UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

RESEARCH TRIANGLE PARK, NC 27711

NOV 29 2018

OFFICE OF AIR QUALITY PLANNING AND STANDARDS

MEMORANDUM

SUBJECT: Model Guidance on Demonstrating Air Quality Goals for Ozone, PM2.5, and Haze

FROM: Richard A. Wayland, Director, Air Quality Assessment Division

TO: Regional Air Division Directors, Regions 1-10

The Environmental Protection Agency (EPA) is issuing *Guidance on Demonstrating Air Quality Goals for Ozone, PM2.5, and Haze*. The document contains recommendations for state, local, and tribal air quality agencies to use when developing air quality management plans to demonstrate attainment of the National Ambient Air Quality Standards (NAAQS) for ozone, particulate matter (PM), and haze. The document takes into account the decisions in the December 4, 2018 D.C. Circuit Court of Appeals decision in *Natural Resources Defense Council v. EPA*, 902 F.3d 1160 (D.C. Cir. 2018).

This guidance supersedes the April 1, 2014 "Guidance for 1-Hour SO2 Nonattainment Area SIP Submissions" and the March 2015 "Guidance for Ozone, Particulate Matter, and Regional Haze SIP Submissions" to the extent that those documents provide guidance on how to meet the requirements of the Clean Air Act (CAA) for air quality goals. This guidance does not impose binding, enforceable requirements on any party. Instead, it describes how EPA intends to exercise its discretion in reviewing state implementation plans (SIPs) and tribal implementation plans (TIPs) in light of the court's decision. The guidance does not substitute for the applicable statutory provisions or regulations. Readers should consult the applicable statutory provisions and regulations when considering how to apply the guidance in specific circumstances. EPA expects to update the guidance as needed to reflect future court decisions or other developments. EPA will post the guidance on its website for public review and comment.

In December 2018, EPA also issued a draft guidance on SIP submittals for the 2015 ozone NAAQS. EPA plans to update the guidance as needed to reflect future court decisions or other developments. EPA expects to update the guidance to reflect the court's decision in the case of the 2015 ozone NAAQS. Implementation guidance for the 2015 ozone NAAQS will be issued in the near future.

ED_002666_00014243-00001

Similar to the draft guidance, the updated guidance is divided into two main sections. The first part describes how to setup and apply a photochemical modeling platform (including meteorological, emissions, and air quality modeling), and the second part describes how to use the air quality modeling results to show whether future attainment of the ozone and/or $PM_{2.5}$ NAAQS is likely. The guidance also describes how to use photochemical grid modeling to evaluate reasonable progress goals for regional haze.

Many of guidance updates are in the form of revised and reorganized text. In addition, numerous references have been added and/or updated. The updated guidance reflects the requirements contained in the 2015 ozone NAAQS SIP requirements rule (November 2018)[1], the 2012 $PM_{2.5}$ NAAQS SIP requirements rule (August 2016)[2], and the regional haze rule (January 2017)[3].

Specific updates to the draft guidance include:

- Reorganization of the document including removal of outdated language and references.
- The 2014 draft guidance updates to the recommended 8-hour ozone attainment test (relative response factors using the top ten modeled days) were finalized, and now also apply to the 2015 ozone NAAQS.
- Consistent with the 2017 regional haze rule revisions, the reasonable progress goals methodology has been updated to reference the 20 percent "most anthropogenically impaired" days and 20 percent "clearest" days.
- The emissions modeling section has been extensively updated to account for new and improved emissions models, techniques, and input data, and has been closely coordinated with the SIP emissions inventory guidance, "*Emissions Inventory Guidance for Implementation of Ozone and Particulate Matter National Ambient Air Quality Standards (NAAQS) and Regional Haze Regulations*," available at: https://www.epa.gov/air-emissions-inventories/air-emissions-inventory-guidance-implementation-ozone-and-particulate.

Please share this guidance with air agencies in your Region. If you have any questions concerning this document, please contact Brian Timin at (919) 541-1850 or *timin.brian@epa.gov*. The guidance document is available electronically on the EPA's website: https://www3.epa.gov/ttn/scram/guidance_sip.htm.

Attachment

[1] *See* https://www.epa.gov/ground-level-ozone-pollution/implementation-2015-national-ambient-air-quality-standards-ozone
[2] *See* 81 FR 58101
[3] *See* 82 FR 3078

Modeling Guidance for Demonstrating Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze

EPA 454/R-18-009
November 2018

Modeling Guidance for Demonstrating Air Quality Goals for Ozone, $PM_{2.5}$, and Regional Haze

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Air Quality Assessment Division
Research Triangle Park, NC

ED_002666_00014243-00005

Table of Contents

Table of Contents 4
1.0 Introduction 8
1.1 What Is the Purpose of This Document? 9
1.2 Does the Guidance in This Document Apply to Me? 9
1.3 Outline 10
2.0 Building a Model Platform 12
2.1 Conceptual Description 12
2.1.1 Example Applications of Conceptual Models 14
2.2 Modeling Protocol and Supporting Documentation 14
2.2.1 Modeling Protocol 14
2.2.2 Attainment Demonstration and/or Reasonable Progress Goals Modeling Documentation Package 16
2.3 Episode Selection 17
2.3.1 Choosing Time Periods to Model 18
2.3.2 Future Year Selection 20
2.4 Modeling Domain Selection 21
2.4.1 Domain Size 21
2.4.2 Vertical Layer Configuration 22
2.4.3 Horizontal Grid Cell Size 23
2.5 Air Quality Model Selection 24
2.6 Meteorological Inputs 27
2.6.1 Developing Base Year Meteorological Fields 28
2.6.1.1 Selecting a Model Domain 29
2.6.1.2 Selecting Physics Options 29
2.6.1.3 Use of Data Assimilation 30
2.6.2 Assessing Impacts of Future Year Meteorology 31
2.6.3 Evaluation of Base Year Meteorological Fields 32
2.6.3.1 Operational Evaluation 33
2.6.3.2 Phenomenological Evaluation 34
2.7 How Are the Emission Inputs Developed? 34
2.7.1 What Emissions Inventory Years Should I use for Base and Future Modeling? 35
2.7.2 What Base Year Emission Inventory Data Are Needed to Support Air Quality Models? 36
2.7.3 What Pollutants Should Be Included in Base Year Inventories? 36
2.7.4 What Temporal Resolution Should Be Used for Base Year Inventories? 37
2.7.5 What Spatial Coverage Is Needed for Base Year Inventories? 38
2.7.5.1 Areas Within the Nonattainment Area 38
2.7.5.2 Areas Outside of the Nonattainment Area 38
2.7.5.3 Areas Outside of the United States 39
2.7.6 How Can the National Emission Inventory Be Used by Air Agencies? 39

ED_002666_00014243-00006

2.7.7 How Should Base Year Emissions Be Created? ... 40
2.7.7.1 Point Sources ... 41
2.7.7.2 Nonpoint Sources ... 42
2.7.7.3 On-Road Mobile ... 42
2.7.7.4 Nonroad Mobile Equipment ... 44
2.7.7.5 Nonroad Mobile: Commercial Marine Vessels, Locomotives, and Aircraft 45
2.7.7.5 Biogenic Emissions ... 46
2.7.7.6 Geogenic and Other Natural Sources ... 47
2.7.7.7 Wildland and Cropland Fires ... 49
2.7.8 What Other Data Are Needed to Support Emissions Modeling? ... 50
2.7.8.1 Temporal Allocation ... 50
2.7.8.2 Chemical Speciation ... 53
2.7.8.3 Spatial Allocation ... 55
2.7.8.4 Other Ancillary Inputs ... 55
2.7.9 How Are Inventory Data Converted into Air Quality Model Input? ... 56
2.7.9.1 Emissions Models ... 56
2.7.9.2 Biogenic Emissions Models ... 57
2.7.10 Other Emissions Modeling Issues ... 58
2.7.10.1 Elevated Point Sources ... 59
2.7.10.2 Treatment of 'Atypical' Sources of Emissions ... 59
2.7.10.3 Transport and Meteorology-based Reductions in Fugitive Dust Emissions ... 59
2.7.10.4 Quality Assurance ... 60
2.7.11 How Are Emissions Estimated for Future Years? ... 61
2.8 Initial and Lateral Boundary Conditions ... 62
3.0 Evaluating Model Performance ... 67
3.1 Overview of Model Performance Evaluation ... 67
3.2 Operational Evaluation ... 69
3.2.1 Metrics ... 70
3.2.2 Plots ... 73
3.2.3 Ozone Model Performance ... 78
3.2.4 $PM_{2.5}$ Model Performance ... 79
3.3 Ambient Measurement Networks ... 81
3.3.1 AQS ... 81
3.3.2 NCore ... 82
3.3.3 IMPROVE ... 82
3.3.4 CASTNet ... 82
3.3.5 CSN ... 83
3.3.6 NADP ... 83
3.3.7 SEARCH ... 84
3.3.8 Speciated $PM_{2.5}$ Data Summary ... 84
3.4 Diagnostic Evaluation ... 90
3.4.1 Process Analysis ... 91

