region 8 sites B-1
B.1 Introduction B-1

B.2 Methodology B-1
B.3 Results/Summary B-5
B.3.1 Temperature B-5
B.3.2 Relative humidity B-6
B.3.3 Wind speed B-7
B.3.4 Wind displacement B-8
B.3.5 Mechanical mixing heights B-9
B.3.6 Convective mixing heights B-10
B.3.7 Surface friction velocity B-11
B.3.8 Convective velocity scale B-12
Appendix C: Evaluations for Baldwin, IL and Martins Creek C-1
C.1 Introduction C-1
C.2 Methodology C-3
C.2.1 Observed meteorological data processing C-3
C.2.2 WRF simulations C-5
C.2.4 Meteorological data evaluation C-8
C.2.5 AERMOD simulations C-9
C.3. Results C-9
C.3.1 Meteorological comparisons C-9
C.3.2 AERMOD results C-20
C.4 Summary and conclusions C-28

Figures

Figure 1. Locations of MMIF evaluation study areas. 3
Figure 2. Location of Gibson facility, SO_2 monitors, and meteorological tower for Gibson. 4
Figure 3. Locations of Gibson meteorological tower (green dot), Evansville (EVV) NWS station (airplane) and 12 km WRF grid cells containing those stations. Panel a is a statewide view and panel b is a closer view of southwest Indiana. 5
Figure 4. Location of Martins Creek study area with meteorological sites, monitors, and emission sources. 7
Figure 5. Detailed view of Martins Creek meteorological sites, monitors, and emissions sources with detailed terrain. 8
Figure 6. Location of the Doe Run lead facility, NWS station (CPS), site-specific tower (Herculaneum), three MMIF output locations (MMIF 4 KM, MMIF 12 KM, and MMIF 36 KM), and lead monitors. 11
Figure 7. As for Figure 6, with more detail around the facility. 12
Figure 8. Detailed view of MMIF grid cell locations, Herculaneum site specific tower, monitors, and Doe Run. 13
Figure 9. Gibson 2010 wind roses for a) GIB OBS, b) GIB MMIF, c) EVV OBS, and d) EVV MMIF. 18
Figure 10. Wind displacement (km) among the Gibson meteorological scenarios. 19
Figure 11. Gibson study monthly surface roughness lengths (m) by 10 degree sectors for a) January, b) February, c) March, and d) April. 20
Figure 12. Gibson study monthly surface roughness lengths (m) by 10 degree sectors for a) May, b) June, c) July, and d) August. 21
Figure 13. Gibson study monthly surface roughness lengths (m) by 10 degree sectors for a) September, b) October, c) November, and d) December. 22
Figure 14. Gibson hourly QQ plots. Concentrations are in $\mu g/m^3$. 26
Figure 15. Gibson 3-hour screening results. 27
Figure 16. Gibson 24-hour screening results. 28
Figure 17. Gibson fractional biases for a) 1-hour, b) 3-hour, c) 24-hour, and d) CPM based on fractional biases. 29
Figure 18. Gibson composite performance metric values with 5th and 95th percentiles of the CPM values from the bootstrap results. 30
Figure 19. Gibson MCM differences with a) 90th percentile and b) 95th confidence intervals. .. 32
Figure 20. May 1992 – May 1993 Martins Creek wind roses for a) Martins Creek, b) ABE, c) MMIF 1 km, and d) MMIF 4 km. 33
Figure 21. Wind displacement (km) among the Martins Creek meteorological scenarios. 34
Figure 22. Martins Creek study monthly surface roughness lengths (m) by 10 degree sectors for a) January, b) February, c) March, and d) April. 35
Figure 23. Martins Creek study monthly surface roughness lengths (m) by 10 degree sectors for a) May, b) June, c) July, and d) August. 36

Figure 24. Martins Creek study monthly surface roughness lengths (m) by 10 degree sectors for a) September, b) October, c) November, and d) December. 37
Figure 25. Martins Creek QQ plots for a) 1-hour, b) 3-hour, c) 24-hour, and d) annual averages. Concentrations are in μg/m^3. 41
Figure 26. Martins Creek 3-hour screening results. 42
Figure 27. Martins Creek 24-hour screening results. 43
Figure 28. Martins Creek fractional biases for a) 1-hour, b) 3-hour, c) 24-hour, and d) CPM based on fractional biases. 44
Figure 29. Martins Creek composite performance metric values with 5th and 95th percentiles of the CPM values from the bootstrap results. 46
Figure 30. Martins Creek MCM differences with a) 90th percentile and b) 95th confidence intervals. 47
Figure 31. 2009 wind roses for a) Herculaneum and b) CPS. 48
Figure 32. 2009 wind roses for a) MMIF 4 km, b) MMIF 12 km, and c) MMIF 36 km. 49
Figure 33. Wind displacement (km) among the Herculaneum meteorological scenarios. 50
Figure 34. Herculaneum study monthly surface roughness lengths (m) by 10 degree sectors for a) January, b) February, c) March, and d) April. 51
Figure 35. Herculaneum study monthly surface roughness lengths (m) by 10 degree sectors for a) May, b) June, c) July, and d) August. 52
Figure 36. Herculaneum study monthly surface roughness lengths (m) by 10 degree sectors for a) September, b) October, c) November, and d) December. 53
Figure 37. Herculaneum 24-hour QQ plots. Concentrations are in μg/m^3. 58
Figure 38. Herculaneum 24-hour screening results. 59
Figure A-1. Gibson wind speed (m/s): a) annual distributions and b) bias distributions. A-3
Figure A-2. Gibson ambient temperature (K): a) annual distributions and b) bias distributions. A-4
Figure A-3. Gibson station pressure (mb): a) annual distributions and b) bias distributions. A-5
Figure A-4. Gibson relative humidity (percent): a) annual distributions and b) bias distributions. A-6
Figure A-5. Gibson daytime albedo (fraction): a) annual distributions and b) bias distributions. A-7
Figure A-6. Gibson Bowen ratio: a) annual distributions and b) bias distributions. A-8
Figure A-7. Gibson heat flux (W/m^2): a) annual distributions and b) bias distributions. A-9
Figure A-8. Gibson surface friction velocity, u* (m/s): a) annual distributions and b) bias distributions. A-10
Figure A-9. Gibson convective velocity scale, w* (m/s): a) annual distributions and b) bias distributions. A-11
Figure A-10. Gibson Monin-Obukhov length (m): a) annual distributions and b) bias distributions. A-12
Figure A-11. Gibson convective mixing height (m): a) annual distributions and b) bias distributions. A-13
Figure A-12. Gibson mechanical mixing height (m): a) annual distributions and b) bias distributions. A-14

Figure A-13. Gibson potential temperature gradient (K/m) above Zic: a) annual distributions and b) bias distributions. A-15
Figure A-14. Gibson cloud cover (tenths*10): a) annual distributions and b) bias distributions. A-16
Figure A-15. Martins Creek wind speed (m/s): a) annual distributions and b) bias distributions. A-19
Figure A-16. Martins Creek ambient temperature (K): a) annual distributions and b) bias distributions A-20
Figure A-17. Martins Creek station pressure (mb): a) annual distributions and b) bias distributions A-21
Figure A-18. Martins Creek relative humidity (percent): a) annual distributions and b) bias distributions A-22
Figure A-19. Martins Creek daytime albedo (fraction): a) annual distributions and b) bias distributions A-23
Figure A-20. Martins Creek Bowen ratio: a) annual distributions and b) bias distributions... A-24
Figure A-21. Martins Creek heat flux (W/m^2): a) annual distributions and b) bias distributions. A-25
Figure A-22. Martins Creek surface friction velocity, u* (m/s): a) annual distributions and b) bias distributions. A-26
Figure A-23. Martins Creek convective velocity scale, w* (m/s): a) annual distributions and b) bias distributions A-27
Figure A-24. Martins Creek Monin-Obukhov length (m): a) annual distributions and b) bias distributions A-28
Figure A-25. Martins Creek convective mixing height (m): a) annual distributions and b) bias distributions A-29
Figure A-26. Martins Creek mechanical mixing height (m): a) annual distributions and b) bias distributions A-30
Figure A-27. Martins Creek potential temperature gradient (K/m) above Zic: a) annual distributions and b) bias distributions. A-31
Figure A-28. Martins Creek cloud cover (tenths*10): a) annual distributions and b) bias distributions A-32
Figure 29. Herculaneum wind speed (m/s): a) annual distributions and b) bias distributions. A-35
Figure A-30. Herculaneum ambient temperature (K): a) annual distributions and b) bias distributions A-36
Figure A-31. Herculaneum station pressure (mb): a) annual distributions and b) bias distributions A-37
Figure A-32. Herculaneum relative humidity (percent): a) annual distributions and b) bias distributions A-38
Figure A-33. Herculaneum daytime albedo (fraction): a) annual distributions and b) bias distributions A-39
Figure A-34. Herculaneum Bowen ratio: a) annual distributions and b) bias distributions. ... A-40
Figure A-35. Herculaneum heat flux (W/m^2): a) annual distributions and b) bias distributions. A-41

Figure A-36. Herculaneum surface friction velocity, u* (m/s): a) annual distributions and b) bias distributions..... A-42
Figure A-37. Herculaneum convective velocity scale, w* (m/s): a) annual distributions and b) bias distributions. A-43
Figure A-38. Herculaneum Monin-Obukhov length (m): a) annual distributions and b) bias distributions..... A-44
Figure A-39. Herculaneum convective mixing height (m): a) annual distributions and b) bias distributions..... A-45
Figure A-40. Herculaneum mechanical mixing height (m): a) annual distributions and b) bias distributions..... A-46
Figure A-41. Herculaneum potential temperature gradient (K/m) above Zic: a) annual distributions and b) bias distributions. A-47
Figure A-42. Herculaneum cloud cover (tenths*10): a) annual distributions and b) bias distributions..... A-48
Figure B-1. Map of sites analyzed for study. B-2
Figure C-1 Baldwin: a) WRF modeling domains; b) locations of emissions, monitors, and meteorological data; c) Baldwin site-specific tower 1992 land use and sectors; and d) BLV 1992 land use and sectors. C-1
Figure C-2 Martins Creek: a) WRF modeling domains; b) locations of emissions, monitors, and meteorological data; c) Baldwin site-specific tower 1992 land use and sectors; and d) BLV 1992 land use and sectors. C-2
Figure C-3. BAL_OBS, BLV_OBS, and MMIF surface wind roses for Baldwin. C-10
Figure C-4. Wind metrics of BLV_OBS and MMIF scenarios compared to BAL_OBS: a) D (km); b) Ds (km); and c) wind direction differences (degrees). C-12
Figure C-5.Wind metrics of MMIF scenarios compared to BLV_OBS: a) D (km); b) Ds (km); and c) wind direction differences (degrees). C-13
Figure C-6. MC_OBS, ABE_OBS, and MMIF surface wind roses for Martins Creek. C-14
Figure C-7. Wind metrics of ABE_OBS and MMIF scenarios compared to MC_OBS: a) D (km); Ds (km); and c) wind direction differences (degrees). C-16
Figure C-8. Wind metrics of MMIF scenarios compared to ABE_OBS: a) D (km); b) Ds (km); and c) wind direction differences (degrees). C-17
Figure C-9. a) Mean fractional biases for selected meteorological variables of MMIF compared to BAL_OBS; b) mean fractional biases for selected meteorological variables of MMIF compared to BLV_OBS; and c) mean fractional biases for mixing heights between PASS and RECALC MMIF output. C-19
Figure C-10. a) Mean fractional biases for selected meteorological variables of MMIF compared to MC_OBS; b) mean fractional biases for selected meteorological variables of MMIF compared to ABE_OBS; and c) mean fractional biases for mixing heights between PASS and RECALC MMIF output. C-20
Figure C-11.CPM results for a) Baldwin and b) Martins Creek. C-21
Figure C-12. Baldwin selected MCM pairings to determine best performing a) grid resolution; b) PBL method; and c) number of levels. C-22

Figure C-13. Martins Creek selected MCM pairings to determine best performing a) grid resolution; b) PBL method; and c) number of levels. C-23

Tables

Table 1. Gibson emissions 4
Table 2. Martins Creek study facility emissions (tons/year) and stack parameters. 9
Table 3. Herculaneum point source emission rates and stack parameters. 14
Table 4. Mean bias, fractional bias, root mean square error, and R^2 for primary meteorological variables. 24
Table 5. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables. 24
Table 6. Mean bias, fractional bias, root mean square error, correlation, and index of agreement for calculated meteorological variables. 25
Table 7. 1-hour, 3-hour, 24-hour absolute fractional biases and composite performance measures for Gibson meteorological scenarios. 30
Table 8. Model comparison measures (MCM) for the four Gibson meteorological scenarios. .. 31
Table 9. Mean bias, fractional bias, root mean square error, and R^2 for primary meteorological variables. 39
Table 10. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables. 39
Table 11. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables. 40
Table 12. 1-hour, 3-hour, 24-hour absolute fractional biases and composite performance measures for Martins Creek meteorological scenarios. 45
Table 13. Model comparison measures (MCM) for the four Gibson meteorological scenarios. 46
Table 14. Mean bias, fractional bias, root mean square error, and R^2 for primary meteorological variables. 55
Table 15. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables. 56
Table 16. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables. 57
Table B-1. Meteorological sites analyzed for study. B-2
Table B-2. Description of NWS/AERMET model cases. B-3
Table B-3. Description of WRF/MMIF model cases. B-4
Table B-4. Monthly averaged temperature (K) across all modeled cases at each site. B-6
Table B-5. Monthly averaged relative humidity across all modeled cases at each site. B-7
Table B-6. Monthly averaged wind speed (m/s) across all modeled cases at each site. B-8
Table B-7. Monthly averaged wind displacement (km) across all modeled cases at each site. . B-9
Table B-8. Monthly averaged mechanical mixing heights (m) across all modeled cases at each site. B-10
Table B-9. Monthly averaged mechanical mixing heights (m) across all modeled cases at each site. B-11

Table B-10. Monthly averaged surface friction velocity (m/s) across all modeled cases at each site. B-12
Table B-11. Monthly averaged convective velocity scale (m/s) across all modeled cases at each site. B-13
Table C-1. BAL_OBS albedo (A), Bowen ratio (Bo) and surface roughness (z_o) by month and sector. C-3
Table C-2. BLV_OBS albedo, Bowen ratio, and surface roughness (zo) by season and sector…C-4
Table C-3. MC_OBS albedo, Bowen ratio, and surface roughness (zo) by season and sector. C-4
Table C-4. ABE_OBS albedo, Bowen ratio, and surface roughness (zo) by season and sector… C-5
Table C-5. MMIF derived albedo (A), Bowen ratio (Bo), and surface roughness (zo) for 1, 4, and 12 km output for Baldwin. C-7
Table C-6. MMIF derived albedo (A), Bowen ratio (Bo), and surface roughness (zo) for 1, 4, and 12 km output for Martins Creek. C-8
Table C-7. Baldwin fractional biases for convective and mechanical mixing heights comparing AERMET generated mixing heights using MMIF output to mixing height from observed data and MMIF provided mixing heights (PASS or RECALC). C-25
Table C-8. Martins Creek fractional biases for convective and mechanical mixing heights comparing AERMET generated mixing heights using MMIF output to mixing height from observed data and MMIF provided mixing heights (PASS or RECALC). C-26
Table C-9. Baldwin MCM and test statistics for MMIF mixing heights – AERMET mixing heights. AERMET mixing height CPM values are subtracted from the listed scenarios. C-27
Table C-10. Martins Creek MCM and test statistics for MMIF mixing heights – AERMET mixing heights. AERMET mixing height CPM values are subtracted from the listed scenarios. C-28

1. Introduction

An important part of dispersion modeling applications is the selection of meteorological data input into the model. The meteorological data input into models such as AERMOD should be adequately representative of conditions of the modeling, as discussed in Section 8.4 of the proposed version of EPA's *Guideline on Air Quality Models* (U.S. EPA 2016a). Specifically, Section 8.4.b states:

> "The meteorological data used as input to a dispersion model should be selected on the basis of spatial and climatological (temporal) representativeness as well as the ability of the individual parameters selected to characterize the transport and dispersion conditions in the area of concern. The representativeness of the measured data is dependent on numerous factors including but not limited to: (1) The proximity of the meteorological monitoring site to the area under consideration; (2) the complexity of the terrain; (3) the exposure of the meteorological monitoring site; and (4) the period of time during which data are collected. The spatial representativeness of the data can be adversely affected by large distances between the source and receptors of interest and the complex topographic characteristics of the area. Temporal representativeness is a function of the year-to-year variations in weather conditions. Where appropriate, data representativeness should be viewed in terms of the appropriateness of the data for constructing realistic boundary layer profiles ..."

Meteorological data often comes from National Weather Service (NWS) or site-specific meteorological monitoring programs. In recent years, interest has grown in the use of prognostic meteorological data, such as the Weather Research and Forecasting (WRF) model to create inputs for dispersion modeling with AERMOD. This is especially true in locations where it can be difficult to find an adequately representative NWS station or cost-prohibitive or infeasible to set up a site-specific meteorological monitoring tower. As part of the 2016 update to Appendix W, EPA is allowing the use of prognostic meteorological data for use in AERMOD for areas where it is cost-prohibitive or not feasible to collect site-specific data and there is no representative NWS or comparable station nearby. EPA has developed the Mesoscale Model Interface Program, or MMIF, for processing prognostic meteorological data for AERMOD (Environ, 2014). For more information on the use of prognostic data in AERMOD see Section 8.4.5 of the proposed version of Appendix W.

The purpose of the prognostic evaluation process was to determine if prognostic meteorological data are suitable for input into AERMOD. The goal of the evaluation was to determine if such prognostic data was comparable to NWS or site-specific meteorological data and, how AERMOD performance to observed concentrations compares based on prognostic meteorological data with simulations based on NWS data or site-specific data.

This report details the evaluation process used to determine the feasibility of prognostic meteorological data for use in AERMOD. Section 2 discusses the methodology of the three case studies: 1) Gibson, IN; 2) Martin's Creek, PA; and 3) Herculaneum, MO. This includes

meteorological data processing (i.e., for both observed and prognostic modeling) and development of AERMOD inputs. Section 3 is a review of the evaluation results, including meteorological data evaluation and AERMOD concentration results for the three locations, while Section 4 is a summary of conclusions and Section 5 provides the references. Appendix A gives more results on the meteorological evaluations of the three case studies and Appendix B is an evaluation of observed and modeled meteorological data for six sites in EPA's Region 8.

2. Methodology

2.1 Study areas

Three case study areas were considered for evaluation (Figure 1). Gibson (Frost, 2014) and Martins Creek were both SO_2 releases while Herculaneum is a lead release. The Gibson modeling presented in this document are based on data presented in Frost (2014). Martins Creek is one of the databases used to evaluate AERMOD during its development (U.S. EPA, 2003; Perry et al., 2005). All three case study area evaluations include site specific meteorological data, a representative NWS station, and multiple prognostic model grid cells. More details about each evaluation are below.

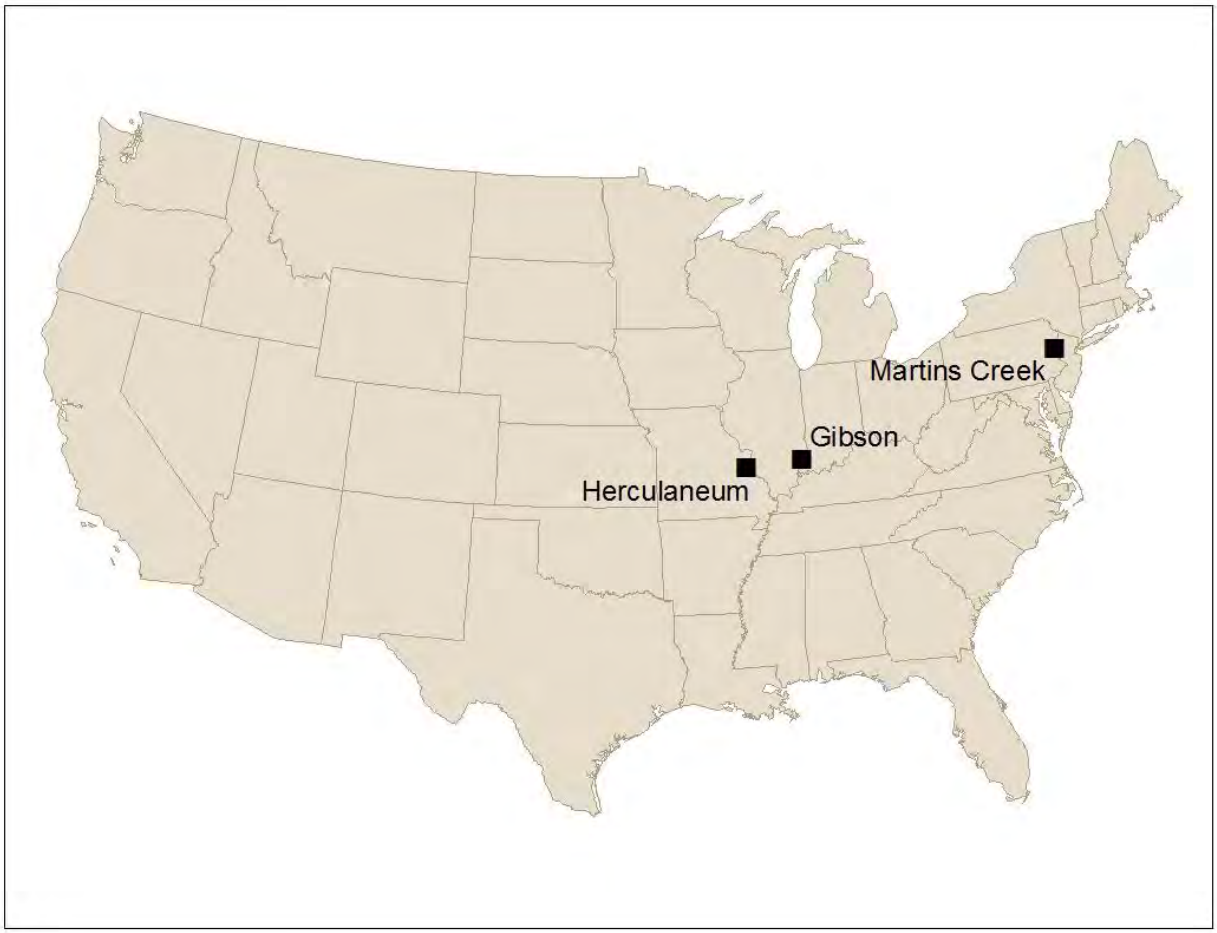

Figure 1. Locations of MMIF evaluation study areas.

2.1.1 Gibson

As shown in Figure 1, the Gibson facility is located in southwest Indiana. The facility is comprised of five units with a 3145 MW capacity (Frost, 2014). Table 1 lists the five units with stack parameters, average annual hourly SO_2 emissions and maximum hourly SO_2 emissions for 2010 as described in Frost (2014).

Table 1. Gibson emissions

Unit	Stack height (m)	Stack diameter (m)	Average annual hourly emissions (g/s)	Maximum hourly emissions (g/s)	Exit temperature (K)	Average stack exit velocity (m/s)
1 and 2	189.0	23.5	116.0	462.2	327.0	15.6
3	189.0	25.0	90.9	794.1	327.0	18.5
4	152.4	25.0	111.3	460.3	322.0	17.4
5	152.4	25.0	290.0	819.8	328.0	12.5

Figure 2 shows the AERMOD study area with the location of stacks, four monitors, and the site-specific meteorological tower.

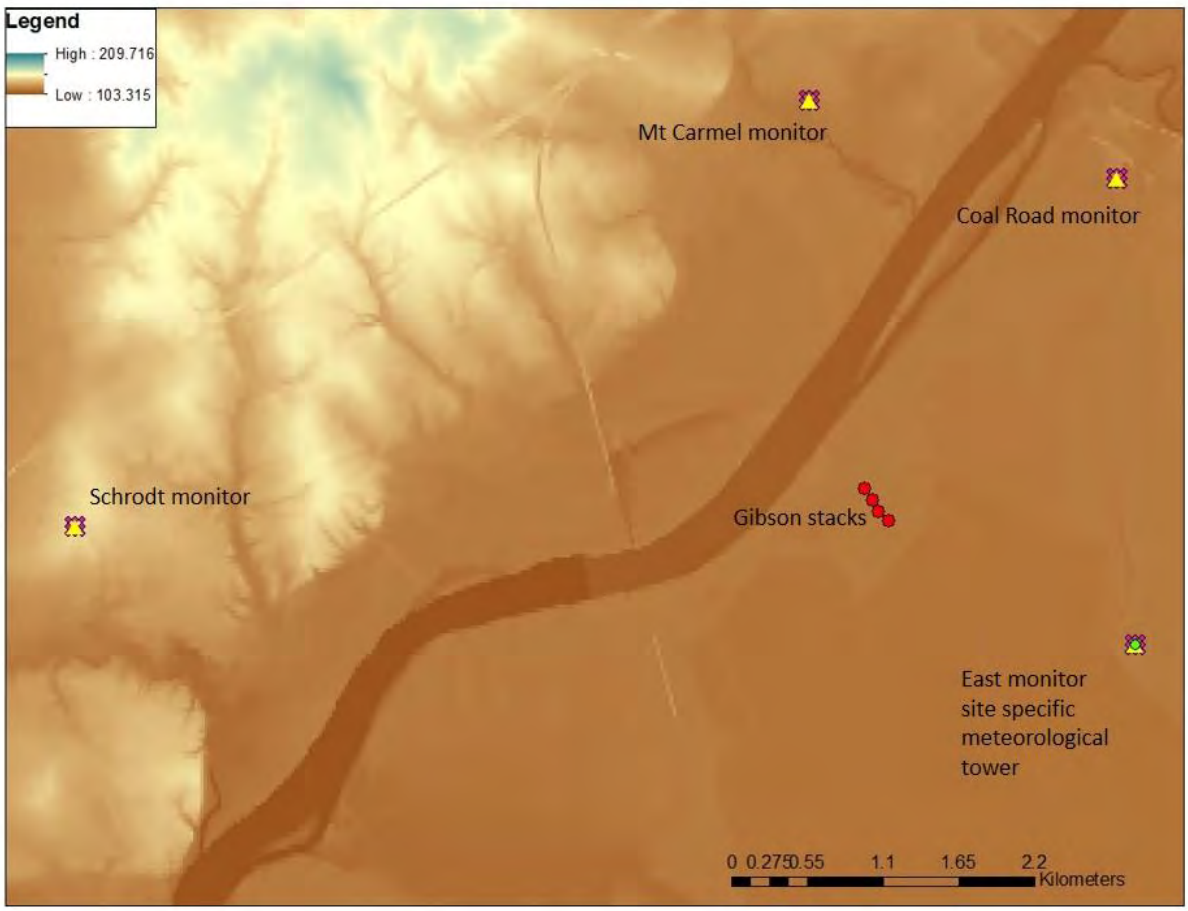

Figure 2. Location of Gibson facility, SO_2 monitors, and meteorological tower for Gibson.

The site-specific tower included multi-level measurements of winds, temperature, pressure, solar radiation, and standard deviation of the horizontal wind direction (σ_Θ) at 10, 25, and 60 m. More

details can be found in Frost (2014). Data from the National Weather Service (NWS) station at Evansville, IN (EVV) were used for substitution of missing hours. Upper air data from Lincoln, IL (ILX) were used for the morning soundings. Figure 3 shows the relationship between the site-specific tower and EVV.

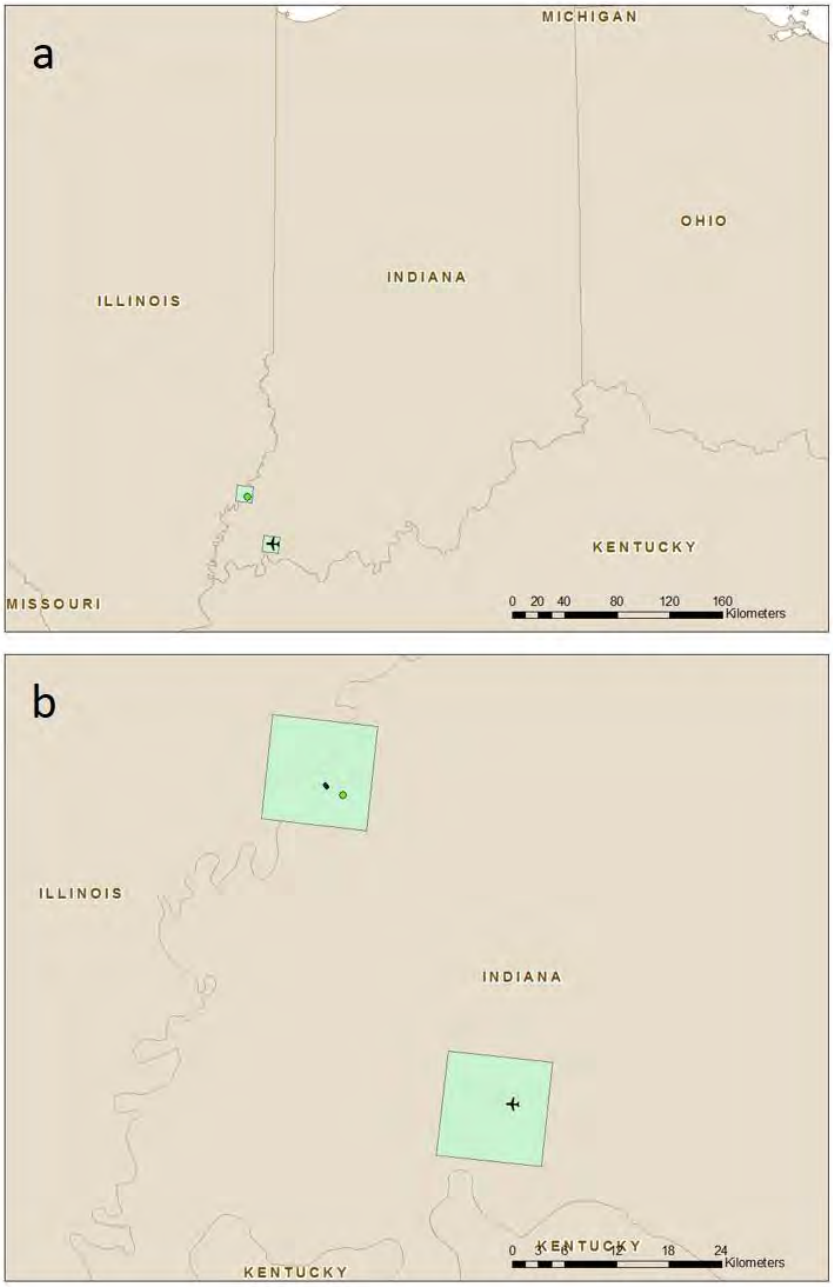

Figure 3. Locations of Gibson meteorological tower (green dot), Evansville (EVV) NWS station (airplane) and 12 km WRF grid cells containing those stations. Panel a is a statewide view and panel b is a closer view of southwest Indiana.

For the purposes of this study, 10 m pressure and insolation were input into AERMET as well as 60 m winds and temperature. The σ_Θ data was not used due to suspected problems with the data (Frost, 2014b).

Four AERMET (version 14134) and AERMOD (version 14134) runs were performed using four sources of meteorological data (meteorological and model run names in parentheses):

- Site-specific data supplemented with EVV surface data (GIB OBS)
- EVV data only (EVV OBS)
- 12 km WRF processed through MMIF and AERMET for the WRF grid cell containing the facility (GIB MMIF)
- 12 km WRF processed through MMIF and AERMET for the WRF grid cell containing EVV (EVV MMIF)

All four scenarios utilized a 0.5 m/s wind speed threshold. In addition to showing the locations of the observed meteorological data sources, Figure 3 shows the locations of the two WRF grid cells processed through MMIF and AERMET. While MMIF can process data for AERMOD input directly, MMIF was run to process MMIF for AERMET in accordance with Section 8.4.2 of Appendix W and MMIF guidance (U.S. EPA, 2016b). Evaluation of the WRF data can be found in U.S. EPA (2014). Note evaluations are for 2011 but are applicable to 2010 as well. MMIF was processed to output 27 layers using the MID interpolation option. The output heights corresponded to the vertical grid structure (25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, and 5000 m). Upper air data in the Forecast Systems Laboratory (FSL) format was output for every hour.

For GIB OBS and EVV OBS surface characteristics, AERSURFACE (U.S. EPA, 2013) was run to determine albedo, Bowen ratio, and surface roughness length for 12 months and 12 30º surface roughness sectors. Monthly Bowen ratios were adjusted using soil moisture and precipitation data from the National Centers for Environmental Information (NCEI)[1] for EVV. Adjustments to albedo, Bowen ratio, and surface roughness length for winter months were determined using snowfall data for EVV from NCEI and ratios of days with at least one inch of snow to days with no snow cover were also used to adjust values (Frost, 2014).For GIB MMIF and EVV MMIF, surface characteristics from the MMIF processor were used in accordance with guidance in Section 8.4.2(b) of Appendix W (U.S, EPA, 2016a) and Section 3.3 of MMIF guidance (U.S. EPA, 2016b). The grid cells for Gibson and Evansville were processed to output both AERMET ready inputs and AERMOD ready inputs. Due to a coding error in the version of MMIF used for this evaluation, when processing multiple grid cells for AERMET input, the surface characteristics were output incorrectly for the AERMET ready files for Gibson and Evansville. The surface characteristics were determined using the values from the AERMOD ready files.

[1] Formerly the National Climatic Data Center (NCDC)

Surface characteristics for 12 months and one 360º sector were calculated. More information about the meteorological data can be found in Section 3.1.1 of this document.

2.1.2 Martins Creek

As previously stated, Martins Creek was one of the evaluation databases used in AERMOD development (U.S, EPA, 2003; Perry et al., 2005). In those evaluations, AERMOD performed better than other models (ISCST3, CTDMPLUS, ISCST3, and RTDM). The Martins Creek Steam Electric Station is located along the Delaware River on the Pennsylvania/New Jersey border, approximately 30 km northeast of Allentown, PA and 95 km north of Philadelphia, PA (U.S. EPA, 2003). See Figure 4 for the general location of Martins Creek and meteorological sites.

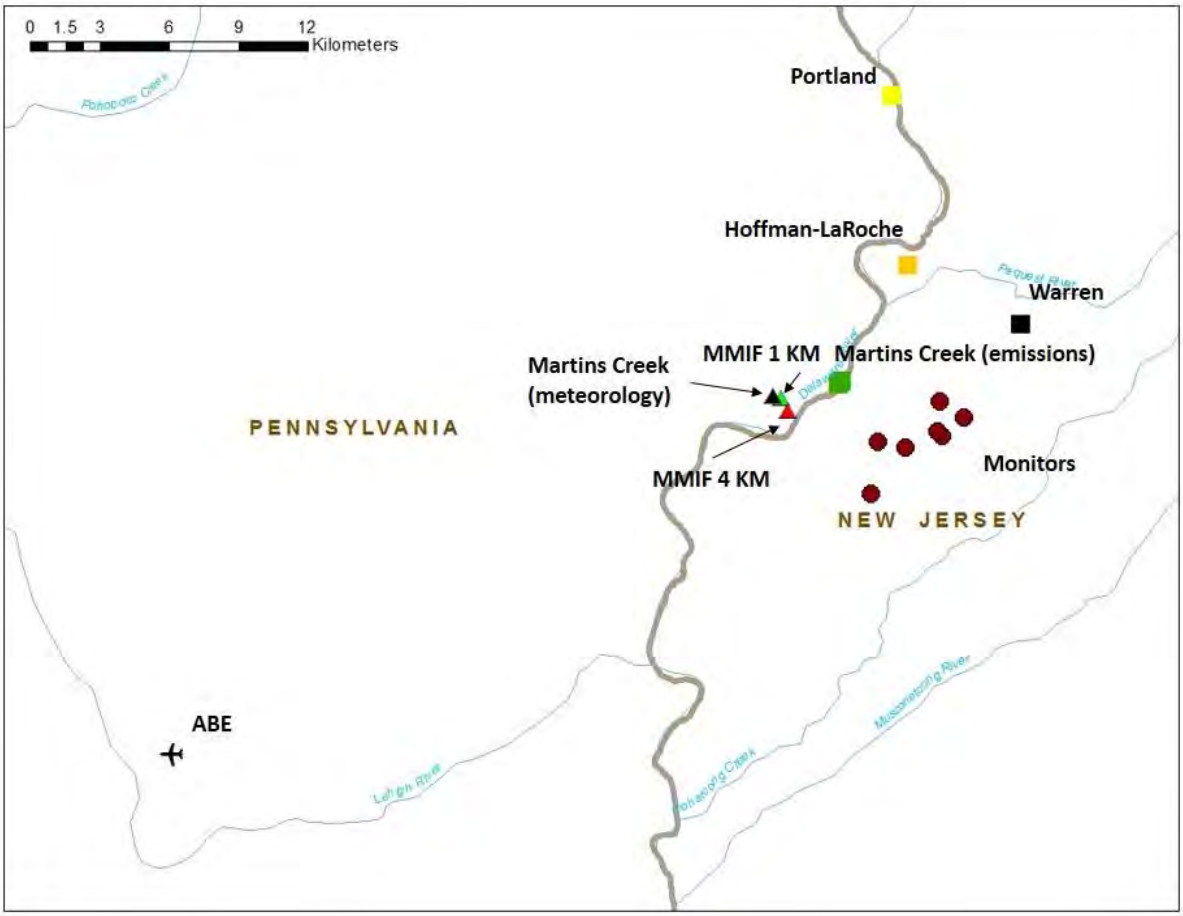

Figure 4. Location of Martins Creek study area with meteorological sites, monitors, and emission sources.

SO_2 measurements were taken at seven monitors east of the facility (Figure 5) on Scott's Mountain. Site-specific meteorology was recorded from May 1, 1992 to May 19, 1993.

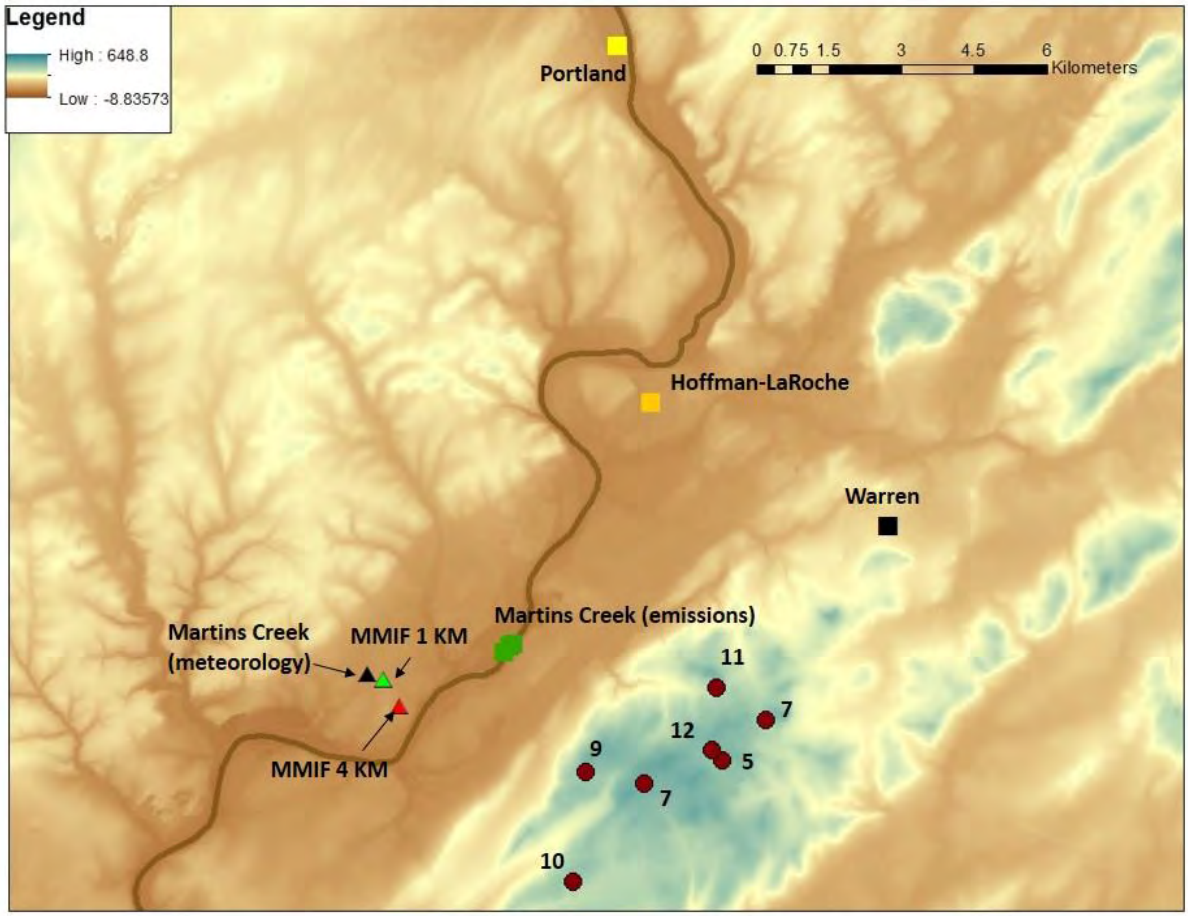

Figure 5. Detailed view of Martins Creek meteorological sites, monitors, and emissions sources with detailed terrain.

Data included temperature, wind speed, wind direction, and σ_Θ at 10 m height west of the facility (Figure 4). In addition, hourly multi-level wind measurements were taken by a sodar located southwest of the facility. Upper air data was from Albany, NY with missing soundings substituted from Sterling, VA.

In addition to emissions from Martins Creek, three other facilities were modeled: 1) Portland, Hoffman-LaRoche, and Warren (Figure 5). Hourly emissions and hourly stack parameters (temperature and exit velocity) at all four facilities were modeled in AERMOD.

Table 2 lists the units at each facility with stack parameters and annual emissions.

Table 2. Martins Creek study facility emissions (tons/year) and stack parameters.

Facility	Stack	Annual SO_2 (tpy)	Stack height (m)	Stack diameter (m)	Avg. exit temperature (K)[1]	Avg. exit velocity (m/s)[1]
Martins Creek	MC12	20272	182.9	5.3	400.8	17.1
	MC3	2923	182.9	6.9	403.9	17.5
	MC4	3395	182.9	6.9	403.0	18.7
Portland	ED1	5459	121.9	3.1	395.0	33.3
	ED2	12939	121.9	3.6	400.0	26.4
Hoffman-LaRoche	HL2	837	59.4	2.7	451.5	6.8
Warren	WC1	0.2	76.2	1.9	404.6	3.3
	WC2	0.2	76.2	1.9	410.3	3.4

1. Average temperature and exit velocity based on hours with non-zero emissions.

Four AERMET (version 14134) and AERMOD (version 14134) simulations were performed:

- Site-specific data supplemented with Allentown/Bethlehem (ABE) data (Martins Creek)
- ABE only data (ABE)
- 1 km WRF simulation of grid cell containing Martins Creek (MMIF 1 km)
- 4 km WRF simulation of grid cell containing Martins Creek (MMIF 4 km)

Figure 5 shows the locations of the 1 km and 4 km WRF grid cells processed through MMIF and AERMET.

For the Martins Creek site-specific data, the surface characteristics supplied with the evaluation database on SCRAM were used. For the ABE surface characteristics, AERSURFACE (U.S. EPA, 2013) was run to determine albedo, Bowen ratio, and surface roughness length for 12 months and three surface roughness sectors (110º-230º, 230º-330º, and 330º-110º). Winter was assumed to be "no snow" and moisture conditions were assumed to be dry based on climatological data for ABE.

For the 1 km and 4 km MMIF generated outputs, surface characteristics from the MMIF processor were used in accordance with guidance in Section 8.4.2(b) of Appendix W (U.S, EPA, 2016a) and Section 3.3 of MMIF guidance (U.S. EPA, 2016b). Surface characteristics for 12 months and one 360º sector were calculated. MMIF was run to output 10 layers (17.13, 51.46, 85.91, 120.48, 207.7, 348.56, 564.66, 936.27, 1566.38, and 2851.29 m) using the TOP interpolations. More information about the meteorological data can be found in Section 3.2.1 of this document.

2.1.3 Herculaneum

The third evaluation site is focused on the Herculaneum, MO area, specifically the Doe Run lead facility (Figures 6 through 8). This evaluation allowed for the use of two observed datasets, site-specific and NWS station and three MMIF outputs at three different horizontal grid resolutions, 4 km, 12 km, and 36 km. Figures 6 through 8 show the spatial relationship between the Doe Run facility, two lead monitors, and the various meteorological data locations. Table 3 lists the major modeled point sources with emission rates. The five scenarios are:

- Site specific data supplemented with St. Louis (CPS) meteorological data (Herculaneum)
- CPS data only (CPS)
- 36 km WRF simulation of grid cell containing Doe Run (MMIF 36 km)
- 12 km WRF simulation of grid cell containing Doe Run (MMIF 12 km)
- 4 km WRF simulation of grid cell containing Doe Run (MMIF 4 km)

The Herculaneum site-specific tower and CPS were processed through AERMET using version 13350 with ILX upper air data. The WRF grid cells were processed through MMIF to generate AERMOD and AERMET ready input files. MMIF was run using the same layer structure and options and FSL formatted output frequency as for GIBSON. All five scenarios were processed in version 13350 of AERMOD.

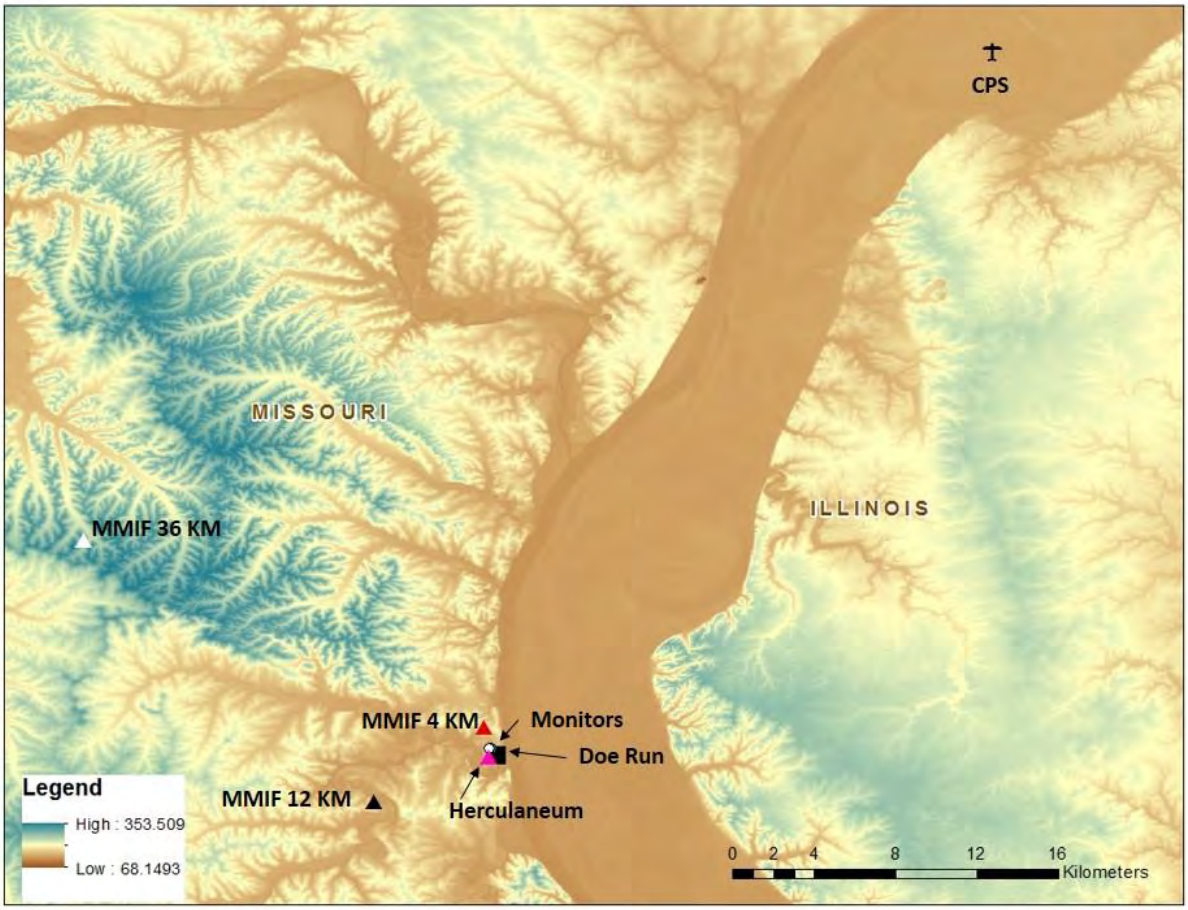

Figure 6. Location of the Doe Run lead facility, NWS station (CPS), site-specific tower (Herculaneum), three MMIF output locations (MMIF 4 KM, MMIF 12 KM, and MMIF 36 KM), and lead monitors.

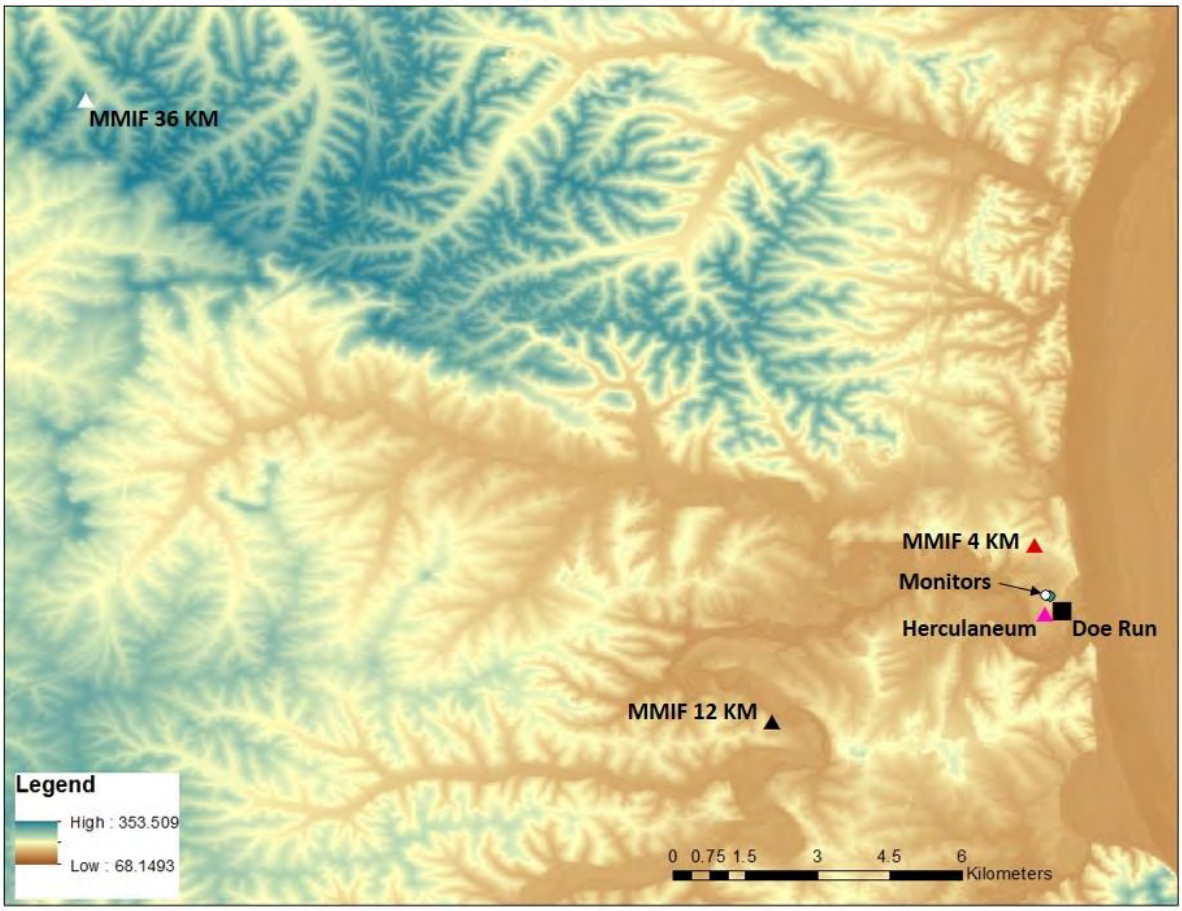

Figure 7. As for Figure 6, with more detail around the facility.

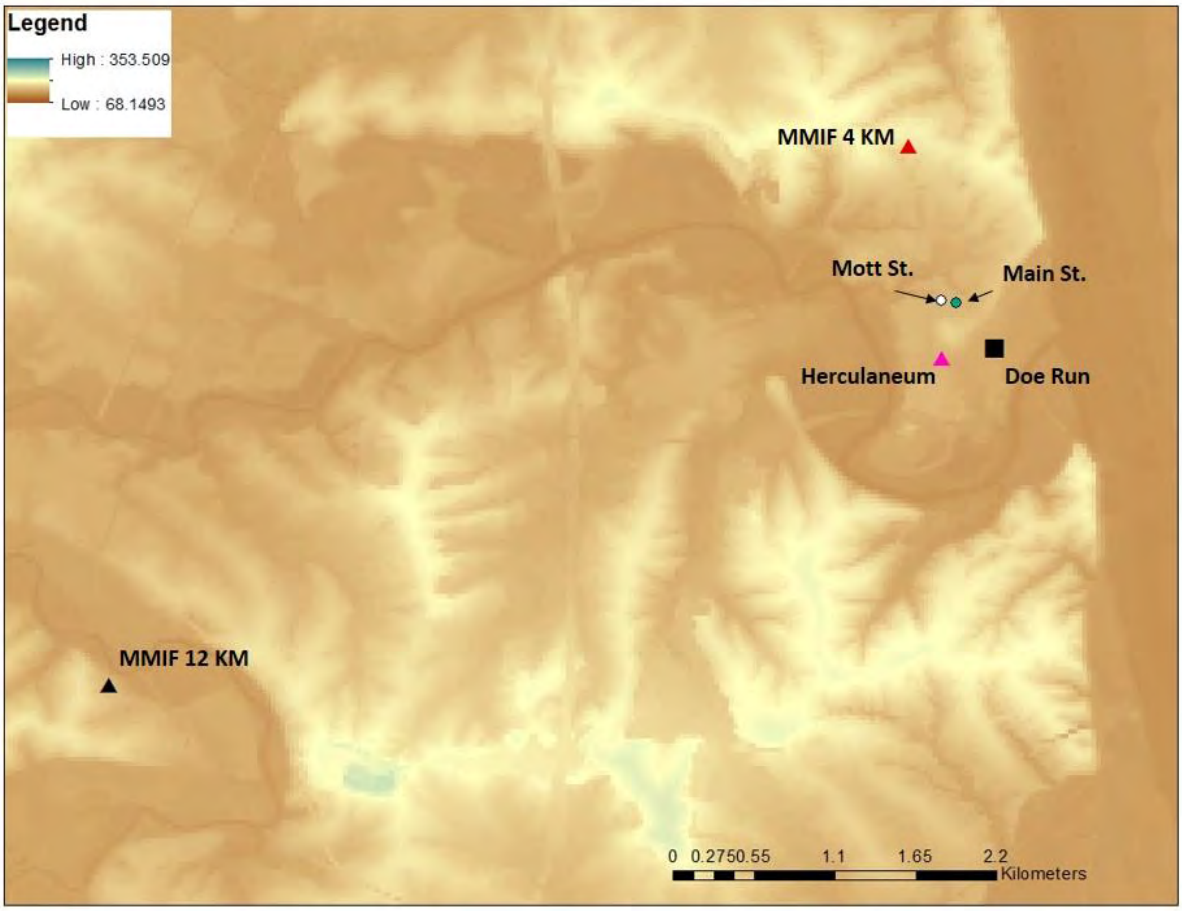

Figure 8. Detailed view of MMIF grid cell locations, Herculaneum site specific tower, monitors, and Doe Run.

Table 3. Herculaneum point source emission rates and stack parameters.

Stack	Emissions (g/s)	Stack height (m)	Avg. exit temperature (K)	Avg. exit velocity (m/s)	Stack diameter (m)
30001	4.17E+00	100.75	346.67	5.81	10.31
40004	8.58E-04	21.3	391.5	0.69	0.76
40005	8.58E-04	21.3	391.5	0.69	0.76
50007	4.31E-02	45.72	285.56	7.13	2.59
50008	2.97E-01	45.72	276.11	34.57	3.05
50011	1.65E-03	18.8	989.3	5.96	0.61
50012	1.65E-03	18.8	989.3	5.96	0.61
50013	1.65E-03	18.8	989.3	5.96	0.61
50014	1.65E-03	18.8	989.3	5.96	0.61
50015	1.65E-03	18.8	989.3	5.96	0.61
50016	1.65E-03	18.8	989.3	5.96	0.61
50017	1.65E-03	18.8	989.3	5.96	0.61
50018	1.65E-03	18.8	989.3	5.96	0.61
60001	1.13E-04	21.3	699.8	2.73	0.56
60002	1.13E-04	21.3	699.8	2.73	0.56
60003	5.93E-06	7.6	297	7.7	1.08
60004	1.80E-03	6.1	327.6	17.5	0.25
60005	1.17E-03	16.8	297	5	0.56
60006	1.17E-03	16.8	297	5	0.56
60007	1.17E-03	16.8	297	5	0.56
60008	1.17E-03	16.8	297	5	0.56

2.2 Meteorological data evaluation

Evaluation among the various meteorological datasets for each study encompassed comparing the hourly distributions of each dataset, comparing the distributions of the hourly differences among pairs of datasets, and calculations of several statistics, including mean bias, fractional bias, root mean square error, correlation, and index of agreement. For the wind direction difference statistics, a difference called displacement, which is the difference in the U and V vectors of the modeled and observed winds and was used. This was used in the assessment of the 2011 12km WRF simulations over the U.S. (US EPA, 2014). The displacement can be calculated as:

$$D = abs((U_M - U_O + V_M - V_O) \times (1\ km/1000\ m) \times (3600\ s/hr) \times 1hr) \quad (1)$$

Where D is the displacement in km, U_M and V_M are the u and v components respectively of the modeled wind vector and U_O and V_O are the u and v components of the observed wind vector.

2.3 Model evaluation methodology

AERMOD output among the different meteorological datasets was evaluated using the EPA Protocol for determining the best performing model, or Cox-Tikvart method (U.S. EPA, 1992; Cox and Tikvart, 1990). The protocol uses a two-step process for determining the better performing model when comparing models. The first step is a screening test that fails to perform at a minimal operational level. The second test applies to those models that pass the screening test that uses bootstrapping to generate a probability distribution of feasible outcomes (U.S. EPA, 1992). This section will discuss the methodology using the evaluation cases as examples.

The first step is to perform a screening test based on fractional bias:

$$FB = 2\left[\frac{OB-PR}{OB+PR}\right] \quad (1)$$

Where FB is the fractional bias, OB is the average of the highest 25 observed concentrations and PR is the average of the highest 25 predicted averages. The fractional bias is also calculated for the standard deviation where OB and PR refer to the standard deviation of the highest 25 observed and predicted concentrations respectively. This is done across all monitors and modeled receptors, unpaired in time and space for the 3-hour and 24-hour averaging periods. The fractional bias of the means is plotted against the fractional bias of the standard deviation. Biases that exceed a factor-of-two under-prediction or over-prediction are considered grounds for excluding a model for further evaluation (U.S. EPA, 1992).

Models that pass the screening test are subjected to a more comprehensive statistical comparison that involves both an operational and scientific component. The operational component is to measure the model's ability to estimate concentration statistics most directly used for regulatory purposes and the scientific component evaluates the model's ability to perform accurately throughout the range of meteorological conditions and the geographic area of concern (U.S. EPA, 1992). The test statistic used for the comparison is the robust highest concentration (RHC) statistic and is given by:

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

Where X(N) is the Nth largest value, $\overline{X}$ is the average of N-1 values, and N is the number of values exceeding the threshold value, usually 26.

The operational component of the evaluation compares performance in terms of the largest network-wide RHC test statistic. The RHC is calculated separately for each monitor within the network for observations and modeled values. The highest observed RHC is then compared to the highest modeled RHC using equation 1, where RHC now replaces the means of the top 25 values of observed or modeled concentrations. Absolute fractional bias (the absolute value of fractional bias), AFB is calculated for 3 and 24-hour averages.

The scientific component of the evaluation is also based on absolute fractional bias but the bias is calculated using the RHC for each meteorological condition and monitor. The meteorological conditions are a function of atmospheric stability and wind speed. For the purposes of these studies, six unique conditions were defined based on two wind speed categories (below and above 2.0 m/s) and three stability categories: unstable, neutral, and stable. [2] In this evaluation, only 1-hour concentrations are used and the AFB is based on RHC values paired in space and stability/wind speed combination.

A composite performance measure (CPM) is calculated from the 1-hour, 3-hour, and 24-hour AFB's:

$$CPM = \frac{1}{3} \times \left(\overline{AFB_{i,j}}\right) + \frac{2}{3} \times \left[\frac{AFB_3 - AFB_{24}}{2}\right] \quad (3)$$

Where $AFB_{i,j}$ is the absolute fractional bias for monitor i and meteorological condition j, $\overline{AFB_{i,j}}$ is the average absolute fractional bias across all monitors and meteorological conditions, AFB_3 is the absolute fractional bias for the 3-hour average, and AFB_{24} is the absolute fractional bias for the 24-hour average. Once CPM values have been calculated for each model, a model comparison measure is calculated to compare the models:

$$MCM_{A,B} = CPM_A - CPM_B \quad (4)$$

Where CPM_A is the CPM for model A and CPM_B is the CPM for model B. When more than two models are being compared simultaneously, the number of MCM values is equal to the total of the number of unique combinations of two models. In the case of these evaluations, it is not the number of models but the number of meteorological scenarios. For Gibson and Martins Creek, there are four meteorological scenarios each, so there were six MCM comparisons for each location. For Herculaneum, CPM values could not be calculated because there were only 24-hour average concentrations available from the two monitors.

In order to determine if the difference between models was statistically significant, the standard error was calculated. A bootstrapping technique was used to create 1000 sample years based on methodology outlined in U.S. EPA (1992). The original data is divided into 3-day blocks. Within each season, the 3-day blocks are sampled with replacement until a total season is created. The process is repeated until a 1000 boot-strap years are created[3]. The standard error is calculated as the standard deviation of the bootstrap generated outcomes for the MCM.

The magnitude and sign of the MCM are indicative of relative performance of each pair of models. The smaller the CPM the better the overall performance of the model. This means that

---

[2] In U.S. EPA (1992), the three stability categories are related to the Pasquill-Gifford categories, unstable being A, B, and C, neutral being D, and stable being E and F. Since AERMOD does not use the stability categories, the stability class was determined using Monin-Obukhov length and surface roughness using methodology from AERMOD subroutine LTOPG.

