Appendix E-1a

Regional Haze Modeling for Southeastern VISTAS II Regional Haze Analysis Project

Final Modeling Protocol

June 27, 2018

Regional Haze Modeling for Southeastern VISTAS II Regional Haze Analysis Project

Final Modeling Protocol

Prepared for:
SESARM, Inc.
205 Corporate Center Dr., Suite D
Stockbridge, GA 30281-7383

Under Contract No. V-2018-03-01

Prepared by:
Alpine Geophysics, LLC
387 Pollard Mine Road
Burnsville, NC 28714

and

Eastern Research Group, Inc.
1600 Perimeter Park Dr., Suite 200
Morrisville, NC 27560

June 27, 2018

Alpine Project Number: TS-527
ERG Project Number: 4133.00.006

This page is intentionally blank.

TITLE and APPROVAL SHEET

Modeling Protocol Plan for Southeastern VISTAS II Regional Haze Analysis Project for SESARM (Final)

This Modeling Protocol is approved by the undersigned and effective on the latest date signed by any party. The organizations implementing the project are Eastern Research Group, Inc. (ERG) and Alpine Geophysics, LLC (Alpine).

6/27/2018

Regi Oommen – ERG Program Manager and Technical Project Coordinator — Date

6/27/2018

Gregory Stella – Alpine Subcontract Manager — Date

John E Hornback

6/27/2018

John Hornback – SESARM Executive Director — Date

This page is intentionally blank.

Contents

TABLES

FIGURES

ABBREVIATIONS / ACRONYM LIST

3DVAR	Three-dimensional variational
AFWA	Air Force Weather Agency
AGL	Above ground level
AL	Alabama
ARW	Advanced research WRF
AIRMoN	Atmospheric Integrated Research Monitoring Network
Alpine	Alpine Geophysics, LLC
AMET	Atmospheric Evaluation Tool
AoI	Area of Influence
AQS	Air Quality System
ARL	Air Resources Laboratory
BC	Boundary conditions
BEIS	Biogenic Emissions Inventory System
b_{ext}	Beta extinction
BNDEXTR	Program used to extract boundary conditions
CAIR	Clean Air Interstate Rule
CAMD	Clean Air Markets Division
CAMx	Comprehensive Air Quality Model with Extensions
CASTNET	Clean Air Status and Trends Network
CB6r4	Version 6 Revision 4 of the Carbon Bond chemical mechanism
CCRS	Coarse PM species (CAMx PM species)
CEM	Continuous Emissions Monitoring
CFR	Code of Federal Regulation
CIRA	Cooperative Institute for Research in the Atmosphere
Cl	Chorine
CMAQ	Community Multiscale Air Quality
CMV	Commercial marine vessel
CO	Carbon monoxide
CONUS	Continental United States
CPRM	Coarse PM
CSAPR	Cross-state Air Pollution Role
CSN	Chemical Speciation Network
CWT	Concentration weighted trajectory
d	Distance
DJF	December, January, and February
dv	Deciview
EC	Elemental carbon
EDAS	Eta data assimilation system
EF	Emission factor
EGU	Electric generating unit
EIS	Emission Inventory System
EPA	United States Environmental Protection Agency
ERG	Eastern Research Group, Inc.
ERTAC	Eastern Regional Technical Advisory Committee

EWRT	Extinction-weighted residency time
f(RH)	Monthly relative humidity function
FAA	Federal Aviation Administration
FCRS	Crustal fraction of PM
FIPS	Federal Information Processing Standard
f_L(RH)	Monthly relative humidity function associated with large size distributions
FLM	Federal Land Manager
FL	Florida
FPRM	Fine Other Primary (diameter $\leq$2.5 µm)
FR	Federal Register
FSL	Forecast Systems Laboratory
f_s(RH)	Monthly relative humidity function associated with small size distribution
f_{ss}(RH)	Monthly relative humidity function associated with sea salt
FTP	File Transfer Protocol
g	Gram
GA	Georgia
GB	Gigabyte
GEOS	Goddard Earth Observing Station
GHRSST	Group for High Resolution Sea Surface Temperatures
GIS	Geographic Information System
H^+ as pH	Free acidity
ha	Hectare
Hg	Total mercury
HGP	Particulate mercury
HNO_3	Nitric acid
HYSPLIT	Hybrid Single Particle Lagrangian Integrated Trajectory
IC	Initial conditions
IMPROVE	Interagency Monitoring of Protected Visual Environments
IPXWRF	Intermediate Processor for Pleim–Xiu for WRF
IPM	Integrated Planning Model
JJA	June, July, and August
K^+	Potassium ion
km	Kilometers
kv	Eddy diffusivity
KY	Kentucky
L	Liter
LAC	Light absorbing carbon
LCP	Landscape
m	Meters
m^2	Square meters
m^3	Cubic meters
MANE-VU	Mid-Atlantic/Northeast Visibility Union
MAM	March, April, and May
MAR	Marine, aircraft, and rail
mb	Millibar
MB	Mean Bias

MDA8	Daily maximum 8-hour average
MDL	Method Detection Level
MDN	Mercury Deposition Network
ME	Mean Error
MFB	Mean Fractional Bias
MFE	Mean Fractional Error
mg	Milligram
Mg^{2+}	Magnesium ion
µg	Microgram
MJO	Multi-Jurisdictional Organization
MLM	Multi-Layer Model
Mm	Megameters
Modeled	Mean Modeled value
MOVES	Motor Vehicles Emissions Simulator
MPE	Model performance evaluation
MPI	Message Passing Interface
MS	Mississippi
Na^+	Sodium ion
NAAQS	National Ambient Air Quality Standards
NADP	National Atmospheric Deposition Program
NAICS	North American Industry Classification System
NAM-12	North American Mesoscale forecast data at the 12-km level
NC	North Carolina
NCAR	National Center for Atmospheric Research
NCDC	National Climatic Data Center
NCEI	National Centers for Environmental Information
NCEP	National Centers for Environmental Prediction
NH_3	Ammonia
NH_4^+	Ammonium ion
NLCD	National Land Cover Database
NMB	Normalized Mean Bias
NME	Normalized Mean Error
NO_2	Nitrogen dioxide
NO	Nitric oxide
NO_3^-	Nitrate
NOAA	National Oceanic and Atmospheric Administration
NODA	Notice of data availability
NO_x	Oxides of nitrogen
O_3	Ozone
OC	Organic carbon
OCM	Organic carbon mass
OM	Organic matter
OMI	Ozone Monitoring Instrument
ORIS	Plant identifier issued by U.S. Department of Energy
PAL	Particulate Aluminum
PAMS	Photochemical Assessment Monitoring System

Pb	Lead
PBL	Planetary Boundary Layer
PCA	Particulate Calcium
PEC	Primary elemental carbon
PFE	Particulate Iron
PGI	Portland Group, Inc.
PGMs	Photochemical grid models
PiG	Plume-in-Grid
PM	Particulate matter
PM_{10}-PRI	Primary particulate matter less than or equal to 10 microns in aerodynamic diameter
$PM_{2.5}$	Fine particle; primary particulate matter less than or equal to 2.5 microns in aerodynamic diameter
$PM_{2.5}$-PRI	Primary particulate matter less than or equal to 2.5 microns in aerodynamic diameter
PNH4	Ammonium
PNO3	Particulate nitrate
POA	Primary organic carbon
POC	Parameter occurrence code
ppb	parts per billion
PSAT	Particulate Matter Source Apportionment Technology
PSCF	Potential Source Contribution Factor
PSI	Particulate Silicon
PSO4	Sulfate
PTI	Particulate Titanium
Q	Emissions
Q/d	Emissions over distance
QA	Quality Assurance
QC	Quality control
r	Pearson correlation coefficient
r^2	Coefficient of Determination
RH	Relative humidity
RHR	Regional Haze Rule
RPG	Reasonable Progress Goals
RRF	Relative response factor
RRTMG	Rapid Radiative Transfer Model-Global
SC	South Carolina
SCC	Source Classification Code
SESARM	Southeastern States Air Resource Managers, Inc.
SIP	State Implementation Plan
SMARTFIRE	Satellite Mapping Automated Reanalysis Tool for Fire Incident Reconciliation
SMAT-CE	Software for Model Attainment Test - Community Edition
SMOKE	Sparse Matrix Operator Kernel Emissions
SO_2	Sulfur dioxide
SO_4^{2-}	Sulfate

SOA	Secondary organic aerosol
SOAP	Secondary organic aerosol partitioning
SON	September, October, and November
TDEP	Total Deposition
TN	Tennessee
TOMS	Total Ozone Mapping Spectrometer
TSD	Technical Support Document
URP	Uniform rate of progress
USDA	U.S. Department of Agriculture
USDI	U.S. Department of the Interior
VA	Virginia
VISTAS	Visibility Improvement - State and Tribal Association of the Southeast
VISTAS_12	12-km modeling domain for the VISTAS study area
VMT	Vehicle Miles Travelled
VOC	Volatile organic compounds
WRF	Weather Research Forecast
WV	West Virginia
YSU	Yonsei University

This page is intentionally blank.

1.0 INTRODUCTION

1.1 Overview

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 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. The 1999 RHR required states to define long-term strategies to improve visibility in Federal Class I national parks and wilderness 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.[1] The RHR requires states to evaluate progress toward visibility improvement goals every five years and submit revised SIPs every ten years.

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

---

[1] RHR summary data is available at: http://vista.cira.colostate.edu/Improve/rhr-summary-data/

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_2) National Ambient Air Quality Standards (NAAQS); the 2012 annual fine particle ($PM_{2.5}$) 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 (dv)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 since the baseline period and ensure no degradation in visibility for the clearest days since the baseline period.

This document serves as the air quality Modeling Protocol for SESARM’s VISTAS II regional haze modeling analysis in support of estimating regional haze and progress goals at southeastern state Class I areas in projection year 2028. The reasonable progress goals must provide for a rate of improvement sufficient to attain “natural conditions” by 2064, or justify any alternative rate. States are to define controls to meet progress goals every 10 years, starting in 2018 that defines progress periods ending in 2018, 2028, 2038, 2048, 2058 and finally 2064. States will determine whether they are meeting their goals by comparing visibility conditions from one five-year period to another (e.g., 2000-2004 to 2013- 2017). As stated in 40 CFR 51.308 (d) (1), baseline visibility conditions, progress goals, and changes in visibility must be expressed in terms of dv units. The dv unit of visibility impairment is derived from beta light extinction (b_{ext}) as follows:

$$dv = 10 \ln (b_{ext}/10)$$

Where b_{ext} is expressed in terms of inverse megameters ($Mm^{-1} = (10^6\ m)^{-1}$).

This Modeling Protocol describes the overall modeling activities to be performed in order to estimate regional haze and progress at southeastern state Class I areas in projection year 2028. This effort is being undertaken working closely with SESARM and other state, local, and tribal agencies.

A comprehensive Modeling Protocol for regional haze study consists of many elements. Its main function is to serve as a means for planning and communicating how a modeled analysis will be performed before it occurs. The protocol guides the technical details of a modeling study and provides a formal framework within which the scientific assumptions, operational details, commitments and expectations of the various participants can be set forth explicitly and means for resolution of potential differences of technical and policy opinion can be worked out openly and within prescribed time and budget constraints.

As noted in the EPA regional haze modeling guidance, the Modeling Protocol serves several important functions (EPA, 2007; 2014e):

- Identify the assistance available to SESARM (the lead agencies) to undertake and evaluate the analysis needed to support the analysis;
- Identify how communication will occur among State, Local, and Federal agencies and stakeholders to develop a consensus on various issues;
- Describe the review process applied to key steps in the analysis; and
- Describe how changes in methods and procedures or in the protocol itself will be agreed upon and communicated with stakeholders and the appropriate EPA Regional Office.

1.2 Study Background

In Section 169A of the Clean Air Act, Congress established a visibility protection goal to prevent future and remedy existing impairment of visibility resulting from manmade pollution in certain national parks and wilderness areas. The statute was codified at 42 U.S. Code §7491. EPA issued regulations implementing the visibility protection mandate that may be found at 40 CFR 51.300 through 51.309. These regulations have been amended several times, most recently as published in the Federal Register on January 10, 2017.

The 1999 RHR (64 FR 35714) identified 156 parks and natural areas as “mandatory Class I Federal areas” for which goals would be established to improve visibility to natural conditions. The 18 Class I areas located in the VISTAS region are tabulated in Table 1-1 that follows. Each row contains the official Class I area name, the state(s) in which it is located, estimated total land area in acres based on current available information, and the designated Federal Land Manager (FLM). A map of the VISTAS Class I areas follows in Figure 1-1.

Table 1-1. VISTAS Region Federal Class I Area List

(State) Class I Area Name	Approx. Acreage	Federal Land Manager
AL – Sipsey Wilderness Area	12,726	USDA Forest Service
FL – Chassahowitzka Wilderness Area	23,579	USDI Fish and Wildlife Service
FL – Everglades National Park	1,399,078	USDI National Park Service
FL – Saint Marks Wilderness Area	17,350	USDI Fish and Wildlife Service
GA/TN – Cohutta Wilderness Area	GA – 35,268 TN – 1,709	USDA Forest Service
GA – Okefenokee Wilderness Area	353,981	USDI Fish and Wildlife Service
GA – Wolf Island Wilderness	5,126	USDI Fish and Wildlife Service
KY – Mammoth Cave National Park	52,830	USDI National Park Service
NC/TN – Great Smoky Mountains National Park	NC – 277,432 TN – 244,645	USDI National Park Service
NC/TN – Joyce Kilmer-Slickrock Wilderness	NC – 13,590 TN – 3,820	USDA Forest Service
NC – Linville Gorge Wilderness Area	11,651	USDA Forest Service
NC – Shining Rock Wilderness Area	18,479	USDA Forest Service
NC – Swanquarter Wilderness Area	8,800	USDI Fish and Wildlife Service
SC – Cape Romain Wilderness	29,000	USDI Fish and Wildlife Service
VA – James River Face Wilderness	8,907	USDA Forest Service
VA – Shenandoah National Park	199,173	USDI National Park Service
WV – Dolly Sods Wilderness	17,776	USDA Forest Service
WV – Otter Creek Wilderness	20,706	USDA Forest Service

USDA – United States Department of Agriculture. USDI – United States Department of Interior.

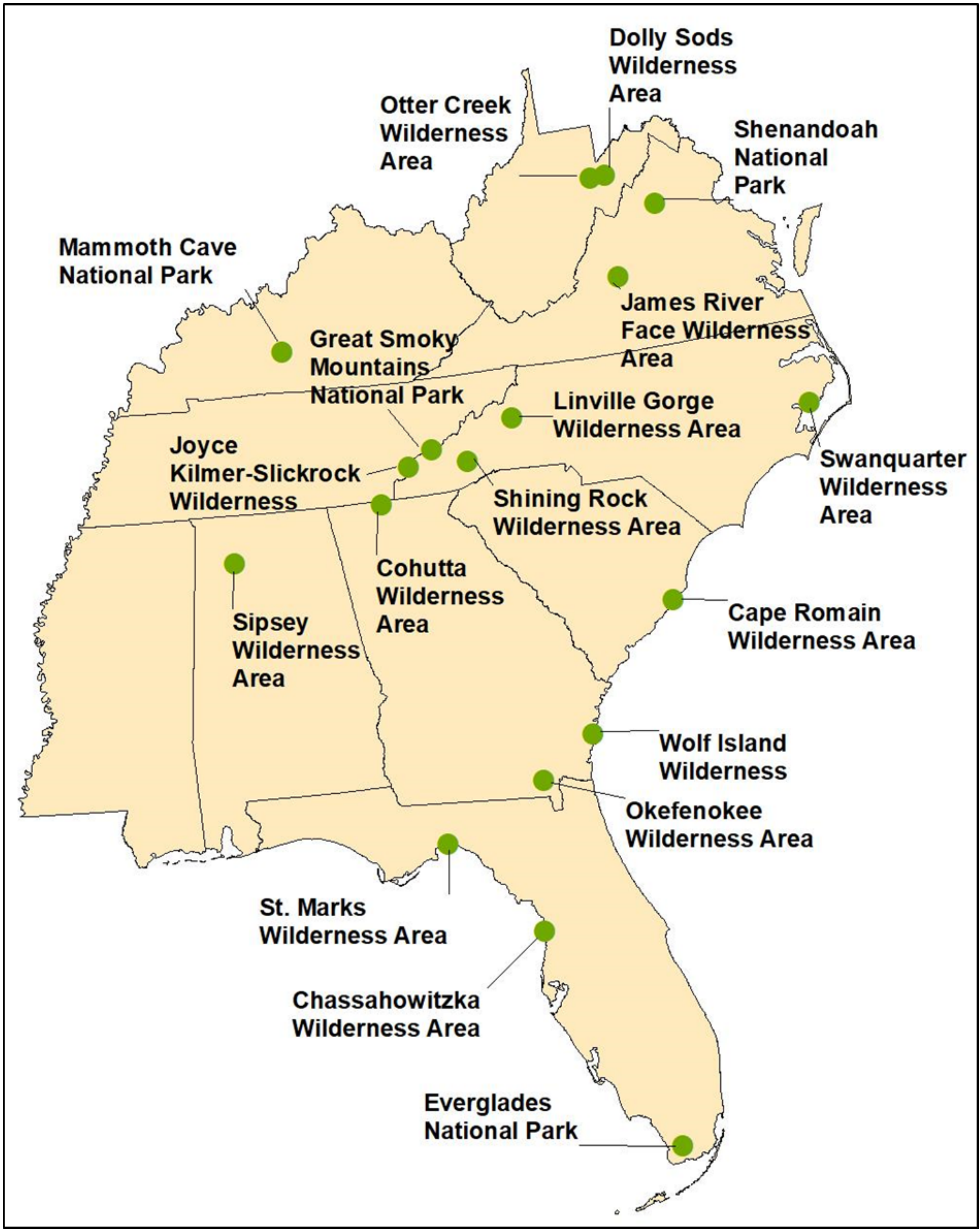


Figure 1-1. VISTAS Region Class I Areas.

The 1999 RHR required states to define long-term strategies to improve visibility in Federal Class I national parks and wilderness 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 improve visibility systematically to reach natural visibility conditions by 2064. The original RHR required states to improve visibility on the 20% most impaired days and protect visibility on the 20% clearest days. States were required to submit SIPs by December 17, 2007 demonstrating reasonable progress to achieve incremental visibility improvements for the 2008-2018 planning period.

The RHR requires states to evaluate progress toward visibility improvement goals every five years and submit revised SIPs every ten years. States are to consult with FLMs in developing the SIPs.

1.2.1 Natural and Current Base Model Period DV Values at the VISTAS Class I Areas

Table 1-2 shows the U.S. EPA calculated natural conditions on the 20% most impaired days and dv values on the 20% clearest and most impaired days at each VISTAS Class I area for the base model period (2009-2013).

Table 1-2. Natural Conditions and Base Year Deciview (dv) Values on the 20% Clearest and 20% Most Impaired Days at each VISTAS Class I Area for the Base Model Period (2009-2013)[2]

Class I Area Site ID	Class I Area Name	IMPROVE Site ID	Natural Conditions 20% Most Impaired Days (dv)	Base Year (2009-2013) 20% Clearest Days (dv)	Base Year (2009-2013) 20% Most Impaired Days (dv)
CHAS	Chassahowitzka	CHAS1	8.97	13.76	19.94
COHU	Cohutta Wilderness	COHU1	9.52	10.94	21.19
DOSO	Dolly Sods Wilderness	DOSO1	8.92	9.03	21.59
EVER	Everglades NP	EVER1	8.34	11.23	16.30
GRSM	Great Smoky Mountains NP	GRSM1	10.05	10.63	21.39
JARI	James River Face Wilderness	JARI1	9.48	11.79	21.37
JOYC	Joyce-Kilmer-Slickrock Wilderness	GRSM1	10.05	10.63	21.39
LIGO	Linville Gorge Wilderness	LIGO1	9.70	9.70	20.39
MACA	Mammoth Cave NP	MACA1	9.79	13.69	24.04
OKEF	Okefenokee	OKEF1	9.47	13.34	20.70
OTCR	Otter Creek Wilderness	DOSO1	8.92	9.03	21.59
ROMA	Cape Romain	ROMA1	9.79	13.59	21.48
SAMA	St. Marks	SAMA1	9.19	13.33	20.11

---

[2] Tables 3-2 and 3-3, EPA, 2017b.

