

# **Appendix D-1**

## **Area of Influence Analyses**

*This page intentionally left blank.*



# Area of Influence Analysis Southeastern VISTAS II Regional Haze Analysis Project

Documentation Report for Task 5

Prepared for:

**Southeastern States Air Resource Managers, Inc.  
(SESARM)**

1252 W. Government St., #1375  
Brandon, MS 39043

Prepared by:

**Eastern Research Group, Inc. (ERG)**

1600 Perimeter Park Drive  
Suite 200  
Morrisville, NC 27560

Revised Final – December 2, 2020

This page is intentionally blank.

**TABLE OF CONTENTS**

|                                               | <b>Page</b> |
|-----------------------------------------------|-------------|
| 1. INTRODUCTION.....                          | 1           |
| 2. APPROACH .....                             | 3           |
| 2.1 Identification of Sites .....             | 3           |
| 2.2 Identification of Dates for Analysis..... | 5           |
| 2.2.1 Emissions Data Collection.....          | 9           |
| 2.3 Trajectories .....                        | 12          |
| 2.3.1 Meteorological Data .....               | 12          |
| 2.3.2 Trajectory Set up .....                 | 13          |
| 2.4 Residence Time.....                       | 17          |
| 2.5 Extinction-Weighted Residence Time .....  | 21          |
| 2.6 Contributor Identification.....           | 27          |

Appendix A: Using Python to Generate HYSPLIT Control Files

Appendix B: Running HYSPLIT with Batch Files

Appendix C: AoI Analysis with R

Appendix D: AoI Analysis with Python and ArcGIS

Appendix E: AoI Analysis with MeteoInfo

## LIST OF TABLES

|                                                                  | Page |
|------------------------------------------------------------------|------|
| Table 2-1. IMPROVE Monitors in the VISTAS_12 Domain .....        | 6    |
| Table 2-2. Revised Facility Emissions for the AoI Analysis ..... | 10   |

## LIST OF FIGURES

|                                                                                                                                                                                                                                                                   | Page |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| Figure 2-1. IMPROVE Monitor Locations and Starting Points for HYSPLIT Trajectories in the VISTAS 12km Domain .....                                                                                                                                                | 7    |
| Figure 2-2. IMPROVE Monitor Locations and Starting Points for HYSPLIT Trajectories in the VISTAS States .....                                                                                                                                                     | 8    |
| Figure 2-3. NAM 12km (Sigma-Pressure Hybrid) Modeling Domain .....                                                                                                                                                                                                | 13   |
| Figure 2-4. Example Trajectory Plot, 100m Trajectories by Year, for Great Smoky Mountain National Park. (Created with R) .....                                                                                                                                    | 15   |
| Figure 2-5. Example Trajectory Plot, 100m Trajectories by Season, for Great Smoky Mountain National Park. (Created with R) .....                                                                                                                                  | 16   |
| Figure 2-6. Example Trajectory Plot for Great Smoky Mountain National Park .....                                                                                                                                                                                  | 17   |
| Figure 2-7. Example Residence Time, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. (Full view (top), Class I zoom (bottom); Plot was Created with R) .....                                                                             | 19   |
| Figure 2-8. Example Residence Time, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park, zoomed in on Class I area. (Full view (top), Class I zoom (bottom); Plot was Created with GIS) .....                                                | 20   |
| Figure 2-9. Example EWRT for Ammonium Nitrate Extinction, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS) .....                                 | 23   |
| Figure 2-10. Example EWRT, as a percentage of the total, for Ammonium Nitrate Extinction, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS) ..... | 24   |
| Figure 2-11. Example EWRT for Ammonium Sulfate Extinction, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS) .....                                | 25   |

## LIST OF FIGURES (Continued)

|                                                                                                                                                                                                                                                                  | <b>Page</b> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| Figure 2-12. Example EWRT, as a percentage of the total, for Ammonium Sulfate Extinction, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS)..... | 26          |
| Figure 2-13. Example EWRT*(Q/d) for Ammonium Sulfate Extinction for 2028, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS).....                 | 28          |

### Abbreviations/Acronym List

|                   |                                                                                                     |
|-------------------|-----------------------------------------------------------------------------------------------------|
| AoI               | Area of Influence                                                                                   |
| ARL               | Air Resources Laboratory                                                                            |
| bext              | Beta extinction                                                                                     |
| CIRA              | Cooperative Institute for Research in the Atmosphere                                                |
| CWT               | Concentration weighted trajectory                                                                   |
| d                 | Distance                                                                                            |
| EIS               | Emission Inventory System                                                                           |
| EPA               | United States Environmental Protection Agency                                                       |
| ERG               | Eastern Research Group, Inc.                                                                        |
| EWRT              | Extinction-weighted residence time                                                                  |
| FIPS              | Federal Information Processing Standard                                                             |
| FLM               | Federal Land Manager                                                                                |
| FR                | Federal Register                                                                                    |
| FTP               | File Transfer Protocol                                                                              |
| GIS               | Geographic Information System                                                                       |
| GUI               | Graphical User Interface                                                                            |
| HYSPLIT           | Hybrid Single Particle Lagrangian Integrated Trajectory                                             |
| IMPROVE           | Interagency Monitoring of Protected Visual Environments                                             |
| km                | Kilometers                                                                                          |
| m                 | Meters                                                                                              |
| NAAQS             | National Ambient Air Quality Standards                                                              |
| NAICS             | North American Industry Classification System                                                       |
| NAM-12            | North American Mesoscale forecast data at the 12-km level                                           |
| NOAA              | National Oceanic and Atmospheric Administration                                                     |
| NO <sub>x</sub>   | Oxides of nitrogen                                                                                  |
| PM <sub>2.5</sub> | Fine particle; primary particulate matter less than or equal to 2.5 microns in aerodynamic diameter |
| PSAT              | Particulate Source Apportionment Technology                                                         |
| Q                 | Emissions                                                                                           |
| Q/d               | Emissions over distance                                                                             |
| Q/d <sup>2</sup>  | Emissions over distance squared                                                                     |
| QA                | Quality assurance                                                                                   |
| R                 | Programming software language called R                                                              |
| RHR               | Regional Haze Rule                                                                                  |
| RT                | Residence Time                                                                                      |
| SCC               | Source Classification Code                                                                          |
| SESARM            | Southeastern States Air Resource Managers, Inc.                                                     |
| SIA               | Second Improve Algorithm                                                                            |
| SIP               | State Implementation Plan                                                                           |
| SO <sub>2</sub>   | Sulfur dioxide                                                                                      |
| TB                | Terabytes                                                                                           |
| tpy               | tons per year                                                                                       |
| UTC               | Universal Time Coordinated                                                                          |
| VISTAS            | Visibility Improvement - State and Tribal Association of the Southeast                              |

**State Abbreviations**

|    |                |
|----|----------------|
| AL | Alabama        |
| AR | Arkansas       |
| FL | Florida        |
| GA | Georgia        |
| KY | Kentucky       |
| LA | Louisiana      |
| ME | Maine          |
| MI | Michigan       |
| MN | Minnesota      |
| MO | Missouri       |
| NC | North Carolina |
| NH | New Hampshire  |
| NJ | New Jersey     |
| NM | New Mexico     |
| OK | Oklahoma       |
| SC | South Carolina |
| TN | Tennessee      |
| TX | Texas          |
| VA | Virginia       |
| VT | Vermont        |
| WV | West Virginia  |

This page is intentionally blank.

## **1. INTRODUCTION**

Southeastern States Air Resource Managers, Inc. (SESARM) has been designated by the United States Environmental Protection Agency (EPA) as the entity responsible for coordinating regional haze evaluations for the ten Southeastern states of Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. The Eastern Band of Cherokee Indians and the Knox County, Tennessee local air pollution control agency are also participating agencies. These parties are collaborating through the Regional Planning Organization known as Visibility Improvement - State and Tribal Association of the Southeast (VISTAS) in the technical analyses and planning activities associated with visibility and related regional air quality issues. VISTAS analyses will support the VISTAS states in their responsibility to develop, adopt, and implement their State Implementation Plans (SIPs) for regional haze.

The state and local air pollution control agencies in the Southeast are mandated to protect human health and the environment from the impacts of air pollutants. They are responsible for air quality planning and management efforts including the evaluation, development, adoption, and implementation of strategies controlling and managing all criteria air pollutants (including fine particles and ozone) as well as regional haze. This project will focus on regional haze and regional haze precursor emissions. Control of regional haze precursor emissions will have the additional benefit of reducing certain criteria pollutants as well.

The 1999 Regional Haze Rule (RHR) identified 18 Class I Federal areas (national parks greater than 6,000 acres and wilderness areas greater than 5,000 acres) in the VISTAS region and required states to define long-term strategies to improve visibility in these areas. States were required to establish baseline visibility conditions for the period 2000-2004, natural visibility conditions in the absence of anthropogenic influences, and an expected rate of progress to reduce emissions and incrementally improve visibility to natural conditions by 2064. The original RHR required states to improve visibility on the 20% most impaired days and protect visibility on the 20% least impaired days.<sup>1</sup> The RHR requires states to evaluate progress toward visibility improvement goals every five years and submit revised SIPs every ten years.

---

<sup>1</sup> RHR summary data is available at: <http://vista.cira.colostate.edu/Improve/rhr-summary-data/>

EPA finalized revisions to various requirements of the RHR in January 2017 (82 FR 3078) that were designed to strengthen, streamline, and clarify certain aspects of the agency's regional haze program including:

- A. Strengthening the Federal Land Manager (FLM) consultation requirements to ensure that issues and concerns are brought forward early in the planning process.
- B. Updating the SIP submittal deadlines for the second planning period from July 31, 2018 to July 31, 2021 to ensure that they align where applicable with other state obligations under the Clean Air Act. The end date for the second planning period remains 2028; that is, the focus of state planning will be to establish reasonable progress goals for each Class I area against which progress will be measured during the second planning period. This extension will allow states to incorporate planning for other federal programs while conducting their regional haze planning. These other programs include: the Mercury and Air Toxics Standards, the 2010 1-hour sulfur dioxide (SO<sub>2</sub>) National Ambient Air Quality Standard (NAAQS); the 2012 annual fine particle (PM<sub>2.5</sub>) NAAQS; and the 2008 and 2015 ozone NAAQS.
- C. Adjusting interim progress report submission deadlines so that second and subsequent progress reports will be due by: January 31, 2025; July 31, 2033; and every ten years thereafter. This means that one progress report will be required midway through each planning period.
- D. Removing the requirement for progress reports to take the form of SIP revisions. States will be required to consult with FLMs and obtain public comment on their progress reports before submission to the EPA. EPA will be reviewing but not formally approving or disapproving these progress reports.

The RHR defines "clearest days" as the 20% of monitored days in a calendar year with the lowest deciview index values. "Most impaired days" are defined as the 20% of monitored days in a calendar year with the highest amounts of anthropogenic visibility impairment. The long-term strategy and the reasonable progress goals must provide for an improvement in visibility for the most impaired days and ensure no degradation in visibility for the clearest days since the baseline period.

To inform the long-term strategy for visibility improvement, states often use an Area of Influence (AoI) analysis to help identify the areas and sources most likely contributing to poor visibility in the Class I areas. The AoI analysis involves running a backward trajectory model to determine the origin of the air parcels affecting visibility. This information is then spatially combined with emissions data to determine the sources or source sectors most likely contributing to the pollutant emissions.

This report describes the AoI Analysis, starting with the technical approach, followed by a walkthrough of the programming code used to generate the images and analysis.

## 2. APPROACH

Under this task, ERG identifies the 20% most impaired days for each Class I area in the VISTAS\_12 modeling domain over the 2011-2016 period based on the IMPROVE monitoring website RHR summary of the 20% most-impaired visibility days.<sup>2</sup> Due to the presence of large SO<sub>2</sub> emission reductions during this six-year period, the AoI analysis was set up to look at: 1) each year individually; 2) two separate periods of 2011-2013 and 2014-2016; and 3) for all years combined.

The initial plan was to use the *R* programming language for all of the AoI analysis, which included running the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model, plotting trajectories, and weighing residence time calculations. However, as the work proceeded, it became apparent that *R* would not provide all the features needed to accomplish this task. Where possible, ERG proceeded with free and open source software and tools to keep the analysis process as reproducible and transparent as possible. As such, several phases of the analysis are completed using multiple methods to provide states options for additional work and analysis they may wish to conduct. The draw backs of each approach used are discussed separately in the subsections for each phase of the analysis.

### 2.1 Identification of Sites

In addition to identifying the areas influencing visibility in Class I areas inside SESARM states, this analysis identifies areas from SESARM states that might have an influence on

---

<sup>2</sup> <http://vista.cira.colostate.edu/Improve/rhr-summary-data/>

visibility in Class I areas outside the SESARM states. Analysis started by examining the eighty-three (83) Interagency Monitoring of Protected Visual Environments (IMPROVE) monitor positions in Class I areas within the VISTAS 12 kilometer (km) modeling domain. The final number of Class I areas was reduced from 83 to 45 due to:

- Exclusion of Central and Western States IMPROVE monitors. To simplify the potential number of trajectories to be generated, SESARM approved a North Carolina Division of Air Quality recommendation which demonstrated that trajectories originating in Class I areas in several western states do not pass over a SESARM state and could be excluded from further analysis. These states are: Kansas; Nebraska; North Dakota; South Dakota; Colorado; Wyoming; and Montana. With the exclusion of these states, the number of IMPROVE monitors in the VISTAS 12km domain reduces to thirty-eight (38).
- Class I Areas With No IMPROVE Monitor. There are six (6) Class I areas in the VISTAS 12km modeling domain that do not have an IMPROVE monitor. Of the six, three are located in the SESARM states (Wolf Island, GA; Joyce Kilmer-Slickrock, NC/TN; and Otter Creek, WV) and three are located outside the SESARM states (Carlsbad Caverns, NM; Pecos Wilderness Area, NM; and Presidential Range-Dry River Wilderness, NH). For these Class I areas, the trajectory origin will be the centroid of the Class I area. Visibility data will be based on an appropriate IMPROVE monitor, as previously determined by the FLMs. This increased the number of trajectory origins to forty-four (44).
- IMPROVE Monitor Relocation. Additionally, the IMPROVE monitor for the Lye Brook Wilderness, VT area moved in 2012. Separate trajectory analysis was completed for each monitoring location. This increased the number of trajectory origins to forty-five (45).

One final adjustment was made to the origin of trajectories for three Class I areas (Breton Island, LA; Guadalupe Mountains National Park, TX; and Isle Royale National Park, MI). For these areas, the IMPROVE monitor lies just outside the Class I area. Thus, the origins of the trajectories were moved to the centroid of the Class I area to better represent the area. This resulted in 36 trajectories originating from the Class I Area IMPROVE monitor (including both of the Lye Brook Wilderness monitors) and nine (9) originating from the Class I area centroids (6 with no monitor, plus three relocated for monitors outside the Class I areas).

Table 2-1 presents the trajectory origins for each Class I area, noting whether the origin is the centroid or IMPROVE monitor location. Figure 2-1 presents the location of the trajectory origins, while Figure 2-2 presents a closer view of the SESARM states. Both figures also denote which Class I areas have trajectories originating from IMPROVE monitors or centroids.

## 2.2 **Identification of Dates for Analysis**

To identify the 20% most impaired days for analysis, ERG downloaded the RHR Summary, “Daily Impairment Values Including Patched Values” (December 2018 version) dataset from the IMPROVE website.<sup>3</sup> IMPROVE is a collaborative association of state, tribal, and federal agencies, and international partners. The EPA is the primary funding source, with contracting and research support from the National Park Service. The Air Quality Group at the University of California, Davis is the central analytical laboratory, with ion analysis provided by Research Triangle Institute, and carbon analysis provided by Desert Research Institute. The RHR Summary data provides the means for the best, middle, and worst 20% visibility days. From this dataset, ERG was able to identify the 20% most impaired days for each Class I area in the VISTAS 12km modeling domain over the 2011-2016 period for AoI analysis. The RHR flags the data to indicate which group (i.e., best, worst). For consistency with the revised EPA guidance on tracking visibility<sup>4</sup>, days were ranked based on the anthropogenic contribution only. The RHR Summary provides this ranking in the “Impairment\_Group” column. Similarly, the anthropogenic portion of nitrate (anthro\_eamm\_no3) and sulfate (anthro\_eamm\_so4) were used for trajectory weighting described in Section 2.5.