[3] The bootstrapping was completed using the SAS© SURVEYSELECT procedure with resampling for 1000 replicates.

for two models, A and B, a negative difference between the CPM for A and CPM for B implies that model A is performing better (Model A has a smaller CPM) while a positive difference indicates that Model B is performing better.

Since more than two scenarios are being evaluated in these studies, simultaneous confidence intervals of 90 and 95 percent were calculated. These were calculated by finding the 90th and 95th percentiles of the distribution across all MCM values from the bootstrapping procedure for all model comparisons. The confidence intervals were then found by:

$$CI_{X,A,B} = MCM_{A,B} \pm c_X s_{A,B} \quad (5)$$

Where $CI_{X,A,B}$ is the confidence interval for X percent (90 or 95th) for models A and B, $MCM_{A,B}$ is as defined in Equation 4, c_X is the X percentile of the MCM values from the bootstrap results and $s_{A,B}$ is the standard deviation of the bootstrap MCM results for models A and B. Note that in Equation 5, $MCM_{A,B}$ is the MCM value from the original data, not the bootstrap results.

For each pair of model comparisons, the significance of the model comparison measure depended on whether the confidence interval overlapped zero. If the confidence interval overlapped zero, then the two models were not performing at a level which was considered statistically different. Otherwise, if they did not overlap zero, then there was a statistically significant difference between the two models.

3. Results

This section provides results for each of the three case study areas. Meteorological data comparisons among the different scenarios for each site are shown as well as the associated AERMOD results for each meteorological scenario.

3.1 Gibson

3.1.1 Meteorological data comparisons

Figure 9 compares the 2010 wind roses for site specific tower (GIB OBS), 12 km WRF cell of the site specific tower (GIB MMIF), Evansville NWS observations (EVV OBS) and the 12 km WRF cell of EVV (EVV MMIF). The wind roses for GIB OBS and GIB MMIF have similarities but GIB OBS exhibits a stronger component of winds from the southwest while GIB MMIF exhibits a stronger components of winds from the northwest. For Evansville, the observed winds (EVV OBS) exhibit dominant directions of southwest, northwest and northeast. The MMIF wind rose (EVV MMIF) exhibits a strong component of winds from the northwest with a secondary component from the southwest.

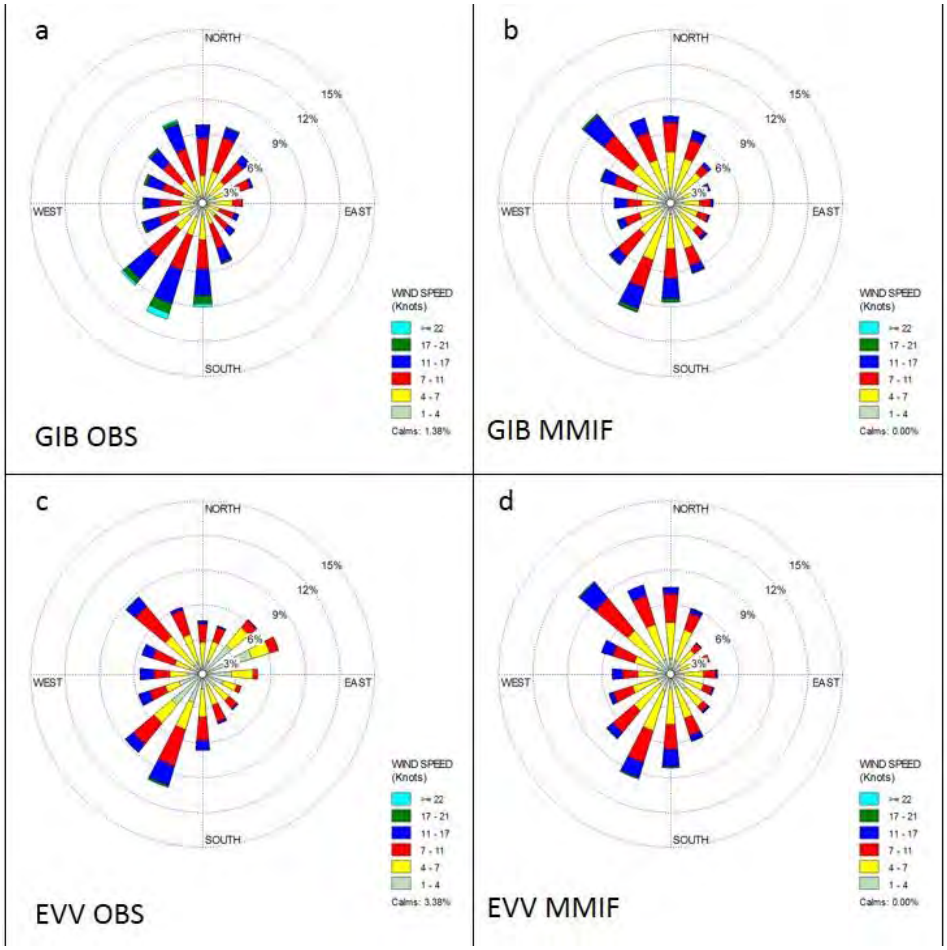

Figure 9. Gibson 2010 wind roses for a) GIB OBS, b) GIB MMIF, c) EVV OBS, and d) EVV MMIF.

Wind displacement (Figure 10), distribution for GIB MMIF – EVV MMIF shows a tighter distribution than the other three indicating wind directions, on an hourly basis, are in more agreement between the two model grid cells. The other three bias distributions are similar, with larger displacement values between the two observed datasets.

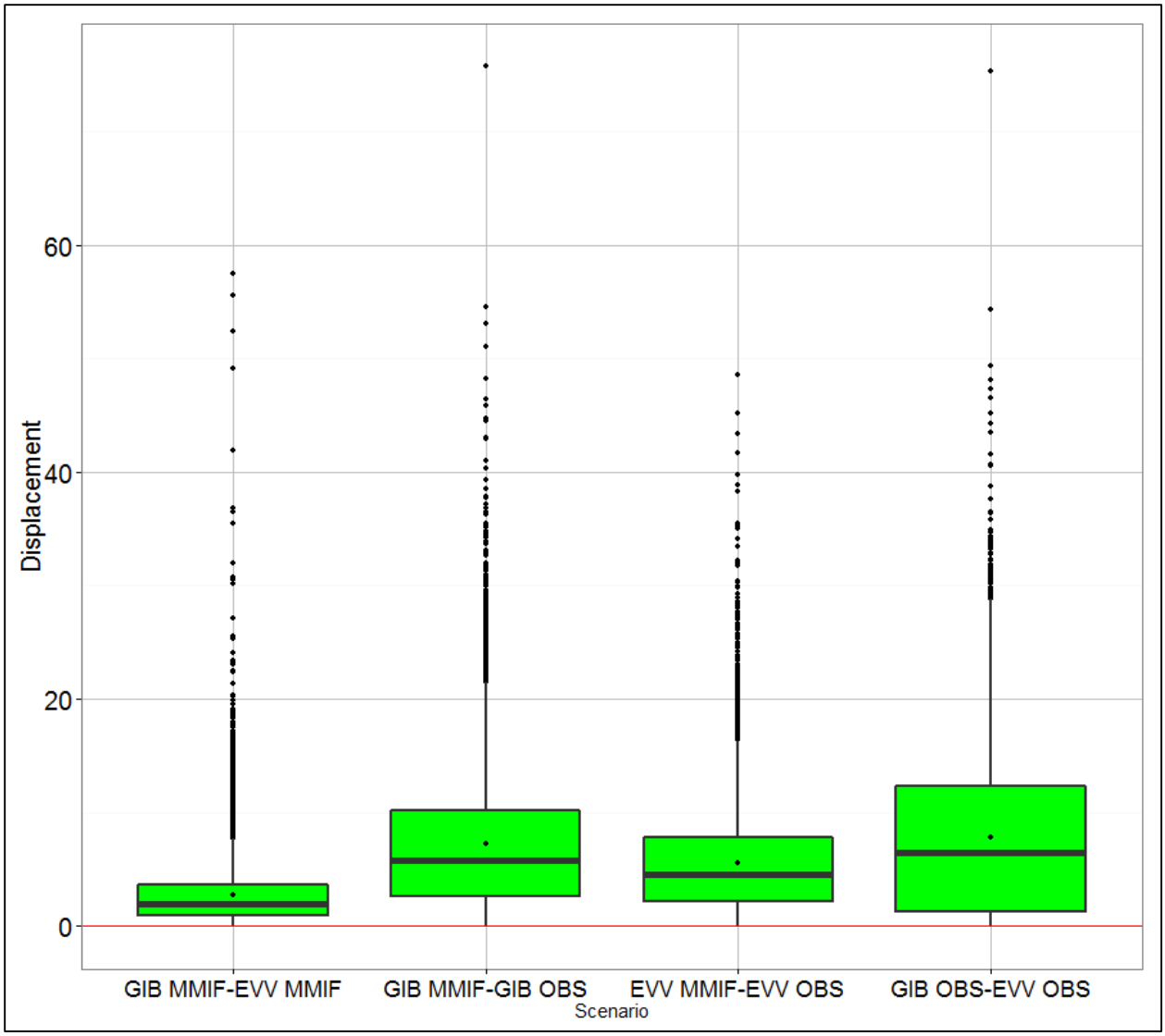

Figure 10. Wind displacement (km) among the Gibson meteorological scenarios.

Figures 11 through 13 show monthly values of surface roughness, z_o (m) by 10º sectors for the four meteorological scenarios. For most of the winter, the EVV MMIF surface roughness is the highest, over 0.3 m. The plots show that GIB OBS can vary significantly depending on the wind direction and month. EVV OBS also vary depending on month and sector.

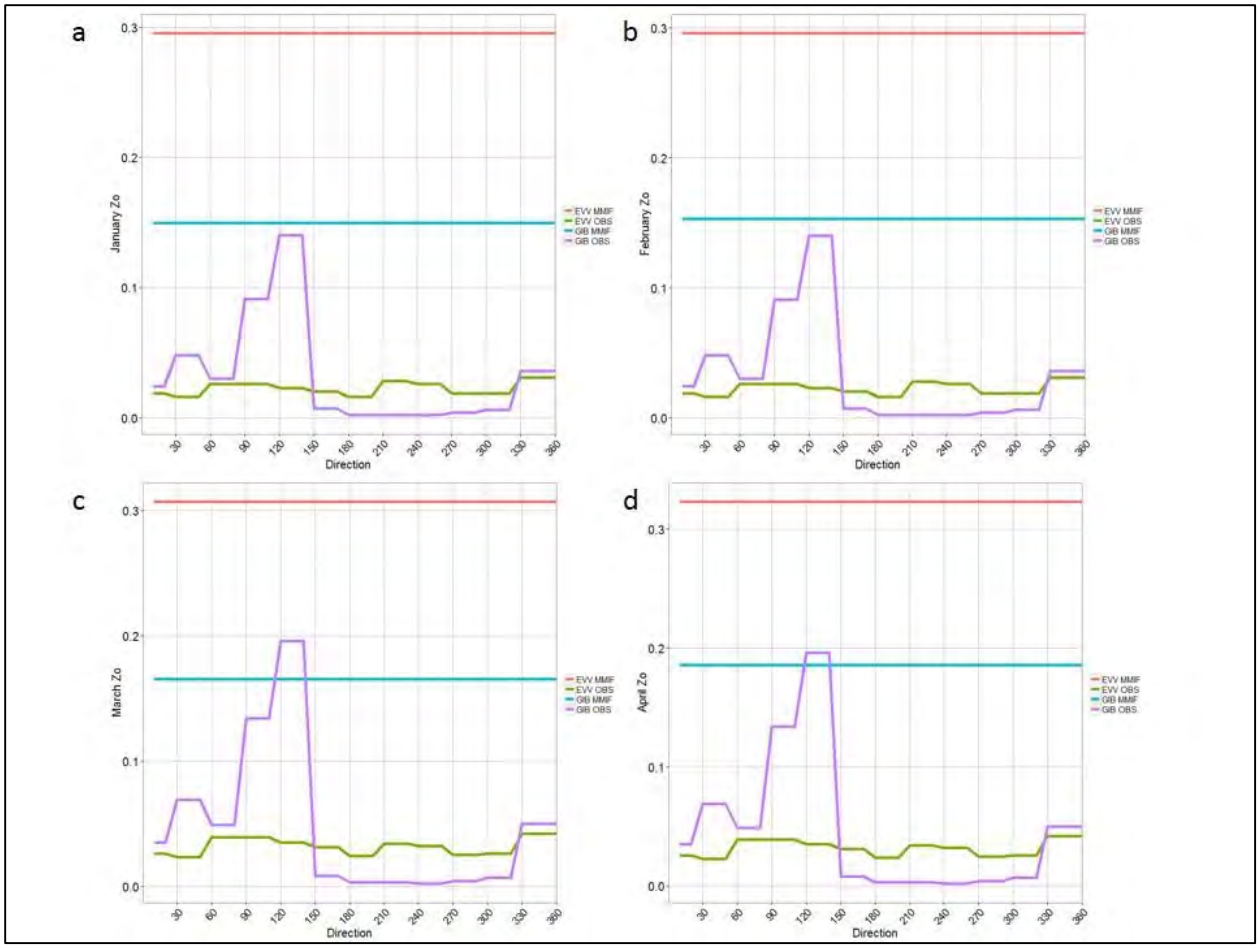

Figure 11. Gibson study monthly surface roughness lengths (m) by 10 degree sectors for a) January, b) February, c) March, and d) April.

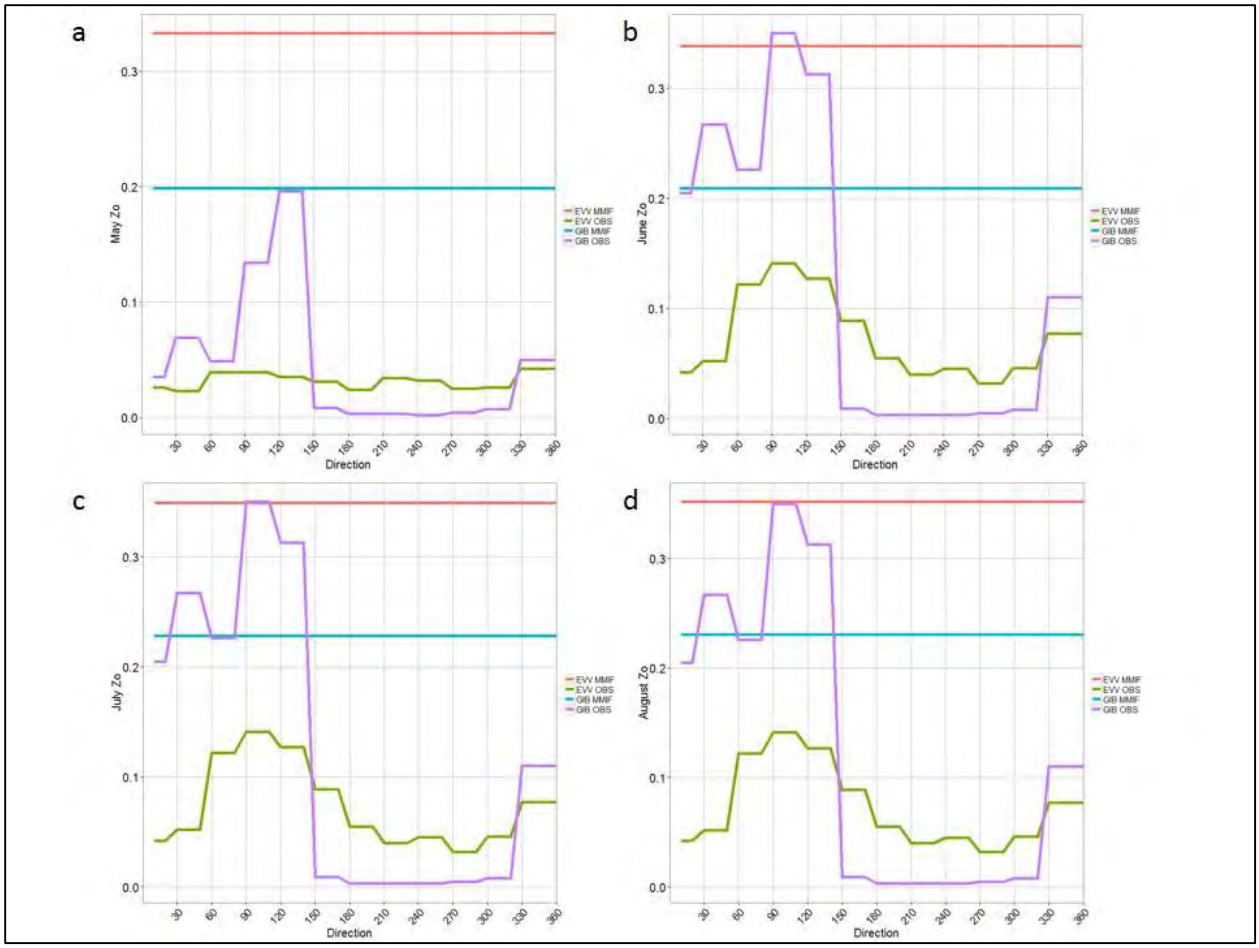

Figure 12. Gibson study monthly surface roughness lengths (m) by 10 degree sectors for a) May, b) June, c) July, and d) August.

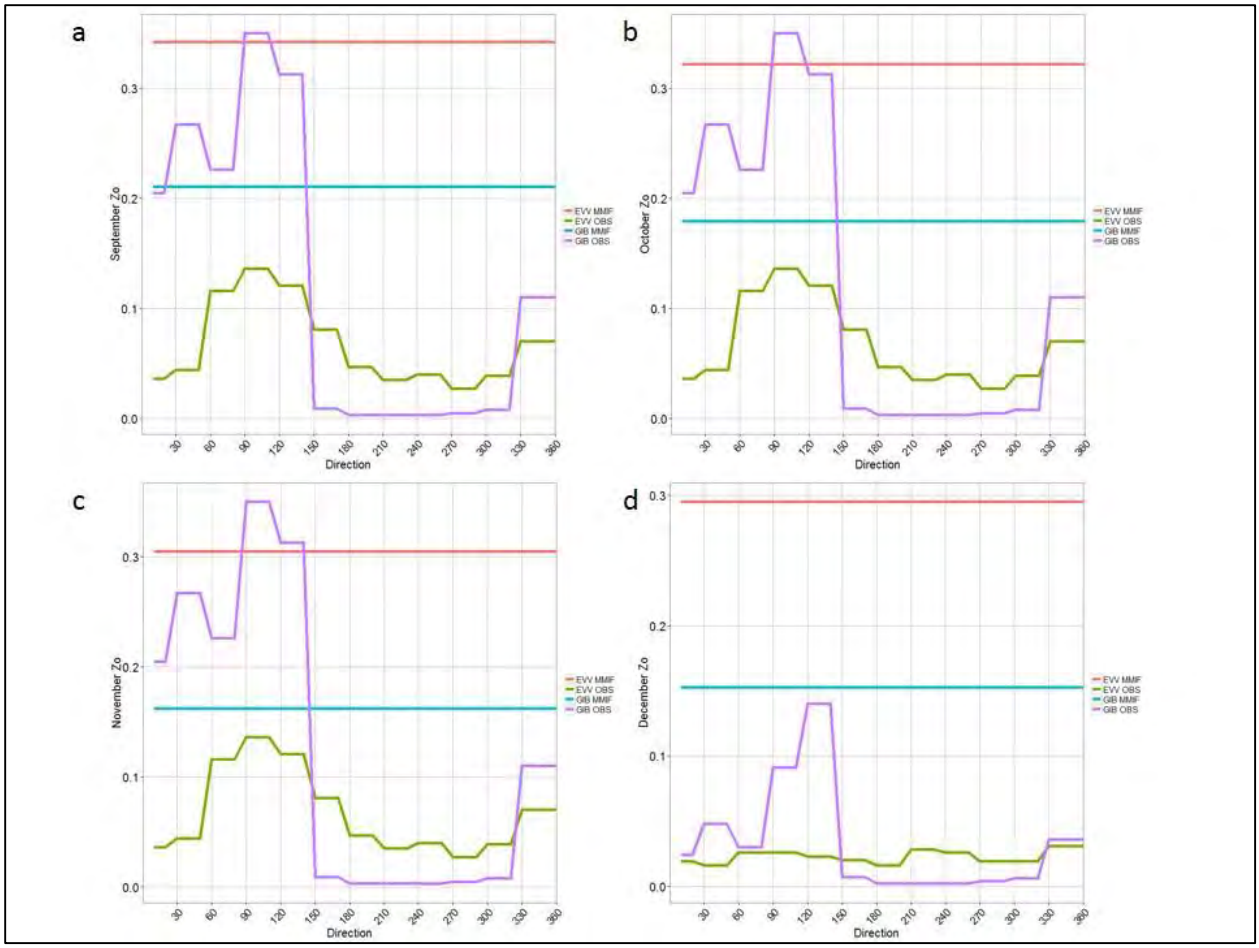

Figure 13. Gibson study monthly surface roughness lengths (m) by 10 degree sectors for a) September, b) October, c) November, and d) December.

Tables 4 through 6 show statistics for several primary variables including wind speed, temperature, pressure, and relative humidity. The tables also show statistics for heat flux, surface friction velocity (u*), convective velocity scale (w*), Monin-Obukhov length (L), convective mixing height (Zi_c), mechanical mixing height (Zi_m), potential temperature gradient ($d\Theta/dz$), and cloud cover. In each table, the GIB OBS – EVV OBS are highlighted as they can be considered a "control" since in the absence of prognostic data, these would be the only available data sources. Box and whisker plots of the variable distributions as well as bias distributions can be found in Appendix A. Findings include:

- Differences between MMIF wind speeds and observed wind speeds appear to be lower than the differences between the two observed datasets. The GIB MMIF scenario under-predicts when compared to the GIB OBS and GIB MMIF over-predicts when compared to EVV MMIF.
- MMIF scenarios tend to under-predict temperatures and pressure when compared to their respective observed counterparts. The GIB MMIF scenario under-predicts (over-

predicts) when compared to EVV MMIF for temperature (pressure), but GIB OBS over-predicts when compared to EVV OBS for both variables.

- MMIF scenarios over-predict relative humidity compared to their observed counterparts. It should be noted that the GIBS-EVV OBS are zero because EVV observed RH is used for the Gibson site-specific tower in AERMET.
- For heat flux, u*, and w*, the MMIF scenarios over-predict compared to their observed counterparts. GIB MMIF under-predicts compared EVV MMIF as does the GIB OBS when compared to EVV OBS.
- For the potential temperature gradient, the MMIF scenarios under-predict when compared to their observed counterparts and the GIB MMIF scenario under-predicts compared against EVV MMIF and GIB OBS under-predicts compared to EVV OBS.
- For mixing heights (both convective and mechanical), the MMIF scenarios over-predict compared to the observed counterparts with low R2 and IOA when compared to the GIB OBS – EVV OBS differences.
- Monin-Obukhov length differences show relatively low agreement between the MMIF and observed scenarios.
- Cloud cover differences also show relatively low agreement but this may be due to the calculation methodology in AERMET when cloud cover is missing.

Overall, while there are differences, the MMIF scenarios appear to show relatively good agreement with the observed data. Differences between the MMIF scenarios and their respective counterparts is usually in line with the differences between the site-specific data and NWS data (i.e., GIB OBS – EVV OBS).

Table 4. Mean bias, fractional bias, root mean square error, and R^2 for primary meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Wind speed	GIB MMIF-EVV MMIF	0.2395	0.0130	0.6873	0.8899
	GIB MMIF-GIB OBS	-0.6808	-0.0309	1.7568	0.5385
	EVV_MMIF-EVV OBS	0.4345	0.0564	1.1826	0.6568
	GIB OBS-EVV OBS	1.3562	0.0943	2.1125	0.5348
Ambient temperature	GIB MMIF-EVV MMIF	-0.3321	-0.0003	0.8500	0.9957
	GIB MMIF-GIB OBS	-0.6568	-0.0006	3.0398	0.9376
	EVV MMIF-EVV OBS	-0.1067	-0.0001	1.9698	0.9736
	GIB OBS-EVV OBS	0.2133	0.0002	3.0920	0.9351
Pressure	GIB MMIF-EVV MMIF	0.4412	0.0001	0.7131	0.9920
	GIB MMIF-GIB OBS	-3.9806	-0.0010	4.1719	0.9658
	EVV MMIF-EVV OBS	-1.0903	-0.0003	1.5625	0.9728
	GIB OBS-EVV OBS	3.3315	0.0008	3.6102	0.9557
Relative humidity	GIB MMIF-EVV MMIF	1.1127	0.0045	6.2611	0.9227
	GIB MMIF-GIB OBS	6.5496	0.0216	14.3825	0.6555
	EVV MMIF-EVV OBS	5.4416	0.0172	14.1551	0.6511
	GIB OBS-EVV OBS	0.0000	0.0000	0.0000	1.0000

Table 5. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Heat flux	GIB MMIF-EVV MMIF	-8.1307	-0.0409	29.7447	0.9256
	GIB MMIF-GIB OBS	16.2727	0.1259	54.6334	0.6997
	EVV MMIF-EVV OBS	11.4507	0.1025	53.7118	0.7413
	GIB OBS-EVV OBS	-13.8544	-0.0673	36.5569	0.8467
u*	GIB MMIF-EVV MMIF	-0.0315	-0.0209	0.0757	0.9050
	GIB MMIF-GIB OBS	0.1660	0.1440	0.2150	0.6087
	EVV MMIF-EVV OBS	0.1858	0.1643	0.2271	0.6816
	GIB OBS-EVV OBS	-0.0159	0.0013	0.0839	0.6753
w*	GIB MMIF-EVV MMIF	-0.1095	-0.0208	0.2595	0.8677
	GIB MMIF-GIB OBS	0.3639	0.1077	0.5728	0.4937
	EVV MMIF-EVV OBS	0.1874	0.0389	0.5022	0.5190
	GIB OBS-EVV OBS	-0.2887	-0.0941	0.4124	0.7618
$d\Theta/dz$	GIB MMIF-EVV MMIF	0.0001	0.0018	0.0013	0.7727
	GIB MMIF-GIB OBS	-0.0016	-0.0375	0.0043	0.2171
	EVV MMIF-EVV OBS	-0.0018	-0.0434	0.0045	0.2009
	GIB OBS-EVV OBS	-0.0001	-0.0040	0.0027	0.6755

Table 6. Mean bias, fractional bias, root mean square error, correlation, and index of agreement for calculated meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Zi_c	GIB MMIF-EVV MMIF	-48.3405	-0.0122	244.6457	0.8405
	GIB MMIF-GIB OBS	250.3559	0.1324	523.6197	0.4137
	EVV MMIF-EVV OBS	66.4046	0.0431	512.7182	0.4251
	GIB OBS-EVV OBS	-236.2854	-0.1140	349.0821	0.8315
Zi_m	GIB MMIF-EVV MMIF	-52.0755	-0.0172	159.9024	0.9166
	GIB MMIF-GIB OBS	249.6961	0.1424	410.5375	0.5749
	EVV MMIF-EVV OBS	274.1437	0.1652	439.7079	0.6514
	GIB OBS-EVV OBS	-35.5563	-0.0001	152.9793	0.6606
L	GIB MMIF-EVV MMIF	24.9063	-0.0197	1592.5484	0.5897
	GIB MMIF-GIB OBS	572.5552	0.1899	2509.5879	0.0475
	EVV MMIF-EVV OBS	336.0061	0.2854	2320.8618	0.0404
	GIB OBS-EVV OBS	-221.9609	0.0651	1389.7583	0.0325
Cloud cover	GIB MMIF-EVV MMIF	0.1830	0.0152	2.2867	0.7180
	GIB MMIF-GIB OBS	1.2902	0.0981	5.7003	0.0307
	EVV MMIF-EVV OBS	2.2222	0.1789	5.4528	0.1123
	GIB OBS-EVV OBS	1.1128	0.1092	3.6298	0.4976

3.1.2 AERMOD results

Figure 14 shows the hourly QQ-plot among all four monitors for the four meteorological scenarios. The GIB OBS concentrations tend to be the highest predicting scenario. The two MMIF driven scenarios and the EVV OBS scenario are visually are in close agreement with each other and tend to under-predict at higher concentrations.

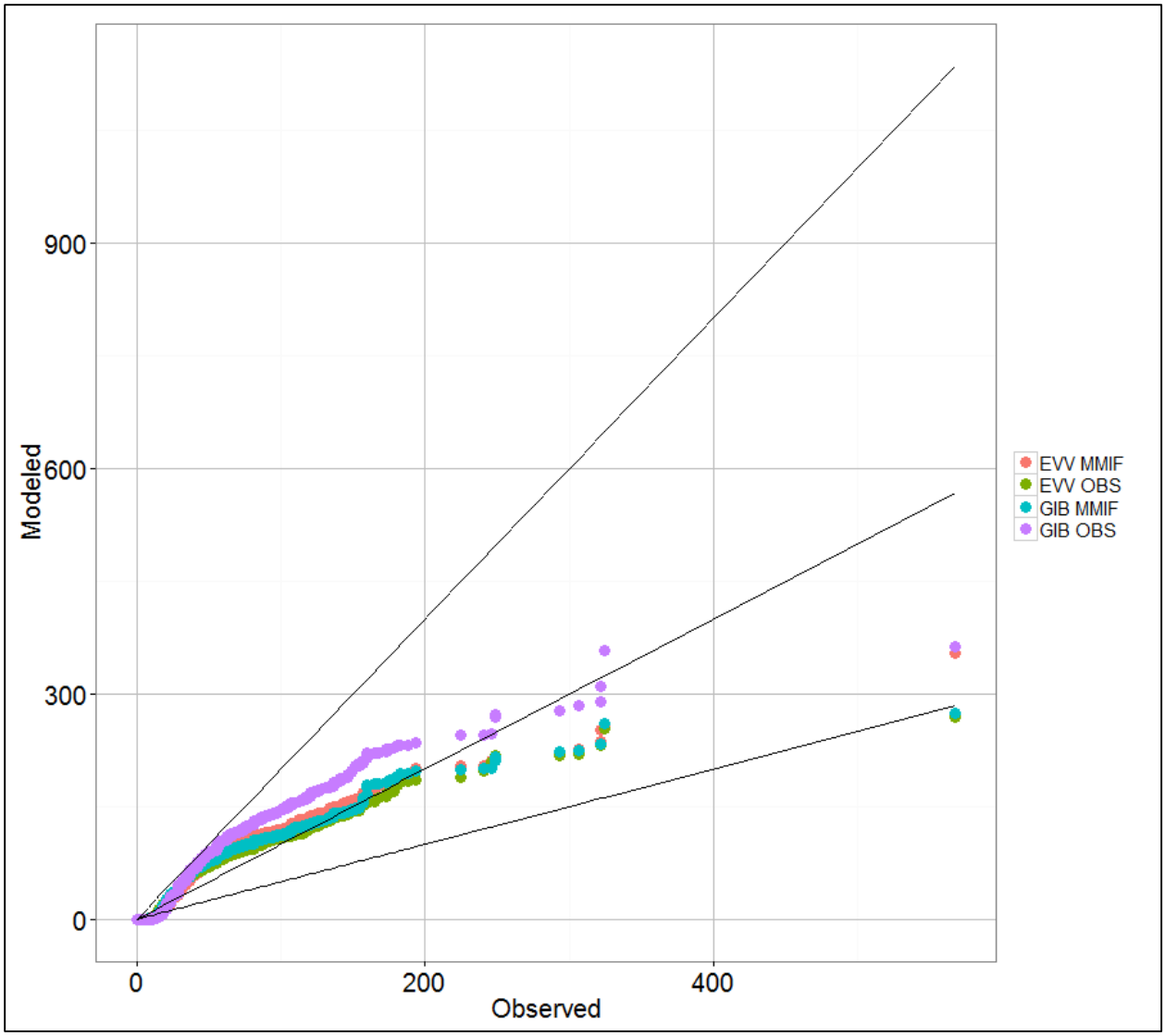

Figure 14. Gibson hourly QQ plots. Concentrations are in μg/m^3.

The modeled scenarios were evaluated using the Cox-Tikvart methodology as discussed in Section 2.3. Figures 29 and 30 show the initial 3-hr and 24-hour screening results. For the 3-hour results, all four scenarios are outside the ±0.67 box that identifies good performance. The 3-hour results correspond with the QQ-plots in Figure 14 in that GIB OBS over-predicts (bias less than 0) and the other scenarios tend to under-predict (bias greater than 0). However, all four scenarios are in close proximity to one another around a bias of zero. For the 24-hour average, all four scenarios fall within the ±0.67 box. For 24-hour averages, all scenarios except for the EVV OBS scenario have slight over-prediction. All four scenarios have a bias close to zero.

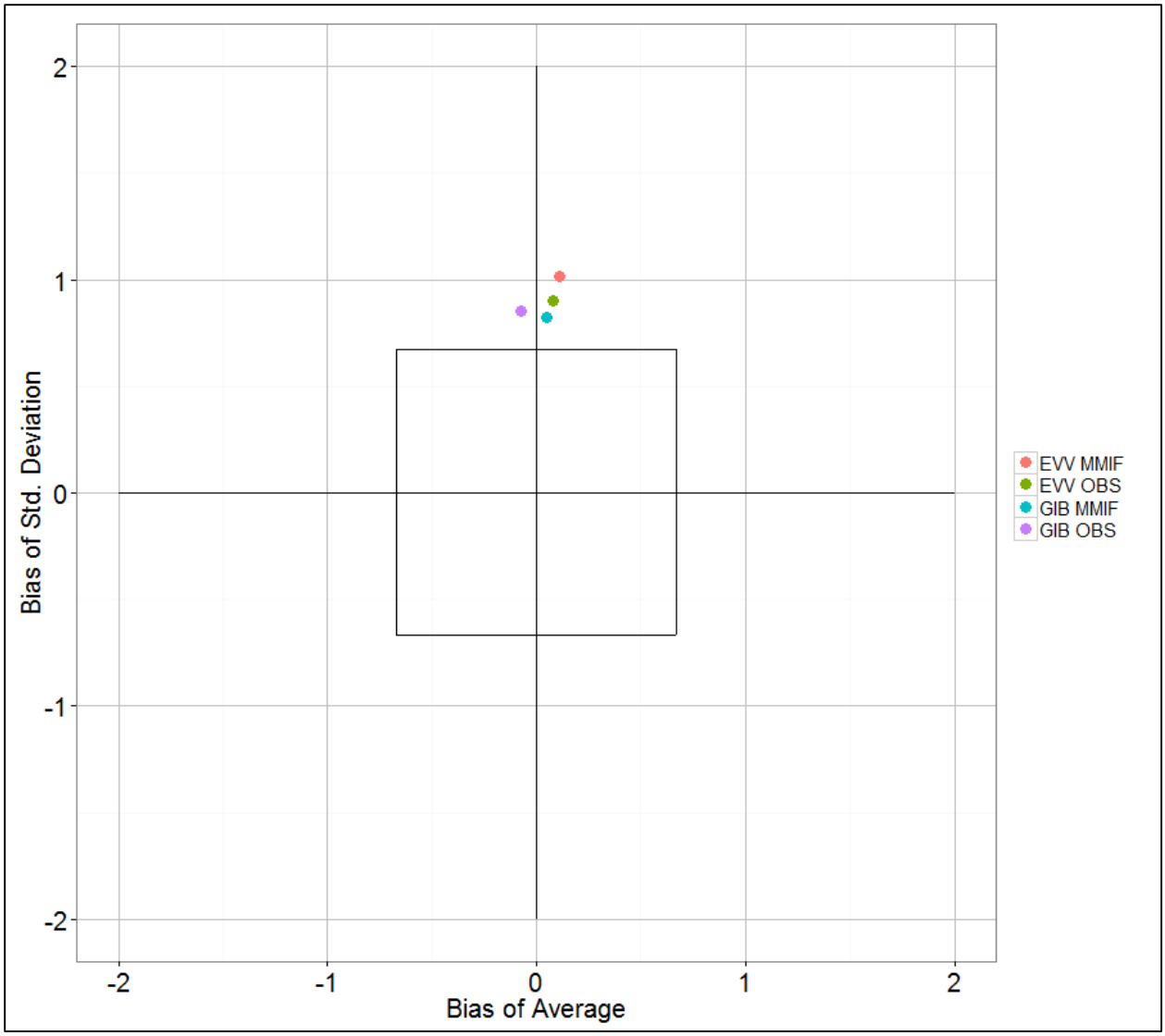

Figure 15. Gibson 3-hour screening results.

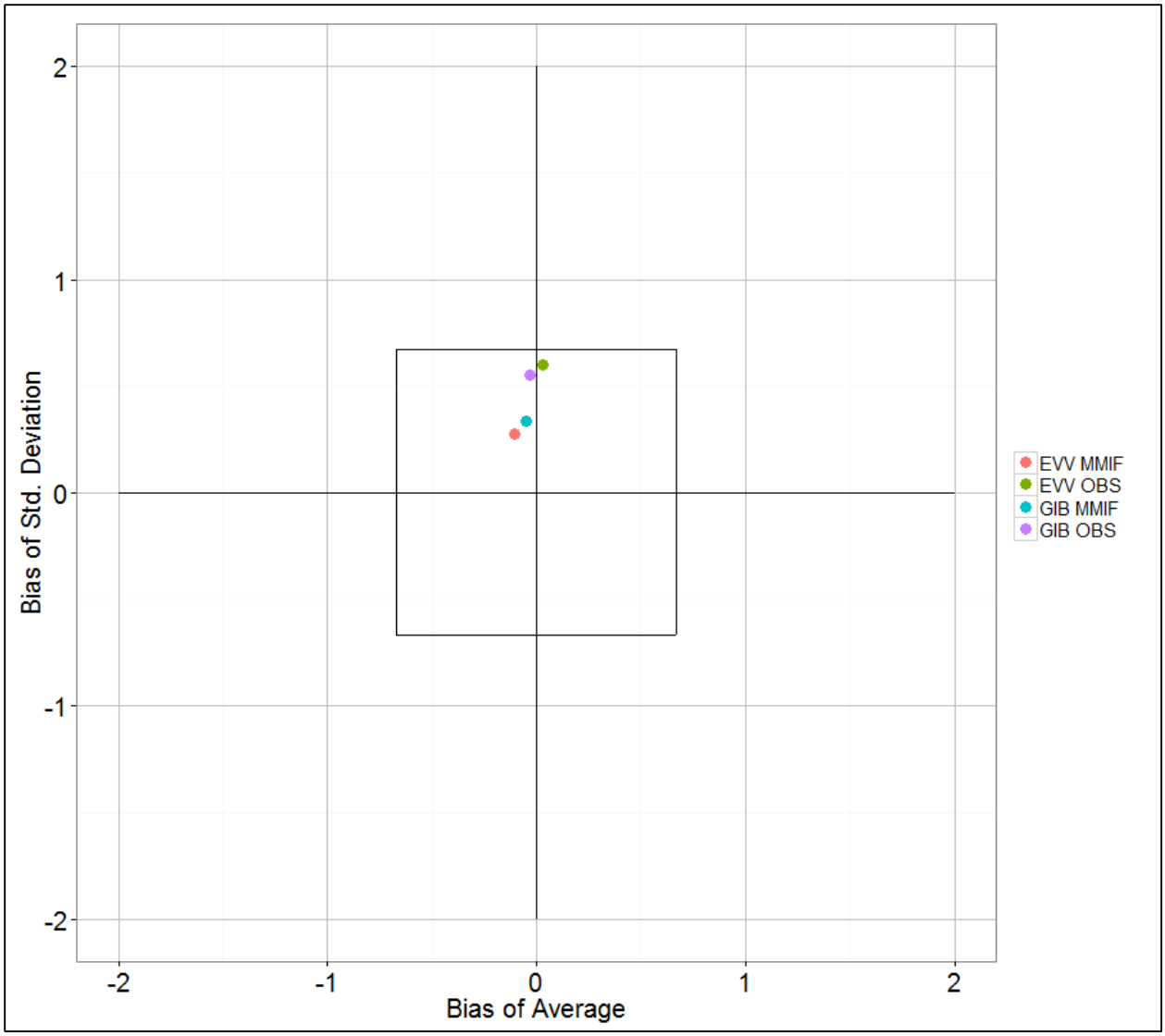

Figure 16. Gibson 24-hour screening results.

Figure 17 compares the 1-hour, 3-hour, and 24-hour fractional biases and a composite of the fraction biases (Equation 3 using fractional biases instead of absolute fractional biases). Shown are the fractional biases from the original data along with the 5th and 95th percentiles of the fractional biases from the bootstrap results. Figure 17.d shows the modified CPM from the original data with the 5th and 95th percentiles of the modified CPM from the bootstrap results. In the plots, positive (negative) fractional biases indicate model under-prediction (over-prediction). Based on the distributions, for 1-hour and 3-hour, all four scenarios tend to under-predict compared to observations. For the 24-hour periods, the models tend to show more over-prediction. The composite shows mostly under-prediction, with the four scenarios showing comparable distributions.

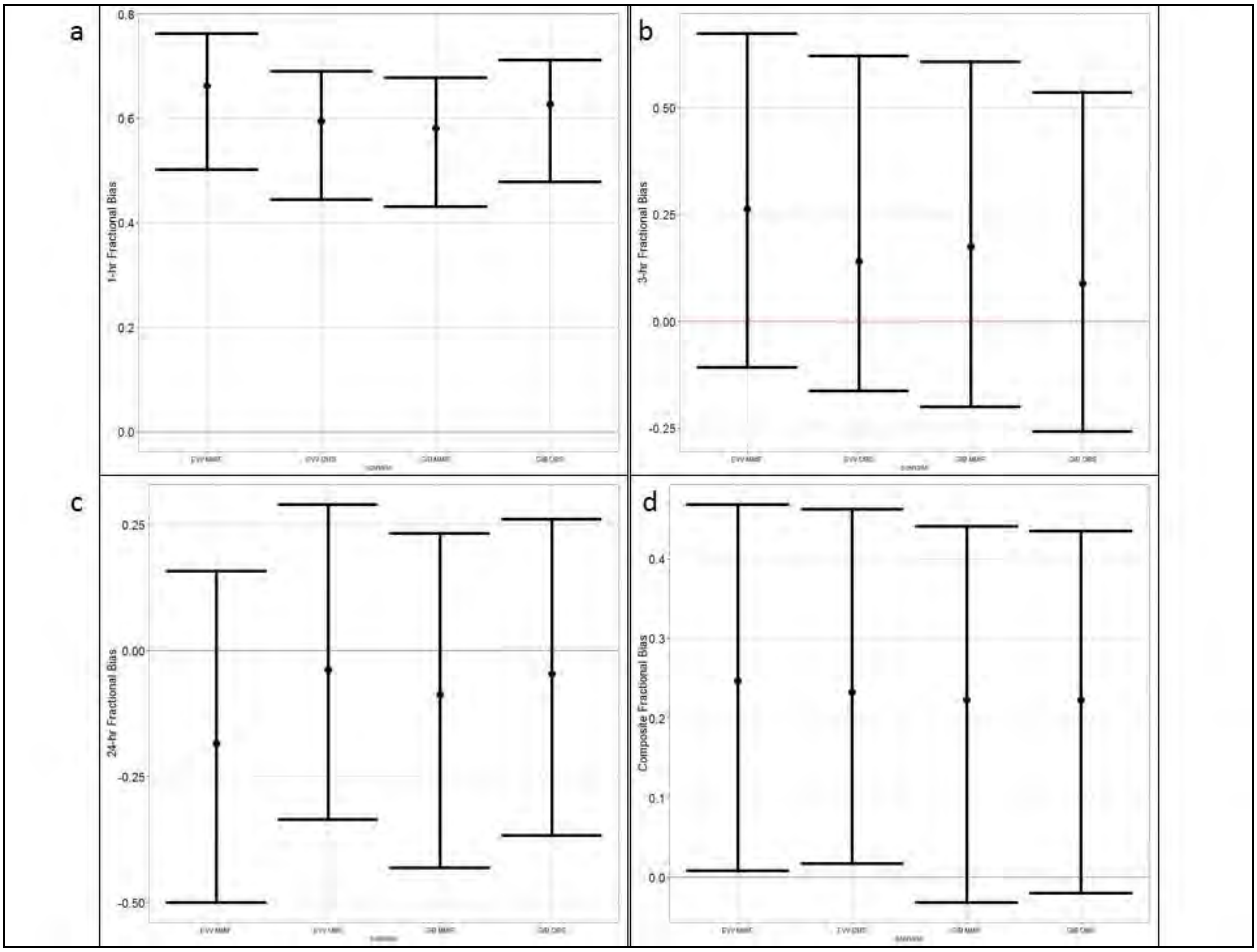

Figure 17. Gibson fractional biases for a) 1-hour, b) 3-hour, c) 24-hour, and d) CPM based on fractional biases.

Table 7 shows the AFB values for the 1-hour, 3-hour, and 24-hour periods as well as the CPM statistics for each meteorological scenario. Figure 18 shows the CPM values with the 5th and 95th percentiles of the CPM's from the bootstrap results. Based on the results shown in Table 7 and Figure 18 the EVV MMIF results were somewhat different compared to the other three scenarios which had similar CPM values.

Table 7. 1-hour, 3-hour, 24-hour absolute fractional biases and composite performance measures for Gibson meteorological scenarios.

Scenario	AFB			CPM
	1-hour	3-hour	24-hour	
EVV MMIF	0.77	0.26	0.19	0.41
EVV OBS	0.75	0.14	0.04	0.31
GIB MMIF	0.77	0.17	0.09	0.34
GIB OBS	0.90	0.09	0.05	0.35

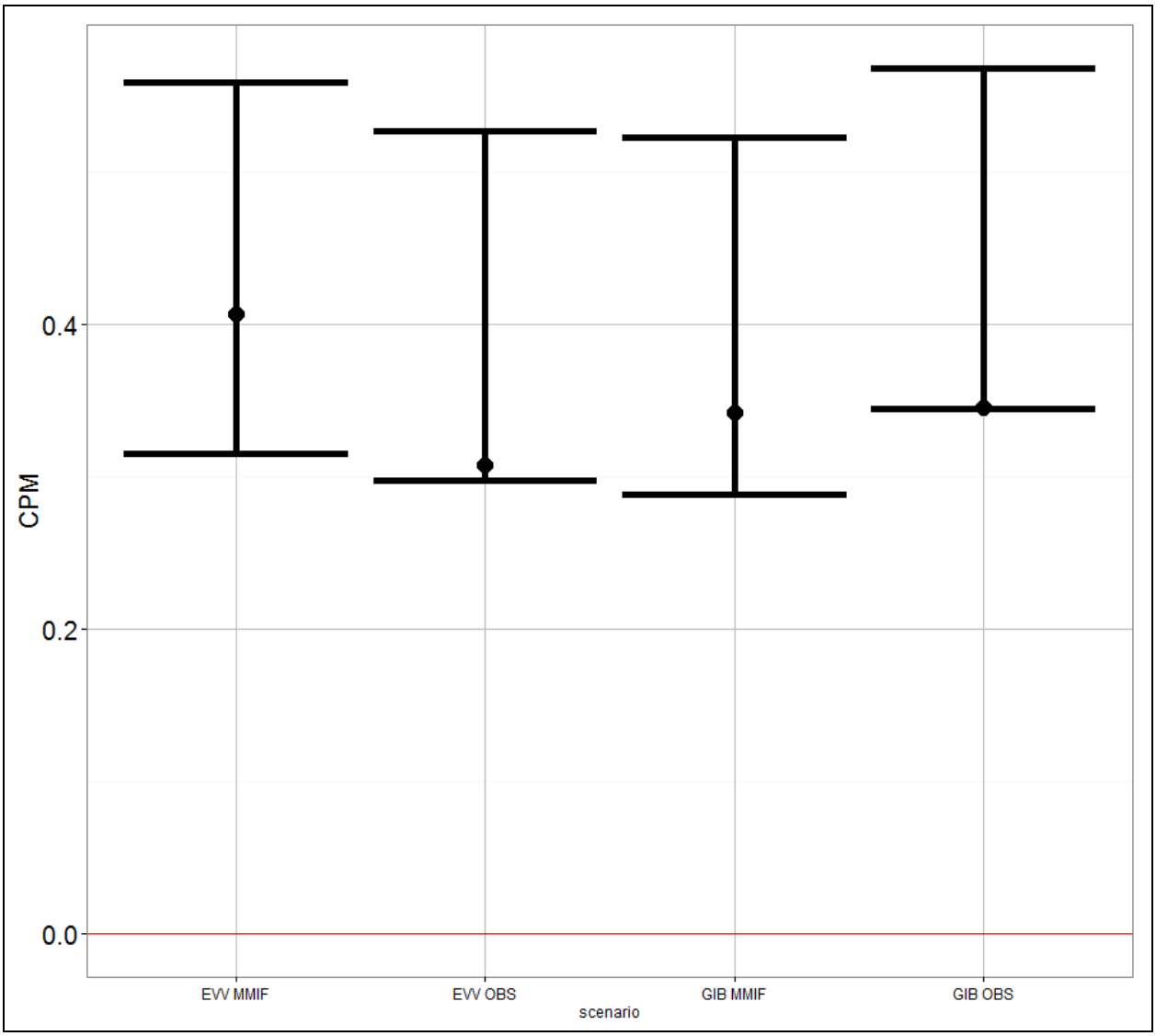

Figure 18. Gibson composite performance metric values with 5th and 95th percentiles of the CPM values from the bootstrap results.

Table 8 lists the MCM values for the six model comparisons and also lists the best performing model for each pair. As discussed in Section 2.2, positive (negative) values of an MCM indicate that the second scenario of the difference (scenario to the right of the "-" sign) is performing better (worse) than the first scenario (scenario to the left of the "-" sign). Figure 19 show the MCM differences among the four scenarios at the 90th and 95th confidence intervals respectively. . The results indicate that the EVV MMIF meteorology performs slightly worse when compared to the other three scenarios. The results also indicate that the GIB MMIF scenario performs worse against EVV OBS but is almost equal to the GIB OBS scenario and that the GIB OBS scenario performs worse than the EVV OBS scenario. While there is differences in performance, the confidence intervals for each difference overlap zero, indicating that differences in performance are not statistically significant. Of the two MMIF scenarios, the GIB MMIF appears to be the better scenario, which is not surprising given that the cell contains the facility.

Table 8. Model comparison measures (MCM) for the four Gibson meteorological scenarios.

MCM Scenario	MCM	Best performing scenario
EVV MMIF – EVV OBS	0.1	EVV OBS
EVV MMIF – GIB OBS	0.06	GIB OBS
GIB MMIF – EVV MMIF	-0.06	GIB MMIF
GIB MMIF – EVV OBS	0.03	EVV OBS
GIB MMIF – GIB OBS	-0.004	GIB MMIF
GIB OBS – EVV OBS	0.04	EVV OBS

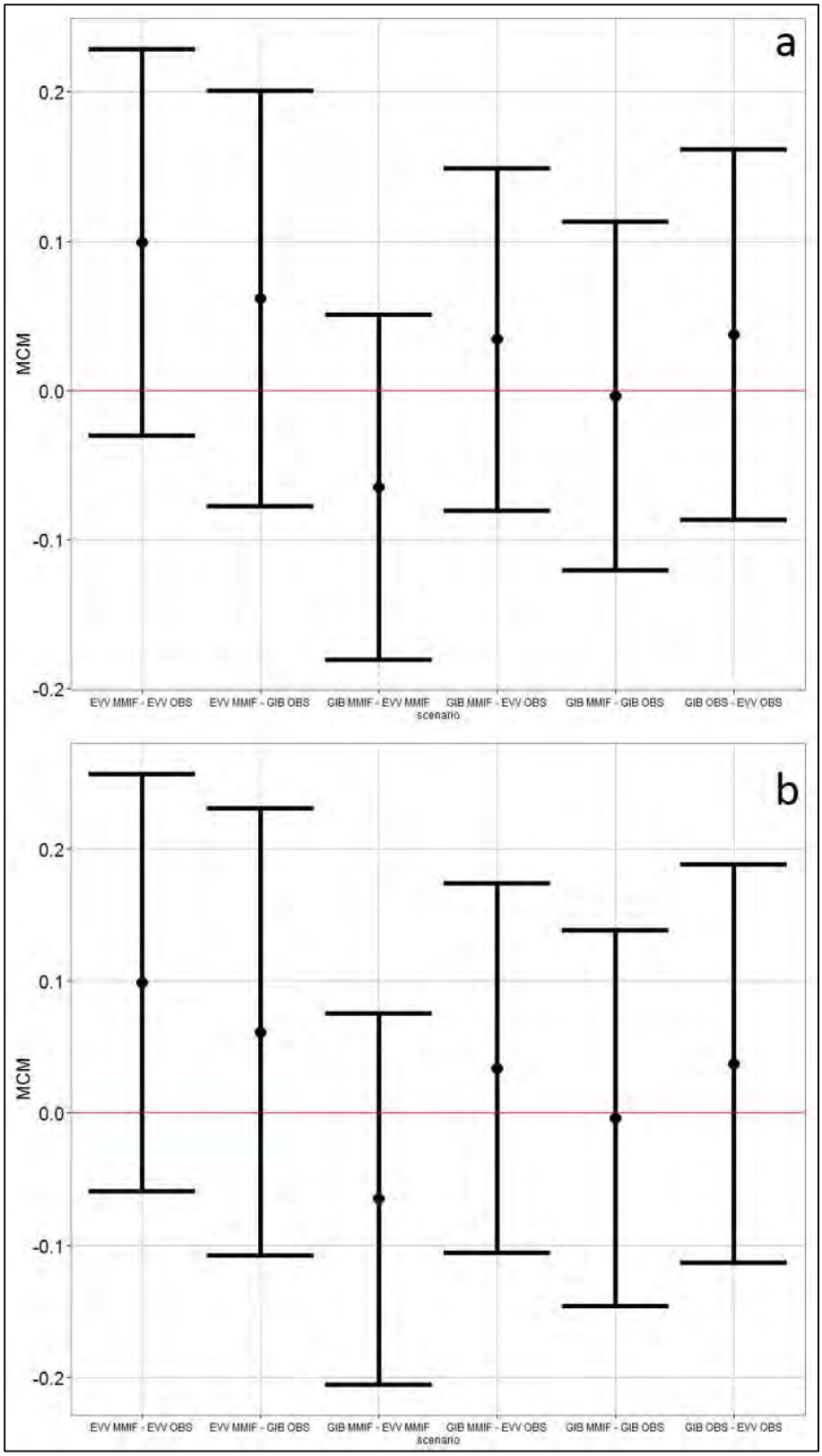

Figure 19. Gibson MCM differences with a) 90th percentile and b) 95th confidence intervals.

3.2 Martins Creek

3.2.1 Meteorological data comparisons

The wind roses for the four meteorological scenarios are presented in Figure 20. The Martins Creek site-specific tower exhibits a strong southwest to northeast pattern. The two MMIF wind roses do exhibit a southwest to northeast pattern, they also have more hours from other directions. The ABE rose shows a strong westerly component to its winds with a secondary maximum from the northeast.

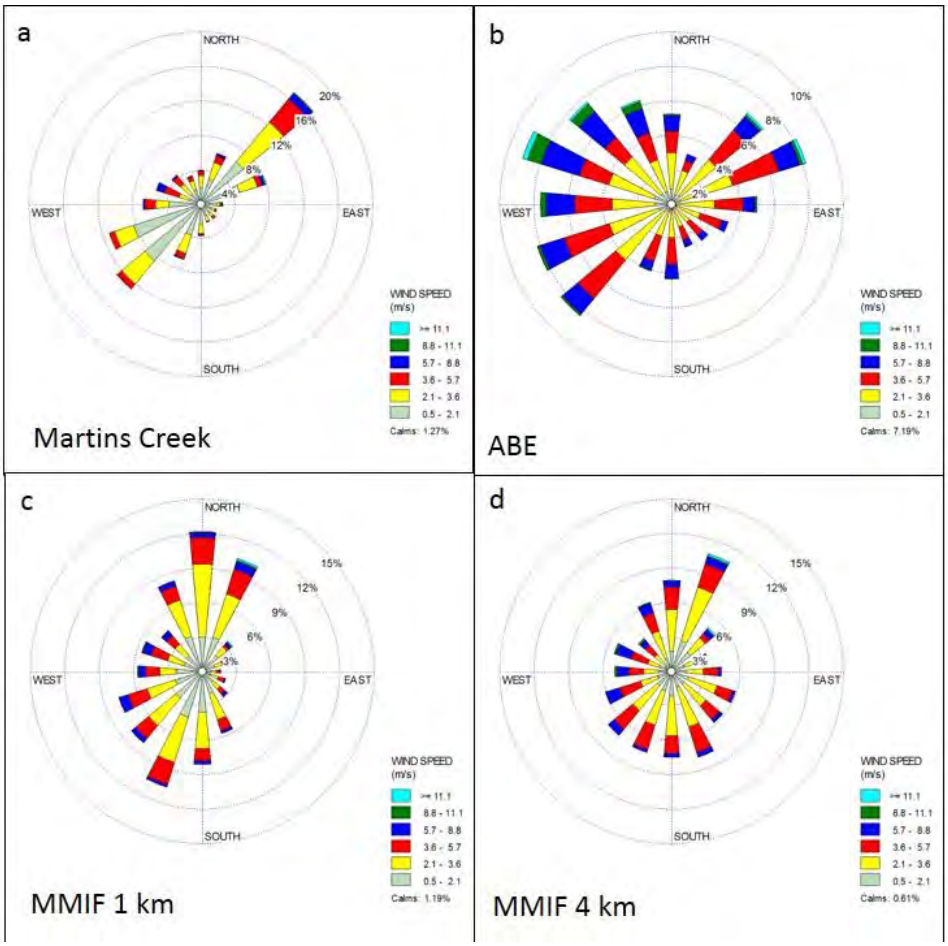

Figure 20. May 1992 – May 1993 Martins Creek wind roses for a) Martins Creek, b) ABE, c) MMIF 1 km, and d) MMIF 4 km.

Figure 21 shows the wind displacement distributions among the scenarios. For the most part, displacements between the scenarios is less than 10 km. Figure 22 through Figure 24 compare the surface roughness values by month and 10º sector. ABE has the lower roughness values throughout the year with Martins Creek spiking between 180º and 260º due to the presence of a stand of trees from that direction.

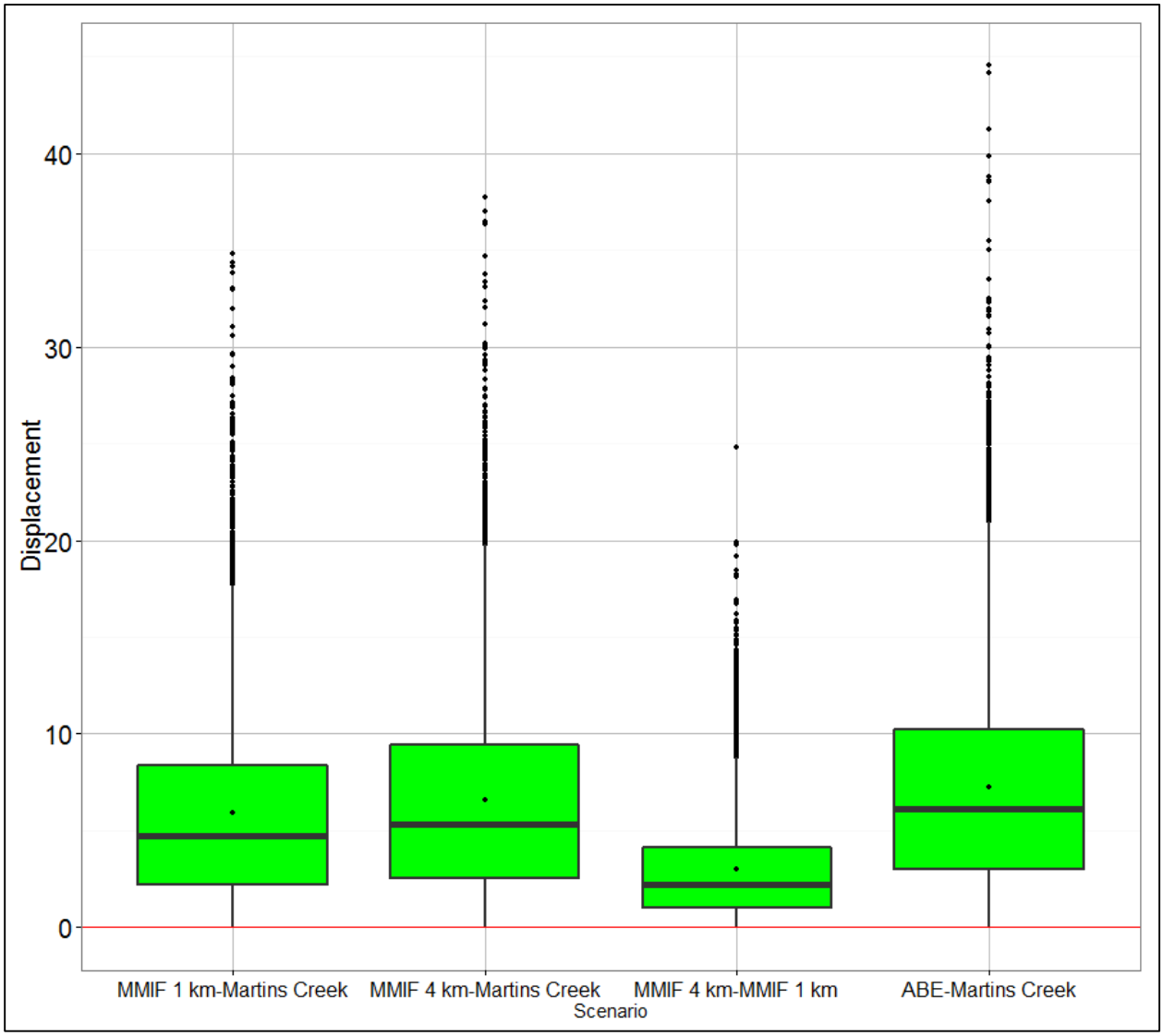

Figure 21. Wind displacement (km) among the Martins Creek meteorological scenarios.