ED_002666_00014243-00007

3.4.2 Indicator Ratios ... 92
3.4.3 Source Apportionment ... 95
3.4.4 Sensitivity Studies ... 96
3.4.5 Dynamic Model Evaluation ... 97
3.5 Evaluation Tools ... 98
4.0 Assessing Modeled Attainment for Ozone and $PM_{2.5}$ ... 99
4.1 Overview of Modeled Attainment Test ... 99
4.1.1 Establishing the Base Design Value ... 101
4.1.2 Calculation of Relative Response Factors ... 103
4.2 Modeled Attainment Test for the Primary Ozone Standard ... 103
4.2.1 Model Values to Use in the RRF Calculation ... 104
4.2.2 Grid Cells to Use in the RRF Calculation ... 106
4.2.3 Sample Modeled Attainment Test Calculation ... 107
4.3 Modeled Attainment Test for the Secondary Ozone Standard ... 110
4.4 What is the Modeled Attainment Tests for the Annual Average $PM_{2.5}$ NAAQS? ... 111
4.4.1 Ambient $PM_{2.5}$ Data Used in the Annual $PM_{2.5}$ Attainment Test ... 113
4.4.2 FRM Monitors that Do Not Have Speciation Data ... 113
4.4.3 $PM_{2.5}$ Species Calculations and Definitions ... 115
4.4.4 Recommended Treatment of Species Data ... 116
4.4.4.1 Retained Nitrate Mass ... 116
4.4.4.2 Ammonium Associated with Sulfates and Retained Nitrates ... 117
4.4.4.3 Particle Bound Water ... 118
4.4.4.4 Salt ... 119
4.4.4.5 Other Primary PM2.5 ... 119
4.4.4.6 Blank mass ... 120
4.4.4.7 Calculation of Carbonaceous Mass ... 121
4.4.4.8 Summary of $PM_{2.5}$ Composition Calculations ... 122
4.5 What Is the Recommended Modeled Attainment Test for the 24-Hour NAAQS? ... 127
4.6 Local Area Analyses - Special Considerations for Primary $PM_{2.5}$ ... 133
4.6.1 Analysis of Primary $PM_{2.5}$ Impacts at Monitors ... 134
4.6.2 Analysis of Primary $PM_{2.5}$ Impacts at Monitors Using a Gaussian Dispersion Model ... 135
4.6.2.1 Double Counting ... 137
4.6.2.2 Analysis of Primary $PM_{2.5}$ at Monitors- Quality Assurance ... 137
4.7 Estimating design values at unmonitored locations ... 138
4.7.1 Why does the unmonitored area analysis need to use both ambient data and model output? ... 139
4.7.2 Implementation of Gradient Adjusted Spatial Fields ... 140
4.7.3 Using the Results of the Unmonitored Area Analysis ... 142
5.0 What Is the Recommended Modeling Method for Developing Air Quality Goals for Regional Haze? ... 143
5.1 Regional Haze Rule Background ... 143
5.2 How is Regional Haze Measured? ... 145
5.2.1 IMPROVE Algorithm ... 146

ED_002666_00014243-00008

5.2.2 Deciview Haze Index 148
5.2.3 Estimating Mass Associated with Components of Particulate Matter 148
5.2.4 Normalizing Trends in Regional Haze - f(rh) Factors 150
5.3 How Do I Apply a Model to Calculate Changes in Visibility? 151
5.4 How Do I Select Appropriate Inputs For Developing RPGs? 163
6.0 How Can Additional Analyses Be Used to Support an Ozone or $PM_{2.5}$ Attainment Demonstration? 169
6.1 What Types of Additional Analyses Should Be Completed as Part of an Ozone or $PM_{2.5}$ Attainment Demonstration? 170
6.1.1 Modeling Analyses 171
6.1.2 Analyses and Trends in Ambient Air Quality and Emissions 173
6.1.3 Additional Emissions Controls/Reductions 175
6.2 Weight of Evidence Summary 176
7.0 What Role Should Additional and/or Supplemental Analyses Play in Regional Haze Modeling Demonstrations? 178
7.1 Additional air quality modeling 178
7.2 Review of trends 178
7.3 Other models and tools 178
7.4 Refinements to the recommended visibility analysis 179
7.5 Concerns about modeling the clearest days 180
8.0 References 181

ED_002666_00014243-00009

1.0 Introduction

This document describes how to estimate the impacts of an emissions control strategy on air quality for purposes of demonstrating attainment of the annual average and 24-hour average national ambient air quality standards (NAAQS) for particles smaller than 2.5 µm in diameter (*$PM_{2.5}$*) and the 8-hour NAAQS for ozone. We also describe how to use modeled and monitored air quality data to estimate future visibility conditions in *Class I areas* (e.g., national parks, wilderness areas) as part of the development of reasonable progress goals (RPGs) that reflect the long-term strategy in a regional haze state implementation plan (SIP).[1]

This document describes how to apply air quality models to generate the predictions used to evaluate attainment and/or to set RPGs for regional haze. Modeling to show attainment of the NAAQS primarily applies to nonattainment areas[2] for which modeling is required, or desired. Modeling to assess reasonable progress for regional haze applies to all states, the District of Columbia, and the U.S. Virgin Islands.[3]

This guidance is designed to implement national policy on air quality modeling requirements as embodied in the Clean Air Act (CAA), the Ozone SIP Requirements Rule[4], the $PM_{2.5}$ SIP Requirements Rule[5], and the Regional Haze Rule[6]. In addition, the United States Environmental Protection Agency (EPA) revised 40 CFR part 51, Appendix W (Guideline on Air Quality Models) (U.S. EPA, 2017a), hereafter referred to as "Appendix W," in 2017 to provide general information about the types of model approaches that may be appropriate for the purposes of demonstrating attainment with the ozone and $PM_{2.5}$ NAAQS (section 5) and projecting visibility impacts at Class I areas (section 6). Appendix W is a regulation and, as such, contains binding requirements. However, the Appendix W photochemical modeling requirements related to SIP

---

[1] Modeling to determine RPGs as described in this document is part of the development of a regional haze SIP. However, unlike the NAAQS, the Regional Haze program does not involve fixed air quality standards that must be achieved. The ways that RPGs are used in the development of a regional haze SIP are described in more detail in section 5.

[2] While this guidance document is primarily directed at modeling applications in nonattainment areas, it may also be useful as a guide for modeling to support NEPA analyses, maintenance plans or to support other rules or provisions of the Clean Air Act.

[3] *See* 40 CFR 51.300(b).

[4] *See* Implementation of the 2015 National Ambient Air Quality Standards for Ozone: Nonattainment Area State Implementation Plan Requirements; [Final rule signed November 7, 2018]

[5] *See* Fine Particulate Matter National Ambient Air Quality Standards: State Implementation Plan Requirements ($PM_{2.5}$ SIP Requirements Rule), 81 FR 58101 (Aug. 24, 2016).

[6] The Regional Haze Rule was revised by a final rule published on January 10, 2017 (82 FR 3078). This version of the modeling guidance reflects that revision of the rule. Additional revisions may be made to this guidance to be consistent with any future regional haze rule and/or guidance issued by EPA.

ED_002666_00014243-00010

attainment demonstration modeling and regional haze modeling do not proscribe any particular models or modeling techniques. As such, this modeling guidance provides additional details about how to set up and run photochemical models that may be useful to air agencies that are required to submit SIP attainment demonstrations and/or regional haze modeling.

This guidance is intended for use by the EPA headquarters and Regional offices; federal land managers of mandatory Class I federal areas; state, local and tribal air quality management authorities, and the general public. This document does not substitute for provisions or regulations of the CAA enumerated above, nor is it a regulation itself. As the term "guidance" suggests, it provides recommendations on how to implement the modeling requirements. Thus, it does not impose binding, enforceable requirements on any party, nor does it assure that the EPA will approve all instances of its application, as the guidance may not apply to a particular situation based upon the circumstances.

The EPA and state, local, and tribal air agencies (hereinafter referred simply as "air agency" or "air agencies") retain the discretion to adopt approaches on a case-by-case basis that differ from this guidance where appropriate. Final decisions by the EPA regarding a particular SIP submission that includes a modeling demonstration will ultimately be made based on the statute and applicable regulations and will only be made following a final submission by air agencies and after notice and opportunity for public review and comment. Interested parties are free to raise questions and objections about the appropriateness of the application of this guidance to a particular situation; the EPA and air agencies should consider whether or not the recommendations in this guidance are appropriate in that situation.

1.1 What Is the Purpose of This Document?

This document has two purposes. The first purpose is to describe how to apply an air quality model to produce results needed to support an attainment demonstration or to calculate RPGs for a regional haze reasonable progress analysis in a scientifically appropriate manner. The second is to explain how to interpret whether results of modeling and other analyses support a conclusion that attainment of the ozone and/or $PM_{2.5}$ NAAQS will occur by the appropriate date for an area, and/or to assess progress towards the national visibility goal of the regional haze program.

1.2 Does the Guidance in This Document Apply to Me?

This guidance applies to all air agencies that are required to submit a State Implementation Plan (SIP), or Tribal Implementation Plan (TIP) submission with an attainment plan designed to

ED_002666_00014243-00011

achieve attainment of the ozone and/or $PM_{2.5}$ NAAQS. The guidance also applies to air agency SIP submissions developed to address Regional Haze Rule requirements. Air agencies required to submit an attainment demonstration and/or a reasonable progress analysis for regional haze are encouraged to follow the procedures described in this document. Details on when a state is required to submit a modeled attainment demonstration can be found in the Ozone SIP Requirements Rule[7] and the $PM_{2.5}$ SIP Requirements Rule[8]. Details on when a state is required to submit a regional haze SIP that includes RPGs can be found in the Regional Haze Rule[9].