In instances where all years are not available at an IMPROVE monitoring site, the 20% most impaired days from the year(s) available were analyzed. For example, Shining Rock Wilderness area does not have data for 2011, as a result, the trajectories for the AoI analysis only covered the 2012 to 2016 period.

---

<sup>3</sup> “Daily Impairment Values Including Patched Values” under the “Regional Haze Rule Summary data through 1988-2017 (posted December 2018)” section. Internet address: <http://vista.cira.colostate.edu/Improve/rhr-archived-data/>

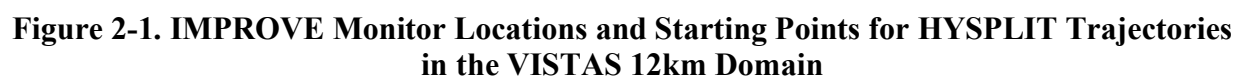
<sup>4</sup> US EPA, 2018. Technical Guidance on Tracking Visibility Progress for the Second Implementation Period of the Regional Haze Program. EPA-454/R-18-010. December 2018 available at: [https://www.epa.gov/sites/production/files/2018-12/documents/technical\\_guidance\\_tracking\\_visibility\\_progress.pdf](https://www.epa.gov/sites/production/files/2018-12/documents/technical_guidance_tracking_visibility_progress.pdf)

Table 2-1. IMPROVE Monitors in the VISTAS\_12 Domain

| Class I Area                               | State | FIPS Code | IMPROVE Site Code  | Trajectory Origin |           |                               |
|--------------------------------------------|-------|-----------|--------------------|-------------------|-----------|-------------------------------|
|                                            |       |           |                    | Latitude          | Longitude | Description                   |
| Acadia National Park                       | ME    | 23009     | ACAD1              | 44.3771           | -68.2610  | IMPROVE Monitor               |
| Bandelier Wilderness Area                  | NM    | 35028     | BAND1              | 35.7797           | -106.2664 | IMPROVE Monitor               |
| Big Bend NP                                | TX    | 48043     | BIBE1              | 29.3027           | -103.1780 | IMPROVE Monitor               |
| Bosque del Apache Wilderness Area          | NM    | 35053     | BOAP1              | 33.8695           | -106.8520 | IMPROVE Monitor               |
| Boundary Waters Canoe Area Wilderness Area | MN    | 27075     | BOWA1              | 47.9466           | -91.4955  | IMPROVE Monitor               |
| Breton Wilderness                          | LA    | 22075     | BRIS1              | 29.8895           | -88.8524  | Class I Centroid <sup>b</sup> |
| Brigantine Wilderness Area                 | NJ    | 34001     | BRIG1              | 39.4650           | -74.4492  | IMPROVE Monitor               |
| Caney Creek Wilderness Area                | AR    | 05113     | CACR1              | 34.4544           | -94.1429  | IMPROVE Monitor               |
| Cape Romain Wilderness                     | SC    | 45019     | ROMA1              | 32.9410           | -79.6572  | IMPROVE Monitor               |
| Carlsbad Caverns NP                        | NM    | 48109     | GUMO1 <sup>a</sup> | 32.1409           | -104.5529 | Class I Centroid              |
| Chassahowitzka Wilderness Area             | FL    | 12017     | CHAS1              | 28.7484           | -82.5549  | IMPROVE Monitor               |
| Cohotta Wilderness Area                    | GA    | 13213     | COHU1              | 34.7852           | -84.6265  | IMPROVE Monitor               |
| Dolly Sods Wilderness                      | WV    | 54093     | DOSO1              | 39.1053           | -79.4261  | IMPROVE Monitor               |
| Everglades NP                              | FL    | 12086     | EVER1              | 25.3910           | -80.6806  | IMPROVE Monitor               |
| Great Gulf Wilderness Area                 | NH    | 33007     | GRGU1              | 44.3082           | -71.2177  | IMPROVE Monitor               |
| Great Smoky Mountains NP                   | NC/TN | 47009     | GRSM1              | 35.6334           | -83.9416  | IMPROVE Monitor               |
| Guadalupe Mountain NP                      | TX    | 48109     | GUMO1              | 31.9236           | -104.8846 | Class I Centroid <sup>b</sup> |
| Hercules-Glades Wilderness Area            | MO    | 29213     | HEGL1              | 36.6138           | -92.9221  | IMPROVE Monitor               |
| Isle Royale NP                             | MI    | 26083     | ISLE1              | 48.0109           | -88.8284  | Class I Centroid <sup>b</sup> |
| James River Face Wilderness                | VA    | 51163     | JARI1              | 37.6266           | -79.5125  | IMPROVE Monitor               |
| Joyce Kilmer-Slickrock Wilderness          | NC/TN | 47009     | GRSM1 <sup>a</sup> | 35.4047           | -83.9762  | Class I Centroid              |
| Linville Gorge Wilderness Area             | NC    | 37011     | LIGO1              | 35.9723           | -81.9331  | IMPROVE Monitor               |
| Lye Brook Wilderness                       | VT    | 50003     | LYBR1              | 43.1482           | -73.1268  | IMPROVE Monitor               |
|                                            |       | 50025     | LYEB1              | 42.9561           | -72.9098  | IMPROVE Monitor               |
| Mammoth Cave NP                            | KY    | 21061     | MACA1              | 37.1318           | -86.1479  | IMPROVE Monitor               |
| Mingo Wilderness Area                      | MO    | 29207     | MING1              | 36.9717           | -90.1432  | IMPROVE Monitor               |
| Moosehorn Wilderness EDM                   | ME    | 23029     | MOOS1              | 44.8362           | -67.2276  | IMPROVE Monitor               |
| Okefenokee Wilderness Area                 | GA    | 13049     | OKEF1              | 30.7405           | -82.1283  | IMPROVE Monitor               |
| Otter Creek Wilderness                     | WV    | 54093     | DOSO1 <sup>a</sup> | 38.9969           | -79.6460  | Class I Centroid              |
| Pecos Wilderness Area                      | NM    | 35055     | WHPE1 <sup>a</sup> | 35.8944           | -105.6453 | Class I Centroid              |
| Presidential Range-Dry River Wilderness    | NH    | 33007     | GRGU1 <sup>a</sup> | 44.1775           | -71.3226  | Class I Centroid              |
| Salt Creek Wilderness Area                 | NM    | 35005     | SACR1              | 33.4598           | -104.4042 | IMPROVE Monitor               |
| San Pedro Parks Wilderness Area            | NM    | 35039     | SAPE1              | 36.0139           | -106.8447 | IMPROVE Monitor               |
| Seney Wilderness Area                      | MI    | 26153     | SENE1              | 46.2889           | -85.9503  | IMPROVE Monitor               |
| Shenandoah NP                              | VA    | 51113     | SHEN1              | 38.5229           | -78.4348  | IMPROVE Monitor               |
| Shining Rock Wilderness Area               | NC    | 37087     | SHRO1              | 35.3937           | -82.7744  | IMPROVE Monitor               |
| Sipsey Wilderness Area                     | AL    | 01079     | SIPS1              | 34.3433           | -87.3388  | IMPROVE Monitor               |
| St Marks Wilderness Area                   | FL    | 12129     | SAMA1              | 30.0926           | -84.1614  | IMPROVE Monitor               |
| Swanquarter Wilderness Area                | NC    | 37095     | SWAN1              | 35.4510           | -76.2075  | IMPROVE Monitor               |
| Upper Buffalo Wilderness Area              | AR    | 05101     | UPBU1              | 35.8258           | -93.2030  | IMPROVE Monitor               |
| Voyageurs National Park                    | MN    | 27137     | VOYA2              | 48.4126           | -92.8286  | IMPROVE Monitor               |
| Wheeler Peak Wilderness                    | NM    | 35055     | WHPE1              | 36.5854           | -105.452  | IMPROVE Monitor               |
| White Mountain Wilderness Area             | NM    | 35027     | WHIT1              | 33.4687           | -105.5349 | IMPROVE Monitor               |
| Wichita Mountains Wilderness               | OK    | 40031     | WIMO1              | 34.7323           | -98.7130  | IMPROVE Monitor               |
| Wolf Island Wilderness                     | GA    | 13049     | OKEF1 <sup>a</sup> | 31.3451           | -81.3058  | Class I Centroid              |

<sup>a</sup> No IMPROVE Monitor is located at this Class I Area. As such, the nearby IMPROVE monitor was used as a surrogate.

<sup>b</sup> This Class I Area does have a dedicated IMPROVE Monitor. However, it is located outside the Class I Area boundary. The trajectory origin was altered to the Class I Area centroid.





**Figure 2-2. IMPROVE Monitor Locations and Starting Points for HYSPLIT Trajectories in the VISTAS States**

### 2.2.1 Emissions Data Collection

The 2028 emissions data used in the AoI analysis came from the Task 2A emissions inventory, which was prepared in August 2018, and was used for the Particulate Source Apportionment Technology (PSAT) tagging. Emissions were split into two distinct groups, by source category: Point and county-level (e.g., nonpoint, onroad, nonroad, prescribed fires, and total point emissions).

Oxides of nitrogen (NO<sub>x</sub>) and SO<sub>2</sub> point source emissions were aggregated at the facility level. The aggregated emissions also included facility latitude and longitude,<sup>5</sup> which was used to calculate the distance between the point source and Class I area. The file also included additional information including Emission Inventory System (EIS) Facility ID, Facility Name, primary North American Industry Classification System (NAICS) code, Tier 1 description, and source classification codes (SCCs) (i.e., all SCCs present at the facility). After initial review of the data, the SESARM partners provided a list of revised facility emissions based on updated information. This included facilities that are scheduled to cease operations (in whole or part) before 2028, or other emissions adjustments that were not included in the original emissions inventory effort. Table 2-2 lists the facilities whose emissions were revised for the AoI analysis.

For source categories within the nonpoint, onroad, nonroad, and prescribed fire sectors,<sup>6</sup> the emissions were calculated at the county-level by summing the emission inventory databases by the state and county FIPS. Similarly, the point emissions were summed by the state and county FIPS reported for each facility. The point source files included airport and other nonroad emission sources reported to a specific latitude and longitude. These sources were separated into two distinct point source categories: 1) NONROAD\_MAR for marine, air, and railroad source, and 2) NONROAD\_OTHER for those sources with traditional nonroad SCC descriptions (e.g., construction). This was done to better separate the nonroad sources, in which NONROAD\_MAR might require different federal actions for controls, as opposed to NONROAD\_OTHER.

---

<sup>5</sup> Emission estimates are often provided at the facility-level (often a facility-centroid), but can also be reported for separate emission release points. In cases where the facility-level coordinates are not reported, the latitude and longitude coordinates by facility emission release points are averaged.

<sup>6</sup> Note that agricultural fires are included in the nonpoint inventory. Additionally, wildfires are not included in this analysis to be consistent with the Second Improve Algorithm (SIA) Impairment data, which excludes the impact of wildfires.

Table 2-2. Revised Facility Emissions for the AoI Analysis

| FIPS Code | State | EIS Facility ID | Facility Name                                                  | Reason for 2028 Emissions Adjustment | NO <sub>x</sub> 2028 (tpy) | NO <sub>x</sub> 2028 Revised (tpy) | SO <sub>2</sub> 2028 (tpy) | SO <sub>2</sub> 2028 Revised (tpy) |
|-----------|-------|-----------------|----------------------------------------------------------------|--------------------------------------|----------------------------|------------------------------------|----------------------------|------------------------------------|
| 01053     | AL    | 7440211         | Escambia Operating Company LLC                                 | Permit Limit                         | 349.3                      | 349.3 <sup>a</sup>                 | 18,974.4                   | 7,963.0                            |
| 01053     | AL    | 985111          | Escambia Operating Company LLC                                 | Facility Shutdown                    | 149.6                      | 0.0                                | 8,589.6                    | 0.0                                |
| 01127     | AL    | 7917311         | Alabama Power - Gorgas                                         | Facility Shutdown                    | 3,976.4                    | 0.0                                | 1,410.9                    | 0.0                                |
| 01129     | AL    | 1028611         | PowerSouth Energy Coop - Lowman                                | Partial Shutdown                     | 2,910.8                    | 300.0                              | 3,805.2                    | 0.0                                |
| 05059     | AR    | 658911          | Lake Catherine                                                 | Facility Shutdown                    | 125.0                      | 0.0                                | 0.4                        | 0.0                                |
| 05069     | AR    | 893911          | White Bluff                                                    | Facility Shutdown                    | 16,179.2                   | 0.0                                | 31,997.1                   | 0.0                                |
| 12047     | FL    | 769711          | White Springs Agricultural Chemicals, Inc                      | Permit Limit                         | 112.4                      | 112.4 <sup>a</sup>                 | 3,197.8                    | 2,745.0                            |
| 12057     | FL    | 716411          | Mosaic Fertilizer, LLLC                                        | Permit Limit                         | 159.7                      | 159.7 <sup>a</sup>                 | 3,034.1                    | 1,890.0                            |
| 12105     | FL    | 717711          | Mosaic Fertilizer LLC                                          | Permit Limit                         | 310.4                      | 310.4 <sup>a</sup>                 | 7,900.7                    | 3,581.0                            |
| 12105     | FL    | 919811          | Mosaic Fertilizer, LLC                                         | Permit Limit                         | 141.0                      | 141.0 <sup>a</sup>                 | 4,425.6                    | 3,614.0                            |
| 13103     | GA    | 3711211         | GA Power Co PLT McIntosh                                       | Facility Shutdown                    | 447.1                      | 0.0                                | 127.3                      | 0.0                                |
| 13115     | GA    | 3713211         | GA Power Company - Plant Hammond                               | Facility Shutdown                    | 864.9                      | 0.0                                | 772.5                      | 0.0                                |
| 21059     | KY    | 5891711         | Owensboro Municipal Utilities - Elmer Smith Station            | Facility Shutdown                    | 23.4                       | 0.0                                | 6.7                        | 0.0                                |
| 21177     | KY    | 5196711         | Tennessee Valley Authority - Paradise Fossil Plant             | Facility Shutdown                    | 2,927.4                    | 0.0                                | 2,990.2                    | 0.0                                |
| 24001     | MD    | 7763811         | Luke Paper Company                                             | Permit Limit                         | 3,607.0                    | 3,607.0 <sup>a</sup>               | 22,660.0                   | 9,876.0                            |
| 28059     | MS    | 8232011         | Mississippi Phosphates Corporation                             | Facility Shutdown                    | 325.7                      | 0.0                                | 1,330.6                    | 0.0                                |
| 28073     | MS    | 7154411         | South Mississippi Electric Power Association, R D Morrow Plant | Fuel Switch                          | 4,219.3                    | 652.6                              | 3,827.5                    | 101.4                              |
| 28121     | MS    | 7288911         | Pursue Energy Corporation Thomasville Gas Plant                | Facility Shutdown                    | 3.9                        | 0.0                                | 8,933.5                    | 0.0                                |
| 28141     | MS    | 17942211        | Mississippi Silicon <sup>b</sup>                               | New Source                           | 0.0                        | 836.0                              | 0.0                        | 648.0                              |
| 31055     | NE    | 6732411         | Omaha Public Power District - North Omaha Power Station        | Units Shutdown                       | 6,961.2                    | 50.0                               | 14,530.0                   | 5.0                                |
| 39081     | OH    | 8190811         | W. H. Sammis Plant (0641160017)                                | Facility Shutdown                    | 3,740.0                    | 0.0                                | 3,184.0                    | 0.0                                |
| 42007     | PA    | 3853711         | FIRSTENERGY GEN LLC/Bruce Mansfield PLT                        | Facility Shutdown                    | 10,707.0                   | 0.0                                | 19,074.0                   | 0.0                                |
| 47001     | TN    | 6196011         | TVA Bull Run Fossil Plant                                      | Facility Shutdown                    | 964.2                      | 0.0                                | 622.5                      | 0.0                                |
| 47105     | TN    | 4129211         | Tate & Lyle, Loudon                                            | Fuel Switch                          | 883.3                      | 252.5                              | 472.8                      | 110.2                              |
| 48161     | TX    | Full_3497       | Big Brown                                                      | Facility Shutdown                    | 4,407.0                    | 0.0                                | 52,307.0                   | 0.0                                |

Table 2-2. Revised Facility Emissions for the AoI Analysis

| FIPS Code             | State | EIS Facility ID | Facility Name                 | Reason for 2028 Emissions Adjustment | NO <sub>x</sub> 2028 (tpy) | NO <sub>x</sub> 2028 Revised (tpy) | SO <sub>2</sub> 2028 (tpy) | SO <sub>2</sub> 2028 Revised (tpy) |
|-----------------------|-------|-----------------|-------------------------------|--------------------------------------|----------------------------|------------------------------------|----------------------------|------------------------------------|
| 48331                 | TX    | 13408411        | Sandow 5 Generating Plant     | Facility Shutdown                    | 0.8                        | 0.0                                | 0.0                        | 0.0                                |
| 48331                 | TX    | 4204811         | Sandow Steam Electric Station | Facility Shutdown                    | 1,055.0                    | 0.0                                | 20,013.0                   | 0.0                                |
| 48331                 | TX    | Full_52071      | Sandow 5                      | Facility Shutdown                    | 1,042.0                    | 0.0                                | 1,255.0                    | 0.0                                |
| 48449                 | TX    | Full_6147       | Monticello                    | Facility Shutdown                    | 7,199.0                    | 0.0                                | 44,287.0                   | 0.0                                |
| 48487                 | TX    | 7927311         | Oklaunion Power Station       | Facility Shutdown                    | 5,513.0                    | 0.0                                | 1,679.0                    | 0.0                                |
| <b>Revised Totals</b> |       |                 |                               |                                      | <b>79,305.0</b>            | <b>6,771.0</b>                     | <b>281,408.4</b>           | <b>30,533.6</b>                    |

<sup>a</sup> No emission changes.