Table 1-2. Natural Conditions and Base Year Deciview (dv) Values on the 20% Clearest and 20% Most Impaired Days at each VISTAS Class I Area for the Base Model Period (2009-2013)[2]

Class I Area Site ID	Class I Area Name	IMPROVE Site ID	Natural Conditions 20% Most Impaired Days (dv)	Base Year (2009-2013) 20% Clearest Days (dv)	Base Year (2009-2013) 20% Most Impaired Days (dv)
SHEN	Shenandoah NP	SHEN1	9.52	8.60	20.72
SHRO	Shining Rock Wilderness*	LIGO1	9.70	9.70	19.91
SIPS	Sipsey Wilderness	SIPS1	9.55	12.84	21.67
SWAN	Swanquarter	SWAN1	9.79	11.76	19.76
WOLF	Wolf Island	OKEF1	9.47	13.34	20.70

* The base year model period dv value for the 20% most impaired days was adjusted to account for the effects that missing IMPROVE monitoring data for 2010 and 2011 may have on the base year model period dv. A ratio calculated for a nearby donor site, Linville Gorge Wilderness Area (i.e., the 5-year average of 2009-2013 IMPROVE data divided by the 3-year average of 2009, 2012, and 2013 IMPROVE data) was applied to the 3-year average of 2009, 2012, and 2013 data for the Shining Rock Wilderness Area to calculate the adjusted base year dv.

1.2.2 Purpose

This document serves as the first draft of the Air Quality Modeling Protocol plan for the VISTAS II Regional Haze modeling analysis to be performed by the contractor team Eastern Research Group, Inc. (ERG) and Alpine Geophysics, LLC (Alpine) with the purpose of estimating regional haze and progress at southeastern state Class I areas in projection year 2028. It is presumed that this information will be used by SESARM-participating states in the regional haze SIP development process.

1.3 Lead Agency and Principal Participants

SESARM is the lead agency in the development of this regional haze modeling analysis. EPA Region 4 in Atlanta, GA is the EPA Regional Office that will take the lead in the review and approval process for this project.

1.4 Related Regional Modeling Studies

There are other regional haze modeling studies related to the VISTAS II regional haze modeling analysis whose results may be useful to SESARM.

1.4.1 EPA Preliminary 2028 Regional Haze Modeling

EPA recently conducted preliminary visibility modeling (EPA, 2017b) for year 2028 with the intention of informing the regional haze SIP development process. For their assessment, air quality modeling was used to project visibility levels at individual Class I areas (represented by Interagency Monitoring of Protected Visual Environments (IMPROVE) monitoring sites) to 2028 and to estimate emissions sector contributions to 2028 particulate matter (PM) concentrations and visibility. The projected 2028 PM concentrations were converted to light extinction coefficients and then to dvs; these values are used to evaluate visibility progress as of

2028. In addition, 2028 visibility contribution information by major emissions source sector was calculated using particulate source apportionment technology (PSAT).

EPA found that visibility at most eastern Class I areas on the 20% most anthropogenically impaired days is projected to be below the unadjusted glidepath in 2028, with a relatively higher percentage of the light extinction due to domestic anthropogenic sources.

Based on their assessment of these results, EPA also identified a number of uncertainties and model performance issues that should be addressed in future EPA, state, multistate, or stakeholder modeling that may be used in SIP development. They have identified several aspects of this initial modeling that should be improved upon through coordination with interested stakeholders. These include, but are not limited to:

- *Expanded domain size* to reduce the impact of the boundary conditions assumptions on predictions, especially near the domain edge.
    - The boundary conditions were found to be the largest contributor to visibility at many Class I areas, especially those near the edge of the modeling domain. Expanding the regional photochemical modeling domain will potentially reduce the influence from global or hemispheric model derived boundary conditions. Those models have much coarser grid resolution and use global emissions inventories which may not be year specific or up to date.
    - There may also be recirculation of U.S. emissions in boundary conditions derived from global models, especially where the boundary is very close to the U.S. mainland. Moving the domain boundary further from the contiguous U.S. will minimize this issue.

- *Updated emission inventory and projections* for certain sectors.
    - More recent nationwide photochemical modeling has incorporated updates in future year emissions inventories that should be considered for 2028.
        - Remove the Clean Power Plan and Texas regional haze FIP from the electric generating unit (EGU) assumptions.
        - Updates to the oil and gas emissions projections.
        - New Canadian base and future year emissions.
        - Other emissions updates based on more recent information.

- *Updated boundary conditions* based on more recent modeling of international emissions as well as additional modeling to help quantify and distinguish anthropogenic and natural international contributions.
    - The 2011 boundary conditions used for the regional haze modeling came from an older version of Goddard Earth Observing System (GEOS)-Chem which did not contain the latest international emissions estimates for 2011.
    - In addition to projecting U.S. emissions to 2028, international emissions are changing between 2011 and 2028 as well. Consideration should be given to estimating future year global emissions to provide an alternate estimate of future year boundary conditions.

  - Global or hemispheric models can potentially be used to adjust the visibility glidepath for impacts from international anthropogenic sources. Sensitivity runs and additional refinements to international inventories may be needed in order to provide more confidence in the model results.

- *Improved treatment of fire and fugitive dust emissions* in the model.
  - The current Comprehensive air quality model with extensions (CAMx) modeling platform does not include estimates of natural windblown dust emissions. Windblown dust (primarily contributing to coarse mass) is an important component of regional haze in some Class I areas.
  - The current modeling used year-specific fire emissions from 2011 which may not be representative of a "typical year" or multi-year period. The IMPROVE measurements used to establish both the base period impairment measurements and progress towards natural conditions, use a five-year average of IMPROVE measurements. Therefore, alternative estimates of fire emissions, which may better represent a longer term average, may be more appropriate for use in visibility projections.
  - Further refinements of fire emissions may also allow exploration of possible adjustments of the glidepath for prescribed fire impacts.

- *Treatment of secondary organic aerosols (SOA)* should be reviewed.
  - In many locations, there is relatively high modeled SOA as a fraction of total organic aerosols. Using the RRF approach, this apportions the modeled SOA as a fraction of the measured total organics. There is considerable uncertainty in the modeled SOA concentrations, which therefore translates into uncertainty in the apportioned SOA mass.
  - Additional information can be gained by running PSAT with SOA source apportionment turned on.

- *Estimation of "natural visibility conditions"* used in the glidepath framework should be further reviewed.
  - Further refinements in the draft methodology can be explored.
  - Further analysis of the IMPROVE data combined with modeled source apportionment information may be useful in evaluating the natural conditions estimates.

All of these options were considered for this project; however due to schedule and resource constraints SESARM chose only to move forward with correcting the emission projections.

1.4.2 Maine Department of Environmental Protection Tracking Visibility Progress 2004-2016

This document (ME DEP, 2018), prepared by the Maine Department of Environmental Protection for the Mid-Atlantic/Northeast Visibility Union (MANE-VU), presents visibility trends at IMPROVE monitoring sites at federal Class I areas in and adjacent to the MANE-VU region that are subject to the USEPA's RHR. This document also presents visibility trends at IMPROVE Protocol monitoring sites in and adjacent to the MANE-VU region. The analyses were performed to determine the extent of progress in meeting short-term and long-term visibility goals for the first RHR SIP period that ends in 2018 using metrics specified in the state SIPs.

The technical document provides an analysis of visibility data collected at the IMPROVE monitoring sites, starting in the baseline period of 2000-2004 through 2012-2016, the most recent five-year period with available data.

The results of the analysis continue to show the following:

- There continue to be definite downward trends in overall haze level at all Class I areas in and adjacent to the MANE-VU region and at IMPROVE Protocol monitoring sites.
- Based on rolling five-year averages demonstrating progress since the 2000-2004 baseline period, the MANE-VU Class I areas are on track to meet their 2018 Reasonable Progress Goals (RPGs) for both 20% clearest and 20% most impaired visibility days.
- The trends are mainly driven by large reductions in sulfate light extinction, and to a lesser extent, nitrate light extinction.
- Levels of organic carbon mass (OCM) and light absorbing carbon (LAC) appear to be approaching natural background levels at most of the MANE-VU Class I areas.
- In all cases, the levels set by 2018 RPG’s have already been met, and progress beyond these goals appears achievable.
- The percent contribution of nitrate light extinction has been significantly increasing at most of the MANE-VU Class I areas.

1.5 Overview of Modeling Approach

The VISTAS II modeling proposed here will include particulate matter simulations and source apportionment studies using the 12 kilometer (km) grid based on EPA’s 2011/2028el modeling platform and preliminary source contribution assessment (EPA, 2017b) updated to include a 12km subdomain over the VISTAS region and augmented with revisions to EGU and non-EGU point source projections.

1.5.1 Episode Selection

Episode selection is an important component of any modeling analysis. EPA guidance recommends that one choose time periods which reflect the variety of meteorological conditions which represent visibility impairment on the 20% clearest and 20% most impaired days in the Class I areas being modeled (high and low concentrations necessary). This is best accomplished by modeling a full year. For this analysis, Alpine will be modeling the full 2011 calendar year with 10 days of model spin-up in 2010.

1.5.2 Model Selection

Details on the rationale for model selection are provided in Section 2. The Weather Research Forecast (WRF) prognostic meteorological model was selected for the VISTAS II modeling. Emissions processing will be completed using the Sparse Matrix Operator Kernel Emissions (SMOKE) model for most source categories. The exceptions are that Biogenic Emissions

Inventory System (BEIS) model was used for biogenic emissions and there are special processors for fires, windblown dust, lightning and sea salt emissions. The 2014 Motor Vehicle Emissions Simulator (MOVES) onroad mobile source emissions model was used by EPA with SMOKE-MOVES to generate onroad mobile source emissions with EPA generated vehicle activity data provided in the 2028 regional haze analysis. The CAMx photochemical grid model, which supports two-way grid nesting will also be used. The setup is based on the same WRF/SMOKE/BEIS/CAMx modeling system used in the EPA 2011/2028el platform modeling.

During the preparation of this Modeling Protocol, it was noted that a newer version of CAMx, Version 6.50, was released (April 30, 2018). After discussions with SESARM, the CC, and the TAWG, it was decided to not use the newer version due to insufficient testing and application history necessary for to ensure confidence in modelling results.

1.5.3 Base and Future Year Emissions Data

A 2011 base year and 2028 future year will be used for the VISTAS II modeling to be consistent with the requirements of EPA's RHR. The 2011 base case and 2028 future year emissions will be founded on EPA's "el" modeling platform with no adjustments made to 2011 estimates. Updates will be made to EGU and non-EGU point source data within the VISTAS states for 2028 and the Eastern Regional Technical Advisory Committee (ERTAC) EGU v.2.7 outputs will be used for non-VISTAS EGUs in the projection year.

1.5.4 Emission Input Preparation and QA/QC

Quality assurance (QA) and quality control (QC) of the emissions datasets are critical steps in performing air quality modeling studies. Because emissions processing is tedious, time consuming and involves complex manipulation of many different types of large databases, rigorous QA measures are a necessity to prevent errors in emissions processing from occurring. The VISTAS II modeling study will utilize methods applied to the emissions platform that follows a multistep emissions QA/QC approach.

1.5.5 Meteorology Input Preparation and QA/QC

The CAMx 2011 12 km meteorological inputs will be based on WRF meteorological modeling conducted by EPA. Details on the EPA 2011 WRF application and evaluation are provided by EPA (EPA 2014d).

1.5.6 Initial and Boundary Conditions Development

Initial concentrations (IC) and Boundary Conditions (BCs) are important inputs to the CAMx model. Alpine intends to run 10 days of model spin-up before the first days occur in the modeling episode so the ICs are washed out of the modeling domain before the first day of the annual 2011 modeling period. The lateral boundary and initial species concentrations are provided by a three-dimensional global atmospheric chemistry model, GEOS-Chem (Yantosca, 2004) standard version 8-03-02 with 8-02-01 chemistry.

1.5.7 Air Quality Modeling Input Preparation and QA/QC

Each step of the air quality modeling will be subjected to QA/QC procedures. These procedures will include verification of model configurations, confirmation that the correct data were used and were processed correctly and other procedures.

1.5.8 Model Performance Evaluation

An operational model performance evaluation will be performed for particulate matter ($PM_{2.5}$ species components and coarse PM) and regional haze to examine the ability of the CAMx v6.40 modeling system to simulate 2011 measured concentrations. This evaluation will focus on graphical analyses and statistical metrics of model predictions versus observations.

1.5.9 Diagnostic Sensitivity Analyses

Depending on the confirmation run results of the CAMx 2011 base case modeling and MPE on Alpine's modeling system, diagnostic sensitivity tests may be conducted to try and improve model performance. The definition of these diagnostic sensitivity tests will depend on the results of the initial MPE for these domains.

1.5.10 Future Year Significant Contribution Modeling

PM predictions from 2011 and 2028 CAMx model simulations will be used to project 2009-2013 IMPROVE visibility data to 2028 following the approach described in EPA's ozone, $PM_{2.5}$ and regional haze modeling guidance (US EPA, 2014b). The guidance describes the recommended modeling analysis used to help set RPGs that reflect emissions controls in a regional haze SIP. The CAMx PSAT method will be utilized for this effort.

1.6 Project Participants and Contacts

SESARM is the lead agency in the development of the VISTAS II modeling analysis. They will work closely with other local agencies, other local cities and agencies, and EPA Region 4 in the study, including the sharing of interim results as they become available. SESARM will also work with local, state, and tribal agencies and stakeholders in the modeling analysis, where stakeholders may include environmental groups and industry. Key participants in the VISTAS II study and their contact information are provided in Table 1-3.

Table 1-3. Key Participants and Contact Information for the VISTAS II Modeling Study

Organization	Individual(s) [Role]	Address	Contact Numbers
Southeastern States Air Resource Managers, Inc. (SESARM)			
Metro 4/ SESARM	Mr. John Hornback SESARM Executive Director	Southeastern States Air Resource Managers, Inc. 205 Corporate Center Drive, Suite D Stockbridge, GA 30281-7383	(404) 361-4000 hornback@metro4-sesarm.org
EPA Region 4			
EPA Region 4	Mr. Rick Gillam Environmental Engineer	EPA Region 4 Sam Nunn Atlanta Federal Center 61 Forsyth St., SW Atlanta, GA 30303	(404) 562-9049 gillam.rick@epa.gov
Contractors (Modeling team)			
Eastern Research Group, Inc.	Mr. Regi Oommen ERG Program Manager	Eastern Research Group, Inc. 1600 Perimeter Park Drive, Suite 200	(919) 468.7829 regi.oommen@erg.com

Table 1-3. Key Participants and Contact Information for the VISTAS II Modeling Study

Organization	Individual(s) [Role]	Address	Contact Numbers
		Morrisville, NC 27560	
Eastern Research Group, Inc.	Ms. Bebhinn Do ERG Senior Scientist	Eastern Research Group, Inc. 1600 Perimeter Park Drive, Suite 200 Morrisville, NC 27560	(919) 468-7840 bebhinn.do@erg.com
Alpine Geophysics, LLC	Mr. Gregory Stella Alpine Subcontract Manager	Senior Scientist 387 Pollard Mine Road Burnsville, NC 28714	(828) 675-9045 gms@alpinegeophysics.com
Alpine Geophysics, LLC	Mr. Dennis McNally Alpine Senior Scientist	Senior Scientist 7341 Poppy Way Arvada, CO 80007	(303) 421- 2211 dem@alpinegeophysics.com

1.7 Communication

Frequent communication between SESARM and the Modeling team and other participants is anticipated. These communications will include e-mails, conference calls, and face-to-face meetings. SESARM envisions that EPA and others will review interim products as they become available so that comments can be received during the study to allow for corrective action as necessary. These interim deliverables would include, but not be limited to, preliminary CAMx model performance evaluation, preliminary current and future-year emissions assumptions and results, and preliminary future year visibility projections and source apportionment results.

1.8 Schedule

Table 1-4 lists the current schedule for the initial key deliverables under the VISTAS II modeling study.

Table 1-4. Initial Key Deliverables and Dates for the VISTAS II Modeling Through the 2028 Source Apportionment Reporting

Task	Subtask	Deliverable	Due Date
2	2.1	2011 Base year emissions inventories	6/30/2018[a]
	2.2a	Projection Year Emissions Inventory Comparisons, draft	5/18/2018
	2.2b	Projection Year Emissions Inventories, final	6/30/2018[a]
	2.3a1	2028 EGU Point Source Emissions, draft	5/18/2018
	2.3a2	2028 EGU Point Source Emissions, final	6/30/2018[a]
	2.3b1	2028 Non-EGU Point Source Emissions, draft	5/18/2018
	2.3b2	2028 Non-EGU Point Source Emissions, final	6/30/2018[a]
	2.3c1	2028 Emissions for Other Categories, draft	5/18/2018
	2.3c2	2028 Emissions for Other Categories, final	6/30/2018[a]
	2.3d1	Emission Comparisons from 2028v6.3el and 2023v6.3en, draft	5/18/2018
	2.3d2	Emission Comparisons from 2028v6.3el and 2023v6.3en, final	6/30/2018[a]
	2.3e1	2028 Documentation, draft	5/18/2018
	2.3e2	2028 Documentation, final	6/30/2018[a]
	2.4	Emissions summaries and Quality Assurance	6/30/2018[a]
3	3.1	Create Photochemical Model-Ready EGU emissions files for 2028	7/13/2018[a]
	3.1.2	Scale Hourly EGU SMOKE emissions to match annual 2028	7/13/2018[a]
	3.2	Create Photochemical Model-Ready Non-EGU emissions files for 2028	7/13/2018[a]
	3.3a	Merge EGU/non-EGU data from Subtasks 3.1 and 3.2 for CAMx Model	7/13/2018[a]
	3.3b*	Merge area/MAR data from Subtasks 3.1 and 3.2 for CAMx Model	7/13/2018[a]
4	4	Data acquisition and preparation	6/1/2018
	4.1	Acid deposition in watersheds	6/1/2018
5	5	Area of Influence Analysis	9/1/2018
	5.1	SO_2 and NO_x emissions contribution rankings	No later than 9/1/2018
6	6.1a	Modeling protocol, draft	5/2/2018
	6.1b	Modeling protocol, final	6/30/2018[a]
	6.2	2011 base year air quality modeling	9/1/2018
	6.3	2028 projection year air quality modeling	12/1/2018
7	7	Source apportionment tagging - 250 tags (final number to be determined)	4/19/2019[a]
8	8	Model performance evaluation	10/1/2018
	8.1	Model performance evaluation (related to Subtask 4.1)	10/1/2018
9	9a	Future-year model projections – 250 tags (final number to be determined)	4/19/2019
	9.1	Calculate Relative Response Factors (related to Subtask 4.1)	5/3/2019
10	10	Website/FTP Site Development; Data Transfer and Archival	Ongoing
11	11.1a	Extraction of state-specific modeling initial conditions/boundary conditions (5 states)	Within 1 week after completion of Task 6.1 and 6.2 activities

Table 1-4. Initial Key Deliverables and Dates for the VISTAS II Modeling Through the 2028 Source Apportionment Reporting

Task	Subtask	Deliverable	Due Date
	11.1b*	Additional extraction of state-specific modeling initial conditions/boundary conditions (5 states)	Within 1 week after completion of Task 6.1 and 6.2 activities
	11.2a	Extraction of state-specific meteorological files (5 states)	Within 1 week after regions are defined by the time the meteorological data is windowed for the VISTAS_12 domain
	11.2b*	Additional extraction of state-specific meteorological files (5 states)	Within 1 week after regions are defined by the time the meteorological data is windowed for the VISTAS_12 domain

[a] Revised dates based on Amendment 1, dated 5/31/2018

* Optional Task, no decision yet whether to perform it.

2.0 MODEL SELECTION

This section introduces the models to be used in the VISTAS II regional haze modeling study. The selection methodology presented in this chapter mirrors EPA's preliminary 2028 regional haze modeling (EPA, 2017b).

40 CFR Part 51 Appendix W does not identify a "preferred model" for use in reasonable progress assessments for regional haze and therefore, EPA does not recommend a specific model for use in these analyses. The latest EPA modeling guidance (EPA, 2014e) explicitly mentions the Community Multiscale Air Quality Modeling System (CMAQ) and CAMx photochemical grid models (PGMs) as the most commonly used PGMs that would satisfy EPA's selection criteria but notes that this is not an exhaustive list and does not imply that they are "preferred" over other PGMs that could also be considered and used with appropriate justification. EPA's current modeling guidelines lists the following criteria for model selection (EPA, 2014e):

- It should not be proprietary;
- It should have received a scientific peer review;
- It should be appropriate for the specific application on a theoretical basis;
- It should be used with data bases which are available and adequate to support its application;
- It should be shown to have performed well in past modeling applications;
- It should be applied consistently with an established protocol on methods and procedures;
- It should have a User's guide and technical description;
- The availability of advanced features (e.g., probing tools or science algorithms) is desirable; and
- When other criteria are satisfied, resource considerations may be important and are a legitimate concern.

Alpine and ERG have been directed by SESARM to use EPA's 2011el-based air quality modeling platform which includes emissions, meteorology, and other inputs for 2011 as the base year for the modeling described in their regional haze TSD (EPA, 2017b). EPA has projected the 2011 base year emissions to a 2028 future year base case scenario. This will be the foundation of the emissions with revisions for this analysis as described elsewhere. The 2011 modeling platform and projected 2028 emissions will be used to drive the 2011 base year and 2028 base case air quality model simulations. As noted in EPA's documentation, the 2011 base year emissions and methods for projecting these emissions to 2028 are in large part similar to the data and methods used by EPA in the final Cross-State Air Pollution Rule (CSAPR) Update[3] and the subsequent notice of data availability (NODA)[4] to support ozone transport for the 2015 ozone NAAQS.