1.3 Outline

Part 1 of this guidance provides an overview of the modeling and attainment demonstration process. Part 2 describes how to build a "modeling platform" and apply air quality models. A "modeling platform" consists of the building blocks of an air quality modeling demonstration. This includes emissions modeling[10], meteorological modeling, and the development of all other inputs needed to run an air quality model. The model platform development process consists of the following steps as outlined in section 2:

Section 2.1	Development of a conceptual description of the problem to be addressed
Section 2.2	Develop a modeling protocol
Section 2.3	Select appropriate meteorological time periods to model
Section 2.4	Choose an appropriate area to model with appropriate horizontal/vertical resolution
Section 2.5	Select an appropriate model to support the demonstration
Section 2.6	Generate meteorological inputs to the air quality model
Section 2.7	Generate emissions inputs to the air quality model
Section 2.8	Develop initial and lateral boundary conditions that are suitable for the application

Typically, the air quality modeling process starts with the development of base year emissions and meteorology for input to an air quality model to evaluate model performance. Section 3

[7] *See* 40 CFR 51.1308.
[8] *See* 40 CFR 51.1011.
[9] *See* 40 CFR 51.308. Sections 308(d) and (e) apply for SIPs for the first implementation period only. States that followed section 51.309 (which is also part of the Regional Haze Rule) in the first implementation period were not required to develop RPGs. Section 51.308(f) applies to all Regional Haze SIPs due in 2021 and later.
[10] Additional emissions inventory requirements and details on emissions modeling are contained in "Emissions Inventory Guidance for Implementation of Ozone and Particulate Matter National Ambient Air Quality Standards (NAAQS) and Regional Haze Regulations."

ED_002666_00014243-00012

describes the process for evaluating model performance and performing diagnostic analyses. After evaluating the model and making any necessary input changes or adjustments, the model is run for a future year, which corresponds to the appropriate attainment year for the area or to the future year modeled for regional haze planning purposes. Sections 2 and 3 apply to both ozone and $PM_{2.5}$ modeling and modeling to set RPGs for regional haze SIPs.

Section 4 describes how the air quality model outputs are then used to apply the modeled attainment test to support an attainment demonstration for ozone or $PM_{2.5}$. We explain what is meant by a *modeled attainment demonstration*, a *modeled attainment test*, and a *weight of evidence demonstration*. Modeled attainment tests are described for the 8-hr ozone NAAQS and the annual and 24-hour $PM_{2.5}$ NAAQS. Section 5 addresses aspects of modeling specific to regional haze SIPs.

Model applications require a substantial effort. Air agencies are encouraged to work closely with the appropriate EPA Regional office(s) in executing each step of the modeling and attainment demonstration or regional haze planning process. This will increase the likelihood of approval of the demonstration at the end of the process.

ED_002666_00014243-00013

2.0 Building a Model Platform

2.1 Conceptual Description

The first step in developing an attainment demonstration is to construct a conceptual description of the problem that is being addressed. Conceptual descriptions, which are also referred to as conceptual models, are comprehensive summaries of the "state of the knowledge" regarding the influence of emissions, meteorology, transport, and other relevant atmospheric processes on air quality in the area (Vickery, 2004). For a conceptual description to be informative, it should identify what processes and sources, in the generic sense, are most responsible for the air quality issue being simulated. Well-constructed conceptual models can substantially inform the design of the attainment demonstration modeling (e.g., episode selection, choice of domain, emissions priorities, evaluation focus) and should be conducted in advance of the development of a modeling protocol. It is worth noting that conceptual descriptions can be valuable in other air quality planning efforts besides attainment demonstrations, such as determining nonattainment area boundaries, investigating emissions control program impacts, interstate transport analyses, and monitoring network design, among others.

The following bullets describe some of the key building blocks in developing a conceptual model. The process steps discussed below are meant to illustrate one possible template for building a conceptual model to inform an air quality modeling demonstration; there may be alternate approaches that better suit individual study areas or issues, which air agencies should discuss with the appropriate EPA Regional office.

- Introduce the general nature of the air quality problem addressed by the conceptual model:
    - What are the pollutants of concern in the area?
    - What are the current air quality levels in the area?
    - What is the attainment/nonattainment status of the area?
    - What is the geographical scope of poor air quality?
    - What is the temporal scope of poor air quality?
    - What are the air quality trends in the area? Is the problem getting better or worse?
    - What are the suspected mechanisms for formation of poor air quality levels?
    - What are the sources of emissions that may contribute to poor air quality?
    - Are there unique meteorological influences on local air quality levels?

- Describe the ambient monitoring network used for the conceptual model:
    - Develop a map of monitor sites and types (e.g., FRM, FEM, CSN, IMPROVE, CASTNET, etc.).

ED_002666_00014243-00014

- Describe characteristics of individual monitoring sites (scale, frequency, etc.).

- Describe the status and trends of air quality in the area:
  - Summarize the relevant monitoring data for the air quality problem being studied.
  - Identify locations that are in violation of the NAAQS.
  - Describe the spatial pattern in pollutant and precursor pollutant levels.
  - Describe the temporal pattern in pollutant and precursor pollutant levels.
  - Describe the annual trends in concentrations over the period of interest.
  - If appropriate, develop fused ambient/model surfaces to fill any gaps within the monitoring network.

- Investigate possible relationships between emissions and air quality:
  - Examine emissions estimates for the main sector/source categories.
  - Compare emission trends for annual and or seasonal/episodic periods to corresponding air quality trends.
  - Identify key emission sources or source categories.
  - Assess the historical effectiveness of control programs.
  - Consider how future emissions growth or reductions may affect air quality.
  - List control programs that are in place, or will soon be implemented, that may impact emissions sources in the area.

- Investigate possible relationships between meteorology and air quality (AQ):
  - Describe meteorological characteristics on poor air quality days (e.g., wind speed/direction, temperatures, relative humidity levels, inversion indicators, etc.).
  - Identify any distinct meteorological phenomena that coincide with poor air quality days.
  - Prepare pollution roses, HYSPLIT back-trajectories (Draxler, 1998), or any other relevant analysis needed to link poor air quality days to specific meteorological patterns and sources of transported pollutants.
  - Assess the impact from pollutant transport into area based on meteorological data and ambient air quality data.

- Synthesize all the relevant information into a detailed conceptual model:
  - Aggregate all elements of the conceptual model and provide the key findings resulting from the analyses completed.
  - Characterize the potential factors that influence air quality in the area and where possible rank the importance of those influences.

- Identify issues that require further investigation, either in terms of data collection or chemical transport modeling.
- Discuss how the subsequent SIP modeling and associated protocol is impacted by this conceptual description.

2.1.1 Example Applications of Conceptual Models

The idea of developing a modeling protocol based upon a conceptual description of the sources and processes that lead to poor air quality is well-established (U.S. EPA, 2007a). Numerous excellent conceptual models of local and regional air quality issues have been developed over the past decade. Table 2-1 provides a non-comprehensive list of example conceptual models that have been developed since the last release of the guidance that may be useful in orienting future SIP modeling applications.

Table 2.1 Examples of recently developed conceptual models

Air quality issue	Reference
$PM_{2.5}$ / Regional Haze (Northeast U.S.)	MANE-VU (2013)
$PM_{2.5}$ (Midwest U.S.)	Lake Michigan Air Directors Consortium (LADCO) (2009)
Ozone (Austin TX)	McGaughey *et al.* (2010)
Ozone (Wyoming)	Stoeckenius (2010)
$PM_{2.5}$ (Southeast Michigan)	Turner (2008)

2.2 Modeling Protocol and Supporting Documentation

As with any technical support document designed to inform air quality planning, an attainment demonstration or regional haze analysis should be supported by documentation that sufficiently describes the procedures used in the analysis. In order to facilitate the process of EPA Regional office review and approval, we recommend the preparation of two separate supporting documents for ozone and $PM_{2.5}$ SIPs: one before the modeling analyses are initiated (modeling protocol) and one after the analyses have been completed (attainment demonstration package). We recommend a similar approach during the development of regional haze SIPs.

2.2.1 Modeling Protocol

Developing and implementing a modeling protocol is an important part of a modeling demonstration. The protocol should detail and formalize the procedures for conducting all

ED_002666_00014243-00016

phases of the modeling study, such as describing the background and objectives for the study, creating a schedule and organizational structure for the study, selection of appropriate period(s) for modeling, developing the input data, conducting model performance evaluations, interpreting modeling results, describing procedures for using the model to demonstrate whether proposed strategies are sufficient to attain the NAAQS and/or regional haze goals, and producing documentation to be submitted for EPA Regional office review and approval. The most important function of the modeling protocol is to serve as a blueprint for planning how the modeled demonstration will be performed. The protocol should be a valuable communication device by which air agencies, EPA, and other stakeholders can assess the applicability of default recommendations and develop area-specific alternatives, where needed, prior to conducting the work to build the modeling system. A suitable protocol should lead to extensive participation by stakeholders in developing the demonstration. It should also reduce the risk of spending time and resources on efforts that are unproductive or inconsistent with EPA rules, policy, and guidance. While the modeling protocol is initially developed at the beginning of a modeling exercise to foster communication, it is advisable to modify the document as needed throughout the modeling process when alterations from the original modeling plan are necessary. Again, any changes to the protocol should be fully communicated between affected air agencies, stakeholders, and the EPA.