<sup>b</sup> New facility added for this analysis, and not included in the final Task 2A point source emissions inventory.

## 2.3 **Trajectories**

For this study, the HYSPLIT model<sup>7</sup> developed by the National Oceanic and Atmospheric Administration's (NOAA) Air Resources Laboratory (ARL) was used to identify areas most likely influencing visibility.

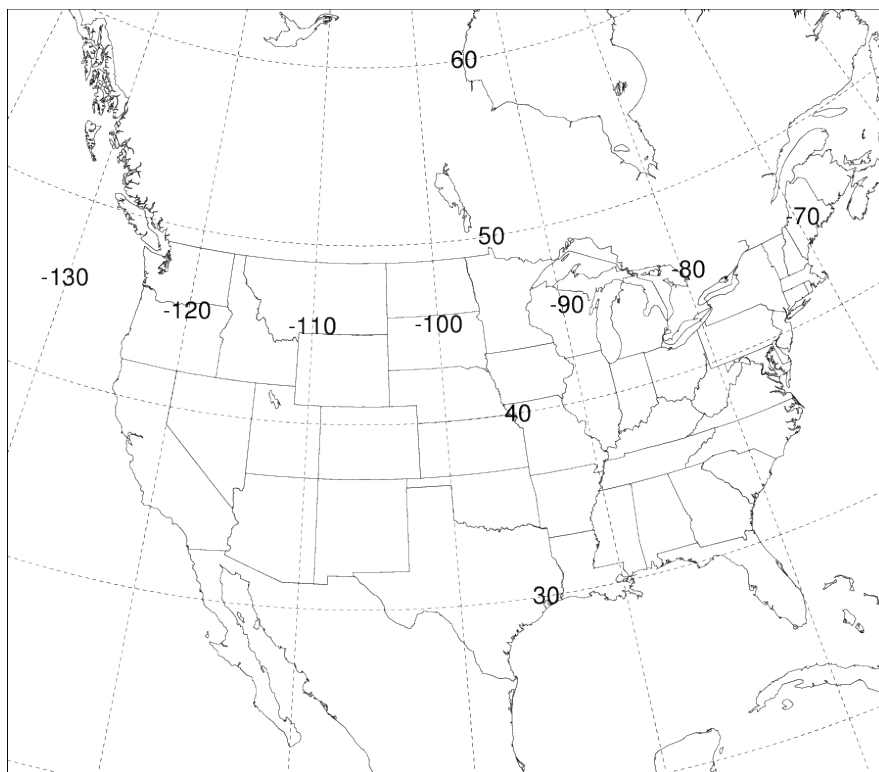
### 2.3.1 ***Meteorological Data***

The analysis started by collecting the meteorological and air quality data needed to complete the task. For the meteorology data, ERG downloaded the North American Mesoscale forecast data at the 12-km level (NAM-12) hybrid sigma-pressure data from NOAA ARL File Transfer Protocol (FTP) Server (<ftp://arlftp.arlhq.noaa.gov/nams>). The extent of the domain is presented in Figure 2-3.

The meteorological files include 40 pressure levels in the vertical, and output for every hour of the day. The files available from NOAA ARL are in a HYSPLIT ready format for individual dates starting in March 2010. ERG downloaded December 29, 2010 through December 31, 2016 for use in the back trajectories. Due to the total size of all the meteorology files (2.23 TB), these data were provided to the SESARM states via hard drive transfer.

---

<sup>7</sup> Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J. B., Cohen, M. D., and Ngan, F., (2015). NOAA's HYSPLIT atmospheric transport and dispersion modeling system, Bull. Amer. Meteor. Soc., 96, 2059-2077, <http://dx.doi.org/10.1175/BAMS-D-14-00110.1>



**Figure 2-3. NAM 12km (Sigma-Pressure Hybrid) Modeling Domain**

### **2.3.2 Trajectory Set up**

The HYSPLIT runs included starting heights of 100 meters (m), 500 m, 1,000 m, and 1,500 m. Trajectories were run 72 hours backwards in time for each height at each location.

Trajectories were run with start times of 12AM (midnight of the start of the day), 6AM, 12PM, 6PM, and 12AM (midnight at the end of the day) local time. ERG converted these times from local time of the Class I area to Universal Time Coordinated (UTC), adjusting the data as necessary, and accounting for the shift from daylight saving time to standard time.

As mentioned in Section 2.0, the initial plan was to run the trajectories from within the *R* programming software. While several packages exist to run HYSPLIT from *R* (e.g., SplitR, opentraj), all were designed to work with reanalysis data<sup>8</sup> and not daily meteorological data files, and the daily NAM 12km hybrid was not an option for the meteorological files in any of the

---

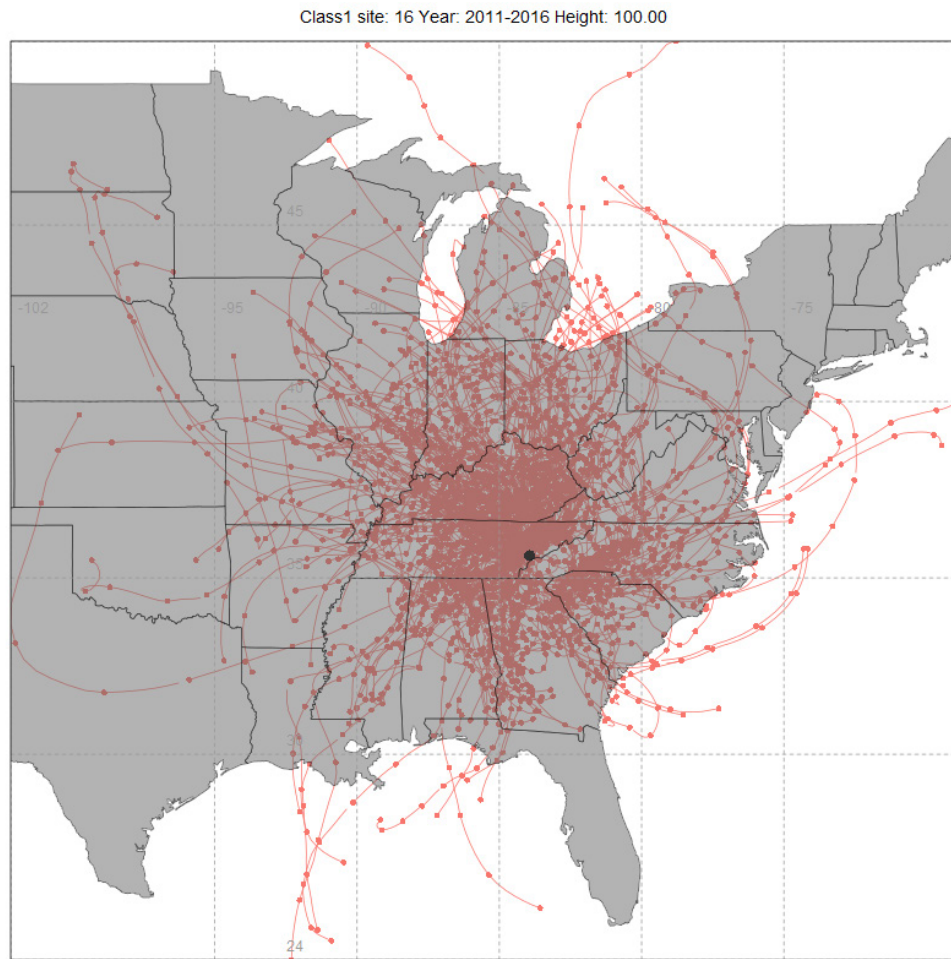
<sup>8</sup> The reanalysis data is the NOAA data reformatted for HYSPLIT.

packages available. In lieu of running HYSPLIT with *R*, ERG developed a python script that created all the necessary control files to run HYSPLIT. HYSPLIT was then run outside its Graphical User Interface (GUI) via a batch script. The python code and batch script used to run HYSPLIT are available in Appendices A and B, respectively, and include instruction on how to run them. This scripted method was necessary for running the 148,468 (37,117 start times modeled, at 4 different heights) trajectories for each site for the AoI analysis.

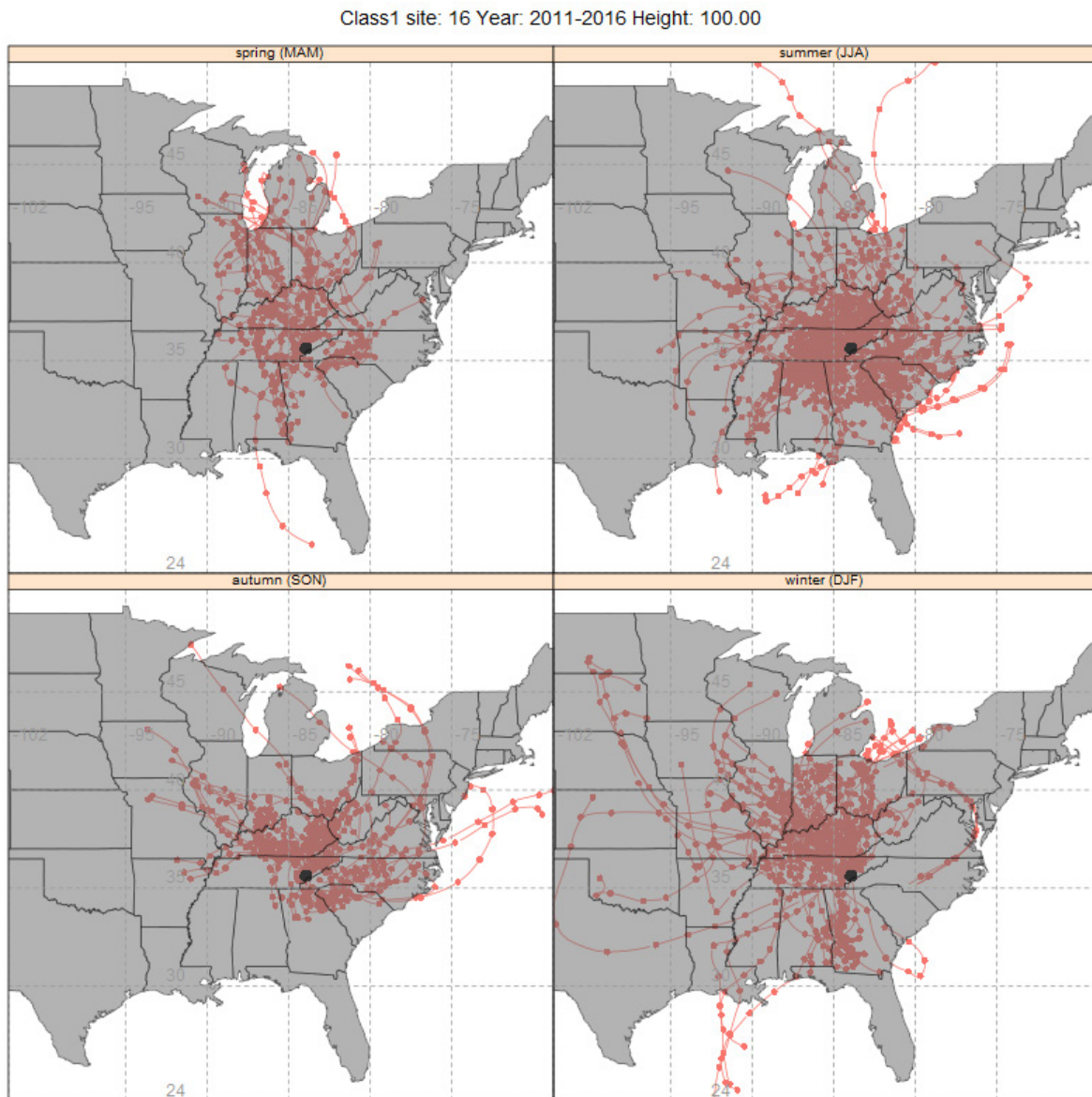
*R* was used to combine the HYSPLIT output files (tdump files) into a single file. Plots of the trajectories by height by year (Figure 2-4) and height by season (Figure 2-5) were generated using *R* (via the *openair*<sup>9</sup> package). The *R* code used to generate the plots can be found in Appendix C. Additional plots can also be generated by modifying the “type” variable.

---

<sup>9</sup> Carslaw, D.C. and Ropkins, K. (2012). “openair — An *R* package for air quality data analysis.” *Environmental Modelling & Software*, 27–28(0), pp. 52–61. ISSN 1364-8152, doi: 10.1016/j.envsoft.2011.09.008.



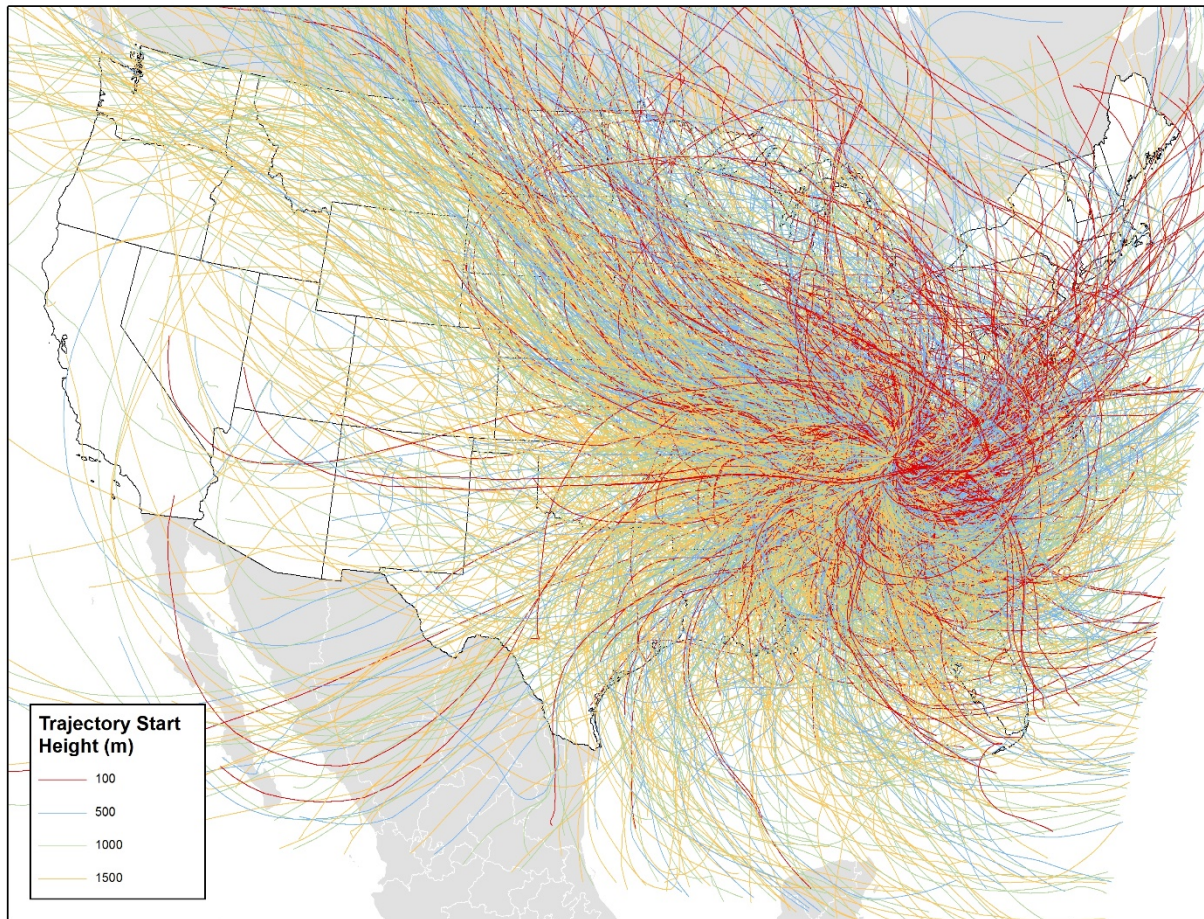
**Figure 2-4. Example Trajectory Plot, 100m Trajectories by Year, for Great Smoky Mountain National Park. (Created with R)**



**Figure 2-5. Example Trajectory Plot, 100m Trajectories by Season, for Great Smoky Mountain National Park. (Created with R)**

The files of the trajectory endpoints for each starting height were also converted to ArcGIS and plotted using ArcGIS (Figure 2-6), and all shapefile and base layers were delivered to SESARM.