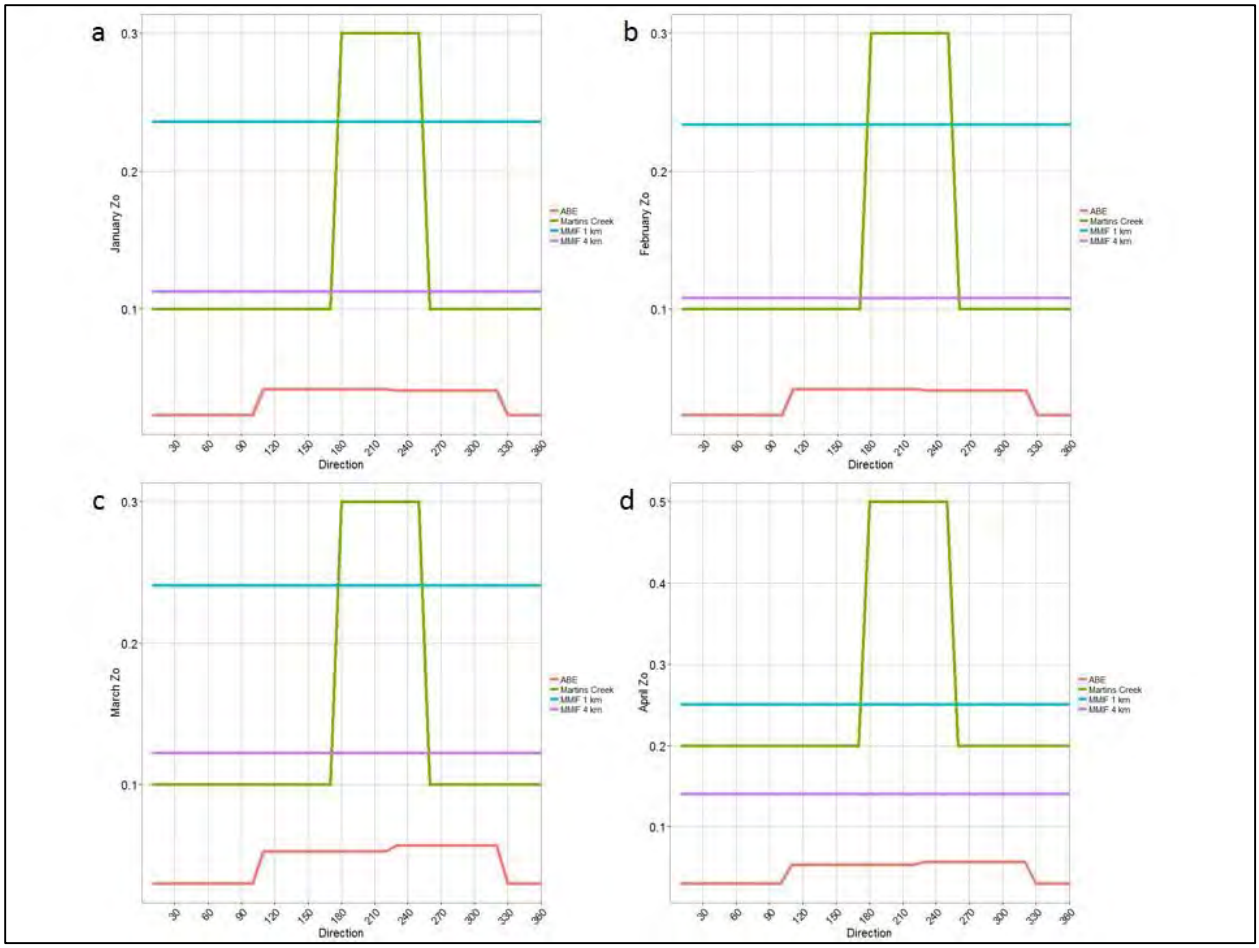

Figure 22. Martins Creek study monthly surface roughness lengths (m) by 10 degree sectors for a) January, b) February, c) March, and d) April.

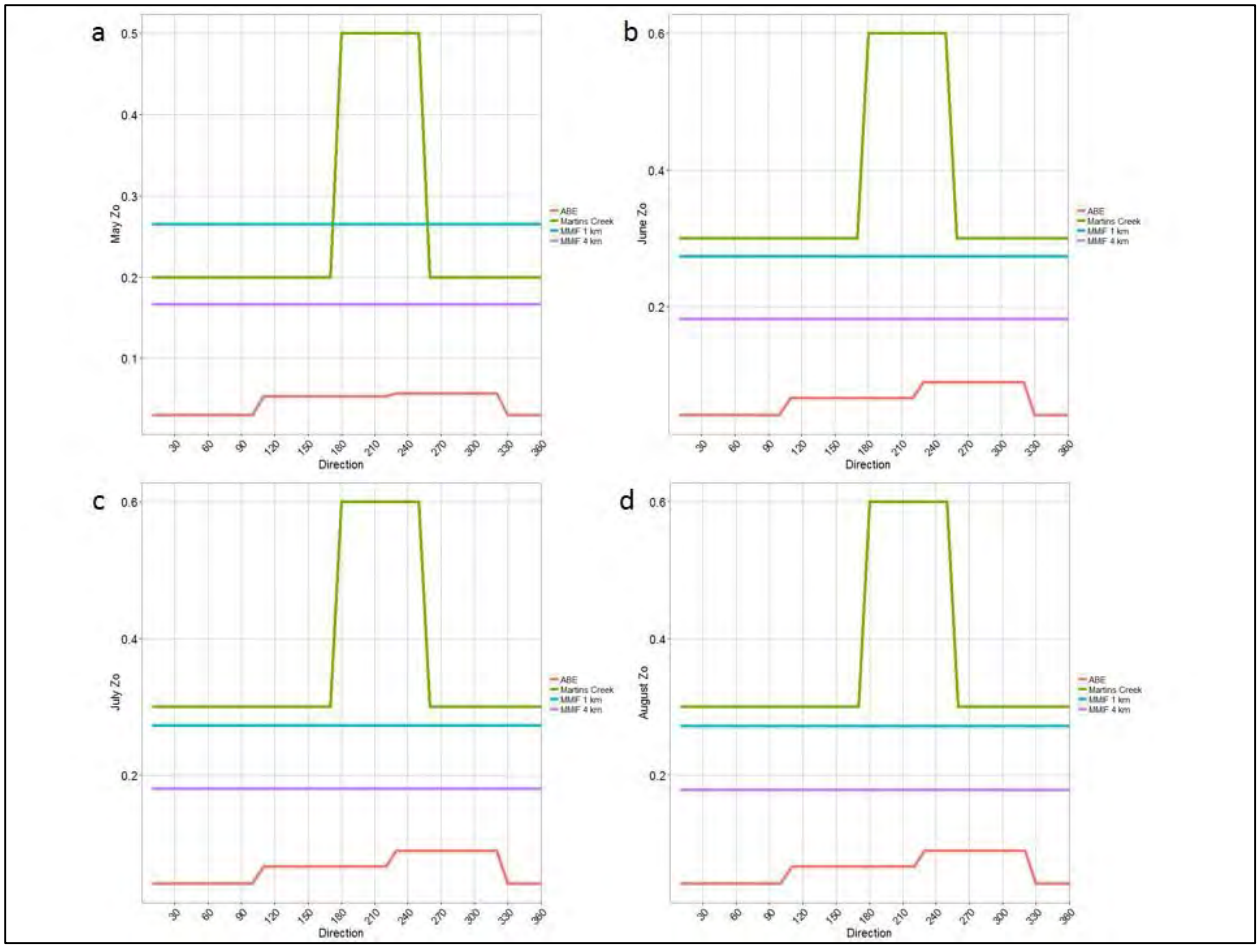

Figure 23. Martins Creek study monthly surface roughness lengths (m) by 10 degree sectors for a) May, b) June, c) July, and d) August.

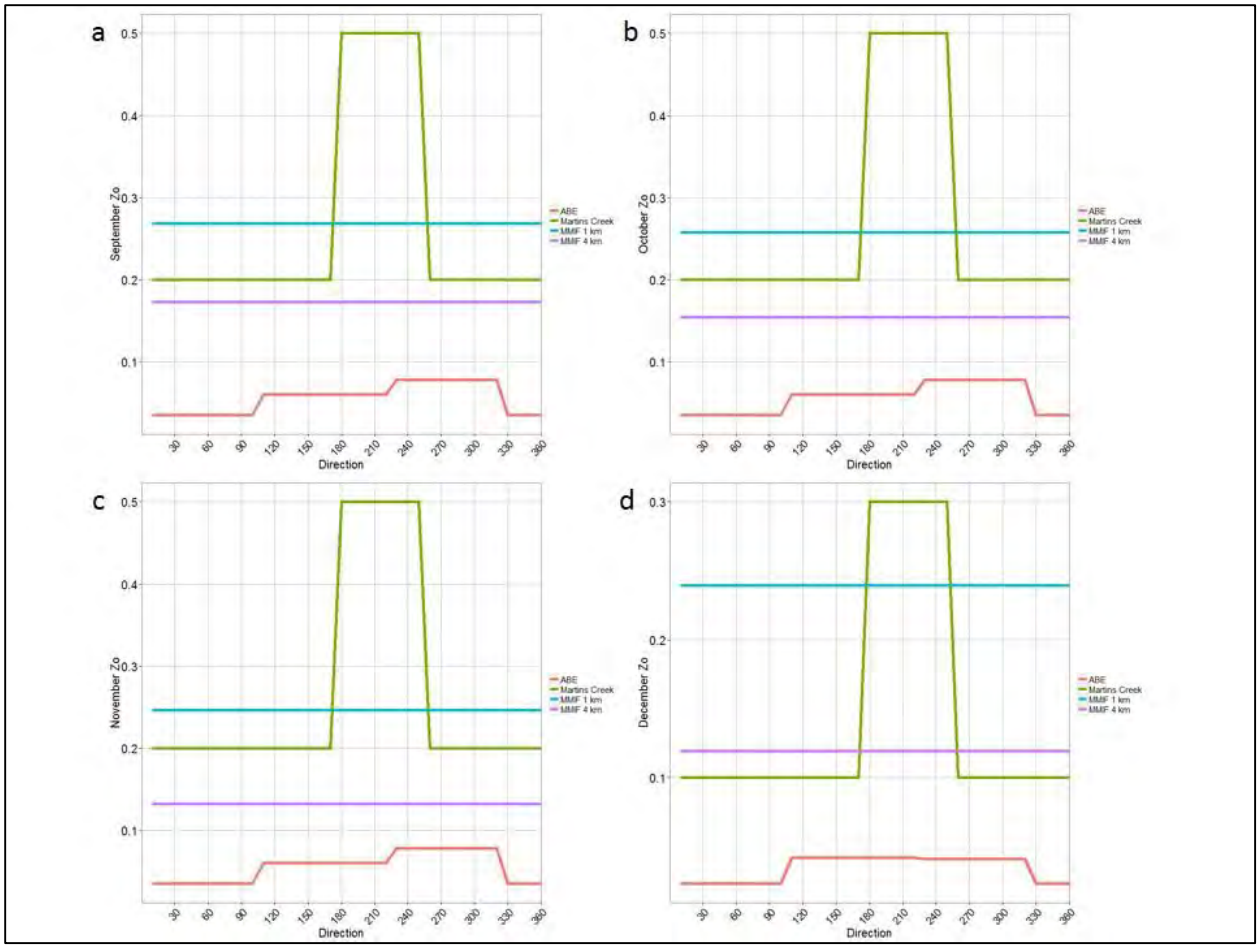

Figure 24. Martins Creek study monthly surface roughness lengths (m) by 10 degree sectors for a) September, b) October, c) November, and d) December.

Tables 9 through 11 present statistics for meteorological variables for the four scenarios. The control difference, ABE – Martins Creek is highlighted in yellow. Box and whisker plots of the variable distributions as well as bias distributions can be found in Appendix A.

- Differences between MMIF wind speeds and observed wind speeds are lower than the differences between the two observed datasets. Both MMIF scenarios over-predict when compared to the Martins Creek winds. The MMIF 4 km simulation over-predicts compared to the 1 km simulation.
- The MMIF scenarios and ABE over-predict temperatures with the MMIF scenarios having less over-prediction.
- For pressures, the MMIF 1 km scenario over-predicts while the MMIF 4 km over-predicts. ABE and Martins Creek do not differ, most likely due to ABE RH values being used at Martins Creek.

- MMIF scenarios over-predict relative humidity and the MMIF 4 km scenario slightly under-predicts compared to MMIF 1 km. ABE and Martins Creek do not differ, most likely due to ABE RH values being used at Martins Creek.
- For heat flux the MMIF 1 km scenario under-predicts while the MMIF 4 km scenario over-predicts as does ABE.
- For u*, both MMIF scenarios and ABE over-predict and are comparable to each other.
- For w*, the MMIF 1 km scenario under-predicts while the MMIF 4 km and ABE over-predict.
- For the potential temperature gradient, the MMIF scenarios under-predict while ABE over-predicts but values are relatively close.
- For mixing heights (both convective and mechanical), the MMIF scenarios and ABE over-predict.
- For Monin-Obukhov length, the MMIF 1 km scenario under-predicts while the MMIF 4 km scenario and ABE over-predict. The MMIF 4 km biases appear to be an outlier when compared to the MMIF 1 km biases and ABE biases.
- Cloud cover differences also show relatively low agreement but this may be due to the calculation methodology in AERMET when cloud cover is missing. ABE and Martins Creek do not differ, most likely due to ABE cloud cover being used for Martins Creek.

Overall, while there are differences, the MMIF scenarios appear to show relatively good agreement with the observed data. Differences between the MMIF scenarios and Martins Creek counterparts is usually in line with the differences in the site-specific data and NWS data (i.e., ABE – Martins Creek).

Table 9. Mean bias, fractional bias, root mean square error, and R^2 for primary meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Wind speed	MMIF 1 km-Martins Creek	0.58962	0.05979	1.36357	0.53139
	MMIF 4 km -Martins Creek	1.06006	0.09979	1.66366	0.54214
	MMIF 4 km -MMIF 1 km	0.47061	0.04196	0.88542	0.83952
	ABE-Martins Creek	1.44933	0.09314	2.17893	0.52709
Ambient temperature	MMIF 1 km -Martins Creek	0.23995	0.00021	2.59116	0.92888
	MMIF 4 km -Martins Creek	0.31796	0.00028	2.42705	0.93764
	MMIF 4 km -MMIF 1 km	0.07750	0.00007	0.62537	0.99582
	ABE-Martins Creek	0.77481	0.00068	1.89823	0.96915
Pressure	MMIF 1 km -Martins Creek	3.24523	0.00081	3.41835	0.98217
	MMIF 4 km -Martins Creek	-1.83670	-0.00046	2.11053	0.98365
	MMIF 4 km -MMIF 1 km	-5.06872	-0.00126	5.09355	1.0
	ABE-Martins Creek	0	0	0	1.0
Relative humidity	MMIF 1 km -Martins Creek	7.73760	0.02848	16.10607	0.55156
	MMIF 4 km -Martins Creek	7.61139	0.02804	15.63945	0.57766
	MMIF 4 km -MMIF 1 km	-0.12598	-0.00049	3.84615	0.96119
	ABE-Martins Creek	0	0	0	1.0

Table 10. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Heat flux	MMIF 1 km-Martins Creek	-10.60956	0.02594	39.27118	0.48069
	MMIF 4 km -Martins Creek	15.76726	0.09672	48.34657	0.75046
	MMIF 4 km -MMIF 1 km	26.36619	0.07375	62.95421	0.56546
	ABE-Martins Creek	19.72457	0.19548	52.67964	0.89900
u*	MMIF 1 km -Martins Creek	0.07382	0.06145	0.15689	0.62727
	MMIF 4 km -Martins Creek	0.09468	0.09864	0.15717	0.62565
	MMIF 4 km -MMIF 1 km	0.02140	0.03829	0.09022	0.85083
	ABE-Martins Creek	0.09746	0.11097	0.16210	0.57773
w*	MMIF 1 km -Martins Creek	-0.03219	-0.02176	0.41303	0.38025
	MMIF 4 km -Martins Creek	0.31158	0.05245	0.52603	0.56513
	MMIF 4 km -MMIF 1 km	0.33947	0.07410	0.49497	0.70120
	ABE-Martins Creek	0.37888	0.08166	0.46271	0.86686
$d\Theta/dz$	MMIF 1 km -Martins Creek	-0.00149	-0.03543	0.00430	0.21148
	MMIF 4 km -Martins Creek	-0.00151	-0.03592	0.00430	0.21398
	MMIF 4 km -MMIF 1 km	-0.00003	-0.00084	0.00087	0.89759
	ABE-Martins Creek	0.00009	0.00311	0.00329	0.51953

Table 11. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Zi_c	MMIF 1 km-Martins Creek	205.20000	0.05694	567.66350	0.27664
	MMIF 4 km -Martins Creek	226.59382	0.06771	526.93158	0.36883
	MMIF 4 km -MMIF 1 km	18.69463	0.01236	216.79283	0.87216
	ABE-Martins Creek	267.46126	0.08113	393.69289	0.81594
Zi_m	MMIF 1 km -Martins Creek	158.29334	0.10105	371.44030	0.48209
	MMIF 4 km -Martins Creek	167.37733	0.13701	353.60026	0.48169
	MMIF 4 km -MMIF 1 km	10.09504	0.03481	145.49705	0.90391
	ABE-Martins Creek	190.78705	0.15120	356.34586	0.55482
L	MMIF 1 km -Martins Creek	-8.03602	0.10690	2087.77469	0.07088
	MMIF 4 km -Martins Creek	360.81687	0.12134	2447.01436	0.04617
	MMIF 4 km -MMIF 1 km	377.92163	-0.00433	1819.59481	0.51254
	ABE-Martins Creek	42.92637	0.09176	667.80236	0.25921
Cloud cover	MMIF 1 km -Martins Creek	0.17313	-0.01996	5.05242	0.08225
	MMIF 4 km -Martins Creek	0.56836	0.00469	4.97054	0.09158
	MMIF 4 km -MMIF 1 km	0.39765	0.03361	3.05554	0.54030
	ABE-Martins Creek	0	0	0	1.0

3.2.2 AERMOD results

QQ plots for various averaging times are shown in Figure 25. For the 1-hour averages (Figure 25.a), ABE tends to over-predict outside of the factor of 2 line except for the highest concentrations. The other scenarios tend to slightly over predict below 1,000 μg/m^3, with the MMIF 4 km showing better agreement with the observed concentrations. For the highest concentrations, the MMIF 4 km and Martins Creek under-predict while the other two scenarios over-predict. All are within a factor of two of the observations.

For the 3-hour averages (Figure 25.b), all scenarios over-predict below 600 μg/m^3. Again, ABE tends to be the higher of the concentrations. As with the 1-hour concentrations, at the higher concentrations, the MMIF 4 km and Martins Creek under-predict. For the 24-hour averages (Figure 25.c), the model values under-predict between 25 and 50 μg/m^3 with over-prediction at concentrations higher than 50 μg/m^3. For the annual averages (Figure 25.d), all scenarios under-predict. Figure 26 and Figure 27 show the 3-hr and 24-hour screening results. Both show that the four scenarios over-predict at 3-hour and 24-hour averages but are within the factor of two limit. Figure 28 shows the fractional biases for 1-hour, 3-hour, and 24-hour as well as a CPM based on those values. ABE tends show over-predictions at all averages, agreeing with the QQ-plots. Fractional biases for Martins Creek and the two MMIF scenarios over-predict but for 1-

hour they do cross over zero, indicating under-prediction. All scenarios over-predict at the 24-hour average and the CPM plot shows that.

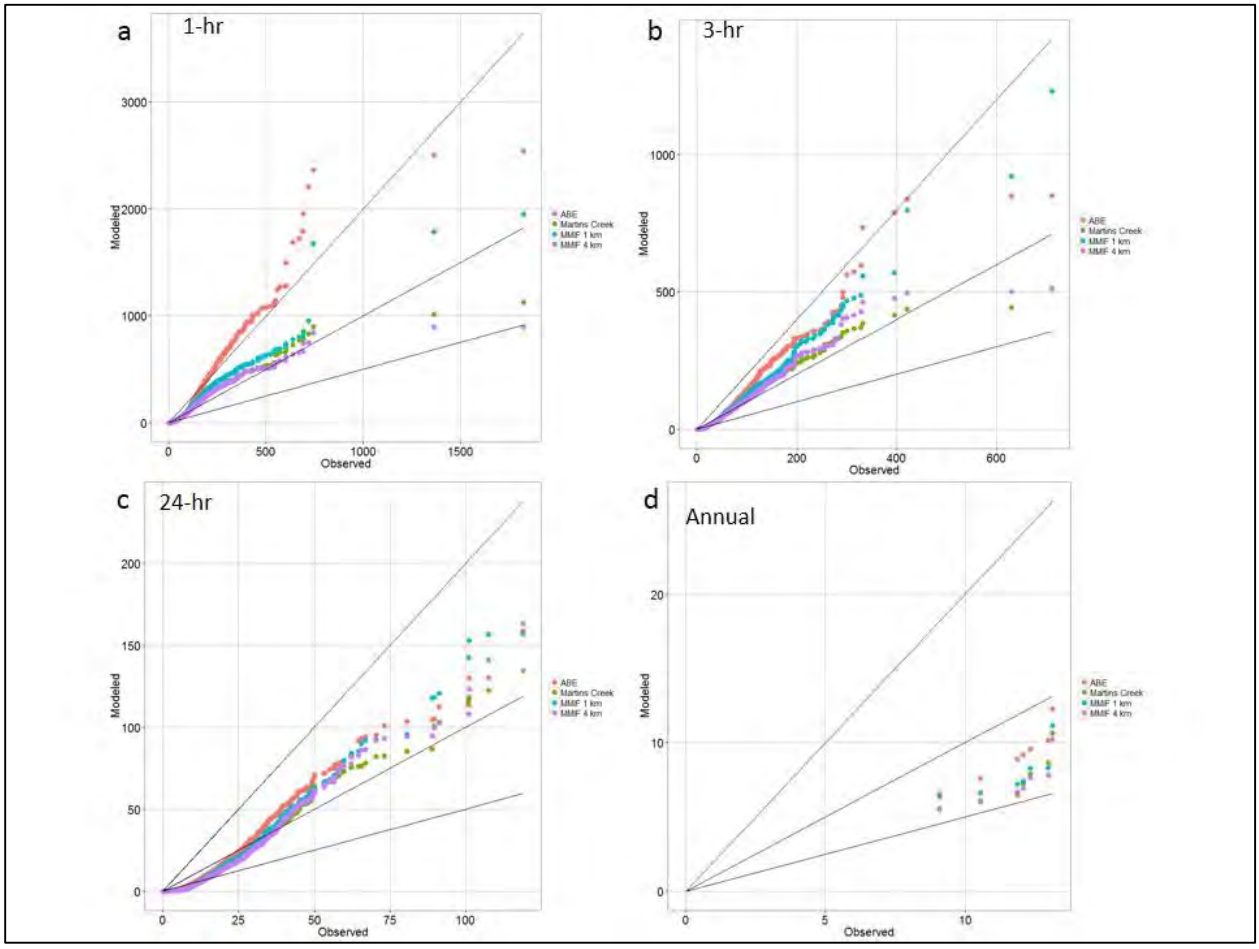

Figure 25. Martins Creek QQ plots for a) 1-hour, b) 3-hour, c) 24-hour, and d) annual averages. Concentrations are in μg/m³.

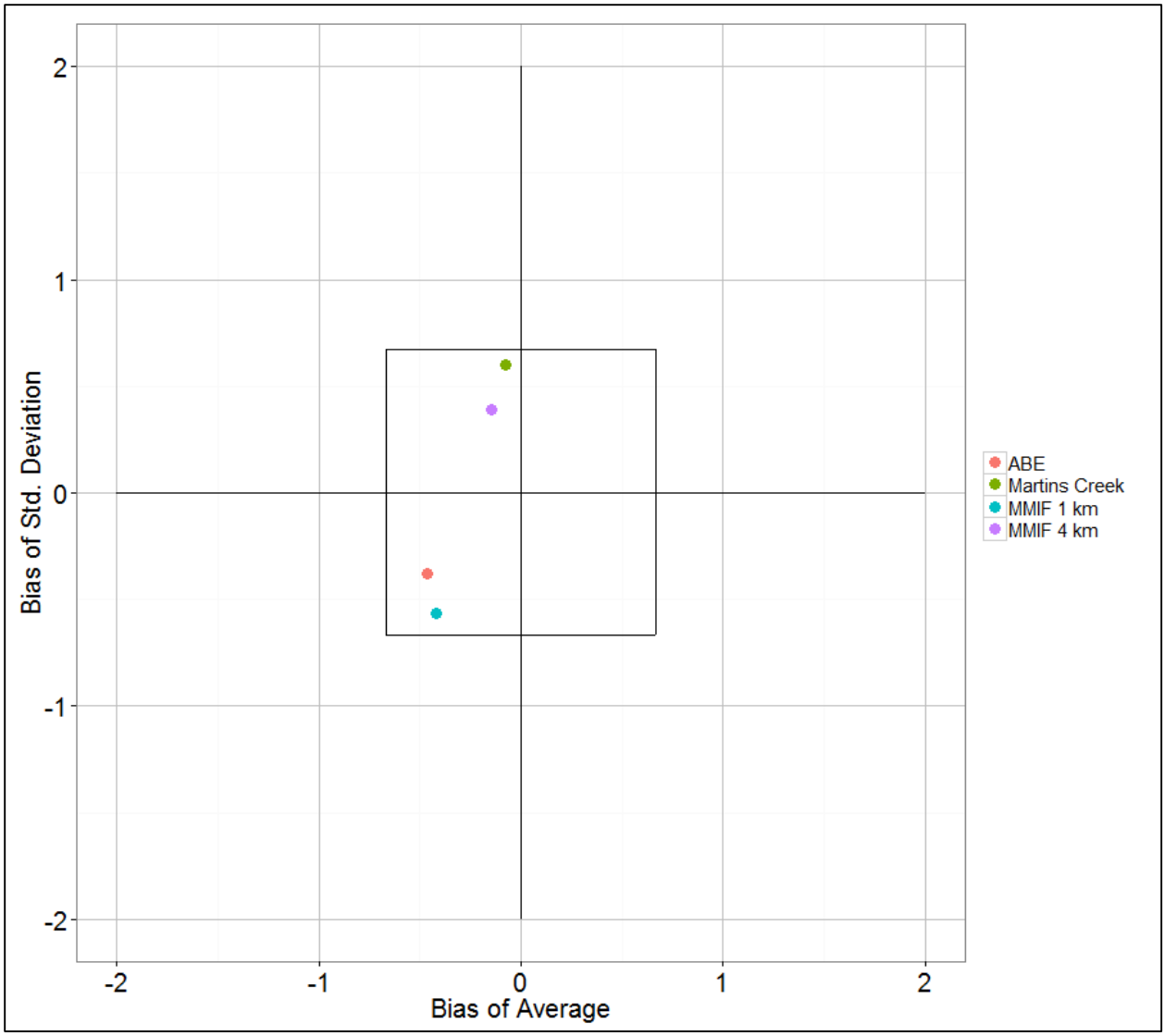

Figure 26. Martins Creek 3-hour screening results.

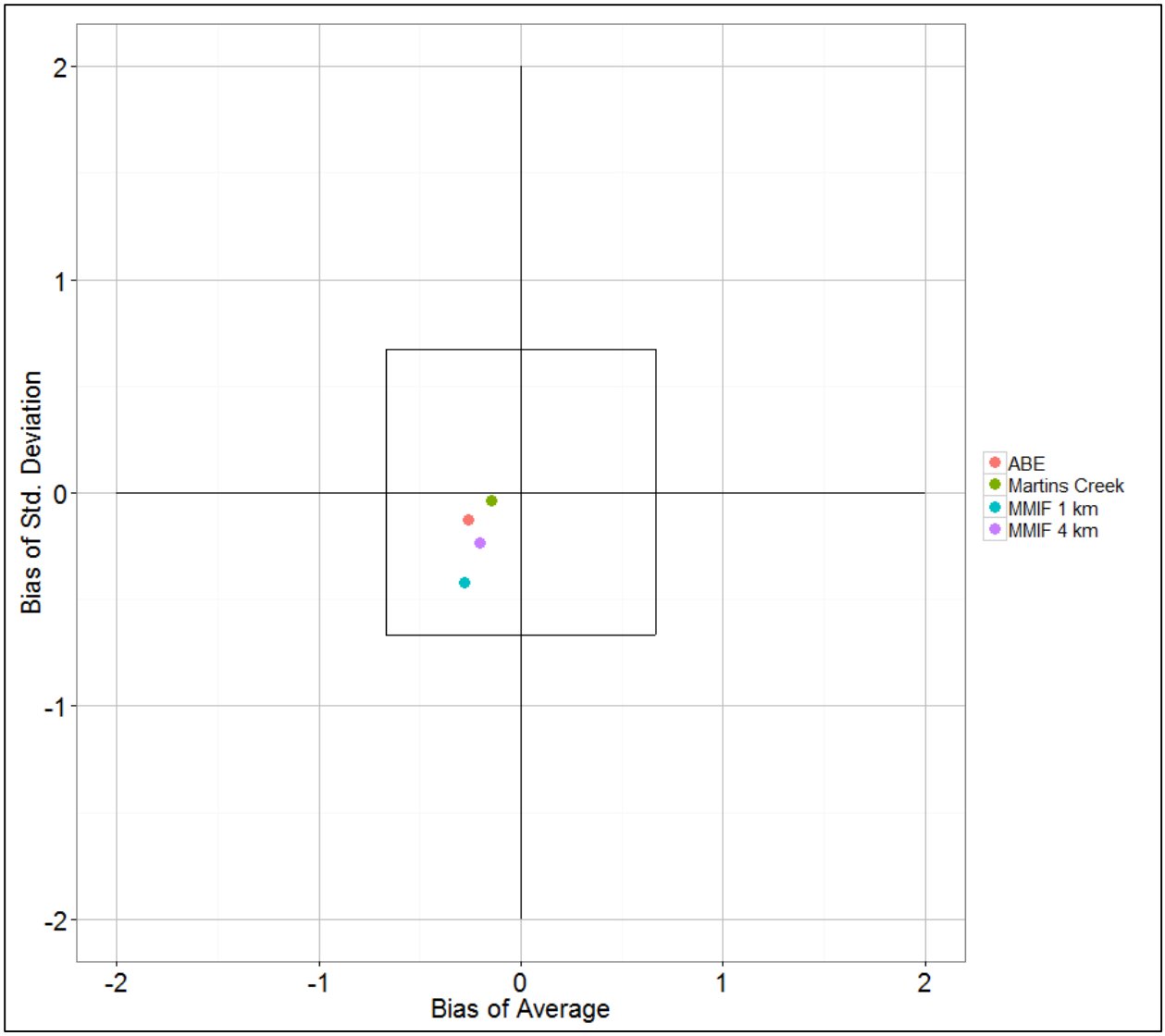

Figure 27. Martins Creek 24-hour screening results.

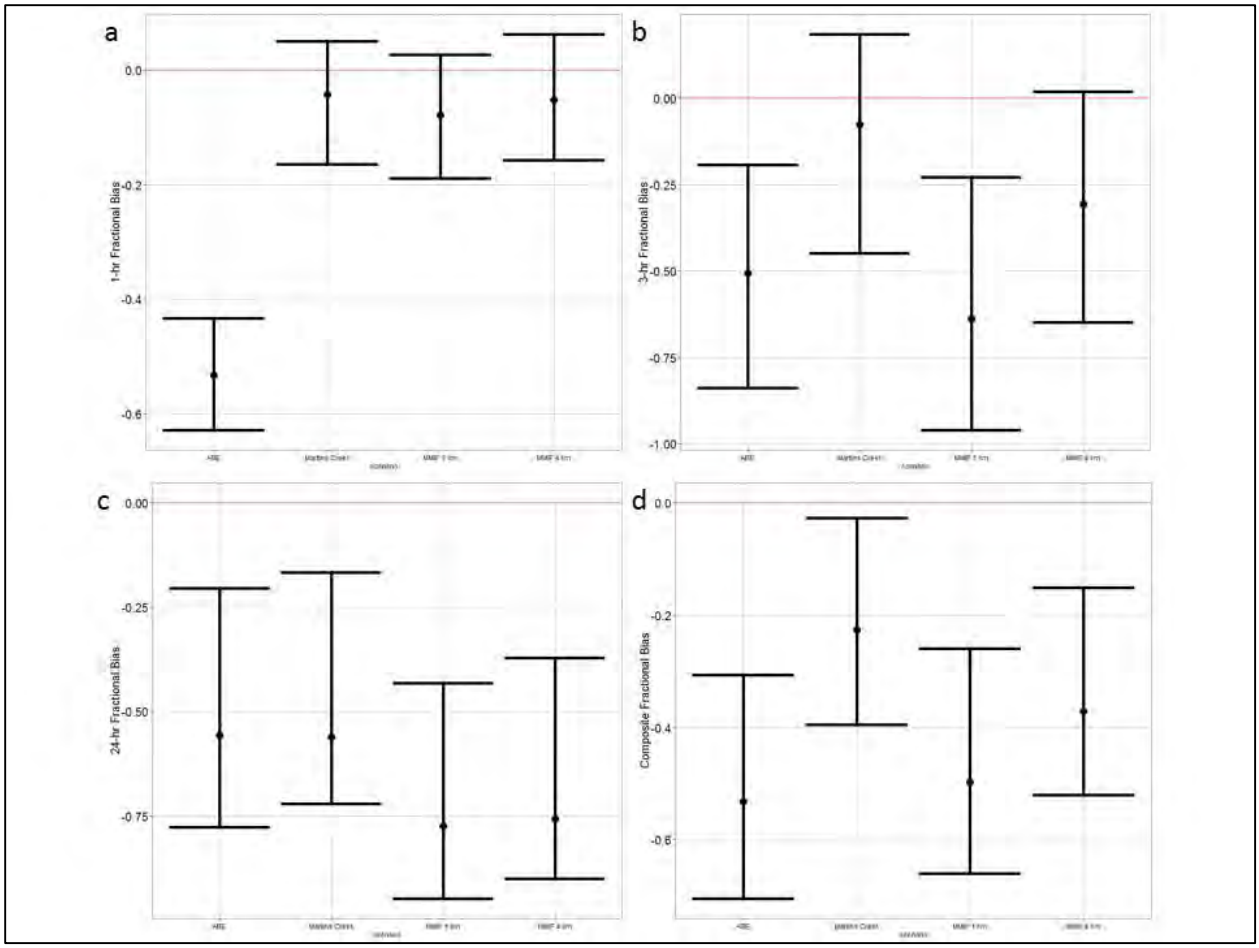

Figure 28. Martins Creek fractional biases for a) 1-hour, b) 3-hour, c) 24-hour, and d) CPM based on fractional biases.

Table 12 shows the absolute fractional biases and CPM for each scenario for Martins Creek and Figure 29 shows the CPM values with the 5th and 95th percentiles of the bootstrap results. The CPM values show that Martins Creek site-specific data performed better with the other three having comparable values. Table 13 shows the MCM values for the different model pair differences and Figure 30 shows the MCM values with the 90th and 95th confidence intervals. Based on Table 13, Martins Creek tends to be the better performing scenario. The plots in Figure 30 show that the differences between the MMIF simulations and Martins Creek are statistically significant.

Table 12. 1-hour, 3-hour, 24-hour absolute fractional biases and composite performance measures for Martins Creek meteorological scenarios.

Scenario	AFB			CPM
	1-hour	3-hour	24-hour	
ABE	0.78	0.51	0.56	0.61
Martins Creek	0.40	0.08	0.56	0.34
MMIF 1 km	0.56	0.64	0.77	0.66
MMIF 4 km	0.50	0.31	0.76	0.52

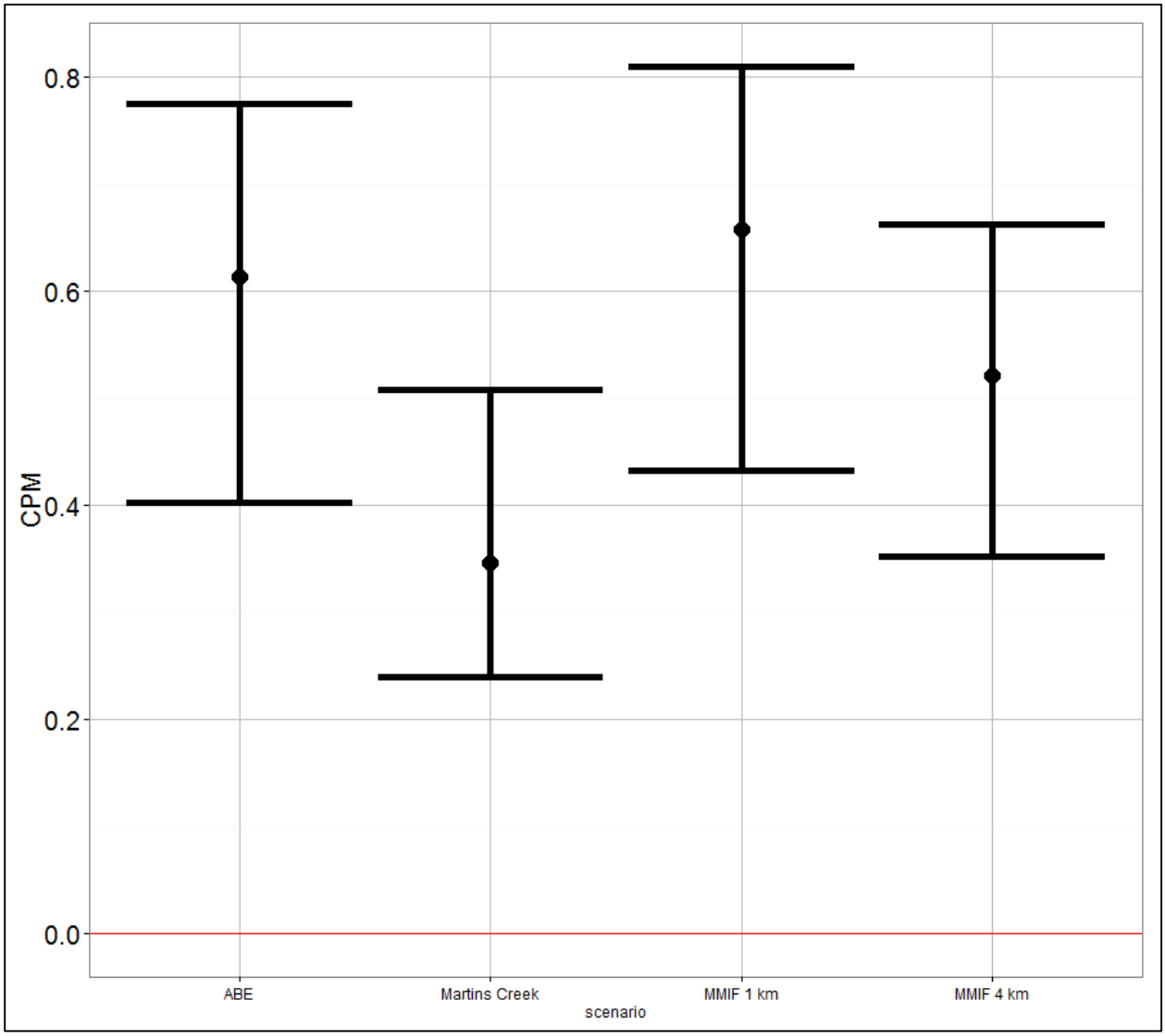

Figure 29. Martins Creek composite performance metric values with 5th and 95th percentiles of the CPM values from the bootstrap results.

Table 13. Model comparison measures (MCM) for the four Gibson meteorological scenarios.

MCM Scenario	MCM	Best performing scenario
ABE - Martins Creek	0.27	Martins Creek
MMIF 1 km – ABE	0.05	ABE
MMIF 1km – Martins Creek	0.31	Martins Creek
MMIF 4 km – ABE	-0.09	MMIF 4 km
MMIF 4 km – Martins Creek	0.18	Martins Creek
MMIF 4 km – MMIF 1km	-0.14	MMIF 1 km

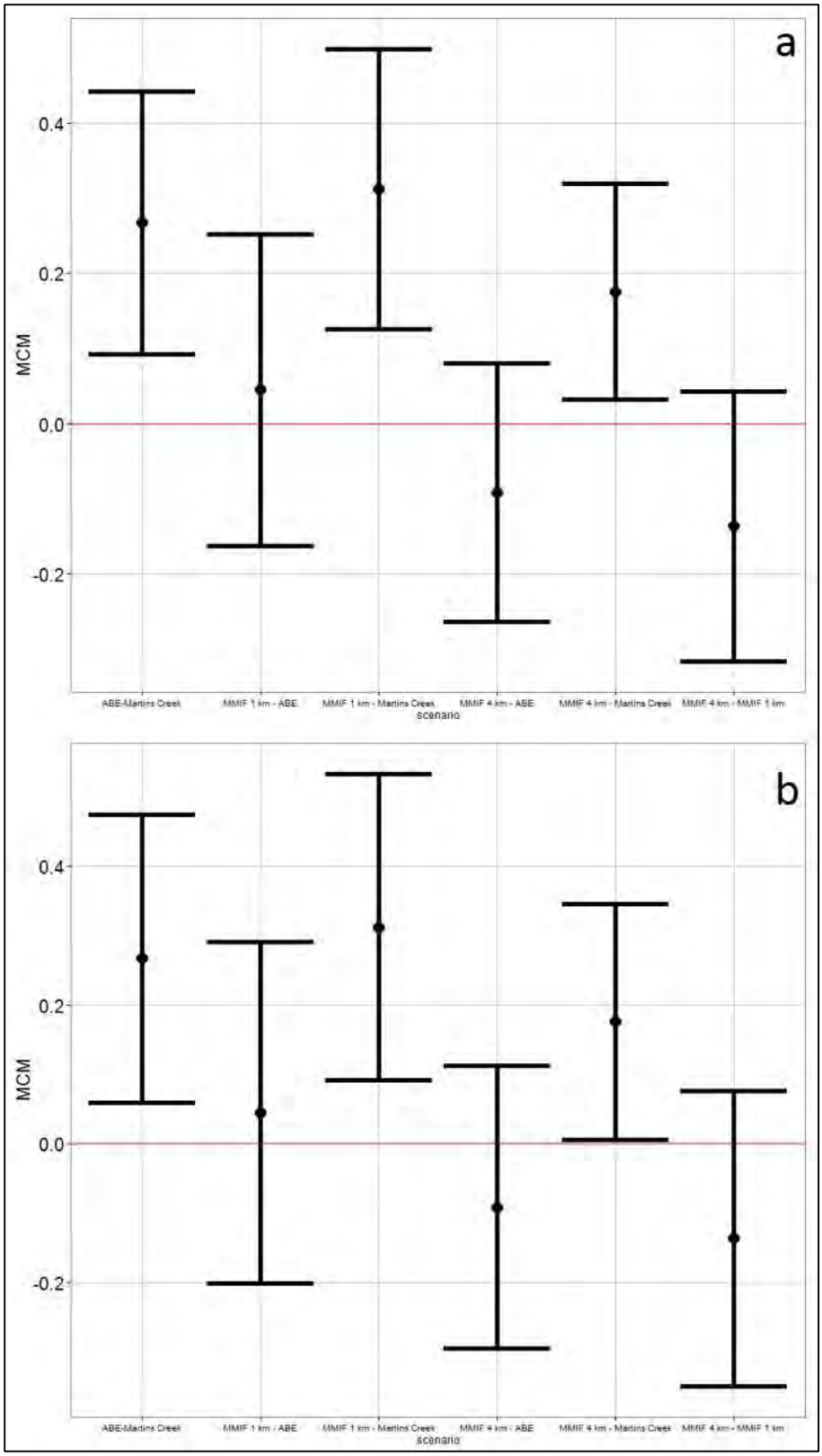

Figure 30. Martins Creek MCM differences with a) 90th percentile and b) 95th confidence intervals.

3.3 Herculaneum

3.3.1 Meteorological data comparisons

Figure 31 and Figure 32 show the wind roses for the Herculaneum site-specific tower, CPS, and three MMIF grid cells. Herculaneum and the three MMIF roses appear to show dominant directions from the northeast, southwest, west, and northwest. CPS is dominated by a southeasterly flow.

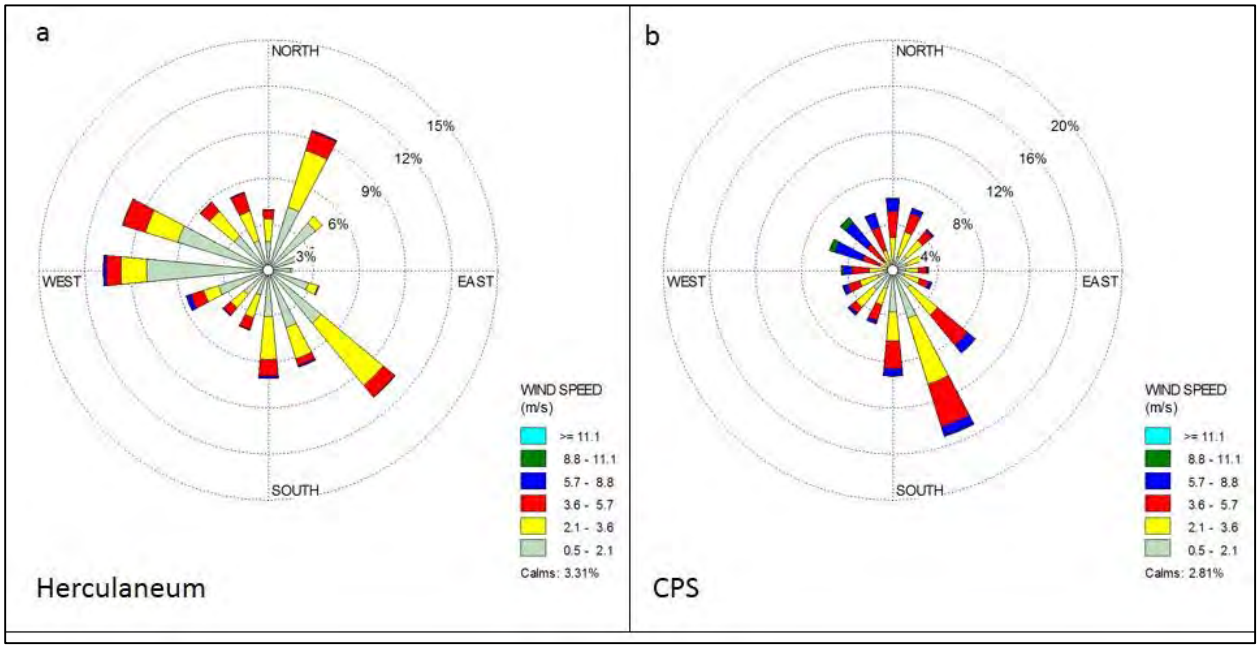

Figure 31. 2009 wind roses for a) Herculaneum and b) CPS.

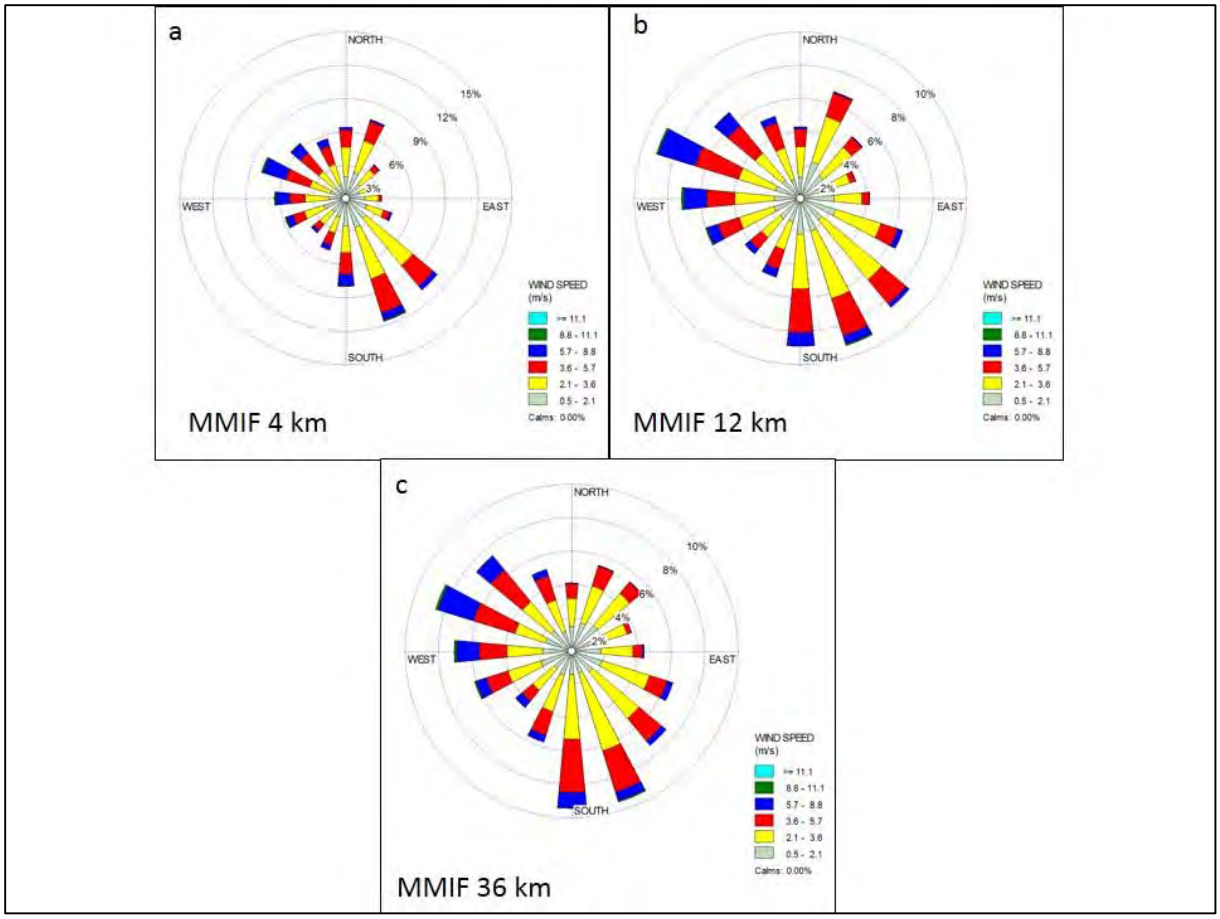

Figure 32. 2009 wind roses for a) MMIF 4 km, b) MMIF 12 km, and c) MMIF 36 km.

Figure 33 compares the wind displacement distributions among several difference among the Herculaneum scenarios. For the most part, displacements are less than 20 km. Figure 34 through Figure 36 compare the surface roughness values by month and wind direction. Surface roughness values for Herculaneum and CPS can vary dramatically by direction and month.

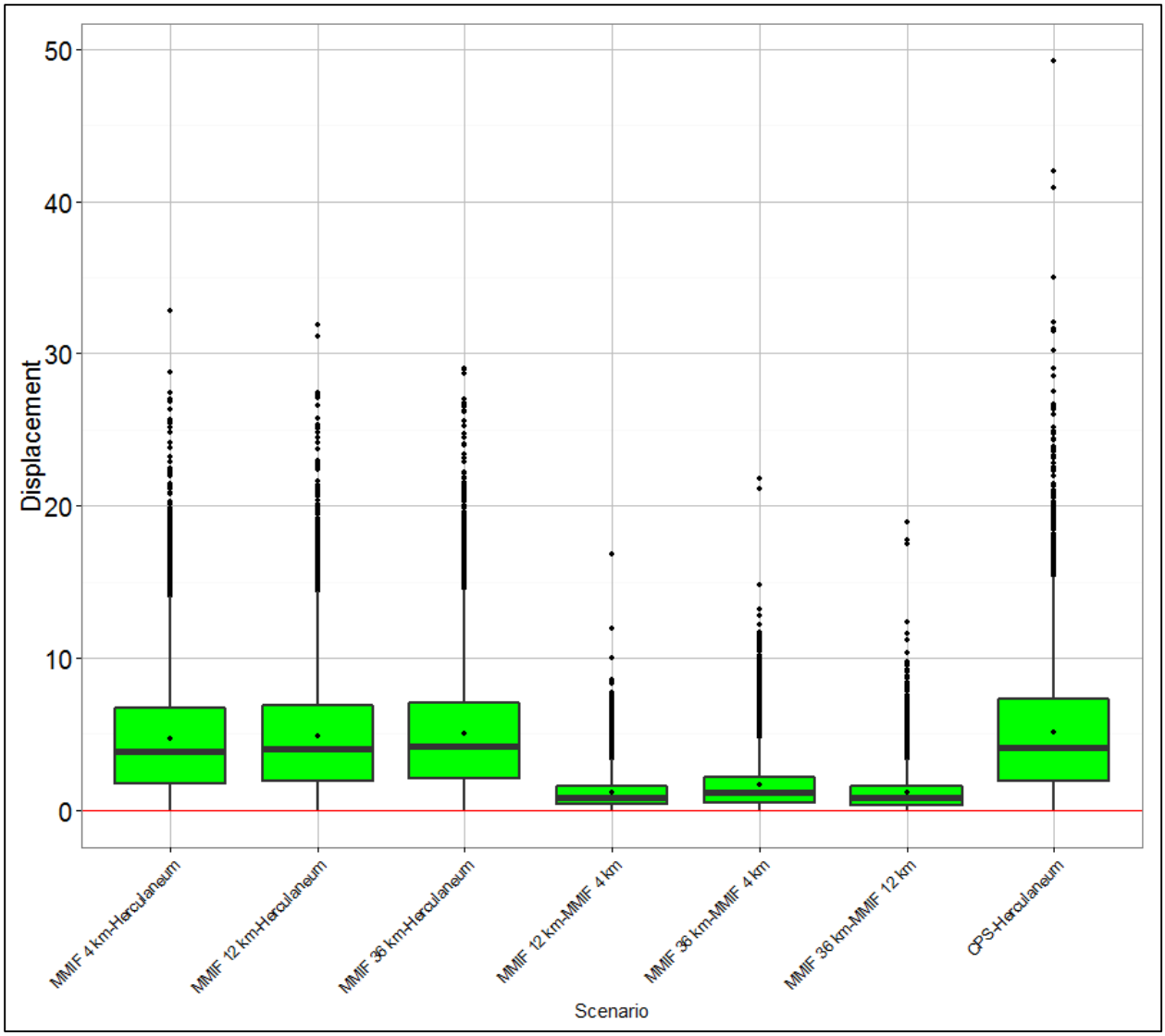

Figure 33. Wind displacement (km) among the Herculaneum meteorological scenarios.

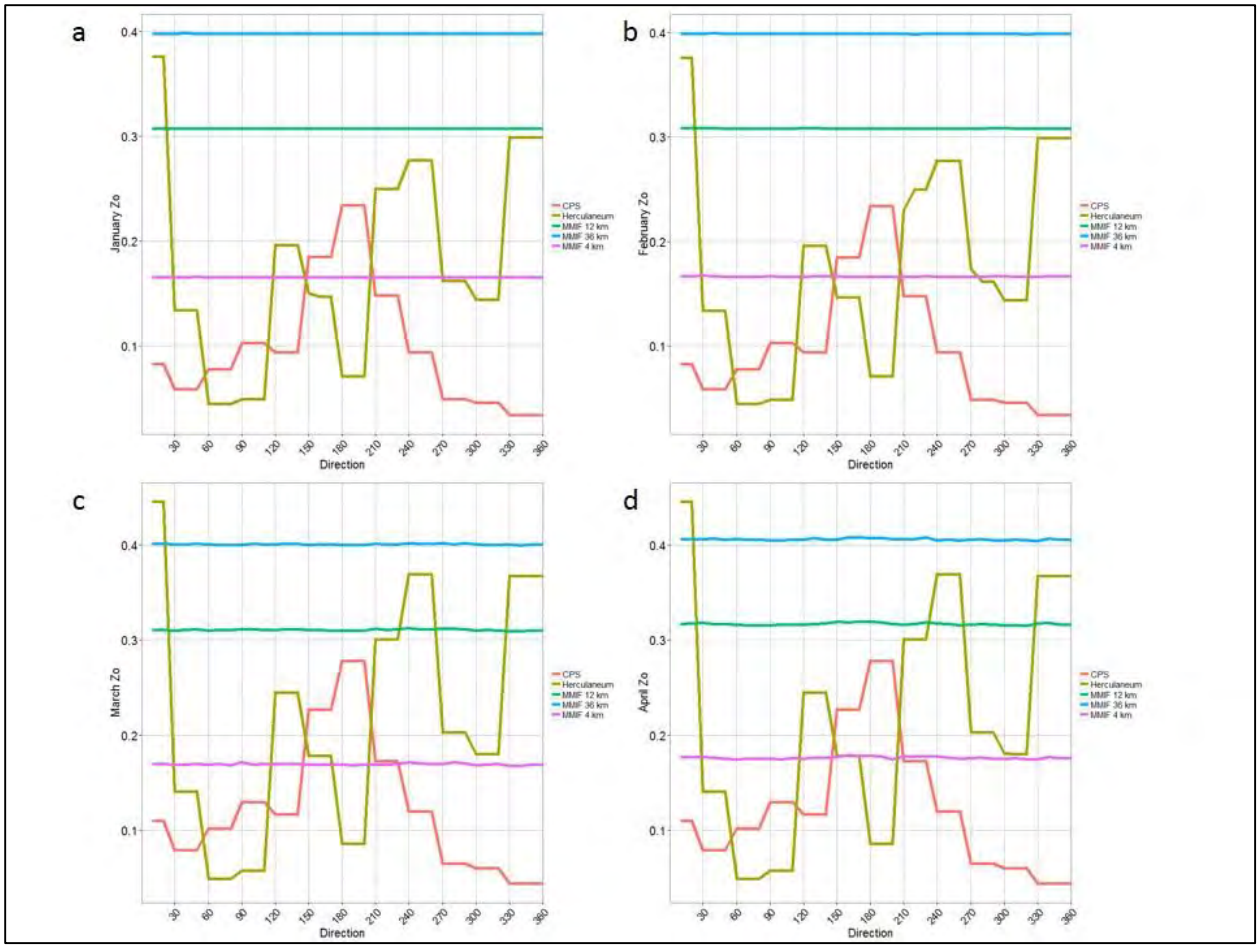

Figure 34. Herculaneum study monthly surface roughness lengths (m) by 10 degree sectors for a) January, b) February, c) March, and d) April.

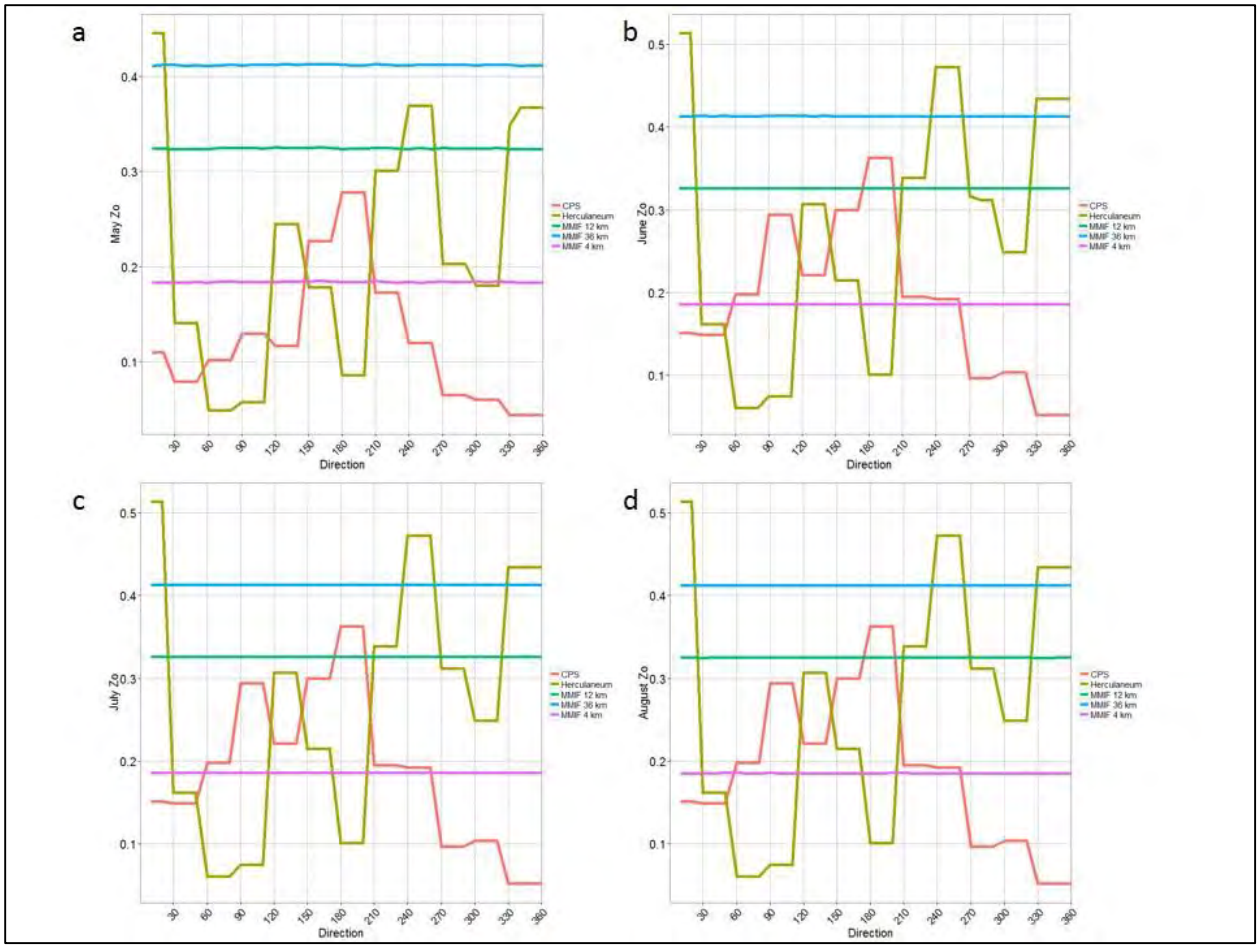

Figure 35. Herculaneum study monthly surface roughness lengths (m) by 10 degree sectors for a) May, b) June, c) July, and d) August.