Meaningful protocols should fully communicate the expected scope of the analysis and provide a blueprint for carrying out the needed analyses. There is no "one-size-fits-all" format for a sufficient modeling protocol, as different individual areas may require specific points of emphasis. However, past attainment demonstrations have yielded several excellent protocol examples that could serve as templates for any groups developing new modeling protocols. While not exhaustive, potentially valuable protocol references include attainment demonstration modeling efforts completed in the Dallas/Fort Worth area (Environ, 2003), the Denver region (Morris, 2011), and the San Joaquin Valley (CARB, 2012). Based on past experience, the EPA recommends that the following topics be core elements of any modeling protocol:

- Overview of the air quality issue being considered including historical background
- List of the planned participants in the analysis and their expected roles
- Schedule for completion of key steps in the analysis and final documentation
- Description of the conceptual model for the area
- Description of periods to be modeled, how they comport with the conceptual model, and why they are sufficient
- Models to be used in the demonstration and why they are appropriate
- Description of model inputs and their expected sources (e.g., emissions, meteorology, etc.)

ED_002666_00014243-00017

- Description and justification of the domain to be modeled (expanse and resolution)
- Process for evaluating base year model performance (meteorology, emissions, and air quality) and demonstrating that the model is an appropriate tool for the intended use
- Description of the future years to be modeled and how projection inputs will be prepared
- Description of the NAAQS attainment test procedures and (if known) planned weight of evidence, and/or description of the procedures for calculating RPGs from the modeling outputs, as applicable.
- Expected diagnostic or supplemental analyses needed to develop weight of evidence analyses
- Commitment to specific deliverables fully documenting the completed analysis

2.2.2 Attainment Demonstration and/or Reasonable Progress Goals Modeling Documentation Package

Whereas the modeling protocol describes the planned scope of the analysis, the final modeling documentation package summarizes the actual analysis conducted (including procedures used) to show that an area will likely meet the NAAQS and/or RPGs under a specific set of future conditions. This document can follow the same basic outline as the modeling protocol, highlighting those aspects of the modeling that may have changed from the original plans. Additionally, the modeling documentation package should have detailed information on any emission reduction strategies that will be implemented as part of an attainment or regional haze SIP. Ultimately, the modeling documentation package should provide a narrative that fully describes the technical rationale behind the projection of a specific air quality goal in an area. Based on past experience, the EPA recommends that the following topics be core elements of any modeling documentation package:

- Executive summary that provides an overview of the analysis and the key conclusions
- Reference to the modeling protocol noting any major deviations from the original plans
- List of the institutional participants in the attainment demonstration and their roles
- Description of air quality in the area and describe how that shaped the analysis
- Justification for the model, episodes, domain, and grid(s) used in the analysis
- Description of the development of the emissions inputs used in the base year modeling, including, at a minimum, tabular summaries by state/county, as appropriate

ED_002666_00014243-00018

- Description of the development of meteorological inputs used in the base year modeling
- Description of all other base year modeling inputs
- Evaluation of base year model performance (meteorology, emissions, and air quality), including a description of the observational database used in the evaluation and any diagnostic or sensitivity tests used to improve the model
- Description of the strategy used to demonstrate attainment and/or RPGs, including speciated emissions summaries for the future year and identification of authority for implementing these strategies
- Description of the attainment test inputs and results
- Description of any supplemental analyses designed to bolster the original attainment test results
- Detailed summary of the entire analysis that leads to the conclusion that the selected attainment demonstration strategy is likely to produce attainment of the NAAQS by the required date and/or leads to the calculation of the RPGs
- Appendices that contain more detailed information on the model inputs and outputs (emissions, meteorology, etc.), model performance, and attainment demonstration, including charts, tables, and descriptive text

All model input and output files (in electronic format) should be made available upon request by the EPA and/or stakeholders.

2.3 Episode Selection

The modeled attainment test and the recommended procedure for setting RPGs for regional haze both adjust observed ambient concentrations during a base case period (e.g., 2012-2016) to a future period (e.g., 2023) using model-derived "relative response factors" (RRFs). It is important that emissions used in the attainment modeling correspond with the period reflected by the chosen baseline design value period (e.g., 2012-2016).[11] Deviations from this constraint will diminish the credibility of the RRFs. The base year modeling inventory typically corresponds to the middle year of the baseline average design value period (e.g. 2014 for a 2012-2016 average design value period). Alternatively, the base year emissions can reflect multi-year average emissions from the base year period (e.g. average emissions for the 2013-2015 period). But in

[11] The regional haze program does not use design values. However, in an analogous approach, this guidance recommends the use of 5 years of historical visibility data (non-weighted) in a manner similar to the approach in the modeled attainment test for ozone and $PM_{2.5}$. For this reason and for brevity, sections 2 and 3 use the "design value" terminology in some explanations that apply to both ozone and $PM_{2.5}$ attainment modeling and to the setting of regional haze RPGs. That is, for regional haze purposes, "design values" can be understood to mean "visibility data."

ED_002666_00014243-00019

either case, the emissions year or years should be representative of the 5-year design value window.

There is no recommended default base year for modeling. However, it is recommended to use a relatively recent base period so that the emissions projection period is as short as possible and the base year ambient data is as current as possible. For example, projecting emissions from 2014 to 2023 (using a 2012-2016 base year average design value) should be less uncertain than projecting emissions from 2005 to 2023 (using a 2003-2007 base year average design value). The most recent ambient design values reflect actual emissions changes that have occurred over time. It is, therefore, better to use recent design values, which are actual measurements, combined with modeled emissions changes to project future concentrations, then to use older ambient data to estimate the change in design values. When selecting a base modeling year, air agencies should review recent ambient data and consider the factors in the following section.

2.3.1 Choosing Time Periods to Model

In the past, the choice of modeled episode days has been limited by the speed of computers and the ability to store model output files. With the advancement in computer technology over the past two decades, computer speed and storage issues are no longer an impediment to modeling long time periods. In fact, the majority of recent regulatory assessment modeling platforms have been inclusive of entire summers and/or full years (as appropriate) for ozone, $PM_{2.5}$, and regional haze (Boylan and Russell, 2006; Morris et al., 2006; Rodriguez et al., 2009; Simon et al., 2012; Tesche et al., 2006; U.S. EPA, 2011a, 2011b, 2016a).

Ozone-based research has shown that model performance evaluations and the response to emissions controls need to consider modeling results from relatively long time periods, in particular, full synoptic cycles or even full ozone seasons (Hogrefe et al., 2000; Vizuete et al., 2011). In order to examine the response to ozone control strategies, it may not be necessary to model a full ozone season (or seasons), but, at a minimum, modeling "longer" episodes that encompass full synoptic cycles is advisable. Time periods which include a ramp-up to a high ozone period and a ramp-down to cleaner conditions allow for a more complete evaluation of model performance under a variety of meteorological conditions.

Most model applications for the annual $PM_{2.5}$ NAAQS have modeled a full year (Morris et al., 2006; Tesche et al., 2006; U.S. EPA, 2011a, 2011b, 2017b). This is a logical goal since every monitored day of the year is included in the calculation of the annual NAAQS. The annual $PM_{2.5}$ NAAQS is unique because each and every ambient observation is included in the average. It is,

ED_002666_00014243-00020

therefore, likely that a full year of modeled data is needed to adequately represent the annual average $PM_{2.5}$ NAAQS.

Eight-hour ozone and 24-hour $PM_{2.5}$ episode selection criteria are similar because both standards are based on short-term peak concentration periods. Regional haze calculations are based on an average of 20 or more days per year (20% clearest and 20% most impaired days). Therefore, regional haze episode selection will likely include more days throughout the year.

At a minimum, several criteria should be used to select time periods that are appropriate to model:

- Model time periods that are close to the most recently compiled and quality assured National Emission Inventory (NEI). These comprehensive inventories are typically generated every 3 years (e.g. 2011, 2014, 2017, etc.), but often contain year-specific data for the intermediate years. Since NEI years will have a nationwide complete and comprehensive inventory, selecting a base modeling year that is also an NEI year may save resources. However, other factors should be considered when selecting a base modeling year, such as the availability and magnitude of observed ambient data, meteorology, and availability of special study data. After consideration of all factors, the most appropriate base year may or may not be an NEI year. See section 2.7 for more information on base year inventory selection issues.

- Model time periods in which observed concentrations are close to the appropriate base year design value or level of visibility impairment and ensure there are a sufficient number of days so that the modeled test applied at each monitor is based on multiple days.

- Model time periods both before and following elevated pollution concentration (poor air quality) episodes to ensure the modeling system appropriately characterizes low pollution periods, development of elevated periods, and transition back to low pollution periods through synoptic cycles.

- Simulate a variety of meteorological conditions conducive to elevated/poor air quality.

Primary Ozone (8-Hour Ozone) - Choose time periods which reflect a variety of meteorological conditions that frequently correspond with observed 8-hour daily maxima concentrations greater than the level of the NAAQS at monitoring sites in the nonattainment area.

ED_002666_00014243-00021

24-Hour $PM_{2.5}$ - Choose time periods that reflect a variety of meteorological conditions that frequently correspond with observed 24-hour average concentrations greater than the level of the NAAQS at monitoring sites in the nonattainment area.
Annual $PM_{2.5}$ - The best way to represent the meteorological variability within a season and over an entire year is to model an entire year that has meteorology generally conducive to elevated $PM_{2.5}$ concentrations.
Regional Haze - Choose time periods that reflect 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 year.
Long-term or seasonal ozone – Since long-term (seasonal) ozone assessments are based on accumulated ozone over multiple months during the ozone season, the most appropriate way to represent the variability in high and low ozone is to model an entire ozone season that is representative of long-term values. [**Note:** This analysis is only applicable if seasonal ozone assessments are necessary (e.g. to support ozone benefits calculations). Currently, there are no long-term or seasonal ozone NAAQS (primary or secondary).]