These back trajectories for the 20% most impaired days were then used to develop residence time (RT) plots.



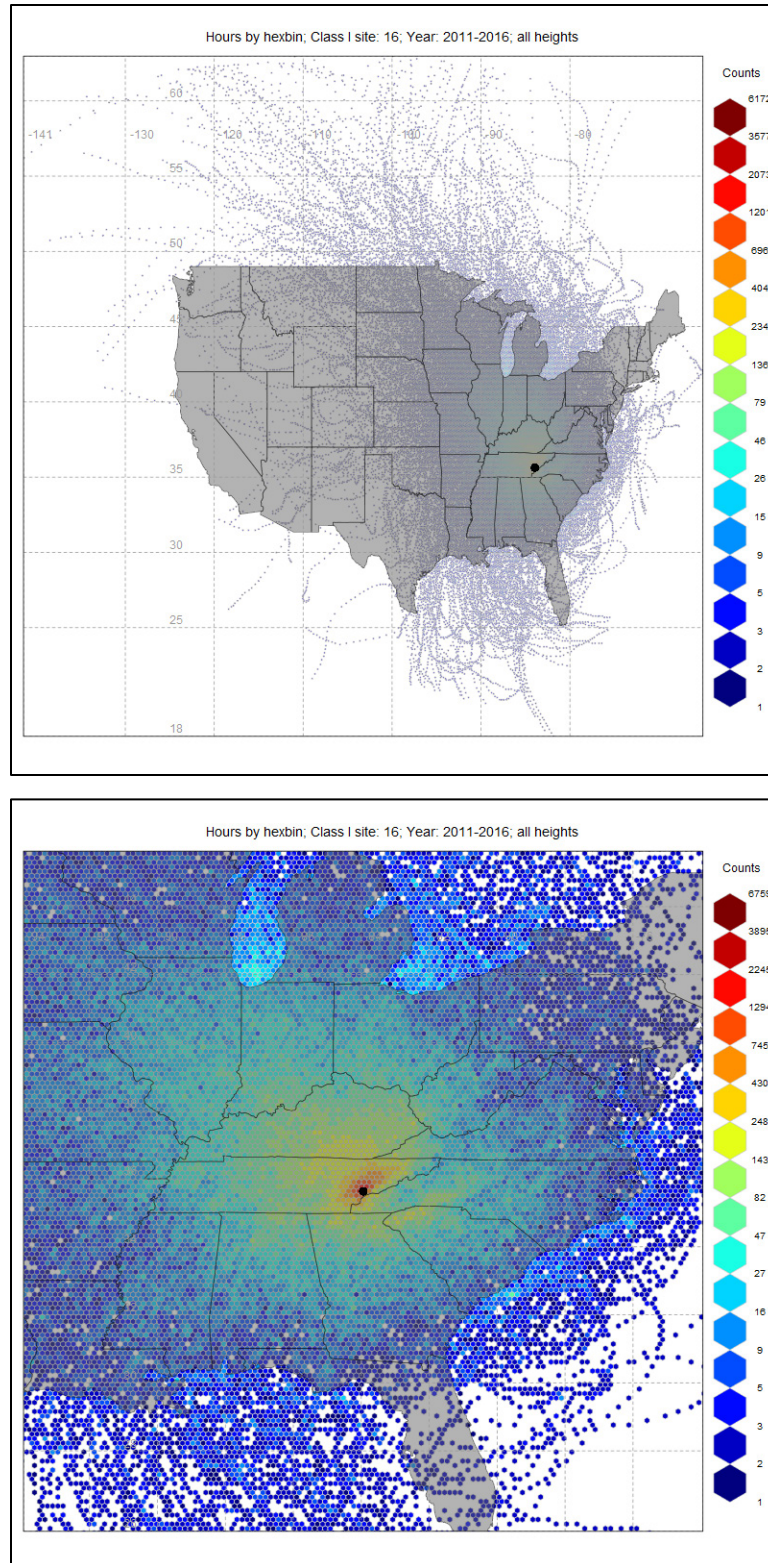
**Figure 2-6. Example Trajectory Plot for Great Smoky Mountain National Park. Plot was Created with GIS, for the 20% Most Impaired Days.**

## **2.4 Residence Time**

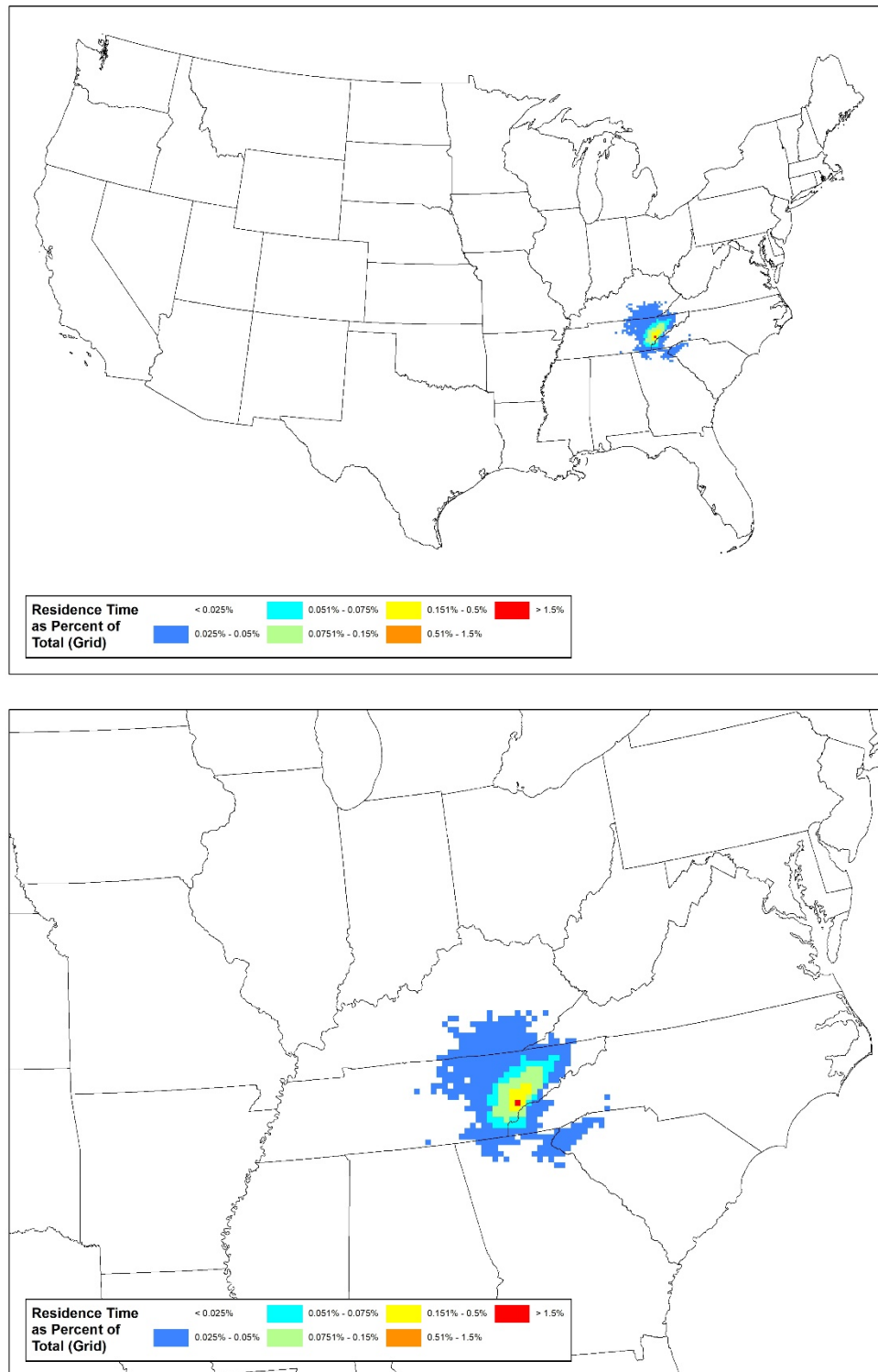
The RT plots define the geographic areas with the highest probability of influencing the monitor on the 20% most impaired visibility days. The RTs for point sources were calculated as the number of trajectory hours that pass through each of the 12km modeling domain grid cell (i.e., the number of hourly trajectory end points in each cell) using *R* (Appendix C) and custom geographic information system (GIS) scripts (Appendix D). The analysis was expanded to the EPA's 12US2 domain, which cover the entire continental United States, from the VISTAS 12km domain. This ensured the analysis for western Class I areas took into account their entire source region, and did not artificially weight results on sources within the VISTAS 12km domain only.

The *R* script was used to look at the various breaks of the data to review for any outlying years, or if emission controls installed midway through the review period affected, extinction-weighted residence time (EWRT). Figure 2-7 shows an example RT plot produced by the *R* code.

ArcGIS was used to develop final RT and EWRT values that were exported to shapefiles and joined with the emissions data for the AoI analysis spreadsheets (Section 2.6). Figure 2-8 shows a similar residence time plot as the *R* files, as plotted in GIS version, with a wide view of the modeling domain at the top, and a close up of the Class I area at the bottom. RT was also calculated for the modeling grid using a custom GIS script, described in Appendix E, for comparison and quality assurance (QA) purposes.



**Figure 2-7. Example Residence Time, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. (Full view (top), Class I zoom (bottom)); Plot was Created with R)**



**Figure 2-8. Example Residence Time, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park, zoomed in on Class I area. (Full view (top), Class I zoom (bottom); Plot was Created with GIS)**

The RTs for county-level source sectors were calculated as the number of trajectory hours that passed through the county, as opposed to grid cells. This was completed in GIS using a custom script, provided in Appendix E.

## 2.5 Extinction-Weighted Residence Time

The trajectory residence time was also weighted by ammonium sulfate and ammonium nitrate and used to produce separate sulfate and nitrate EWRT plots. This allows separate analysis for sulfate and nitrate that is weighted toward the days influenced most by those constituents and not days most influenced by other constituents, like organic carbon.

In this project, the Concentration Weighted Trajectory (CWT)<sup>10</sup> approach was used to develop the EWRT, substituting the extinction values for the concentration. The extinction attributable to each pollutant is paired with the trajectory for that day. The mean weighted extinction of the pollutant species for each grid cell is calculated according to:

$$\bar{E}_{ij} = EWRT = \frac{1}{\sum_{k=1}^N \tau_{ijk}} \sum_{k=1}^N (b_{ext_k}) \tau_{ijk}$$

Where:

$i$  and  $j$  are the indices of grid

$k$  the index of trajectory

$N$  the total number of trajectories used in analysis

$b_{extk}$  is the 24-hour extinction attributed to the pollutant measured upon arrival of trajectory  $k$ , and

$\tau_{ijk}$  the number of trajectory hours that pass through each grid cell ( $i, j$ ) (where “ $i$ ” is the row and “ $j$ ” is the column).<sup>11</sup>

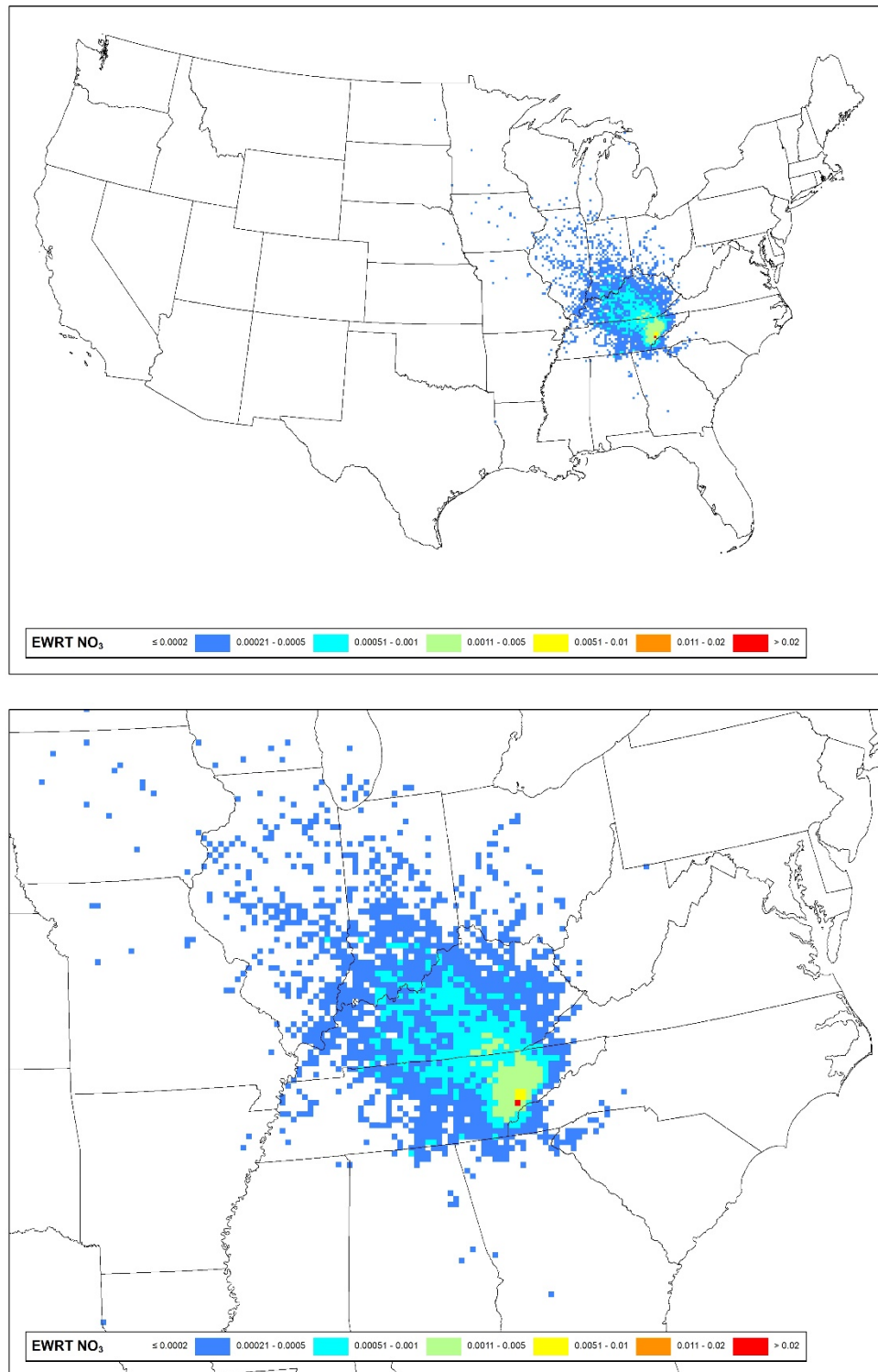
<sup>10</sup> Hsu, Y.-K., T. M. Holsen and P. K. Hopke (2003). “Comparison of hybrid receptor models to locate PCB sources in Chicago”. In: Atmospheric Environment 37.4, pp. 545–562. DOI: 10.1016/S1352-2310(02)00886-5