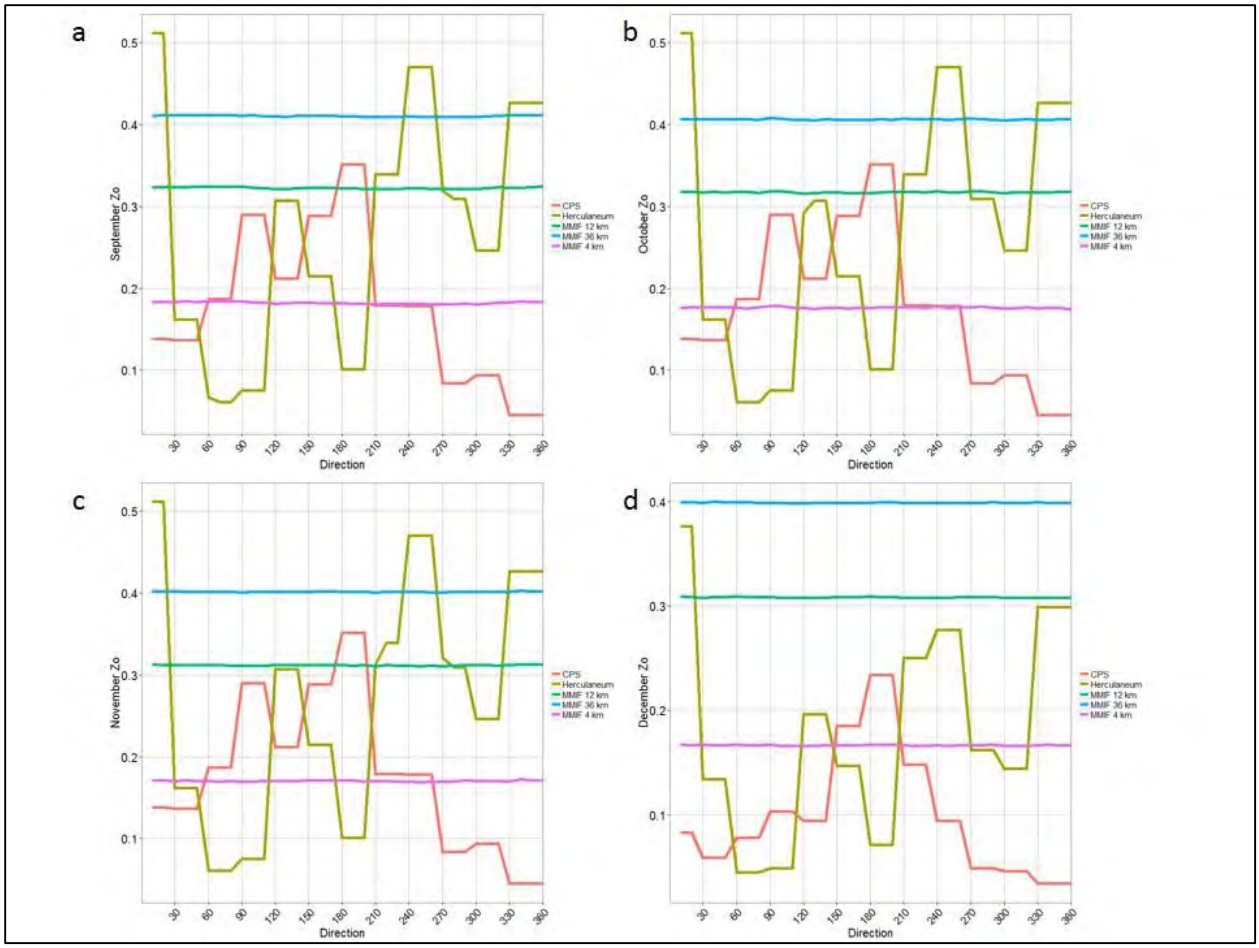

Figure 36. Herculaneum study monthly surface roughness lengths (m) by 10 degree sectors for a) September, b) October, c) November, and d) December.

Table 14 through Table 16 show statistics for the meteorological variables for several scenario differences. Box and whisker plots of the variable distributions as well as bias distributions can be found in Appendix A.

- Differences between MMIF wind speeds and observed wind speeds are lower than the differences between the two observed datasets but are comparable.
- The MMIF 4 and 12 km scenarios over-predict temperatures while the MMIF 36 km and CPS over-predict temperatures with the MMIF scenarios having less over-prediction.
- For pressures, the MMIF scenario under-predict. CPS and Herculaneum do not differ, most likely due to CPS values being used at Herculaneum in AERMET processing.
- MMIF scenarios and CPS over-predict relative humidity and heat flux.
- For u*, both the MMIF 4 km under-predicts while the MMIF 12 and 36 km scenarios over-predict.
- MMIF scenarios and CPS over-predict w*.

- For the potential temperature gradient, the MMIF scenarios over-predict while the CPS over-prediction is essentially zero.
- For mixing heights (both convective and mechanical) and Monin-Obukhov length, the MMIF scenarios and CPS over-predict.
- Cloud cover differences also show relatively low agreement but this may be due to the calculation methodology in AERMET when cloud cover is missing. ABE and Martins Creek do not differ, most likely due to ABE cloud cover being used for Martins Creek.

Overall, while there are differences, the MMIF scenarios appear to show relatively good agreement with the observed data. Differences between the MMIF scenarios and Herculaneum are usually in line with the CPS – Herculaneum biases.

Table 14. Mean bias, fractional bias, root mean square error, and R^2 for primary meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Wind speed	MMIF 4 km-Herculaneum	1.11391	0.10602	1.60036	0.58907
	MMIF 12 km-Herculaneum	1.10712	0.10918	1.56114	0.57557
	MMIF 36 km-Herculaneum	1.15359	0.11262	1.60752	0.56262
	MMIF 12 km-MMIF 4 km	-0.00679	0.00354	0.30025	0.97318
	MMIF 36 km-MMIF 4 km	0.03969	0.00759	0.46448	0.93287
	MMIF 36 km-MMIF 12 km	0.04647	0.00411	0.30716	0.96801
	CPS-Herculaneum	1.37366	0.10400	1.93392	0.59851
Ambient temperature	MMIF 4 km-Herculaneum	-0.03556	-0.00002	2.65892	0.93389
	MMIF 12 km-Herculaneum	0.00787	0.00002	2.66439	0.93362
	MMIF 36 km-Herculaneum	-0.05384	-0.00004	2.55690	0.93889
	MMIF 12 km-MMIF 4 km	0.04342	0.00004	0.34129	0.99886
	MMIF 36 km-MMIF 4 km	-0.01828	-0.00002	0.63005	0.99602
	MMIF 36 km-MMIF 12 km	-0.06170	-0.00006	0.47229	0.99780
	CPS-Herculaneum	-0.38928	-0.00035	2.09388	0.96207
Pressure	MMIF 4 km-Herculaneum	-0.72089	-0.00018	1.42029	0.97096
	MMIF 12 km-Herculaneum	-2.80479	-0.00070	3.05647	0.97200
	MMIF 36 km-Herculaneum	-6.45605	-0.00162	6.56495	0.97386
	MMIF 12 km-MMIF 4 km	-2.08390	-0.00052	2.10606	0.99801
	MMIF 36 km-MMIF 4 km	-5.73516	-0.00144	5.75746	0.99454
	MMIF 36 km-MMIF 12 km	-3.65126	-0.00092	3.68374	0.99470
	CPS-Herculaneum	0	0	0	1.0
Relative humidity	MMIF 4 km-Herculaneum	2.81507	0.00556	14.68699	0.58847
	MMIF 12 km-Herculaneum	1.88505	0.00103	14.86928	0.59163
	MMIF 36 km-Herculaneum	1.06450	-0.00176	14.50951	0.59984
	MMIF 12 km-MMIF 4 km	-0.93002	-0.00463	2.81453	0.98725
	MMIF 36 km-MMIF 4 km	-1.75057	-0.00746	4.48268	0.96733
	MMIF 36 km-MMIF 12 km	-0.82055	-0.00284	3.12158	0.98291
	CPS-Herculaneum	0.81473	0.00436	9.40803	0.75540

Table 15. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Heat flux	MMIF 4 km-Herculaneum	18.72392	0.03237	56.91295	0.58847
	MMIF 12 km-Herculaneum	30.29101	0.00573	79.39036	0.59163
	MMIF 36 km-Herculaneum	31.98523	0.01935	82.15212	0.59984
	MMIF 12 km-MMIF 4 km	11.55608	0.08710	35.43597	0.98725
	MMIF 36 km-MMIF 4 km	13.21507	0.06863	51.23716	0.96733
	MMIF 36 km-MMIF 12 km	1.65898	0.01543	38.98357	0.98291
	CPS-Herculaneum	6.65078	0.04541	23.37517	0.75540
u*	MMIF 4 km-Herculaneum	-0.01028	-0.02039	0.14290	0.58847
	MMIF 12 km-Herculaneum	0.05683	0.03896	0.16563	0.59163
	MMIF 36 km-Herculaneum	0.10084	0.07518	0.18834	0.59984
	MMIF 12 km-MMIF 4 km	0.06712	0.06483	0.08534	0.98725
	MMIF 36 km-MMIF 4 km	0.11108	0.10116	0.13904	0.96733
	MMIF 36 km-MMIF 12 km	0.04397	0.03955	0.07255	0.98291
	CPS-Herculaneum	0.01861	0.02198	0.12423	0.75540
w*	MMIF 4 km-Herculaneum	0.44585	0.11391	0.65988	0.58847
	MMIF 12 km-Herculaneum	0.70574	0.15819	0.88205	0.59163
	MMIF 36 km-Herculaneum	0.72036	0.16160	0.88035	0.59984
	MMIF 12 km-MMIF 4 km	0.34209	0.06955	0.42287	0.98725
	MMIF 36 km-MMIF 4 km	0.33219	0.06015	0.50998	0.96733
	MMIF 36 km-MMIF 12 km	-0.00180	-0.00490	0.30334	0.98291
	CPS-Herculaneum	0.17976	0.08154	0.33074	0.75540
dΘ/dz	MMIF 4 km-Herculaneum	0.00314	0.07345	0.00647	0.58847
	MMIF 12 km-Herculaneum	0.00298	0.06886	0.00649	0.59163
	MMIF 36 km-Herculaneum	0.00280	0.06495	0.00620	0.59984
	MMIF 12 km-MMIF 4 km	0.00004	0.00012	0.00151	0.98725
	MMIF 36 km-MMIF 4 km	-0.00028	-0.00648	0.00214	0.96733
	MMIF 36 km-MMIF 12 km	-0.00035	-0.00710	0.00201	0.98291
	CPS-Herculaneum	5.6×10^{-7}	0.00078	0.00304	0.75540

Table 16. Mean bias, fractional bias, root mean square error, and R^2 for calculated meteorological variables.

Variable	Scenario	Mean bias	Fractional bias	RMSE	R^2
Zi_c	MMIF 4 km-Herculaneum	566.12886	0.21689	689.85788	0.55479
	MMIF 12 km-Herculaneum	570.47601	0.21825	699.26281	0.55115
	MMIF 36 km-Herculaneum	555.89452	0.21464	682.41628	0.57861
	MMIF 12 km-MMIF 4 km	25.53641	0.00685	82.12590	0.98265
	MMIF 36 km-MMIF 4 km	10.05379	0.00148	134.17900	0.94961
	MMIF 36 km-MMIF 12 km	-17.08090	-0.00619	113.15424	0.96540
	CPS-Herculaneum	178.88253	0.11968	288.74963	0.79372
Zi_m	MMIF 4 km-Herculaneum	89.93761	0.03008	523.68956	0.20612
	MMIF 12 km-Herculaneum	111.37459	0.04574	534.07238	0.20443
	MMIF 36 km-Herculaneum	107.75352	0.04503	544.49645	0.18224
	MMIF 12 km-MMIF 4 km	21.22066	0.01780	105.80767	0.96336
	MMIF 36 km-MMIF 4 km	17.59007	0.01751	151.40997	0.92336
	MMIF 36 km-MMIF 12 km	-3.63059	-0.00001	122.67669	0.94935
	CPS-Herculaneum	28.58368	0.02725	276.80805	0.62259
L	MMIF 4 km-Herculaneum	249.61225	-0.10265	1681.53419	0.00120
	MMIF 12 km-Herculaneum	231.00914	-0.06609	1699.98627	0.00065
	MMIF 36 km-Herculaneum	288.82372	-0.45512	1823.38370	0.00419
	MMIF 12 km-MMIF 4 km	-16.39179	0.06853	1356.42715	0.00001
	MMIF 36 km-MMIF 4 km	39.35898	0.09618	1458.10173	0.01241
	MMIF 36 km-MMIF 12 km	55.75078	0.07233	1570.10520	0.00005
	CPS-Herculaneum	141.90371	0.01109	1276.36930	0.20768
Cloud cover	MMIF 4 km-Herculaneum	3.71020	0.23923	5.81028	0.13757
	MMIF 12 km-Herculaneum	3.67508	0.23811	5.78924	0.13824
	MMIF 36 km-Herculaneum	3.55515	0.23551	5.72211	0.14100
	MMIF 12 km-MMIF 4 km	-0.03505	-0.00294	0.45694	0.98193
	MMIF 36 km-MMIF 4 km	-0.15502	-0.00932	0.94022	0.92781
	MMIF 36 km-MMIF 12 km	-0.11998	-0.00655	0.88577	0.93537
	CPS-Herculaneum	0	0	0	1.0

3.3.2 AERMOD results

Figure 37 shows the 24-hour QQ plots for Herculaneum. Since the pollutant is lead, daily averages are the only averaging period available at the monitors. Figure 37 shows that all five scenarios are under-predicting concentrations. Figure 38 shows the 24-hour screening results and the results concur with the QQ-plots. With the exception of CPS, the scenarios are outside the factor of two limit but the MMIF scenarios are comparable to the Herculaneum site specific scenario.

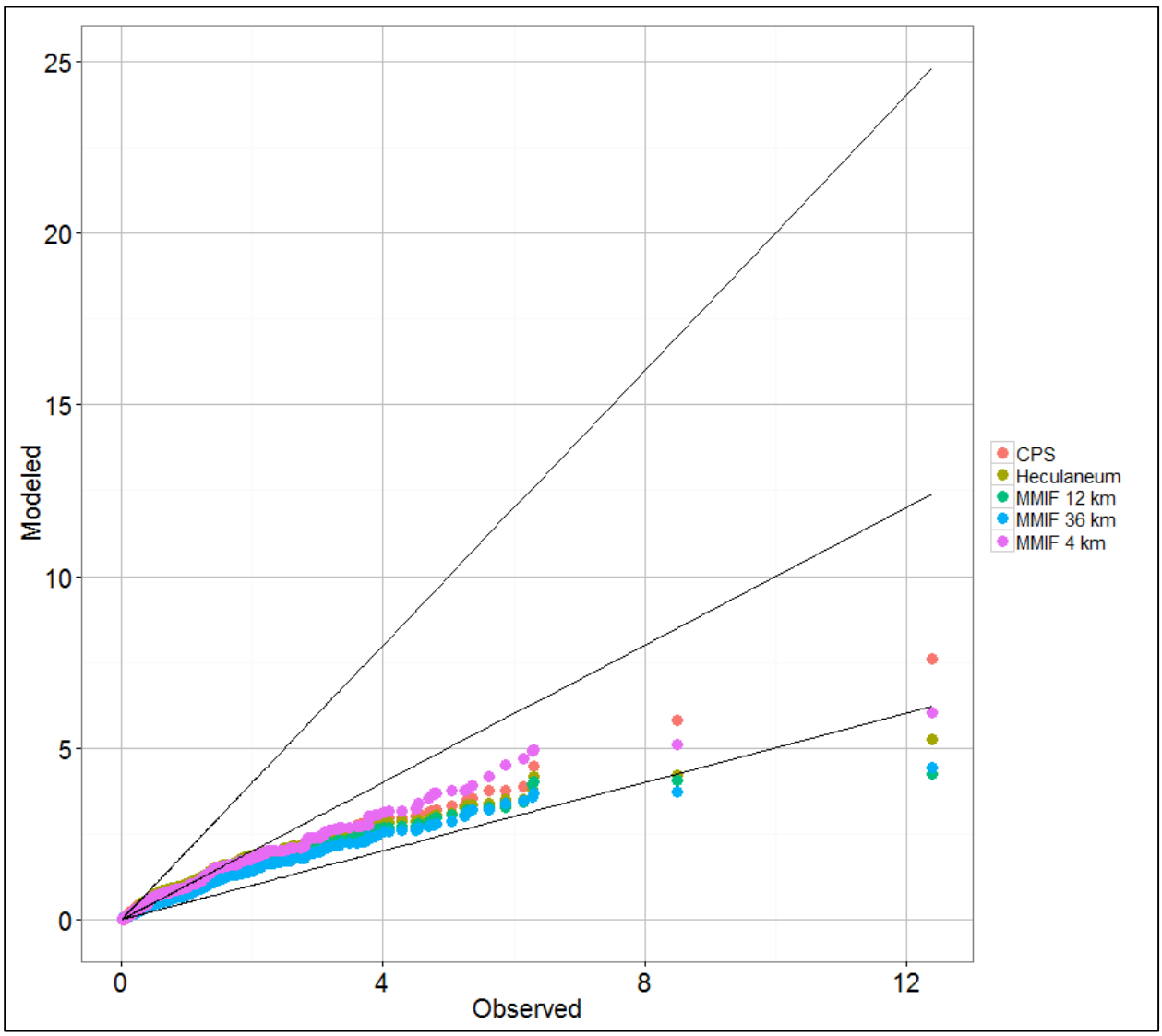

Figure 37. Herculaneum 24-hour QQ plots. Concentrations are in µg/m^3.

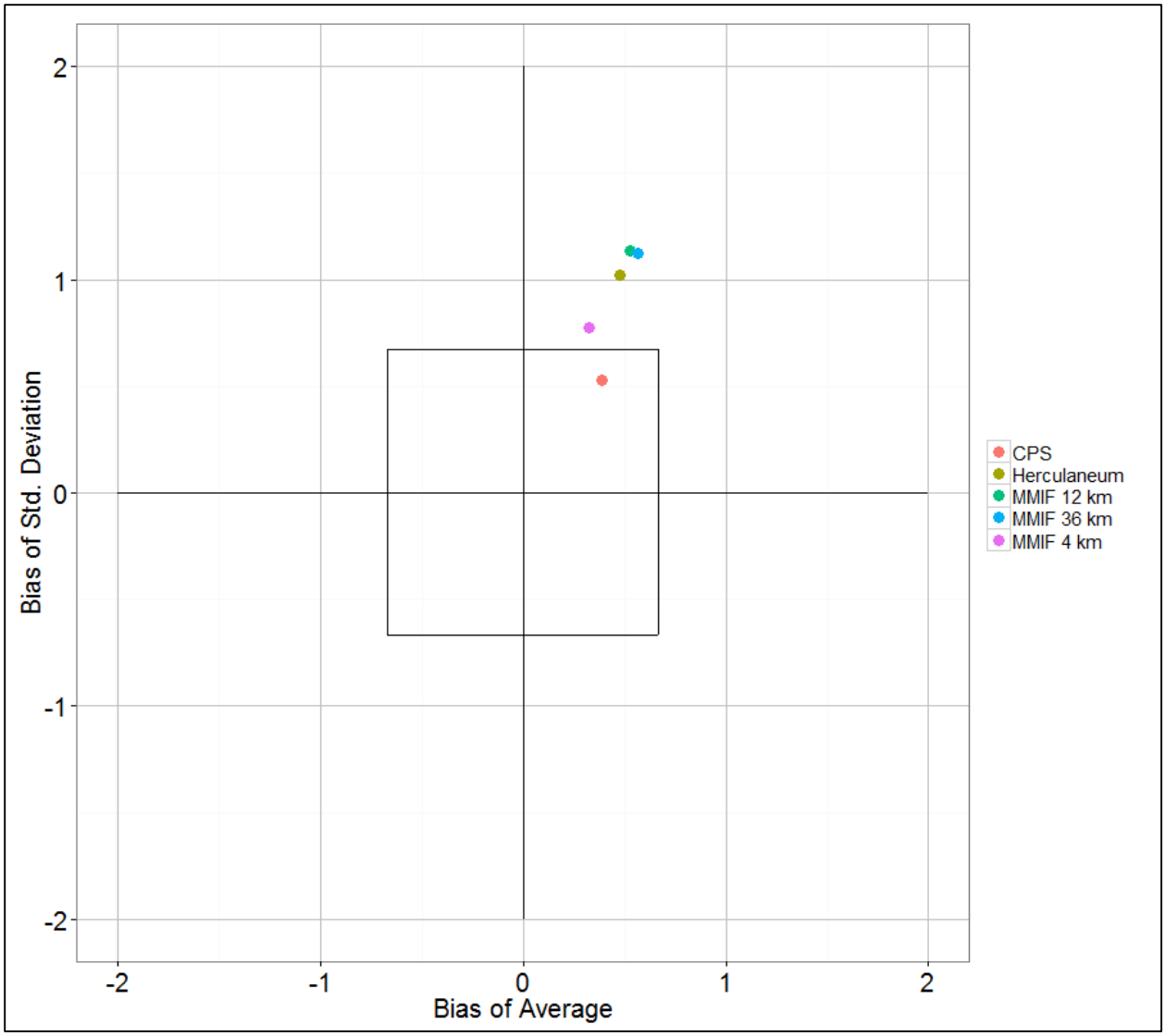

Figure 38. Herculaneum 24-hour screening results.

4. Summary and Conclusions

MMIF output was evaluated for three areas involving differing complexities of terrain from relatively flat to complex, differing resolutions of prognostic model output, and two different pollutants, SO_2 and lead. Evaluation of the meteorological data for several variables indicated that while there are differences between the prognostic model output and observations, the prognostic model output was not unreasonable. However, it should be noted that small differences in meteorological variables, such as temperature, wind speed or direction, can lead to vastly different air quality results.

Statistical evaluations of the resulting AERMOD results from observed meteorological data and prognostic data exhibited differing behaviors.

For Gibson, all meteorological scenarios, observed or prognostic, tended toward under-prediction of SO_2 concentrations at higher concentrations but the screening analyses indicated biases close to zero. At the 3-hour averaging time, all scenarios exceeded the factor-of-two limit in the screening analysis but were within a factor-of-two for the 24-hour averages. Evaluations of the MCM values indicated the best performing scenarios tended to be the observed meteorological data but that the prognostic outputs were not statistically different. Of the two prognostic outputs, the MMIF output of the grid cell of Gibson performed better than the grid cell of the Evansville NWS station.

For the Martins Creek case study, the meteorological scenarios tended to over-predict and the CPM and MCM analyses revealed that the Martins Creek site-specific observed data performed better. From the MCM analysis, the prognostic meteorological scenarios were statistically different from the Martins Creek site-specific scenario. Compared to the ABE NWS data, both prognostic scenarios were not statistically different, which was a goal of this evaluation; prognostic meteorological data performs no worse than a representative NWS station.

Finally, for Herculaneum, detailed statistical analyses could not be performed due to the lack of hourly lead observations, but a screening analysis of the outputs indicated that while the prognostic scenarios were not within the factor-of-two agreement for 24-hour averages, their performance was actually comparable to the Herculaneum site-specific tower.

While there is a need for more evaluations in more challenging environments, i.e. complex terrain or meteorological conditions, these results indicate promise for the use of prognostic meteorological data in AERMOD applications.

5. References

Cox, W.M. and J.A. Tikvart, 1990: A statistical Procedure for Determining the Best Performing Air Quality Simulation Model. *Atmos. Environ.,* **24A (9)**: 2387-2395.

Environ, 2014: The Mesoscale Model Interface Program (MMIF) Version 3.1 User's Manual.

Environ, 2016: The Mesoscale Model Interface Program (MMIF) Version 3.1 User's Manual.

Frost, Kali, 2014: AERMOD Performance Evaluation for Three Coal-fired Electrical Generating Units in Southwest Indiana. *Journal of the Air & Waste Management Association*, **64:3** 280-290.

Perry Steven G., Alan J. Cimorelli, Robert J. Paine, Roger W. Brode, Jeffrey C. Weil, Akula Venkatram, Robert B. Wilson, , Russell F. Lee., and Warren D. Peters, 2005: AERMOD: A Dispersion Model for Industrial Source Applications. Part II: Model Performance Against 17 Field Study Databases. Journal of Applied Meteorology, 44, 694-708.

U.S. EPA, 1992: Protocol for Determining the Best Performing Model, EPA-454/R-92-025. U.S. Environmental Protection Agency, Research Triangle Park, NC.

U.S. EPA, 2003: AERMOD: Latest Features and Evaluation Results, EPA-454/R-03-003. U.S. Environmental Protection Agency, Research Triangle Park, NC.

U.S. EPA, 2013: AERSURFACE User's Guide. EPA-454/B-08-001. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711.

U.S. EPA, 2014: Meteorological Model Performance for Annual 2011 WRF v3.4 Simulation. http://www.epa.gov/ttn/scram/reports/MET_TSD_2011_final_11-26-14.pdf

U.S. EPA, 2016a: *Guideline on Air Quality Models.* 40 CFR Part 51 Appendix W.

U.S. EPA, 2016b: Guidance on the Use of the Mesoscale Model Interface Program (MMIF) for AERMOD Applications. EPA-454/B-16-003. U.S. Environmental Protection Agency, Research Triangle Park, NC 27711.

Appendix A. Meteorological data comparisons

A.1 Gibson

Figures A-1 through A-6 compare the distributions of several meteorological variables among the four meteorological datasets (panel a of the figures). These variables are the input variables into AERMET: wind speed, temperature, pressure, relative humidity, albedo, and Bowen ratio. Comparisons of key meteorological variables calculated by AERMET including: heat flux, u*, w*, Monin-Obukhov length, and mixing heights can be found in Figures A-7 through A-14. In panel b of each figure are distributions of the biases of each variable among the different scenarios. The first bias distribution is between the two MMIF outputs, GIB MMIF and EVV MMIF. This bias shows how the meteorology can vary in a distance of a few grid cells. The next two bias distributions show the differences between MMIF and observations for the Gibson and Evansville grid cells respectively. The final distribution shows the bias between the two observed datasets and acts as a control since, in the absence of prognostic data, the two observed datasets would be the only ones available for consideration. If the differences between the MMIF outputs and the two observed datasets are similar, then from qualitative standpoint, the MMIF output is reasonable for use.

Based on the figures the following can be seen:

- For wind speed (Figure A-1), EVV OBS tends to show lower wind speeds (Figure A-1.a) among the four datasets. Each locations' MMIF – OBS bias tends to be more positive, correlating with the distributions in Figure A-1.a. The GIB MMIF – EVV MMIF bias trends positive as does the GIB OBS- EVV OBS, indicating that the MMIF output appears reasonable when compared to the observed datasets.
- Temperature (K) distributions (Figure A-2) indicate very similar patterns for all scenarios with the bias distributions tending to show fairly unbiased results among the scenario (median bias near zero degrees).
- The GIB OBS pressure distribution (Figure A-3) shows higher pressures than over the other scenarios which also bears out in the bias distributions. The two MMIF scenarios are similar. The differences between the two observed datasets indicate higher pressures at Gibson over Evansville.
- Relative humidity (Figure A-4) is identical for GIB OBS and EVV OBS. This is due to EVV being the source of RH for GIB in AERMET. The two MMIF scenarios are very similar. Their bias distributions when compared to their respective OBS dataset are very similar.
- Daytime albedo (Figure A-5) somewhat similar distributions among the scenarios with GIB MMIF – EVV MMIF, GIB MMIF – GIB OBS exhibiting positive bias and the other two being negative.

- Bowen ratios (Figure A-6) vary widely among the scenarios. This could be due to the subjectivity of the average, dry, and wet selections in AERSURFACE.
- Heat flux (Figure A-7) shows GIB OBS with lower values compared to the other scenarios. The differences between the two MMIF sites and the differences between the observed heat fluxes were small compared to the differences between the respective MMIF and observed datasets.
- Surface friction velocity, u* (Figure A-8) shows that the MMIF scenarios tended to have higher values than the observed scenarios. The GIB MMIF – EVV MMIF bias distribution is similar to the GIB OBS – EVV OBS bias distribution.
- Convective velocity scale, w* (Figure A-9) show a similar trend as u*, the MMIF scenarios exhibited higher values than the observed scenarios.
- Monin-Obukhov length, L (Figure A-10) distributions appear very similar across the scenarios with bias distributions being similar as well.
- Convective mixing height, Zic (Figure A-11) distributions and bias distributions indicate that the MMIF mixing heights were higher than their observed counterparts. Both the MMIF and observed bias plots indicate a tendency for Gibson mixing heights to be lower than the Evansville mixing heights.
- Mechanical mixing height, Zim (Figure A-12) distributions show similar behavior as the convective mixing heights.
- Potential temperature gradient ($d\Theta/dz$), (Figure A-13), distributions indicate smaller lapse rates for the MMIF data compared to the observed data.
- Cloud cover (Figure A-14) distributions show no differences between EVV and GIB for the observed datasets since EVV cloud cover is used at Gibson. Cloud cover estimates for MMIF are calculated in AERMET. Differences between the MMIF and OBS for Gibson and Evansville show a large swing in biases.

For the most part, the meteorological data processed through MMIF for Gibson and Evansville appear reasonable when compared to the Gibson and Evansville observations.

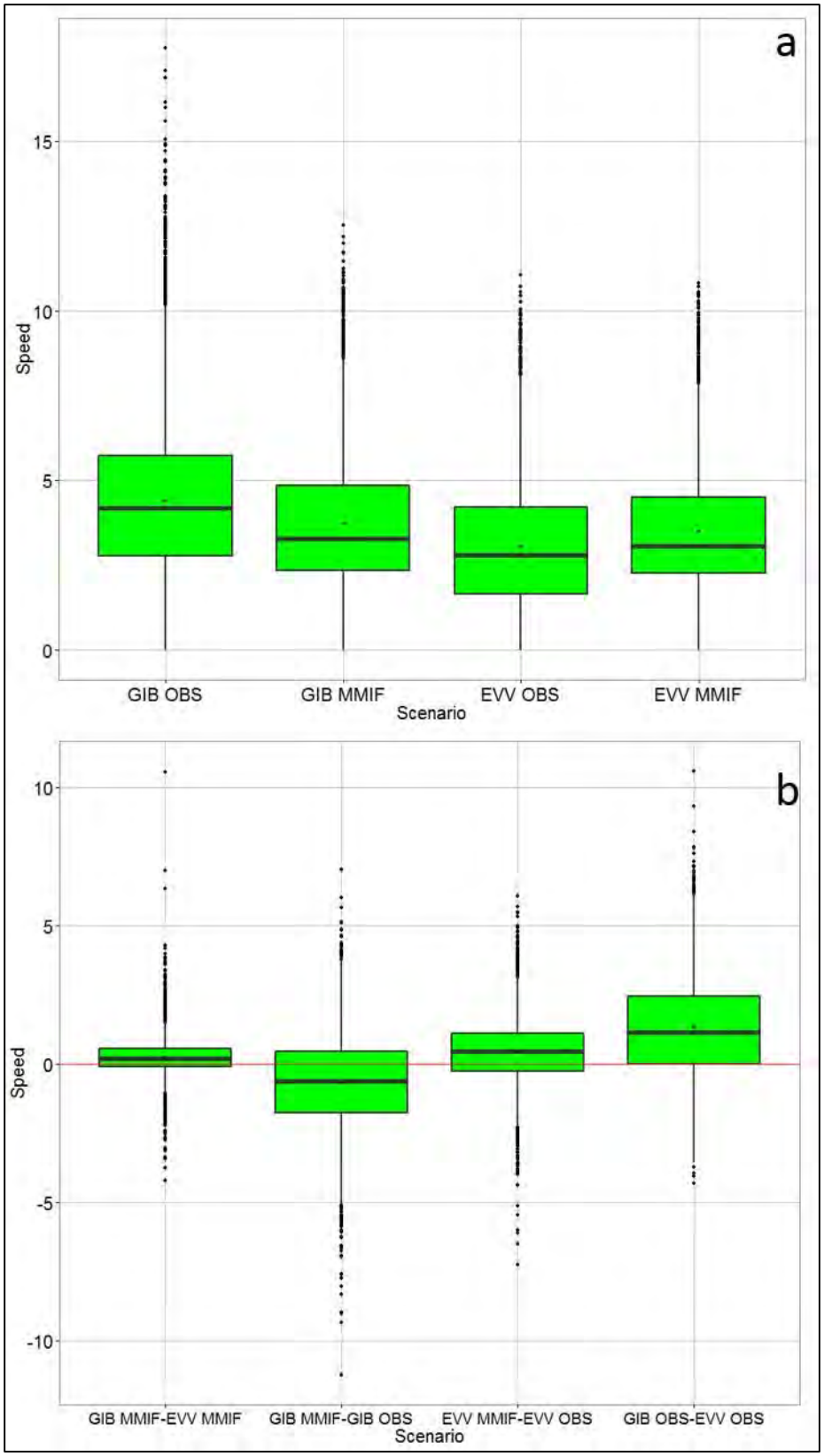

Figure A-1. Gibson wind speed (m/s): a) annual distributions and b) bias distributions.

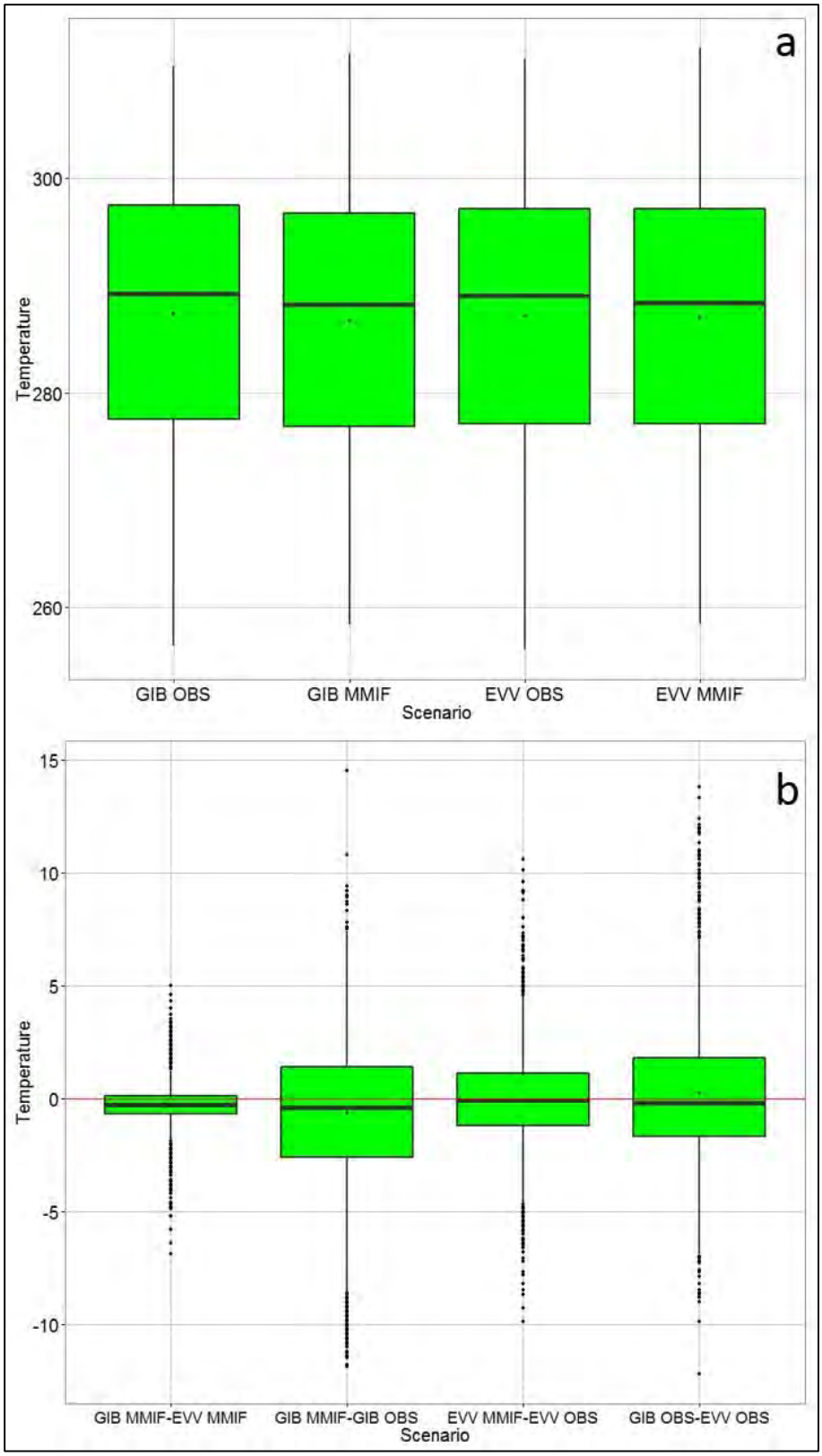

Figure A-2. Gibson ambient temperature (K): a) annual distributions and b) bias distributions.

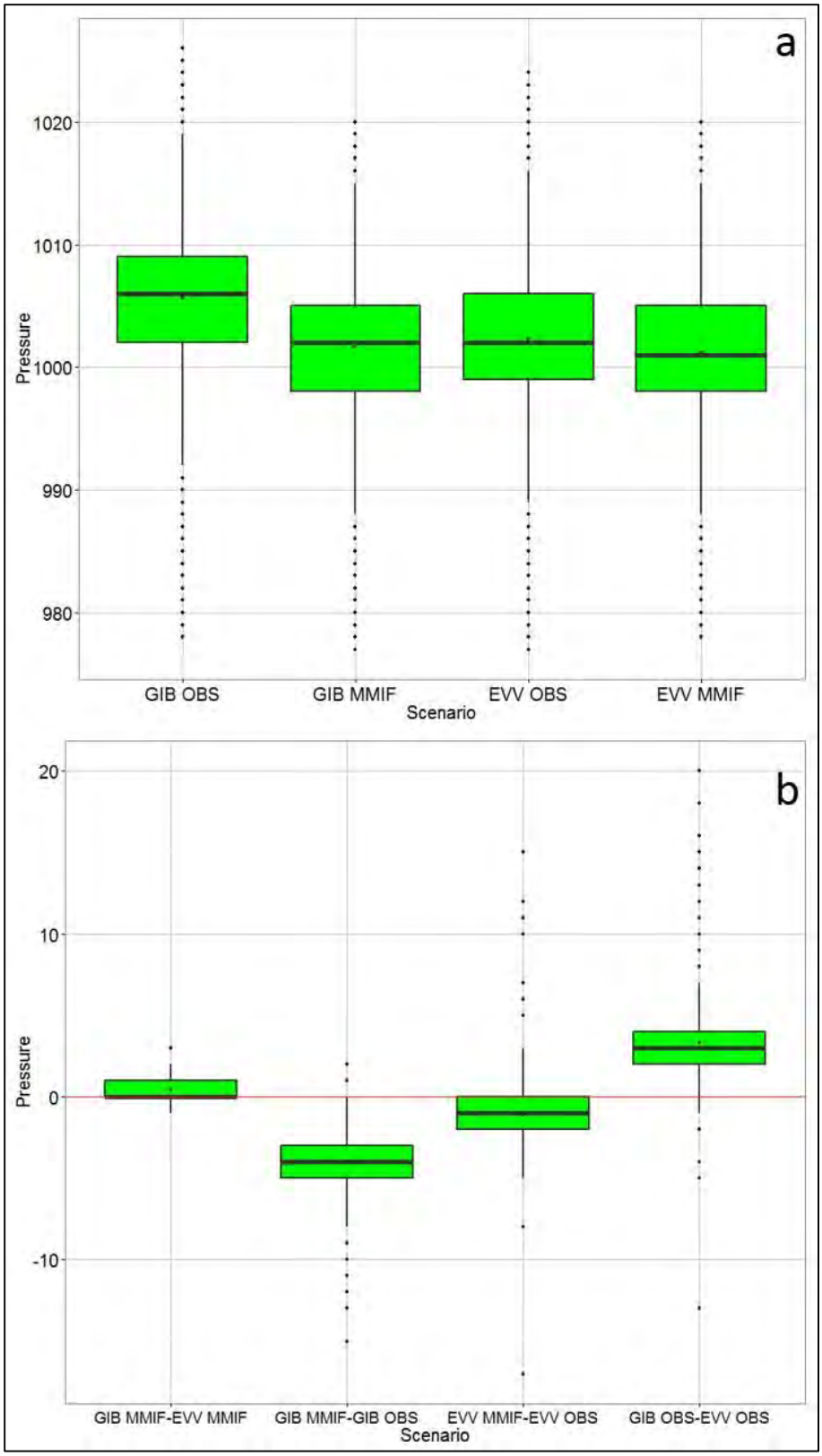

Figure A-3. Gibson station pressure (mb): a) annual distributions and b) bias distributions.

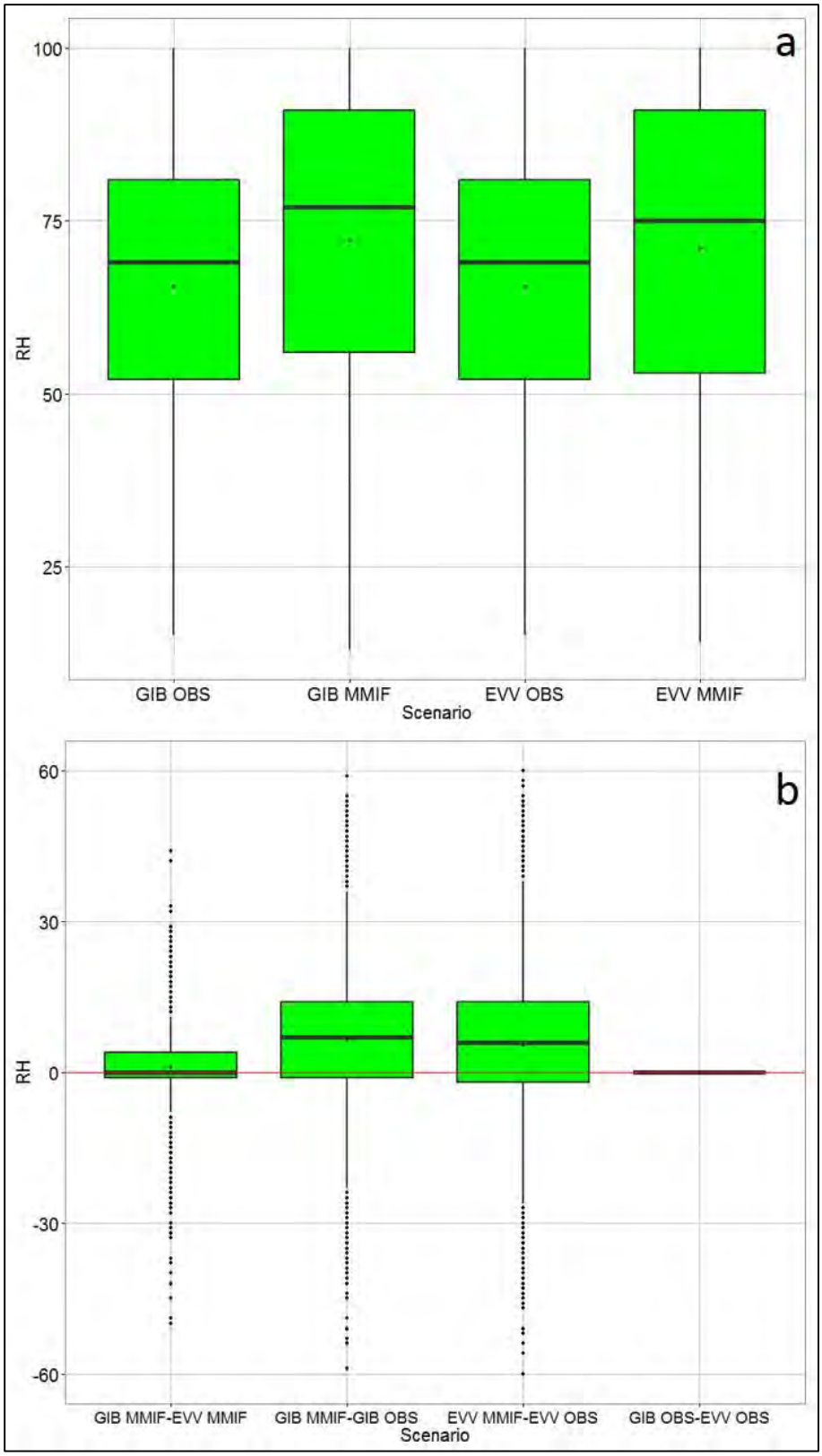

Figure A-4. Gibson relative humidity (percent): a) annual distributions and b) bias distributions.

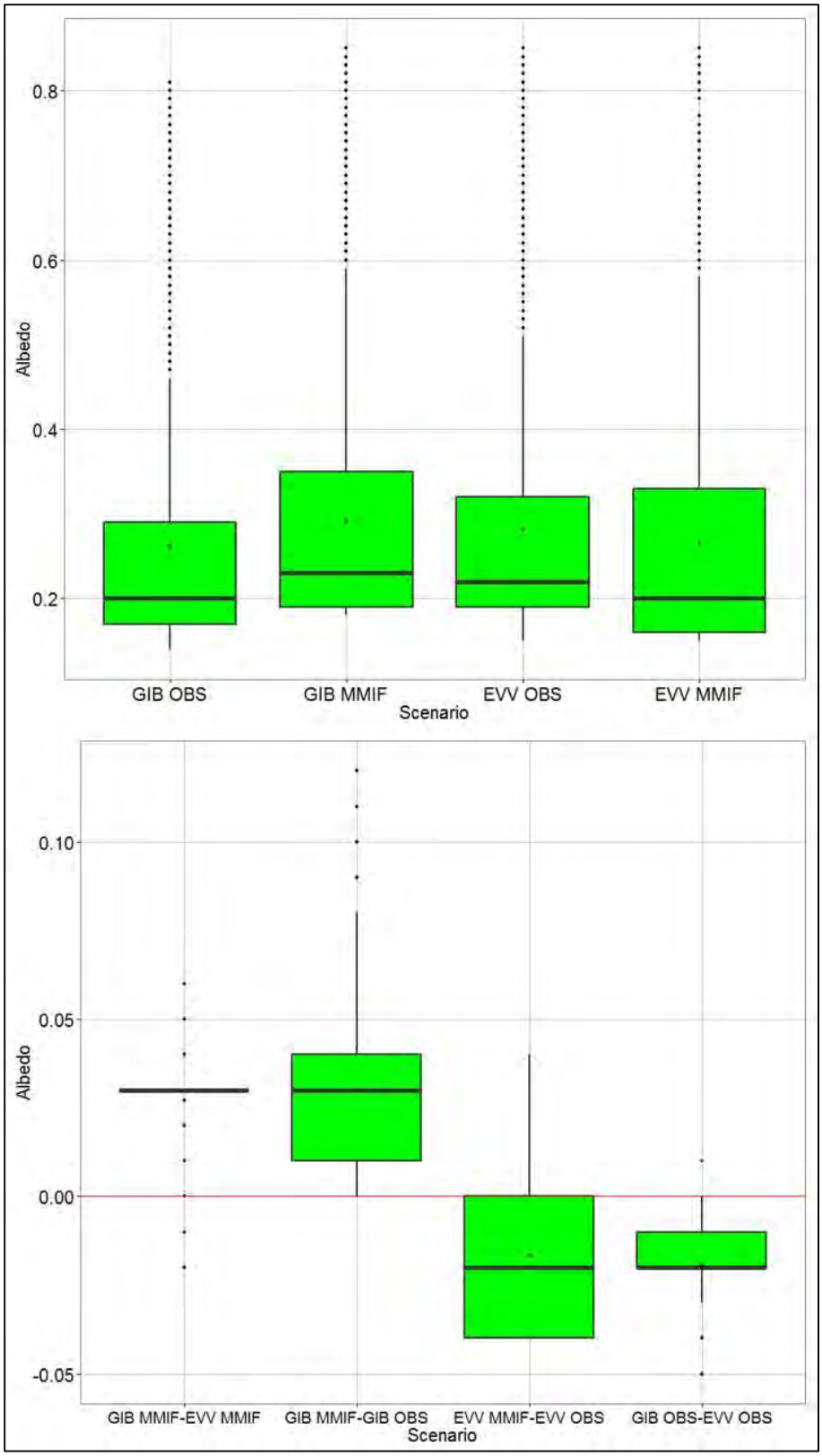

Figure A-5. Gibson daytime albedo (fraction): a) annual distributions and b) bias distributions.

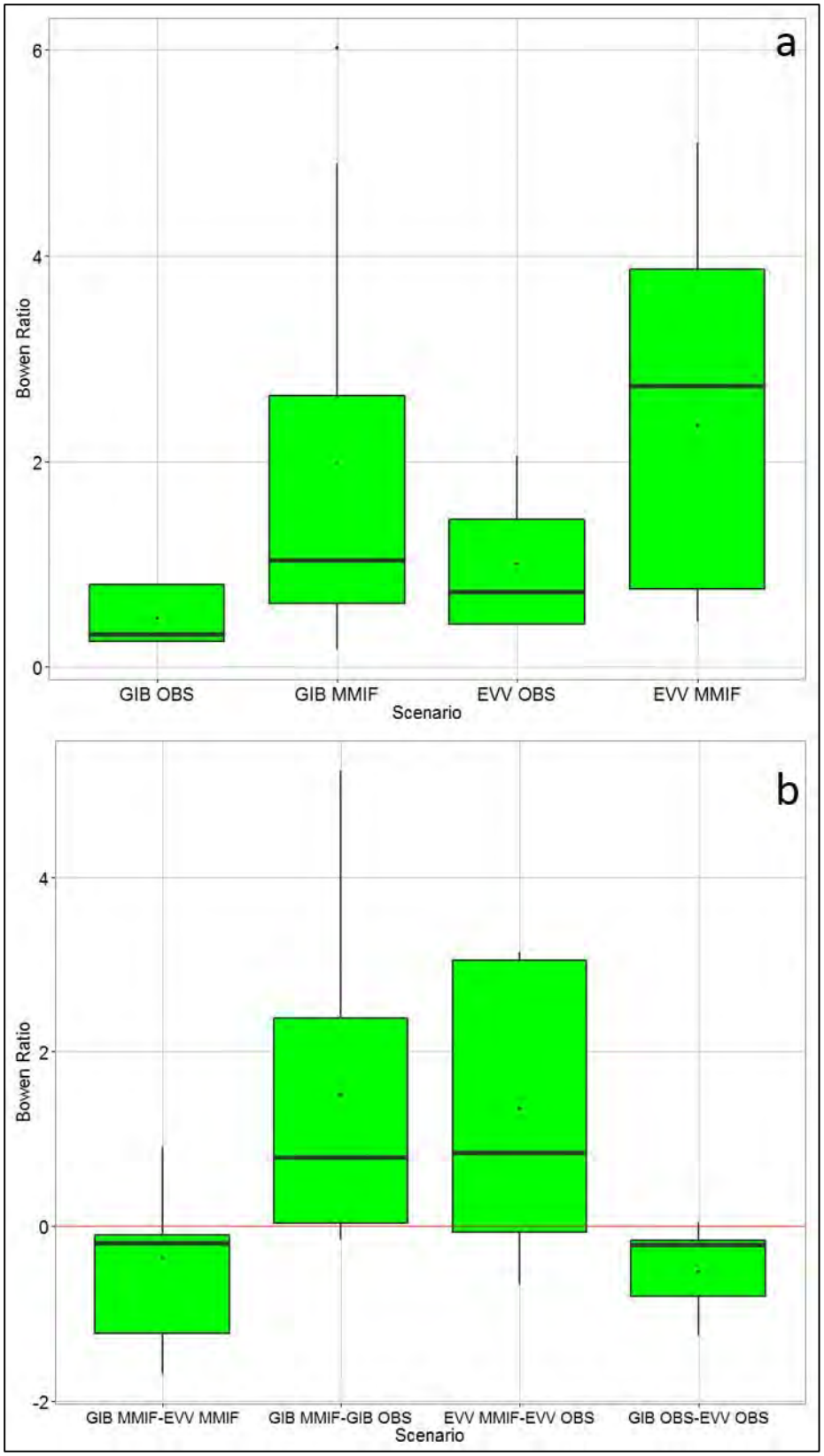

Figure A-6. Gibson Bowen ratio: a) annual distributions and b) bias distributions.

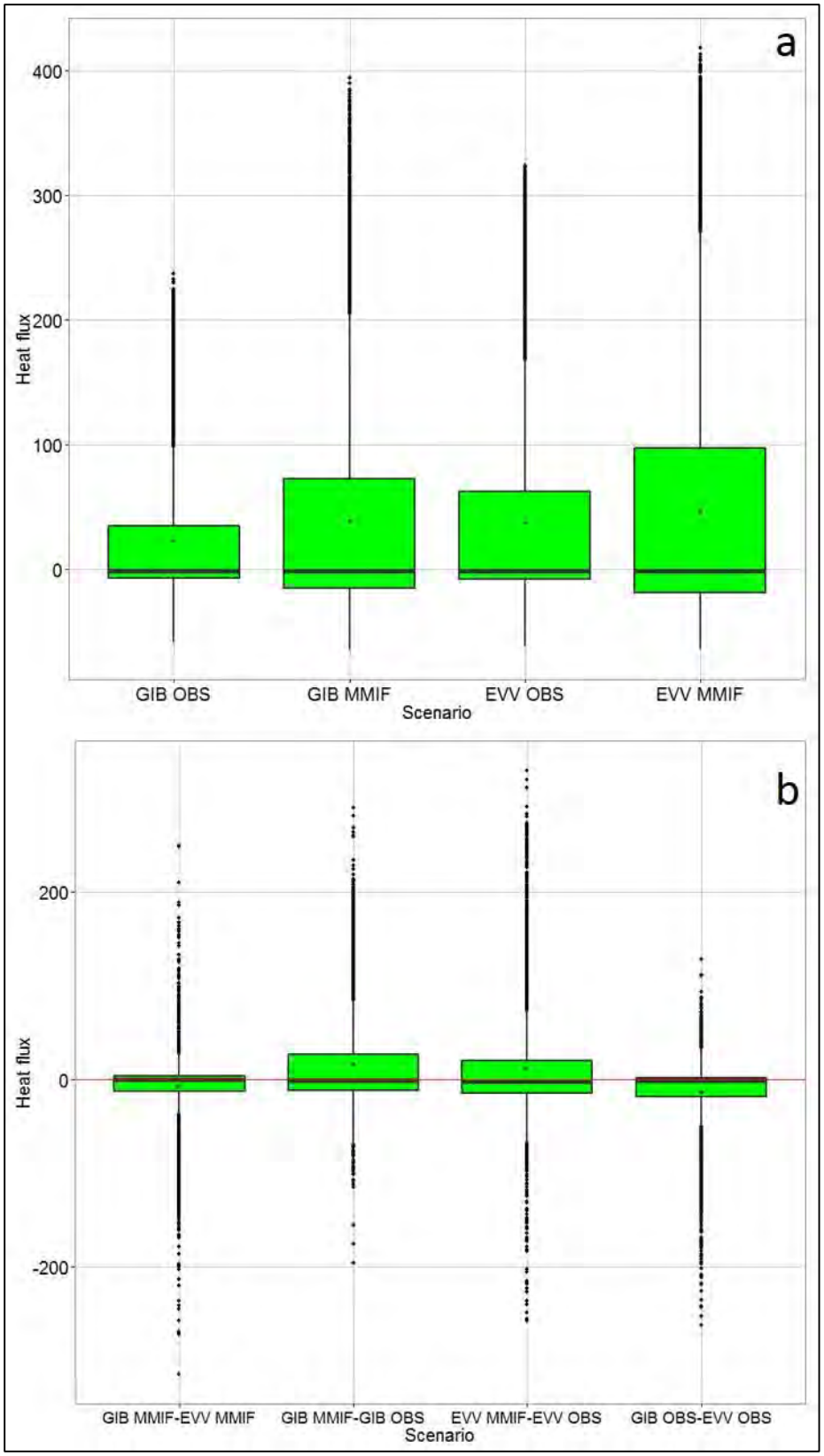

Figure A-7. Gibson heat flux (W/m^2): a) annual distributions and b) bias distributions.

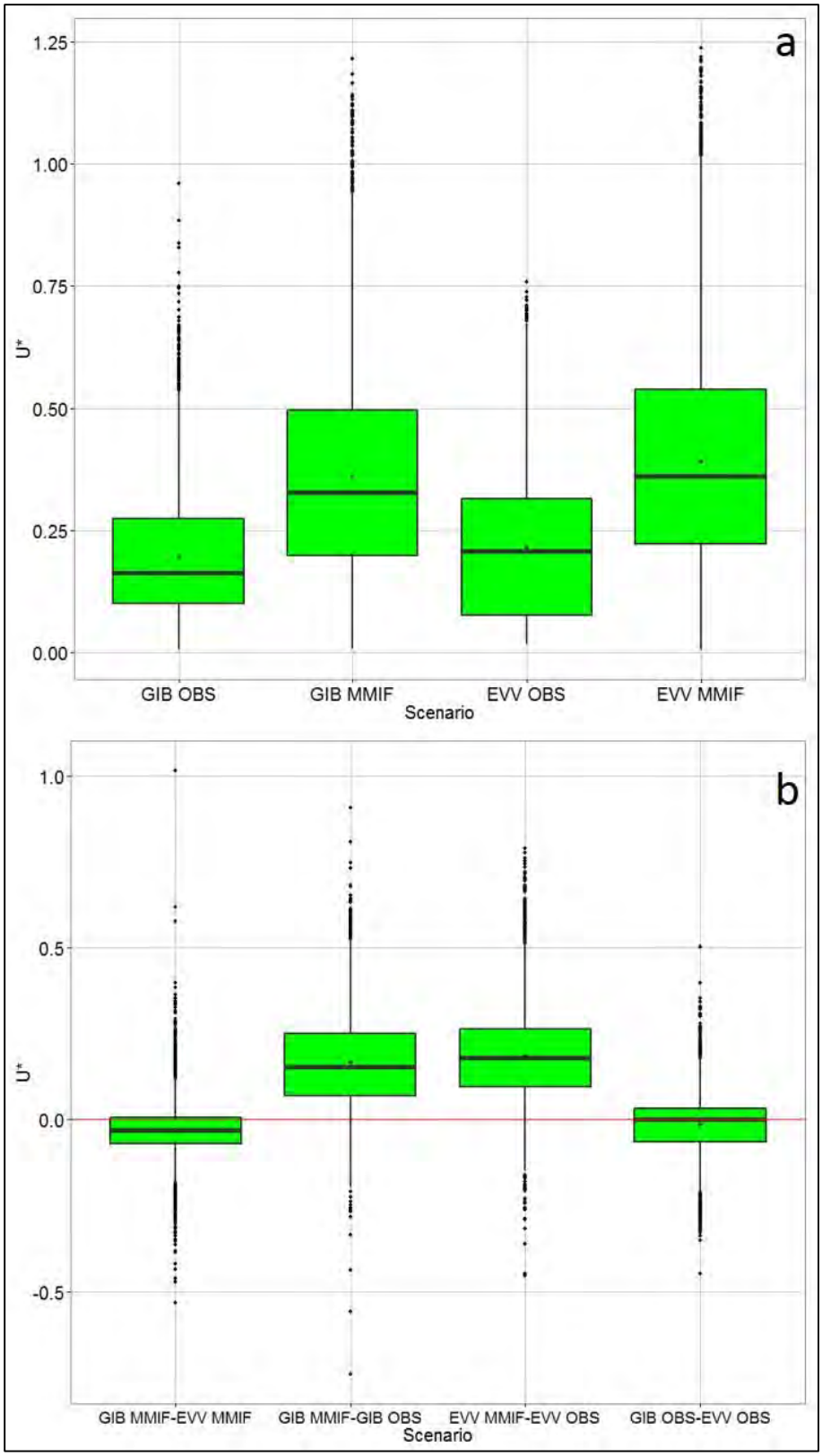

Figure A-8. Gibson surface friction velocity, u* (m/s): a) annual distributions and b) bias distributions.

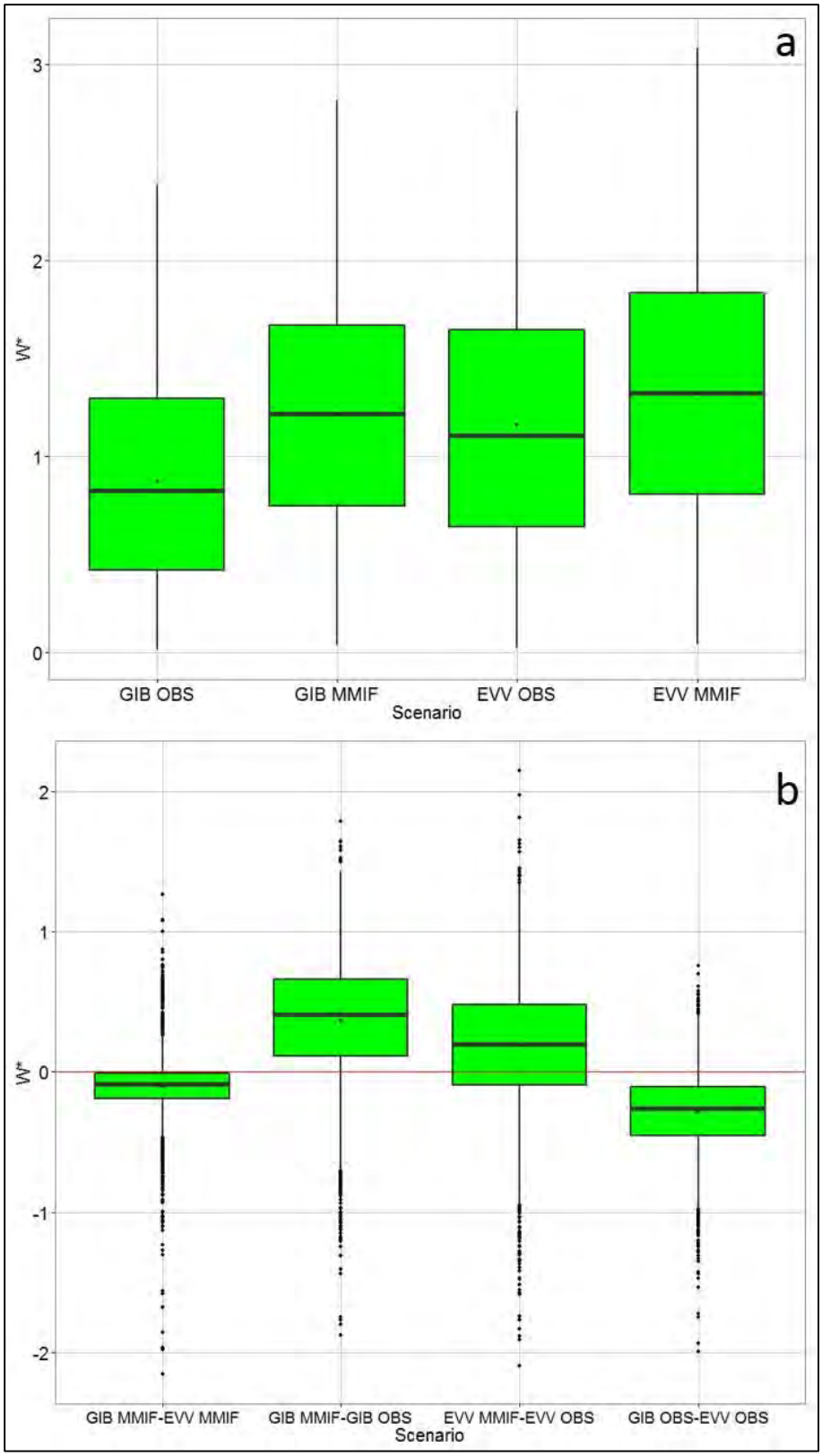

Figure A-9. Gibson convective velocity scale, w* (m/s): a) annual distributions and b) bias distributions.