2.3.2 Future Year Selection

For ozone and $PM_{2.5}$, future emissions should be projected to the attainment year or appropriate time period based on the area's classification. The Ozone SIP Requirements Rule provides a schedule for implementing emission reductions needed to ensure attainment by the area's attainment date. Specifically, it states that emission reductions needed for attainment must be implemented no later than the beginning of the attainment year ozone season.[12] The $PM_{2.5}$ Implementation Rule contains similar provisions. It states that emissions reductions should be in place no later than the beginning of the year containing the applicable attainment date.[13]

As part of demonstrating attainment of the ozone and $PM_{2.5}$ NAAQS, air agencies are required to conduct a Reasonably Available Control Measures (RACM) analysis to determine if the attainment date can be advanced by at least a year. Since areas are required to attain as expeditiously as practicable, results of the RACM analysis may indicate attainment can be achieved earlier.

[12] *See* 40 CFR 51.1308(d).
[13] *See* 40 CFR 51.1011(a)(5) and 40 CFR 51.1011(b)(5).

ED_002666_00014243-00022

There are varying requirements related to attainment year and RACM analyses for ozone and $PM_{2.5}$ nonattainment areas, depending on classifications and other factors. Please see the appropriate ozone and/or $PM_{2.5}$ implementation rule for more details and information.[14]

For regional haze SIPs, RPGs must reflect the visibility conditions that are projected to be achieved by the end of the applicable implementation period.[15] Therefore, the modeling should be based on emissions forecasted for the end year of the implementation period addressed by the SIP. These end years are 2018, 2028, etc.

2.4 Modeling Domain Selection

A modeling domain identifies the geographical bounds of the area to be modeled. Horizontal resolution is the geographic size of individual grid cells within the modeling domain. Vertical resolution is specified in terms of multiple layers of the atmosphere between the surface and top of the model (usually near the tropopause).

2.4.1 Domain Size

The principal determinants of model domain size are the nature of the ozone, $PM_{2.5}$ and/or regional haze problem being modeled, and the spatial scale of the emissions that impact the nonattainment or Class I area. Establishment of a sufficiently large model domain that utilizes the output from a larger regional or global modeling simulation to feed hourly lateral boundary conditions is a preferred approach (see section 2.8 for more information on boundary conditions). However, regardless of the size of the modeling domain, sources outside the modeling domain may have an important influence on concentrations within nonattainment areas or Class I areas. Therefore, boundary conditions need to be well characterized. The grid containing the key emissions sources and receptors needs to have sufficiently fine-scale to resolve local gradients and have a large enough spatial extent to capture recirculation due to shifting wind directions. An alternative to a single large domain to capture recirculation is to apply a chemical transport model with a large coarse domain with a smaller-domain, nested fine grid with feedback between the grids (i.e., 2-way nesting). This option allows pollutants from the fine grid to move into the coarser domain and back into the fine grid depending on wind patterns.

Global models are routinely used to supply regional chemical transport models with temporally and spatially variant lateral boundary conditions. Since regional and urban model applications

---

[14] *See* 40 CFR 51.1009 (moderate $PM_{2.5}$), 40 CFR 51.1010 (serious $PM_{2.5}$) and 40 CFR 51.1312 (ozone).
[15] See 40 CFR 51.308(f)(3)(i).

ED_002666_00014243-00023

for attainment demonstration purposes are typically less than or equal to 12 km sized grid cells, it may be worthwhile to apply the regional scale model at a coarser grid resolution (e.g. 36 km) to downscale global model estimates to the regional model. The coarser simulation results provide a better transition from the global simulation to the nested urban or urban/regional area of interest both at the boundaries near the surface and in the free troposphere. See section 2.8 for more information on using global and/or regional models to derive initial and lateral boundary conditions.

Isolated nonattainment areas that have limited impacts from regional transport of ozone and/or $PM_{2.5}$ and its precursors may be able to use a relatively small domain. The modeling domain should be designed so that all major upwind source areas that influence the downwind nonattainment area are included in the modeling domain. It is also important to be able to directly capture the impact of future year emissions increases or decreases from upwind areas. In addition, all monitors that are currently or recently (within any of the years included in the base year period) violating the NAAQS or close to violating[16] the NAAQS in the nonattainment area should be contained in the modeling domain far enough from the edge to minimize complications associated with air mass recirculation and to minimize lateral boundary influence. Similarly, all Class I areas to be evaluated in a regional haze modeling application should be sufficiently distant from the edge of the modeling domain.

2.4.2 Vertical Layer Configuration

There is no correct maximum or minimum number of vertical layers needed in attainment demonstration modeling. However, the specification of the air quality model vertical layer structure should closely match the vertical layer structure of the meteorological model used to generate inputs for the air quality model. It is best to have an air quality model's vertical layers align with the layers in the meteorological model matching one-to-one. However, resource constraints sometimes preclude direct vertical layer mapping. When vertical layer collapsing is necessary, use the highest resolution where the conceptual model suggests it is most needed.

The top of the modeling domain should typically be set above the tropopause at the 50 or 100 millibar level. The lowest layer (surface layer) in the air quality model should be no more than ~40 meters thick. The vertical resolution between the surface layer and model top will vary

[16] A monitor that is "close to violating" the NAAQS in a base year may be modeled to violate in the future due to projected emissions increases or modeled "dis-benefits" due to non-linear chemistry. Air agencies should consult with the appropriate EPA Regional office if there are questions regarding the interpretation of "close to violating" and which monitors should be included in the modeling analysis.

ED_002666_00014243-00024

depending on the application. Layer thickness should monotonically increase or be the same size as altitude increases.

In view of the importance of carefully specifying the temporal variation in mixing height, high precision below and near the anticipated maximum afternoon mixing height is ideal. In addition, near-surface vertical resolution is important to capture overnight stable conditions. Layers above the boundary layer are important for characterizing clouds and precipitation and layers near the tropopause are important to create a realistic stratification of strongly variant pollutant concentrations in the stratosphere and free troposphere (Emery et al., 2011).

There are some model applications that may not need to consider the full extent of the troposphere. These applications typically use vertical domains, which extend up to 4 or 5 km. These types of applications may cover episodes of elevated pollutant concentrations associated with strong surface-based inversions that occur in isolated areas. However, in almost all cases, the EPA encourages the use of full-scale one-atmosphere models that account for all atmospheric processes throughout the extent of the troposphere.

2.4.3 Horizontal Grid Cell Size

Recent chemical transport modeling applications for ozone, $PM_{2.5}$, and regional haze have commonly used grid cell sizes ranging from 4 km to 12 km resolution (Morris et al., 2006; Rodriguez et al., 2009; Simon et al., 2012; U.S. EPA, 2011b, 2016a; Vizuete et al., 2010). Urban scale assessments have used chemical transport grid models at 1 to 2 km resolution (Vizuete et al., 2011). Intuitively, one would expect to get more accurate results in urban applications with smaller grid cells (e.g., ~1 or 4 km) provided the spatial details in the emissions and meteorological inputs support making such predictions. However, model performance at 4 km resolution is not always better than 12 km resolution (Simon et al., 2012).

For coarse portions of the regional grids, a grid cell size of 12 km is generally recommended. For urban areas, it may be desirable to use grid cells sized ~1 to 4 km, but not larger than 12 km. The relative importance of using a domain with grid cells as small as ~1 km should be weighed on a case-by-case basis. If the conceptual model indicates that there are atmospheric, physical, or chemical processes related to the air quality issue that require a specific minimum resolution to simulate (e.g., localized emissions, complex terrain, land-water interfaces, etc.), then that resolution should be utilized in the model. For example, it is likely that spatial gradients in concentrations are higher for primary pollutants (e.g., primary particulate matter, or PM) than for secondary pollutants because physical processes dominate concentration rather than chemical processes. Although it is clear that spatial resolution of primary pollutants will impact

ED_002666_00014243-00025

the predicted concentrations, it is not clear how it will impact the relative change in concentrations due to emissions changes. Areas that have large gradients in primary $PM_{2.5}$ may need to use finer resolution (approximately ~1 to 4 km) or may need to supplement the grid modeling with dispersion modeling. This is particularly true if violating monitors are strongly impacted by local sources of primary $PM_{2.5}$ emissions.

The most important factor to consider when establishing grid cell size is model response to emissions controls. Analysis of ambient data, sensitivity modeling, and past modeling results can be used to evaluate the expected response to emissions controls at various horizontal resolutions for ozone, $PM_{2.5}$, and regional haze. If model response is expected to be different (and presumably more accurate) at higher resolution, then higher resolution modeling should be considered. If model response is expected to be similar at both high and low(er) resolution, then high resolution modeling may not be necessary. The use of grid resolution finer than 12 km would generally be more appropriate for areas with a combination of complex meteorology, strong gradients in emissions sources, and/or land-water interfaces in or near the nonattainment area(s).

Sensitivity tests comparing relative response factors in predicted 8-hour daily maximum ozone at sites in the eastern United States indicate relatively small unbiased differences (< 4%, in RRF difference in 95% of the comparisons) using a grid with 12 km versus 4 km sized grid cells (Arunachalam, 2006). The largest difference in the relative response of secondary pollutants at varying resolution is likely to occur in oxidant limited areas (areas that are more sensitive to VOC reductions than NOx reductions). In such areas, horizontal resolution may have a large impact on the spatial distribution and magnitude of modeled NOx "disbenefits" (i.e., ozone increases in oxidant limited areas when NOx emissions are reduced).