<sup>11</sup> Carslaw, D.C. (2015). The openair manual — open-source tools for analyzing air pollution data. Manual for Version 1.1-4, King’s College London. [http://www.openair-project.org/PDF/OpenAir\\_Manual.pdf](http://www.openair-project.org/PDF/OpenAir_Manual.pdf)

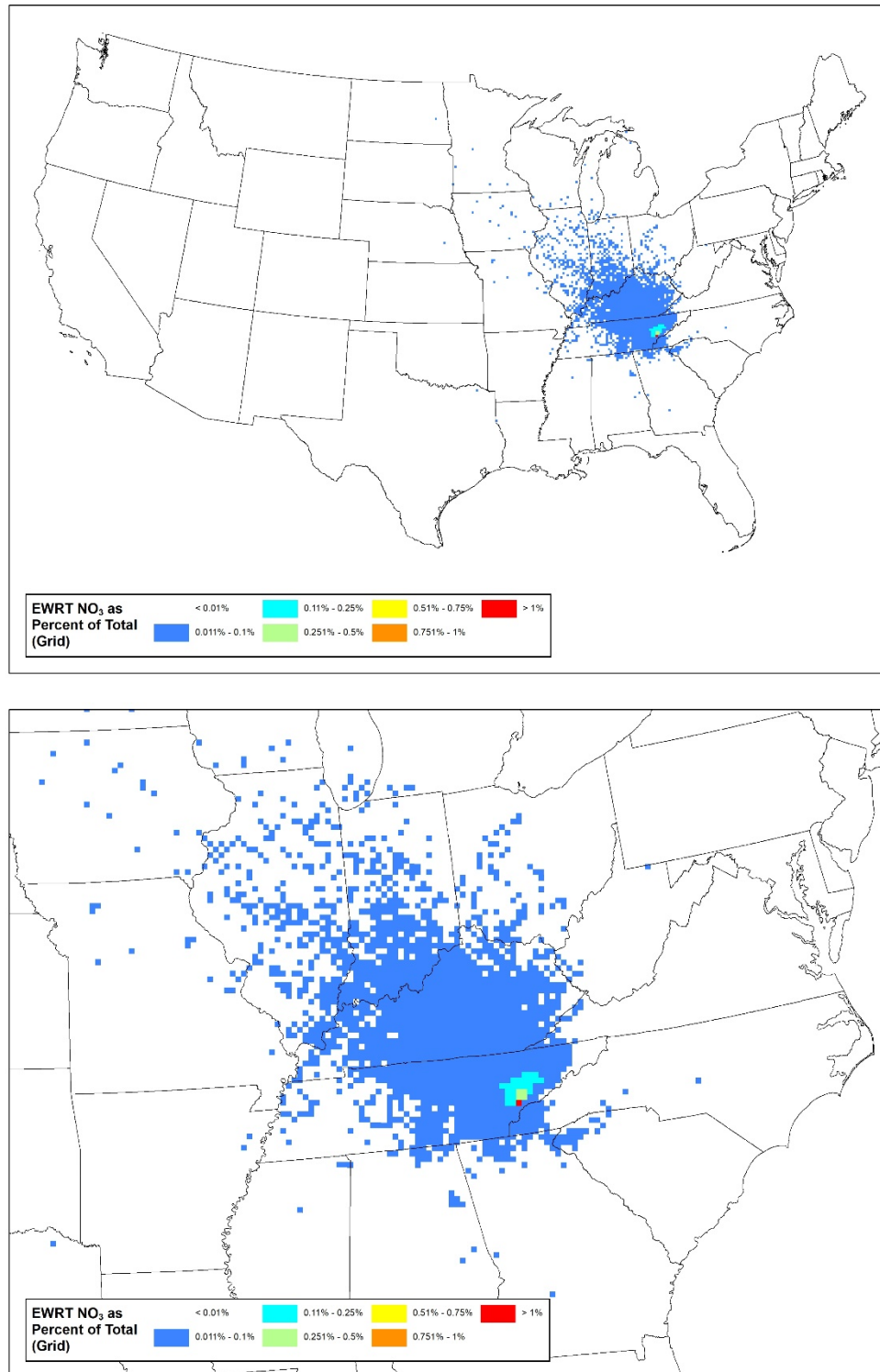
The higher the value of the EWRT ( $\bar{E}_{ij}$ ), the more likely that the air parcels passing over cell (i, j) would cause higher extinction at the receptor site for that light extinction species. Since this method uses the extinction value for weighting, trajectories passing over large sources are more discernible from those passing over moderate sources.

Figure 2-9 presents the EWRT for ammonium nitrate extinction for the Great Smoky Mountain National Park for the 20% most impaired days from 2011 to 2016. The figure provides a view of the full modeling domain (top) and an image focused on the Class I Area (bottom). Figure 2-10 show the EWRT for ammonium nitrate extinction as a percentage of the total ammonium nitrate, with a full domain view and Class I area view. Figure 2-11 shows the weighted ammonium sulfate EWRT, and Figure 2-12 shows the weighted ammonium sulfate EWRT as a percent of the total ammonium sulfate.

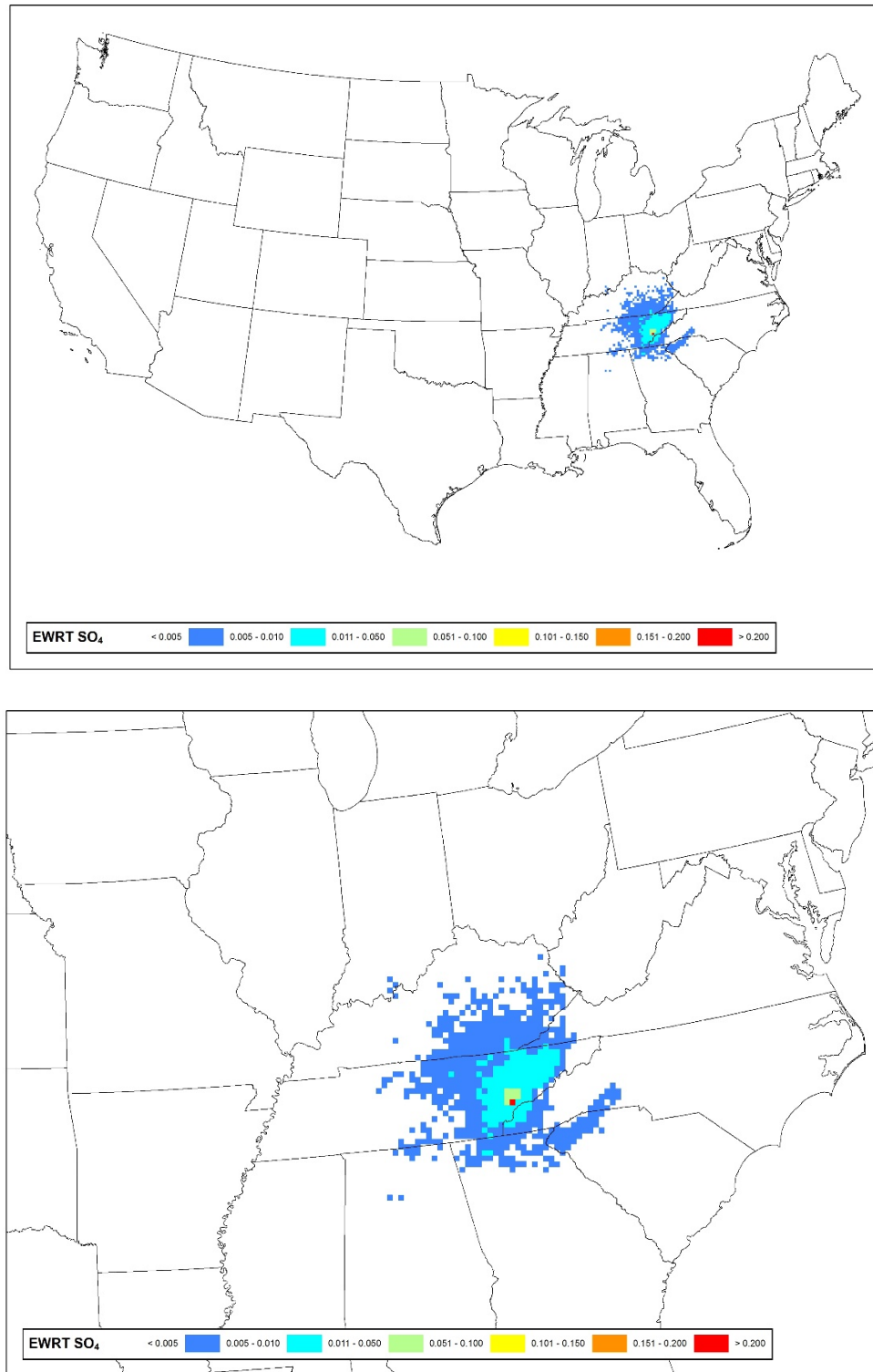
EWRT were performed using *R* (Appendix C) and custom GIS scripts (Appendix D). The *R* calculation process is memory intensive and was only run at a slightly coarser grid resolution for initial analysis to determine if any time period displayed a significantly different EWRT pattern. The GIS produced EWRT values were used to develop the spreadsheet of source contribution.



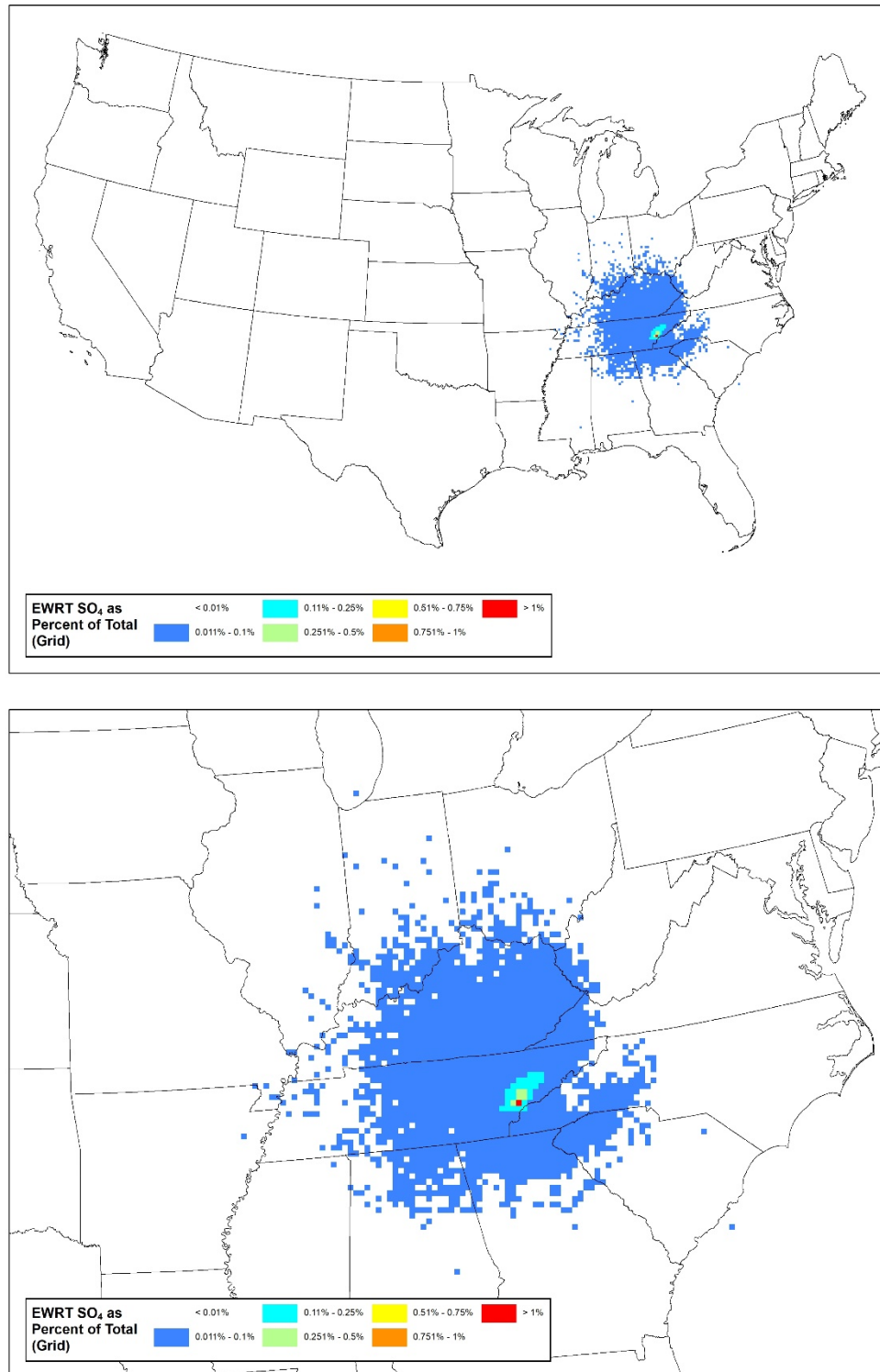
**Figure 2-9. Example EWRT for Ammonium Nitrate Extinction, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS)**



**Figure 2-10. Example EWRT, as a percentage of the total, for Ammonium Nitrate Extinction, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS)**



**Figure 2-11. Example EWRT for Ammonium Sulfate Extinction, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS)**



**Figure 2-12. Example EWRT, as a percentage of the total, for Ammonium Sulfate Extinction, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS)**

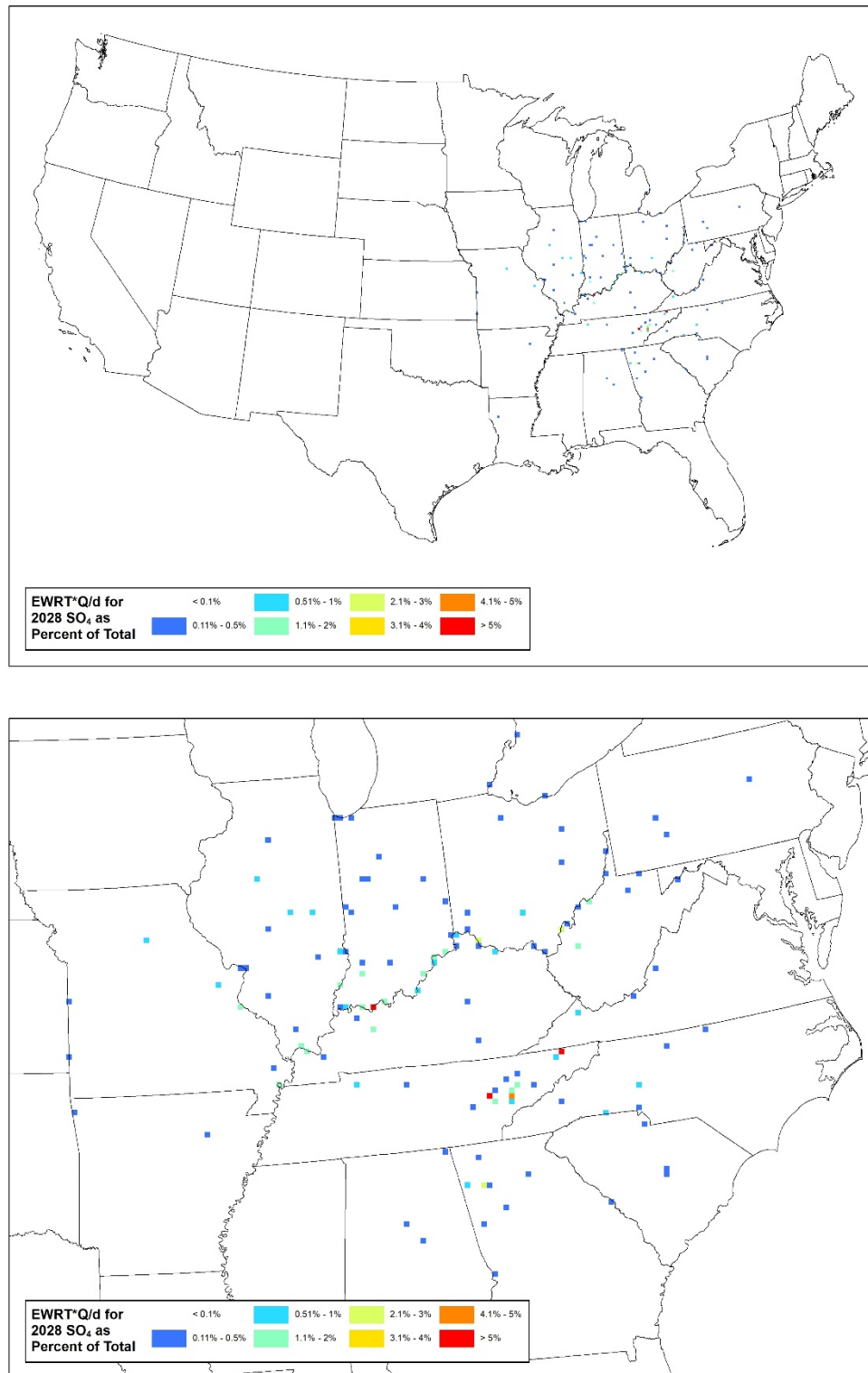
## 2.6 Contributor Identification

The final phase of the analysis combined the EWRT values with the emission data to determine the sources most likely contributing to elevated extinction levels. ArcGIS was used to spatially join the emissions data with the gridded EWRT data to obtain the EWRT value corresponding to each point source. This new layer was extracted from ArcGIS and brought into an Access database to be paired with other data elements.