---

[3] https://www.epa.gov/airmarkets/final-cross-state-air-pollution-rule-update

[4] https://www.epa.gov/airmarkets/notice-data-availability-preliminary-interstate-ozone-transport-modeling-data-2015-ozone

The 2011 and 2028 emissions used for EPA's regional haze modeling are described in the documents:

- "Preparation of Emissions Inventories for the Version 6.3, 2011 Emissions Modeling Platform";[5]
- "Technical Support Document (TSD) Updates to Emissions Inventories for the Version 6.3 2011 Emissions Modeling Platform for the Year 2028";[6] and
- "EPA Base Case v.5.16 for 2023 Ozone Transport NODA Using IPM Incremental Documentation."[7]

The meteorological data and initial and boundary concentrations used for this regional haze assessment, as described below, are the same as those used for the Final CSAPR Update air quality modeling and the 2015 ozone transport NAAQS NODA.

For the VISTAS II regional haze modeling analysis, Alpine and ERG propose to use the WRF/SMOKE/MOVES2014/BEIS/CAMx-PSAT modeling system as the primary tool for modeling PM concentrations and visibility and in the calculation of significant contribution to downwind Class I areas. The proposed modeling system satisfies all of EPA's selection criteria. The key models to be used in this regional haze modelling effort are described below:

- WRF/ARW: The Weather Research and Forecasting (WRF)[8] Model is a mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs (Skamarock, 2004; 2006; Skamarock et al., 2005). The Advanced Research WRF (ARW) version of WRF will be used in this regional haze analysis study. It features multiple dynamical cores, a 3-dimensional variational (3DVAR) data assimilation system, and a software architecture allowing for computational parallelism and system extensibility. WRF is suitable for a broad spectrum of applications across scales ranging from meters to thousands of kilometers. The effort to develop WRF has been a collaborative partnership, principally among the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (NOAA), the National Centers for Environmental Prediction (NCEP) and the Forecast Systems Laboratory (FSL), the Air Force Weather Agency (AFWA), the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration (FAA). WRF allows researchers the ability to conduct simulations reflecting either real data or idealized configurations. WRF provides operational forecasting a model that is flexible and efficient computationally, while offering the advances in physics, numerics, and data assimilation contributed by the research community.
- MOVES2014: MOVES2014[9] is EPA's latest onroad mobile source emissions model that was first released in July 2014 (EPA, 2014a; 2014b; 2014c). MOVES2014 includes the

---

[5] https://www.epa.gov/air-emissions-modeling/2011-version-63-technical-support-document

[6] https://www.epa.gov/air-emissions-modeling/2011-version-63-platform

[7] https://www.epa.gov/airmarkets/epa-base-case-v516-2015-ozone-naaqs-transport-noda-using-ipm-incremental-documentation

[8] http://www.wrf-model.org/index.php

[9] http://www.epa.gov/otaq/models/moves/

latest onroad mobile source emissions factor information. Emission factors developed by EPA will be used in this analysis.

- SMOKE: The SMOKE[10] modeling system is an emissions modeling system that generates hourly gridded speciated emission inputs of mobile, nonroad, nonpoint area, point, fire and biogenic emission sources for photochemical grid models (Coats, 1995; Houyoux and Vukovich, 1999). As with most 'emissions models', SMOKE is principally an emission processing system and not a true emissions modeling system in which emissions estimates are simulated from 'first principles'. This means that, with the exception of mobile and biogenic sources, its purpose is to provide an efficient, modern tool for converting an existing base emissions inventory data into the hourly gridded speciated formatted emission files required by a photochemical grid model. SMOKE will be used to prepare emission inputs for nonroad mobile, nonpoint area, and point sources.
- SMOKE-MOVES[11]: SMOKE-MOVES uses an Emissions Factor (EF) Look-Up Table from MOVES, county-level gridded vehicle miles travelled (VMT) and other activity data, and hourly gridded meteorological data (typically from WRF) to generate hourly gridded speciated onroad mobile source emissions inputs.
- ERTAC EGU Forecasting Tool: The ERTAC EGU Forecasting Tool[12] was developed through a collaborative effort to improve emission inventories among the Northeastern, Mid-Atlantic, Southeastern, and Lake Michigan area states; other member states; industry representatives; and multi-jurisdictional organization (MJO) representatives. The tool can be used to grow base year hourly EGU emissions inventories into future projection years. The tool uses base year hourly EPA Clean Air Markets Division (CAMD) data, fuel specific growth rates, and other information to estimate future emissions.
- BEIS: Biogenic emissions were modeled by EPA using version 3.61 of BEIS. First developed in 1988, BEIS estimates volatile organic compound (VOC) emissions from vegetation and nitric oxide (NO) emissions from soils. Because of resource limitations, recent BEIS development has been restricted to versions that are built within the SMOKE system.
- CAMx: The CAMx[13] model is a state-of-science "One-Atmosphere" photochemical grid model capable of addressing ozone, PM, visibility, and acid deposition at a regional scale for periods up to one year (Ramboll Environ, 2016). CAMx is a publicly-available open-source computer modeling system for the integrated assessment of gaseous and particulate air pollution. Built on today's understanding that air quality issues are complex, interrelated, and reach beyond the urban scale, CAMx is designed to: (a) simulate air quality over many geographic scales; (b) treat a wide variety of inert and chemically active pollutants including ozone, inorganic and organic $PM_{2.5}$ and PM_{10} and mercury and toxics; (c) provide source-receptor, sensitivity, and process analyses; and (d) be computationally efficient and easy to use. The U.S. EPA has approved the use of CAMx for numerous ozone, PM, and regional haze SIPs throughout the U.S. and has used this model to evaluate regional mitigation strategies including those for most recent

---

[10] http://www.smoke-model.org/index.cfm

[11] https://www.epa.gov/sites/production/files/2016-06/documents/smoke-moves-2011.pdf

[12] http://www.marama.org/2013-ertac-egu-forecasting-tool-documentation

[13] http://www.camx.com

regional-scale rules (e.g., Transport Rule, Clean Air Interstate Rule (CAIR), NO_X SIP Call, etc.). CAMx Version 6.40 will be used in this study, with the secondary organic aerosol partitioning (SOAP) algorithm module as the default.

- During the preparation of this Modeling Protocol, it was noted that a newer version of CAMx, Version 6.50, was released (April 30, 2018). After discussions with SESARM, the CC, and the TAWG, it was decided to not use the newer version due to insufficient testing and practical usage necessary for to ensure confidence in potential modelling results.
- PSAT: The PSAT tool of CAMx was selected to develop source contribution and significant contribution calculations.

3.0 EPISODE SELECTION

EPA's most recent regional haze modeling guidance (EPA, 2014e) contains recommended procedures for selecting modeling episodes. The VISTAS II regional haze modeling will use the annual calendar year 2011 modeling period because it satisfies the most criteria in EPA's modeling guidance episode selection discussion and is consistent with the base year modeling platform. Specifically, EPA's guidance recommends choosing a time period which reflects the variety of meteorological conditions that represent visibility impairment on the 20% clearest and 20% most-impaired days in the Class I areas being modeled (high and low concentrations necessary). This is best accomplished by modeling a full calendar year.

4.0 MODELING DOMAIN SELECTION

This section summarizes the modeling domain definitions for the VISTAS II regional haze modeling, including the domain coverage, resolution, and map projection. It also discusses emissions, aerometric, and other data available for use in model input preparation and performance testing.

4.1 Horizontal Domains

The VISTAS II modeling will use a 12 km continental U.S. (CONUS_12 or 12US2) domain. The 12 km nested grid modeling domain configuration is shown in Figure 4-1. The 12 km domain shown in Figure 4-1 represents the CAMx 12km air quality and SMOKE/BEIS emissions modeling domain. The WRF meteorological modeling was run on larger 12 km modeling domains than used for CAMx as demonstrated in EPA's meteorological model performance evaluation document (EPA, 2014d). The WRF meteorological modeling domains are defined larger than the air quality modeling domains because meteorological models can sometimes produce artifacts in the meteorological variables near the boundaries as the prescribed boundary conditions come into dynamic balance with the coupled equations and numerical methods in the meteorological model.

An additional VISTAS_12 domain will be prepared that is a subset of the CONUS_12 domain, with dimensions for both provided in Table 4-1. Development of the VISTAS_12 domain (also presented in Figure 4-1) will require the EPA CONUS_12 simulation to be run using CAMx Version 6.40 modeling saving 3-dimensional concentration fields for extraction using the CAMx BNDEXTR program.

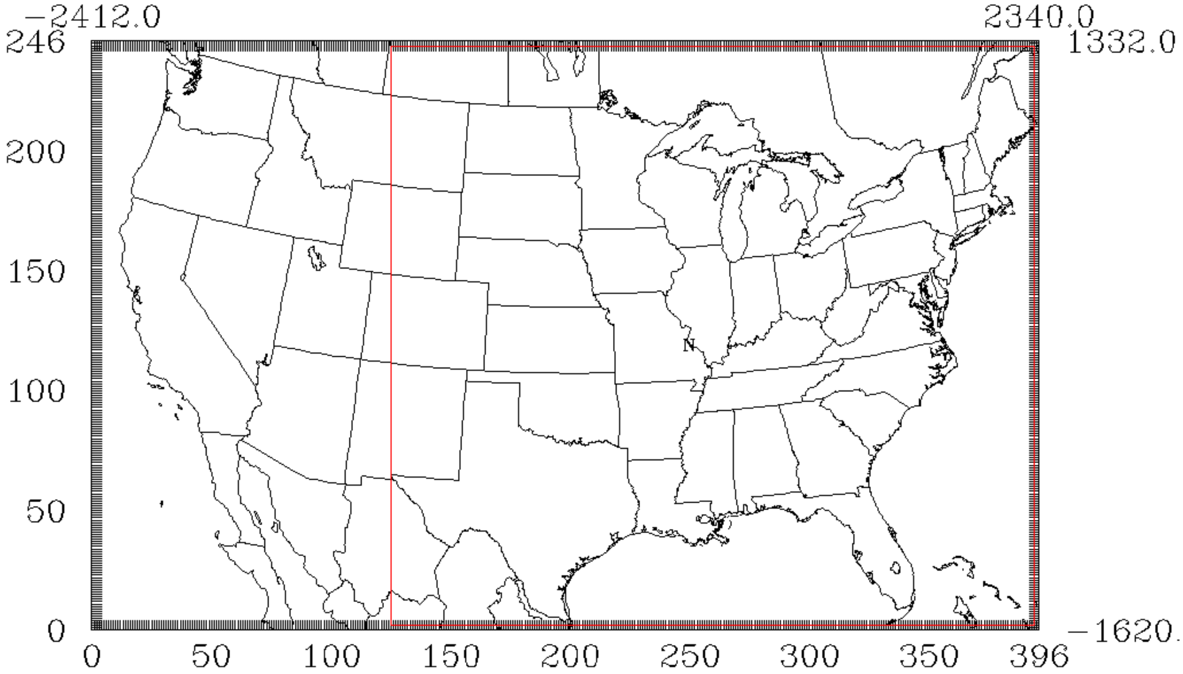


Figure 4-1. Map of 12km CAMx Modeling Domains. VISTAS_12 Domain Represented as Inner Red Domain.

Table 4-1. VISTAS II Modeling Domain Specifications

Domain	Columns	Rows	Vertical Layers	X Origin (km)	Y Origin (km)
CONUS_12	396	246	25	-2,412	-1,620
VISTAS_12	269	242	25	-912	-1,596

4.2 Vertical Modeling Domain

The CAMx vertical structure is primarily defined by the vertical layers used in the WRF meteorological modeling. The WRF model employs a terrain following coordinate system defined by pressure, using multiple layer interfaces that extend from the surface to 50 mb (approximately 19 km above sea level). EPA ran WRF using 35 vertical layers. A layer averaging scheme is adopted for CAMx simulations whereby multiple WRF layers are combined into one CAMx layer to reduce the air quality model computational time. Table 4-2 displays the approach for collapsing the WRF 35 vertical layers to 25 vertical layers in CAMx and is consistent with EPA's draft 2028 regional haze modeling.[14]

Table 4-2. WRF and CAMx Layers and Their Approximate Height Above Ground Level

CAMx Layer	WRF Layers	Sigma P	Pressure (mb)	Approx. Height (m AGL)
25	35	0.00	50.00	17,556
	34	0.05	97.50	14,780
24	33	0.10	145.00	12,822
	32	0.15	192.50	11,282
23	31	0.20	240.00	10,002
	30	0.25	287.50	8,901
22	29	0.30	335.00	7,932
	28	0.35	382.50	7,064
21	27	0.40	430.00	6,275
	26	0.45	477.50	5,553
20	25	0.50	525.00	4,885
	24	0.55	572.50	4,264
19	23	0.60	620.00	3,683
18	22	0.65	667.50	3,136
17	21	0.70	715.00	2,619
16	20	0.74	753.00	2,226
15	19	0.77	781.50	1,941
14	18	0.80	810.00	1,665
13	17	0.82	829.00	1,485
12	16	0.84	848.00	1,308
11	15	0.86	867.00	1,134
10	14	0.88	886.00	964
9	13	0.90	905.00	797
	12	0.91	914.50	714
8	11	0.92	924.00	632

---

[14] Table 2-2, EPA, 2017b.

Table 4-2. WRF and CAMx Layers and Their Approximate Height Above Ground Level

CAMx Layer	WRF Layers	Sigma P	Pressure (mb)	Approx. Height (m AGL)
	10	0.93	933.50	551
7	9	0.94	943.00	470
	8	0.95	952.50	390
6	7	0.96	962.00	311
5	6	0.97	971.50	232
4	5	0.98	981.00	154
	4	0.99	985.75	115
3	3	0.99	990.50	77
2	2	1.00	995.25	38
1	1	1.00	997.63	19

5.0 DATA AVAILABILITY

The CAMx modeling systems requires emissions, meteorology, surface characteristics, initial and boundary conditions (IC/BC), and ozone column data for defining the inputs.

5.1 Emissions Data

Without exception, as directed by SESARM, the 2011 base year emissions inventories for this analysis will be based on emissions obtained from the EPA's "el" modeling platform. This platform was prepared by EPA and used in the regional haze modeling for 2028. Emissions for 2028 will be based on a mixture of EPA's 2028el sectors and updated emissions as prepared by SESARM for inclusion in this study.

5.2 Ambient Air Quality Observations

Year 2011 data from all available ambient air monitoring networks for gas and PM species are used in the model performance evaluation. Table 5-1 summarizes routine ambient gaseous and PM monitoring networks available in the U.S. Alpine will focus on the ambient data collected from the IMPROVE network. This network began in 1985 as a cooperative visibility monitoring effort between EPA, federal land management agencies, and state air agencies (IMPROVE, 2011). Data are collected at Class I areas across the United States mostly at National Parks, National Wilderness Areas, and other protected pristine areas. Currently, there are approximately 181 IMPROVE sites that have complete annual $PM_{2.5}$ mass and/or $PM_{2.5}$ species data. There are 110 IMPROVE monitoring sites which represent air quality at the 156 designated Class I areas. The 71 additional IMPROVE sites are "IMPROVE protocol" sites which are generally located in rural areas throughout the U.S. Although these sites use the IMPROVE monitoring samplers and collection routines, they are not located at Class I areas.

5.3 Meteorological Data

The 2011 meteorological data for the air quality modeling of 2011 and 2028 will be derived from EPA's run of Version 3.4 of the WRF Model (Skamarock, et al., 2008) as completed for their regional haze modeling analysis for 2028 (EPA, 2017b).

5.4 Ozone Column Data

Ozone column data were derived by EPA using the Total Ozone Mapping Spectrometer (TOMS) with aboard Ozone Monitoring Instrument (OMI) satellite platform[15] for 2011 at a daily, 1 degree resolution. These data will be used as obtained directly from EPA for the 12US2 simulations. For the VISTAS_12 domain simulations, the data will be windowed onto the smaller VISTAS_12 domain.

5.5 Photolysis Rates

Photolysis rates were calculated using the NCAR Tropospheric Ultraviolet and Visible (TUV) Radiation Model Version 4.8. These data will be used as obtained directly from EPA.

5.6 Land Use

Land use and land cover data is based on the 2006 National Land Cover Database (NLCD2006) data.[16] These data will be used as obtained directly from EPA for the 12US2 simulations. For the VISTAS_12 domain simulations, the data will be windowed onto the smaller region.

5.7 Initial and Boundary Conditions Data

The lateral boundary and initial species concentrations are provided by a three-dimensional global atmospheric chemistry model, GEOS-Chem (Yantosca, 2004) standard version 8-03-02 with 8-02-01 chemistry. The global GEOS-Chem model simulates atmospheric chemical and physical processes driven by assimilated meteorological observations from the NASA's GEOS-5.[17] This model was run for 2011 with a grid resolution of 2.0 degrees x 2.5 degrees (latitude-longitude). The predictions were used to provide one-way dynamic boundary concentrations at one-hour intervals and an initial concentration field for the CAMx simulations. The 2011 boundary concentrations from GEOS-Chem were used for the 2011 and 2028 model simulations. The procedures for translating GEOS-Chem predictions to initial and boundary concentrations are described elsewhere (Henderson, 2014). More information about the GEOS-Chem model and other applications using this tool is available at: http://www-as.harvard.edu/chemistry/trop/geos.

---

[15] https://ozoneaq.gsfc.nasa.gov/data/

[16] The 2006 NLCD data are available at http://www.mrlc.gov/nlcd06_data.php

[17] Additional information available at: http://gmao.gsfc.nasa.gov/GEOS/ and http://wiki.seas.harvard.edu/geos-chem/index.php/GEOS-5

Table 5-1. Overview of Routine Ambient Data Monitoring Networks

Monitoring Network	Chemical Species Measured	Sampling Period	Data Availability/Source
The Interagency Monitoring of Protected Visual Environments (IMPROVE)	Speciated $PM_{2.5}$ and PM_{10} (see species mappings); light extinction data	1 in 3 days; 24-hour average	http://vista.cira.colostate.edu/improve/Data/IMPROVE/improve_data.htm
Clean Air Status and Trends Network (CASTNET)	Speciated $PM_{2.5}$, and O_3 (see species mappings)	Approximately 1-week average	http://www.epa.gov/castnet/data.html
National Atmospheric Deposition Program (NADP)	Wet deposition (hydrogen (acidity as pH), sulfate, nitrate, ammonium, chloride, and base cations (such as calcium, magnesium, potassium and sodium)), Mercury	1-week average	http://nadp.sws.uiuc.edu/
Air Quality System (AQS)	CO, NO_2, O_3, SO_2, $PM_{2.5}$, PM_{10}, lead (Pb), HAPs	Typically hourly average to 24-hour average	http://www.epa.gov/air/data/
Chemical Speciation Network (CSN)	Speciated $PM_{2.5}$	24-hour average	http://www.epa.gov/ttn/amtic/amticpm.html
Photochemical Assessment Monitoring Stations (PAMS)	Varies for each of 4 station types.	Varies by station	http://www.epa.gov/ttn/amtic/pamsmain.html
National Park Service Gaseous Pollutant Monitoring Network	Acid deposition (Dry; SO_4^{2-}, NO_3^-, HNO_3, NH_4, SO_2), O_3, meteorological data	Hourly	http://www2.nature.nps.gov/ard/gas/netdata1.htm

6.0 MODEL INPUT PREPARATION PROCEDURES

This section summarizes the procedures to be used in developing the meteorological, emissions, and air quality inputs to the CAMx model for the VISTAS II modeling on the 12 km grids for the annual 2011 period. The 12 km CAMx modeling databases are largely based on the EPA "el" platform databases.

6.1 Meteorological Inputs

The meteorological inputs in EPA's "el" platform will be used directly without revision. Details of EPA's annual 2011 meteorological model simulation and evaluation are provided in a separate technical support document (EPA, 2014a), which can be obtained at: http://www.epa.gov/ttn/scram/reports/MET_TSD_2011_final_11-26-14.pdf.

6.1.1 WRF Model Science Configuration

Version 3.4 of the WRF model ARW core (Skamarock, 2008) was used for generating the 2011 simulations. Selected physics options include Pleim-Xiu land surface model, Asymmetric Convective Model Version 2 planetary boundary layer scheme, KainFritsch cumulus parameterization utilizing the moisture-advection trigger (Ma and Tan, 2009), Morrison double moment microphysics, and Rapid Radiative Transfer Model-Global (RRTMG) longwave and shortwave radiation schemes (Iacono et al., 2008). The WRF model configuration was prepared by EPA (EPA, 2014d).