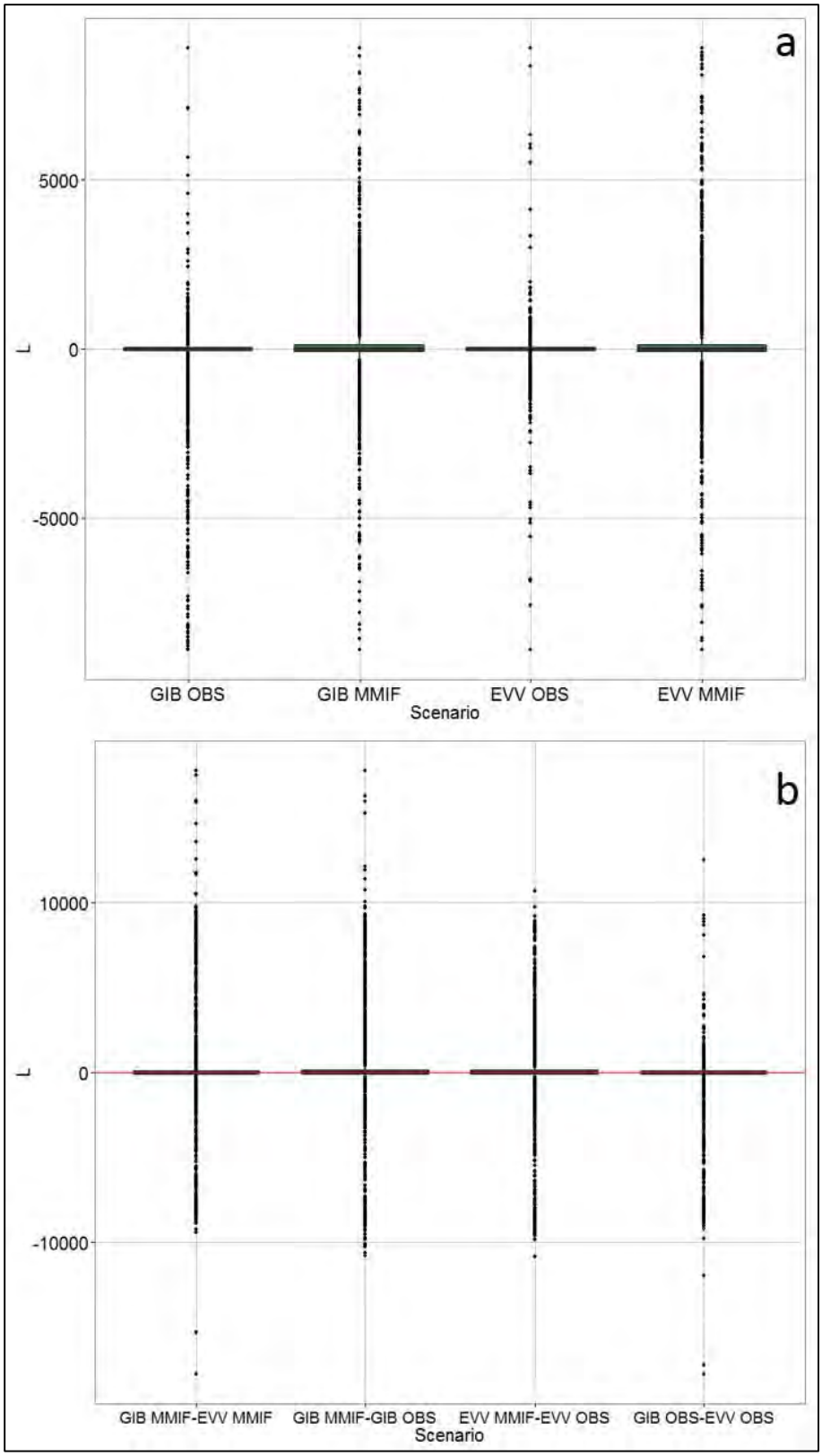

Figure A-10. Gibson Monin-Obukhov length (m): a) annual distributions and b) bias distributions.

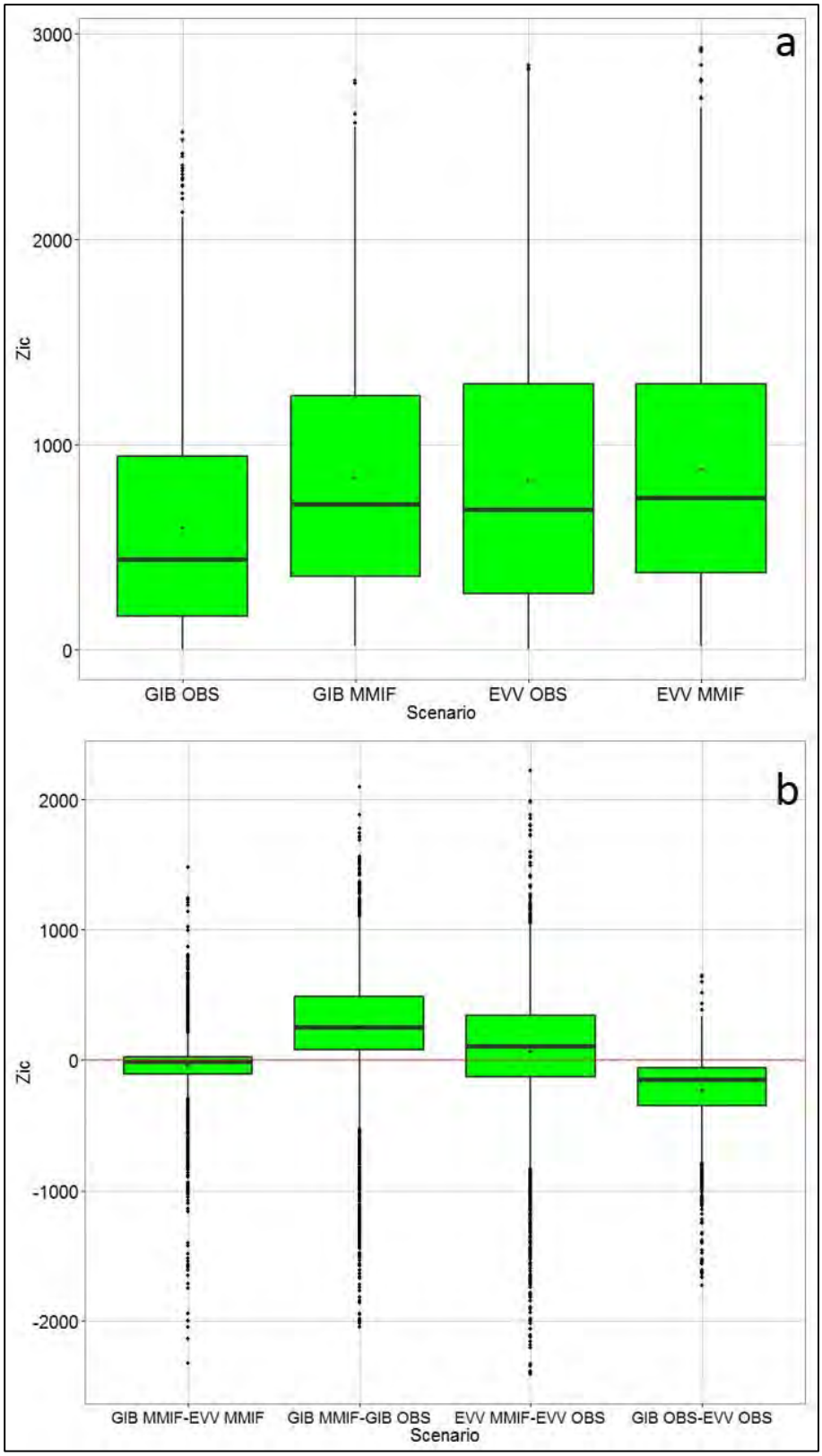

Figure A-11. Gibson convective mixing height (m): a) annual distributions and b) bias distributions.

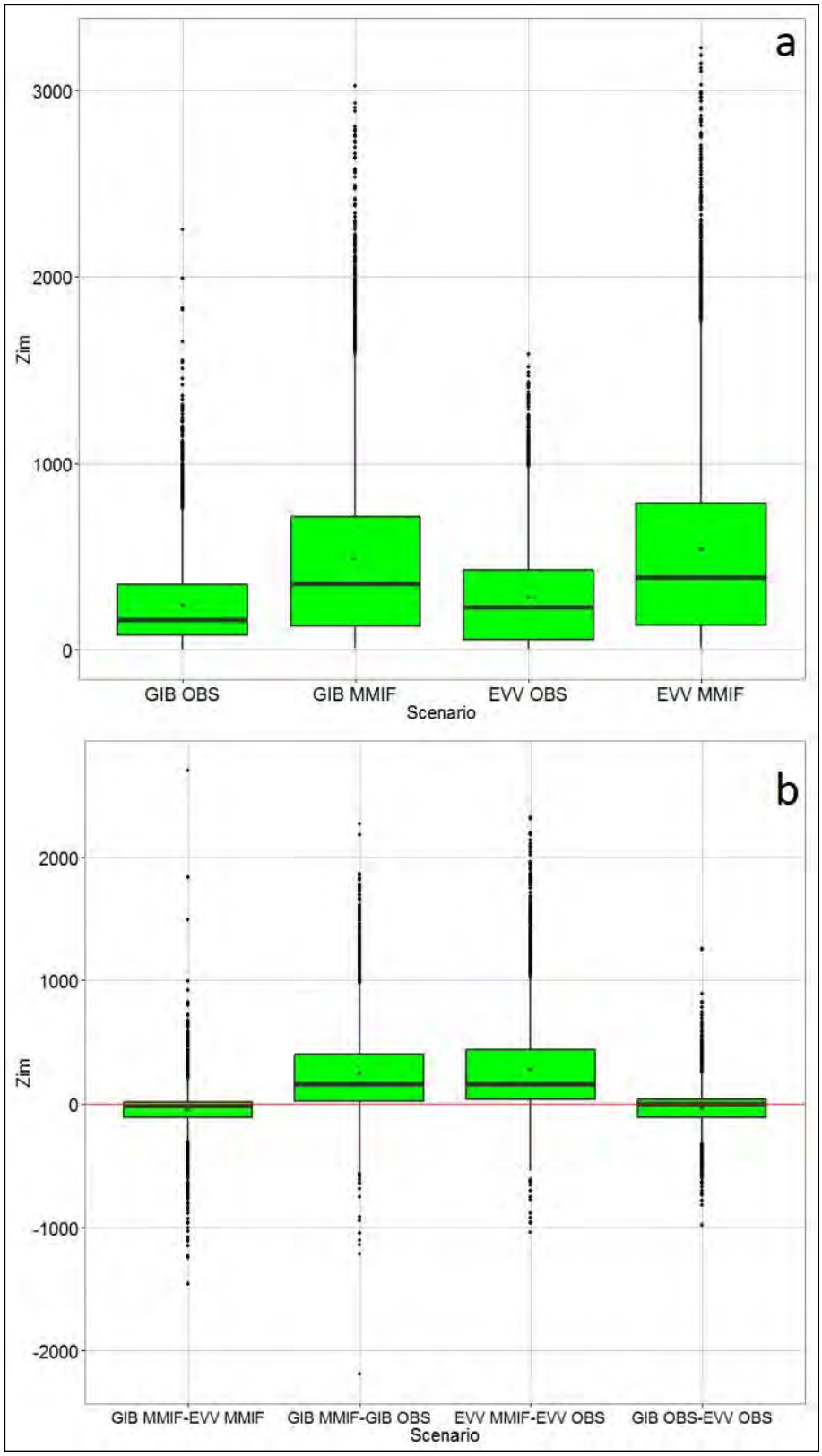

Figure A-12. Gibson mechanical mixing height (m): a) annual distributions and b) bias distributions.

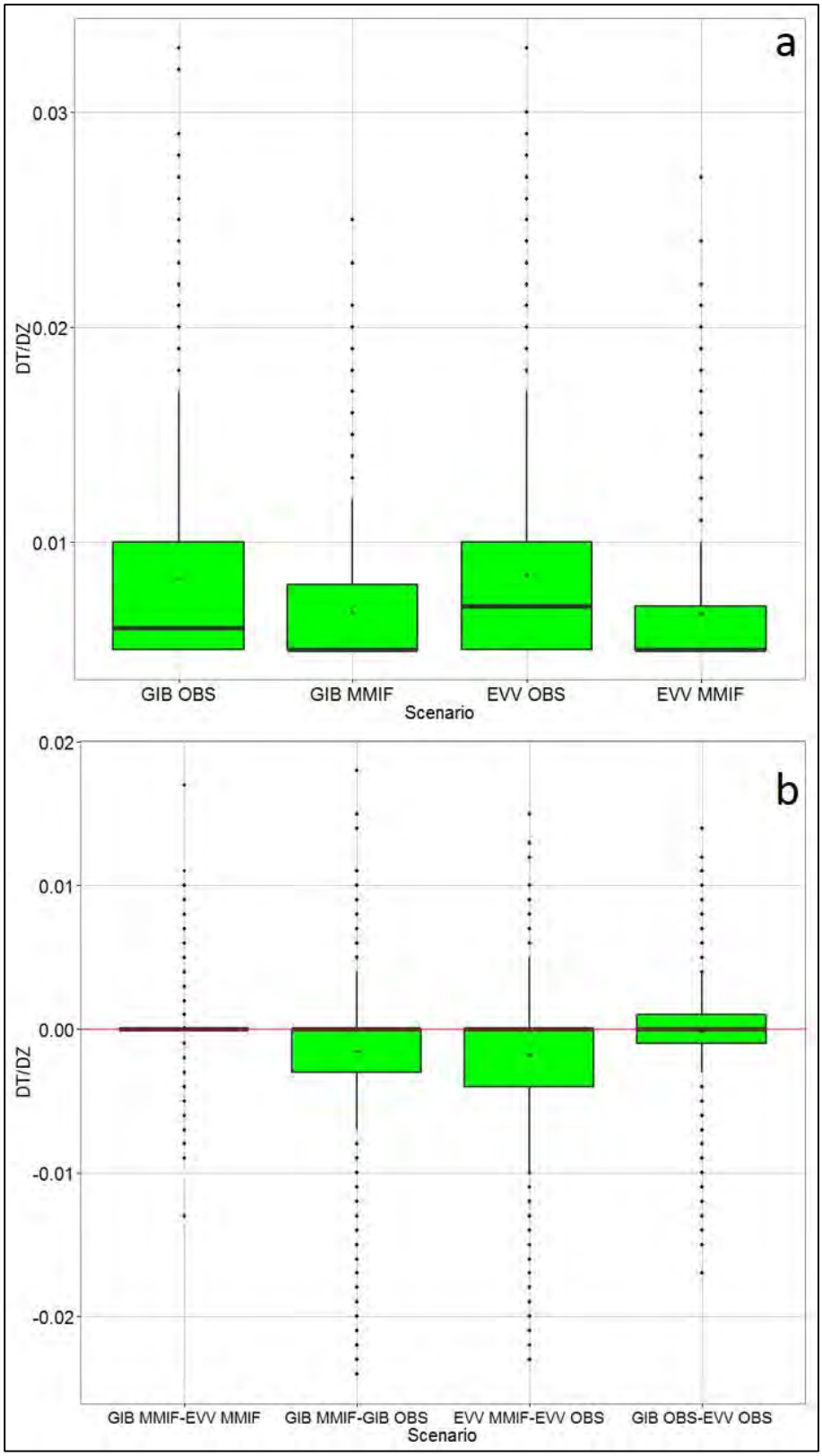

Figure A-13. Gibson potential temperature gradient (K/m) above Zic: a) annual distributions and b) bias distributions.

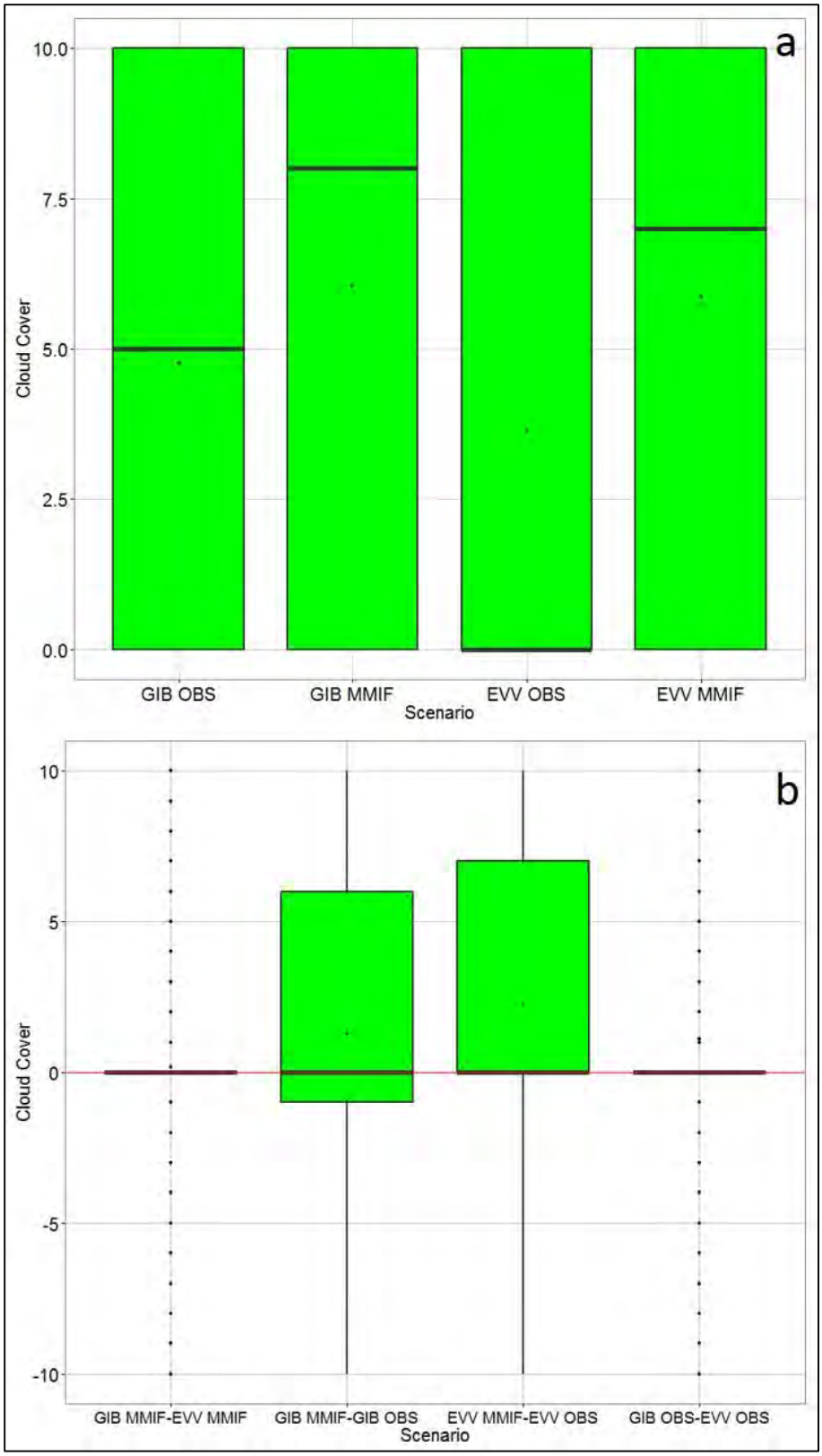

Figure A-14. Gibson cloud cover (tenths*10): a) annual distributions and b) bias distributions.

A.2 Martins Creek

Figures A-15 through A-20 show the distributions of the AERMET input variables for the Martins Creek study and Figures A-21-27 show the calculated variables from AERMET. As with the Gibson plots, panel b of each figure shows key biases. The first two bias distributions show the biases of the two MMIF scenarios against the Martins Creek site-specific data. The third fourth distribution compares the two MMIF scenarios against each other, and the fourth distribution compares the two observed datasets and acts as a control. The following can be seen in the plots:

- For wind speed (Figure A-15), Martins creek tends to show lower wind speeds (Figure A-15.a) among the four datasets. The bias distributions appear very similar between the MMIF scenarios and Martins Creek and ABE and Martins creek.
- Temperature (K) distributions (Figure A-16) indicate very similar patterns for all scenarios with the bias distributions tending to show fairly unbiased results among the scenario (median bias near zero degrees).
- The pressure distribution (figure A-17) appear to be similar with the MMIF 1 km scenario exhibiting slightly higher values.
- Relative humidity (Figure A-18) is identical for Martins Creek and ABE are identical being due to ABE being the source of RH for Martins Creek in AERMET. The two MMIF scenarios are very similar. Their bias distributions are very similar.
- Daytime albedo (Figure A-19) show comparable values for ABE and the MMIF scenarios. Martins Creek has a wider distribution of values. Bias distributions for the two MMIF scenarios are tighter than for the ABE – Martins Creek bias distribution.
- Bowen ratios (Figure A-20) vary widely among the scenarios. This could be due to the subjectivity of the average, dry, and wet selections in AERSURFACE.
- Heat flux (Figure A-21) distributions show comparable distributions between Martins Creek and the MMIF 1 km scenario while ABE and the MMIF 4 km scenario appear similar. The bias distributions appear comparable to each other.
- Surface friction velocity (Figure A-22) distributions show comparable distributions among the scenarios. The bias distributions appear comparable to each other.
- Convective velocity scale (Figure A-23) distributions show comparable distributions between Martins Creek and the MMIF 1 km scenario while ABE and the MMIF 4 km scenario appear similar. The bias distribution for the MMIF 1km and Martins Creek is noticeably lower than the others which appear comparable to each other.
- Monin-Obukhov length, L (Figure A-24) distributions appear very similar across the scenarios with bias distributions being similar as well. It is difficult to detect differences, which could be due to large absolute values for near neutral conditions.
- Convective mixing height, Zic (Figure A-25) distributions and bias distributions indicate that the MMIF mixing heights were comparable to ABE and Martins Creek
- Mechanical mixing height, Zim (Figure A-26) distributions show similar behavior as the convective mixing heights.

- Potential temperature gradient ($d\Theta/dz$), (Figure A-27), distributions indicate smaller lapse rates for the MMIF data compared to the observed data.
- Cloud cover (Figure A-28) distributions show no differences between ABE and Martins Creek for the observed datasets since ABE cloud cover is used at Martins Creek. Cloud cover estimates for MMIF are calculated in AERMET. Differences between the MMIF and Martins Creek show a large swing in biases.

For the most part, the meteorological data processed through MMIF for Martins Creek appear reasonable when compared to Martins Creek observations.

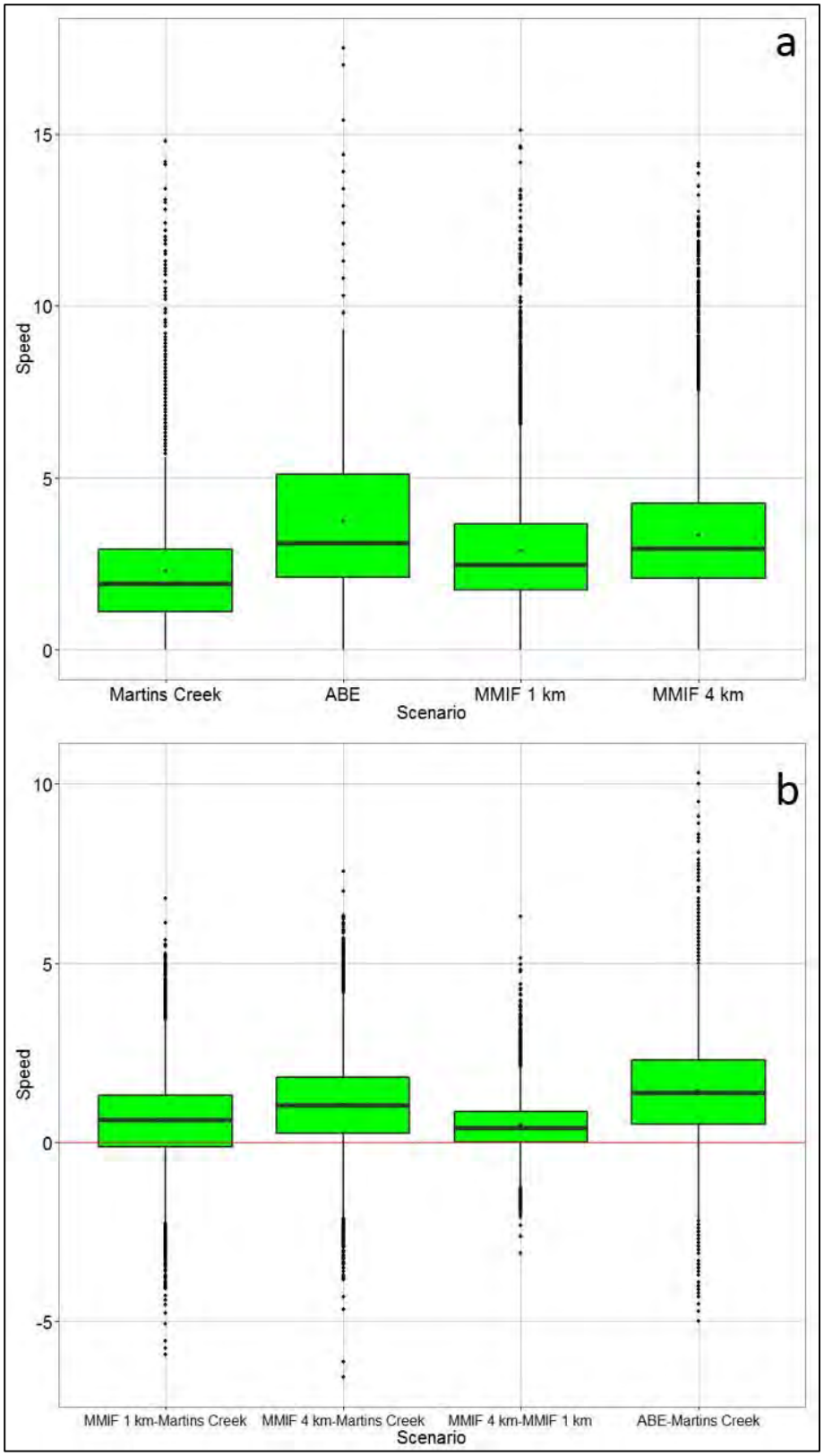

Figure A-15. Martins Creek wind speed (m/s): a) annual distributions and b) bias distributions.

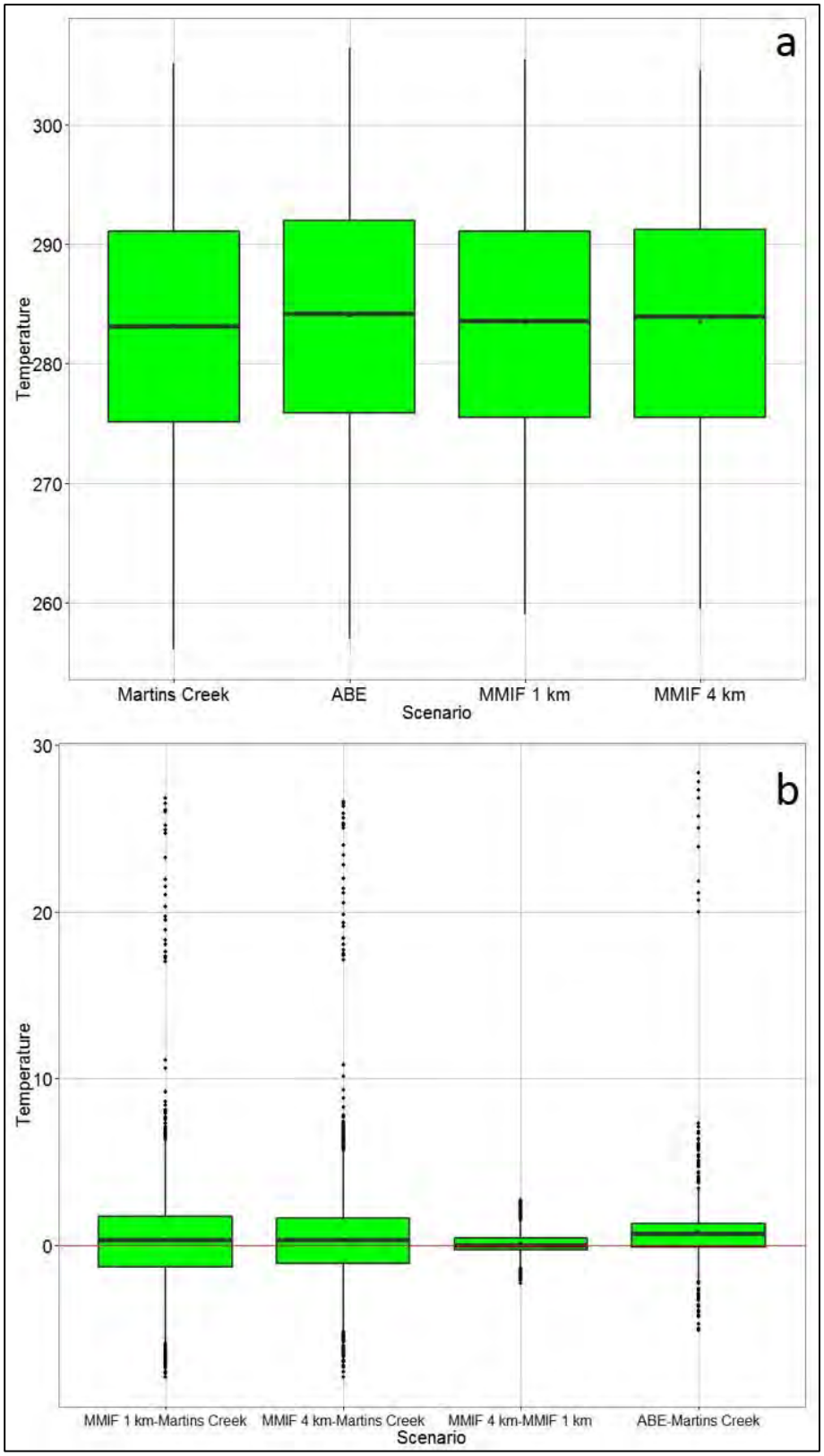

Figure A-16. Martins Creek ambient temperature (K): a) annual distributions and b) bias distributions.

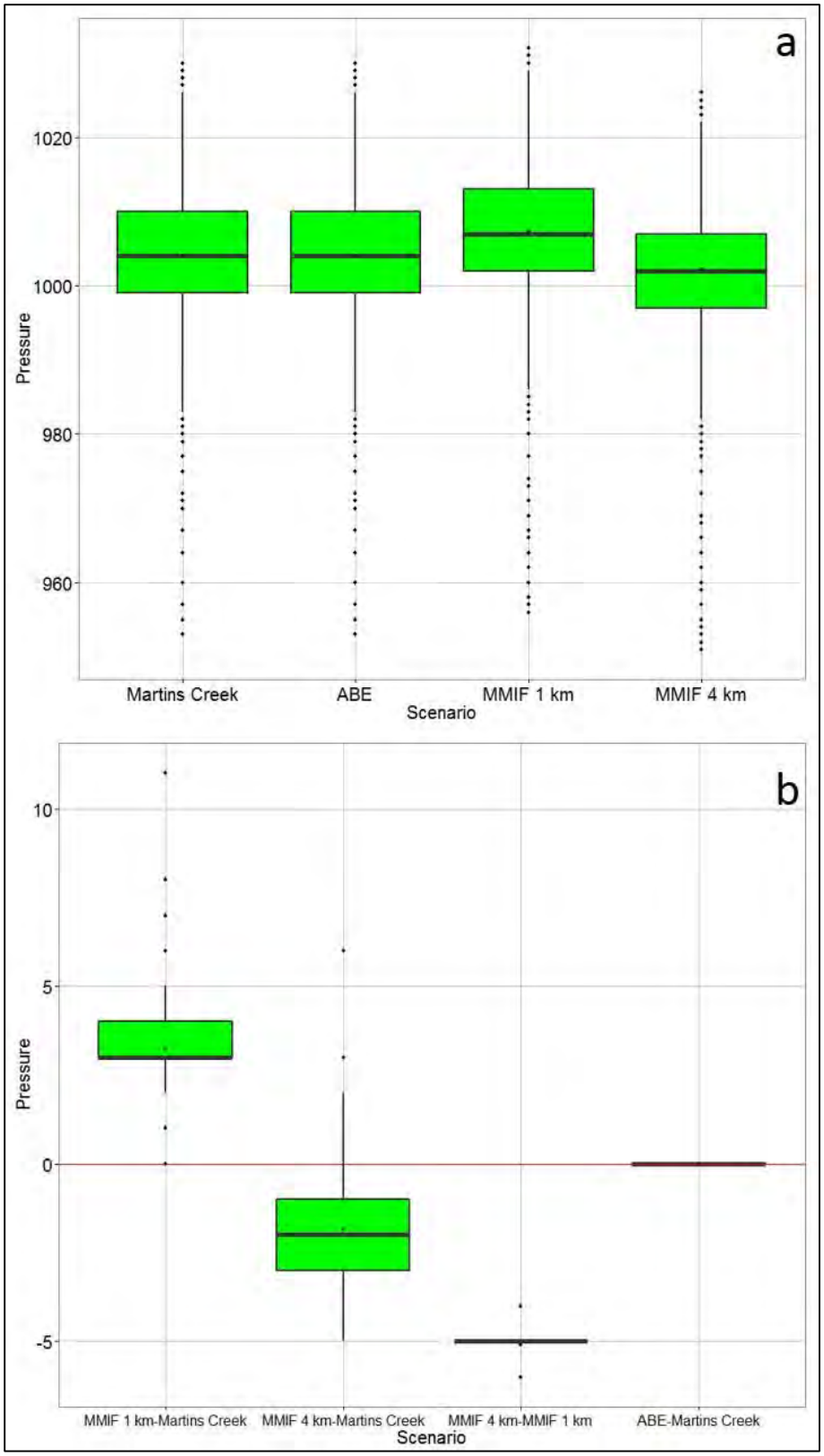

Figure A-17. Martins Creek station pressure (mb): a) annual distributions and b) bias distributions.

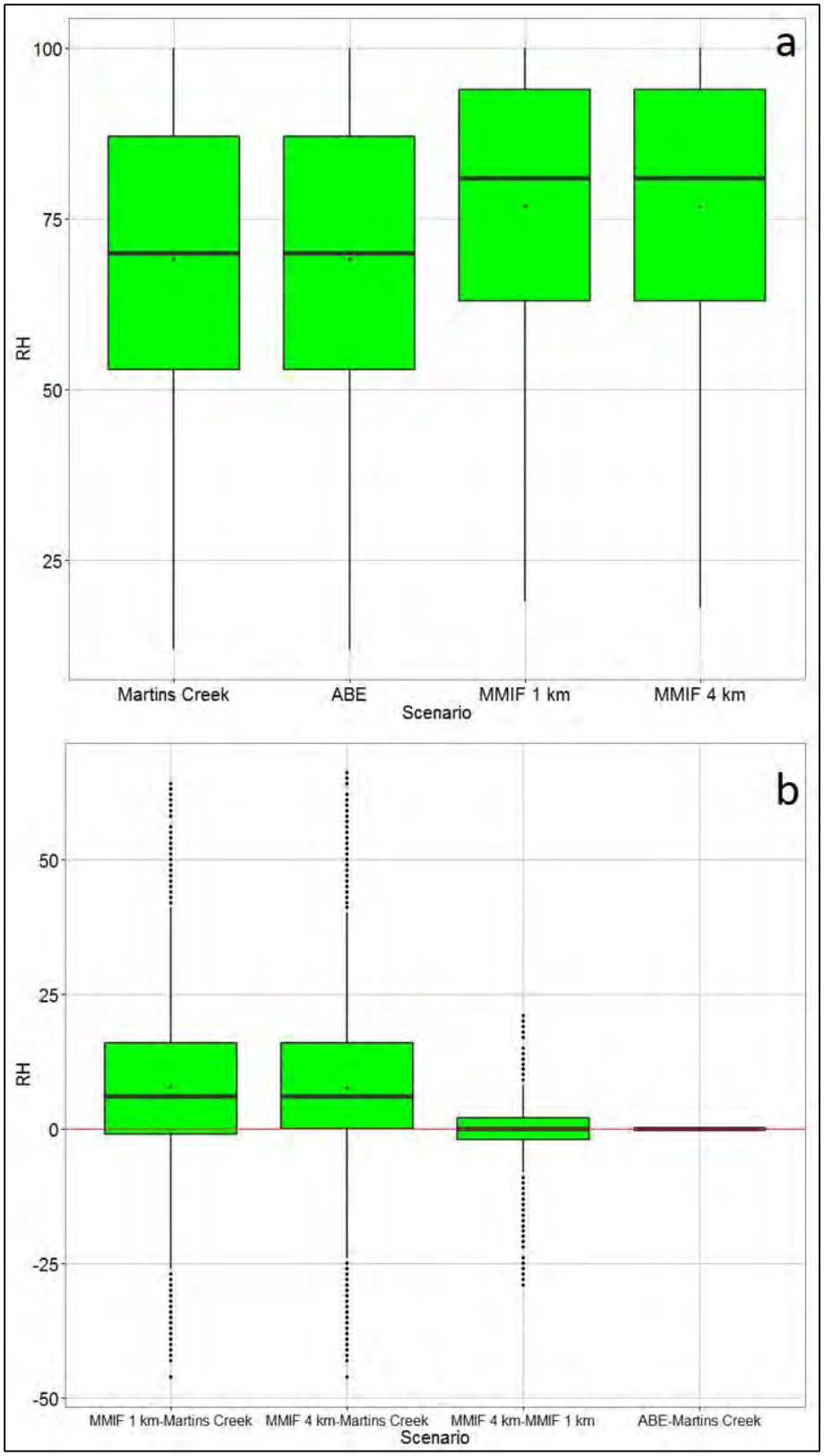

Figure A-18. Martins Creek relative humidity (percent): a) annual distributions and b) bias distributions.

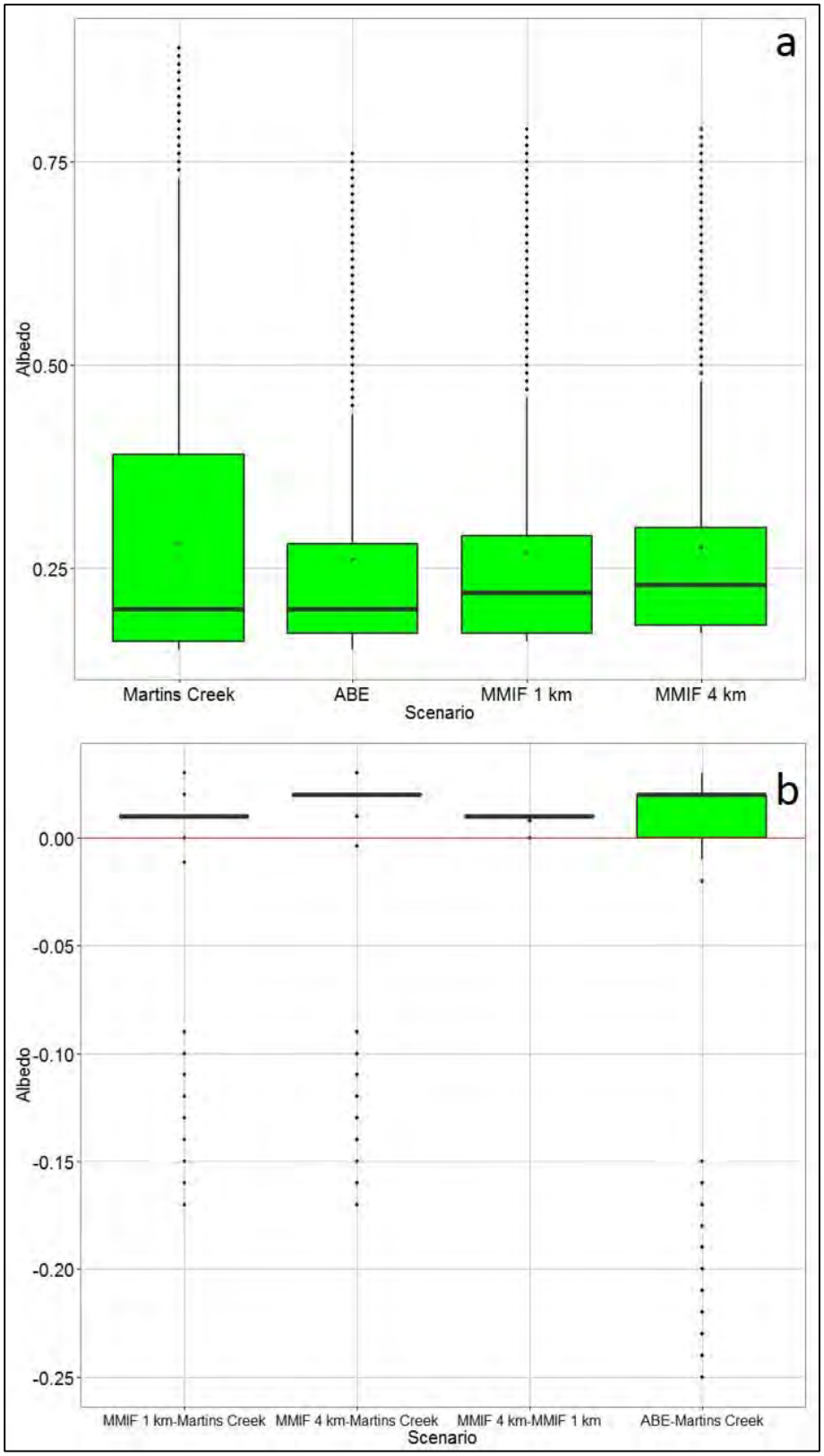

Figure A-19. Martins Creek daytime albedo (fraction): a) annual distributions and b) bias distributions.

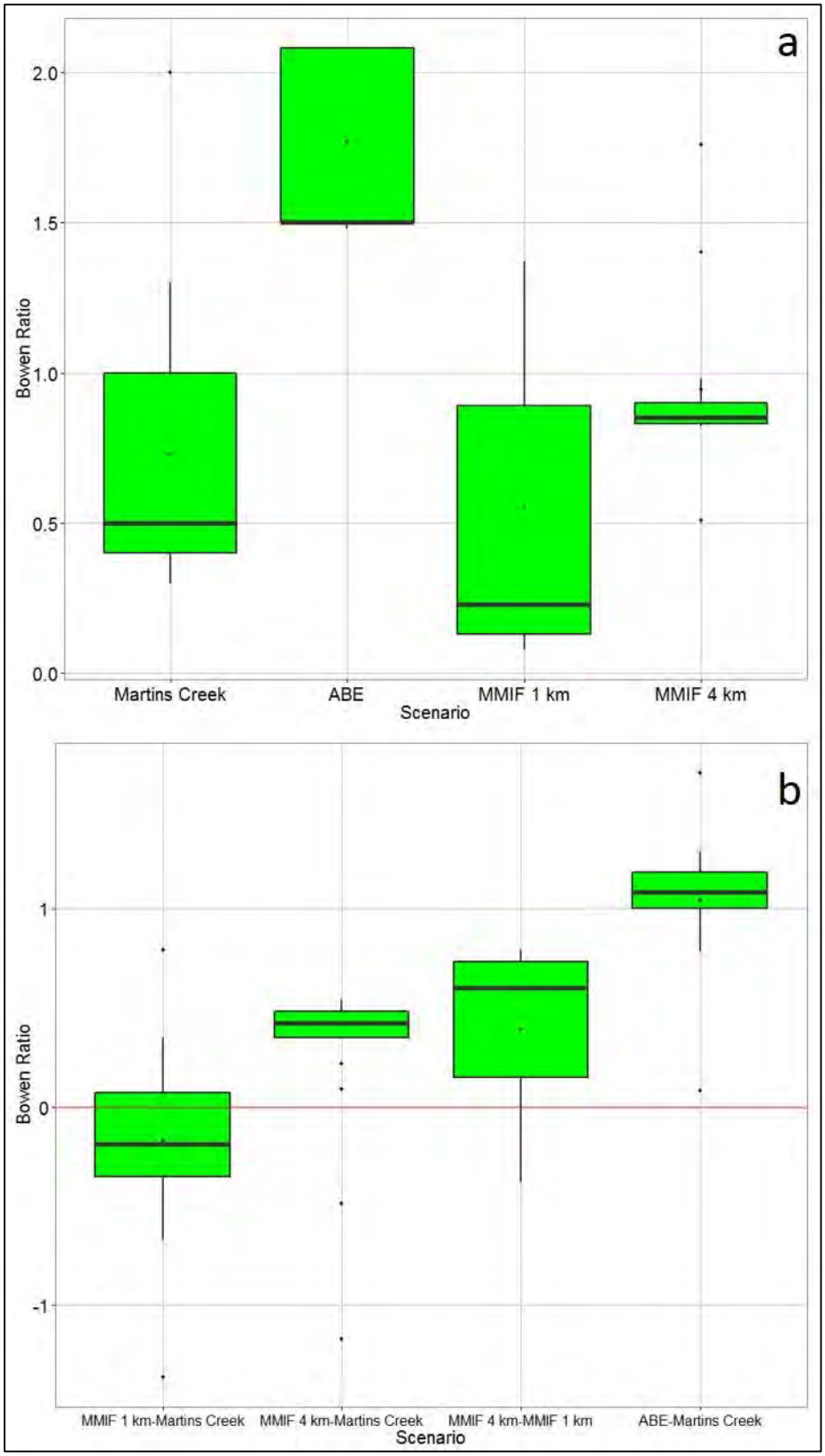

Figure A-20. Martins Creek Bowen ratio: a) annual distributions and b) bias distributions.

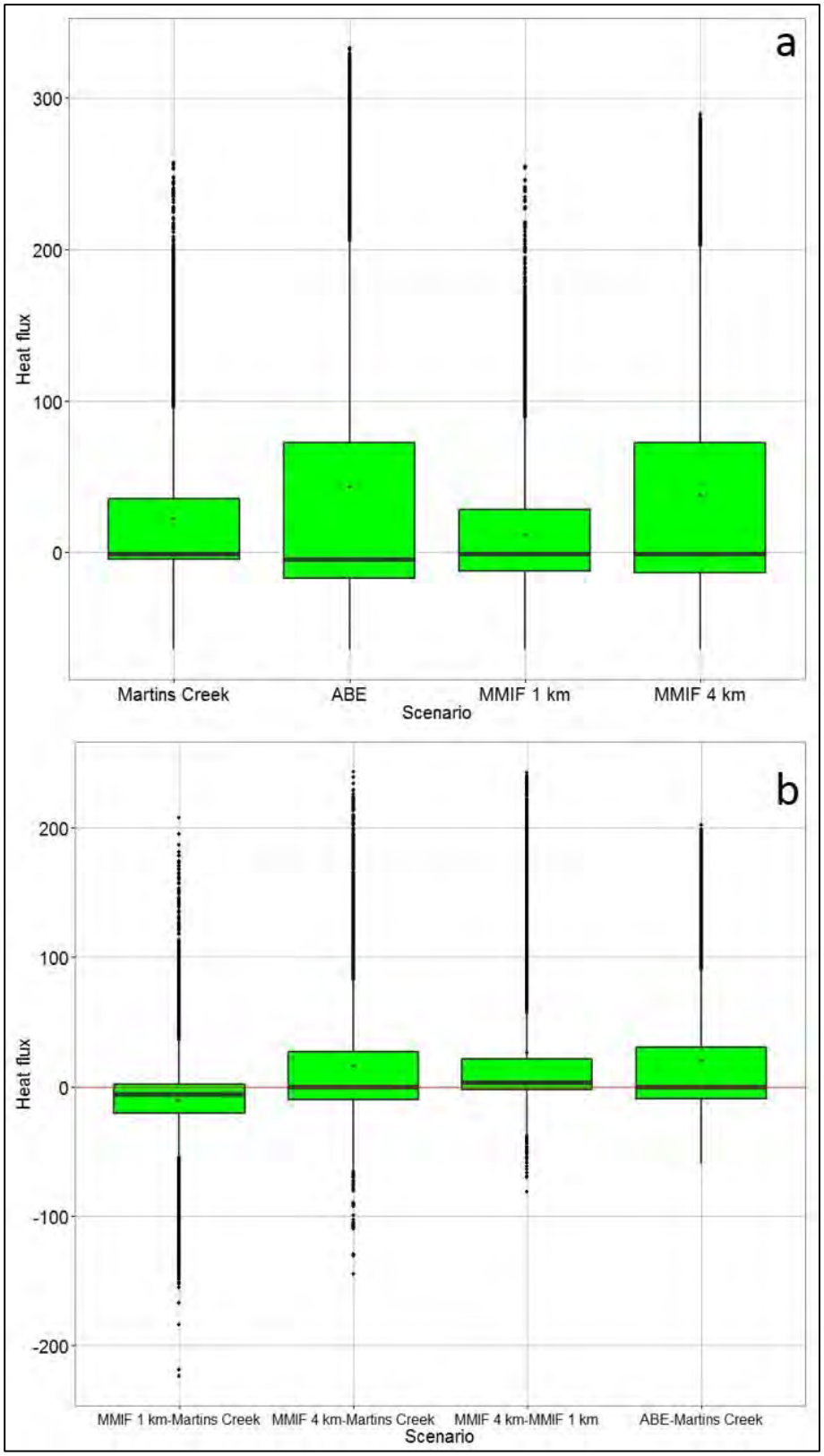

Figure A-21. Martins Creek heat flux (W/m^2): a) annual distributions and b) bias distributions.

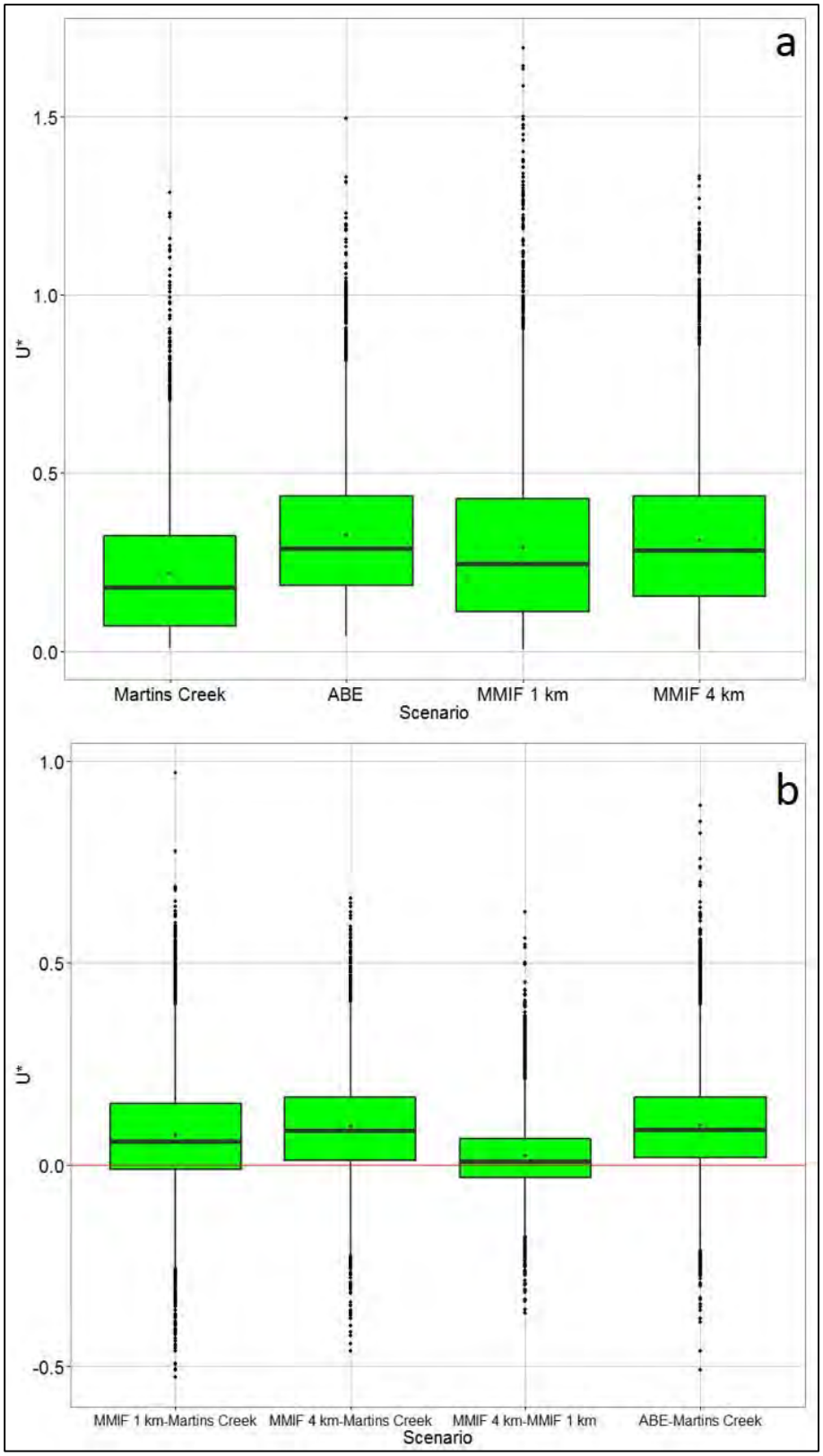

Figure A-22. Martins Creek surface friction velocity, u* (m/s): a) annual distributions and b) bias distributions.

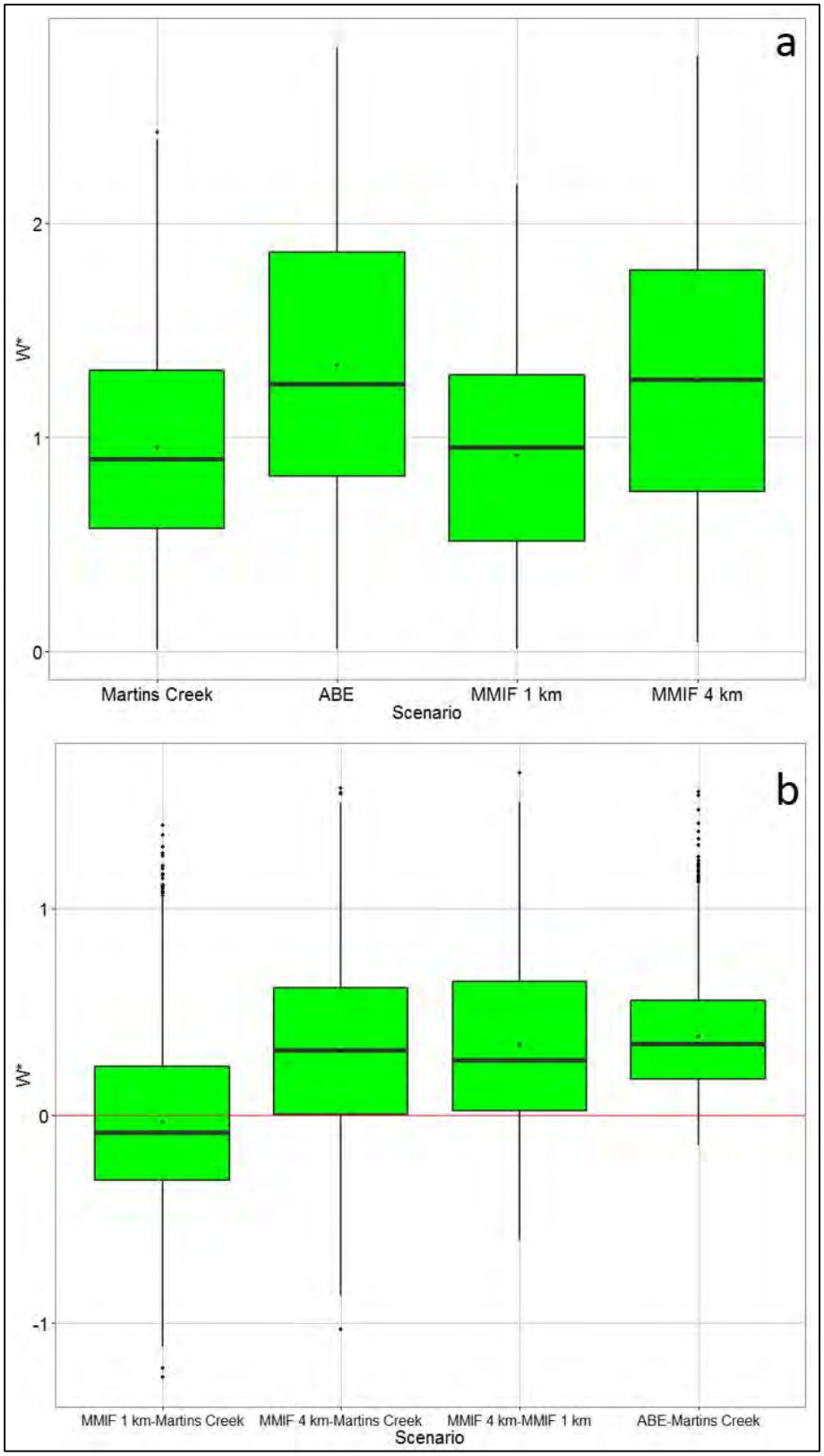

Figure A-23. Martins Creek convective velocity scale, w* (m/s): a) annual distributions and b) bias distributions

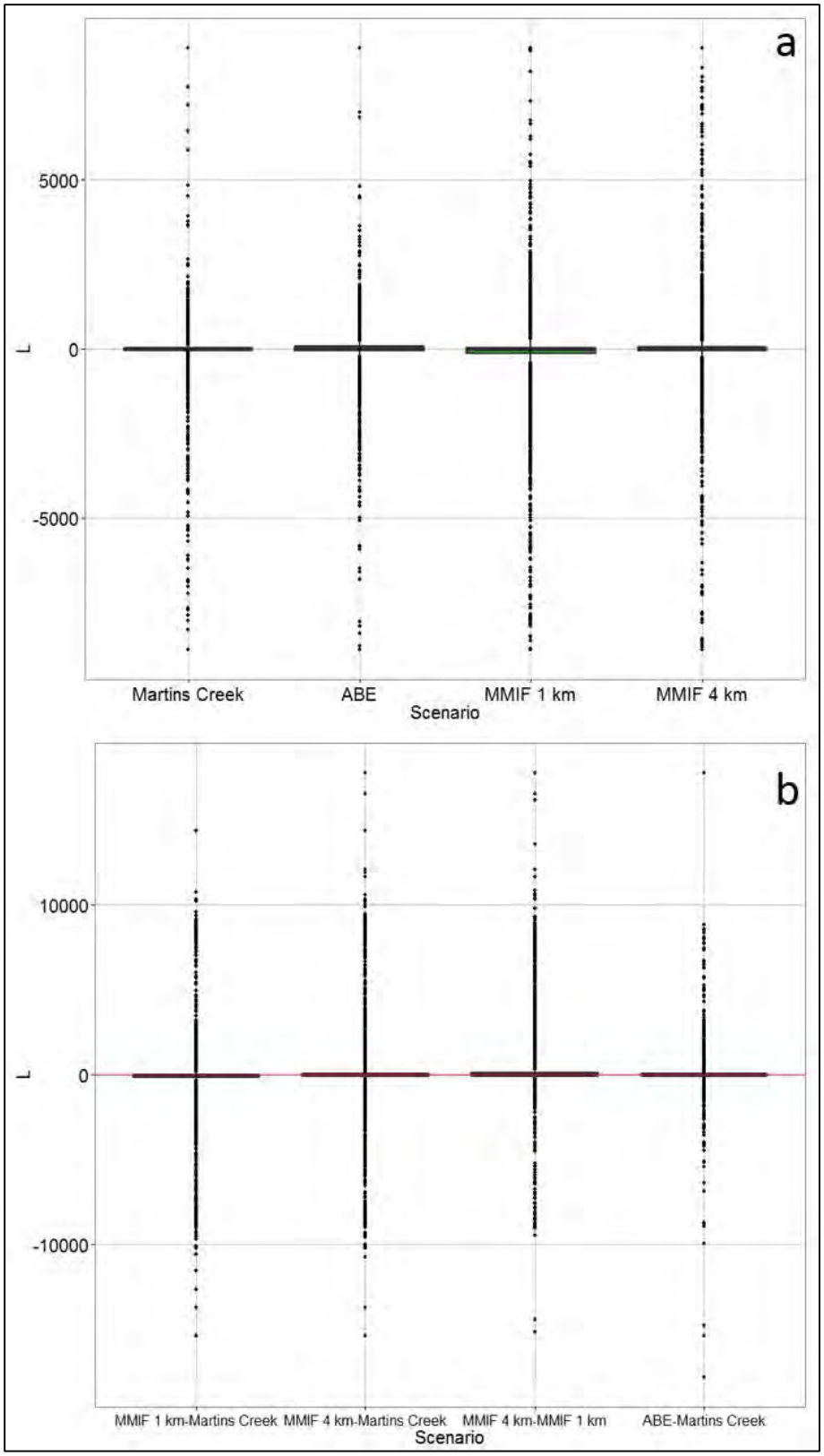

Figure A-24. Martins Creek Monin-Obukhov length (m): a) annual distributions and b) bias distributions.