2.5 Air Quality Model Selection

Chemical transport models address the physical processes and chemistry that form ozone and PM. Air quality models continue to evolve and each has its own strengths and weakness (Galmarini et al., 2012; Pirovano et al., 2012; Russell, 2008; Simon et al., 2012). The most commonly used chemical transport models for attainment demonstrations are the Community Multiscale Air Quality Model (CMAQ) (Byun and Schere, 2006; Foley et al., 2010) and the Comprehensive Air Quality Model with Extensions (CAMx) (Baker and Scheff, 2007; Vizuete et al., 2011). The mention of CMAQ and CAMx is not intended to be a comprehensive list of available chemical transport models and omittance from this discussion does not imply that a different model cannot be used to support a modeled attainment demonstration or reasonable

ED_002666_00014243-00026

progress assessment. In the same way, inclusion in this discussion does not imply that a model is "preferred" for a particular type of application.

A modeling-based demonstration of the impacts of an emissions control scenario for attainment of the ozone or $PM_{2.5}$ NAAQS, or as part of a regional haze assessment, usually necessitates the application of a chemical transport grid model. However, it is not necessary to use a model that considers atmospheric chemistry in addressing changes in primary $PM_{2.5}$ components. Either a numerical grid or a Lagrangian (such as a Gaussian dispersion) model or other empirical techniques could potentially be used for this purpose. In general, modeling primary $PM_{2.5}$ components with a grid model is acceptable, but dispersion modeling may be necessary in areas with large spatial gradients of primary $PM_{2.5}$. Depending on the nature of the problem, an air agency may choose to use a regional chemical transport grid model to address both primary and secondary components of PM or they may need to use a chemical transport model to address secondary $PM_{2.5}$ and regional primary PM and an inert model applied over a more limited domain to address fine scale primary components of PM. In addition, the EPA recognizes that in some cases, more simplistic modeling techniques (such as dispersion, receptor, rollback, and/or box models) may suffice to demonstrate that an area will attain the $PM_{2.5}$ NAAQS, especially in areas that are dominated by primary $PM_{2.5}$ emissions (e.g., residential wood smoke).

A model should meet several general criteria in order to be considered for use in an attainment demonstration or reasonable progress assessment. These general criteria are consistent with requirements in Appendix W (U.S. EPA, 2017a). Appendix W does not identify a "preferred model" for use in attainment demonstrations of the NAAQS for ozone or $PM_{2.5}$ or reasonable progress assessments for regional haze. Therefore, at this time, the EPA is not recommending a specific model for use in ozone or $PM_{2.5}$ attainment demonstrations or regional haze uniform rate of progress assessments. Instead, models used for these purposes should meet requirements for "alternative models" in Appendix W, section 3.2. Air agencies should use a non-proprietary model, which is a model whose source code is available for free (or for a reasonable cost). Furthermore, the user should be able to revise the source code to perform diagnostic analyses and/or to improve the model's ability to describe observations in a credible manner. Several additional prerequisites should be met for a model to be used to support an attainment demonstration or reasonable progress assessment (see Appendix W, section 3.2.2(e), as well as sections 5 and 6 for more general secondary pollutant modeling requirements):

1) The model or technique has received a scientific peer review.

   An air quality model may be considered to have undergone "scientific peer review" if each of the major components of the modeling system has been

ED_002666_00014243-00027

described and tested, and the results have been documented and reviewed by one or more disinterested third parties, and the model has been revised in response to the review. This review is not the responsibility of the EPA. Air agencies should reference available documentation to gain acceptance of an air quality model for use in a modeled attainment demonstration.

2) The model or technique can be demonstrated to be applicable to the problem on a theoretical basis.
    - The model should be scientifically appropriate and applicable for the intended purpose.

3) Databases, which are necessary to perform the analysis, are available and adequate.

4) Appropriate performance evaluations of the model or technique have shown that the model or technique is not inappropriately biased for regulatory application.
    - Prior to use of a selected model's results in an attainment demonstration or regional haze analysis, the model should be shown to perform adequately for the specific application.
    - If the application is the first for a particular model, then the air agency should demonstrate that the new model is expected to perform sufficiently. For a model to be used in an attainment demonstration, evidence should be presented that it has been found acceptable for estimating hourly and 8-hourly ozone concentrations and/or hourly and 24-hour average $PM_{2.5}$ and $PM_{2.5}$ component concentrations. In addition to ozone and/or $PM_{2.5}$, the model should be acceptable for estimating precursor species, important chemical intermediates, and in the case of $PM_{2.5}$, the main chemical constituents. Preference should be given to models exhibiting satisfactory past performance under a variety of conditions.

5) A protocol on methods and procedures to be followed has been established (such as described in section 2.2.1).

In addition to the criteria in Appendix W, the EPA also recommends the following criteria and considerations:

1) A user's guide (including a benchmark example and outputs) and technical description of the model should be available.

ED_002666_00014243-00028

2) The advanced technical features available in a model could be a consideration when other criteria are satisfied. Models are often differentiated by their available advanced science features and tools. For example, some models include advanced probing tools that allow tracking of downwind impacts from upwind emissions sources. Availability of probing tools and/or science algorithms is a legitimate reason to choose one equally capable model over another.

3) When other criteria are satisfied, resource considerations may be important. This is a legitimate criterion provided the other listed criteria are met.

The EPA has prepared an "alternative model" demonstration for both CMAQ and CAMx for the purposes of supporting ozone or $PM_{2.5}$ attainment demonstrations and regional haze assessments (Fox, 2017). This document is also a useful template for others seeking to provide a similar type of demonstration for another photochemical modeling system.

2.6 Meteorological Inputs

In order to solve for the pollutant concentrations over time and space, air quality models require accurate meteorological inputs to properly simulate the formation, transport, and removal of pollutant material. The required meteorological inputs can vary by air quality model, but all models require parameters such as wind, vertical mixing, temperature, humidity, and solar radiation. While model inputs can be created strictly from ambient measurements, a more credible technical approach is to use off-line, dynamic, meteorological grid models to provide the necessary inputs. When these models are applied retrospectively (i.e., for historical time periods), they are able to blend ambient data with model predictions via four-dimensional data assimilation (FDDA), thereby providing temporally and spatially complete data sets that are grounded by actual observations.

This section provides recommendations for generating the meteorological data sets needed for regional air quality modeling purposes. In many ways, the development of meteorological inputs parallels the steps needed to conduct the air quality modeling. A meteorological model platform (episodes, domain, model selection, model configuration, input data, etc.) must be established and then evaluated. Because of the strong sensitivity of the eventual air quality results to the input meteorology (Appel, 2007; Appel, 2010), it is recommended that air agencies spend extensive effort in developing and evaluating the meteorological inputs.

ED_002666_00014243-00029

2.6.1 Developing Base Year Meteorological Fields

The recommended approach for generating the base year meteorological data needed to conduct the attainment or regional haze demonstration is to apply off-line, dynamic meteorological models with FDDA. These models use the fundamental equations of momentum, thermodynamics, and moisture to determine the evolution of specific meteorological variables from a given initial state. When modeling historic episodes, the use of data assimilation helps to "nudge" solutions so that they do not diverge greatly from the actual observed meteorological fields. A major benefit of using dynamic meteorological models is that they provide a way of consistently characterizing conditions at times and locations where observations do not exist.

The Weather Research and Forecasting (WRF) (Skamarock, 2008) model is a community supported mesoscale numerical prediction model that replaced the previous community supported mesoscale meteorological model v5 (MM5) model. The MM5 model is no longer being updated or supported. Meteorological models other than the WRF can be used as long as they are appropriate for the situation being modeled and are properly evaluated. A number of recent studies (Appel et al., 2010; de Meij et al., 2009; Gilliam et al., 2010; Matichuk et al., 2017; Pleim et al., 2016) have compared the ability of the WRF and MM5 models to reproduce past meteorological conditions and their suitability for use in air quality modeling. Generally speaking, comparative evaluation efforts have shown that the WRF model outputs represent equivalent or better statistical skill over a variety of meteorological conditions. An additional advantage of the WRF model is that there is a large community of WRF users and developers. As a result, the code is both maintained and frequently updated as new science emerges. The EPA recommends the use of a well-supported gridded mesoscale meteorological model for generating meteorological inputs to regional air quality model attainment demonstrations.

In some cases, however, there may be legitimate reasons for using an approach other than well-supported gridded mesoscale meteorological models to provide the requisite meteorological data. For instance, there may be long-standing local experience for a given domain and/or episode with a different meteorological model. In other cases, dynamic meteorological models may not adequately capture key meteorological elements of an airshed's conceptual model (e.g., source-receptor transport vectors to key monitoring locations). In rare cases such as these, it may be appropriate to blend the dynamic model data with wind data from an objective analysis of observed wind fields. The guiding factor in determining which meteorological model or approach to use should be an assessment of which set of inputs will best capture the key meteorological conditions that led to poor air quality. It is recommended that a description of the methods used to generate the meteorological fields be

ED_002666_00014243-00030

included in the modeling protocol. For those cases in which an off-line prognostic meteorological model is not used, it is recommended that a detailed description of the alternate technique that will be used to generate the three-dimensional meteorological fields be shared with the appropriate EPA Regional office(s) prior to conducting the meteorological analysis and air quality modeling analysis.

As with other parts of the air quality modeling system, choices made regarding how to configure the meteorological modeling can affect the quality and suitability of the air quality model predictions. Decisions regarding the configuration of complex dynamic meteorological models can be particularly challenging because of the amount of flexibility available to the user. The goal in establishing the proper configuration for meteorological modeling should be to obtain the best possible meteorological model performance, especially for those parameters identified as most important in the conceptual description. As part of the overall chemical transport modeling system evaluation process, performance issues in the chemical transport model can sometimes be explained by options chosen in the meteorological simulation. Therefore, a feedback between the meteorological model evaluation and the air quality model evaluation can be highly beneficial.