6.1.2 WRF Input Data Preparation Procedures

The WRF model simulation was initialized using the 12km North American Model (NAM-12) analysis product provided by the National Climatic Data Center (NCDC). Where NAM-12 data were unavailable, the 40km Eta Data Assimilation System (EDAS) analysis (ds609.2) from the National Center for Atmospheric Research (NCAR) was used. Analysis nudging for temperature, wind, and moisture was applied above the boundary layer only. The model simulations were conducted in 5.5 day blocks with soil moisture and temperature carried from one block to the next via the Intermediate Processor for Pleim–Xiu for WRF (IPXWRF) program (Gilliam and Pleim, 2010). Land use and land cover data were based on the 2006 National Land Cover Database (NLCD2006) data.[18] Sea surface temperatures at 1 km resolution were obtained from the Group for High Resolution Sea Surface Temperatures (GHRSST) (Stammer, et al., 2003).

As shown in Table 4-2, the WRF simulations were performed with 35 vertical layers up to 50 mb, with the thinnest layers being nearest the surface to better resolve the planetary boundary layer (PBL). The WRF 35-layer structure was collapsed to 25 layers for the CAMx air quality model simulations, as shown in Table 4-2.

6.1.3 WRF Model Performance Evaluation

The WRF model performance evaluation is provided in EPA's Meteorological Model Performance for Annual 2011 WRF v3.4 Simulation documentation report (EPA, 2014d). The WRF model evaluation approach was based on a combination of qualitative and quantitative analyses. The quantitative analysis was divided into monthly summaries of 2-m temperature, 2-m

---

[18] The 2006 NLCD data are available at http://www.mrlc.gov/nlcd06_data.php

mixing ratio, and 10-m wind speed using the boreal seasons to help generalize the model bias and error relative to a set of standard model performance benchmarks. The qualitative approach was to compare spatial plots of model estimated monthly total precipitation with the monthly PRISM precipitation.

6.1.4 WRFCAMx/MCIP Reformatting Methodology

The WRF meteorological model output data was processed to provide inputs for the CAMx photochemical grid model. The WRFCAMx processor maps WRF meteorological fields to the format required by CAMx. It also calculates turbulent vertical exchange coefficients (Kv) that define the rate and depth of vertical mixing in CAMx. The methodology used by EPA to reform the meteorological data into CAMx format is provided in documentation provided with the wrfcamx conversion utility[19].

The meteorological data generated by the WRF simulations were processed by EPA using WRFCAMx v4.3 (Ramboll Environ, 2014) meteorological data processing program to create model-ready meteorological inputs to CAMx. In running WRFCAMx, vertical eddy diffusivities (Kv) were calculated using the Yonsei University (YSU) (Hong and Dudhia, 2006) mixing scheme. Alpine used a minimum Kv of 0.1 m^2/sec except for urban grid cells where the minimum Kv was reset to 1.0 m^2/sec within the lowest 200 m of the surface in order to enhance mixing associated with the nighttime “urban heat island” effect. In addition, EPA invoked the subgrid convection and subgrid stratoform stratiform cloud options in our wrfcamx run for 2011.

6.1.5 Windowing from EPA 12US2 to VISTAS_12

The meteorological data will be windowed from the EPA 12US2 domain onto the VISTAS_12 domain using a slightly modified version of the CAMx utility program “window”.[20] The only required change to the distributed version program is to allow the program to window three-dimension files instead of just two-dimensional files.

6.2 Emission Inputs

6.2.1 Available Emissions Inventory Datasets

The base year emission inventories to be used in the VISTAS II modeling study will be based on EPA’s 2011el modeling platform without exception. Complete documentation for the 2011 emissions used for EPA’s regional haze modeling are described in the documents, “Preparation of Emissions Inventories for the Version 6.3, 2011 Emissions Modeling Platform.”

Emissions for the 2028 base year will include EPA 2028el projections for most sectors (onroad and nonroad mobile sources, marine, aircraft, railroad, fires, nonpoint area, biogenic, and international sources) augmented with updated EGU and non-EGU point source emission estimates provided by SESARM.

Documentation on EPA’s 2028el platform can be found in “Technical Support Document (TSD) Updates to Emissions Inventories for the Version 6.3 2011 Emissions Modeling Platform for the

---

[19] http://www.camx.com/getmedia/7f3ee9dc-d430-42d6-90d5-dedb3481313f/wrfcamx-11jul17.tgz

[20] http://www.camx.com/getmedia/88755b80-6992-4f07-bcaa-596d05e1b4b8/window-6may13_1.tgz

Year 2028". Adjustments planned for the SESARM adjusted source categories are documented below.

Outside of the VISTAS states, source categories and their associated emissions will be taken directly from the EPA 2028el modeling platform. The exception to this will be for EGU sources where data obtained from the projections to 2028 for EGU point source emissions by the ERTAC EGU projection tool[21] from the most recent CONUS 2.7 run will be used.

6.2.2 2011 Base Year Emissions

The emissions data in the 2011 platform are primarily based on the 2011NEIv2 for point sources, nonpoint sources, commercial marine vessels (CMV), nonroad mobile sources and fires. The onroad mobile source emissions are similar to those in the 2011NEIv2, but were generated using the released 2014a version of the Motor Vehicle Emissions Simulator (MOVES2014a). Fugitive dust emissions from anthropogenic sources (i.e., agricultural tilling and unpaved roads) are included in the nonpoint sector of the inventory, but wind-blown dust from natural sources is not accounted for in the inventory.

2011 CAMx-ready emission inputs were generated by EPA mainly by the SMOKE and BEIS emissions models. CAMx requires two emission input files for each day: (1) low level gridded emissions that are emitted directly into the first layer of the model from sources at the surface with little or no plume rise; and (2) elevated point sources (stacks) containing stack parameters from which the model can calculate plume rise. As noted earlier, EPA's 2011el emission platform in CAMx-ready format will be used without exception.

6.2.3 2028 Projection Year Emissions

Certain 2011 emission sectors were also projected by EPA to 2028 using various sector dependent methodologies. Onroad and nonroad mobile source emissions were created for 2028 using the MOVES and NONROAD models, respectively. Nonpoint area source emissions were prepared using growth and control factors simulating changes in economic conditions and environmental regulation anticipated to be fully implemented by calendar year 2028. For these categories, Alpine will be using EPA's estimates from the 2028el platform, unless a VISTAS state provides an update that is authorized by the SESARM contract.

Projections to 2028 for EGU and non-EGU point source emissions for 2028 will be derived from files produced by the ERTAC EGU projection tool from the most recent CONUS 2.7 run and from SESARM review of EPA 2028el emissions augmented with state-provided growth and control adjustments, respectively. ERG will review the emissions adjustments and work with states to ensure that the revised values are reasonable. ERG will also document the source of the revised 2028 emissions.

ERG will prepare state-specific summary comparisons of EPA's 2028v6.3el modeling platform emissions to the 2023v6.3en modeling platform emissions for stationary electricity generating unit (EGU) and non-EGU stationary point sources to facilitate review by each VISTAS state. The summaries between the two inventories will be grouped by Emission Inventory System

---

[21] http://www.marama.org/2013-ertac-egu-forecasting-tool-documentation

(EIS) facility, emissions unit, process, and release point identifiers and source classification code (SCC) for annual emissions of:

- Oxides of nitrogen (NO_x);
- Volatile organic compounds (VOC);
- Primary particulate matter less than or equal to 2.5 microns in aerodynamic diameter ($PM_{2.5}$-PRI);
- Primary particulate matter less than or equal to 10 microns in aerodynamic diameter (PM_{10}-PRI);
- Carbon monoxide (CO);
- Sulfur dioxide (SO_2); and
- Ammonia (NH_3).

If the 2023v6.3en emissions are selected for the final 2028 emissions inventory, ERG will document these updates.

ERG will work with SESARM on the final format of the comparison tables, including additional fields that may be useful for review, such as: facility information; SCC descriptions; unit, process, and release point descriptions; ORIS boiler identifiers; control information; and absolute and percentage differences between the two emissions inventories. All data will be provided in a single Excel file, unless the file size is prohibitive, at which point ERG will work with SESARM on the best way to divide the data across multiple files.

6.2.3.1 EGU Point Source Emissions

For EGU sources only, the ERG/Alpine team will use an already-obtained version of the 2028 emissions forecast and associated files produced by the ERTAC EGU projection tool from the most recent CONUS 2.7 run available. The team will prepare state-specific summary comparisons of EPA's 2028v6.3el modeling platform emissions to EPA's 2023v6.3en modeling platform emissions to the ERTAC 2028 modeling platform emissions for EGU point sources to facilitate the VISTAS state review. The summaries will be produced in Microsoft Excel format grouped by EIS facility, emissions unit, process, and release point identifiers, ORIS ID, and SCC for annual emissions of NO_X, VOC, $PM_{2.5}$-PRI, PM_{10}-PRI, CO, SO_2, and NH_3 between the two emissions inventories.

ERG will work with SESARM on the final format of the comparison tables, including additional fields that may be useful for review, such as: facility information; SCC descriptions; unit, process, and release point descriptions; ORIS boiler identifiers; control information; and absolute and percentage differences between the two emissions inventories.

SESARM will identify for ERG in the final comparison tables which emissions projection platform (e.g., EPA 2023en, EPA 2028el, ERTAC EGU, or state provided) should be used in the final 2028 modeling file preparation. For any one EGU source, only a single platform should be selected. In other words, emissions from one platform cannot be mixed with emissions from another platform at the same unit.

6.2.3.2 Non-EGU Point Source Emissions

Similar to the work being conducted for EGU point sources, ERG will update the 2028 non-EGU point source projection year mass emissions inventories based on the information collected from SESARM's review of prepared comparison tables. All revisions will be documented to account for changes in emissions due to retirements, control enhancements, and/or fuel switches, as well as any additional metadata to describe the data. For certain situations, a state may wish to develop their own revised 2028 point sources emissions inventory using updated growth and/or control factors on the 2011 point sources emissions inventory. In these cases, ERG will work with the state agencies to provide the data into the format needed for integration.

6.2.3.3 Nonpoint Area, Onroad, and Nonroad Mobile Source Emissions

Emissions data for 2028 nonpoint area, onroad mobile, and nonroad mobile sources will be used from the 2011/2028el modeling platform. These sources will be spatially allocated to the grid using an appropriate surrogate distribution (e.g., population for home heating, etc.) and will be temporally allocated by month, by day of week, and by hour of day using the EPA source-specific temporal allocation factors. The SMOKE source-specific CB6r4 speciation allocation profiles will also be used for all categories.

6.2.3.4 Biogenic Source Emissions

Biogenic emissions were generated by EPA using the BEIS biogenic emissions model within SMOKE. BEIS uses high resolution GIS data on plant types and biomass loadings and the WRF surface temperature fields, and solar radiation (modeled or satellite-derived) to develop hourly emissions for biogenic species on the 12 km grids. BEIS generates gridded, speciated, temporally allocated emission files. EPA's 2011 biogenic emissions will be used for the 2028 modeling platform.

6.2.3.5 Wildfires, Prescribed Burns, Agricultural Burns

Fire emissions in 2011NEIv2 were developed based on Version 2 of the Satellite Mapping Automated Reanalysis Tool for Fire Incident Reconciliation (SMARTFIRE) system (Sullivan, et al., 2008). SMARTFIRE2 was the first version of SMARTFIRE to assign all fires as either prescribed burning or wildfire categories. In past inventories, a significant number of fires were published as unclassified, which impacted the emissions values and diurnal emissions pattern. Recent updates to SMARTFIRE include improved emission factors for prescribed burning. These emissions as prepared by EPA will be included in both the 2011 and 2028 emission platforms. EPA's 2011 fire emissions will be used for the 2028 modeling platforms. It should be noted that there were large wildfires in the Okefenokee Swamp and Eastern North Carolina in 2011. However, contributions from these wildfires should not impact the development of reasonable progress goals since EPA's new 20% most impaired days metric removes IMPROVE days that were significantly impacted by these wildfires from the analysis.

6.3 Emissions Processing

CAMx requires detailed emissions inventories containing temporally allocated (i.e., hourly) emissions for each grid-cell in the modeling domain for a large number of chemical species that act as primary pollutants and precursors to secondary pollutants. Annual emission inventories for 2011 and 2028 will be preprocessed into SMOKE-ready modeling system (Houyoux et al., 2000) inputs for eventually additional processing and use in CAMx. For this analysis, CAMx will be operated using Version 6 revision 4 of the Carbon Bond chemical mechanism (CB6r4).

For emission segments that are unchanged from the EPA distribution, the emission will be windowed from the EPA 12US2 domain onto the VISTAS_12 domain using the CAMx "window"[22] utility program. For VISTAS state specific emission segments, the emissions will be developed on the VISTAS_12 domain.

Steps necessary to prepare SMOKE-ready input files from the mass emissions data is outlined in the VISTAS II workplan, Task 2. In that task, Alpine will ensure that annual emission changes submitted or authorized by SESARM are carried through to all relevant input files consistent with EPA's 2011 and 2028 v6.3el platform processing. Upon the completed development of these new input files, Alpine will use EPA's modeling platform scripts, with the updated input files from this task, to generate CAMx photochemical model Version 6.40-ready inputs using the SMOKE Modeling System.

The CMAQ and CAMx models require hourly emissions of specific gas and particle species for the horizontal and vertical grid cells contained within the modeled region (i.e., modeling domain). To provide emissions in the form and format required by the model, it is necessary to "pre-process" the "raw" emissions (i.e., emissions input to SMOKE) for the sectors described above. In brief, the process of emissions modeling transforms the emissions inventories from their original temporal resolution, pollutant resolution, and spatial resolution into the hourly, speciated, gridded resolution required by the air quality model.

The emission processing for this project will take two different "paths" as shown in Figure 6-1.

For emission segments that are being updated for the project, the emissions will be processed using SMOKE version 3.7 with the emissions being output directly into CAMx format for the VISTAS_12 domain. For emission segments that are unchanged from the EPA simulation the emissions will be converted from the EPA supplied CMAQ format into CAMx format using the CMAQtoCAMx processor and windowed onto the VISTAS_12 domain using the "WINDOW" program. The CMAQ2CAMx converter reformats the files, converts units, and transforms the CMAQ in-line elevated emissions into CAMx elevated format and low level files into CAMx format. As a final step before running the CAMx model, the low-level emissions will be merged into the single low level and elevated files.

---

[22] http://www.camx.com/getmedia/88755b80-6992-4f07-bcaa-596d05e1b4b8/window-6may13_1.tgz

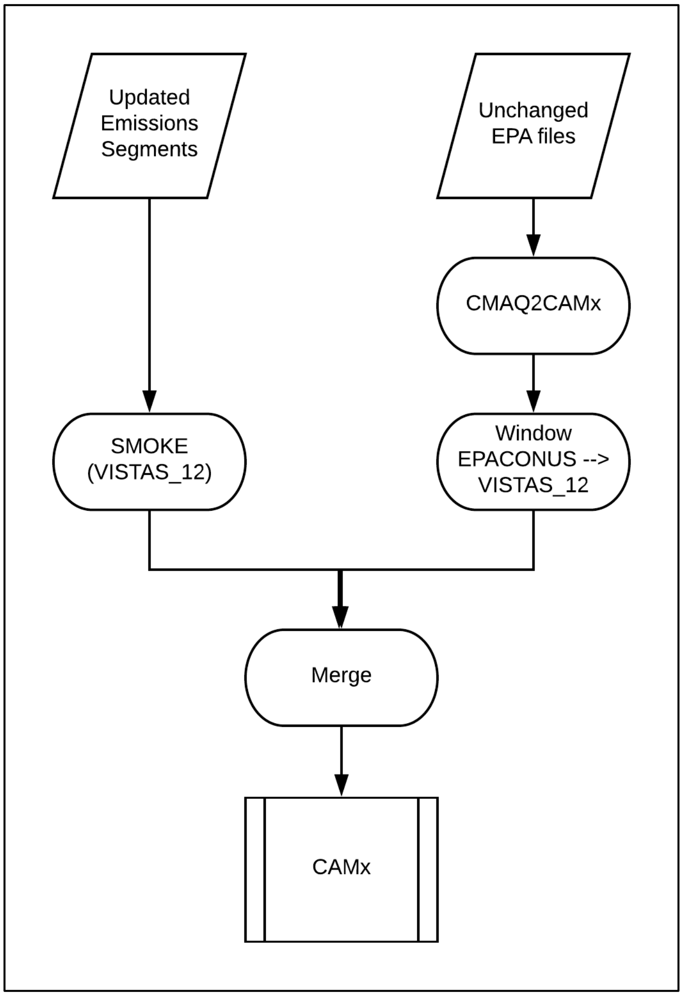


Figure 6-1. Emission Processing Paths

While the 2028 projection year non-EGU and EGU point source inventory will be updated to reflect requested changes by SESARM, the steps associated with the emissions processing of these files remains the same as EPA's methods for the 2028 regional haze modeling analysis. Details of the temporal, spatial, and speciation data and methods are described in EPA's technical support document for the development of the 2028 platform.[23]

Other than the preparation of data for an additional modeling domain (VISTAS_12) for this analysis, all other emissions processing steps, methods, and ancillary data and associations will be identical to EPA's documented processing steps.

No plume-in-grid (PiG) subgrid-scale plume treatment will be used in this project.

Scripts to perform the emissions merging of the appropriate biogenic, onroad, nonroad, nonpoint area, low-level, fire, and point emission files will be written to generate the CAMx-ready two-dimensional day and domain-specific hourly speciated gridded emission inputs. The point source

---

[23] https://www.epa.gov/sites/production/files/2017-11/documents/2011v6.3_2028_update_emismod_tsd_oct2017.pdf

and, as available elevated fire, emissions would be processed into the day-specific hourly speciated emissions in the CAMx-ready point source format.

6.3.1 QA/QC of CAMx-Ready Emission Files

For quality assurance of the emissions modeling steps, emissions totals for all species across the entire model domain will be output as reports that are then compared to reports generated by SMOKE on the input inventories to ensure that mass is not lost or gained during the emissions modeling process.

In addition to the CAMx-ready emission input files generated for each hour of all days modeled in the annual 2011 or 2028 modeling period, a number of QA files may be prepared and used to check for gross errors in the emissions inputs. The model-ready emissions will be imported into visualization tools and Alpine will examine both the spatial and temporal distribution of the emission to investigate the quality and accuracy of the emissions inputs.

- Visualizing the model-ready emissions with the scale of the plots set to a very low value, Alpine can determine whether there are areas omitted from the raw inventory or if emissions sources are erroneously located in water cells;
- Spot-checking the holiday emissions files to confirm that they are temporally allocated like Sundays;
- Producing pie charts emission summaries that highlight the contribution of each emissions source component (e.g. nonroad mobile); and
- Normalizing the emissions by population for each state will illustrate where the inventories may be deficient and provide a reality check of the inventories.

State inventory summaries prepared prior to the emissions processing will be used to compare against SMOKE output report totals generated after each major step of the emissions generation process. To check the chemical speciation of the emissions to CB6 species, Alpine will compare reports generated with SMOKE to target these specific areas of the processing. For speciation, the inventory state import totals will be compared against the same state totals with the speciation matrix applied.

The quantitative QA analyses often reveal significant deficiencies in the input data or the model setup. It may become necessary to tailor these procedures to track down sources of any identified major problem. As such, Alpine can only outline the basic quantitative QA steps that are performed by Alpine to reveal the underlying problems with the inventories or processing.

Standard inventory assessment methods may be employed to generate the future year emissions data QA files including, but not limited to: (a) visualizing the model-ready emissions graphically, (b) spot-checking the holiday emissions files to confirm that they are temporally allocated like Sundays, (c) producing pie charts emission summaries for each source category, and (d) normalizing the emissions by population for each state to reveal where the future year inventories may be suspect. Of particular importance will be the comparison of the 2011 base year and 2028 future year emissions by source category to make sure the expected changes have been appropriately accounted for in the modeling inventories.

The resultant CAMx model-ready emissions will be subjected to a final QA using spatial maps to assure that: (1) the emissions were merged properly; (2) CAMx inputs contain the same total emissions; and (3) to provide additional QA/QC information.

Emissions are processed by major source category in several different "streams", including nonpoint area sources, onroad mobile sources, nonroad mobile sources, biogenic sources, non-CEM point sources, CEM point sources using day-specific hourly emissions, and emissions from fires. Separate Quality Assurance (QA) and Quality Control (QC) will be performed for each stream of emissions processing and in each step following the procedures outlined in the project QAPP. SMOKE includes advanced quality assurance features that include error logs when emissions are dropped or added. In addition, Alpine will generate visual displays that include:

- Spatial plots of the hourly emissions for each major species (e.g., NO_X, VOC, some speciated VOC, SO_2, NH_3, PM and CO).
- Summary tables of emissions for major species for each grid and by major source category.
- This QA information will be examined against the original point and nonpoint area source data and summarized in an overall QA/QC assessment.

6.4 Photochemical Modeling Inputs

6.4.1 CAMx Science Configuration and Input Configuration

This section describes the model configuration and science options to be used in the VISTAS II modeling effort. Version 6.40 of CAMx will be used for this modeling.