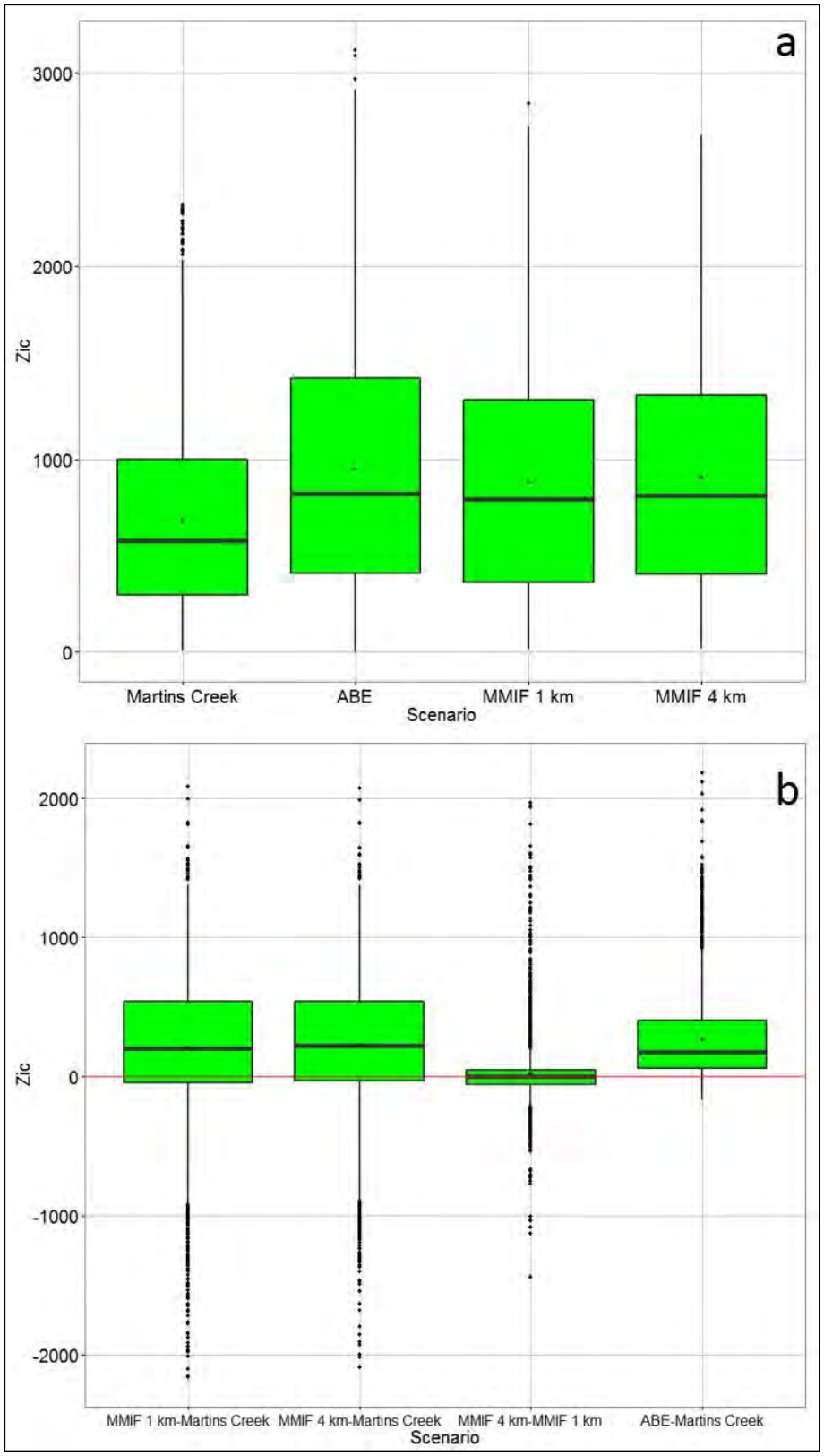

Figure A-25. Martins Creek convective mixing height (m): a) annual distributions and b) bias distributions.

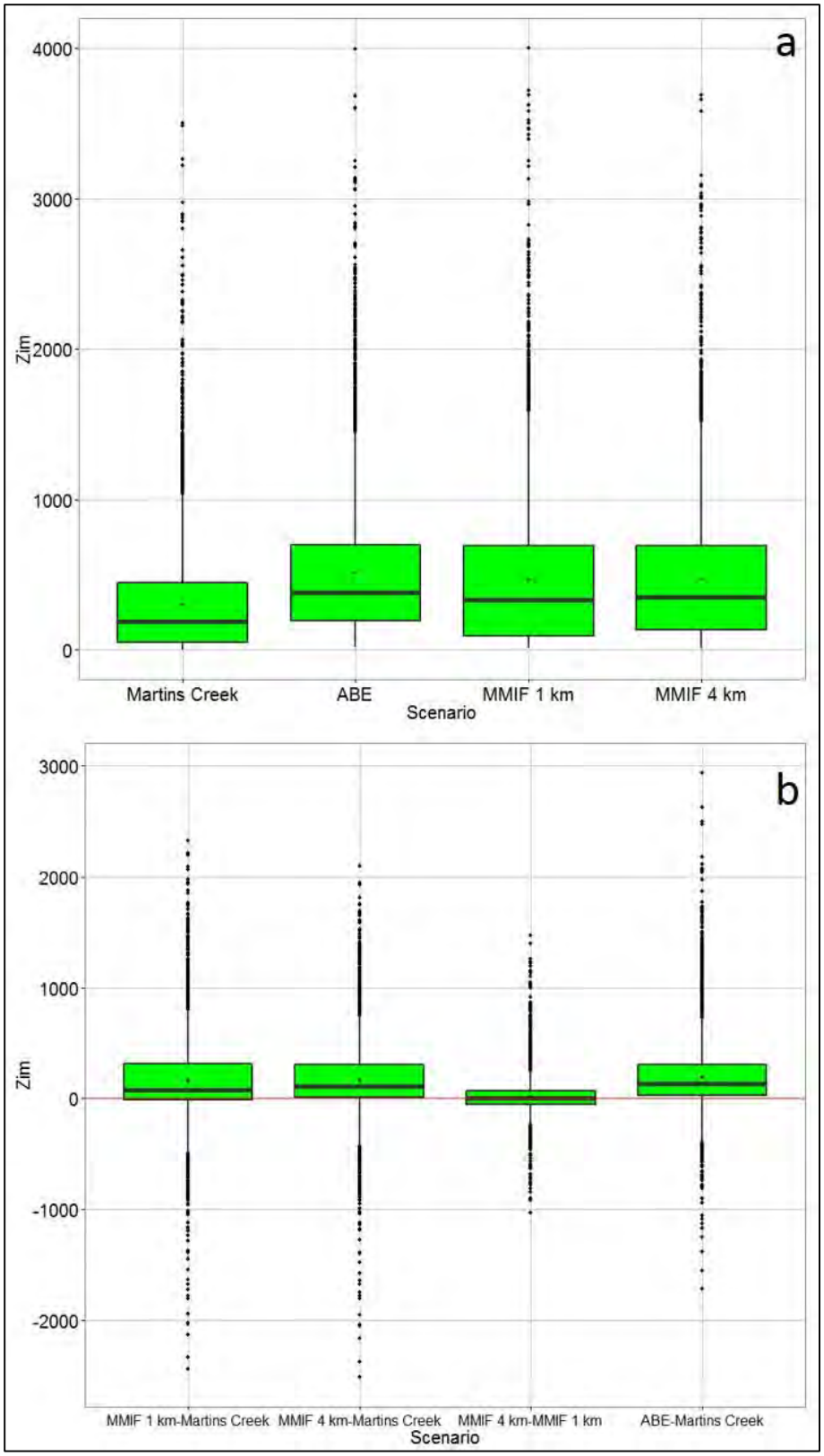

Figure A-26. Martins Creek mechanical mixing height (m): a) annual distributions and b) bias distributions.

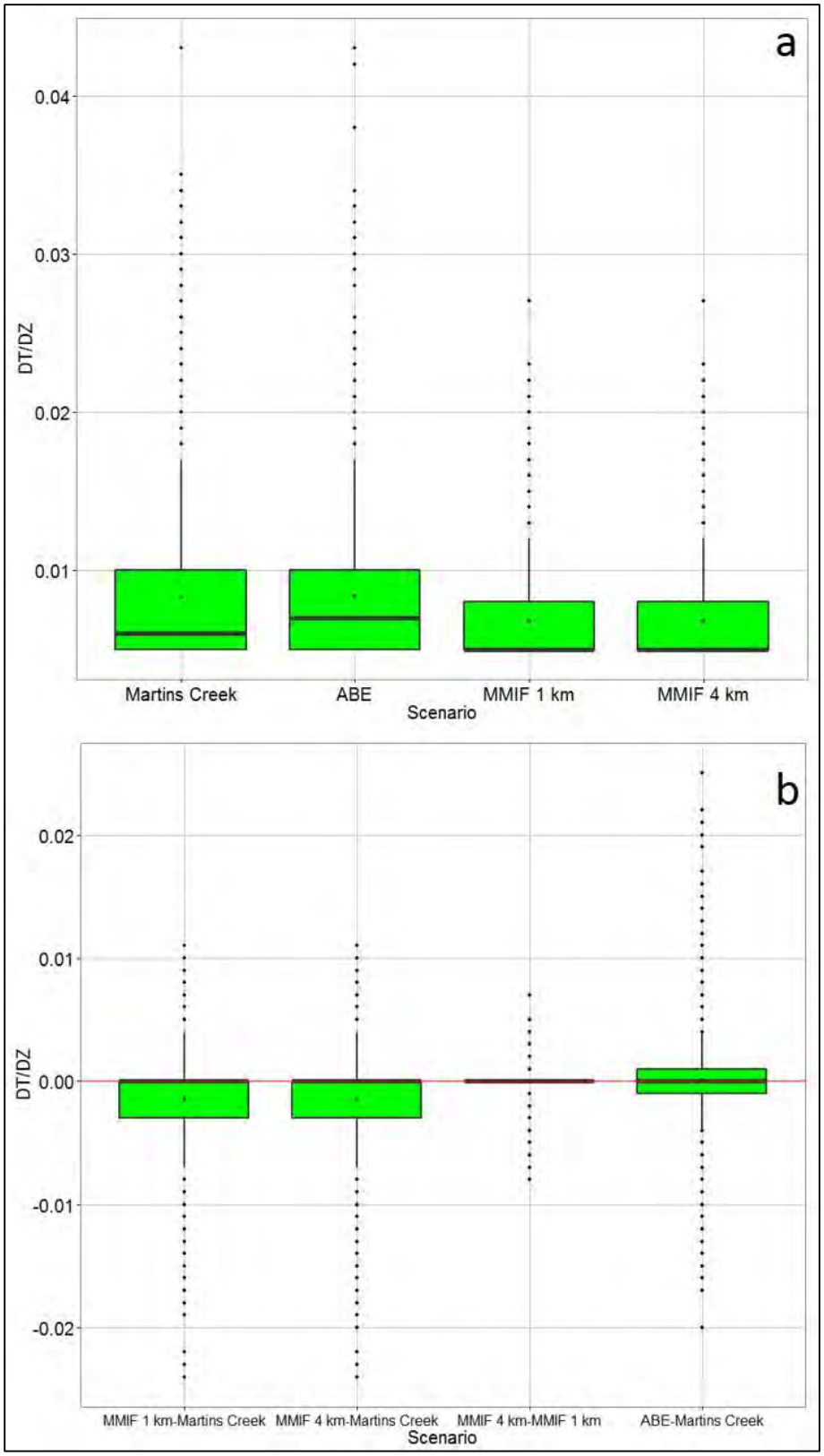

Figure A-27. Martins Creek potential temperature gradient (K/m) above Zic: a) annual distributions and b) bias distributions.

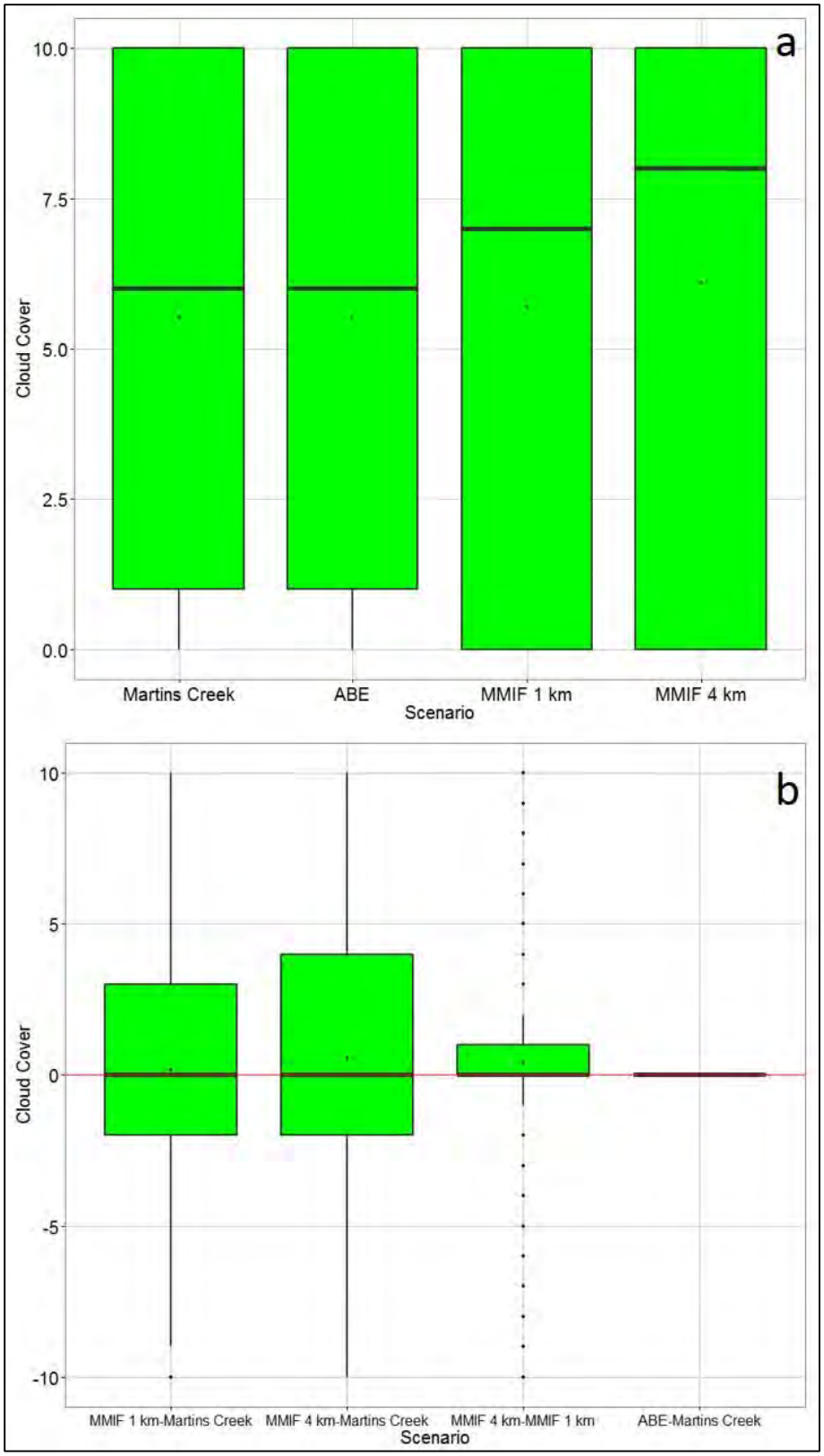

Figure A-28. Martins Creek cloud cover (tenths*10): a) annual distributions and b) bias distributions.

A.3 Herculaneum

Figures A-29 through A-34 show the distributions of the AERMET input variables for the Herculaneum study and Figures A-35 through A-41 show the calculated variables from AERMET. As with the other evaluation plots, panel b of each figure shows key biases. The first three bias distributions show the biases of the three MMIF scenarios against the Herculaneum site-specific data. The fourth through sixth distribution compares the three MMIF scenarios against each other, and the final distribution compares the two observed datasets and acts as a control. The following can be seen in the plots:

- For wind speed (Figure A-29), Herculaneum tends to show lower wind speeds (Figure A-29.a) among the datasets. The bias distributions appear very similar between the MMIF scenarios and Herculaneum and CPS and Herculaneum. Biases between the three MMIF scenarios are relatively low and comparable to each other.
- Temperature (K) distributions (Figure A-30) indicate very similar patterns for all scenarios with the bias distributions tending to show fairly unbiased results among the scenario (median bias near zero degrees).
- The pressure distribution (Figure A-31) appear to be similar among the two observed datasets and the MMIF 4 km scenario. Pressures tend to be lower with the 12 and 36 km MMIF scenarios.
- Relative humidity (Figure A-32) are comparable for Herculaneum and CPS. The three MMIF scenarios are very similar. Their bias distributions are also very similar and comparable to the CPS – Herculaneum bias distribution.
- Daytime albedo (Figure A-33) show comparable values among the scenarios. The bias distributions for the MMIF 4 km scenario shows a more positive bas than the other two MMIF scenarios.
- Bowen ratios (Figure A-34) are comparable for the two observed datasets and comparable among the three MMIF scenarios. MMIF bias distributions show more differences than the CPS –Herculaneum biases but are comparable.
- Heat flux (Figure A-35) distributions show comparable among the scenarios. The bias distributions appear comparable to each other.
- Surface friction velocity (Figure A-36) distributions show comparable distributions among the scenarios. For the most part, the bias distributions appear comparable to each other with differences in the ranges of the biases.
- Convective velocity scale (Figure A-37) distributions show comparable distributions between Herculaneum and CPS as well as the 12 and 36 km MMIF scenarios. The bias distributions for the MMIF scenarios (compared to Herculaneum) are slightly higher than the CPS – Herculaneum bias distribution.
- Monin-Obukhov length, L (Figure A-38) distributions appear very similar across the scenarios with bias distributions being similar as well. It is difficult to detect differences, which could be due to large absolute values for near neutral conditions.
- Convective mixing height, Zic (Figure A-39) distributions and bias distributions indicate that the MMIF mixing heights were higher than the two observed datasets. The MMIF

bias distributions (compared to Herculaneum) are noticeably higher than the CPS – Herculaneum bias distribution.

- Mechanical mixing height, Zim (Figure A-40) distributions show comparable distributions and bias distributions among the scenarios.
- Potential temperature gradient ($d\Theta/dz$), (Figure A-41), indicate lower lapse rates for the two observed datasets when compared to the MMIF scenarios. The bias distributions for the MMIF scenarios (compared to Herculaneum) show a more positive bias when compared to the CPS – Herculaneum bias distribution.
- Cloud cover (Figure A-42) distributions show no differences between CPS and Herculaneum since CPS cloud cover is used at Herculaneum. Cloud cover estimates for MMIF are calculated in AERMET. Differences between the MMIF and Herculaneum show a large swing in biases with a tendency to positive bias.

For the most part, the meteorological data processed through MMIF for Herculaneum appear reasonable when compared to Herculaneum and CPS observations.

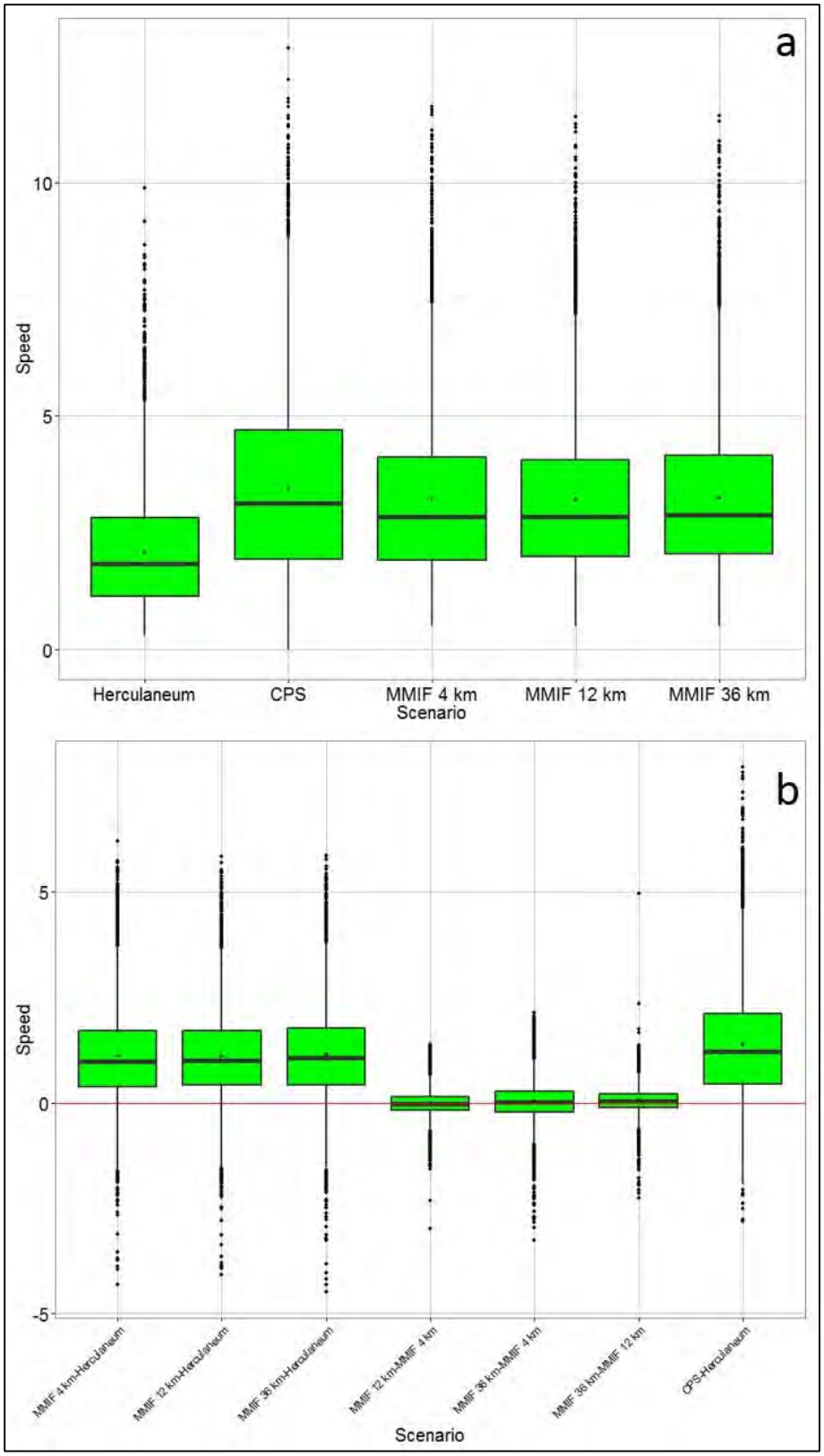

Figure 29. Herculaneum wind speed (m/s): a) annual distributions and b) bias distributions.

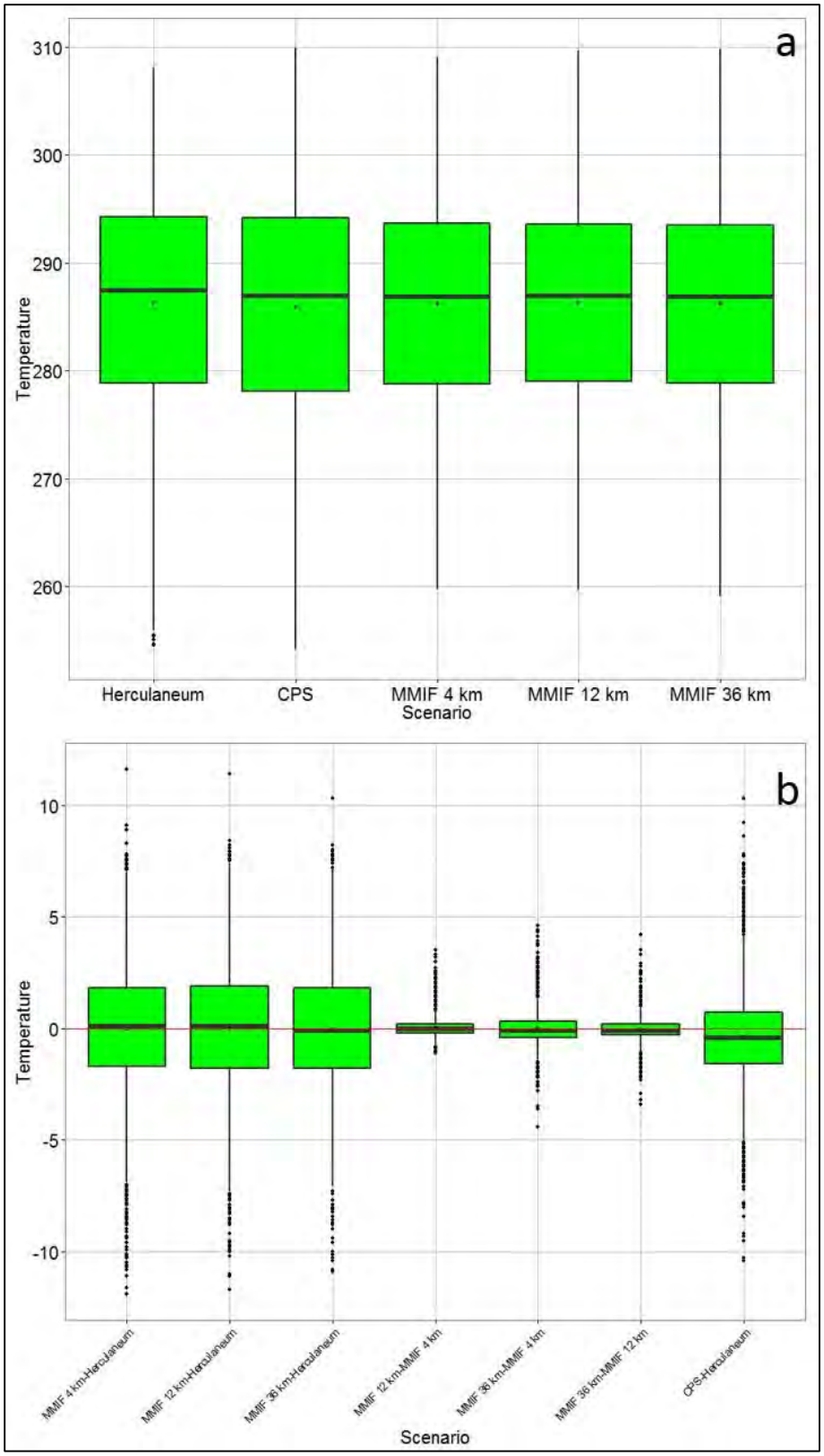

Figure A-30. Herculaneum ambient temperature (K): a) annual distributions and b) bias distributions.

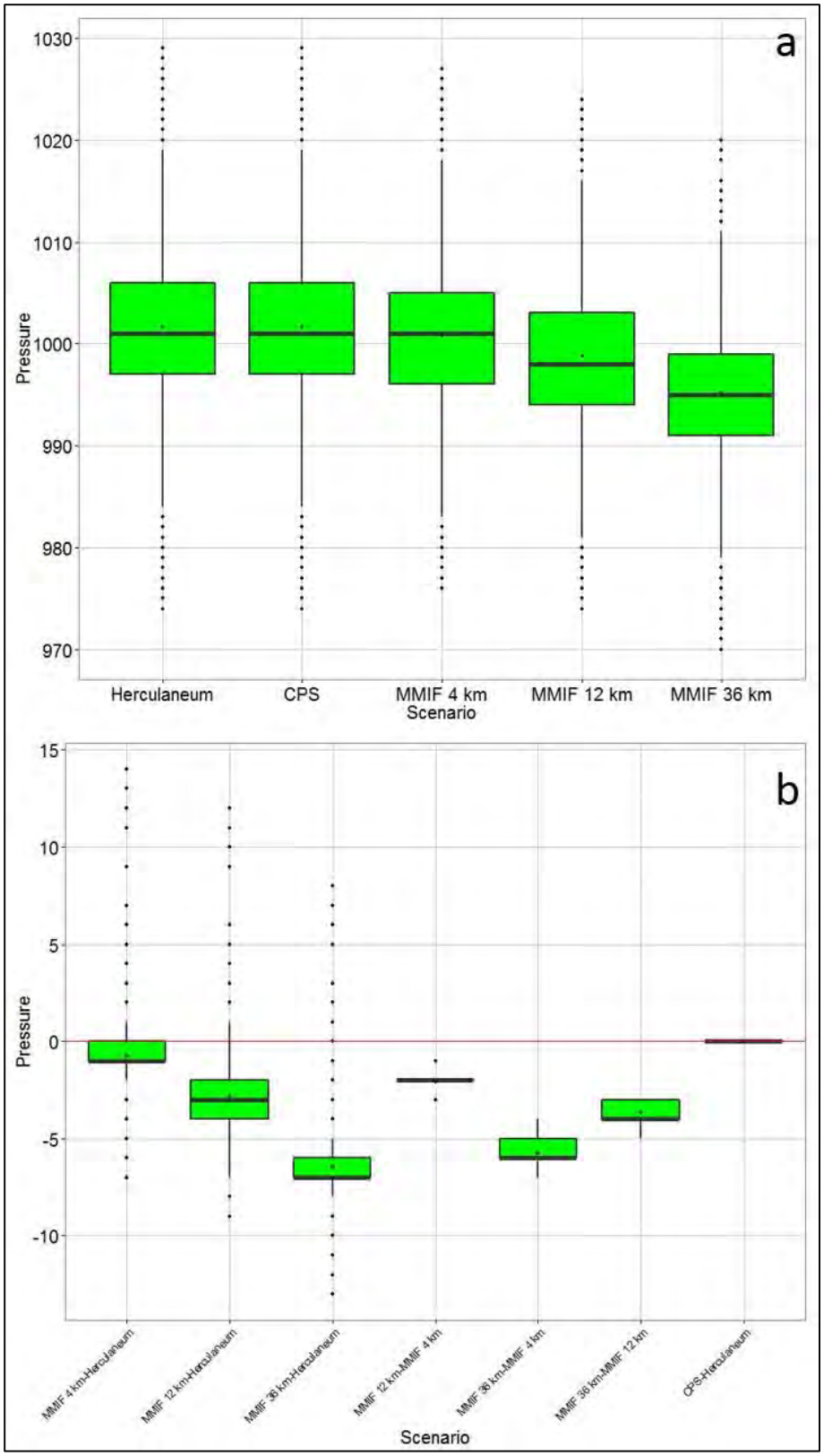

Figure A-31. Herculaneum station pressure (mb): a) annual distributions and b) bias distributions.

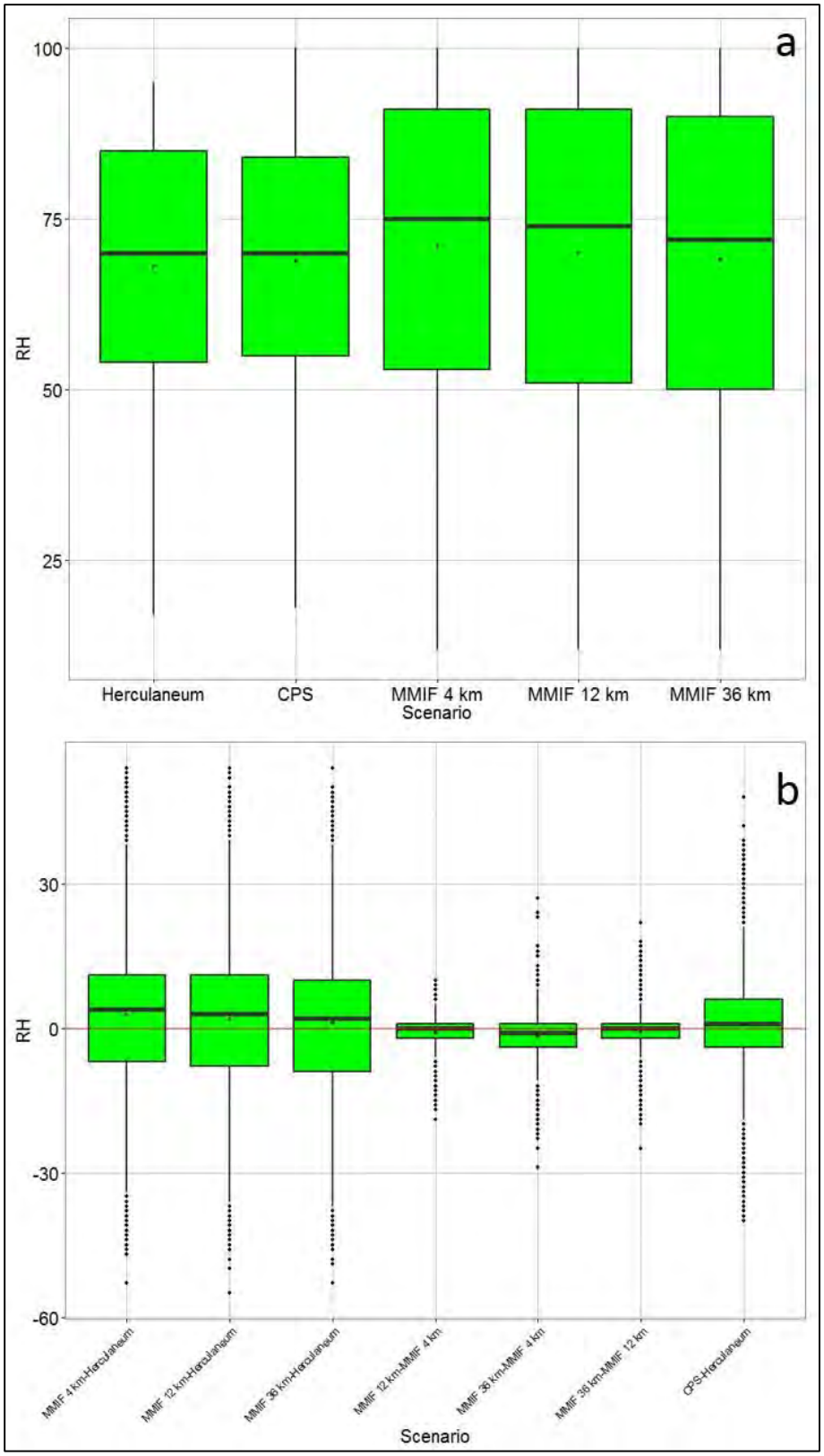

Figure A-32. Herculaneum relative humidity (percent): a) annual distributions and b) bias distributions.

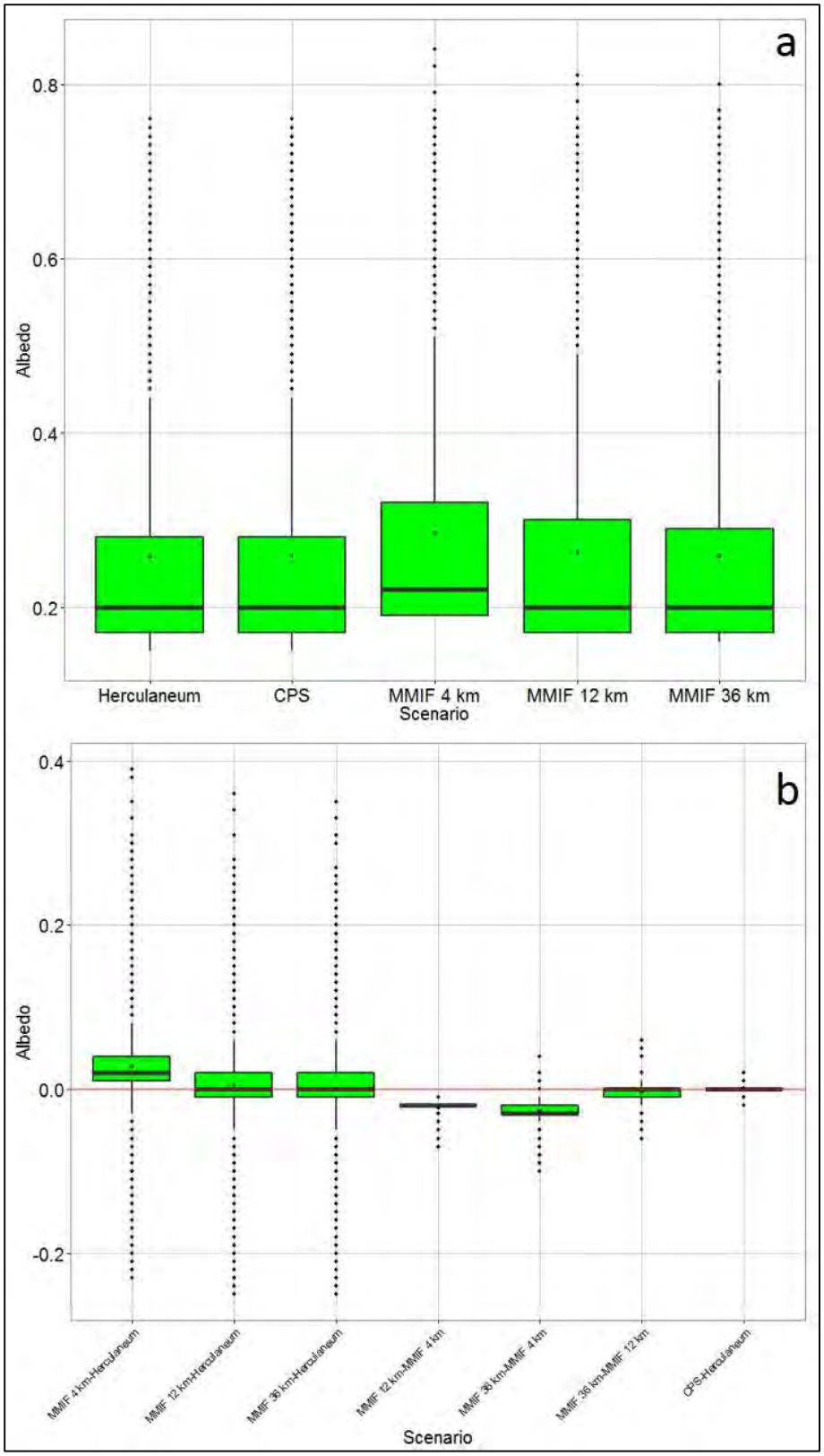

Figure A-33. Herculaneum daytime albedo (fraction): a) annual distributions and b) bias distributions.

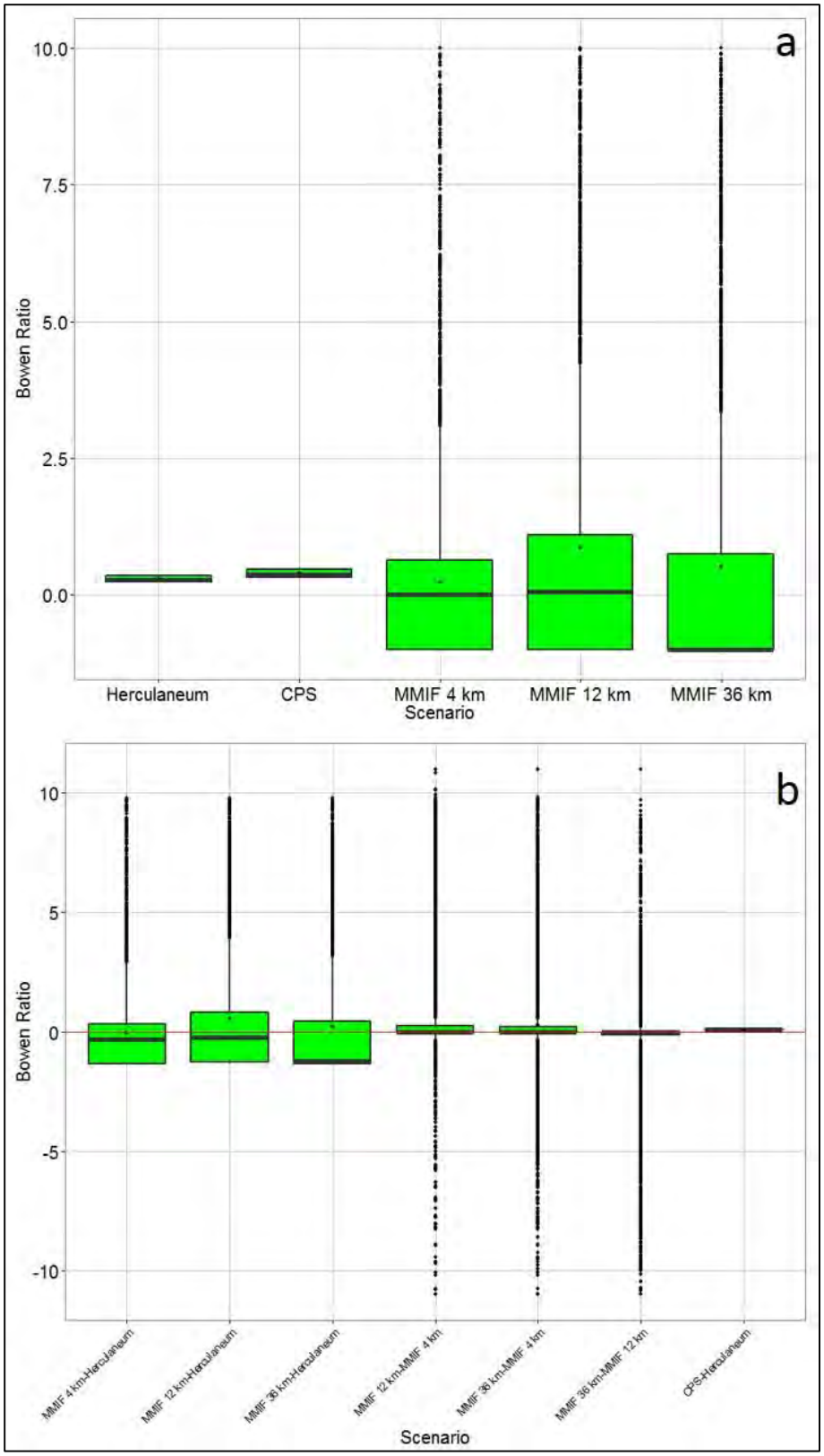

Figure A-34. Herculaneum Bowen ratio: a) annual distributions and b) bias distributions.

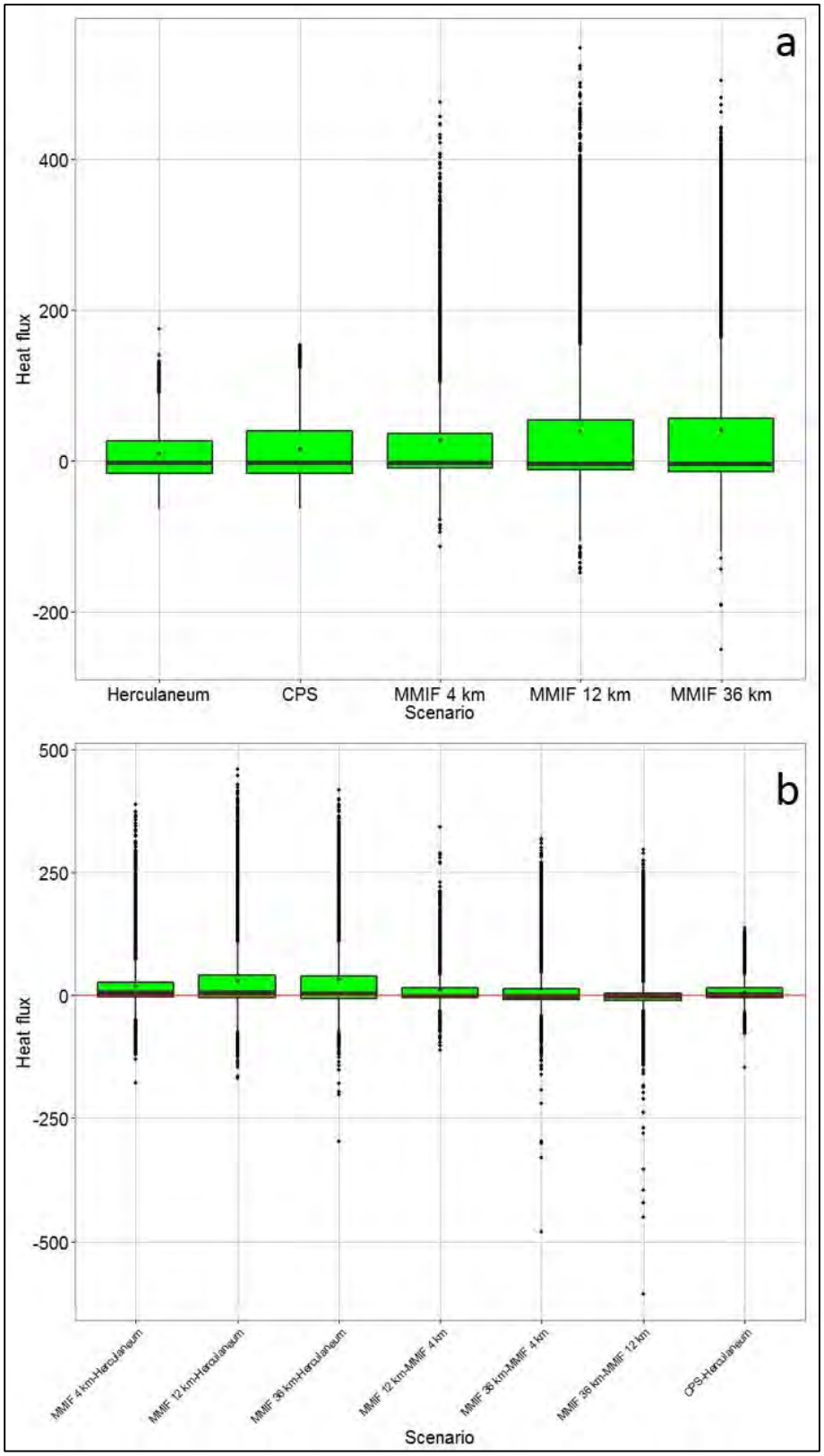

Figure A-35. Herculaneum heat flux (W/m^2): a) annual distributions and b) bias distributions.

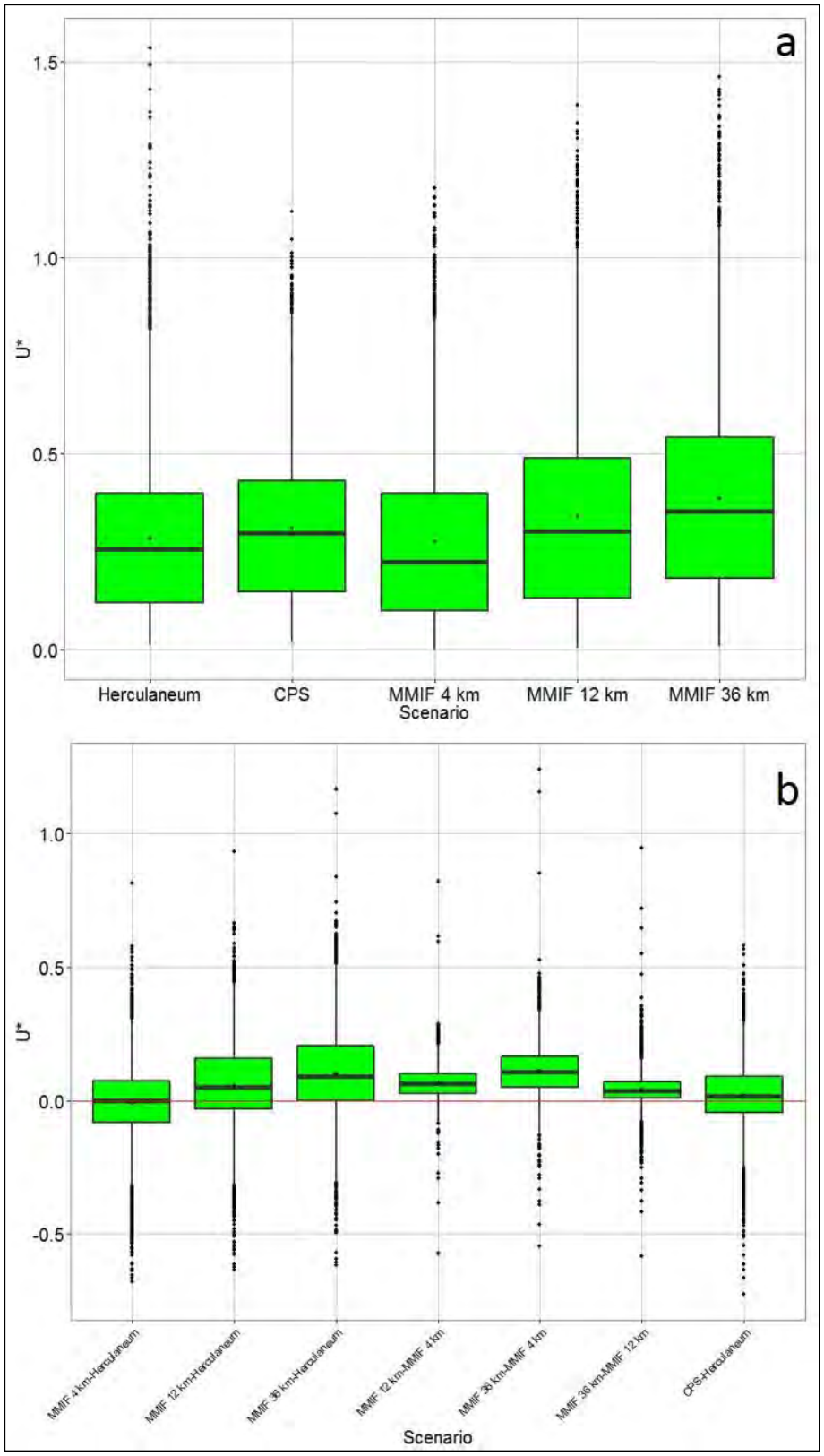

Figure A-36. Herculaneum surface friction velocity, u* (m/s): a) annual distributions and b) bias distributions.

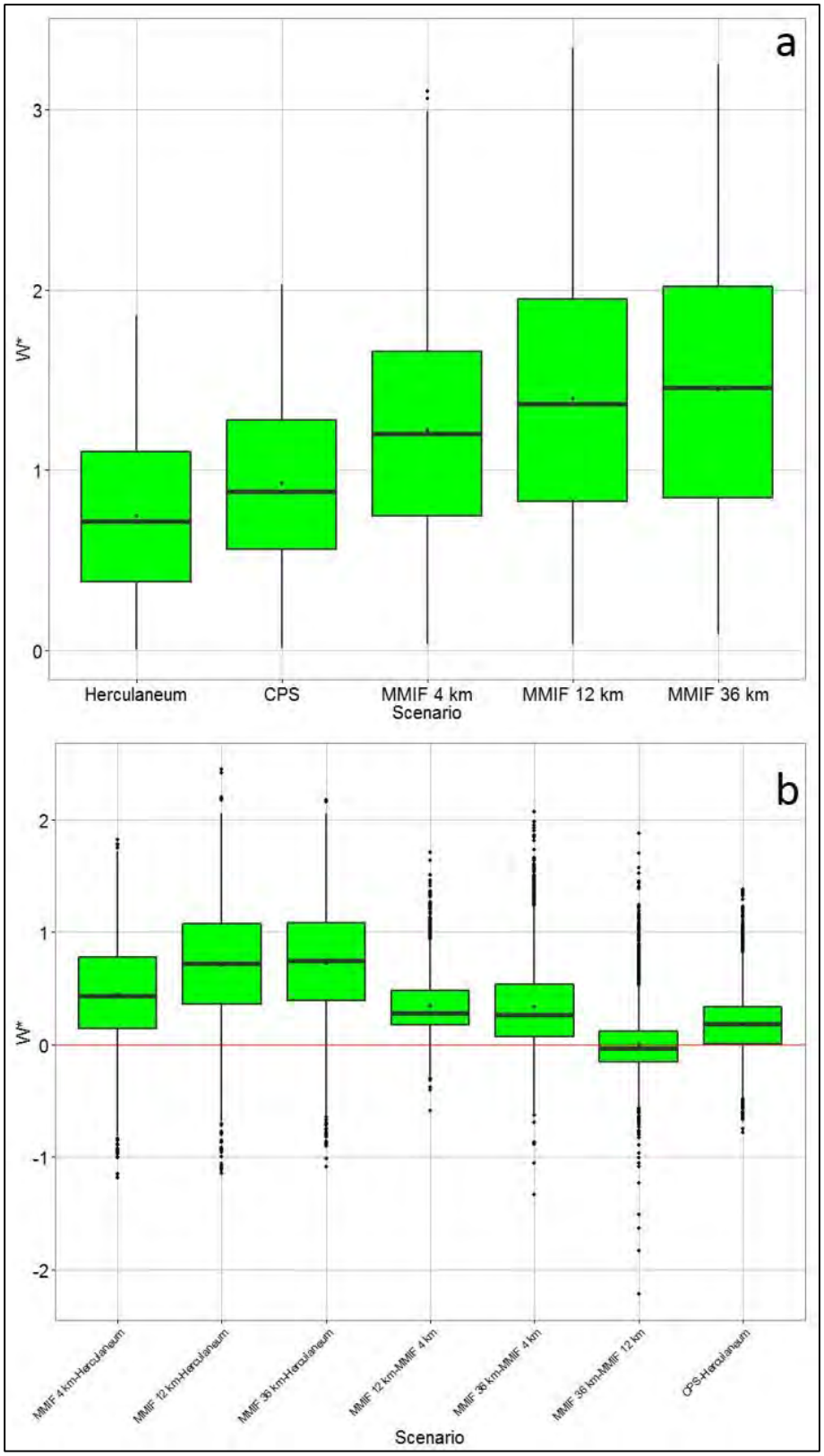

Figure A-37. Herculaneum convective velocity scale, w* (m/s): a) annual distributions and b) bias distributions.

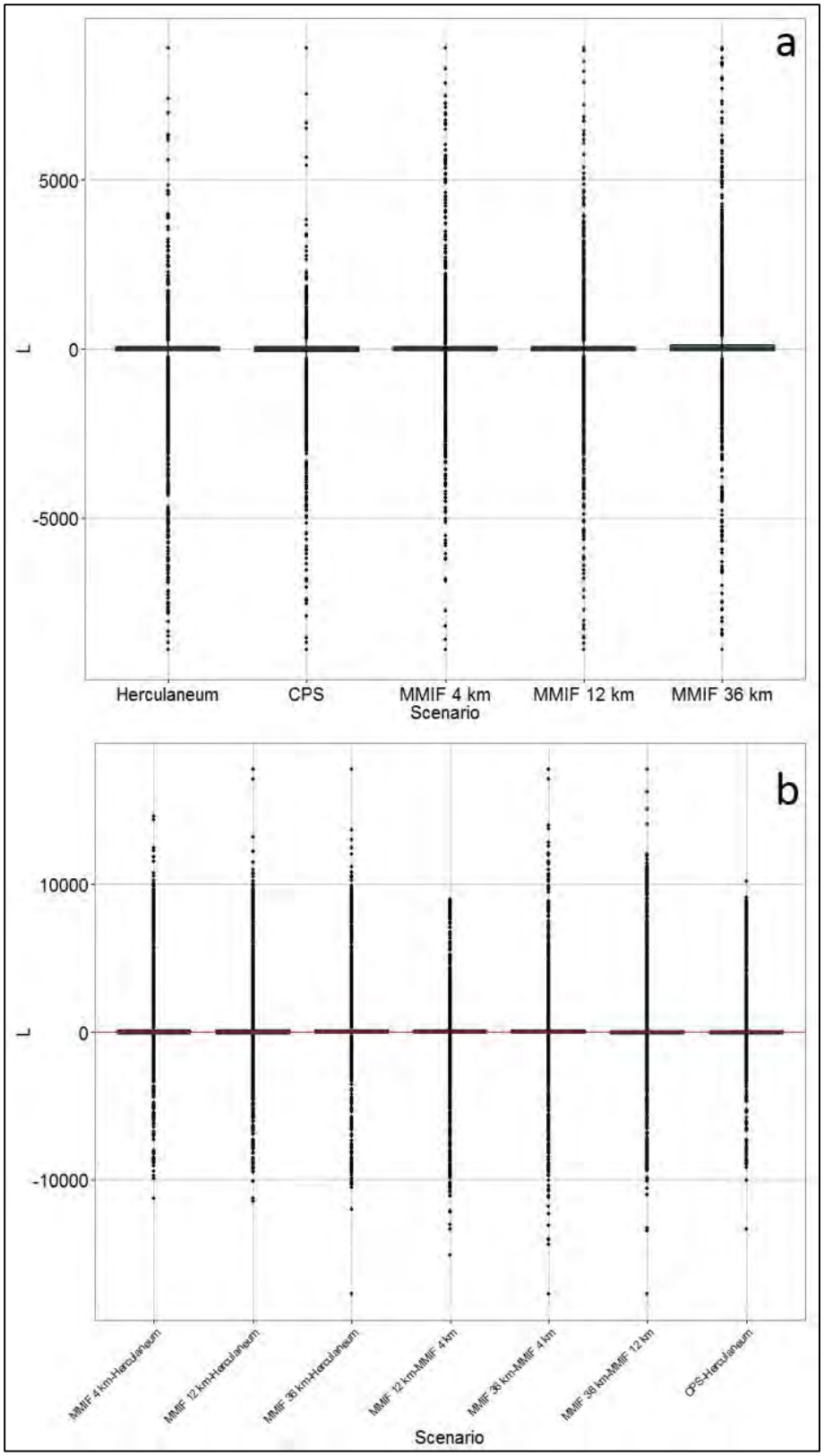

Figure A-38. Herculaneum Monin-Obukhov length (m): a) annual distributions and b) bias distributions.

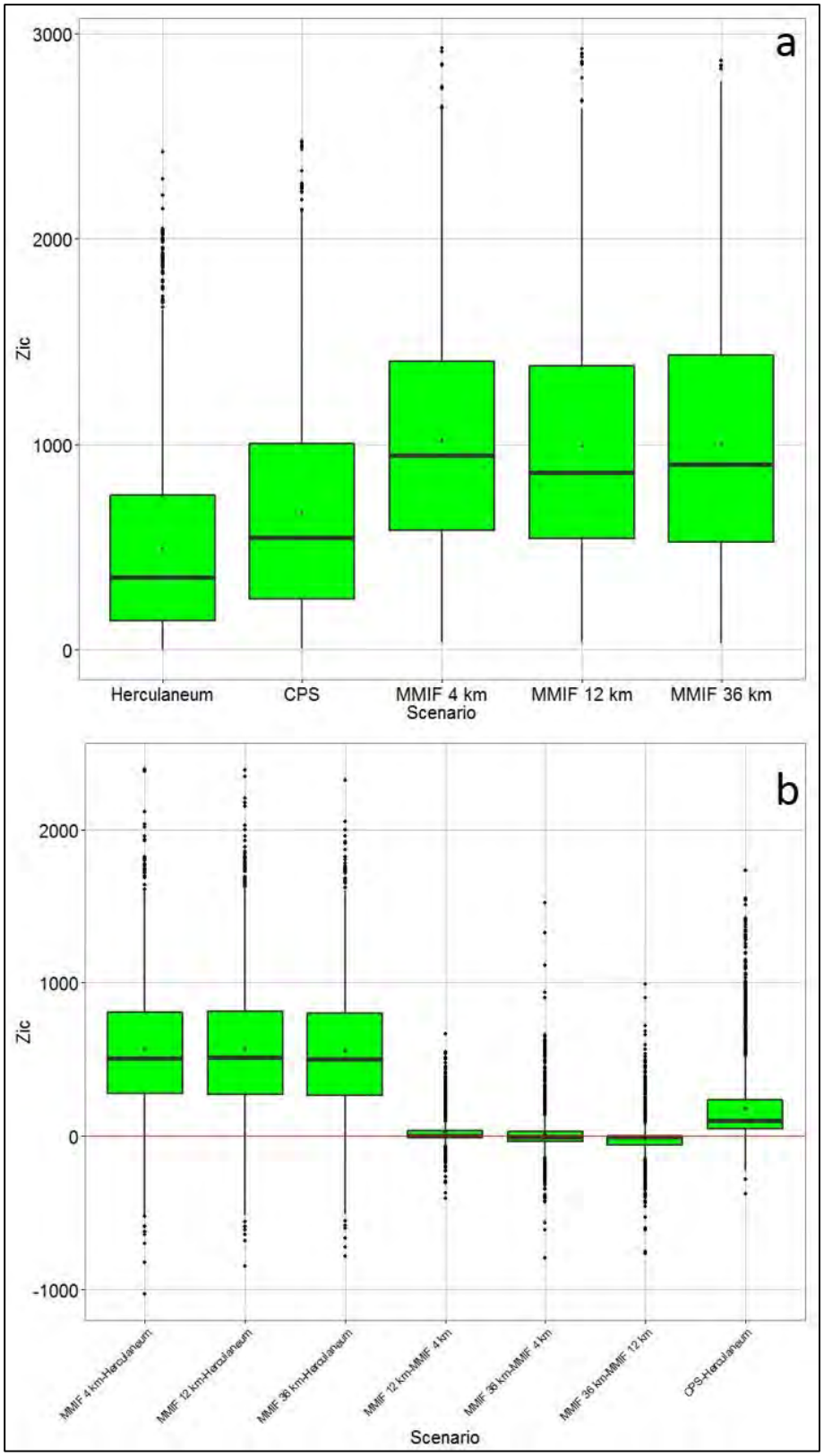

Figure A-39. Herculaneum convective mixing height (m): a) annual distributions and b) bias distributions.

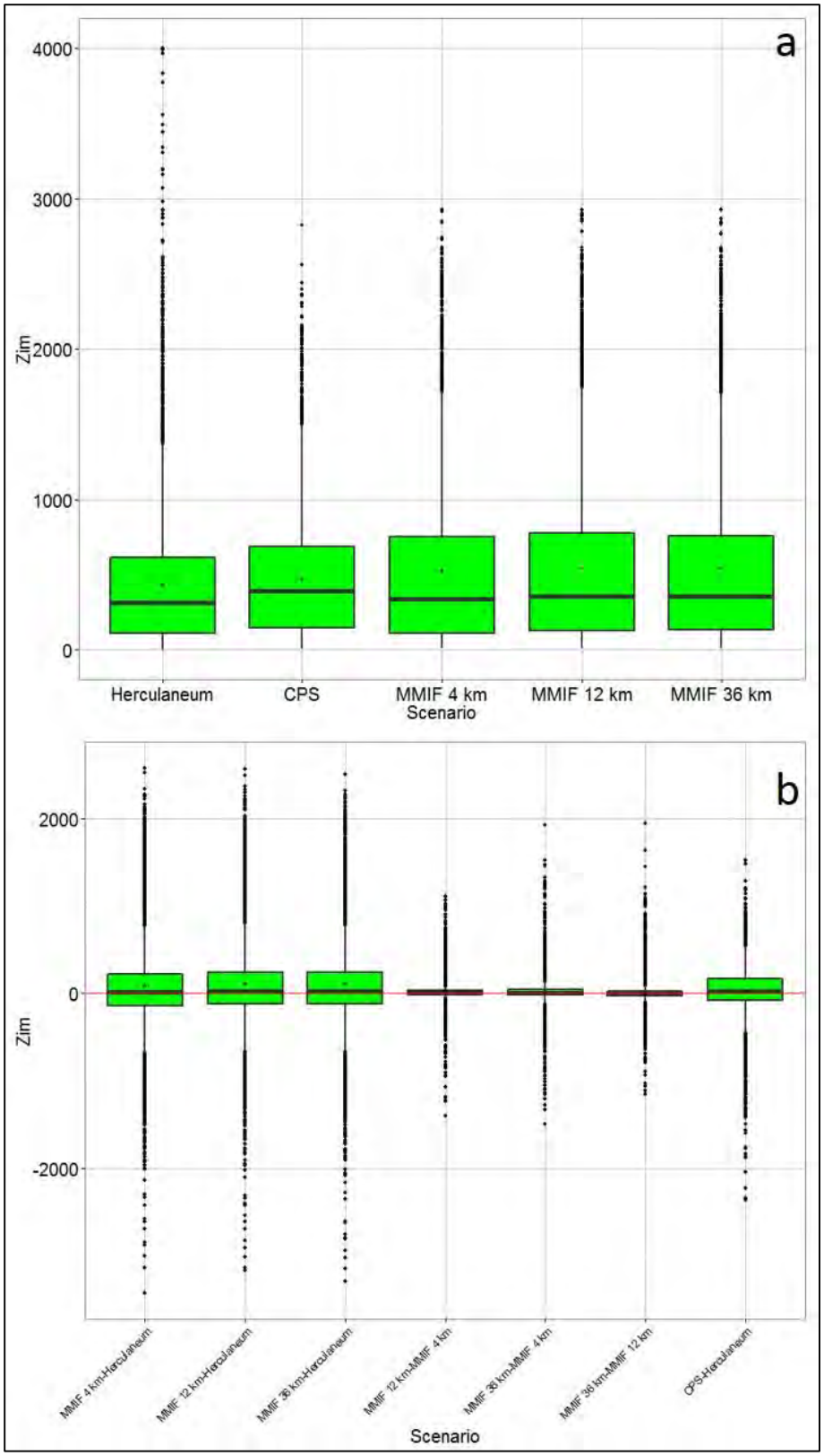

Figure A-40. Herculaneum mechanical mixing height (m): a) annual distributions and b) bias distributions.