2.6.1.1 Selecting a Model Domain

The selection of the meteorological modeling domain should closely match the air quality domain. The outermost grid should capture all upwind areas that can reasonably be expected to influence local concentrations of ozone and/or $PM_{2.5}$. In terms of selecting an appropriate meteorological modeling domain, one should extend the grid at least 3 to 6 cells beyond the domains of each air quality modeling grid to avoid boundary effects. It is recommended that the vertical and horizontal grid structures, including geographic datum and projected coordinate definitions, be generally consistent within the meteorological and air quality models to minimize interpolation and aggregation issues associated with the post-processing of meteorological model outputs into air quality model inputs.

2.6.1.2 Selecting Physics Options

Meteorological models have a suite of physics options that allow users to configure how a given meteorological effect will be simulated. For example, there may be several options for specifying the planetary boundary layer scheme or how sub-grid cumulus clouds will be handled. In many situations, the "optimal" configuration cannot be determined without performing an initial series of sensitivity tests, which consider various combinations of physics options over specific time periods and regions. While these tests may not ultimately conclude that any one configuration is clearly superior at all times and in all areas, it is recommended

ED_002666_00014243-00031

that these sensitivity tests be completed, as they should lead to a modeling analysis that is best suited for the domain and period being simulated. An optimal starting point of physics options may be selected from another modeling application that has a similar set of meteorological conditions that lead to elevated pollutant levels. Typically, the model configuration that yields predictions that provide the best statistical match with observed data over the most cases (episodes, regions, etc.) is the one that should be chosen, although other more qualitative information can also be considered. Additionally, model configurations should be designed to account for the pollutants and time periods that are of most interest, per the conceptual description.

2.6.1.3 Use of Data Assimilation

As noted above, the use of FDDA helps to keep the model predictions from widely diverging from what was actually observed to occur at a particular point in time/space. Studies have shown that the incorporation of FDDA into atmospheric models improves meteorological simulations (Otte, 2008a and b; Pleim and Gilliam, 2009; Gilliam and Pleim, 2010). In particular, these studies have shown that better representation of the atmospheric fields aloft through assimilation of data from nationwide atmospheric profiler networks can aid in reproducing certain meteorological features (e.g., low-level jets) (Godowitch et al., 2011; Gilliam et al., 2012). However, if used improperly, FDDA can significantly degrade overall model performance and introduce computational artifacts (Tesche and McNally, 2001). Inappropriately strong nudging coefficients can distort the magnitude of the physical terms in the underlying atmospheric thermodynamic equations and result in "patchwork" meteorological fields with strong gradients between near-site grid cells and the remainder of the grid. Additionally, if specific meteorological features are expected to be important for predicting the location and amount of pollution formed, based on an area's conceptual model, then the meteorological modeling should be set up to ensure that FDDA does not prevent the model from forming these features (e.g., lake/sea breeze circulations).

In general, analysis nudging strengths should be no greater than 3.0×10^{-4} for winds and temperatures and 1.0×10^{-5} for water vapor mixing ratio. In the case of observation nudging (i.e., FDDA based on individual observations as opposed to analysis fields), it is recommended that the resultant meteorological fields be examined to ensure that the results over the entire domain are still consistent. Further, based on past experience, we recommend against using FDDA within the boundary layer for thermodynamic variables like temperature and water vapor mixing ratio because of the potential for spurious convection. If the dynamic model is applied without FDDA, it is suggested that the simulation durations be no longer than 24 hours. As with selecting physics options, it is often valuable to conduct sensitivity testing of various FDDA approaches to help determine the optimal nudging configuration for any specific case.

Along with application of the meteorological modeling, users should be careful when translating the meteorological model outputs into air quality modeling inputs. A number of air quality model pre-processors (which are meteorological model post-processors) exist to conduct this off-line coupling, including the Meteorology-Chemistry Interface Processor (MCIP) (Otte and Pleim, 2010) and wrfcamx (Ramboll-Environ, 2017). These meteorological post-processors generate the complete set of meteorological data needed for the air quality simulation by accounting for issues related to:

1) Data format translation;
2) Conversion of parameter units;
3) Extraction of data for appropriate window domains;
4) Reconstruction of the meteorological data on different grid and layer structures; and
5) Calculation of additional variables.

As noted above, most meteorological modeling applications should be set up to minimize the influence of the pre-processing software. For instance, it is recommended that the horizontal grid structures be consistent within the meteorological and air quality models to minimize issues associated with interpolation and aggregation. While it is often not feasible to run the air quality model with the same vertical layer structure as the meteorological model, it is recommended that vertical layer collapsing between the models be minimized (preferably with no layer collapsing), to the extent possible.

While the traditional approach for attainment and regional haze demonstration modeling is to convert "off-line" meteorological model outputs into air quality model inputs via specific pre-processing tools, there have been recent efforts to couple "on-line" meteorological models within air quality model applications, such as WRF-Chem (Grell et al., 2005; Wilczak et al., 2009), NOAA-EPA CMAQ (Davidson et al., 2007), and WRF-CMAQ (Wong et al., 2012). The advantage of an on-line meteorological model is that one can simulate impacts that air quality changes might have on meteorological conditions (and vice versa) through direct and indirect feedback mechanisms. In most cases, the EPA does not expect emissions reductions from local SIP reductions to appreciably affect future meteorological conditions. Therefore, off-line meteorological approaches, where the same meteorological fields are used for multiple emissions scenarios, are expected to be sufficient for the majority of SIP modeling exercises.

2.6.2 Assessing Impacts of Future Year Meteorology

Traditionally, SIP meteorology simulations have not accounted for the impact of global emissions of carbon dioxide and other climate forcers on future meteorological conditions. However, ozone concentrations have a high correlation with daily maximum temperatures over many areas in the U.S. (Camalier et al., 2007), suggestive of a potential climate connection.

ED_002666_00014243-00033

Recent research, taking into account temperature and other meteorological variables, indicates that in certain populated regions of the country, climate change could lead to higher future ozone concentrations (often called a "climate penalty") (Jacob and Winner, 2009; Bloomer et al., 2009; U.S. EPA, 2009; CARB 2010; Fiore, 2012; Fann, 2016). Assuming climate change does lead to higher ozone concentrations, there could potentially be a need for more stringent emissions reductions to counteract the higher ozone potential from warmer conditions. However, there are significant uncertainties regarding the precise location and timing of climate change impacts on air quality. Generally, climate projections are more robust for periods at least several decades in the future because the forcing mechanisms that drive near-term natural variability in climate patterns (e.g., El Nino, North American Oscillation) have substantially larger signals over short time spans than the driving forces related to long-term climate change. In contrast, projections for SIP purposes are generally for time spans of less than 20 years. Given the relatively short time span between base and future year meteorology in most SIP demonstrations, the EPA does not recommend that air agencies explicitly account for long-term climate change in attainment demonstrations. However, air agencies are welcome to consider potential climate impacts in their specific areas, especially where and when there is evidence of significant potential impacts.

2.6.3 Evaluation of Base Year Meteorological Fields

While the air quality models used in attainment and regional haze demonstrations have consistently been subjected to a rigorous performance assessment, in many cases the meteorological inputs to these models have received less rigorous evaluation, even though this component of the modeling is quite complex and has the potential to substantially affect air quality predictions (Tesche, 2002). The EPA recommends that air agencies devote appropriate effort to the process of evaluating the meteorological inputs to the air quality model as we believe good meteorological model performance will yield more confidence in predictions from the air quality model. One of the objectives of this evaluation should be to determine if the meteorological model output fields represent a reasonable approximation of the actual meteorology that occurred during the modeling period. Further, because it will never be possible to exactly simulate the actual meteorological fields at all points in space/time, a second objective of the evaluation should be to identify and quantify the existing biases and errors in the meteorological predictions in order to allow for a downstream assessment of how the air quality modeling results are affected by issues associated with the meteorological data. To address both objectives, it will be necessary to complete both an operational evaluation (i.e., quantitative, statistical, and graphical comparisons) as well as a more phenomenological assessment (i.e., generally qualitative comparisons of observed features versus their depiction in the model data).

ED_002666_00014243-00034

2.6.3.1 Operational Evaluation

The operational evaluation results should focus on the values and distributions of specific meteorological parameters as paired with and compared to observed data. It is recommended that the observation-model matching be paired as closely as possible in space and time. Typical statistical comparisons of the key meteorological parameters will include: comparisons of the means, mean bias, normalized mean bias, mean absolute error, normalized mean error, and root mean square error. For modeling exercises over large domains and entire ozone seasons or years, it is recommended that the operational evaluation be broken into individual segments such as geographic subregions and/or months/seasons to allow for a more comprehensive assessment of the meteorological strengths and weaknesses. It may also be useful to break out model performance aloft, at the surface, during individual episodes (e.g., high ozone/$PM_{2.5}$ days), over the diurnal cycle, and as a function of synoptic regime. Modelers are also encouraged to set aside a portion of the ambient data strictly for evaluation purposes (i.e., data not used in the FDDA). Examples of observed meteorological data sets available for operational evaluations include (some of these data sets are not typically used in FDDA and, therefore, may be serve as an "independent" source of evaluation data):

1. TDL U.S and Canada Surface Hourly Observations (http://rda.ucar.edu/datasets/ds472.0);
2. Cooperative Agency Profilers (https://madis-data.noaa.gov/cap/);
3. NOAA Profiler Network (http://profiler.noaa.gov/npn/);
4. Individual State Climate Office Observation Networks; and
5. PRISM Climate Group Precipitation Analysis (http://www.prism.oregonstate.edu/).