CAMx is a three-dimensional grid-based Eulerian air quality model designed to simulate the formation and fate of oxidant precursors, primary and secondary particulate matter concentrations, and deposition over regional and urban spatial scales (e.g., the contiguous U.S.). Consideration of the different processes (e.g., transport and deposition) that affect primary (directly emitted) and secondary (formed by atmospheric processes) pollutants at the regional scale in different locations is fundamental to understanding and assessing the effects of emissions on air quality concentrations.

Figure 4-1 presented the geographic extent of the modeling domains (12US2 and VISTAS_12) that will be used for air quality modeling in this analysis. The 12US2 domain covers the 48 contiguous states along with the southern portions of Canada and the northern portions of Mexico. The VISTAS_12 domain covers the continental US eastward from the western extent of Texas along with the southern portions of Canada and the northern portions of Mexico. As discussed later, the limited coverage of Canada and Mexico is an important consideration when interpreting the modeling results. This modeling domain contains 25 vertical layers with a top at about 17,550 meters, or 50 millibars (mb), and horizontal grid resolution of 12 km x 12 km. The model simulations produce hourly air quality concentrations for each 12 km grid cell across the modeling domain.

For the simulations that use the EPA 12US2 domain, the model will be applied using the same time segments as EPA. The model will be applied from December 22, 2010 through April 30, 2011 and from April 21, 2011 through December 31, 2011. The beginning of each segment, December 21-31, 2010 for the first segment and April 21-30, 2011 for the second segment are used as spin-up periods and will not be analyzed.

For the simulation on the VISTAS_12 domain, the model will be applied quarterly using the definitions in Table 6-1. The initial conditions for each quarter, and hourly boundary conditions will be extracted from the CAMx version 6.40 simulations performed over the EPA 12US2 CONUS domain.

Table 6-1. VISTAS II Simulation Periods

Quarter Number	Starting Date	Ending Date
1	December 22, 2010	March 31, 2011
2	March 15, 2011	June 30, 2011
3	June 15, 2011	September 30, 2011
4	September 15, 2011	December 31, 2011

CAMx requires a variety of input files that contain information pertaining to the modeling domain and simulation period. These include gridded, hourly emissions estimates and meteorological data, and boundary concentrations. Separate emissions inventories will be prepared for the 2011 base year and the 2028 base case. Meteorological fields will be specified for the 2011 base year model application and remained unchanged for the future-year model simulations.[24] The IC/BC for VISTAS_12 in 2028 will be based on 12US_2 domain as mentioned in section 5.5.2.

Ten CAMx model runs, associated with up to 250 tagged sources, are currently planned for this analysis. The simulations are summarized in Table 6-2.

Table 6-2. CAMx Simulations for the VISTAS II Project

Simulation Number	Description	CAMx Version	Grid	Task	Time
1	EPA 2011el Confirmation	6.32	12US2	6.2	EPA 1,2*
2	EPA 2011el saving 3-D average for VISTAS12 IC/BC Extracts	6.40	12US2	6.2	EPA 1,2
3	EPA 2028el Confirmation	6.32	12US2	6.3	EPA 1,2
4	VISTAS 2028 saving 3-D average for VISTAS12 IC/BC extracts	6.40	12US2	6.3	EPA 1,2
5	EPA 2011el saving 3-D average for state domains IC/BC extracts	6.40	VISTAS_12	6.2	Q1,2,3,4
6	VISTAS 2028 saving 3-D average for state domain IC/BC extracts	6.40	VISTAS_12	6.3	Q1,2,3,4
7	VISTAS 2028 PSAT with selected sources 1-50 tagged	6.40	VISTAS_12	7	Q1,2,3,4

---

[24] The CAMx annual simulations for 2011 and 2028 were each performed using two time segments (January 1 through April 30, 2011 with a 10-day ramp-up period at the end of December 2010 and May 1 through December 31, 2011 with a 10-day ramp-up period at the end of April 2011).

Table 6-2. CAMx Simulations for the VISTAS II Project

Simulation Number	Description	CAMx Version	Grid	Task	Time
8	VISTAS 2028 PSAT with selected sources 51-100 tagged	6.40	VISTAS_12	7	Q1,2,3,4
9	VISTAS 2028 PSAT with selected sources 101-150 tagged	6.40	VISTAS_12	7	Q1,2,3,4
10	VISTAS 2028 PSAT with selected sources 151-200 tagged	6.40	VISTAS_12	7	Q1,2,3,4
11	VISTAS 2028 PSAT with selected sources 201-250 tagged	6.40	VISTAS_12	7	Q1,2,3,4

*EPA twice annual segments from December 22, 2010 through April 30, 2011 and April 21, 2011 through December 31, 2011.

6.4.2 VISTAS_12 Boundary And Initial Conditions

Boundary and initial conditions for the VISTAS_12 domain will be extracted from simulations using the EPA 12US2 domain. For the 2011 baseline simulation, the EPA 2011el platform for the 12US2 domain will be run using CAMx 6.40 and the resultant three-dimensional outputs will be saved. For the 2028 future year simulation, the EPA 2028el point source emissions will be reprocessed to include updated point source emissions, and CAMx 6.40 will be run saving three-dimensional outputs. The boundary and initial conditions for the VISTAS_12 domain will be extracted from the respective three-dimensional output files using the CAMx BNDEXTR program.

6.4.3 VISTAS_12 Ozone Column

The ozone column data for the VISTAS_12 domain will be windowed from the EPA 12US2 domain. The ozone column data is a simple text file and new code will be developed to extract data for the smaller VISTAS_12 domain.

6.4.4 Photolysis Rates

The EPA photolysis rates used will be unchanged in the VISTAS_12 domain.

6.4.5 CAMx Land Use

The land use file for the VISTAS_12 domain will be windowed from the EPA 12US2 domain. The land use file is a simple text file and new code will be developed to extract data for the smaller VISTAS_12 domain.

6.5 EPA 2011 and 2028 Base Case Confirmation

The numerics in photochemical grid models are very complex and it is typical to get slightly different model concentrations based on the version of the computer and compilers. When comparing simulations, it is critical to isolate the changes in concentrations to the changes in the model inputs, and not on the computing details (i.e., compiler version, computer architecture, parallelization options). This is especially problematic when looking at particulate matter, since the particulate treatments have multiple pathways, and small concentration differences can lead to different pathways through the code and different concentrations.

Sources of the difference can come from the options used in CAMx compilation, the version of the compiler, the compiler vendor, and how the model calculation is split onto different processors (parallelization).

Alpine will execute two confirmation runs, one for the 2011el base year and one for the 2028el base case, to confirm the contract team's ability to replicate EPA's results and to ensure that the EPA data, models, and scripts operated in a consistent manner as EPA's procedure.

6.5.1 Differences Between EPA And VISTAS Simulations

EPA ran the 2011v6.3el platform on EPA's supercomputer with the model configured to use four (4) processor nodes with 16 processors per node. The use of multiple processor nodes with multiple processors per node is efficient on the EPA supercomputer due to the low latency interconnect between the nodes. On more typical computer clusters with the nodes interconnected with Ethernet, like the Alpine cluster and most likely the State and stakeholder clusters, the latency between nodes is sufficiently high that it is inefficient to spread processing between nodes. Our experience with the EPA platform has shown that on an Ethernet connected cluster with 12 Intel XEON processors per node and hyperthreading enabled it is most efficient to use a single node configured with 10 Message Passing Interface (MPI) instances, each with 2 OpenMP threads.

EPA used the Intel FORTRAN compiler. Alpine, and the CAMx developers, use the Portland Group (PGI) FORTRAN compiler. The PGI compiler has been the standard compiler for CAMx applications for many years and it's anticipated this compiler will be more widely used by the States and stakeholders. The version of CAMx 6.32 EPA distributed with the 2011el platform will be recompiled on the Alpine computer system and used for the confirmation.

As noted in Section 5.5.1, EPA ran the model in two time segments. The first segment, typically used only for PM applications, runs from December 22, 2010 through April 30, 2011. The second segment runs from April 21, 2011 through December 31, 2011. The VISTAS confirmation run will use the same two segments. December 22-31, 2010 and the April portion of the second segment are spin-ups and will not be analysed.

6.5.2 Confirmation Methodology

The simulations on the Alpine computer cluster and the EPA computer will be based on hourly differences in ozone, $PM_{2.5}$, POC, Particulate Nitrate and Particulate Sulfate. The metric for comparison will be the absolute difference (Equation 1) and percent difference (Equation 2) defined as:

(Equation 1) $$\left(C_{vistas} - C_{epa}\right)$$

(Equation 2) $$\frac{\left(C_{vistas} - C_{epa}\right)}{\left(C_{epa}\right)}$$

Where C_{epa} is the concentration at each grid cell hour for the EPA simulation and C_{vistas} is the concentration at each grid cell hour for the simulation on the Alpine computers.

The comparison will be done both graphically (e.g., scatter density plots) and quantitatively (e.g., residual distributions) for reviewed concentrations. Analysis products will be hourly spatial plots of the absolute differences. Should significant differences be noted between the confirmation runs and EPA's original simulations, Alpine will generate an electronic appendix of the spatial plots and discuss with SESARM and others as requested. If it is determined that noted differences are the result of modeling errors, Alpine will propose to SESARM recommended approaches to correct the issues and will correct configurations based on negotiation action.

The following confirmation runs will be performed:

1. EPA 2011 with CAMx_6.32 (CONUS) vs. Alpine 2011 with CAMx_6.32 (CONUS)
2. EPA 2028 with CAMx_6.32 (CONUS) vs. Alpine 2028 with CAMx_6.32 (CONUS)
3. Alpine 2011 with CAMx_6.32 (CONUS) vs. Alpine 2011 with CAMx_6.40 (CONUS)
4. Alpine 2028 with CAMx_6.32 (CONUS) vs. Alpine 2028 with CAMx_6.40 (CONUS)
5. Alpine 2011 with CAMx_6.40 (CONUS) vs. Alpine 2011 with CAMx_6.40 (VISTAS)
6. Alpine 2028 with CAMx_6.40 (CONUS) vs. Alpine 2028 with CAMx_6.40 (VISTAS)

7.0 MODEL PERFORMANCE EVALUATION

An operational model performance evaluation will be conducted for total fine particulate matter ($PM_{2.5}$), $PM_{2.5}$ species components (e.g., SO_4, NO_3, NH_4, EC, and OC), measured gas phase species (e.g., O_3, HNO_3, SO_2, VOCs, NO_2, and NO), and speciated components of light extinction to examine the ability of the CAMx v6.40 modeling system to simulate 2011 measured concentrations. This evaluation will focus on graphical analyses and statistical metrics of model predictions versus observations. Note that as EPA did not publish an evaluation of $PM_{2.5}$ gas phase species in their 2028 regional haze modeling we will not have a basis for comparison of these metrics.

The CAMx 2011 base case model estimates will be compared against the observed concentrations to establish that the model is capable of reproducing the current year observed concentrations so that it can be considered a reliable tool for estimating future year PM and regional haze levels.

The model evaluation will focus on the ability of the model to predict visibility-reducing PM at VISTAS Class I areas (represented by IMPROVE monitoring sites). The analysis will look at monthly and seasonal average PM species component performance at IMPROVE and other PM monitoring networks, and performance on the 20% most impaired (and 20% clearest) days at individual IMPROVE sites. This will provide a comprehensive assessment of the components that make up visibility performance.

The measured concentrations of PM components such as nitrate on the 20% most impaired days at many Class I areas are extremely small. Numerous Class I areas in the Eastern U.S. have average nitrate observations (on the 20% most impaired days) of less than 1 $\mu g/m^3$. This makes it challenging to correctly model observed visibility. Assumptions regarding particular emissions categories and boundary conditions can have a large impact on model performance. Even when model performance appears to be accurate, it is difficult (without further modeling and analysis) to determine if the answers observed are right for the right reasons (for example, when the extinction is dominated by modeled boundary conditions).

7.1 Model Performance Evaluation

7.1.1 Overview of EPA Model Performance Evaluation Recommendations

EPA current (EPA, 2007) and draft (EPA, 2014e) ozone, PM, and regional haze modeling guidance recommendations for model performance evaluation (MPE) describes a MPE framework that has four components:

- Operation evaluation that includes statistical and graphical analysis aimed at determining how well the model simulates observed concentrations (i.e., does the model get the right answer).
- Diagnostic evaluation that focuses on process-oriented evaluation and whether the model simulates the important processes for the air quality problem being studied (i.e., does the model get the right answer for the right reason).

- Dynamic evaluation that assess the ability of the model air quality predictions to correctly respond to changes in emissions and meteorology.
- Probabilistic evaluation that assess the level of confidence in the model predictions through techniques such as ensemble model simulations.

EPA's guidance notes that there is no single definitive test for evaluating model performance. All of the tests mentioned here have strengths and weaknesses. Further, even with a single performance test, it is not appropriate to assign "bright line" criteria that distinguish between adequate and inadequate model performance. In this regard, EPA recommends that a "weight of evidence" approach be used to determine whether a particular modeling application is valid for assessing the future attainment status of an area.

EPA recommends that air agencies conduct a variety of performance tests and weigh them qualitatively to assess model performance. Provided suitable databases are available, greater weight should be given to those tests which assess the model capabilities most closely related to how the model is used in the analysis (i.e. tests that provide insight into the accuracy of the model's relative response to emissions reductions). Generally, additional confidence should be attributed to model applications in which a variety of the tests described here are applied and the results indicate that the model is performing well.

From an operational standpoint, EPA recommends that air agencies compare their evaluation results against similar modeling exercises to ensure that the model performance approximates the quality of other applications. Recent literature reviews (Simon et al, 2012, Emery 2017) summarized photochemical model performance for applications published in the peer-reviewed literature between 2006 and 2012.

Because this 2011 VISTAS II modeling is using a CAMx 2011 modeling database developed by EPA, Alpine will also include by reference the air quality modeling performance evaluation as conducted by EPA (EPA, 2017b) on the national 12km domain.

Alpine will review EPA's current operational MPE for particulate matter ($PM_{2.5}$ species components, coarse PM, and total $PM_{2.5}$) and light extinction (total and species components) to compare the ability of the CAMx v6.40 modeling system to simulate 2011 measured concentrations. Using a combination of the Atmospheric Model Evaluation Tool (AMET) and internal scripts used by Alpine, comprehensive MPE statistics and graphics from the 2011 CAMx simulation using data from the IMPROVE network will be prepared in formats that will be accessible for stakeholder review and use. Alpine will use the current IMPROVE equation (see Section 8) inclusive of the monthly relative humidity function [f(RH)] values for both observed and modeled data to develop performance statistics at each IMPROVE monitor in the VISTAS_12 domain.

7.1.2 VISTAS II Calculated Model Evaluation Statistics

In order to estimate the ability of CAMx to replicate the 2011 base year concentrations of particulate matter, an operational model performance evaluation will be conducted. For this evaluation, mean bias and normalized mean bias, mean error and normalized mean error, and Pearson's correlation coefficient will be used and directly compared to EPA's results using these same statistics. Mean fractional bias, mean fractional error, and root mean squared error may also be calculated but as these statistics were not calculated by EPA in its evaluation, there will be no basis for comparison for these metrics.

Mean bias (MB) is the average difference between predicted (P) and observed (O) concentrations for a given number of samples (n):

$$MB(\mu g\ m^{-3}\ or\ Mm^{-1}) = \frac{1}{n}\sum\nolimits_{i=1}^{n}(P_i - O_i)$$

Mean error (ME) is the average absolute value of the difference between predicted and observed concentrations for a given number of samples:

$$ME(\mu g\ m^{-3}\ or\ Mm^{-1}) = \frac{1}{n}\sum\nolimits_{i=1}^{n}|P_i - O_i|$$

Normalized mean bias (NMB) is the sum of the difference between predicted and observed values divided by the sum of the observed values:

$$NMB(\%) = \frac{\sum_1^n(P - O)}{\sum_1^n(O)} * 100$$

Normalized mean error (NME) is the sum of the absolute value of the difference between predicted and observed values divided by the sum of the observed values:

$$NME(\%) = \frac{\sum_1^n|P - O|}{\sum_1^n(O)} * 100$$

Pearson's correlation coefficient (r) is defined as:

$$r = \frac{\sum_{i=1}^{n}(P_i - \overline{P})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{n}(P_i - \overline{P})^2}\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2}}$$

Mean Fractional Bias (MFB) is defined as:

$$MFB(\%) = \frac{2}{N}\sum_{1}^{N}\left(\frac{P - O}{P + O}\right) \times 100$$

Mean Fractional Error (MFE) is defined as:

$$MFE(\%) = \frac{2}{N}\sum_{1}^{N}\left(\frac{|P - O|}{P + O}\right) \times 100$$

Root Mean Squared Error (RMSE) is defined as:

$$RSME = \sqrt{\frac{\sum_1^N (P - O)^2}{N}}$$

Model predictions of PM species will be paired in space and time with observational data from the IMPROVE monitoring network. These results will be organized by IMPROVE monitor and season (winter (DJF), spring (MAM), summer (JJA), and fall (SON)).

ERG/Alpine expects the model performance of the replicated 2011 CAMx run to be slightly different from EPA published MPE metrics due to the differences in the version of CAMx being used and the modeling domain. If performance is not comparable to EPA's MPE, then data files will be reviewed to determine the cause. If the difference is not explainable by the changes in domain or model version, a call will be convened between ERG, Alpine, SESARM, and the appropriate EPA staff to identify any inconsistencies and come to consensus on appropriate corrective actions. Any of the metrics outside published proposed criteria levels will be noted as part of the uncertainty associated with the modeling.

Additional statistical analysis may also be performed, as determined necessary. All statistics will be calculated consistent with the respective pollutants NAAQS averaging time.

Tables and plots will be prepared for VISTAS and identified non-VISTAS IMPROVE monitors as directed to demonstrate light extinction model performance in a graphical manner. Scatter (with linear regression and r^2 value), bugle, and soccer plots for all light extinction and speciated components will be developed for the 20% most impaired and 20% clearest days for each IMPROVE monitor in the VISTAS_12 modeling domain.

Alpine will develop the individual day-by-day and site-by-site stacked bar plots of total b_{ext} and speciated components of b_{ext} for these most impaired and clearest days. Metrics and plots generated for the most impaired days will be consistent with the latest definition of this classification as "anthropogenically impaired" days as defined by EPA.

7.1.3 Model Performance Evaluation For Weekly Wet And Weekly Dry Deposition Species

The modeling team will also perform a MPE for weekly wet deposition and weekly dry deposition species collected in workplan Subtask 4.1. For this MPE, VISTAS CAMx deposition values will be aggregated to appropriate time periods to match the various NADP monitoring network's concentration collection times. To prevent confounding the MPE, the networks with different collection time (i.e., biweekly versus weekly) will be examined separately.

For wet deposition, NADP networks typically present measurements as concentration in mg/L, which is equivalent to g/m^3. These concentrations are then multiplied by the precipitation in meters to yield wet deposition rates in units of g/m^2. The CAMx wet deposition outputs are provided in grams per hectare (g/ha), which will be converted to grams per meter squared (g/m^2), using the conversion of 1 ha = 10,000 m^2, to have consistent units with the NADP monitoring networks. CAMx estimates of wet deposition can also be adjusted to account for the error present in the model estimated precipitation through a ratio of the observed to estimated precipitation.[25] Dry deposition values from CASTNET can be developed from the observed concentration multiplied by a deposition velocity generated by the Multi-Layer Model (MLM)[26] for the site. The MLM generated deposition velocities are available for download with the CASTNET observations.[27]

Annual mean MPE statistics, like the statistics for the base year MPE, will be developed for the wet deposition and dry deposition species available. Analysis will also include scatter plots of NADP network observations versus CAMx predictions, and their correlation (r), both annually and by season. Statistical and scatter plots will also be examined by VISTAS states to provide more refined MPE information to facilitate further use by the states.

Additionally, annual deposition totals will be produced from the VISTAS II base year modeling and compared to the annual Total Deposition Maps developed by the NADP and EPA. These total deposition maps are produced via a hybrid approach that combines the monitored data with modeled data to produce a gridded map of total sulfate and nitrate depositions. While not entirely observed truth, these hybrid estimates could provide the ability to evaluate generally the MPE for the entire domain in areas where data availability is limited due to incomplete records from the monitoring sites.

7.2 Performance Goals and Benchmarks

Establishment of performance goals and benchmarks for regulatory modeling is a necessary but difficult activity. Here, performance goals refer to targets that Alpine believes a good performing model should achieve, whereas performance benchmarks are based on historical model performance measures for the best performing simulations. Performance goals are necessary in order to provide consistency in model applications and expectations across the region and to provide standardization in how much weight may be accorded modeling study results in the decision-making process. These performance goals should not be interpreted as a “bright line” where metrics below the value are judged as adequate and values above the line are unusable. Rather they should be interpreted as the performance for a particular metric with good performance should be more highly relied upon for decision making where a less well performing pollutant may be less relied upon.

It is a problematic activity, though, because many areas present unique challenges and no one set of performance goals is likely to fit all needs. Equally concerning is the very real danger that modeling studies will be truncated when the ‘statistics look right’ before full assessment of the model’s reliability is made. This has the potential from breeding built-in compensating errors as modelers strive to get good statistics as opposed to searching for the explanations for poor performance and then rectifying them.