In the database, the data were further joined with the distance (d) from the point source (average of all emission release points) to the trajectory origin in kilometers, which was calculated using R. The facility emissions (Q, in tons per year) were then divided by the distance (d, in kilometers) to the trajectory origin; for a final value (Q/d). This was then multiplied by the facility's sulfate or nitrate EWRT grid values (i.e.,  $\text{EWRT} \times (Q/d)$ ). Next, the sulfate and nitrate  $\text{EWRT} \times (Q/d)$  values were summed for all point sources at each Class I area and used to normalize the sulfate and nitrate contributions from each individual source. This information allows the individual facilities to be ranked from highest to lowest based on sulfate and/or nitrate contributions.

This data was further paired with additional point source metadata that defined the facility (i.e., Facility ID, Facility Name, State, County, Federal Information Processing Standard (FIPS), NAICS, and industry description). Spreadsheets for individual Class I areas were then exported from the database for further analysis by the states. Plots of the  $\text{EWRT} \times (Q/d)$ , as a percentage of the total, were generated based on the grid total emissions and distance to the class I area. Figure 2-13 shows an example of these plots for the Great Smoky Mountain National Park. The figure includes a wide view of the entire continental United States, and a closer view in the vicinity of the Class I area.

Users of the data should note that while point sources account for most of the sulfate extinction, these sources only account for a portion of the nitrate extinction. Much of the nitrate extinction can be attributable to the onroad and nonpoint sectors. As such, a similar analysis for county level data was conducted, that included county total point source contributions. This allows the point source contribution to be directly compared to the other source categories.



**Figure 2-13. Example EWRT\*(Q/d) for Ammonium Sulfate Extinction for 2028, Per 12km Modeling Grid Cell, Plot for Great Smoky Mountain National Park. Based on All Years. (Full view (top), Class I zoom (bottom); Plot was Created with GIS)**

Similar analysis was conducted to rank SO<sub>2</sub> and NO<sub>x</sub> emissions contributions for the county-level sources (nonpoint, onroad, nonroad, fires, and total point source sectors). The process was similar to the process for point sources previously described, except calculations of RT and EWRT were completed at the county-level as opposed to grids.

The analysis also added new columns to normalize the  $EWRT \cdot (Q/d)$  by the area of each county to develop a metric to compare the contributions from counties on a relative basis. The existing calculation had a propensity to attribute higher contributions to larger counties simply because they typically contained more emission sources and more hourly trajectory end points. Normalizing the contribution by the area of the county (i.e.,  $EWRT \cdot (Q/d)$  per square kilometer) provides a sense of the source emission density within the county. This allows county contributions to be directly compared, without large counties being weighted more heavily by simply having more emission sources and more hourly trajectory end points.

This analysis was completed in GIS using the same calculation method as the point sources. The calculation of “d” was from the centroid of the county to the trajectory origin, in km. Similar to point sources, the final spatial join was made between the county-level EWRT, emissions, and source information for each sector. All county and emissions source identifying information was joined in an Access database with calculations of  $Q/d$  and  $Q/d^2$  values, EWRT,  $EWRT \cdot (Q/d)$ , fraction and sum contributions, and any other source information. The database was then used to generate individual spreadsheets for each Class I area.

**Appendix A**  
**Using Python to Generate HYSPLIT Control Files**

This page is intentionally blank.

## **INTRODUCTION**

Producing Control files through the HYSPLIT GUI involves manual input of modeling parameters for each trajectory run. Python scripts provide an efficient way to create large numbers of HYSPLIT bat files in a fraction of the time. These files can be used to generate trajectory outputs in succession, rather than one at a time.

## **STEP BY STEP INSTRUCTIONS**

This section goes through the Python script for generating Control files for each run.

1. Set paths for inputs and outputs
  - a. Lines 20-31
  - b. These lines should be adjusted to reflect the location of input/output files on your computer or networks.
2. Define model variables for HYSPLIT runs
  - a. Lines 34-43
  - b. These lines indicate heights to be modeled and other parameters.
  - c. Lines 41-43 are standard lines in the Control file and should not be adjusted.
3. getNewDate function
  - a. Returns the date values for the additional days of meteorology files. This corresponds to the hours indicated on line 37 and the number of met files set on line 40. Adjustments to this function would need to be made if these lines changed.
4. getCoords function
  - a. Returns the latitude and longitude coordinates of a site by looking up the classID in a file defined on line 31 as “siteLookup”
5. Writing the bat file
  - a. The script loops through the dayOf and nextDay files and assigns the input values needed and calls the two functions described in steps 3 and 4 above.
  - b. A bat file is created and opened, and the input values are written to the file.
  - c. The file is then closed.
  - d. Note: If a file already exists in the directory, it will be overwritten.

This page is intentionally blank.

**Appendix B**  
**Running HYSPLIT with Batch Files**

This page is intentionally blank.

## 1. RUNNING HYSPLIT OUTSIDE OF THE GUI

After the batch files to run HYSPLIT have been created, a batch file can be used to run HYSPLIT outside of its graphical user interface (GUI). First, the generated batch file must be placed in the HYSPLIT working directory on the computer.

Next, the user needs to generate a list of the batch file to run. The file can be developed in a text editor and needs to contain the following line on separate lines for each file that needs to be run: `call <batch file name>`. Where `<batch file name>` is the name of the batch file, without the brackets. This file can then be saved as with the `.bat` extension and subsequently run by double clicking on the file.

## 2. INCLUDED CODE

**GetBatchFileNames.bat** and **RunAll\_HYSPLIT.bat**. **GetBatchFileNames.bat** is a simple script that obtains a specific set of batch files from the current directory and places them in a list. **RunAll\_HYSPLIT.bat** is a batch file that then runs all the files in the list created by **GetBatchFileNames.bat**.

**GetBatchFileNames.bat**: This code, shown in Figure B-1, utilizes a windows command (`dir /b`) to list the files in a directory and pipe the results into a file named *All\_years.txt*. The flag `"/b"` with the `dir` command lists bare format for the command, which omits headers and summary information. The script also utilizes the wildcard `"*"` to limit the additions to the file to ensure only the files desired are added to the list. For this project, all batch files names start with the year. In line 1 of the code, all the files that start with `"11"` are added to the file. These lines can be edited to accommodate any regular pattern of filename used. The first line contains a single `">"` to write to the file, and a double `">>"` needs to be used to append to the same file without overwriting what was previously there.

```
GetBatchFileNames.bat
1 |dir /b 11* > All_years.txt
2 |dir /b 12* >> All_years.txt
3 |dir /b 13* >> All_years.txt
4 |dir /b 14* >> All_years.txt
5 |dir /b 15* >> All_years.txt
6 |dir /b 16* >> All_years.txt
```

Figure B-1. Contents of the GetBatchFileName.bat file

**RunAll\_HYSPLIT.bat**: This code, shown in Figure B-2, performs actions based on the file names in *All\_years.txt*. The code runs HYSPLIT run batch file (line 5), and then moves the plot (line 7), tdump (lines 9-10), and CONTROL file (lines 12-13) created during the run from the HYSPLIT working directory to a project directory. Users will need to adjust the directories to match their HYSPLIT working directory and working directory.

To run the code, the user will need to first open *All\_years.txt* and remove the .bat from the file names in the list. This can be accomplished by hitting the Control and H keys at the same time to open the “Find and Replace” window. The phrase .bat can be entered in the find box, and the replace box left blank for the task. Next users will need to edit **RunAll\_HYSPLIT.bat** to reflect their directory structure. Once that is complete, the user can close the editor, and double click the file to run.

```
RunAll_HYSPLIT.bat
1 setlocal
2 echo starting HYSPLIT runs
3 for /F "tokens=*" %%a in (All_years.txt) do (
4     echo %%a
5     call %%a.bat
6
7     move C:\hysplit4\working\*.ps C:\hysplit4\working\SESARM_OUT\
8
9     ren %%a %%a.tdump
10    move C:\hysplit4\working\*.tdump C:\hysplit4\working\SESARM_OUT\
11
12    copy CONTROL %%a.CONTROL
13    move C:\hysplit4\working\*.CONTROL C:\hysplit4\working\SESARM_OUT\
14 )
```

Figure B-2. Contents of the script RunAll\_HYSPLIT.bat

Specifically, each line does the following:

Line 2: echoes “starting HYSPLIT runs” to the command window

Line 3: starting line of the for loop

Line 4: echoes the file name that is being run

Line 5: runs HYSPLIT via the previously created batch files

Line 7: moves the plot file to the project directory

Line 9: renames the tdump file to the run *<file name>.tdump*

Line 10: moves the renamed file to the working directory

Line 12: copies the created control file to *<file name>.CONTROL*

Line 13: moves the created control file to the working directory

**Appendix C**  
**Aol Analysis with *R***

This page is intentionally blank.

## INTRODUCTION

*R* is an open source statistical computing and graphing software.<sup>12</sup> The *R* software's base functionality is enhanced by the use of custom functions or collection of functions known as packages. The HYSPLIT\_plots\_09092018.R script utilizes several packages to process and visualize the trajectory data. Appendix E contains the MeteoInfo procedure for analyzing the meteorological information with the spatial allocation information.

## STEP BY STEP INSTRUCTIONS

The following section goes through the process of importing HYSPLIT output into *R* through plotting the trajectory images for several subsets of the data. The plotting is mainly achieved via the function included in the openair package<sup>13</sup>, which includes several options for plot types.

The openair package is capable of completing the residence time and concentration weighted trajectory (CWT) analysis as well. However, an update to a package it was dependent on at the start of the project in summer 2018 created a bug in this processing in the latest version of openair. ERG is currently working on applying the intermediate patch and will provide an update of this code when it has been implemented and tested.

1. Installation
  - a. Download base *R* from <https://cloud.r-project.org/> for your operating system.
    - i. Installation instructions are included on the download site.
    - ii. It is recommended that you download RStudio desktop (<https://www.rstudio.com/products/rstudio/download/#download>) to have a graphical user interface (GUI) to run *R*. This software is also free to download and use.
2. Open HYSPLIT\_plots\_12162018.R
  - a. At line 30, correct the path name to match your system.
    - i. This should be the path to the folders for the project.
    - ii. The script assumes this is the directory with a subfolder for each Class I area.
    - iii. Each Class I folder should have the following folders:
      1. /HYSPLIT\_OUTPUT/02\_TDUMP/: folder with the HYSPLIT tdump files
      2. /ANALYSIS\_PLOTS/: folder to save generated images
3. If using *R* Studio:
  - a. Highlight the lines 10 through 22, and click run at the top left to install the need packages and pull in the air quality data.
  - b. Next lines 38-130 can be run to regenerate all the analysis.

---

<sup>12</sup> <https://www.r-project.org/>; Wang, Y.Q., 2014. MeteoInfo: GIS software for meteorological data visualization and analysis. Meteorological Applications, 21: 360-368.

<sup>13</sup> <http://davidcarslaw.github.io/openair/>; Carslaw, D.C., Ropkins, K. (2012). "openair — An *R* package for air quality data analysis." Environmental Modelling & Software, 27–28(0), 52–61. ISSN 1364-8152, doi: 10.1016/j.envsoft.2011.09.008.

- i. This code loops through all the Class I area folders (01-43). Merges all the tdump files together, and then generates the trajectory plots.
- c. If you want to run new plots for one area:
  - i. Skip down to line 166 and change the list of sites to the desired Class I area ID.
  - ii. Highlight line 160 through the end, and click run.
  - iii. The highlight the code with the plot you would like to generate.
  - iv. The code can also be adjusted to run various iteration of the plots:
    - 1. Type can be changes to any of the variables to generate plot by those categories. (as in line 167).
    - 2. Multiple values of type can be specified (as in line 157).

**Appendix D**  
**Aol Analysis with Python and ArcGIS**

This page is intentionally blank.

## INTRODUCTION

*R* is used to format data from HYSPLIT tdump into yearly CSV files by Class 1 site IDs. This combined file can be processed using a Python script to find the count of trajectories by county, and to calculate the sulfate and nitrate extinction-weighted residence time (EWRT) for each county within the area of interest. A second version of the code was produced to complete the same calculations for the modeling grid.

## STEP BY STEP INSTRUCTIONS

This section goes through, in order, the *R* and Python scripts used in the county analysis. ArcPy is used in the Python script and requires an ESRI users license.

1. Open HYSPLIT\_combineYears.R
  - a. This script assumes that the HYSPLIT .tdump files are separated by year into individual folders.
  - b. Place the *HYSPLIT\_combineYears.R* script in the folder with the files and run. CSV files will be written to the same folder. File naming convention:  
***Rcombined\_<ID>\_<YR>.csv***
  - c. Repeat this process year ran in HYSPLIT.
  - d. Future Update: This script could be modified to run without having to separate the TDUMP files by year.
2. Open HYSPLIT\_AllYears.R
  - a. This file combines all the year files created using HYSPLIT\_combineYears.R. The output csv file created in step 1 should be moved to a new folder with the *HYSPLIT\_AllYears.R* script.
  - b. All files for each of the Class 1 IDs will be read in a written to one file, combining all of the years into the same file for processing. File naming convention: ***AllYears\_Rcombined\_<ID>.csv***
  - c. Future Update: Combine the scripts used in steps 1 and 2 into a single *R* script.
3. **County\_PointCounts\_EWRT.py**
  - a. This final step of the analysis is running the *County\_PointCounts\_EWRT.py* Python script. This script pulls in the all years combined text files for each of the Class 1 areas and spatially relates the trajectory outputs to the counties within the area of interest using arcpy.
  - b. The arcpy workspace is specified on line 15 and the overwrite is set to True so that existing files will be overwritten.
  - c. Directories should be changed to reflect your own working directories, or paths created to match what is currently in the script.
  - d. Check the county shapefile and variable names on lines 23-25.
  - e. Line 30 assigns an excel file containing the beta extinction. Check the file name and directory. This file must be converted to a table so that it may be used as a look up table later in the script.
  - f. The script then iterates through each file in the file list and creates shapefiles and joins with the county to get information on a county basis.

- g. Lines 94-112 execute the EWRT calculations for each county and export the final data to a CSV file. File naming convention:  
AllYears\_Rcombined\_<ID>\_EWRT\_final.csv

**Appendix E**  
**Aol Analysis with MeteolInfo**

This page is intentionally blank.