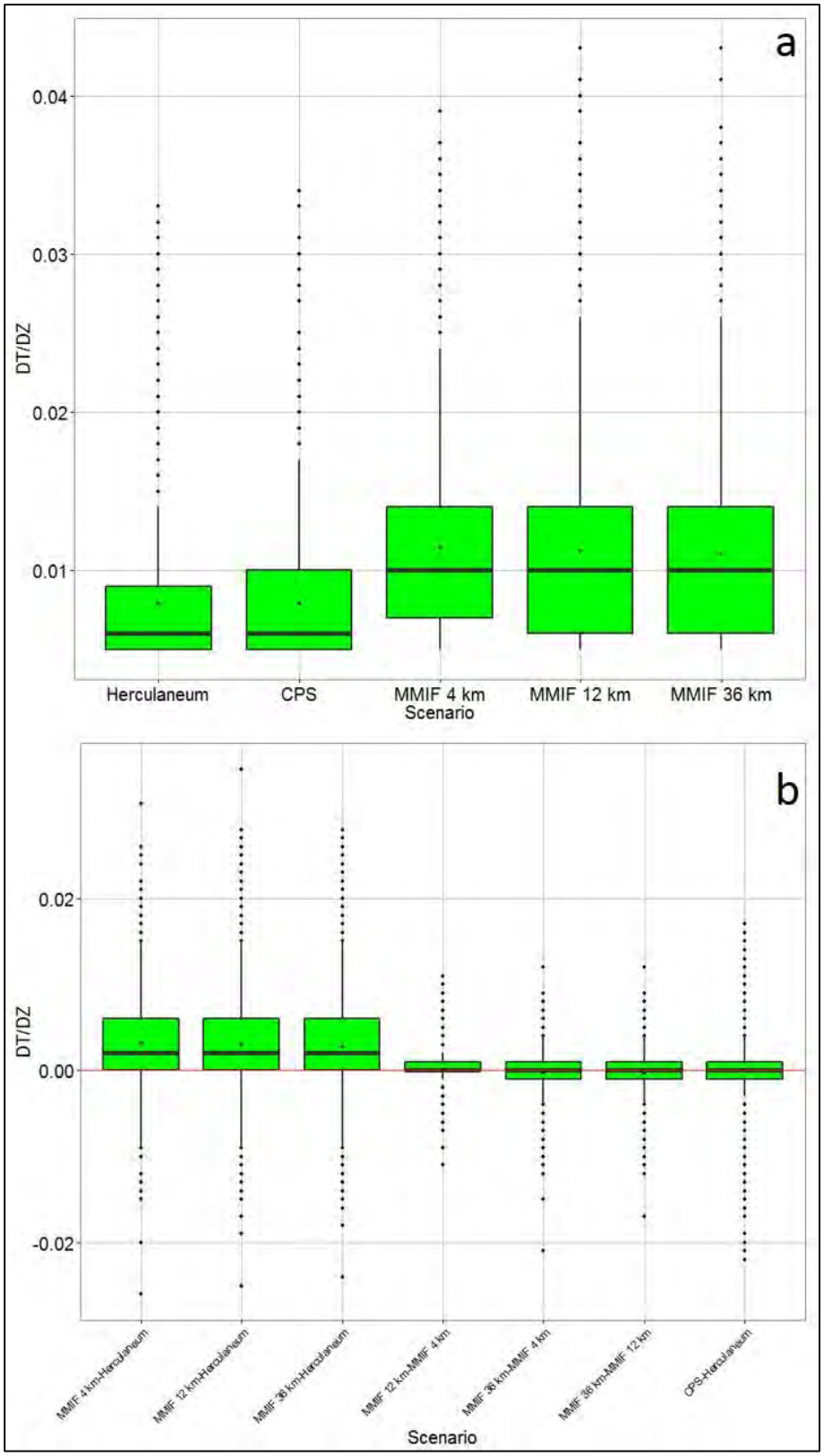

Figure A-41. Herculaneum potential temperature gradient (K/m) above Zic: a) annual distributions and b) bias distributions.

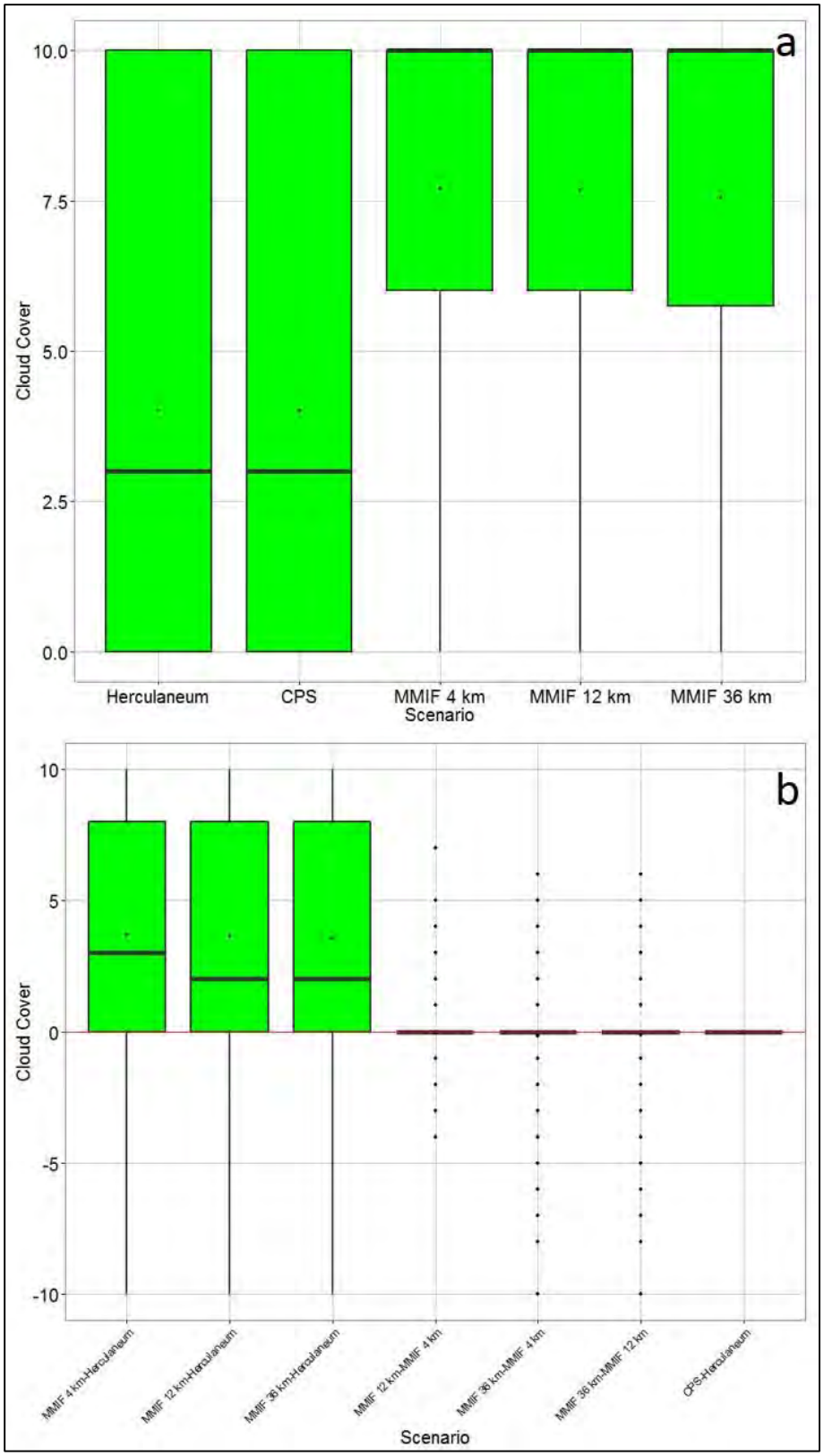

Figure A-42. Herculaneum cloud cover (tenths*10): a) annual distributions and b) bias distributions.

Appendix B. NWS/AERMET and WRF/MMIF analyses for Region 8 sites.

B.1 Introduction

The AERMOD Meteorological Data workgroup selected a number of sites within EPA Region 8 to examine and better understand the issues of utilizing various types of meteorological data and post-processors for air quality dispersion modeling. The workgroup focused on meteorological data reported through the National Weather Service (NWS) and processed through AERMET, as well as meteorological data generated by a the Weather Research and Forecasting (WRF) prognostic model and processed through the Mesoscale Model Interface Program (MMIF). This section summarizes the results of an analysis that evaluated the differences among NWS/AERMET and WRF/MMIF output for various surface meteorological parameters at sites within EPA Region 8. This work does not comment on the accuracy of the results because of the lack of observational data (i.e., monitored air quality concentrations) available for model validation at the selected sites.

The basic purpose of AERMET is to use meteorological measurements to compute certain boundary layer parameters used to estimate profiles of wind, turbulence and temperature. The depth of this layer and the dispersion of pollutants within it are influenced on a local scale by surface characteristics, such as surface roughness, reflectivity (albedo), and the availability of surface moisture. Surface characteristics in the form of albedo, surface roughness and Bowen ratio, plus standard meteorological observations (wind speed, wind direction, temperature, and cloud cover), are input to AERMET. AERMET then calculates the boundary layer parameters, including the Monin-Obukhov Length (L), surface friction velocity (u*), surface roughness length (z_o), surface heat flux (H), and the convective scaling velocity (w*). AERMET also provides estimates of the convective and mechanical mixed layer heights, z_{ic} and z_{im}, respectively. These parameters are then passed to AERMOD, where similar expressions (in conjunction with measurements) are used to calculate vertical profiles of wind speed, lateral and vertical turbulent fluctuations, potential temperature gradient, and potential temperature. Although AERMOD is capable of estimating meteorological profiles with data from as little as one measurement height, it will use as much data as the user can provide for defining the vertical structure of the boundary layer. In addition to the boundary layer parameters, AERMET passes all measurements of wind, temperature, and turbulence in a form AERMOD needs.

B.2 Methodology

Based on the data available during the time of this study, this work evaluated a total of five model cases covering year 2010 and 2011 and eight meteorological parameters at six sites with EPA Region 8. This work selected sites based on proximity to flat terrain, valleys, and mountains to understand the impacts various types of terrain may have on the meteorological data. Table B-1 and Figure B-1 presents information about the sites selected for this work.

Table B-1. Meteorological sites analyzed for study.

Site	Latitude	Longitude	NLCD Codes	Surface Site Codes	Upper Air Site Codes (2010/2011)
Lamar, CO	38.07 N	102.69 W	090900	KLAA (03013)	KDDC (31484/17452)
Miles City, MT	46.43 N	105.89 W	050600	KMLS (24037)	KGGW (31986/24880)
Minot, ND	48.26 N	101.28 W	083000	KMOT (24013)	KBIS (18650/15494)
Rapid City, SD	44.05 N	103.05 W	101100	KRAP (24090)	KRAP (19280/17035)
Vernal, UT	40.44 N	109.51 W	081500	KVEL (94030)	KGJT (1005/29527)
Riverton, WY	43.06 N	108.46 W	081700	KRIW (24061)	KRIW (1410/29950)

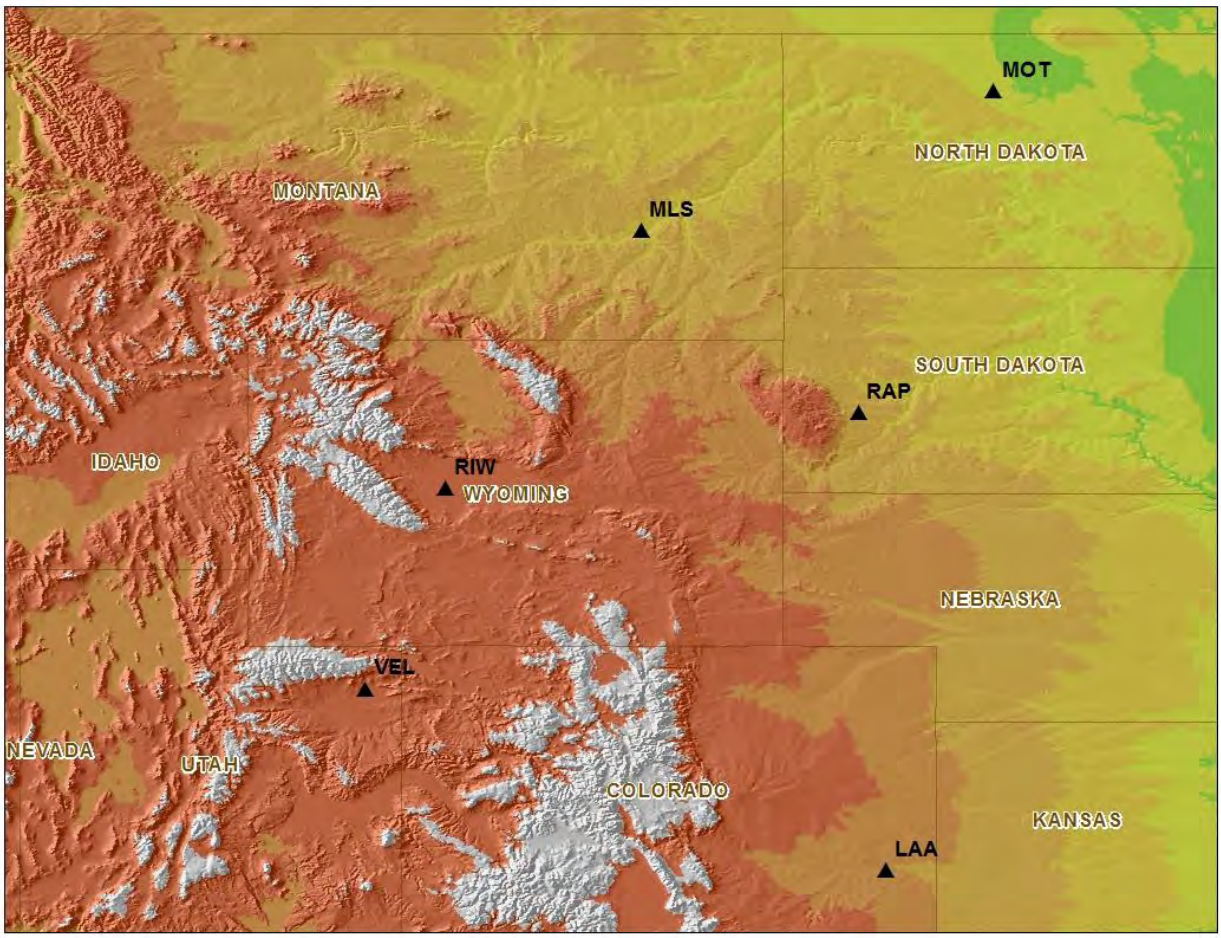

Figure B-1. Map of sites analyzed for study.

The work also focused on meteorological information generated from two meteorological platforms. The first platform included EPA's preferred meteorological model programs, comprising of AERSURFACE, AERMINUTE, and AERMET. This platform, referenced as NWS/AERMET, was used to model calendar 2010 and 2011 at the selected sites. The second platform included an alternative meteorological model and post-processor, comprising of the

WRF model and the MMIF post-processor. This platform, referenced as WRF/MMIF, was used to model calendar 2010 and 2011 using 4 kilometer (km) and 12 km grid resolutions. The details of the model descriptions and model assumptions for each model case are outlined in Table B-2 for NWS/AERMET and Table B-3 for the WRF/MMIF cases.

Table B-2. Description of NWS/AERMET model cases.

Case Name	Year	Model Platforms	Model Assumptions	Input Data
2010 NWS/AERMET	2010	AERSURFACE v13016	Datum: NAD83 Study radius for surface roughness: 1 Vary By Sector: Yes Number of Sectors: 12 Temporal Resolution: Monthly Snow Cover: Yes Months to Seasons: No Airport: Yes Arid: N Surface Moisture: Average	USGS NLCD92 Data
		AERMINUTE v14237	Default	ASOS 1-minute Data [6405 Datasets]
		AERMET v14134	Default	NWS Hourly Surface Data [ISHD Format]
				NWS Hourly Upper Air Data [FSL Format]
2011 NWS/AERMET	2011	AERSURFACE v13016	Datum: NAD83 Study radius for surface roughness: 1 Vary By Sector: Yes Number of Sectors: 12 Temporal Resolution: Monthly Snow Cover: Yes Months to Seasons: No Airport: Yes Arid: N Surface Moisture: Average	USGS NLCD92 Data
		AERMINUTE v14237	Default	ASOS 1-minute Data [6405 Datasets]
		AERMET v14134	Default	NWS Hourly Surface Data [ISHD Format]
				NWS Hourly Upper Air Data [FSL Format]

Table B-3. Description of WRF/MMIF model cases.

Case Name	Year	Model Platforms	Model Assumptions	Input Data
2010 4-km WRF/MMIF	2010	WRF v3.5.1	Resolution: 4km Configuration: Details included in WestJumpAQMS WRF Evaluation Report[4]	Details included in WestJumpAQMS WRF Evaluation Report[5]
		MMIF v3.0		
2010 12-km WRF/MMIF	2010	WRF v3.5.1	Resolution: 12km Configuration: Details included in WestJumpAQMS WRF Evaluation Report[6]	Details included in WestJumpAQMS WRF Evaluation Report[7]
		MMIF v3.0		
2011 12-km WRF/MMIF	2011	WRF v3.3.1	Resolution: 12km Configuration: Details included in EPA's Evaluation Report[8]	Details included in WestJumpAQMS WRF Evaluation Report[9]
		MMIF v3.0		

Using these two platforms, this study analyzed eight meteorological surface parameters, including temperature, relative humidity, wind speed, wind displacement, mechanical mixing height, convective mixing height, surface friction velocity, and convective velocity. A number of statistical metrics and graphical displays were also generated and reviewed to evaluate the differences among the platforms at the selected sites for the various meteorological parameters and model years. The statistical metrics included hourly, monthly or annual averages, bias, fractional bias, error, fractional error, and coefficient of determination. Time series, scatter and box plots were also generated as part of the graphical displays. This work only includes a summary of the statistical results. However, all of the graphical displays and statistical analyses can be provided upon request.

---

[4] Western Regional Air Partnership (WRAP) West-wide Jump Start Air Quality Modeling Study (WestJumpAQMS) WRF Application/Evaluation, UNC, ENVIRON, Alpine Geophysics, February 29, 2012, http://www.wrapair2.org/pdf/WestJumpAQMS_2008_Annual_WRF_Final_Report_February29_2012.pdf

[5] Ibid.

[6] Ibid.

[7] Ibid.

[8] XXX

[9] XXX

Given that the NWS/AERMET platform may be more representative of the meteorological conditions at the selected sites, the statistical analysis were based on comparing the model cases that used the WRF/MMIF platform relative to the model cases that used the NWS/AERMET platform. For instance, the statistics were based on comparing the 2010 NWS/AERMET case to the 2010 4-km WRF/MMIF case or the 2010 NWS/AERMET case to the 2010 12-km WRF/MMIF case. Some of the analyses or graphical displays also compared the 2010 4-km WRF/MMIF case to the 2010 12-km WRF/MMIF case. Although the WRF model utilizes observational datasets, the observations used for this study were prescribed to the associated grid box with a resolution of 4 km or 12 km, which may not be representative of the actual conditions that cover the entire grid box or at the selected sites. Therefore, the meteorological information generated by the NWS/AERMET platform may be more representative at the selected sites than the WRF/MMIF platform.

B.3 Results/Summary

A summary of the results by meteorological parameter and the significant differences among the model cases are outlined below. In general, careful evaluation and consideration is recommended when selecting a meteorological model platform because of the range and variability observed in the results of this analysis. Although additional evaluation is needed for model validation, the results from the WRF/MMIF platform for temperature, relative humidity, and wind speed were found to be similar to the NWS/AERMET platform. However, mechanical and convective mixing heights and surface friction and convective velocities varied significantly among the platforms, with notable differences in magnitudes, temporal variability, and low correlations. Generally, the WRF/MMIF platform predicted significantly higher mechanical and convective mixing heights and convective velocities and lower surface friction velocities relative to the NWS/AERMET platform. The results of the wind displacement analysis also suggest that the WRF/MMIF 4 km model case could potentially displace the plume significantly, relative to the NWS/AERMET platform. These results raise concerns on the adequacy of the WRF/MMIF platform for generating meteorological information for the dispersion models and whether the information would generate conservative results.

This work also found that the 4 km and 12 km WRF/MMIF model cases compared well to one another, with correlations greater than 0.70 for all meteorological, except wind speed at the Lamar, CO, Vernal, UT, and Riverton, WY sites and surface friction velocity at the Vernal, UT and Riverton, WY sites. These results suggests systematic difference among the WRF/MMIF and NWS/AERMET platforms. Without additional evaluation, this work did not find that the WRF/MMIF 4 km case out-performed the 12 km case relative to the NWS/AERMET platform.

B.3.1 Temperature

Table B-4 shows statistics for monthly average temperatures. The findings for temperature are:

- The WRF/MMIF results for both model years and all sites are similar to the NWS/AERMET results. In general, the WRF/MMIF platforms slightly over-estimated the values relative to the NWS/AERMET platform at all sites, except at the Minot, ND and Riverton, WY sites. Although minor, this work found that the most significant

differences among the platforms were observed at the Lamar, CO, Vernal, UT and Riverton, WY sites during January and February.

- Relative to NWS/AERMET, the WRF/MMIF results show similar temporal variability for both model years and at all sites.
- The WRF/MMIF 4 km and 12 km are comparable to one another, with correlations greater than 0.98. This work did not find that the WRF/MMIF 4 km case out-performed the 12 km case relative to the NWS/AERMET platform.

Table B-4. Monthly averaged temperature (K) across all modeled cases at each site.

Site	Temperature (K)	Mean	Bias	Fractional Bias	Error	Fractional Error	R^2
LAA	Minimum	267.45	-0.61	-0.20	1.67	0.56	0.76
	Maximum	301.41	3.40	1.27	4.40	1.63	0.94
	Average	285.93	1.43	0.52	2.48	0.88	0.88
MLS	Minimum	263.54	-0.73	-0.26	1.16	0.41	0.72
	Maximum	297.35	2.24	0.78	3.88	1.49	0.95
	Average	280.54	0.65	0.23	1.93	0.70	0.88
MOT	Minimum	258.97	-2.71	-1.05	1.01	0.35	0.51
	Maximum	295.48	0.99	0.35	3.33	1.29	0.95
	Average	278.48	-0.73	-0.27	1.82	0.66	0.86
RAP	Minimum	266.64	-0.35	-0.12	1.24	0.44	0.24
	Maximum	297.69	2.39	0.81	3.22	1.22	0.97
	Average	281.40	0.88	0.31	1.96	0.70	0.82
VEL	Minimum	261.15	-1.17	-0.41	1.33	0.45	0.24
	Maximum	297.30	5.70	2.17	5.87	2.22	0.94
	Average	281.17	0.98	0.37	2.30	0.83	0.80
RIW	Minimum	262.49	-3.65	-1.39	1.27	0.43	0.61
	Maximum	296.40	1.29	0.45	4.07	1.54	0.93
	Average	279.94	-0.07	-0.03	1.97	0.71	0.83

B.3.2 Relative humidity

Table B-5 shows statistics for monthly average relative humidity. Findings are:

- The WRF/MMIF results for both model years and all sites are similar to the NWS/AERMET results. In general, the WRF/MMIF platforms slightly over-estimated the values relative to the NWS/AERMET platform at the Miles City, MT, Minot, ND, and Riverton, WY sites, and slightly under-estimated the values at the Lamar, CO, Rapid City, SD, and Vernal, UT sites. Although minor, this work found that the most significant differences between the platforms were observed at the Miles City, Montana site during the summer and the Riverton, Wyoming site during January and February.

- Relative to NWS/AERMET, the WRF/MMIF results show similar temporal variability for both model years and at all sites.
- The WRF/MMIF 4 km and 12 km are comparable to one another, with correlations greater than 0.86. This work found that the WRF/MMIF 4 km case compared better to NWS/AERMET platform than the WRF/MMIF 12 km case at all sites.

Table B-5. Monthly averaged relative humidity across all modeled cases at each site.

Site	Temperature (K)	Mean	Bias	Fractional Bias	Error	Fractional Error	R^2
LAA	Minimum	39.41	-12.13	-23.03	6.02	11.96	0.51
	Maximum	81.49	7.37	19.72	13.23	32.68	0.88
	Average	53.25	-1.43	-0.36	9.50	20.61	0.74
MLS	Minimum	39.92	-16.08	-27.27	6.90	8.94	0.10
	Maximum	92.10	14.35	17.19	16.48	28.04	0.84
	Average	65.58	0.88	-0.54	10.61	17.60	0.60
MOT	Minimum	55.54	-5.88	-10.53	6.28	9.89	0.26
	Maximum	92.44	15.53	20.93	17.84	25.37	0.84
	Average	71.98	5.72	7.28	10.02	14.62	0.62
RAP	Minimum	41.94	-14.78	-26.91	7.15	11.42	0.03
	Maximum	89.09	13.54	17.18	15.40	28.27	0.88
	Average	65.51	-0.44	-1.22	10.35	17.40	0.55
VEL	Minimum	26.57	-16.09	-29.53	7.35	10.60	0.03
	Maximum	85.06	8.25	22.66	17.05	32.55	0.79
	Average	54.70	-1.54	-1.41	10.20	20.93	0.53
RIW	Minimum	33.20	-3.85	-8.23	6.42	15.88	0.08
	Maximum	93.77	20.10	25.03	20.35	27.53	0.80
	Average	55.96	5.40	9.03	11.29	21.31	0.56

B.3.3 Wind speed

Table B-6 shows statistics for monthly average wind speed. Findings are:

- The WRF/MMIF results for both model years and all sites are similar to the NWS/AERMET results. In general, the WRF/MMIF platforms slightly under-estimated the values relative to the NWS/AERMET platform at all sites, except at the Vernal, UT site. Although minor, this work found that the most significant differences between the platforms were observed at the Vernal, UT site.
- Relative to NWS/AERMET, the WRF/MMIF results show similar temporal variability for both model years and at all sites.
- The WRF/MMIF 4 km and 12 km are comparable to one another at the Miles City, MT, Minot, ND, and Rapid City, SD sites, with correlations greater than 0.70. Generally, the

WRF/MMIF 12 km case compared better to NWS/AERMET platform than the WRF/MMIF 4 km case at all sites except at the Vernal, UT site.

Table B-6. Monthly averaged wind speed (m/s) across all modeled cases at each site.

Site	Temperature (K)	Mean	Bias	Fractional Bias	Error	Fractional Error	R^2
LAA	Minimum	2.75	-1.44	-32.47	0.94	33.80	0.24
	Maximum	6.07	-0.12	3.83	2.06	49.37	0.75
	Average	4.09	-0.75	-15.23	1.44	39.25	0.55
MLS	Minimum	2.61	-1.29	-30.36	0.98	23.99	0.30
	Maximum	5.88	-0.17	-5.49	1.69	48.03	0.83
	Average	4.03	-0.66	-16.70	1.28	36.80	0.55
MOT	Minimum	3.52	-1.36	-26.36	0.87	21.93	0.55
	Maximum	6.17	-0.14	0.77	1.62	39.41	0.90
	Average	4.58	-0.45	-6.98	1.07	27.89	0.74
RAP	Minimum	3.21	-1.04	-22.30	1.21	33.42	0.02
	Maximum	5.77	0.29	14.60	1.92	50.93	0.90
	Average	4.39	-0.47	-3.24	1.49	39.58	0.34
VEL	Minimum	1.35	-0.57	-18.96	0.80	41.04	0.02
	Maximum	4.04	1.38	55.55	1.72	69.55	0.71
	Average	2.75	0.46	15.58	1.30	51.28	0.26
RIW	Minimum	1.59	-1.59	-40.45	0.96	39.88	0.04
	Maximum	5.39	0.30	11.52	2.20	57.21	0.59
	Average	3.25	-0.99	-27.16	1.59	49.53	0.35

B.3.4 Wind displacement

Wind displacement, as defined in Section 2.2, is shown in Table B-7 for each site. Findings are:

- The calculation of wind displacement is important for determining how far a plume may be displaced relative to its actual location. Given that this work is assuming that the NWS/AERMET platform is more representative that the WRF/MMIF platform, this means that the wind displacement results will determine how far the WRF/MMIF platform will potentially displace the plume relative to the NWS/AERMET platform. The wind displacement results can also determine whether the resolution of the prognostic model could be sufficient for capturing the plume. For instance, when the wind displacement is within 4 km, these results could suggest that the WRF/MMIF 4 km model case could potentially capture the plume within that selected grid box. However, careful consideration of the adequacy of the WRF/MMIF 4 km model platform should be considered when the wind displacement is greater than 4 km.

- The results of the wind displacement analysis suggests that the WRF/MMIF 12 km model case could potentially capture the plume given the resolution and distance of less than 12 km. However, careful consideration is needed when determining the adequacy of the WRF/MMIF 4 km model case because the distances are more than 4 km. Further, the most notable differences were found at the Vernal, UT site.

Table B-7. Monthly averaged wind displacement (km) across all modeled cases at each site.

Site	Minimum	Maximum	Average
LAA	6.32	13.13	9.23
MLS	5.58	8.64	7.08
MOT	6.32	11.03	9.04
RAP	7.13	11.33	8.11
VEL	5.47	14.42	10.65
RIW	6.60	10.84	9.33

B.3.5 Mechanical mixing heights

Table B-8 shows statistics for monthly average mechanical mixing heights. The findings for mechanical heights are:

- The WRF/MMIF results for both model years and all sites are not similar to the NWS/AERMET results, including notable differences in magnitudes, temporal variability, and low correlations. In general, the WRF/MMIF platforms significantly over-estimated the values relative to the NWS/AERMET platform at all sites, except at the Minot, ND site.
- The WRF/MMIF 4 km and 12 km are comparable to one another, with correlations greater than 0.80. This work did not find that the WRF/MMIF 4 km case out-performed the 12 km case relative to the NWS/AERMET platform.

Table B-8. Monthly averaged mechanical mixing heights (m) across all modeled cases at each site.

Site	Temperature (K)	Mean	Bias	Fractional Bias	Error	Fractional Error	R^2
LAA	Minimum	166.76	-234.50	-46.96	130.18	69.77	0.03
	Maximum	1125.91	360.06	60.41	844.73	92.30	0.47
	Average	637.47	128.03	5.69	482.44	78.08	0.22
MLS	Minimum	114.02	-368.02	-90.02	102.72	58.59	0.06
	Maximum	974.30	443.90	19.53	706.95	99.80	0.61
	Average	520.02	24.49	-24.62	384.27	77.71	0.27
MOT	Minimum	112.38	-1168.20	-110.96	358.86	72.09	0.06
	Maximum	1469.21	-120.02	-24.69	1200.46	117.07	0.37
	Average	679.38	-506.86	-66.89	696.03	95.24	0.20
RAP	Minimum	201.56	-261.57	-35.49	150.71	63.89	0.03
	Maximum	858.03	442.28	31.27	646.23	93.02	0.61
	Average	490.40	23.49	-2.58	354.88	74.73	0.25
VEL	Minimum	45.52	12.35	-19.08	34.11	68.10	0.16
	Maximum	988.12	661.08	92.88	761.05	102.31	0.47
	Average	394.62	275.89	24.32	372.65	83.12	0.31
RIW	Minimum	31.61	-299.73	-113.72	97.79	79.51	0.05
	Maximum	1036.19	379.77	-4.74	836.78	118.22	0.44
	Average	539.10	7.36	-45.85	489.74	98.40	0.18

B.3.6 Convective mixing heights

Table B-9 shows statistics for monthly average convective mixing heights. The findings for convective heights are:

- The WRF/MMIF results for both model years and all sites are not similar to the NWS/AERMET results, including notable differences in magnitudes, temporal variability, low correlations. In general, the WRF/MMIF platforms significantly over-estimated the values relative to the NWS/AERMET platform at all sites.
- The WRF/MMIF 4 km and 12 km are comparable to one another, with correlations greater than 0.75. Generally, the WRF/MMIF 12 km case compared slightly better to NWS/AERMET than the WRF/MMIF 4 km case at all sites except at the Vernal, UT site.

Table B-9. Monthly averaged mechanical mixing heights (m) across all modeled cases at each site.

Site	Temperature (K)	Mean	Bias	Fractional Bias	Error	Fractional Error	R^2
LAA	Minimum	121.70	108.28	15.73	228.66	39.38	0.06
	Maximum	1932.54	993.78	114.41	1035.89	114.41	0.71
	Average	947.26	510.25	60.82	605.79	71.23	0.39
MLS	Minimum	68.00	-117.41	-11.38	74.57	34.06	0.00
	Maximum	1576.02	619.99	147.31	672.85	151.82	0.71
	Average	767.75	249.92	43.43	363.57	62.45	0.43
MOT	Minimum	94.26	-22.78	-30.46	62.00	30.96	0.03
	Maximum	1227.90	420.21	118.19	491.13	123.38	0.74
	Average	640.71	149.98	24.62	271.34	52.45	0.46
RAP	Minimum	132.75	-144.11	-7.36	106.70	31.95	0.00
	Maximum	1548.62	479.86	95.05	628.09	95.68	0.76
	Average	781.64	139.84	24.46	341.08	50.63	0.38
VEL	Minimum	96.43	-200.12	-18.94	81.84	30.55	0.00
	Maximum	1981.48	493.66	98.98	711.62	103.24	0.65
	Average	996.73	86.06	21.06	425.01	53.47	0.35
RIW	Minimum	68.18	-294.78	-49.76	50.88	29.24	0.02
	Maximum	1919.11	311.17	75.16	698.34	86.45	0.76
	Average	971.69	31.04	-1.52	401.16	51.31	0.41

B.3.7 Surface friction velocity

Table B-10 shows statistics for monthly average surface friction velocity (u*). The findings for u* are:

- The WRF/MMIF results for both model years and all sites not similar to the NWS/AERMET results, including notable differences in magnitudes, temporal variability, and low correlations. In general, the WRF/MMIF platforms significantly under-estimated the values relative to the NWS/AERMET platform at all sites, except at the Rapid City, SD and Vernal, UT sites.
- The WRF/MMIF 4 km and 12 km are comparable to one another, with correlations greater than 0.75, at all site except at the Vernal, UT and Riverton, WY sites. This work did not find that the WRF/MMIF 4 km case out-performed the 12 km case relative to the NWS/AERMET platform.

Table B-10. Monthly averaged surface friction velocity (m/s) across all modeled cases at each site.

Site	Temperature (K)	Mean	Bias	Fractional Bias	Error	Fractional Error	R^2
LAA	Minimum	0.15	-0.11	-29.97	0.07	39.07	0.20
	Maximum	0.49	0.12	45.34	0.18	67.36	0.75
	Average	0.33	-0.01	-2.21	0.13	48.56	0.51
MLS	Minimum	0.16	-0.09	-47.52	0.07	35.23	0.34
	Maximum	0.48	0.07	23.88	0.15	69.38	0.78
	Average	0.31	-0.01	-15.17	0.12	49.08	0.55
MOT	Minimum	0.16	-0.31	-73.39	0.14	35.89	0.26
	Maximum	0.67	-0.07	-10.17	0.34	76.68	0.73
	Average	0.40	-0.16	-36.58	0.20	53.49	0.51
RAP	Minimum	0.20	-0.02	-17.23	0.06	40.28	0.01
	Maximum	0.43	0.09	44.04	0.14	64.88	0.84
	Average	0.32	0.03	9.28	0.11	48.19	0.38
VEL	Minimum	0.06	-0.02	-29.02	0.04	47.41	0.12
	Maximum	0.36	0.14	74.01	0.18	89.71	0.80
	Average	0.21	0.05	19.17	0.11	62.22	0.39
RIW	Minimum	0.05	-0.19	-80.06	0.07	48.41	0.01
	Maximum	0.47	-0.02	-5.00	0.20	89.44	0.46
	Average	0.26	-0.10	-41.32	0.15	64.77	0.30

B.3.8 Convective velocity scale

Table B-10 shows statistics for monthly average convective velocity scale (w*). The findings for w* are:

- The WRF/MMIF results for both model years and all sites are not similar to the NWS/AERMET results, including notable differences in magnitudes, temporal variability, and low correlations. In general, the WRF/MMIF platforms significantly over-estimated the values relative to the NWS/AERMET platform at all sites.
- The WRF/MMIF 4 km and 12 km are comparable to one another, with correlations greater than 0.85. This work did not find that the WRF/MMIF 4 km case out-performed the 12 km case relative to the NWS/AERMET platform.

Table B-11. Monthly averaged convective velocity scale (m/s) across all modeled cases at each site.

Site	Temperature (K)	Mean	Bias	Fractional Bias	Error	Fractional Error	R^2
LAA	Minimum	0.32	-0.02	-5.37	0.25	21.82	0.07
	Maximum	2.83	1.91	111.72	1.92	111.72	0.70
	Average	1.50	0.88	57.65	0.95	64.59	0.42
MLS	Minimum	0.18	-0.46	-35.61	0.33	24.65	0.00
	Maximum	2.68	1.58	114.61	1.58	117.97	0.61
	Average	1.27	0.67	53.13	0.78	65.53	0.31
MOT	Minimum	0.19	-0.13	-13.22	0.25	26.07	0.00
	Maximum	2.23	1.25	99.65	1.29	100.92	0.64
	Average	1.09	0.47	40.50	0.58	56.00	0.34
RAP	Minimum	0.29	-0.45	-34.27	0.38	29.55	0.00
	Maximum	2.49	1.22	86.88	1.24	98.56	0.74
	Average	1.29	0.57	43.62	0.73	60.05	0.41
VEL	Minimum	0.25	-0.32	-25.71	0.27	25.65	0.00
	Maximum	2.90	1.11	99.42	1.17	108.27	0.67
	Average	1.46	0.50	35.60	0.66	50.93	0.39
RIW	Minimum	0.25	-0.25	-36.53	0.16	23.05	0.00
	Maximum	2.92	1.16	64.27	1.20	84.81	0.74
	Average	1.41	0.37	19.37	0.58	44.81	0.43

Appendix C: Evaluations for Baldwin, IL and Martins Creek

C.1 Introduction

This appendix discusses evaluations of MMIF and AERMOD sensitivity to grid resolution, number of output vertical levels, and PBL calculation methods. Two cases from the initial evaluation and promulgation of AERMOD (U.S. EPA, 2003; Perry et al., 2005) were modeled: a flat terrain case, Baldwin, IL (Figure C-1), and a further evaluation of Martins Creek (Figure C-2, which represents a case involving complex terrain.

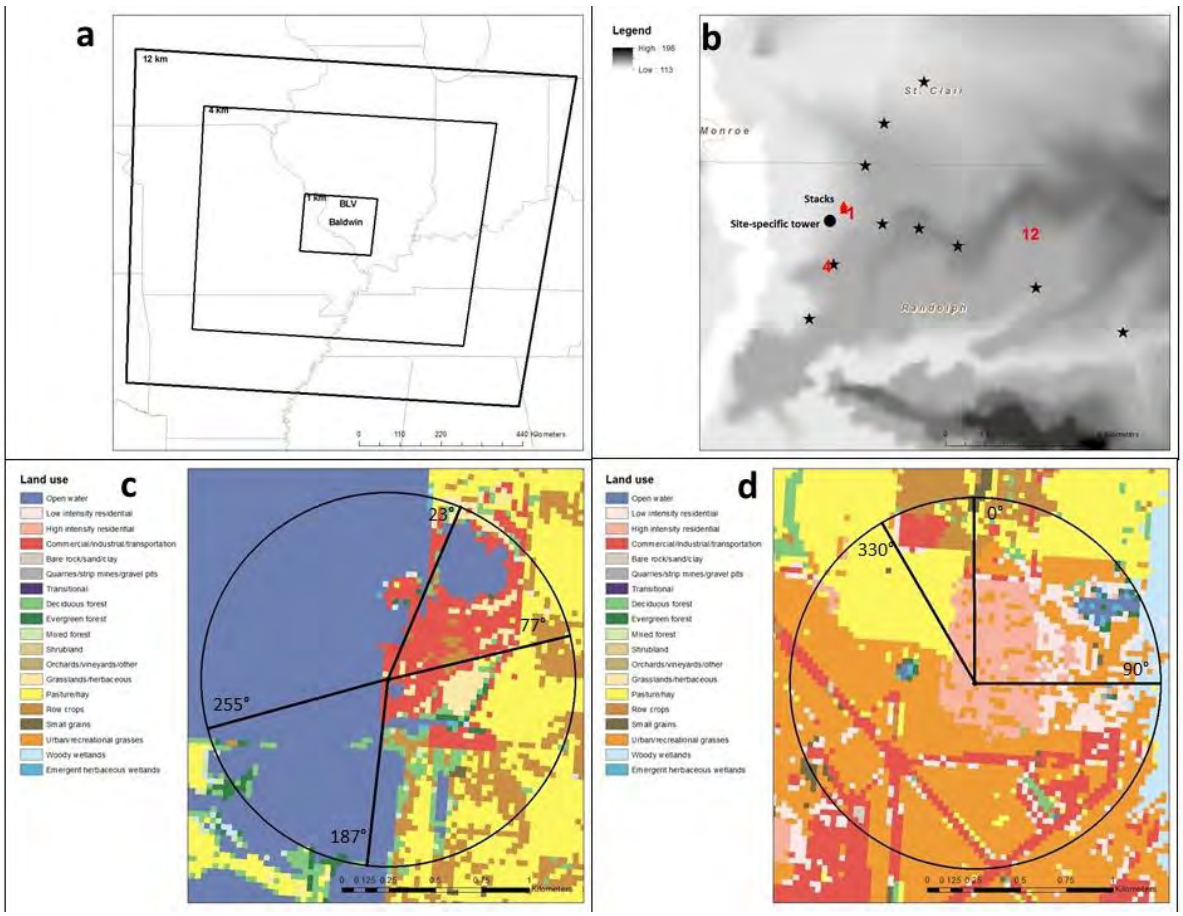

Figure C-1 Baldwin: a) WRF modeling domains; b) locations of emissions, monitors, and meteorological data; c) Baldwin site-specific tower 1992 land use and sectors; and d) BLV 1992 land use and sectors.

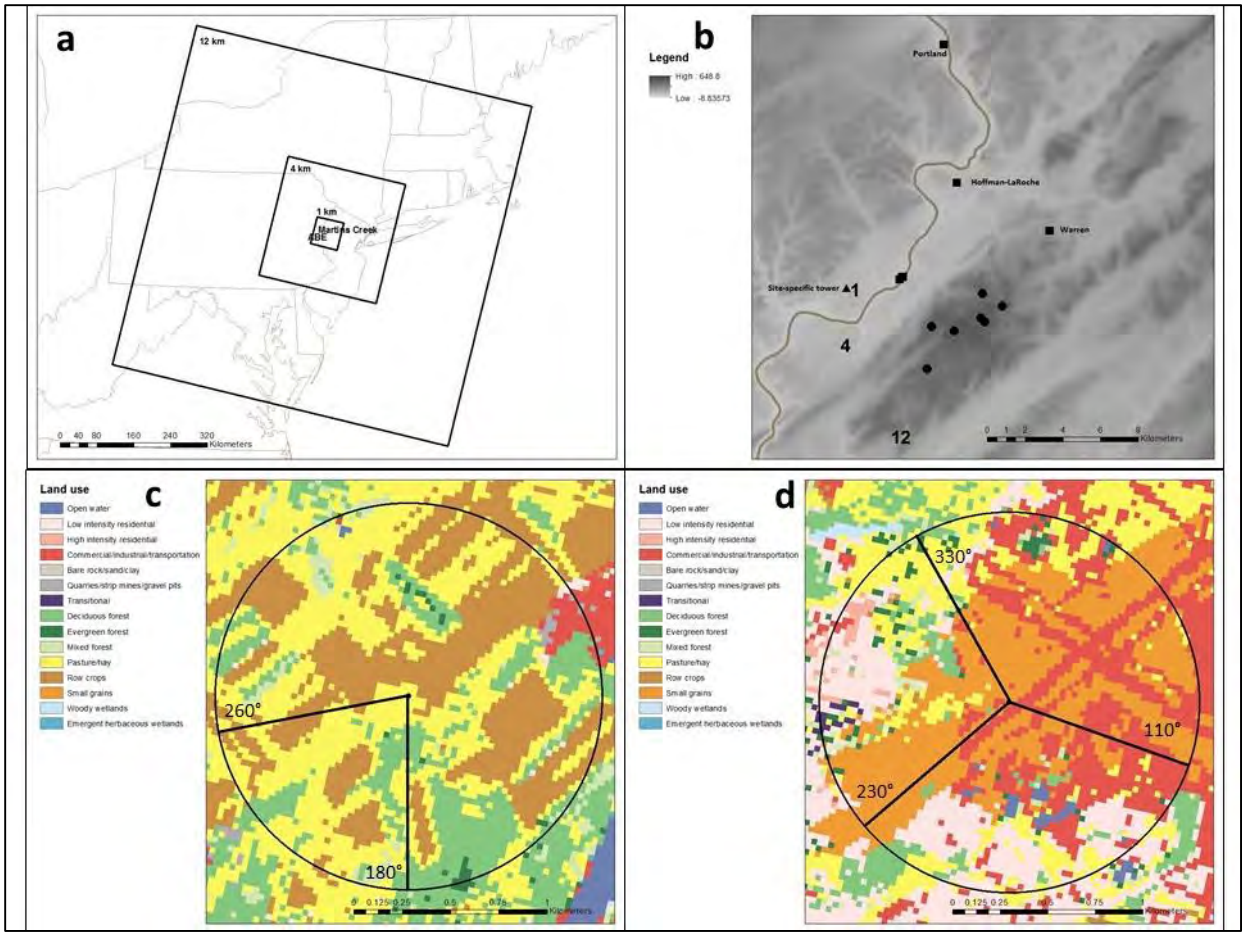

Figure C-2 Martins Creek: a) WRF modeling domains; b) locations of emissions, monitors, and meteorological data; c) Baldwin site-specific tower 1992 land use and sectors; and d) BLV 1992 land use and sectors.

As discussed in U.S. EPA (2003) and Perry et al. (2005), the Baldwin Power Plant is located in a rural, flat terrain setting of southwestern Illinois and has three identical 184-m stacks (5.94 m diameter) aligned approximately north-south with a horizontal spacing of about 100 m with annual (April 1982-March 1983) sulfur dioxide (SO_2) emissions of 85,241 tons, 100,473 tons, and 88,555 tons. Ten SO_2 monitors surrounded the facility, ranging in distance from 2-10 km (Figure 1b). Site-specific meteorological data were available during the study period of April 1982 through March 1983, and consisted of hourly averaged wind speed, wind direction, and temperature at 10 m, and wind speed and wind direction at 100 m. Information regarding Martins Creek can be found in Section 2.1.2 of this document.

C.2 Methodology

C.2.1 Observed meteorological data processing

Observed meteorological data were processed for each of the site-specific towers and the representative NWS stations for the case studies. The site-specific data will be denoted as BAL_OBS and MC_OBS for Baldwin and Martins Creek, respectively. The NWS stations are Scott AFB (BLV) for Baldwin and Lehigh Valley International Airport (ABE) for Martins Creek. These will be denoted as BLV_OBS and ABE_OBS[10], respectively. As with the other evaluations in this document, the NWS stations were processed to determine how MMIF compares to concentrations when only representative NWS data were available for each site.

For the observed site-specific data, surface characteristics included in the evaluation database files were used. For Baldwin, the sectors for the site-specific data, BAL_OBS, are shown in Figure C-1c and the surface characteristic values are shown in Table 1. For BLV_OBS, AERSURFACE (U.S. EPA, 2013) was used to calculate surface characteristics for three sectors (Figure 1d) by season. Surface characteristics for BLV_OBS are shown in Table C-2.

Table C-1. BAL_OBS albedo (A), Bowen ratio (Bo) and surface roughness (z_o) by month and sector.

Month	23°-77°			77°-187°			187°-255°			255°-23°		
	A	Bo	z_o	A	Bo	z_o	A	Bo	z_o	A	Bo	z_o
Jan	0.18	2.0	0.01	0.18	2.0	0.01	0.17	1.7	0.01	0.15	0.4	0.003
Feb	0.18	2.0	0.01	0.18	2.0	0.01	0.17	1.7	0.01	0.15	0.4	0.003
Mar	0.14	0.3	0.03	0.14	0.3	0.03	0.14	0.25	0.02	0.13	0.15	0.009
Apr	0.14	0.3	0.03	0.14	0.3	0.03	0.14	0.25	0.02	0.13	0.15	0.009
May	0.14	0.2	0.03	0.14	0.2	0.03	0.14	0.17	0.02	0.13	0.13	0.009
Jun	0.20	0.3	0.19	0.20	0.3	0.20	0.18	0.25	0.15	0.13	0.15	0.06
Jul	0.20	0.3	0.19	0.20	0.3	0.20	0.18	0.25	0.15	0.13	0.15	0.06
Aug	0.20	0.3	0.19	0.20	0.3	0.20	0.18	0.25	0.15	0.13	0.15	0.06
Sep	0.20	0.3	0.19	0.20	0.3	0.20	0.18	0.25	0.15	0.13	0.15	0.06
Oct	0.18	0.7	0.05	0.18	0.7	0.05	0.17	0.6	0.04	0.15	0.3	0.015
Nov	0.18	0.7	0.05	0.18	0.7	0.05	0.17	0.6	0.04	0.15	0.3	0.015
Dec	0.18	0.4	0.05	0.18	0.4	0.05	0.17	0.3	0.04	0.15	0.2	0.015

[10] MC_OBS corresponds to observed data denoted as Martins Creek in the main document. ABE_OBS corresponds to data denoted as ABE in the main document.

Table C-2. BLV_OBS albedo, Bowen ratio, and surface roughness (zo) by season and sector.

Season	Albedo	Bowen ratio	z_o		
			0°-90°	90°-330°	330°-360°
Winter	0.17	0.64	0.114	0.041	0.098
Spring	0.14	0.33	0.142	0.054	0.121
Summer	0.18	0.47	0.183	0.076	0.258
Fall	0.18	0.64	0.168	0.066	0.253

The Martins Creek surface characteristics are shown in Table C-3 for MC_OBS and in Table C-4 for ABE_OBS. Figure C-2c and C-2d show the locations of the surface roughness sectors discussed in the main document. For a discussion of the calculation of the surface roughness see Section 2.1.2.

Table C-3. MC_OBS albedo, Bowen ratio, and surface roughness (zo) by season and sector.

Month	Albedo	Bowen ratio	z_o	
			180°-260°	260°-180°
Jan	0.15	1.00	0.10	0.30
Feb	0.40	1.30	0.10	0.30
Mar	0.40	0.50	0.10	0.30
Apr	0.15	0.30	0.20	0.50
May	0.15	0.50	0.20	0.50
Jun	0.15	0.40	0.30	0.60
Jul	0.15	0.30	0.30	0.60
Aug	0.15	0.40	0.30	0.60
Sep	0.15	0.80	0.20	0.50
Oct	0.15	2.00	0.20	0.50
Nov	0.15	0.40	0.20	0.50
Dec	0.15	1.00	0.10	0.30

Table C-4. ABE_OBS albedo, Bowen ratio, and surface roughness (zo) by season and sector.

Month	Albedo	Bowen ratio	z_o		
			110°-230°	230°-330°	330°-110°
Jan	0.17	2.08	0.042	0.041	0.023
Feb	0.17	2.08	0.042	0.041	0.023
Mar	0.15	1.50	0.053	0.057	0.030
Apr	0.15	1.50	0.053	0.057	0.030
May	0.15	1.50	0.053	0.057	0.030
Jun	0.17	1.48	0.066	0.089	0.041
Jul	0.17	1.48	0.066	0.089	0.041
Aug	0.17	1.48	0.066	0.089	0.041
Sep	0.17	2.08	0.060	0.078	0.035
Oct	0.17	2.08	0.060	0.078	0.035
Nov	0.17	2.08	0.060	0.078	0.035
Dec	0.17	2.08	0.042	0.041	0.023

C.2.2 WRF simulations

For each of the Baldwin and Martins Creek scenarios, WRF was applied using the Asymmetric Convective Model 2 (ACM2) (Pleim, 2007a; Pleim, 2007b), the Pleim-Xiu land-surface model (Xiu and Pleim, 2001; Pleim and Xiu, 2003), the Morrison double-moment microphysics scheme (Morrison et al., 2009), the rapid radiative transfer model for global climate models (RRTMG) (Iacono et al., 2008), and the Kain-Fritsch cumulus parameterization (Kain, 2004). Two-way nesting without feedback was employed for both scenarios, using objective analysis (nudging) to constrain the model simulations. Nudging of temperature, winds, and moisture was done above the boundary layer only. The North American Regional Reanalysis product (http://nomads.ncdc.noaa.gov) was used for initialization. All simulations utilized 35 vertical layers with a model top of 100 mb and a lowest layer depth of approximately 20 meters.

C.2.2.1 Baldwin

The Baldwin WRF domains were 12-, 4-, and 1-km domains centered over the Baldwin facility (Figure 1a). An annual simulation beginning on March 22, 1982 through April 1, 1983, was generated.

C.2.2.2 Martins Creek

For Martins Creek, a 36-, 12-, 4-, and 1-km setup was used and centered over the Martins Creek plant (Figure 2a). A 2-year WRF simulation spanning the entirety of 1992 and 1993 was

generated. Note, this simulation differs from that discussed in the main document.C.2.3 MMIF processing

For each of the case studies, 24 MMIF scenarios based on grid resolution, PBL treatments, and number of output levels were processed for input into AERMET (U.S. EPA, 2004b). MMIF can process data for input into AERMET for processing for AERMOD, or create AERMOD-ready surface and profile data without having to run AERMET. As discussed in the final 2016 Guideline, for regulatory applications, MMIF should be run to process data for input into AERMET as AERMET is the regulatory meteorological processor for AERMOD. For non-regulatory applications, MMIF can be run to process data for either AERMET or AERMOD input. MMIF outputs for AERMET include a pseudo upper air sounding file, site-specific profile file (ONSITE pathway of AERMET), and surface characteristics (albedo, Bowen ratio, and surface roughness). The site-specific file contained modeled insolation, mixing heights, precipitation, and surface pressure. In AERMET processing, the mixing height is designated as the convective or mechanical mixing height based on the Monin-Obukhov length. Also, at each level output as defined below, temperature, wind speed, wind direction, and relative humidity were output. AERMET was run with a 0.5 m/s wind speed threshold, the Bulk Richardson Number approach for stable conditions, and using the sunrise option for upper air sounding time. These are the default options output by MMIF for AERMET and can be changed by the user.

For both case studies, grid resolutions of one, four, and twelve kilometers were output (denoted as 1, 4, or 12). The 36 km grid for Martins Creek was not used as it would most likely be too coarse given the terrain. The grid cell containing the site-specific meteorological tower was chosen as the representative grid cell for the AERMOD simulations. For each grid resolution, two PBL or mixing height treatments were processed in MMIF, a pass through of the WRF PBL height (PASS) or recalculation of the PBL height using a Bulk Richardson approach (RECALC), based on Vogelezang and Holtslag (1996) and Louis (1979). For details on PBL treatment in MMIF, see the MMIF User's Guide (Environ, 2016). In addition to the PBL treatments, four options for output layers were processed. The first option was to use the MMIF default levels of 20, 30, 60, 120, 240, 480, 920, 1600, 2500, and 3500 m (DEFAULT) which are based on EPA/FLM CALMET guidance.[11] The second option (GUIDANCE) was to output levels recommended in the MMIF guidance (U.S. EPA, 2016b). These levels were 10 m, 25-200 m every 25 m, 200-500 m every 50 m, 500-1000 m every 100 m, and 1000-5000 m every 500 m. These levels correspond to the vertical grid structure used in AERMOD. The third option was to output levels corresponding to the site-specific data (ONSITE), which varied by case study. For Baldwin, this was 10 and 100 m. For Martins Creek, the heights were 10 m, and 90-420 m every 30 m. The fourth option was to output levels near the stack heights (STACK). For Baldwin these were heights of 10, 100 and 180 m and for Martins Creek, 10, 180 and 210 m. For the DEFAULT and GUIDANCE cases, the upper air file and site-specific meteorological file contained all the requested levels. For the ONSITE and STACK cases, the upper air files

[11] https://www3.epa.gov/ttn/scram/guidance/clarification/CALMET%20CLARIFICATION.pdf.

contained the same levels as the GUIDANCE cases and the site-specific profile files contained the heights defined for the ONSITE and STACK cases above.

For MMIF output, the surface characteristics are calculated by MMIF for input into AERMET. The surface characteristics are based on the land use of the processed grid cell. Surface characteristics are monthly and for one sector, i.e., the grid cell. Tables C-5 and C-6 show the MMIF generated surface characteristics for Baldwin and Martins Creek respectively.

Table C-5. MMIF derived albedo (A), Bowen ratio (Bo), and surface roughness (zo) for 1, 4, and 12 km output for Baldwin.

Month	1 km			4 km			12 km		
	A	Bo	z_o	A	Bo	z_o	A	Bo	z_o
Jan	0.26	1.18	0.00044	0.26	1.16	0.00042	0.19	0.91	0.17259
Feb	0.22	1.01	0.00035	0.22	1.07	0.00034	0.26	1.20	0.17393
Mar	0.12	0.20	0.00030	0.12	0.23	0.00026	0.18	1.02	0.17737
Apr	0.09	0.01	0.00028	0.09	0.01	0.00024	0.16	1.10	0.18440
May	0.08	0.02	0.00023	0.08	0.01	0.00019	0.16	0.15	0.18822
Jun	0.08	0.01	0.00018	0.08	0.01	0.00016	0.16	0.01	0.18906
Jul	0.08	0.01	0.00019	0.08	0.01	0.00016	0.16	0.01	0.19430
Aug	0.08	0.05	0.00017	0.08	0.14	0.00015	0.16	0.01	0.19650
Sep	0.08	0.17	0.00019	0.08	0.13	0.00016	0.16	0.03	0.19138
Oct	0.09	0.26	0.00022	0.09	0.24	0.00019	0.16	0.44	0.18221
Nov	0.14	0.35	0.00033	0.14	0.42	0.00029	0.16	1.08	0.17746
Dec	0.14	1.17	0.00034	0.14	1.06	0.00029	0.16	0.53	0.17365

Table C-6. MMIF derived albedo (A), Bowen ratio (Bo), and surface roughness (zo) for 1, 4, and 12 km output for Martins Creek.

Month	1 km			4 km			12 km		
	A	Bo	z_o	A	Bo	z_o	A	Bo	z_o
Jan	0.17	1.25	0.23621	0.18	1.40	0.11273	0.17	1.50	0.22066
Feb	0.23	1.37	0.23383	0.23	1.76	0.10830	0.21	1.89	0.21795
Mar	0.24	1.29	0.24109	0.24	0.98	0.12238	0.22	1.09	0.23165
Apr	0.16	0.22	0.25079	0.17	0.03	0.14067	0.16	0.21	0.25073
May	0.16	0.17	0.26523	0.17	0.07	0.16664	0.16	0.18	0.27683
Jun	0.16	0.08	0.27385	0.17	0.01	0.18149	0.16	0.01	0.29045
Jul	0.16	0.01	0.27323	0.17	0.01	0.18033	0.16	0.01	0.28899
Aug	0.16	0.08	0.27193	0.17	0.01	0.17812	0.16	0.01	0.28697
Sep	0.16	0.13	0.26866	0.17	0.09	0.17255	0.16	0.04	0.28184
Oct	0.16	0.64	0.25813	0.17	0.53	0.15406	0.16	0.75	0.26446
Nov	0.16	0.75	0.24616	0.17	0.9	0.13203	0.16	0.96	0.24165
Dec	0.18	0.89	0.23948	0.18	0.51	0.11916	0.17	0.55	0.22763

For Baldwin, surface roughness outputs by MMIF for 1 and 4 km were much lower than the values used for the site-specific tower because the WRF grid cells were considered all-water cells due to a large cooling lake north of the site-specific tower (see Figure C-1c). For Martins Creek, MMIF output surface roughness values were on par with the site-specific values.

C.2.4 Meteorological data evaluation

Evaluation among the various meteorological datasets for each study encompassed comparing the hourly distributions of each dataset; comparing the distributions of the hourly differences among pairs of datasets; and calculations of several statistics, including mean bias, fractional bias, root mean square error, correlation, and index of agreement. For the wind direction difference statistics, two distance variables, in addition to wind direction differences, were calculated between the modeled and observed winds. The first distance, D, is the difference between two parcels of air that start at the same position and accounts for speed and direction differences. A second distance, Ds, was calculated as simply the distance between two parcels after 1 hour just based on wind speed (observed - modeled), regardless of wind direction. A negative difference means the observed wind speed was over predicted. This provides an operational metric to measure the regulatory aspect of a dispersion model. Finally, the absolute value of the wind direction difference between the observed and modeled winds was calculated.

C.2.5 AERMOD simulations

AERMOD simulations were made for each of the MMIF output scenarios described in Section 2.2. These will be denoted by the grid resolution (1, 4, or 12), PBL treatment (PASS or RECALC) and number of levels option (DEFAULT, GUIDANCE, ONSITE, or STACK). Additionally, AERMOD simulations were completed for the site-specific observed meteorological data and representative NWS data. The AERMOD simulations included hourly emissions for each modeled stack. AERMOD input files were based on the input files provided for the Baldwin and Martins Creek evaluation databases with modification to include the BETA model option to run with AERMET-processed MMIF data. AERMOD output would be evaluated using the protocol outlined in Section 2.3 of this document.

C.3. Results

C.3.1 Meteorological comparisons

Figure C-3 shows the observed 10 m wind roses for the Baldwin site-specific meteorological data (BAL_OBS) and BLV (BLV_OBS) as well as the 1, 4, and 12 km MMIF output for the different levels (DEFAULT, GUIDANCE, ONSITE, and STACK). There were no differences in surface winds between the PBL methodology, PASS and RECALC, so the MMIF wind roses represent both PASS and RECALC. In Figure C-3, the tower at Baldwin and BLV show similar patterns, a strong southeast to southern component of the winds, with more low speed winds at Baldwin. Baldwin also exhibits a stronger northeast component while BLV exhibits a stronger northern component. The 1 km, 4 km, and 12 km MMIF wind roses show similar patterns, with little difference in winds due to grid resolution or number of output levels.