---

[25] Appel, K. W., et al. 2011. "A multi-resolution assessment of the Community Multiscale Air Quality (CMAQ) model v4. 7 wet deposition estimates for 2002–2006." Geoscientific Model Development 4.2 (2011): 357-371.

[26] Meyers, T. P., Finkelstein, P., Clarke, J., Ellestad, T.G., and Sims, P.F. 1998. A Multilayer Model for Inferring Dry Deposition Using Standard Meteorological Measurements. J. Geophys. Res., 103(D17): 22,645-22,661, DOI: 10.1029/98jd01564.

[27] https://java.epa.gov/castnet/clearsession.do

For this analysis, Alpine will pair model predictions in space and time with observational data from the IMPROVE, CSN, and CASTNET monitoring networks. These results will be compared by network and season (winter (Dec, Jan, Feb), spring (Mar, Apr, May), summer (Jun, Jul, Aug), and fall (Sep, Oct, Nov)).

Recommended benchmarks for photochemical model performance statistics (Boylan, 2006; Emery, 2017) will be used to assess the applicability of this modeled simulation for regulatory purposes. The goal and criteria values noted in Table 7-1 below will be used for this study.

Table 7-1. Fine Particulate Matter Performance Goals and Criteria

	NMB		NME	
Species	**Goal**	**Criteria**	**Goal**	**Criteria**
24-hr $PM_{2.5}$ and Sulfate	<± 10%	<± 30%	< 35%	< 50%
24-hr Nitrate	<± 10%	<± 65%	< 65%	< 115%
24-hr OC	<± 15%	<± 50%	< 45%	< 65%
24-hr EC	<± 20%	<± 40%	< 50%	< 75%

In addition to these goals, Alpine will compare performance evaluation metrics generated using CAMx 6.40 for each modeled species to EPA's evaluation metrics for the same simulation using CAMx 6.32. In cases where the VISTAS II performance metrics differ by more than 10% from EPA's regional haze metrics, Alpine will document the differences and provide summaries to SESARM for discussion and resolution, as necessary.

Because $PM_{2.5}$ is a mixture, current EPA PM modeling guidance (EPA, 2014e) recommends that a meaningful performance evaluation should include an assessment of how well the model is able to predict individual chemical components that constitute $PM_{2.5}$.

Consistent with EPA's performance evaluation of the regional haze 2028 analysis, in addition to total $PM_{2.5}$, components of $PM_{2.5}$ Alpine will assess:

- Sulfate ion (SO_4^{2-})
- Nitrate ion (NO_3^-)
- Ammonium ion (NH_4^+)
- Elemental Carbon (EC)
- Organic Carbon (OC) and/or Organic Carbon Mass (OCM)
- Crustal (weighted average of the most abundant trace elements in ambient air)
- Sea salt constituents (Na and Cl)

The mapping of the CAMx species into the observed species are presented in Table 7-2.

Table 7-2. Species Mapping from CAMx into Observation Network

Network	Observed Species	CAMx Species
IMPROVE	NO_3	PNO3
	SO_4	PSO4
	NH_4	PNH4
	OM = 1.8*OC	SOA1+SOA2+SOA3+SOA4+SOA5+SOA6+SOA7+SOPA+SOPB +POA
	EC	PEC
	SOIL	FPRM+FCRS
	$PM_{2.5}$	PSO4+PNO3+PNH4+SOA1+SOA2+SOA3+SOA4+SOA5+SOA6+ SOA7+SOPA+SOPB+POA+PEC+FPRM+FCRS+NA+PCL
CSN	$PM_{2.5}$	PSO4+PNO3+PNH4+SOA1+SOA2+SOA3+SOA4+SOA5+SOA6+ SOA7+SOPA+SOPB+POA+PEC+FPRM+FCRS+NA+PCL
	NO_3	PNO3
	SO_4	PSO4
	NH_4	PNH4
	OM = 1.4*OC	SOA1+SOA2+SOA3+SOA4+SOA5+SOA6+SOA7+SOPA+SOPB +POA
	EC	PEC

7.2.1 Diagnostic Evaluation

If the annual CAMx model base case simulations present performance challenges, it may necessitate focused diagnostic and sensitivity testing in order for them to be resolved. If needed, it is hopeful that these diagnostic and/or sensitivity tests can be adequately carried out within the resources and schedule of this contract. If not, then it may be necessary to draw upon the Optional Task 11 resources to conduct the necessary work. Where practical, diagnostic or sensitivity analyses, if needed, could be performed on selected episodes within the annual cycle, thereby avoiding the time-consuming task of running CAMx for the full 2011 period. Upon identification of performance evaluation failure, Alpine will identify the types of diagnostic and sensitivity testing methods that might be employed in diagnosing inadequate model performance and devising appropriate methods for improving the model response. Upon discussion, negotiation, and approval with and by SESARM and EPA of evaluation procedures to conduct, Alpine will execute the tests, identify the issues, propose and resolution to the problem, and apply any changes as approved.

8.0 AREA OF INFLUENCE

Under this task, ERG will identify 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.[28] Due to the presence of large SO_2 emission reductions during this six-year period, the area of influence (AoI) analysis will be 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 Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model[29] developed by the National Oceanic and Atmospheric Administration's (NOAA) Air Resources Laboratory (ARL) will be run for each of these days to identify areas most likely influencing visibility. The HYSPLIT runs will use the NAM-12 hybrid meteorology and trajectory will be run for starting heights of 100 meters (m), 500 m, 1,000 m, and 1,500 m. Trajectories will be run 72 hours backwards in time at each height and location. The AoI analysis will be set up to look at: 1) each starting height individually and 2) all starting heights combined.

Trajectories will be 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. Trajectories will originate from each of the forty (40) IMPROVE monitors in the VISTAS_12 domain (Table 8-1 and Figure 8-1). Based on analysis from NC, ERG omitted several western states, as trajectories originating in Class I areas in those states do not pass over a SESARM state (hatched shading in Figure 8-1).

In certain instances, the trajectory origin will be the centroid of the Class I area. Class I areas without a dedicated IMPROVE monitor will have their trajectories originate from the centroid of the Class I area. Visibility data will be based on an appropriate IMPROVE monitor, as previously determined by the federal Land Managers. Another example when the monitor is remove from the Class I area. For example, the Breton Island trajectory will start from the centroid of the Class I, as the monitor has been moved further on shore in Louisiana. There are nine (9) such areas identified in the VISTAS 12km domain.

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

Three VISTAS region Class I areas do not have an IMPROVE monitor: Wolf Island, Joyce Kilmer-Slickrock and Otter Creek. For these sites, trajectories will originate from the centroid of the Class I area (Figure 8-2). The 20% most impaired dates will be based off a representative IMPROVE monitor, which are listed in Table 8-2.

---

[28] http://vista.cira.colostate.edu/Improve/rhr-summary-data/

[29] 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

Table 8-1. IMPROVE Monitors in the VISTAS_12 Domain

IMPROVE Site	IMPROVE Site Code	State	FIPS Code	Latitude	Longitude	Start Date	End Date
Sipsey Wilderness	SIPS1	AL	01079	34.3433	-87.3388	03/04/1992	04/28/2017
Caney Creek	CACR1	AR	05113	34.4544	-94.1429	06/24/2000	04/28/2017
Upper Buffalo Wilderness	UPBU1	AR	05101	35.8258	-93.2030	12/04/1991	04/28/2017
Chassahowitzka NWR	CHAS1	FL	12017	28.7484	-82.5549	03/03/1993	04/28/2017
Everglades NP	EVER1	FL	12086	25.3910	-80.6806	09/03/1988	04/28/2017
St. Marks	SAMA1	FL	12129	30.0926	-84.1614	08/16/2000	04/28/2017
Cohutta	COHU1	GA	13213	34.7852	-84.6265	06/03/2000	04/28/2017
Okefenokee NWR	OKEF1	GA	13049	30.7405	-82.1283	09/04/1991	04/28/2017
Mammoth Cave NP	MACA1	KY	21061	37.1318	-86.1479	09/04/1991	04/28/2017
Breton Island	BRIS1	LA	22075	30.1086	-89.7617	01/16/2008	04/28/2017
Acadia NP	ACAD1	ME	23009	44.3771	-68.2610	03/02/1988	04/28/2017
Moosehorn NWR	MOOS1	ME	23029	45.1259	-67.2661	12/03/1994	04/28/2017
Isle Royale NP	ISLE1	MI	26083	47.4596	-88.1491	11/17/1999	04/28/2017
Seney	SENE1	MI	26153	46.2889	-85.9503	11/17/1999	04/28/2017
Boundary Waters Canoe Area	BOWA1	MN	27075	47.9466	-91.4955	06/01/1991	04/28/2017
Voyageurs NP #2	VOYA2	MN	27137	48.4126	-92.8286	03/02/1988	04/28/2017
Hercules-Glades	HEGL1	MO	29213	36.6138	-92.9221	03/02/2001	04/28/2017
Mingo	MING1	MO	29207	36.9717	-90.1432	06/03/2000	04/28/2017
Linville Gorge	LIGO1	NC	37011	35.9723	-81.9331	04/01/2000	04/28/2017
Shining Rock Wilderness	SHRO1	NC	37087	35.3937	-82.7744	06/01/1994	04/28/2017
Swanquarter	SWAN1	NC	37095	35.4510	-76.2075	06/10/2000	04/28/2017
Great Gulf Wilderness	GRGU1	NH	33007	44.3082	-71.2177	06/03/1995	04/28/2017
Brigantine NWR	BRIG1	NJ	34001	39.4650	-74.4492	09/04/1991	04/28/2017
Bandelier NM	BAND1	NM	35028	35.7797	-106.2664	03/02/1988	04/28/2017
Bosque del Apache	BOAP1	NM	35053	33.8695	-106.8520	04/15/2000	04/28/2017
Salt Creek	SACR1	NM	35005	33.4598	-104.4042	04/08/2000	04/28/2017
San Pedro Parks	SAPE1	NM	35039	36.0139	-106.8447	08/16/2000	04/28/2017
Wheeler Peak	WHPE1	NM	35055	36.5854	-105.4520	08/16/2000	04/28/2017
White Mountain	WHIT1	NM	35027	33.4687	-105.5349	12/03/2001	04/28/2017
Wichita Mountains	WIMO1	OK	40031	34.7323	-98.7130	03/02/2001	04/28/2017
Cape Romain NWR	ROMA1	SC	45019	32.9410	-79.6572	09/03/1994	04/28/2017
Great Smoky Mountains NP	GRSM1	TN	47009	35.6334	-83.9416	03/02/1988	04/28/2017
Big Bend NP	BIBE1	TX	48043	29.3027	-103.1780	03/02/1988	04/28/2017
Guadalupe Mountains NP	GUMO1	TX	48109	31.8330	-104.8094	03/02/1988	04/28/2017
James River Face Wilderness	JARI1	VA	51163	37.6266	-79.5125	06/03/2000	04/28/2017
Shenandoah NP	SHEN1	VA	51113	38.5229	-78.4348	03/02/1988	04/28/2017
Lye Brook Wilderness	LYBR1	VT	50003	43.1482	-73.1268	09/04/1991	09/30/2012
Lye Brook Wilderness	LYEB1	VT	50025	42.9561	-72.9098	01/01/2012	04/28/2017
Dolly Sods Wilderness	DOSO1	WV	54093	39.1053	-79.4261	09/04/1991	04/28/2017

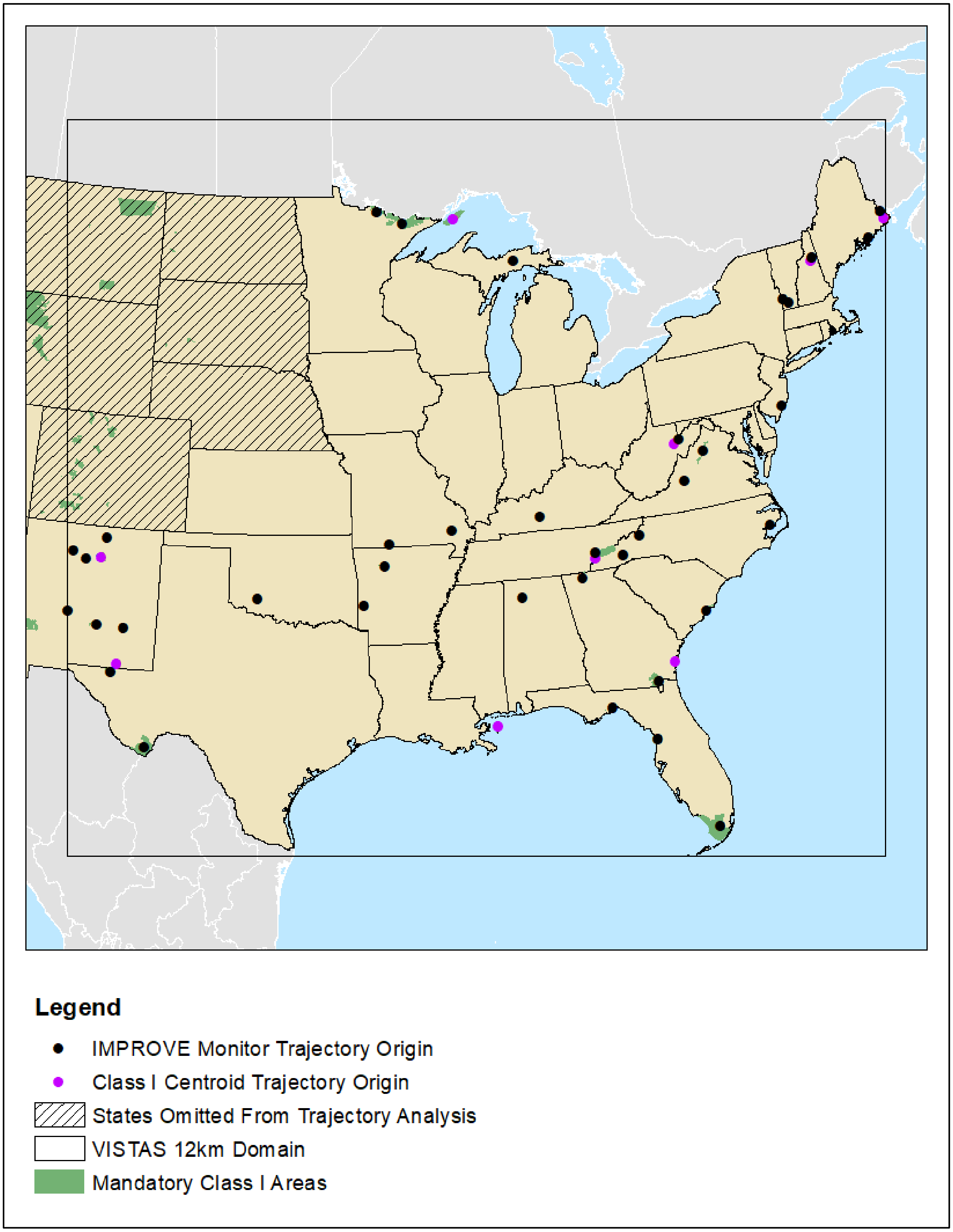


Figure 8-1. IMPROVE Monitor Locations and Starting Points for HYSPLIT Trajectories in the VISTAS 12km Domain

Figure 8-2. IMPROVE Monitor Locations and Starting Points for HYSPLIT Trajectories in the VISTAS States

Table 8-2. Representative IMPROVE Monitor for Each VISTAS Class I Area

Class I Area	Representative IMPROVE Site	IMPROVE Site Code	State	FIPS County Code	Latitude	Longitude
AL - Sipsey Wilderness Area	Sipsey Wilderness	SIPS1	AL	01079	34.3433	-87.3388
FL - Chassahowitzka Wilderness Area	Chassahowitzka NWR	CHAS1	FL	12017	28.7484	-82.5549
FL - Everglades National Park	Everglades NP	EVER1	FL	12086	25.391	-80.6806
FL - St. Marks Wilderness Area	St. Marks	SAMA1	FL	12129	30.0926	-84.1614
GA - Cohutta Wilderness Area	Cohutta	COHU1	GA	13213	34.7852	-84.6265
GA - Okefenokee Wilderness Area	Okefenokee NWR	OKEF1	GA	13049	30.7405	-82.1283
GA - Wolf Island Wilderness Area	Okefenokee NWR	OKEF1	GA	13049	30.7405	-82.1283
KY - Mammoth Cave National Park	Mammoth Cave NP	MACA1	KY	21061	37.1318	-86.1479
NC/TN - Great Smoky Mountains National Park	Great Smoky Mountains NP	GRSM1	TN	47009	35.6334	-83.9416
NC/TN - Joyce Kilmer-Slickrock Wilderness	Great Smoky Mountains NP	GRSM1	TN	47009	35.6334	-83.9416
NC - Linville Gorge Wilderness Area	Linville Gorge	LIGO1	NC	37011	35.9723	-81.9331
NC - Shining Rock Wilderness Area	Shining Rock Wilderness	SHRO1	NC	37087	35.3937	-82.7744
NC - Swanquarter Wilderness Area	Swanquarter	SWAN1	NC	37095	35.451	-76.2075
SC - Cape Romain Wilderness Area	Cape Romain NWR	ROMA1	SC	45019	32.941	-79.6572
VA - James River Face Wilderness Area	James River Face Wilderness	JARI1	VA	51163	37.6266	-79.5125
VA - Shenandoah National Park	Shenandoah NP	SHEN1	VA	51113	38.5229	-78.4348
WV - Dolly Sods Wilderness Area	Dolly Sods Wilderness	DOSO1	WV	54093	39.1053	-79.4261
WV - Otter Creek Wilderness Area	Dolly Sods Wilderness	DOSO1	WV	54093	39.1053	-79.4261

Trajectories will be run utilizing the SplitR package (https://github.com/rich-iannone/SplitR), which allows the control of HYSPLIT through the R statistical software. This allows for automation of the HYSPLIT runs for each location, while still generating the GIS shapefiles and separate files of the endpoint for further use in R.

The back trajectories for the 20% most impaired days will then be used to develop residency time (RT) plots via the openair[30] package for R. The RT plots define the geographic areas with the highest probability of influencing the monitor on the 20% most impaired visibility days.

The RT is calculated as the number of trajectory hours that pass through each grid cell. This can be presented as a percentage of the total number of trajectory hours. The grid used would align with the photochemical modeling 12km grid for consistency with emission analysis. For further analysis, R allows the residence time plots to be split by time increments (i.e., year, season). Images of the RT plots will be generated for QA and review purposes. Images will at least cover the VISTAS 12-km domain and include outlines of states and counties.

The trajectory data will also be weighted by ammonium sulfate and ammonium nitrate and used to produce separate sulfate and nitrate extinction weighted residency time (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)[31] approach will be used to develop the EWRT, substituting the extinction value for the concentration. The extinction attributable to each pollutant is paired with the trajectory for that day. R then calculates the mean weighted extinction of the pollutant species for each grid cell. The mean weighted extinction is calculated by:

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

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 τ_{ijk} the number of trajectory hours that pass through each grid cell (i, j) (where “i” is the row and “j” is the column).[32] 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. A point filter can be used to eliminate grid cell with few trajectory hours from the

---

[30] Carslaw DC 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.

[31] 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

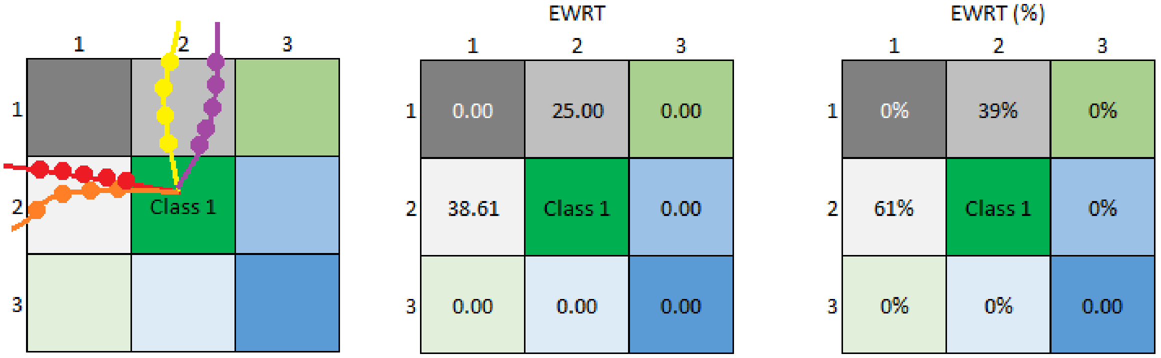
[32] Carslaw, D.C. (2015). The openair manual — open-source tools for analyzing air pollution data. Manual for Version 1.1-4, King’s College London. http://www.openair-project.org/PDF/OpenAir_Manual.pdf

analysis. As this distinction can help refine control strategy development, using the CWT method to create the EWRT is ERG's recommendation for the analysis.