---

## INTRODUCTION

MeteoInfo is an open source software that was developed to view and analyze meteorological and spatial information.<sup>14</sup> MeteoInfo has a “trajStat” plugin that performs the AoI trajectory analysis (i.e., residence time and concentration weighted trajectory (CWT)). MeteoInfo also has additional spatial analysis functions that are not discussed here, but may be of interest to states looking for an open source alternative to ArcGIS.

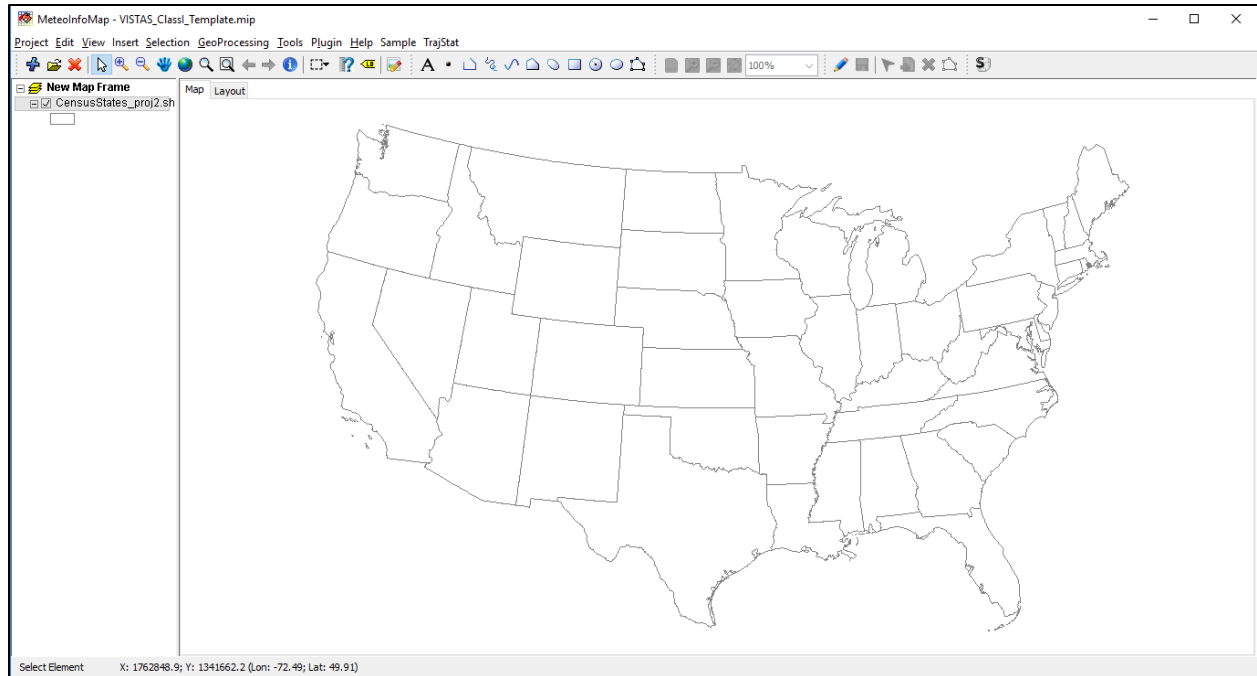
## STEP BY STEP INSTRUCTIONS

The following section goes through the process of importing HYSPLIT output into MeteoInfo through the CWT calculations used in the AoI analysis.

1. Installation
  - a. Download files from <http://www.meteothinker.org/downloads/index.html>
  - b. Select the directory for installation.
  - c. Unzip MeteoInfo in install directory
  - d. Unzip TrajStat in to MeteoInfo\plugins\
  - e. Open MeteoInfo by double clicking MeteoInfoMap.exe in the install directory
  - f. Go to Plugin>Plugin Manager> to add the TrajStat plugin to MeteoInfo
2. Set map base layer
  - a. Add a new layer
    - i. \BaseLayers\CensusStates\_proj2.shp is the continental US on the standard modeling projection.
  - b. For convenience, VISTAS\_Template.mip has been created as a starting point.
    - i. Go to Project> Open and navigate to the save location of the template.
    - ii. Click Open to open the template map (Figure E-1).

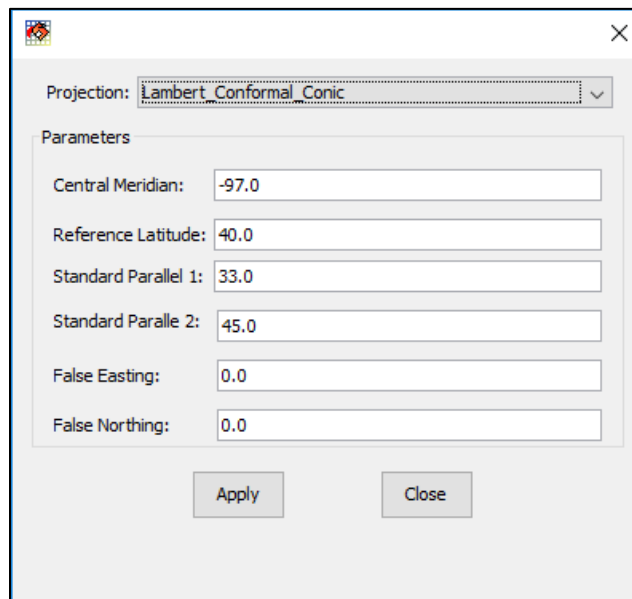
---

<sup>14</sup> <http://www.meteothinker.org/#>; Wang, Y.Q., 2014. MeteoInfo: GIS software for meteorological data visualization and analysis. Meteorological Applications, 21: 360-368.



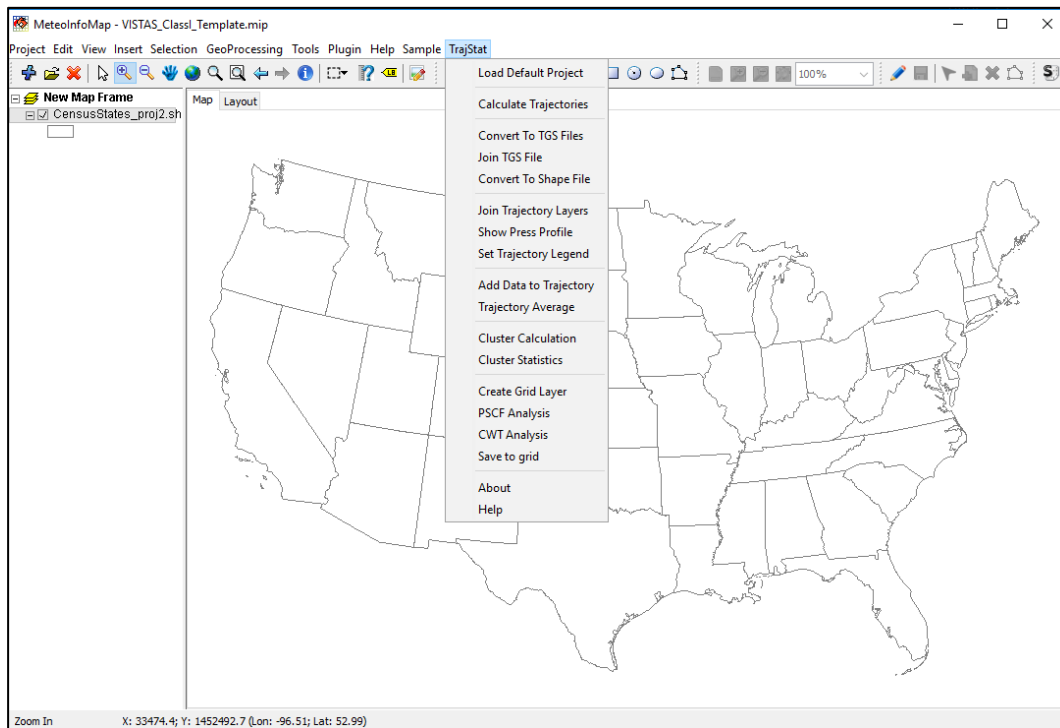
**Figure E-1. MeteoInfo with the VISTAS Template Map File (VISTAS\_Template.mip) Opened**

3. Check projection
  - a. Select View > Projection to open the projection menu.
  - b. From this menu you can adjust the projection of the map (Figure E-2).



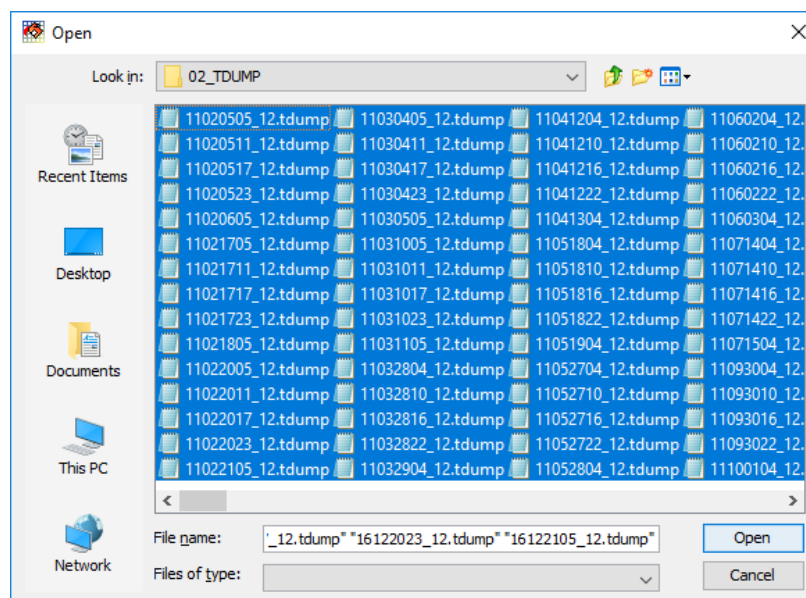
**Figure E-2. MeteoInfo Projection Menu**

4. Importing HYSPLIT output (converting .tdump files to .TGS)
  - a. Select “Convert to TGS Files” from the TrajStat Menu (Figure E-3)



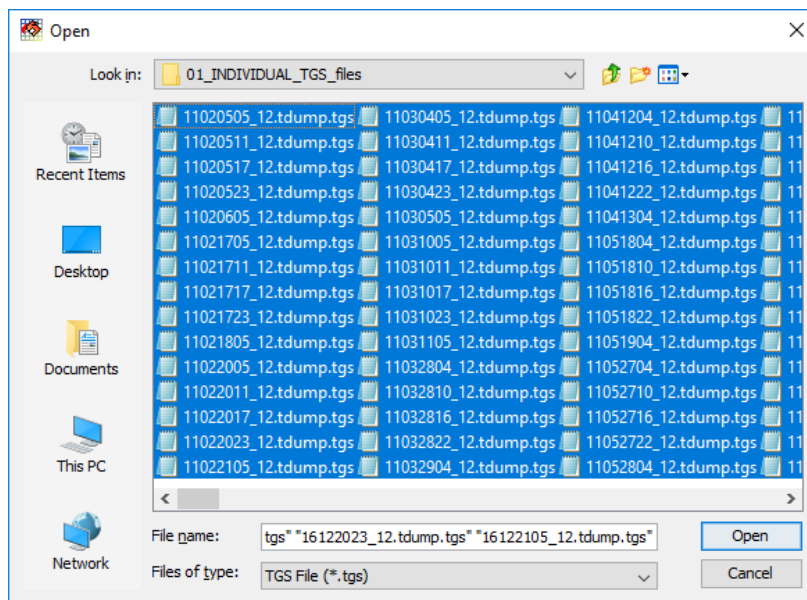
**Figure E-3. MeteInfo TrajStat Menu**

- b. Navigate to folder with the .tdump files from HYSPLIT
  - c. Select all trajectories for to be considered in the analysis
    - i. You can hold control while you click to select multiple files
    - ii. Or hit control+A to select all files in a folder (Figure E-4)



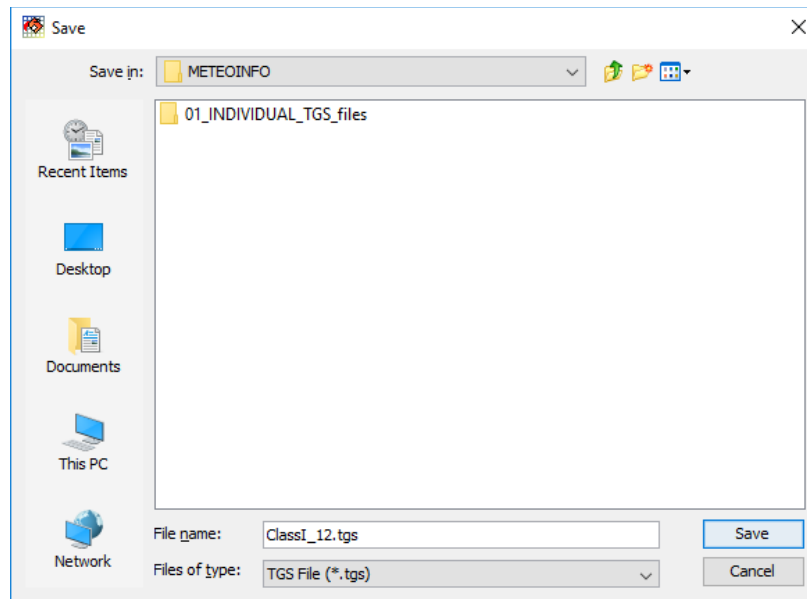
**Figure E-4. Selecting HYSPLIT Files for Import into MeteInfo**

- d. Select 'Open'
  - e. It will create the TGS files in the same directory as the TDUMP files.
5. Combine into a single TGS
- a. Select "Join TGS File" from TrajStat Menu (Figure E-5)
  - b. Navigate to the folder with the TGS files
  - c. Select all TGS for to be considered in the analysis
    - i. You can hold control while you click to select multiple files
    - ii. Or hit control+A to select all files in a folder



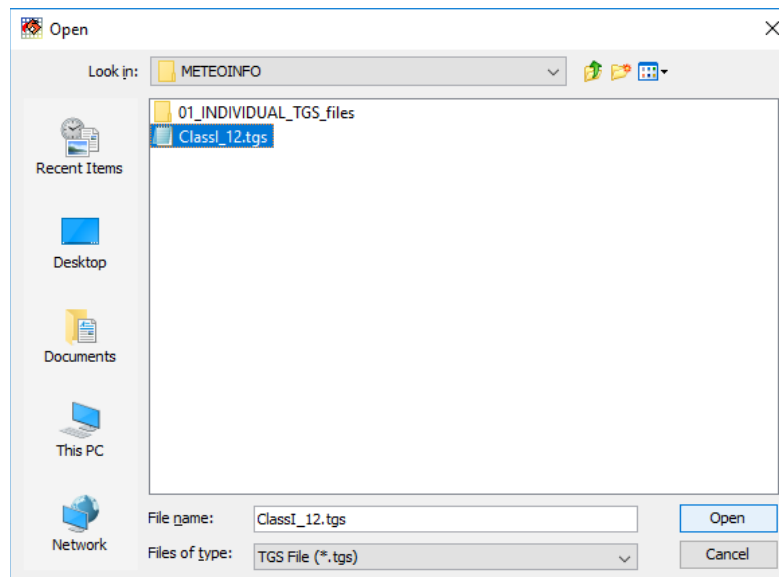
**Figure E-5. Selecting tgs Files to Join**

- d. Click 'Open'
- e. In the next window, enter the name you want to save the file as, including the extension .tgs (Figure E-6)



**Figure E-6. Saving Joined tgs File**

6. Convert to shapefile
  - a. Select “Convert to Shape File” from TrajStat menu (Figure E-7)



**Figure E-7. Converting Joined tgs File to a Shapefile**

- b. Will save as the shapefile with the same name as the TGS file.
7. Add extinction data
  - a. Select “Add Data to Trajectory” from the TrajStat menu (Figure E-8)
  - b. Navigate to the file with the extinction information
    - i. File should be either a csv or txt file

- ii. Date should be included, and formatted as yyyyMMDD or YYYY-MM-DD (option shown in)

**Add data to trajectories**

This tool will add data from a comma-delimited text file to trajectories according to the date field. The file must contain column titles as the first row and one or two date columns.