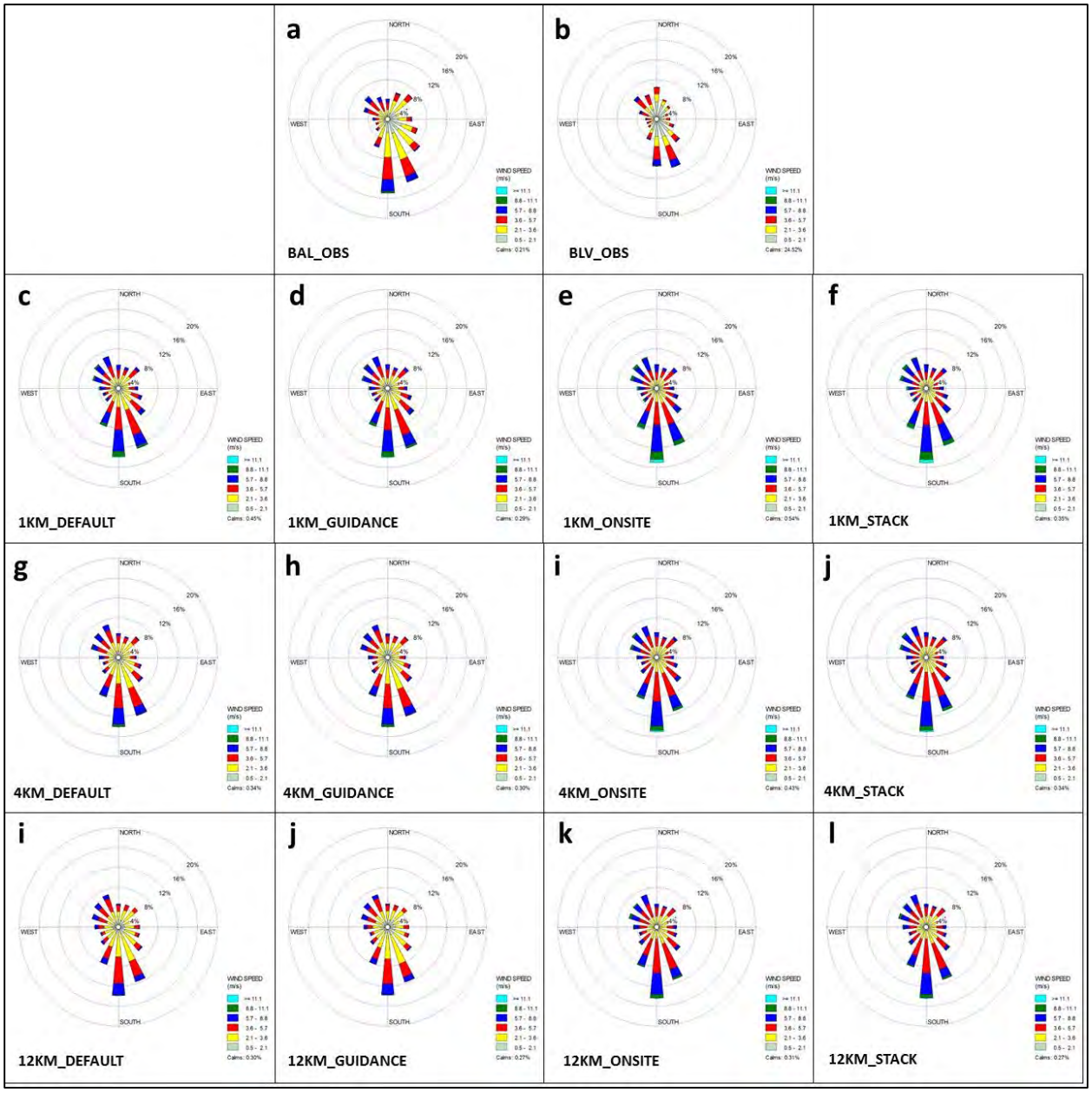

Figure C-3. BAL_OBS, BLV_OBS, and MMIF surface wind roses for Baldwin.

Figures C-4 and C-5 show the wind metrics for BAL_OBS and BLV_OBS, respectively. Shown are the distributions of the hourly distances D and Ds, distance D, and direction differences. For BAL, D is usually less than 20 km. In Figure C-4b, the distance between parcels after 1 hour due to wind speed, Ds, are shown. The distances represent BAL_OBS minus the scenarios on the x-axis. Therefore, a negative (positive) distance indicates that the BAL observed wind speeds are lower (higher) than either BLV_OBS or the MMIF scenarios. This metric agrees well with D in that the magnitude was generally less than 20 km. Figure C-4c shows the magnitude of the wind direction differences. As with the wind speed distance metric, the difference indicates BAL_OBS minus the scenarios on the x-axis. The MMIF scenarios tend to be 30º or less while the differences between BAL_OBS and BLV_OBS approaches 60º on average.

Figure C-5 shows the wind metrics for BLV_OBS. As with BAL_OBS, D is generally less than 20 km. The Ds distances indicated that the BLV_OBS winds were lower than the MMIF scenarios and wind direction differences are generally less than 60º on average.

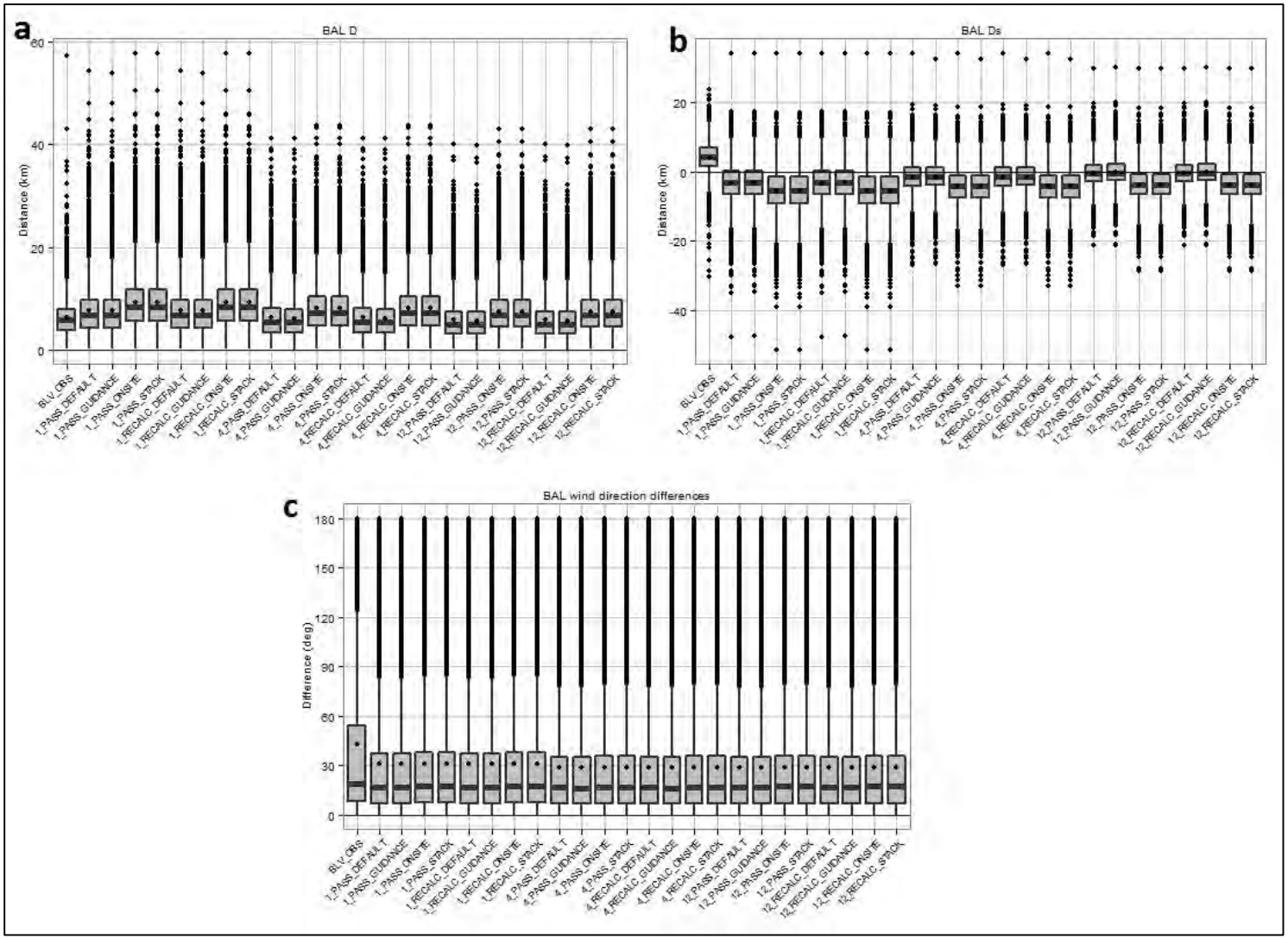

Figure C-4. Wind metrics of BLV_OBS and MMIF scenarios compared to BAL_OBS: a) D (km); b) Ds (km); and c) wind direction differences (degrees).

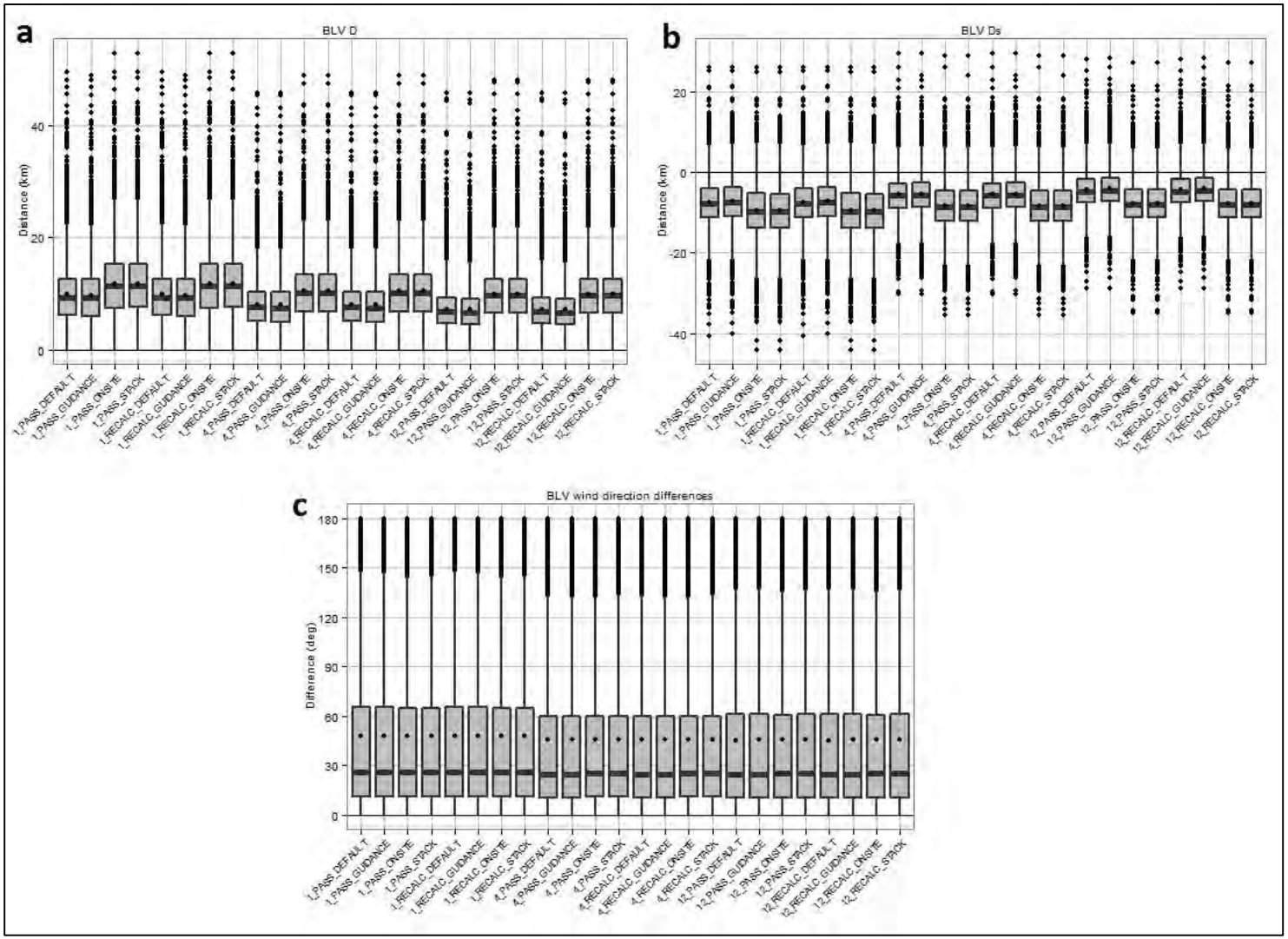

Figure C-5.Wind metrics of MMIF scenarios compared to BLV_OBS: a) D (km); b) Ds (km); and c) wind direction differences (degrees).

Figure C-6 shows the observed wind roses for the Martins Creek site-specific tower (MC_OBS) and ABE (ABE_OBS) as well as the 1, 4, and 12 km MMIF output. The MC_OBS wind rose shows a clear southwest to northeast orientation consistent with the terrain shown in Figure C-2b. The ABE_OBS wind rose shows predominant directions from the southwest, west, northwest and northeast. The 1 km wind roses generally match the pattern of the MC_OBS wind rose. The 4 km and 12 km wind roses appear to agree more with the ABE_OBS wind rose rather than the 1 km wind rose.

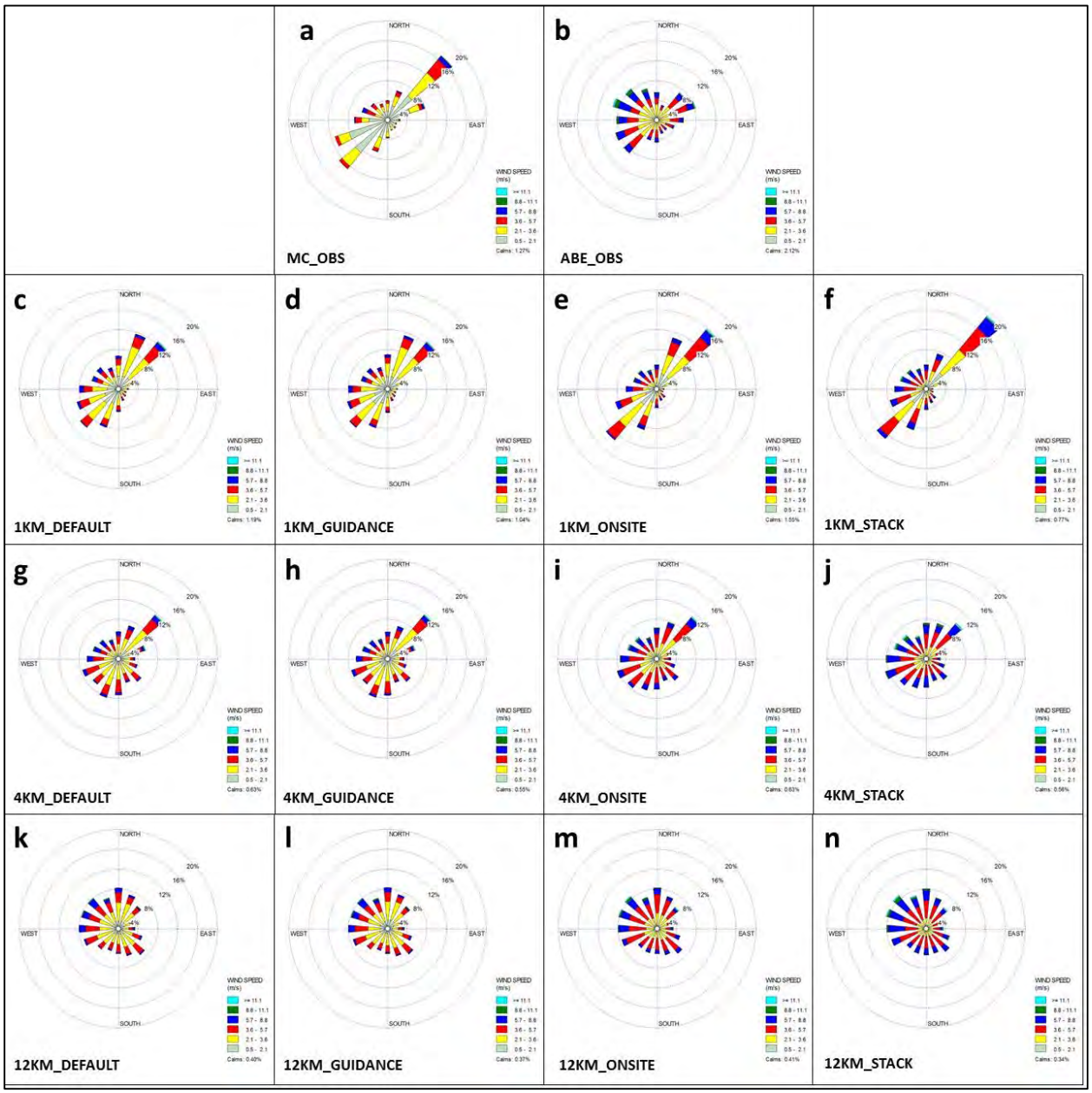

Figure C-6. MC_OBS, ABE_OBS, and MMIF surface wind roses for Martins Creek.

Figures C-7 and C-8 show the wind metric plots for MC_OBS and ABE_OBS. As with the Baldwin cases, the distance D distributions are generally less than 20 km. The Ds plots show that the MC_OBS winds are generally lower than the ABE_OBS or MMIF scenarios. Wind direction differences are generally less than 60º.

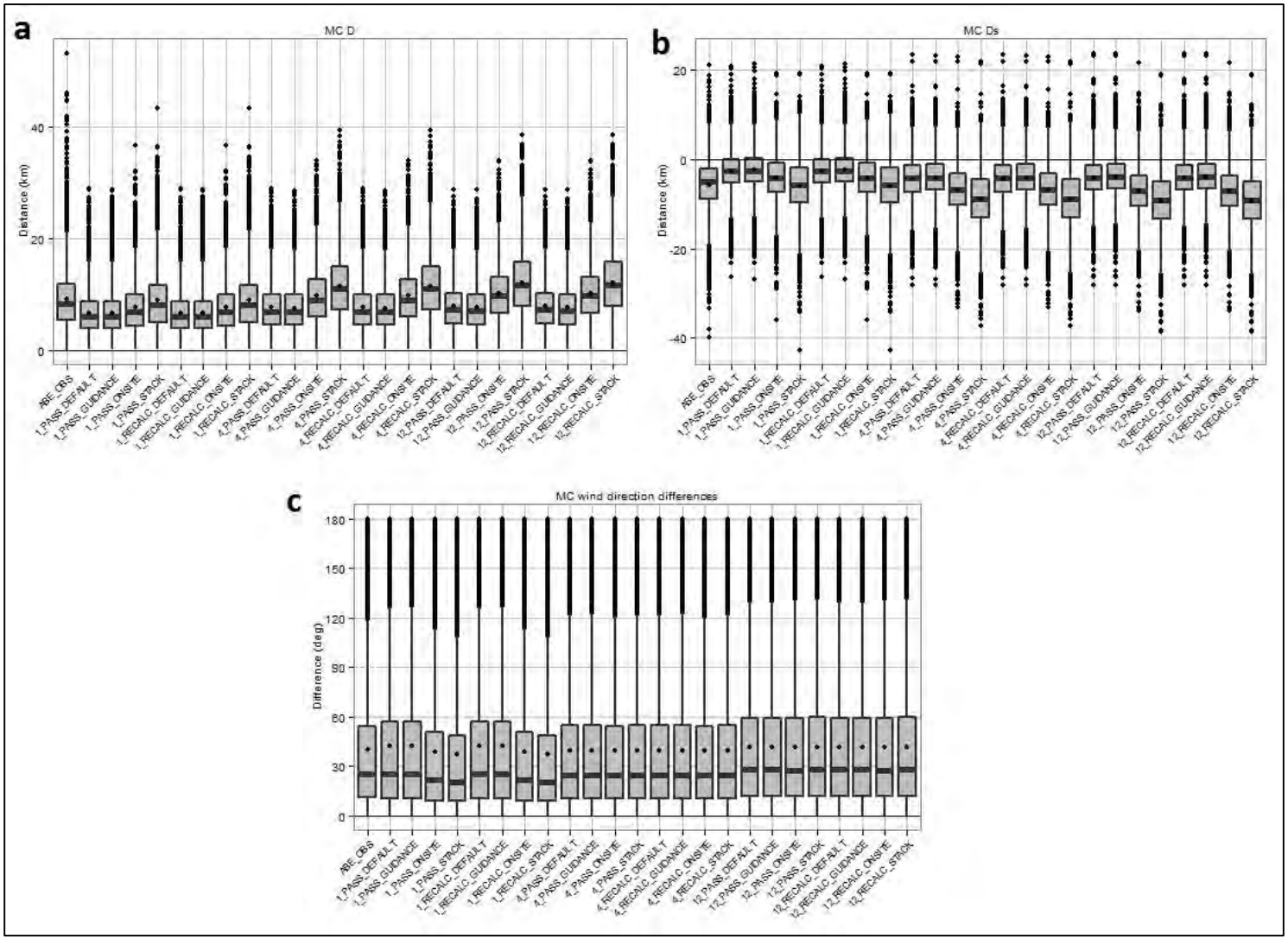

Figure C-7. Wind metrics of ABE_OBS and MMIF scenarios compared to MC_OBS: a) D (km); Ds (km); and c) wind direction differences (degrees).

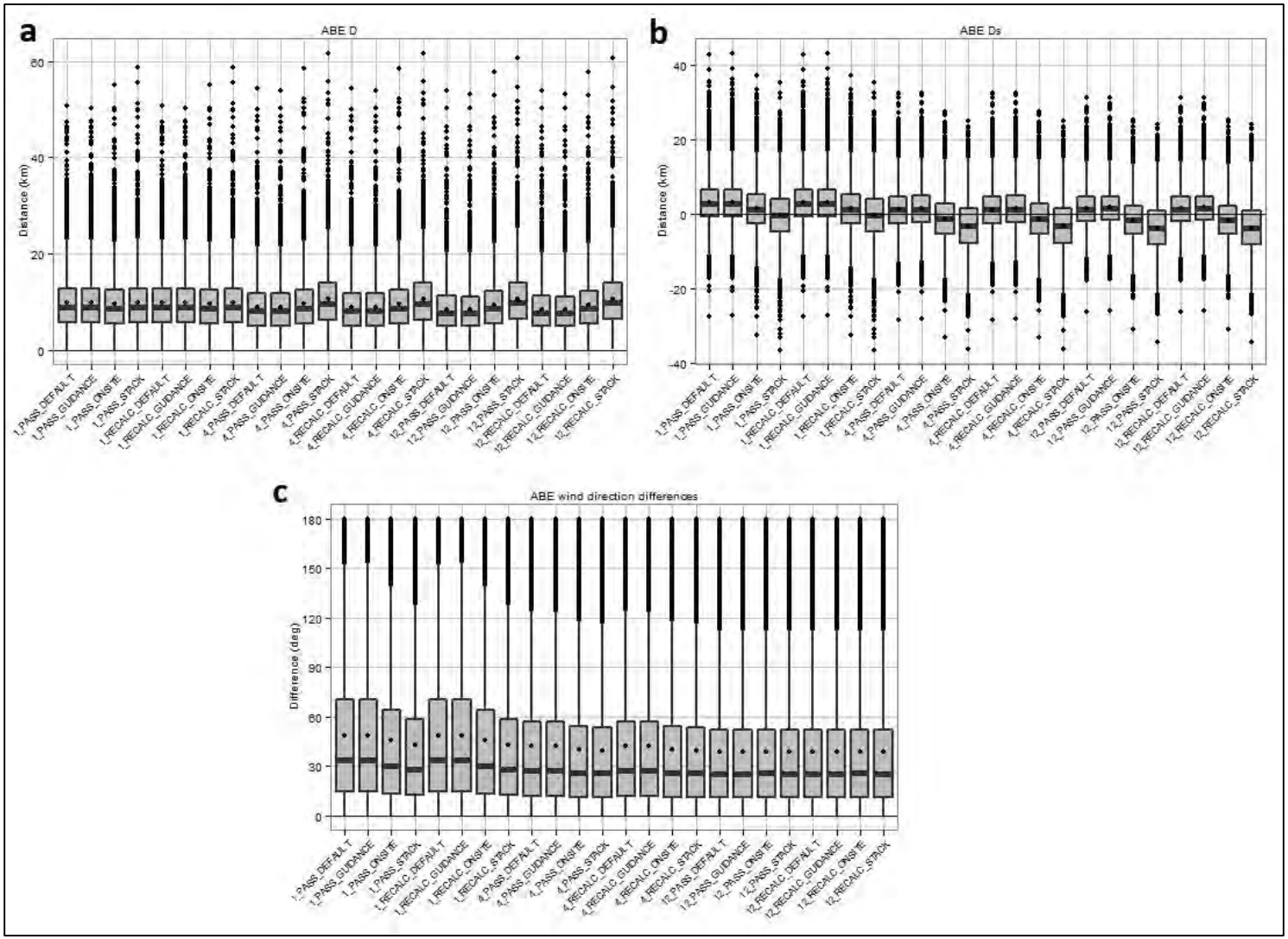

Figure C-8. Wind metrics of MMIF scenarios compared to ABE_OBS: a) D (km); b) Ds (km); and c) wind direction differences (degrees).

Figures C-9 and C-10 show fractional biases for surface temperature, convective mixing height, and mechanical mixing heights compared to the observations. It should be noted that there are no direct comparisons of mixing height and the fractional biases for mixing heights are comparing more of the methodologies used in AERMET and MMIF to calculate mixing heights. As with the wind displacement plots, the fractional biases are BAL_OBS (Figure C-9a) or BLV_OBS (Figure C-9b) minus the scenarios on the x-axis. Negative (positive) fractional biases indicate over prediction (under prediction) by the scenarios on the x-axes of Figure C-9. Temperature fractional biases are negligible. The wind speed fractional biases support the D plots of Figures C-4 and C-8. The 1 km and 4 km MMIF scenarios under predict mechanical mixing heights compared to both BAL_OBS and BLV_OBS. The 12 km MMIF scenarios over predict mechanical mixing heights compared to BAL_OBS and BLV_OBS. Convective mixing heights are over predicted by the BLV_OBS (Figure C-9a) data and the 1 and 4 km PBL PASS through scenarios (compared to BAL_OBS and BLV_OBS). All 12 km MMIF scenarios over predict the convective mixing heights. For the 1 and 4 km scenarios, the PBL RECALC scenarios under predict the convective mixing heights, compared to both BAL_OBS and BLV_OBS. Figure C-9c shows the fractional biases of convective and mechanical mixing heights between the PASS and RECALC scenarios for each grid resolution and number of levels, e.g., 1_DEFAULT is the fractional bias of 1_PASS_DEFAULT – 1_RECALC_DEFAULT. A negative (positive) fractional bias indicates over prediction (under prediction) by the RECALC methodology relative to the PASS methodology. The 1 and 4 km results show similar fractional biases between the grid resolutions, with a smaller bias, i.e., closer to zero, for the 12 km results. This may be due to the surface characteristics between the 12 km and 1 and 4 km resolutions used to recalculate the mixing heights. All scenarios indicate a bias to over predict by the RECACL methodology, with the smallest biases (close to zero) for the 12 km scenarios.

Figure C-10 shows the fractional biases for MC_OBS (Figure C-10a) and ABE_OBS (Figure C-10b). As with Figure C-9, negative (positive) fractional biases indicate over prediction (under prediction) by the scenarios on the x-axes of Figure C-10. As with the Baldwin study, temperature fractional biases are negligible. Compared to MC_OBS, ABE_OBS and all MMIF scenarios over predict for the mixing heights. Compared to ABE_OBS, all scenarios under predict the convective mixing heights. MMIF scenarios under predict for mechanical mixing heights, with the exception of 4_PASS_ONSITE; 4_PASS_STACK; 12_PASS_ONSITE; and 12_PASS_STACK.

Analogous to Figure C-9c, Figure C-10c shows the fractional biases of convective and mechanical mixing heights between the PASS and RECALC scenarios for each grid resolution and number of levels. All scenarios indicate a bias to over predict by the RECACL methodology, but not as dramatic as the Baldwin cases, possibly due to more similar surface characteristics among the resolutions. Across the three grid resolutions, the fractional biases are fairly constant for the convective mixing heights. For mechanical mixing heights, the 12 km biases are closer to zero than the 1 and 4 km.

Overall, for both scenarios and the presented variables, the fractional biases are within acceptable bounds.

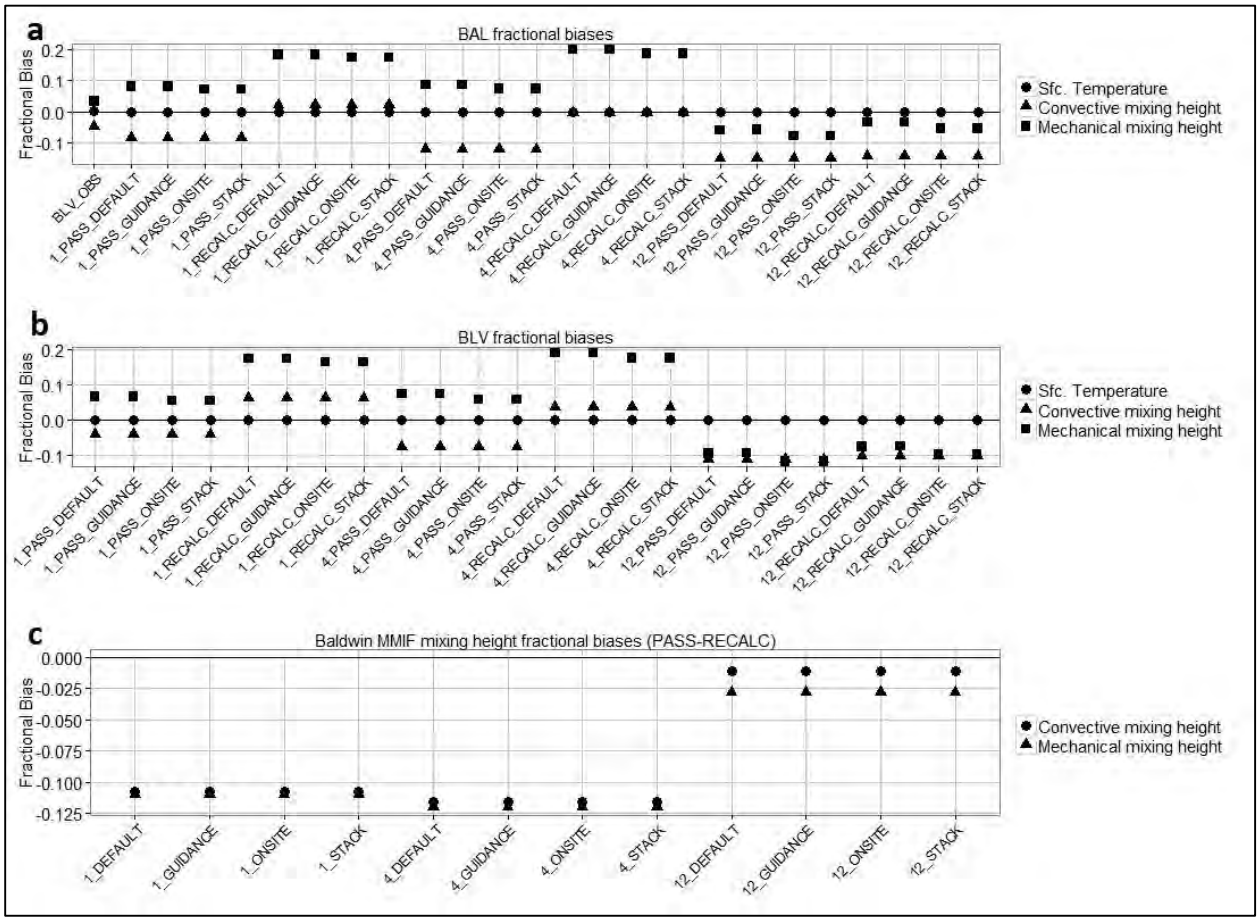

Figure C-9. a) Mean fractional biases for selected meteorological variables of MMIF compared to BAL_OBS; b) mean fractional biases for selected meteorological variables of MMIF compared to BLV_OBS; and c) mean fractional biases for mixing heights between PASS and RECALC MMIF output.

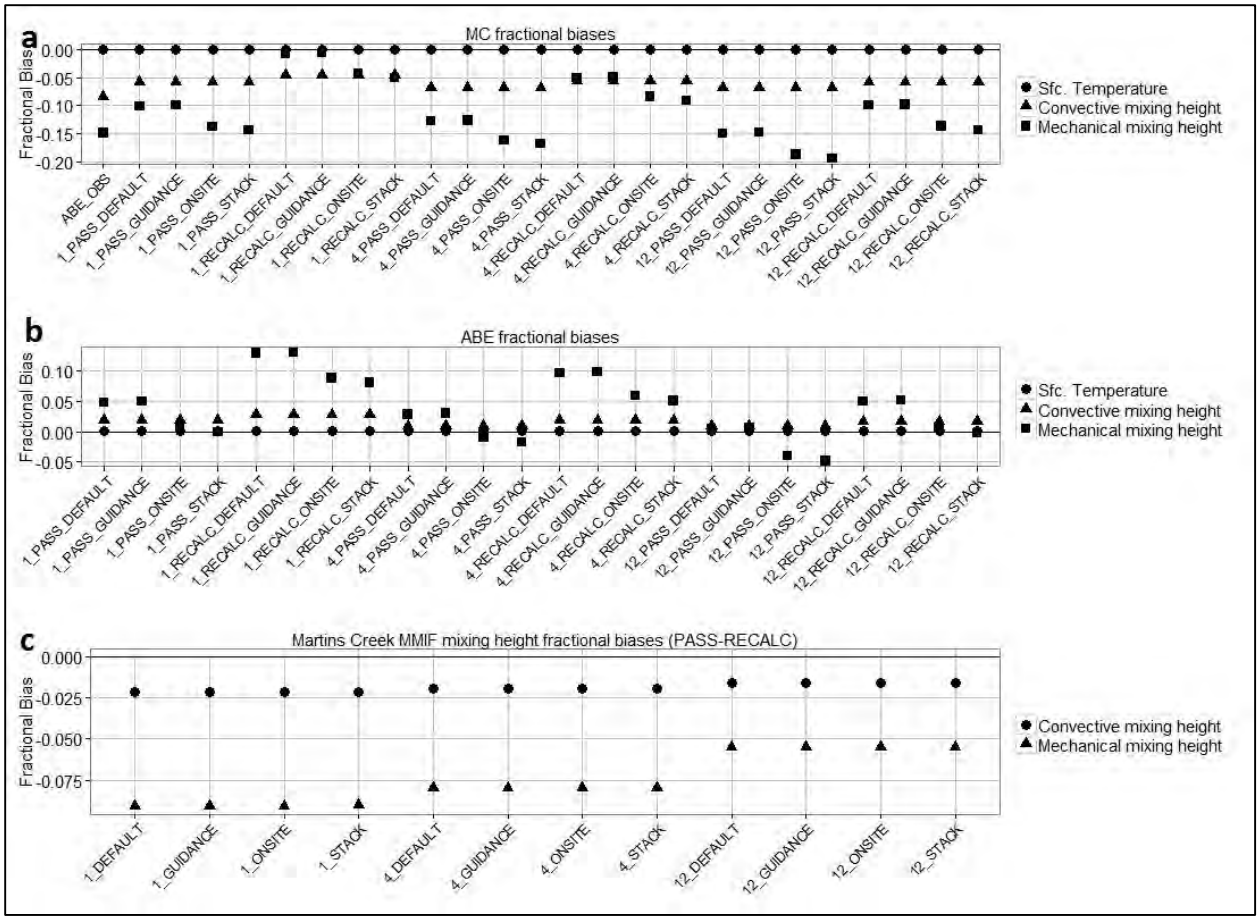

Figure C-10. **a) Mean fractional biases for selected meteorological variables of MMIF compared to MC_OBS; b) mean fractional biases for selected meteorological variables of MMIF compared to ABE_OBS; and c) mean fractional biases for mixing heights between PASS and RECALC MMIF output.**

C.3.2 AERMOD results

C.3.2.1 MMIF versus MMIF evaulations

Figures C-11 through C-13 summarize the results of the statistical tests of the observed meteorological datasets and the MMIF scenarios. Figure C-11 shows the CPMs for Baldwin (Figure C-11a) and Martins Creek (Figure C-11b). Also shown are the 95 percent confidence intervals for the CPM based on the bootstrapping results. The confidence intervals indicate that the modeled results are probably statistically different than the observed concentrations. For Baldwin, the CPM values are relatively close among all scenarios. The best performing scenario is 1_RECALC_DEFAULT (0.36) followed closely by 4_RECALC_GUIDANCE (0.38). These two scenarios are the only two scenarios below 0.4 for the CPM. Ranking the scenarios by CPM, eight MMIF scenarios performed better than BAL_OBS (0.42), in order of lowest to

highest CPM: 1_RECALC_DEFAULT; 4_RECALC_GUIDANCE; 12_RECALC_DEFAULT; 12_PASS_DEFAULT; 1_PASS_DEFAULT; 4_PASS_GUIDANCE; 12_PASS_GUIDANCE; and 4_RECALC_DEFAULT. The worst performing scenario was BLV_OBS.

For Martins Creek, the best performing scenario was the 12_PASS_STACK scenario (0.36) and was the only one to perform better than the MC_OBS scenario (0.37). This was surprising in that one would expect the 1 or 4 km scenarios to perform better than the 12 km scenario given the wind roses. Two MMIF scenarios performed worse than the NWS data, ABE_OBS (0.66), though only slightly worse: 4_RECALC_GUIDANCE and 4_RECALC_ONSITE.

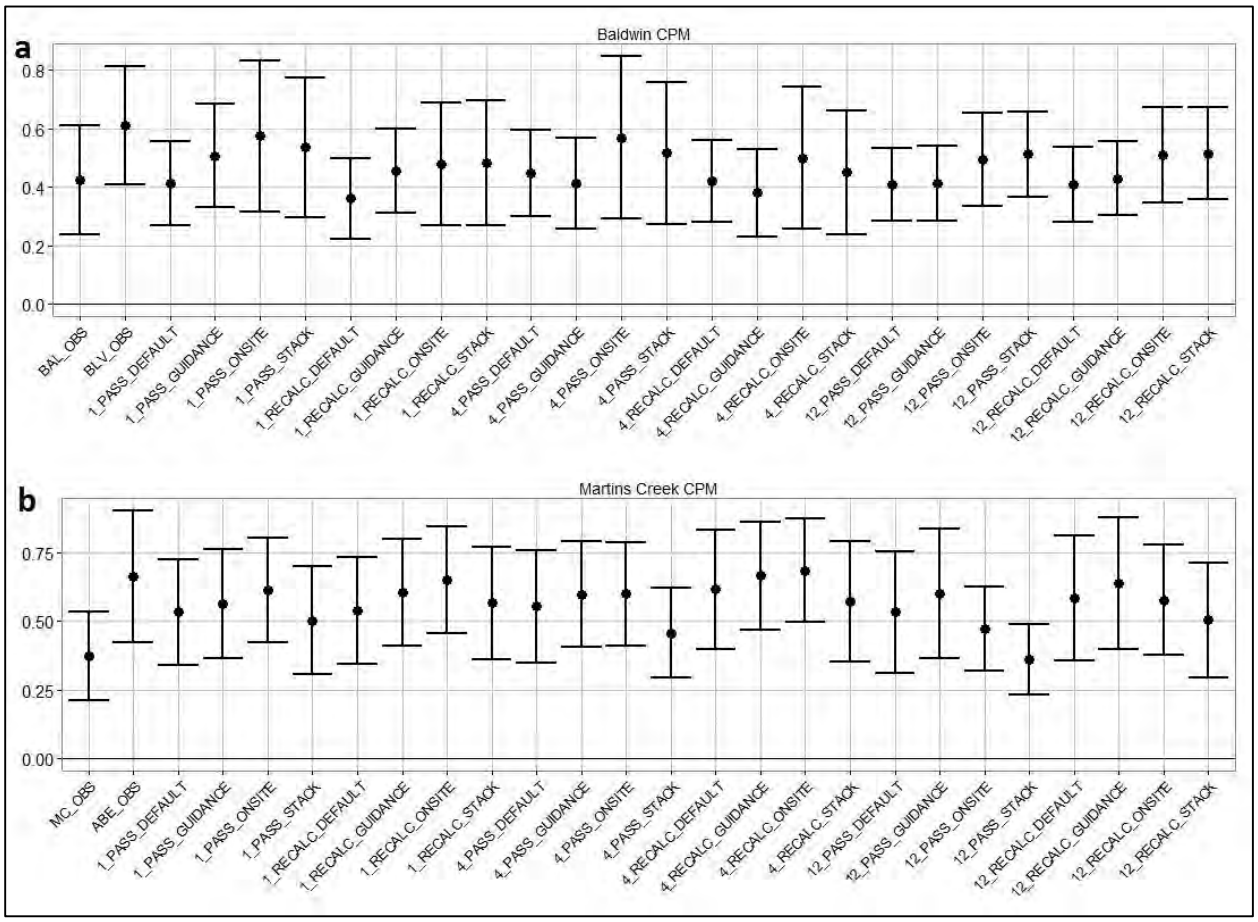

Figure C-11.CPM results for a) Baldwin and b) Martins Creek.

After completing the analyses, the question remained: Did grid resolution, PBL treatment, and the number of levels make a difference in results? To determine if grid resolution made a difference, the best performing resolution was determined by analyzing the MCM pairings where the grid resolution differed but PBL treatment and number of levels were kept the same (e.g., MCM pairings OF 1_PASS_DEFAULT with 4_PASS_DEFAULT, 1_PASS_GUIDANCE with 4_PASS_GUIDANCE, etc.). A similar approach was done for the PBL treatment and number of levels. The results are shown in Figures C-12 for Baldwin and Figure C-13 for Martins Creek. Shown are the MCM for each scenario pair as well as the 95 percent confidence intervals.

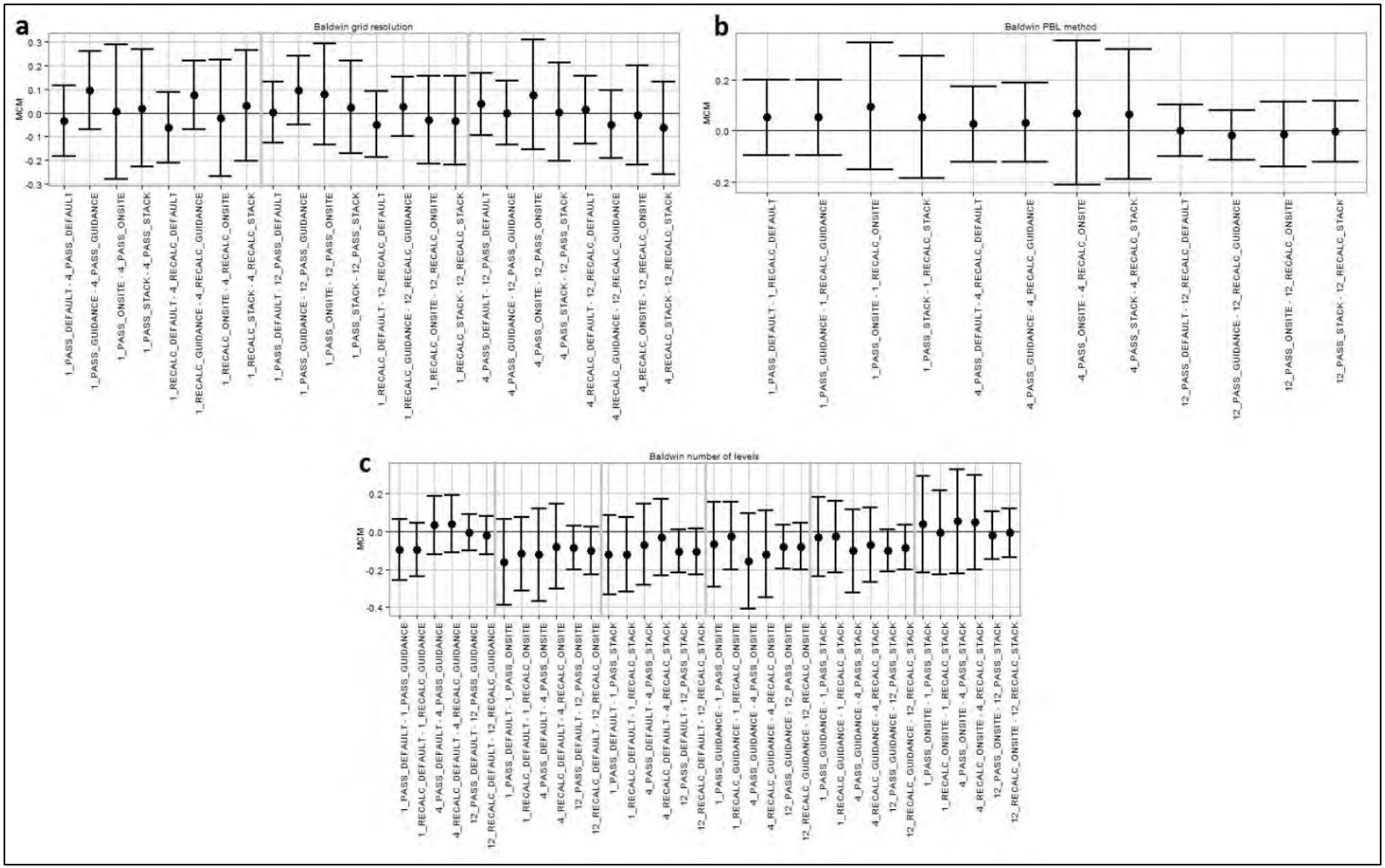

Figure C-12. Baldwin selected MCM pairings to determine best performing a) grid resolution; b) PBL method; and c) number of levels.

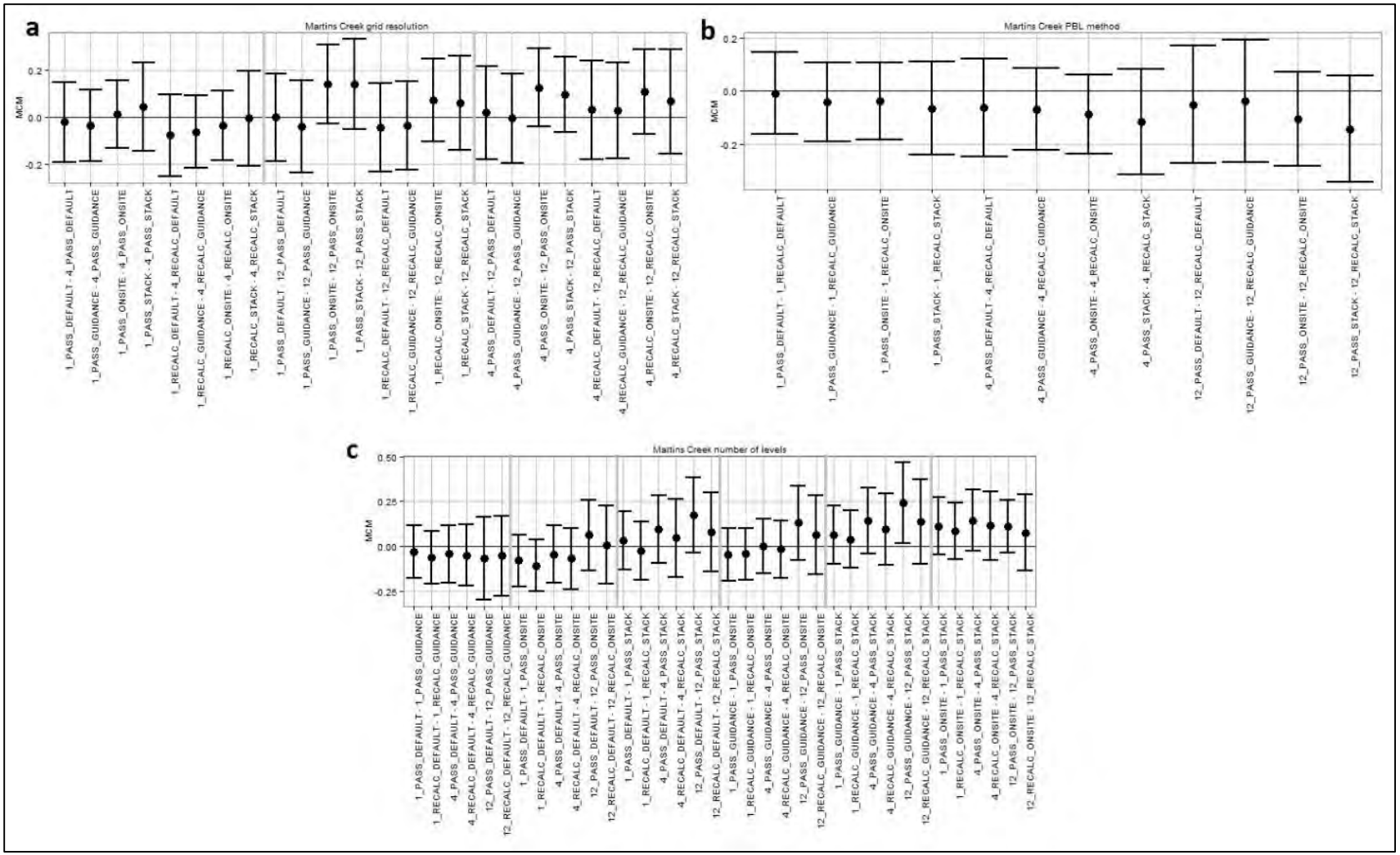

Figure C-13. **Martins Creek selected MCM pairings to determine best performing a) grid resolution; b) PBL method; and c) number of levels.**

For Baldwin, the MCM results showed that the differences in grid resolution were not statistically different. There was no clear better performer between 1 and 4 km resolution. In eight comparisons, the 4 km results were better (as indicated by positive MCMs) five times, but always insignificant. Somewhat surprisingly, 12 km results were usually better than 1 km results; five of the comparisons were better for 12 km. For 4 versus 12 km, the results were four to four. The lack of a clear better performer for grid resolution is not too surprising given the flat nature of the terrain around Baldwin. However, the better performance for 12 km may have been due to surface characteristics of the 12 km MMIF output being more representative of the observed surface characteristics, while the 1 and 4 km surface characteristics were more indicative of over water surface characteristics.

The PBL comparisons for Baldwin showed that the RECALC method performed better than the PASS method; nine of the twelve comparisons showed RECALC was better. For the number of levels, the DEFAULT and GUIDANCE were the better performing scenarios compared to ONSITE and STACK, with DEFAULT performing better against GUIDANCE for four of six cases. However, all comparisons were insignificant at the 95 percent confidence level.

The Martins Creek analyses in Figure C-13 showed that 1 km resolution outperformed 4 km resolution for the most part and was even when compared to 12 km resolution. But 12 km resolution outperformed 4 km resolution in all cases, though not statistically different. For the PBL methodology, the PASS methodology was better in all scenarios, but not statistically different. For the number of levels, the better performing level was the STACK scenario, though mostly insignificant. This may have been the case as the winds representing the stack level may have been representative of the monitors given the complex terrain.

C.3.2.2 MMIF mixing heights versus AERMET mixing heights

As part of the update to MMIF, the option to allow AERMET to calculate mixing heights instead of using MMIF supplied mixing heights (recalculated or pass through). At the time of these evaluations, this option was still being implemented in MMIF, but the ability to use AERMET mixing heights can still be utilized with MMIF inputs to AERMET that already include MMIF mixing heights. This was done by eliminating the MIXHTS variable in the READ statement of the AERMET stage 1 input control file (i.e. AERMET.INP file). The FORMAT statement in the stage 1 input file was also modified to skip the mixing height data in the surface file (Figure C-14).

AERMET was run with the MMIF data and mixing heights calculated. Table C-7 shows the fractional biases of mixing height comparisons between the observed data (BAL and BLV) and the AERMET generated mixing heights using MMIF output and the fractional biases between the MMIF generated mixing heights (PASS or RECALC) and the AERMET generated mixing heights using MMIF output. Fractional biases are calculated as the difference between columns listed in Table C-7 and the AERMET generated mixing heights using MMIF output (i.e. BAL convective mixing heights – AERMET mixing heights from MMIF). Table C-7 shows that

AERMET generated mixing heights from MMIF output are less than the mixing heights from observed data and MMIF provided mixing heights. For mechanical mixing heights, the AERMET generated estimates from MMIF output are higher than the mixing heights from observed data and MMIF provided mixing heights for all 12 km scenarios and the AERMET generated mixing heights from MMIF output are higher than the RECALC mixing heights for all scenarios.

Table C-7. Baldwin fractional biases for convective and mechanical mixing heights comparing AERMET generated mixing heights using MMIF output to mixing height from observed data and MMIF provided mixing heights (PASS or RECALC).

	Convective mixing heights				Mechanical mixing heights			
Scenario	BAL	BLV	PASS	RECALC	BAL	BLV	PASS	RECALC
1_DEFAULT	0.09	0.13	0.16	0.05	0.11	0.09	0.04	-0.08
1_GUIDANCE	0.10	0.13	0.16	0.06	0.11	0.09	0.03	-0.08
1_ONSITE	0.09	0.13	0.15	0.05	0.07	0.06	0.00	-0.11
1_STACK	0.09	0.12	0.15	0.05	0.07	0.06	0.00	-0.11
4_DEFAULT	0.10	0.13	0.19	0.08	0.15	0.13	0.07	-0.06
4_GUIDANCE	0.10	0.13	0.19	0.08	0.16	0.13	0.07	-0.06
4_ONSITE	0.09	0.13	0.18	0.07	0.10	0.08	0.02	-0.10
4_STACK	0.09	0.13	0.18	0.07	0.10	0.08	0.02	-0.10
12_DEFAULT	0.06	0.10	0.18	0.17	-0.13	-0.15	-0.07	-0.10
12_GUIDANCE	0.06	0.10	0.18	0.17	-0.13	-0.15	-0.08	-0.10
12_ONSITE	0.05	0.09	0.18	0.17	-0.20	-0.21	-0.13	-0.15
12_STACK	0.05	0.09	0.18	0.17	-0.20	-0.21	-0.13	-0.15

Table C-8 shows the mixing height fractional biases for Martins Creek. For convective mixing heights, the AERMET generated mixing heights from MMIF output are lower than the mixing heights based on observed data and MMMIF supplied mixing heights (PASS or RECALC). For mechanical mixing heights, the AERMET generated mixing heights from MMIF output are higher than the mixing heights based on observed data or MMIF supplied mixing heights (PASS or RECALC).

Table C-8. Martins Creek fractional biases for convective and mechanical mixing heights comparing AERMET generated mixing heights using MMIF output to mixing height from observed data and MMIF provided mixing heights (PASS or RECALC).

	Convective mixing heights				Mechanical mixing heights			
Scenario	MC	ABE	PASS	RECALC	MC	ABE	PASS	RECALC
12_DEFAULT	0.12	0.18	0.17	0.16	-0.18	-0.03	-0.04	-0.09
12_GUIDANCE	0.12	0.18	0.17	0.16	-0.18	-0.03	-0.04	-0.09
12_ONSITE	0.11	0.17	0.17	0.15	-0.29	-0.19	-0.15	-0.18
12_STACK	0.11	0.17	0.16	0.15	-0.32	-0.22	-0.18	-0.20
1_DEFAULT	0.08	0.15	0.13	0.11	-0.09	0.06	0.01	-0.08
1_GUIDANCE	0.09	0.16	0.13	0.12	-0.11	0.04	-0.01	-0.11
1_ONSITE	0.08	0.15	0.13	0.11	-0.21	-0.08	-0.08	-0.15
1_STACK	0.08	0.14	0.12	0.11	-0.24	-0.12	-0.12	-0.18
4_DEFAULT	0.15	0.21	0.20	0.18	-0.12	0.03	0.00	-0.08
4_GUIDANCE	0.15	0.21	0.20	0.18	-0.13	0.02	-0.01	-0.09
4_ONSITE	0.14	0.20	0.20	0.18	-0.24	-0.11	-0.10	-0.16
4_STACK	0.14	0.20	0.19	0.17	-0.27	-0.16	-0.14	-0.19

AERMOD was then run and results compared against the existing AERMOD output for the MMIF supplied mixing heights. In this comparison, there were only two model outputs compared at one time: AERMOD using PASS mixing heights against AERMOD using AERMET mixing heights or AERMOD using RECALC mixing heights against AERMOD using AERMET mixing heights. This was done for all three grid resolutions and four number of level options. Since only two models are being compared, the test statistic is the MCM divided by the standard error (standard deviation) from bootstrapped results. If the ratio is ±1.7 then the models are not considered statistically significant (U.S. EPA, 1992).

Table C-9 shows the MCM, standard error, and ratios for Baldwin and Table C-10 shows the same for Martins Creek. For Baldwin, the results show that for the most part for 1 and 4 km resolution that the AERMET mixing heights perform slightly better than MMIF supplied mixing heights while at 12 km, the MMIF mixing heights perform better, although all ratios are within the ±1.7 bounds for significance, meaning AERMET mixing heights and MMIF mixing heights are not statistically different. Table C-8 shows the same statistics for Martins Creek. Those results appear to indicate that for the most part, MMIF performs better than the AERMET mixing heights, although generally not at a statistically significant level. The two exceptions are for 4_RECALC_ONSITE and 4_RECALC_STACK, indicating that the 4_RECALC_ONSITE AND 4_RECALC_STACK are statistically better than AERMOD results using AERMET mixing heights. However, the DEFAULT and GUIDANCE cases indicate insignificance and all PASS cases indicate insignificance.

Table C-9. Baldwin MCM and test statistics for MMIF mixing heights – AERMET mixing heights. AERMET mixing height CPM values are subtracted from the listed scenarios.

Scenario	MCM	Standard error	Ratio
1_PASS_DEFAULT	0.04	0.09	0.44
1_PASS GUIDANCE	-0.07	0.09	-0.78
1_PASS ONSITE	0.01	0.16	0.06
1_PASS STACK	0.01	0.16	0.06
1_RECALC_DEFAULT	0.10	0.09	1.11
1_ RECALC GUIDANCE	-0.02	0.08	-0.25
1_ RECALC ONSITE	0.11	0.15	0.73
1_ RECALC STACK	0.06	0.16	0.38
4_PASS_DEFAULT	0.03	0.08	0.38
4_PASS GUIDANCE	0.08	0.10	0.80
4_PASS ONSITE	0.01	0.17	0.06
4_PASS STACK	0.03	0.15	0.2
4_RECALC_DEFAULT	0.05	0.09	0.56
4_ RECALC GUIDANCE	0.11	0.10	1.1
4_ RECALC ONSITE	0.08	0.16	0.50
4_ RECALC STACK	0.10	0.16	0.63
12_PASS_DEFAULT	-0.03	0.06	-0.50
12_PASS GUIDANCE	-0.04	0.06	-0.67
12_PASS ONSITE	-0.05	0.07	-0.71
12_PASS STACK	-0.06	0.07	-0.86
12_RECALC_DEFAULT	-0.03	0.06	-0.50
12_ RECALC GUIDANCE	-0.06	0.06	-1.00
12_ RECALC ONSITE	-0.07	0.07	-1.00
12_ RECALC STACK	-0.06	0.08	-0.75

Table C-10. Martins Creek MCM and test statistics for MMIF mixing heights – AERMET mixing heights. AERMET mixing height CPM values are subtracted from the listed scenarios.

Scenario	MCM	Standard error	Ratio
1_PASS_DEFAULT	-0.02	0.09	-0.22
1_PASS GUIDANCE	-0.02	0.09	-0.22
1_PASS ONSITE	-0.04	0.09	-0.44
1_PASS STACK	-0.11	0.11	-1.00
1_RECALC_DEFAULT	-0.03	0.09	-0.33
1_ RECALC GUIDANCE	-0.06	0.08	-0.75
1_ RECALC ONSITE	-0.08	0.08	-1.00
1_ RECALC STACK	-0.17	0.11	-1.55
4_PASS_DEFAULT	0.04	0.10	0.4
4_PASS GUIDANCE	0.01	0.08	0.13
4_PASS ONSITE	-0.07	0.08	-0.88
4_PASS STACK	-0.12	0.10	-1.2
4_RECALC_DEFAULT	-0.02	0.11	-0.18
4_ RECALC GUIDANCE	-0.06	0.09	-0.67
4_ RECALC ONSITE	-0.16	0.09	-1.78
4_ RECALC STACK	-0.23	0.12	-1.92
12_PASS_DEFAULT	0.01	0.13	0.08
12_PASS GUIDANCE	-0.01	0.13	-0.08
12_PASS ONSITE	-0.04	0.09	-0.44
12_PASS STACK	0.11	0.09	1.22
12_RECALC_DEFAULT	-0.04	0.13	-0.31
12_ RECALC GUIDANCE	-0.04	0.14	-0.29
12_ RECALC ONSITE	-0.14	0.12	-1.17
12_ RECALC STACK	-0.04	0.14	-0.29

C.4 Summary and conclusions

MMIF processing and AERMOD simulations were done for two of the AERMOD evaluation databases: Baldwin, IL and Martins Creek, PA. The resulting analyses using the EPA protocol for best performing model were done to determine if MMIF output performed as well as site-specific data and representative NWS data. Three grid resolutions, two PBL methodologies, and four height level options were assessed. The two PBL methodologies were also compared against model results using the MMIF data but allowing AERMET to calculate mixing heights. Analyses of the meteorological data indicated good performance against the site-specific and NWS data. For air quality concentrations, the findings for Baldwin indicate there was no clear better performance of either grid resolution over the other. The results do indicate that the PBL recalculation performed better than the PBL pass through and, using the default MMIF (i.e., FLM CALMET) or EPA draft MMIF guidance, levels performed better than only including the onsite or stack height levels. However, the results were not statistically significant. For Martins

Creek, 12 km appeared to be the better performing resolution with the PBL pass through the better PBL option. The STACK level option performed better for the number of levels. When comparing the MMIF mixing heights (PASS or RECALC) to AERMET mixing heights, Baldwin showed slightly better performance for the AERMET mixing heights, though not statistically significant. Martins Creek comparisons indicated better performance for the MMIF mixing heights over AERMET mixing heights and two of the cases using PBL recalculation and onsite or stack heights for the levels indicated statistically better performance for MMIF mixing heights.

United States
Environmental Protection
Agency

Office of Air Quality Planning and Standards
Air Quality Assessment Division
Research Triangle Park, NC

Publication No. EPA-454/R-16-004
December, 2016