The EWRT results can be normalized by the domain total to present the results as a percentage in images. Images of the extinction weighted RT plots will be generated for QA and review purposes. Images will at least cover the VISTAS 12-km domain and include outlines of states and counties. An example calculation is provided in Figure 8-3, which has been simplified to only four trajectories. In this example, two trajectories pass through the cell (1,2). The first of these trajectories (yellow) has an extinction value $(bext_k)$ of 50 Mm^{-1} associated with it. The four yellow dots in the cell denote that four hours of the back trajectory are spent in the cell. This yields "$\tau_{1,2,yellow}$" equal to 4 hours. Multiplying by the extinction $(bext_k)$ for the cell yields a trajectory weight of 200 Mm^{-1}hrs. This is added to the weight of the second (purple) trajectory for a total extinction weight of 450 Mm^{-1}hrs. This is then divided by the total number of trajectory hours in the analysis (18 hours) to yield the final EWRT for the cell of 25 Mm^{-1}. The same calculation is done for cell (2,1). While the cells have the same number of trajectories and endpoints, the weight given to cell (2,1) is higher due to the higher extinctions associated with the trajectories.

The next phase of the analysis will combine the EWRT values with the distance weighted gridded emission data to determine the sources most likely contributing to the elevated extinction levels. Distances (d) for the weighting, calculated using ArcGIS, will be calculated from the location of the point source to the trajectory origin in kilometers. The weighted emission file is comprised of the EGU and non-EGU point source emissions value for each grid cell (Q, in tons per year) divided by the distance (d, in kilometres) to the trajectory origin; that is the final value is (Q/d). Each of these grid cell values is multiplied by its respective sulfate or nitrate EWRT plot values (i.e., EWRT *(Q/d)). A simplified example calculation for a single point source per cell is provided in Figure 8-4. Continuing with weights calculated in Figure 8-3, point sources in cell (1,2) emit 100 tons per year (Q=100 tpy), with the centroid 12km away from the trajectory origin. This yields a "Q/d" of 8.33 tpy/km for the source. This "Q/d" term is multiplied by the EWRT previously calculated at 25 Mm^{-1} to yield a source weighted EWRT value that indicates the potential importance of the sources in this cell to impaired visibility days in the Class I area. Alternately, the EWRT*(Q/d) can be calculated for each source in a cell, and totalled for a grid cell total. Similar to the EWRT plots, these contribution plots can be normalized by the domain total to present them as a percentage in plots. Images of the results will be mapped over the VISTAS 12-km modeling domain, with state and county boundaries for review and QA purposes.

These gridded results will then be linked with the 2011 and 2028 point source inventories to calculate the emission contribution from each source. ArcGIS will be used to spatially join the gridded information with shapefiles the point source information. This will create a dataset that combines the point source metadata facility identifying information (i.e., Facility ID, Facility Name, State, County, Federal Information Processing Standard (FIPS), North American Industry Classification System (NAICS), and industry description), and the gridded information (i.e., SO_2 and NO_2 emissions, d, Q/d, EWRT, EWRT *(Q/d)). The Q/d values will be calculated by dividing facility-wide emissions (tons per year) by distance (km) to the Class I area. In the alternate strategy proposed above, this step would be completed first, then totalled for visualization. Additional information can be added as deemed necessary in making control strategy decisions. The information from these spatial files can then be exported to separate Excel spreadsheets for each Class I area in the VISTAS_12 domain for further review by the states.



EWRT

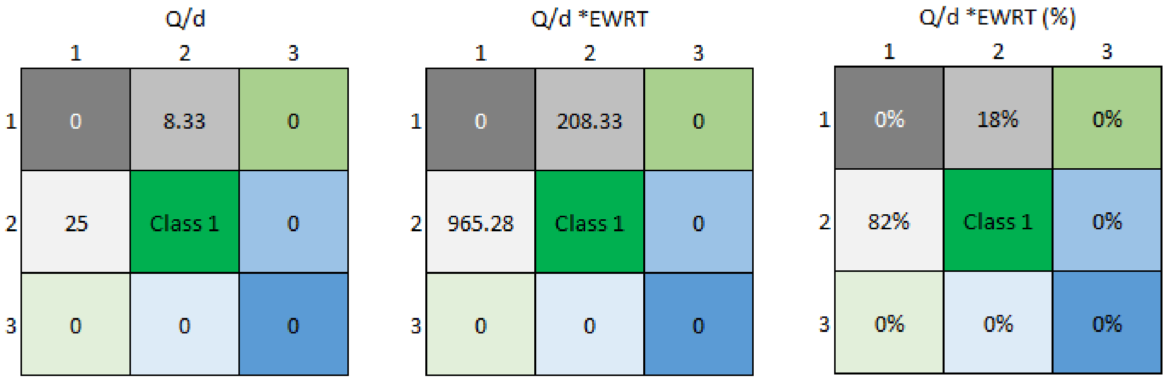
	1	2	3
1	0.00	25.00	0.00
2	38.61	Class 1	0.00
3	0.00	0.00	0.00

EWRT (%)

	1	2	3
1	0%	39%	0%
2	61%	Class 1	0%
3	0%	0%	0.00

Row (i)	Col (j)	# Trajectory hours	Trajectory (k)	Extinction ($bext_k$) [Mm^{-1}]	τ_{ijk} [hrs]	($bext_k$)τ_{ijk} [Mm^{-1} hrs]	$\Sigma\tau_{ij}$ [hrs]	$\Sigma(bext_k)\tau_{ijk}$ [Mm^{-1} hrs]	EWRT [Mm^{-1}]
2	1	5	1 (Red)	75	5	375	18	695	38.61
		4	2 (Orange)	80	4	320			
1	2	4	1 (Yellow)	50	4	200	18	450	25
		5	2 (Purple)	50	5	250			

Figure 8-3. Example EWRT Calculations



Q/d

	1	2	3
1	0	8.33	0
2	25	Class 1	0
3	0	0	0

Q/d *EWRT

	1	2	3
1	0	208.33	0
2	965.28	Class 1	0
3	0	0	0

Q/d *EWRT (%)

	1	2	3
1	0%	18%	0%
2	82%	Class 1	0%
3	0%	0%	0%

Row (i)	Col (j)	ΣQ [tpy]	d [km]	Q/d [tpy km^{-1}]	Q/d*EWRT [tpy km^{-1} Mm^{-1}]	Q/d*EWRT [%]
2	1	300	12	25	965.28	82%
1	2	100	12	8.33	208.33	11%

Figure 8-4. Example (Q/d)*EWRT Calculations

Similar analysis will be conducted to rank SO_2 and NO_2 emissions contributions for onroad, nonroad, fires, and area source sectors from each county. The process will be similar to the process for point sources previously described, except calculations of RT and EWRT will be done to counties as opposed to grids. ERG will determine if the trajectories can be weighted further for the time spent in the county. The length of the trajectory within the county would be used as a proxy for time, so that trajectories that only cross a small corner of the county are not weight as much as trajectories passing through the center of the county. This will be done in GIS using the same calculation method as in *R* (i.e., CWT). The calculation of d would then be from the centroid of the county to the trajectory origin, in kilometers. Similar to point sources, the final spatial join would be to the county level EWRT and a shapefile of the source information at the county level, for each sector. All county and emissions source identifying information will be provided along with inventory emissions, distance, Q/d and Q/d^2 values, EWRT, EWRT*(Q/d), fraction and sum contributions, and any other information deemed necessary in making control strategy decisions for each source.

All images, shapefiles, and spreadsheets will be uploaded to the files sharing platform, as designed in Task 10, in separate folders for each Class I area or IMPROVE monitor in the VISTAS_12 domain. Each analysis element (e.g., RT plots, summary spreadsheets) will be contained in separate subfolders for ease of navigation. SESARM will be notified when the files are available and ready for use. A technical memorandum/interim report describing the area of influence calculations, and the results, will be prepared for SESARM and other stakeholders for use in their implementation plans.

9.0 FUTURE YEAR MODELING

This chapter discusses the future year modeling using the annual modeling databases and how these results will be used to project 2009-2013 IMPROVE visibility data to 2028 following the approach used in EPA's regional haze modeling guidance (EPA, 2014e).

9.1 Regional Haze Rule Requirements

As required by the RHR, RPGs must provide for an improvement in visibility for the 20 percent most anthropogenically impaired days relative to baseline visibility conditions and ensure no degradation in visibility for the 20 percent clearest days relative to baseline visibility conditions. The baseline for each Class I area is the average visibility (in dvs) for the years 2000 through 2004. The visibility conditions in these years are the benchmark for the "provide for an improvement" and "no degradation" requirements. In addition, states are required to determine the rate of improvement in visibility needed to reach natural conditions by 2064 for the 20 percent most anthropogenically impaired days. A line drawn between the end of the 2000-2004 baseline period and 2064 (dv/year) shows a uniform rate of progress (URP) between these two points. This "glidepath" is the amount of visibility improvement needed in each implementation period, starting from the baseline period, to stay on a linear path towards visibility improvement to natural conditions by 2064. The glidepath represents a linear or uniform rate of progress. This is a framework for consideration but there is no requirement to be on or below the glidepath.

The RHR requires states to submit an implementation plan that evaluates reasonable progress for implementation periods in approximately ten year increments. The next regional haze SIP is due in 2021, for the implementation period which ends in 2028 (period of 2019-2028).

Therefore, modeling is being conducted to project visibility to 2028 using a 2028 emissions inventory with "on-the-books" controls. The EPA Software for Model Attainment Test-Community Edition (SMAT-CE) tool[33] will be used to calculate 2028 dv values on the 20% most impaired and 20% clearest days at each Class I Area (IMPROVE site). SMAT-CE is an EPA software tool which implements the procedures in the modeling guidance to project visibility to a future year.

9.2 Future Year to be Simulated

As discussed in Section 1, to support the preliminary 2028 regional haze modeling requested by SESARM, Alpine will conduct air quality modeling to project particulate matter concentrations at individual monitoring sites to 2028 and to estimate changes in regional haze as a result of those concentrations.

9.3 Future Year Baseline Air Quality Simulations

The 2028 future year base case CAMx simulation will be conducted and particulate matter concentration calculations made following the procedures in Section 5 and that are consistent with EPA's latest regional haze modeling analysis (EPA, 2017b).

---

[33] https://www.epa.gov/scram/photochemical-modeling-tools

9.4 Calculation of 2028 Visibility

The visibility projections will follow the procedures in EPA's modeling guidance (EPA, 2014e). Based on the recommendation in the modeling guidance, the observed base period visibility data will be linked to the base modeling year. This is the 5-year ambient data base period centered about the base modeling year. In this case, for a base modeling year of 2011, the ambient IMPROVE data will be from the 2009-2013 period.

The visibility calculations will use the "revised" IMPROVE equation (Hand, 2006); (Pitchford, 2007), which has replaced the original IMPROVE equation and has been used in most regional haze SIPs over the last 10 years. The IMPROVE equation (or algorithm), which uses PM species concentrations and relative humidity data to calculate visibility impairment, or b_{ext}, in units of inverse megameters (Mm^{-1}), is presented below:

$$b_{ext} \approx 2.2 \times f_s(RH) \times [\text{Small Sulfate}] + 4.8 \times f_L(RH) \times [\text{Large Sulfate}] + 2.4 \times f_s(RH) \times [\text{Small Nitrate}] + 5.1 \times f_L(RH) \times [\text{Large Nitrate}] + 2.8 \times [\text{Small Organic Mass}] + 6.1 \times [\text{Large Organic Mass}] + 10 \times [\text{Elemental Carbon}] + 1 \times [\text{Fine Soil}] + 1.7 \times f_{SS}(RH) \times [\text{Sea Salt}] + 0.6 \times [\text{Coarse Mass}] + \text{Rayleigh Scattering (Site Specific)} + 0.33 \times [NO_2\ (ppb)]$$

The total sulfate, nitrate, and organic carbon compound concentrations are each split into two fractions, representing small and large size distributions of those components. Site-specific Rayleigh scattering is calculated based on the elevation and annual average temperature of each IMPROVE monitoring site.

The 2028 future year visibility on the 20% most impaired days and 20% clearest days at each Class I area will be estimated using the observed IMPROVE data (2009-2013) and the relative percent modeled change in PM species between 2011 and 2028. The process is described in the following six steps.

1. For each Class I area (IMPROVE site), estimate anthropogenic impairment[34] on each day using observed speciated $PM_{2.5}$ data plus PM_{10} data (and other information) for each of the 5 years comprising the base period (2009-2013 in this case) and rank the days on this indicator. This ranking will determine the 20 percent most anthropogenically impaired days. For each Class I area, also rank observed visibility (in dvs) on each day using observed speciated $PM_{2.5}$ data plus PM_{10} data for each of the 5 years comprising the base period. This ranking will determine the 20 percent clearest days.

2. For each of the 5 years comprising the base period, calculate the mean dvs for the 20 percent most anthropogenically impaired days and 20 percent clearest days. For each Class I area, calculate the 5 year mean dvs for most impaired and clearest days from the 5 year-specific values.

3. Use an air quality model to simulate air quality with base period (2011) emissions and future year (2028) emissions. Use the resulting information to develop site-specific relative response factors (RRFs) for each component of PM identified in the "revised"

---

[34] http://vista.cira.colostate.edu/Improve/rhr-summary-data/

IMPROVE equation. The RRFs are an average percent change in species concentrations based on the measured 20% most impaired and 20% clearest days from 2011 (the calendar days identified from the IMPROVE data above are matched to the same modeled days).

4. Multiply the species-specific RRFs by the measured daily species concentration data during the 2009-2013 base period (for each day in the measured 20% most impaired day set and each day in the 20% clearest day set), for each site. This results in daily future year 2028 PM species concentration data.

5. Using the results in Step 4 and the IMPROVE algorithm, calculate the future daily extinction values for the previously identified 20 percent most impaired days and 20 percent clearest days in each of the five base years.

6. Calculate daily dv values (from total daily extinction) and then compute the future year (2028) average mean dvs for the 20 percent most impaired days and 20 percent clearest days for each year. Average the five years together to get the final future mean dv values for the 20 percent most impaired days and 20 percent clearest days.

The SMAT-CE tool will be used to generate individual year and 5-year average base year and future year dv values on the 20% most impaired days and 20% clearest days. Additional SMAT output variables include the results of intermediate calculations such as species specific extinction values (both base and future year) and species specific RRFs (on the 20% most impaired and clearest days).

Table 9-1 details the settings to be used for the SMAT runs to generate the 2028 future year dv projections:

Table 9-1. SMAT-CE Settings for 2028 Visibility Calculations

SMAT-CE Option	Setting or File Used
IMPROVE algorithm	Use new version
Grid cells at monitor or Class I area centroid?	Use grid cells at monitor
IMPROVE data file	ClassIareas_NEWIMPROVEALG_2000to2015_2017april27_IMPAIRMENT.csv
Temporal adjustment at monitor	3 x 3
Start monitor year	2009
End monitor year	2013
Base Model year	2011
Minimum years required for a valid monitor	3

In order for Alpine to conduct the source apportionment runs, the CAMx 2011 and 2028 model output will be post-processed using a “species definition file” that cross references raw CAMx output species names with PM species needed for SMAT. The results of the post-processing are 24-hour average PM species with the “combine file” output names. These are matched to the SMAT species as shown in Table 10-1.

Table 9-2. Matching of CAMx Raw Output Species to SMAT Input Variables

SMAT Species	"Combine File" Output Name	Raw CAMx 6.40 Species
Sulfate	PM25_SO4	PSO4
Nitrate	PM25_NO3	PNO3
Ammonium	PM25_NH4	PNH4
Organic carbon	PM25_OM	POA+SOA1+SOA2+SOPA+SOA3+SOA4+SOA5 +SOA6+SOA7+SOPB
Elemental carbon	PM25_EC	PEC
Crustal	CRUSTAL	FCRS+FPRM
Coarse PM	PMC_TOT	CCRS+CPRM
PM2.5	PM25_SMAT	CRUSTAL+PSO4+PNO3+PNH4+PEC+NA+PCL+ SOA1+SOA2+SOA3+SOA4+SOA5+SOA6+SOA7 +SOPA+SOPB+POA

9.5 Comparison to Regional Haze "Glidepath"

The future year 2028 dv projections will be compared to the unadjusted visibility "glidepath" at each Class I area. The unadjusted "glidepath" represents the amount of visibility improvement needed in each implementation period, starting from the baseline 2000-2004 period, to stay on a linear path to natural visibility conditions by 2064. Visibility on the 20% most impaired days is compared to the relevant value of the glidepath, in this case for a future year of 2028. Since the glidepath is a linear path between 2004 and 2064, a glidepath value (in dvs) can be calculated for any future year, using a simple equation. The following formula will be used to calculate the 2028 glidepath value:

$$Glidepath_{2028} = Baseline\ average\ dv - (((Baseline\ average\ dv - Natural\ conditions)/60)*24)$$

Where,

Baseline average dv = average observed dv value on the 20% most impaired days for 2000-2004

Natural conditions= Natural conditions on the 20% most impaired days at the Class I area (in dv)

Once calculated, the 2028 future year projected dv values can be compared to the unadjusted glidepath for 2028 to determine if the Class I area is projected to be above, below, or on the glidepath. While the RHR requires future year projected visibility impairment be compared to the glidepath, it does not require the RPGs be on or below the glidepath. However, the rule has different requirements depending on whether the projected value (RPG) is above or below the glidepath.

10.0 PSAT SOURCE APPORTIONMENT

In order to gain a better understanding of the source contributions to modeled visibility, Alpine will use CAMx Particulate Source Apportionment Technology (PSAT) modeling. PSAT uses multiple tracer families to track the fate of both primary and secondary PM (Yarwood et al., 2004). PSAT is designed to apportion the following classes of CAMx PM species:

- Sulfate (PSO_4)
- Particulate nitrate (PNO_3)
- Ammonium (PNH_4)
- Secondary organic aerosol (SOA)
- Primary PM (PEC, POA, FCRS, FPRM, CCRS, and CPRM)
- Particulate mercury (HGP)

PSAT allows emissions to be tracked (tagged) by various combinations of sectors and geographic areas (e.g., by state or facility). For this application, 2028 emissions will be tagged (“tag”) using SESARM-identified combinations of region, facilities, and/or source category. Each combination accounts for a single “tag” with SESARM planning to identify up to 250 individual tagged combinations. Each of these emissions combinations will be processed separately through SMOKE and tracked in PSAT as individual source tags. Receptors, identified as all the Class I areas in the VISTAS_12 modeling domain, will be used to analyze the results and impacts of each tagged combination.

For this application, only sulfate and nitrate will be tracked using PSAT. Tracking of other contributions may also be of use, but is not requested in this analysis. In the PSAT post processing the data will be converted from the UTC based modeling data into 24-hour average grid cell local standard time using the HR2DAY[35] utility program. The HR2DAY program matches each grid cell in the model to a timezone and performs the 24-hour average on the timezone adjusted hourly data. This is the same program that EPA uses in their national modeling.

In order for Alpine to conduct the source apportionment runs, the CAMx 2011 and 2028 model output will be post-processed using a “species definition file” that cross references raw CAMx output species names with PM species needed for SMAT. The results of the post-processing are 24-hour average PM species with the “combine file” output names. These are matched to the SMAT species as shown in Table 10-1.

Table 10-1. Matching of CAMx Raw Output Species to SMAT Input Variables

SMAT Species	Raw CAMx 6.40 Species
Sulfate (SO4)	PSO4
Nitrate (NO3)	PNO3
Ammonium (NH4)	PNH4
Organic carbon (OC)	POA+SOA1+SOA2+SOPA+SOA3+SOA4+SOA5+SOA6+SOA7+SOPA +SOPB

---

[35] https://www.airqualitymodeling.org/index.php/CMAQv5.1_Tools_and_Utilities#HR2DAY_utility_program

Table 10-1. Matching of CAMx Raw Output Species to SMAT Input Variables

SMAT Species	Raw CAMx 6.40 Species
Elemental carbon (EC)	PEC
Crustal (CRUSTAL)	FCRS+CFRM
Coarse PM (CM)	CPRM+CCRS
PM2.5 (PM25)	PSO4+PNO3+PNH4+POA+PEC+FCRS+FPRM+SOA1+SOA2+SOA3+SOA4+SOA4+SOA5+SOA6+SOA7+SOPA+SOPB

10.1 Process for Creating PSAT Contributions for Class I Areas

The CAMx hourly concentration data will be post-processed to create SMAT input files. This will involve processing both the 2028 "full model" and the specific source apportionment outputs. The "full model" results are the total PM species concentrations (e.g. sulfate, nitrate) and are identical to the total species concentrations from the non-source apportionment model run for 2028 (e.g., future year base case). The source apportionment outputs contain the sulfate and/or nitrate contributions for each tagged source.

The PSAT source apportionment tracking uses slightly different variables names for the source apportionment variables. Table 10-2 below shows the SMAT species definition matching to be used for the 2028 full model and 2028 source apportionment results in the VISTAS II analysis.