Input File:

Select Fields From Input File

Start Date:

End Date:

Date Format:

Time Zone:

Data Field:

Data Field Property

Name:

Type:

Width:

Precision:

Missing Value:

Add Data Cancel

**Figure E-8. Screen to Add Data to the Trajectory Shapefile**

- c. Set date field, and note the format in the “Date Format” drop down.
- d. Select field to add, adjusting the field format and Time Zone as necessary (Figure E-9).
  - i. You can add air quality information or other identifying fields (e.g., Class I area name) to the file at this point.

- ii. If using the AQ files provided, the “MeteoInfoDate” field is the trajectory start time in GMT.

**Add data to trajectories**

This tool will add data from a comma-delimited text file to trajectories according to the date field. The file must contain column titles as the first row and one or two date columns.

Input File: Apache Wilderness Area\_TrajDatesWithBext.csv

Select Fields From Input File

Start Date: MeteoInfoDate

End Date: Null

Date Format: yyyyMMddHH

Time Zone: GMT+0

Data Field: EAmn\_SO4

Data Field Property

Name: EAmn\_SO4

Type: Double

Width: 10

Precision: 4

Missing Value: -9999.0

Add Data Cancel

**Figure E-9. Screen to Add Data to the Trajectory Shapefile, Addition of Ammonium Nitrate Extinction Values**

8. Add grid
  - a. Select “Create Grid Layer” from TrajStat menu to open the grid menu (Figure E-10).
    - i. If starting from the VISTAS\_template.mip, the file is in the same projection that is used for modeling. This will allow you to use the same grid specification as the modeling to define the grid.

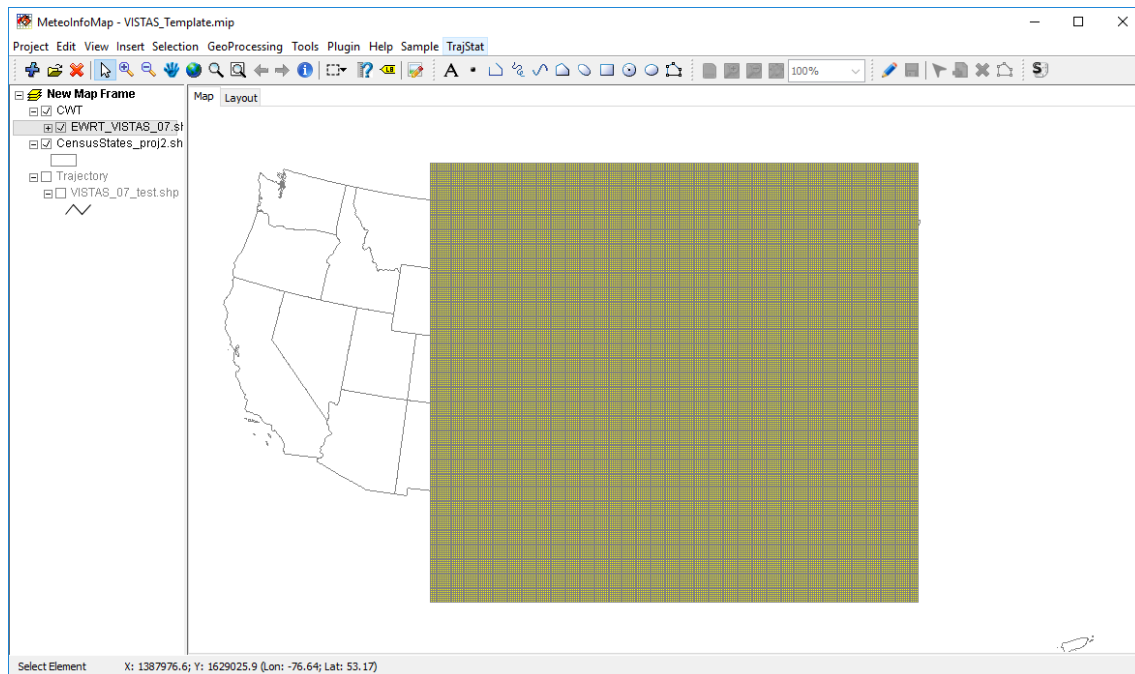
- b. Select “Same as Map” and enter the bottom left and top right coordinates (in meters) and then enter the cell size in meters.
- c. Select CWT as type.
- d. Set the desired precision for the CWT field.
- e. Click ‘Create Layer’ to generate the grid.
  - i. A navigation window will open, select save location and name.
  - ii. After it is generated, you can click “Create Layer” to create a second grid for a second weighting scheme.
  - iii. The grids are automatically added to the maps. (Figure E-11)

The screenshot shows a dialog box titled "Add Uniform Grid" with a close button (X) in the top right corner. The dialog is divided into three main sections: "Projection", "Setting", and "CWT Field".

- Projection:** Contains two radio buttons. "Longitude/Latitude" is unselected, and "Same as Map" is selected.
- Setting:** Contains a dropdown menu "Extent same as:" with "None" selected. Below this are input fields for "Left: -912000", "Right: 2316000", "Cell Size: 12000", "Bottom: -1596000", and "Top: 1308000". To the right of these is a dropdown menu "Type:" with "CWT" selected.
- CWT Field:** Contains two spinners. "Width:" is set to 10, and "Precision:" is set to 2.

At the bottom center of the dialog is a button labeled "Create Layer".

**Figure E-10. Screen to Add Uniform Grid to the Map File**



**Figure E-11. Resulting Grid is Automatically Added to the Map**

9. Concentration/Extinction weighted trajectory (CWT) Analysis

- a. Select CWT Analysis from the TrajStat menu to open the CWT analysis window (Figure E-12)
- b. For CWT Layer - Select layer to generate the CWT based on (i.e., the grid layer)
- c. For Field: Select the field for the Concentration/Extinction weighting)
- d. Click “Get Nij” to get trajectory counts per grid cell
- e. Click “Cal CWT” to calculate the Extinction Weighted Trajectory
- f. Optional: Click “Weight CWT” to calculate a weighted CWT (weighted even more to cell with higher counts)
- g. Close window when completed
- h. The data will have been added to the grid shapefiles.
- i. Adjust the map to show the CWT analysis
  - i. Right Click on the layer to open the layer menu (Figure E-13) and select properties to open the layer property window (Figure E-14).
  - ii. From the layer property window, select “Graduated Color” as the legend type and the desired variable ( $N_{ij}$  = residence time, CWT, or weighted CWT). Then Click OK.
  - iii. The map will update to show your final plot (Figure E-15), which can be saved (via the save picture Icon on the tool bar). Otherwise the data is in a shapefile format, which can be brought into ArcGIS.

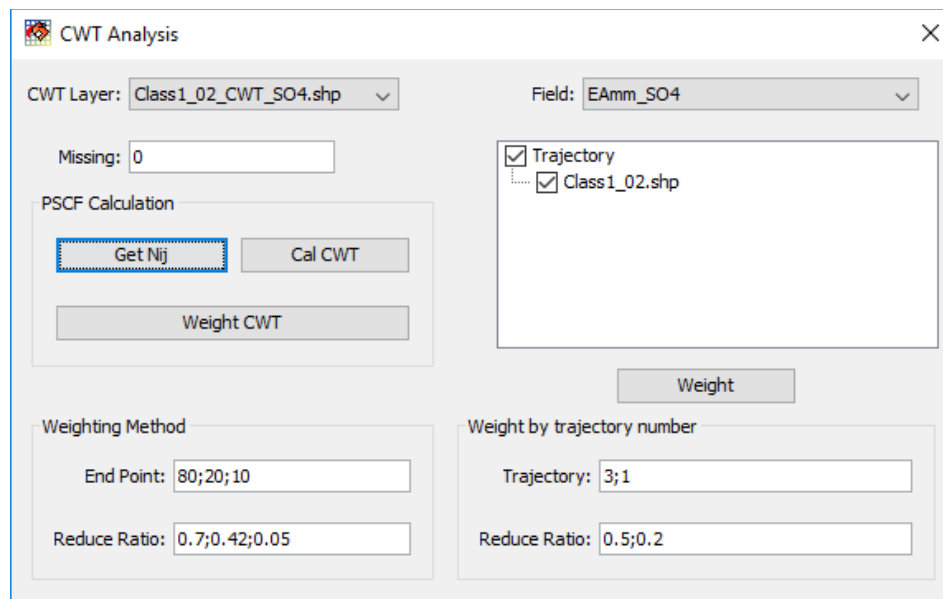


Figure E-12. CWT Analysis Window

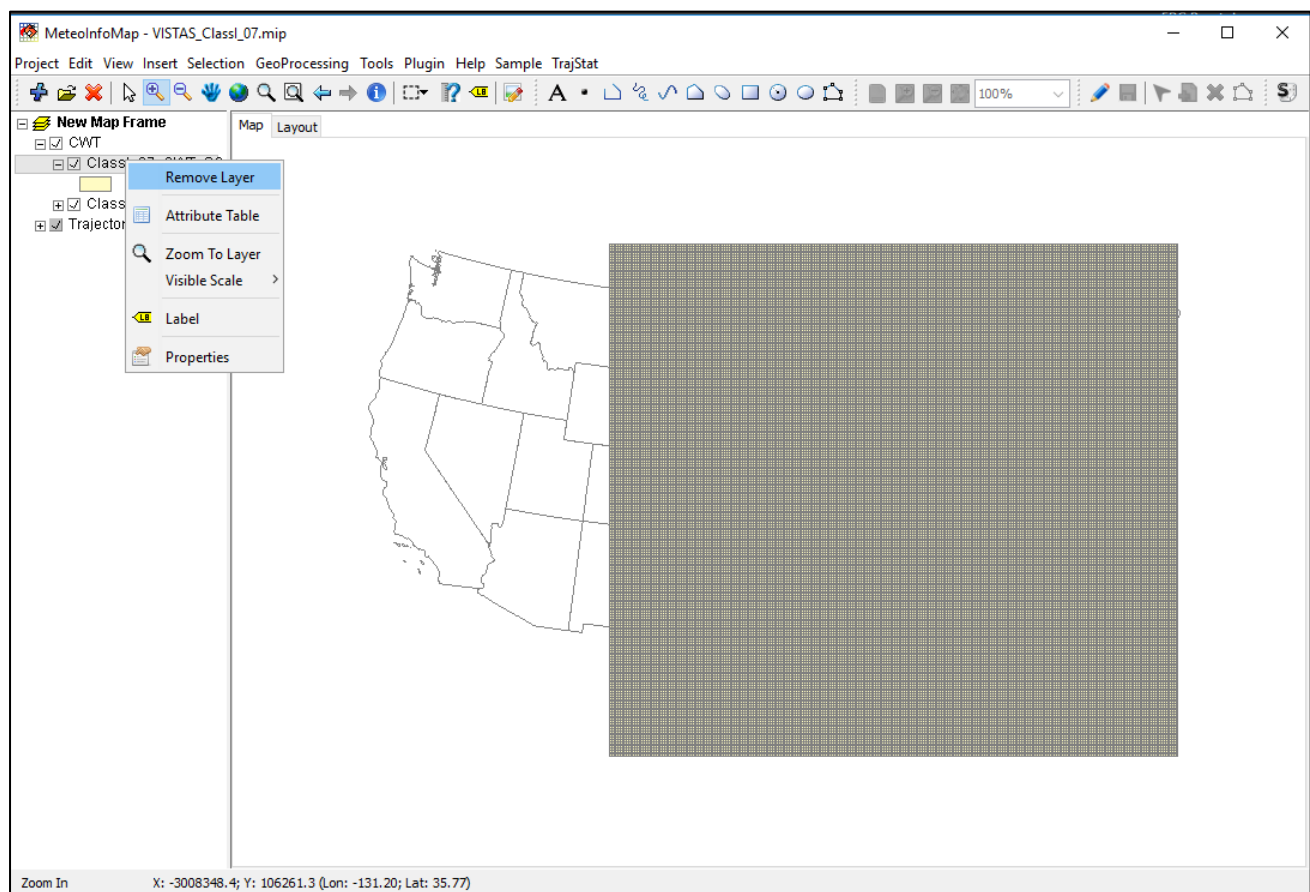
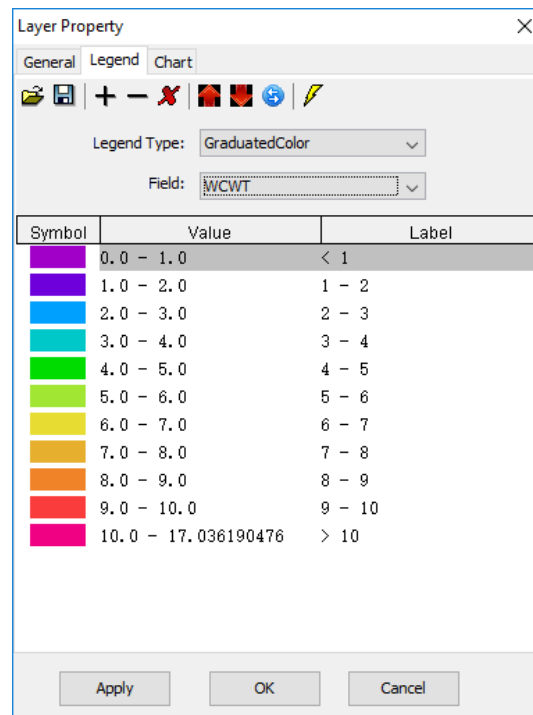
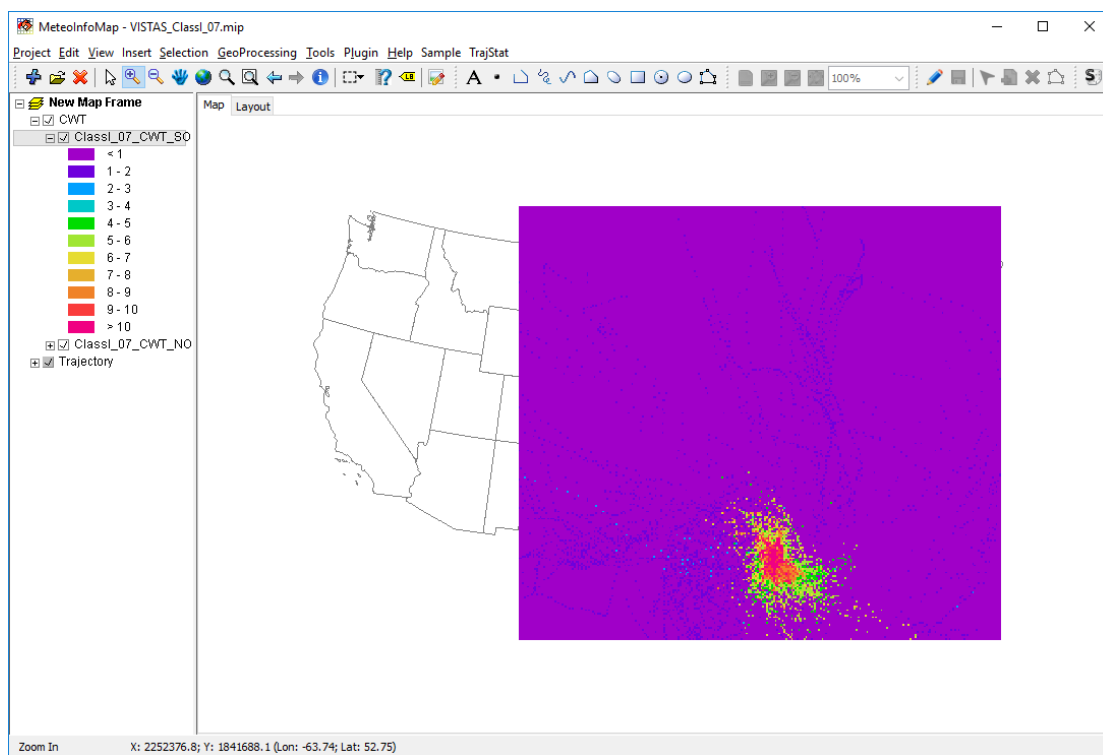


Figure E-13. Select Layer Properties to Shade Grid According to the CWT Results



**Figure E-14. Select Layer Properties Window to Apply Color Shading Based on Value to the Layer**



**Figure E-15. Gradient Color Plot of the Weighted CWT**

This page is intentionally blank.

*(This page intentionally left blank)*