Table 10-2. Matching of "Bulk Raw Species", PSAT Output Species, and SMAT Input Variables

SMAT Species	2028 Full Model Species	2028 PSAT Tag Raw Species
Sulfate	PSO4	PS4
Nitrate	PNO3	PN3

This analysis will use a slightly different method than was documented by EPA in the regional haze modeling for 2028. For example, in this study we are looking at the SMAT-CE generated visibility/extinction deltas. The EPA approach, however, was designed for a different purpose than just to estimate emissions sector contributions to 2028 particulate matter concentrations and visibility. Their configuration of the SMAT simulations and post-processed calculations allowed EPA to better understand the sources of future visibility impairment (including domestic anthropogenic, domestic natural, international anthropogenic, and natural sources). Since these latter issues are not being investigated by SESARM in this study, our presented approach is conceptually simpler and more straightforward to implement and to document. SESARM is only looking for sector contributions to visibility impairment based on a set of to-be-defined sulfate and nitrate tags and not looking to establish species-based contribution metrics.

The following approach will be used in preparing the SMAT input files, running the SMAT software, and analysing the results:

1. Regional haze SMAT will be run for the 2028 future case using "standard" 2011 and 2028 full model SMAT input files. This prepares the 2028 output files which will be used as the basis for comparison with the "tagged" SMAT output described below.

2. Alpine with then create future year, tag-specific SMAT input files by subtracting the 2028 hourly tags from the hourly full model concentration files. This simple arithmetic will be implemented using standard IOAPI utility programs and will generate files similar to EPA's source category-based tagged SMAT input files. Once the hourly files are created, the same processing stream as was used in Step 1 will be used create the tagged SMAT input files from the hourly model concentration files.

3. SMAT will be run (in batch mode) for each future year tag-specific input file generated in Step 2 using the "standard" 2011 SMAT input file as the base year. In these runs, SMAT will be configured identically as in Step 1 except for using the future year "tagged" input files. These individual runs will generate SMAT output files that contain the forecasted extinction data *absent* the tagged contribution.

4. The total extinction (on the 20% most impaired days) for each tag will be calculated from the SMAT bulk output file and each of the tag output files. The visibility impacts of each tag will be computed by subtracting the SMAT output *absent* the tag (created in Step 3) from the full model SMAT output file (created in Step 1).

An example calculation for Cohutta using EPA's 2028 draft regional haze modeling SMAT input and generated output files is provided in Table 10-3 below. In this calculation, we use the 2011 base year and 2028 base case simulation compared to the 2011 base year and Sector 9 (no EGUs) simulation. Visibility (in dv), large sulfate, small nitrate, and bext are provided as examples; numbers are rounded in this example and may not sum due to rounding.

Table 10-3. Tagged Contribution Calculation Example

SMAT Calculated Value	2028 Base Case		2028 Sector 9 (Absent EGUs)		Delta (Base – Sector 9) Tag Impact	
	Best 20% Days	**Most Impaired 20% Days**	**Best 20% Days**	**Most Impaired 20% Days**	**Best 20% Days**	**Most Impaired 20% Days**
Visibility (dv_f)	9.11	17.69	8.07	16.07	1.04	1.62
$SO_{4,Large}$_f (µg/m^3)	0.037	0.326	0.017	0.148	0.020	0.178
$NO_{3,Small}$_f (µg/m^3)	0.208	0.238	0.187	0.221	0.021	0.018
Bext_f (Mm^{-1})	25.41	59.75	22.79	50.9	2.62	8.84

In summary, since SMAT directly provides visibility metrics for the 20% most impaired and 20% clearest days, our proposed application is to subtract the visibility results of the tagged run (deciviews absent the tag) from the 2028 base case run to determine to contribution of each tag to the total base case visibility value.

Alpine will continue to review this method with SESARM and EPA and should a change be warranted in the calculation steps, we will change the calculation to meet the needs of SESARM prior to running the contribution steps.

Finally, using data from the resulting calculations, Alpine will create the day-by-day stacked bar charts of total and speciated component b_{ext} (ammonium sulfate, ammonium nitrate, OC, EC, crustals, CM) for the 20% most impaired and 20% clearest days per site for each IMPROVE monitor in the VISTAS_12 modeling domain and as directed by SESARM will also modify the similar 2011 site-by-site charts to include 2028 modeled data.

11.0 MODELING DOCUMENTATION AND DATA ARCHIVE

EPA recommends that certain types of documentation be provided along with a photochemical modeling attainment demonstration. ERG and Alpine Geophysics are committed to supplying the material needed to ensure that the technical support for this analysis is understood by all stakeholders, EPA, and SESARM.

Alpine plans to archive all documentation and modeling input/output files generated as part of the VISTAS II modeling analysis and will provide a copy to SESARM for permanent archival, additional internal use, and public distribution. Key participants in this modeling effort will be given data access to the archived modeling information.

12.0 REFERENCES

Abt. 2014. Modeled Attainment Test Software – User's Manual. Abt Associates Inc., Bethesda, MD. April. (http://www.epa.gov/ttn/scram/guidance/guide/MATS_2-6-1_manual.pdf).

Arnold, J.R., R.L. Dennis, and G.S. Tonnesen. 2003. "Diagnostic Evaluation of Numerical Air Quality Models with Specialized Ambient Observations: Testing the Community Multiscale Air Quality Modeling System (CMAQ) at Selected SOS 95 Ground Sites", *Atmos. Environ.*, Vol. 37, pp. 1185-1198.

Arnold, J.R.; R.L. Dennis, and G.S. Tonnesen. 1998. Advanced techniques for evaluating Eulerian air quality models: background and methodology. In: Preprints of the 10th Joint Conference on the Applications of Air Pollution Meteorology with the Air & Waste Management Association, January, Phoenix, Arizona. American Meteorological Society, Boston, Massachusetts, paper no. 1.1, pp. 1-5.

Arunachalam, S. 2009. Peer Review of Source Apportionment Tools in CAMx and CMAQ. University of North Carolina Institute for the Environment, Chapel Hill, NC. August 31. (http://www.epa.gov/scram001/reports/SourceApportionmentPeerReview.pdf).

Boylan, J. W. 2004. "Calculating Statistics: Concentration Related Performance Goals", paper presented at the EPA PM Model Performance Workshop, Chapel Hill, NC. 11 February.

Boylan, J.W., and A.G. Russell. 2006. PM and light extinction model performance metrics, goals, and criteria for three dimensional air quality models. Atmos. Environ. 40:4946–59. doi:10.1016/j.atmosenv.2005.09.087.

Byun, D.W., and J.K.S. Ching. 1999. "Science Algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System", EPA/600/R-99/030.

Coats, C.J. 1995. Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System, MCNC Environmental Programs, Research Triangle Park, NC.

Colella, P., and P.R. Woodward. 1984. The Piecewise Parabolic Method (PPM) for Gasdynamical Simulations. *J. Comp. Phys.*, **54**, 174201.

Emery, C.A., E. Tai, E., R. E. Morris, G. Yarwood. 2009a. Reducing Vertical Transport Over Complex Terrain in CMAQ and CAMx; AWMA Guideline on Air Quality Models Conference, Raleigh, NC, October 26-30, 2009.

Emery, C.A., E. Tai, R.E. Morris, G. Yarwood. 2009b. Reducing Vertical Transport Over Complex Terrain in Photochemical Grid Models; 8th Annual CMAS Conference, Chapel Hill, NC, October 19-21, 2009.

Emery, C.A., E. Tai, G. Yarwood and R. Morris. 2011. Investigation into approaches to reduce excessive vertical transport over complex terrain in a regional photochemical grid model. *Atmos. Env.*, Vol. 45, Issue 39, December 2011, pp. 7341-7351. (http://www.sciencedirect.com/science/article/pii/S1352231011007965).

Emery, C.A., Z. Liu, A. Russell, M. Odman, G. Yarwood and N. Kumar. 2017. Recommendations on statistics and benchmarks to assess photochemical model performance, Journal of the Air & Waste Management Association, 67:5, 582-598, DOI: 10.1080/10962247.2016.1265027

EPA. 2001. "Guidance Demonstrating Attainment Air Quality Goals for $PM_{2.5}$ and Regional Haze". Draft Final (17 February 2005), U.S. Environmental Protection Agency, Atmospheric Sciences Modeling Division, Research Triangle Park, N.C.

EPA. 2003a. A Conceptual Model to Adjust Fugitive Dust Emissions to Account for Near Source Particle Removal in Grid Model Applications, prepared by Tom Pace, U.S. EPA, August. http://www.epa.gov/ttn/chief/emch/invent/statusfugdustemissions_082203.pdf.

EPA. 2005a. Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-hr Ozone NAAQS -- Final. U.S. Environmental Protection Agency, Atmospheric Sciences Modeling Division, Research Triangle Park, N.C. October.

EPA. 2005b. "Regional Haze Regulations and Guidelines for Best Available Technology (BART) Determinations". Fed. Reg./Vol. 70, No. 128/Wed. July, Rules and Regulations, pp. 39104-39172. 40 CFR Part 51, FRL-7925-9, RIN AJ31.

EPA. 2007. Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, $PM_{2.5}$ and Regional Haze. U.S. Environmental Protection Agency, Research Triangle Park, NC. EPA-454/B-07-002. April. (http://www.epa.gov/ttn/scram/guidance/guide/final-03-pm-rh-guidance.pdf).

EPA. 2014a. Motor Vehicle Emissions Simulator (MOVES) – User Guide for MOVES2014. Assessment and Standards Division, Office of Transportation and Air Quality, U.S. Environmental Protection Agency. (EPA-420-B-14-055). July. (http://www.epa.gov/oms/models/moves/documents/420b14055.pdf).

EPA. 2014b. Motor Vehicle Emissions Simulator (MOVES) –MOVES2014 User Interface Manual. Assessment and Standards Division, Office of Transportation and Air Quality, U.S. Environmental Protection Agency. (EPA-420-B-14-067). July. (http://www.epa.gov/oms/models/moves/documents/420b14057.pdf).

EPA. 2014c. Motor Vehicle Emissions Simulator (MOVES) –MOVES2014 Software Design Reference Manual. Assessment and Standards Division, Office of Transportation and Air Quality, U.S. Environmental Protection Agency. (EPA-420-B-14-058). December. (http://www.epa.gov/oms/models/moves/documents/420b14056.pdf).

EPA. 2014d. Meteorological Model Performance for Annual 2011 WRF v3.4 Simulation, U.S. Environmental Protection Agency. November 2014.

EPA. 2014e. Draft Modeling Guidance for Demonstrating Attainment of Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze, U.S. Environmental Protection Agency. December 2014.

EPA. 2015. Air Quality Modeling Technical Support Document for the 2008 Ozone NAAQS Transport Assessment. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. January. (http://www.epa.gov/airtransport/O3TransportAQModelingTSD.pdf).

EPA. 2016a. Technical Support Document (TSD) Preparation of Emissions Inventories for the Version 6.3, 2011 Emissions Modeling Platform. U.S. Environmental Protection Agency. August 2016.

EPA. 2016b. Air Quality Modeling Technical Support Document for the 2015 Ozone NAAQS Preliminary Interstate Transport Assessment. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. December 2016.

EPA. 2017a. Use of Photochemical Grid Models for Single-Source Ozone and secondary PM2.5 impacts for Permit Program Related Assessments and for NAAQS Attainment Demonstrations for Ozone, $PM_{2.5}$ and Regional Haze. Memorandum from Tyler Fox, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. August 2017. https://www3.epa.gov/ttn/scram/guidance/clarification/20170804-Photochemical_Grid_Model_Clarification_Memo.pdf

EPA. 2017b. Documentation for the EPA's Preliminary 2028 Regional Haze Modeling. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. October.

Gery, M. W., G.Z. Whitten, J.P. Killus, and M.C. Dodge. 1989. A photochemical mechanism for urban and regional-scale computer modeling. *J. Geophys. Res.* 94, 12925-12956.

Gilliam, R. 2010. Evaluation of Multi-Annual CONUS 12 km WRF Simulations. U.S. Environmental Protection Agency, NREL, Atmospheric Modeling and Analysis Division. (http://epa.gov/scram001/adhoc/gilliam2010.pdf).

Gilliam, R.C. and J.E. Pleim, 2010. Performance Assessment of New Land Surface and Planetary Boundary Layer Physics in the WRF-ARW. J. Appl. Meteor. Climatol., 49, 760–774.

Hand, J. L., and W. C. Malm (2006), Review of the IMPROVE equation for estimating ambient light extinction coefficients, CIRA Report, ISSN: 0737-5352-71, Colo. State Univ., Fort Collins. http://vista.cira.colostate.edu/improve/Publications/GrayLit/016_IMPROVEeqReview/IMPROVEeqReview.htm

Henderson, B.H., F. Akhtar, H.O.T. Pye, S.L. Napelenok, W.T. Hutzell, 2014. A Database and Tool for Boundary Conditions for Regional Air Quality Modeling: Description and Evaluations, Geoscientific Model Development, 7, 339-360.

Hong, S-Y, Y. Noh, and J. Dudhia, 2006. A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. Mon. Wea. Rev., 134, 2318–2341.

Houyoux, M.R., Vukovich, J.M., Coats, C.J., Wheeler, N.J.M., Kasibhatla, P.S.,2000. Emissions Inventory Development and Processing for the Seasonal Model for Regional Air Quality. (SMRAQ) project, Journal of Geophysical Research – Atmospheres, 105(D7), 9079-9090.

Iacono, M.J., J.S. Delamere, E.J. Mlawer, M.W. Shephard, S.A Clough, and W.D. Collins, 2008. Radiative Forcing by Long-Lived Greenhouse Gases: Calculations with the AER Radiative Transfer Models, J. Geophys. Res., 113, D13103.

Kain, J.S., 2004. The Kain-Fritsch Convective Parameterization: An Update, J. Appl. Meteor.,43, 170-181.

Ma, L-M. and Tan Z-M, 2009. Improving the Behavior of Cumulus Parameterization for Tropical Cyclone Prediction: Convective Trigger, Atmospheric Research, 92, 190-211.

ME DEP, 2018. Tracking Visibility Progress, 2004-2016, (1st RH SIP Metrics). Maine Department of Environmental Protection for the Mid-Atlantic/Northeast Visibility Union. February.

Michalakes, J., J. Dudhia, D. Gill, J. Klemp and W. Skamarock. 1998. Design of a Next-Generation Regional Weather Research and Forecast Model. Mesoscale and Microscale Meteorological Division, National Center for Atmospheric Research, Boulder, CO. (http://www.mcs.anl.gov/~michalak/ecmwf98/final.html).

Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff and W. Skamarock. 2001. Development of a Next-Generation Regional Weather Research and Forecast Model. Developments in Teracomputing: Proceedings of the 9th ECMWF Workshop on the Use of High Performance Computing in Meteorology. Eds. Walter Zwieflhofer and Norbet Kreitz. World Scientific, Singapore. Pp. 269-276. (http://www.mmm.ucar.edu/mm5/mpp/ecmwf01.htm).

Michalakes, J., J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock and W. Wang. 2004. The Weather Research and Forecast Model: Software Architecture and Performance. Proceedings of the 11th ECMWF Workshop on the Use of High Performance Computing in Meteorology. October 25-29, 2005, Reading UK. Ed. George Mozdzynski. (http://wrf-model.org/wrfadmin/docs/ecmwf_2004.pdf).

Morrison, H.J., A. Curry, and V.I. Khvorostyanov, 2005. A New Double-Moment Microphysics Parameterization for Application in Cloud and Climate Models. Part I: Description, J. Atmos. Sci., 62, 1665–1677.

Morrison, H. and A. Gettelman, 2008. A New Two-Moment Bulk Stratiform Cloud Microphysics Scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I: Description and Numerical Tests, J. Climate, 21, 3642-3659.

Pleim, J.E. and A. Xiu, 2003. Development of a Land-Surface Model. Part II: Data Assimilation, J. Appl. Meteor., 42, 1811–1822

Pleim, J.E., 2007a. A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing, J. Appl. Meteor. Climatol., 46, 1383-1395.

Pleim, J.E., 2007b. A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part II: Application and Evaluation in a Mesoscale Meteorological Model, J. Appl. Meteor. Climatol., 46, 1396–1409.

Ramboll Environ, 2014. wrfcamx Version 4.3 Release Notes. December 17, 2014.www.camx.com. Ramboll Environ International Corporation, Novato, CA.

Ramboll Environ, 2016. User's Guide Comprehensive Air Quality Model with Extensions Version 6.40, www.camx.com. Ramboll Environ International Corporation, Novato, CA.

Simon, H., K. Baker and S. Phillips. 2012. Compilations and Interpretation of Photochemical Model Performance Statistics Published between 2006 and 2012. *Atmos. Env.* 61 (2012) 124-139. December. (http://www.sciencedirect.com/science/article/pii/S135223101200684X).

Skamarock, W. C. 2004. Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra. *Mon. Wea. Rev.*, Volume 132, pp. 3019-3032. December. (http://www.mmm.ucar.edu/individual/skamarock/spectra_mwr_2004.pdf).

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang and J. G. Powers. 2005. A Description of the Advanced Research WRF Version 2. National Center for Atmospheric Research (NCAR), Boulder, CO. June. (http://www.mmm.ucar.edu/wrf/users/docs/arw_v2.pdf)

Skamarock, W. C. 2006. Positive-Definite and Monotonic Limiters for Unrestricted-Time-Step Transport Schemes. *Mon. Wea. Rev.*, Volume 134, pp. 2241-2242. June. (http://www.mmm.ucar.edu/individual/skamarock/advect3d_mwr.pdf).

Stammer, D., F.J. Wentz, and C.L. Gentemann, 2003. Validation of Microwave Sea Surface Temperature Measurements for Climate Purposes, J. of Climate, 16(1), 73-87.

Sullivan D.C., Raffuse S.M., Pryden D.A., Craig K.J., Reid S.B., Wheeler N.J.M., Chinkin L.R., Larkin N.K., Solomon R., and Strand T. (2008) Development and applications of systems for modeling emissions and smoke from fires: the BlueSky smoke modeling framework and SMARTFIRE: 17th International Emissions Inventory Conference, Portland, OR, June 2-5. Available at: http://www.epa.gov/ttn/chief/conferences.html.

UNC. 2008. Atmospheric Model Evaluation Tool (AMET) User's Guide. Institute for the Environment, University of North Carolina at Chapel Hill. May 30. (https://www.cmascenter.org/amet/documentation/1.1/AMET_Users_Guide_V1.1.pdf).

UNC and ENVIRON, 2014a. Three-State Air Quality Modeling Study (3SAQS) – Draft Modeling Protocol 2011 Emissions & Air Quality Modeling. University of North Carolina at Chapel Hill and ENVIRON International Corporation, Novato, CA. July. (http://vibe.cira.colostate.edu/wiki/Attachments/Modeling/3SAQS_2011_Modeling_Protocol_Finalv2.pdf).

UNC. 2015. SMOKE v3.6.5 User's Manual. University of North Carolina at Chapel Hill, Institute for the Environment. (https://www.cmascenter.org/smoke/documentation/3.6.5/html/).

UNC and ENVIRON, 2015a. Three-State Air Quality Modeling Study (3SAQS) – Weather Research Forecast 2011 Meteorological Model Application/Evaluation. University of North Carolina at Chapel Hill and ENVIRON International Corporation, Novato, CA. March 5. (http://vibe.cira.colostate.edu/wiki/Attachments/Modeling/3SAQS_2011_WRF_MPE_v05Mar2015.pdf).

Wesely, M.L. 1989. Parameterization of Surface Resistances to Gaseous Dry Deposition in Regional-Scale Numerical Models. *Atmos. Environ.*, 23, 1293-1304.

Xiu, A. and J.E. Pleim. 2000. Development of a land surface model. Part I: application in a mesoscale meteorology model. *J. App. Met.*, 40, pp. 192-209.

Yantosca, B. 2004. GEOS-CHEMv7-01-02 User's Guide, Atmospheric Chemistry Modeling Group, Harvard University, Cambridge, MA.

Yarwood, G., R.E. Morris, G.M. Wilson. 2004. Particulate Matter Source Apportionment Technology (PSAT) in the CAMx Photochemical Grid Model. Proceedings of the 27th NATO/ CCMS International Technical Meeting on Air Pollution Modeling and Application. Springer Verlag (Available from http://camx.com/publ/pdfs/yarwood_itm_paper.pdf).

Yarwood, G., J. Jung, G. Z. Whitten, G. Heo, J. Mellberg and M. Estes. 2010. Updates to the Carbon Bond Mechanism for Version 6 (CB6). 2010 CMAS Conference, Chapel Hill, NC. October. (http://www.cmascenter.org/conference/2010/abstracts/emery_updates_carbon_2010.pdf)

Zhang, L., S. Gong, J. Padro, L. Barrie. 2001. A size-segregated particle dry deposition scheme for an atmospheric aerosol module. *Atmos. Environ.*, **35**, 549-560.

Zhang, L., J. R. Brook, and R. Vet. 2003. A revised parameterization for gaseous dry deposition in air-quality models. *Atmos. Chem. Phys.*, **3**, 2067–2082.