While these data are readily available, it is up to the air agency to determine if the data have undergone appropriate quality assurance/quality control (QA/QC) procedures through the provider.

It may be helpful when calculating domainwide and/or regional summary statistics to compare the results against previously generated meteorological model performance "benchmarks" (Emery et al., 2001). However, because of concerns about potentially misleading comparisons of model performance findings across different analyses with differing model configurations and FDDA strengths, the EPA does not recommend using these benchmarks as a "pass/fail" indicator of the acceptability of a model simulation. The benchmarks should only be used as a means of assessing general confidence in the meteorological model data. Statistical results that are outside the range of the compiled benchmarks may indicate a performance issue that should be given further examination. In some cases, adjustment of input data and/or model settings may be needed in order to improve model performance. The Atmospheric Model

ED_002666_00014243-00035

Evaluation Tool (AMET) was designed for use in the assessment of meteorological and chemical transport models. AMET provides a method of pairing observed and simulated values and performing common analysis techniques to evaluate the performance and suitability of atmospheric models.

2.6.3.2 Phenomenological Evaluation

As discussed in section 2.1, it is recommended that a conceptual description of the area's air quality problem be developed prior to the initiation of any air modeling study. Within the conceptual description, it is recommended that the specific meteorological parameters that influence air quality be identified and qualitatively ranked in importance. When evaluating meteorological models or any other source of meteorological data, the focus of the phenomenological evaluation should be on those specific meteorological phenomena that are thought to strongly affect air pollution formation and transport within the geographic area of interest. This event-oriented evaluation should summarize model performance in terms of statistical metrics such as probability of detection and false alarm rate. As an example of a potential phenomenological analysis, many regional air quality modeling exercises attempt to assess the effects of transport of pollutants from one area to a downwind area with the intent of establishing source-receptor relationships. For these types of analyses, accurate transport wind trajectories are needed to properly establish these source-receptor linkages. A useful event-based meteorological evaluation would be to compare model-derived trajectories versus those based on ambient data to determine what error distance can be associated with the model fields. Other examples of phenomenological evaluations would be the replication by the model of low level nocturnal jets, cold pool conditions, fog events, or land/sea breezes.

2.7 How Are the Emission Inputs Developed?

Air quality modeling for 8-hour ozone, $PM_{2.5}$, and regional haze requires emissions data for a base year and future modeling years. While not explicitly a requirement of the implementation rules, the emissions data for modeling are a necessary input to the modeling that is required by the implementation rules as part of the modeled attainment demonstration and for regional haze analyses.

A separate EPA guidance document describes the planning inventories that are specifically required for SIPs and provides an overview of all emission inventory requirements, resources, and techniques. That guidance is the "Emissions Inventory Guidance for Implementation of Ozone and Particulate Matter National Ambient Air Quality Standards (NAAQS) and Regional Haze Regulations" (U.S. EPA, 2017b), and is used as a key reference throughout this section.

ED_002666_00014243-00036

Many of the issues associated with the planning inventories are also relevant for modeling inventories. We will refer to that guidance here as the "EI SIP Guidance."

Table 1 of the EI SIP Guidance provides a list of emissions estimation resources for emissions inventories. In addition to the resources listed there, the EPA also provides emissions data used in modeling analyses in support of rulemakings. The current available "emissions platforms" are based on 2005, 2007, 2008, 2011, and 2014 base years. These platforms are available publicly on the Emissions Modeling Clearinghouse website. The emissions modeling data and other data included in these platforms can be a starting point for air agencies in developing their own capabilities for preparing emissions for use in air quality models.

2.7.1 What Emissions Inventory Years Should I use for Base and Future Modeling?

For the base year, the EPA does not require a particular modeling year for $PM_{2.5}$ and ozone SIP attainment plans, and regional haze plans. However, for several reasons, it is technically appropriate to use a recent year as the base year. In addition, there are certain regulatory requirements for selecting an appropriate base year for planning inventories. See ozone NAAQS inventory requirements at 40 CFR 51.1315 and $PM_{2.5}$ NAAQS inventory requirements at 40 CFR 51.1008.[17] Currently, the most recent base National Emissions Inventory (NEI) is 2014 (U.S. EPA 2016c), and NEI data are developed every 3 years (the next NEI year will be 2017). In most cases, the most recent NEI year will be the most appropriate year to use for base case modeling. In some cases, other years (either before or after the most recent NEI year) may be more appropriate for modeling in view of unusual meteorology, transport patterns, exceptional events (i.e. large wildfires), lack of air pollution events, or other factors that may vary from year to year. See section 2.3.1 for more information on episode selection. The choice of the base year should be discussed with the appropriate EPA Regional office as part of the SIP planning process for the attainment demonstration.

For attainment planning in particular, the choice of a future year for an inventory depends on the nonattainment classification of the nonattainment area, as described in section 2.3.2. For ozone and $PM_{2.5}$, in most cases the future modeling year will be the expected attainment year. Requirements for selecting appropriate future years for modeling are provided in the ozone and $PM_{2.5}$ NAAQS implementation rules. For regional haze, the future year will depend on the end-date of the planning period being examined.

[17] There are additional regulations related to reasonable further progress (RFP) requirements for both the ozone and $PM_{2.5}$ NAAQS.

ED_002666_00014243-00037

2.7.2 What Base Year Emission Inventory Data Are Needed to Support Air Quality Models?

Air quality models require hourly emissions of modeled species for each grid cell and model layer in the modeling domain. For the base year, the inventories are called the *base year inventory for modeling.* To meet the input requirements of models, the inventory must have pollutants that are the scientific precursors (pollutants that impact the relevant chemistry) to the ozone, $PM_{2.5}$, and/or regional haze being addressed in the modeling demonstration. The time frame of the emissions inventory should be consistent with the period being modeled. The inventory should include states, counties, and tribes that are covered by the spatial extent of the modeling domain. Additionally, the emissions sources included should be comprehensive, including emissions from all source categories. As explained with additional detail below, this includes point sources, non-point stationary sources, on-road and nonroad mobile sources, fires, and biogenic sources.

2.7.3 What Pollutants Should Be Included in Base Year Inventories?

Because many sources emit more than one of the precursor pollutants, and because the precursor pollutants have the potential to be transported across state boundaries, the EPA encourages air agencies to develop comprehensive inventories to support integrated, regional-scale modeling, and control strategy development for ozone, $PM_{2.5}$, and regional haze. However, there are some considerations for which pollutants to include in inventories for the purpose of supporting air quality modeling.

For 8-hour ozone NAAQS modeling, the pollutants to be inventoried for modeling are VOC, NO_X, and CO. For $PM_{2.5}$ NAAQS modeling, the pollutants to be inventoried are primary emissions of $PM_{2.5}$ and PM_{10} (including both the filterable and condensable portions) and emissions of SO_2, NH_3, VOC, NO_X, and CO. While certain $PM_{2.5}$ precursors may not be considered precursors for planning purposes in some nonattainment areas[18], inclusion of all precursors is a prerequisite for appropriate application of air quality models for modeled attainment demonstrations. For regional haze, the pollutants to be inventoried include all of the pollutants and precursor pollutants identified for ozone and $PM_{2.5}$. While CO does not need to be included in the *planning inventories* associated with ozone and/or $PM_{2.5}$ SIP submissions (as described in the EI SIP Guidance), it should be included for modeling purposes because CO plays an important role in atmospheric chemistry.

[18] See the $PM_{2.5}$ SIP requirements rule for detailed information on the definition of "planning" precursors for $PM_{2.5}$ nonattainment areas.

ED_002666_00014243-00038

For modeling purposes, emissions inventories of these pollutants undergo many transformations to prepare the inventories for use in air quality models. These transformations include chemical speciation, which is described below in section 2.7.8.2. As part of the chemical transformations of VOC, in particular, total organic gases (TOG) are calculated to include lower reactivity compounds that are not included in the regulatory definition of VOC, but nonetheless have some role in ozone and/or $PM_{2.5}$ formation.

2.7.4 What Temporal Resolution Should Be Used for Base Year Inventories?

The emissions inventories should cover the time periods associated with the modeled time period. Different approaches to accomplishing this goal arise because of the nature of ozone, $PM_{2.5}$, and regional haze problems.

For ozone, the episodic nature of 8-hour ozone exceedances leads to inventories that may reflect a specific period being modeled. For this reason, the *planning* inventories should be provided as an ozone season day inventory[19] (as described in the EI SIP Guidance). For the nonattainment area, the planning ozone season day inventories are also often used for modeling inventories.

For $PM_{2.5}$, both annual and daily NAAQS result in different inventory temporal possibilities for the modeling episode, depending on which NAAQS a particular nonattainment area must address in the SIP. As a result, the *planning* inventories may be annual, seasonal, or both. Like ozone, the planning inventories may also be used as the starting point for the modeling inventories. For example, an annual planning inventory will need to be temporally allocated to the days and hours of the time period used for $PM_{2.5}$ modeled attainment demonstrations. This is the case for both annual modeling and when the air agency has chosen to use a seasonal demonstration as part of a 24-hour demonstration.

Modeling inventories also include areas outside the nonattainment area, for which ozone season day or average season day (for $PM_{2.5}$) emissions may not be readily available. In these instances, air agencies may elect to start with an annual inventory and use an emissions model to adjust the annual emissions to reflect the modeling period. This may not be advisable in every case; in particular, mobile source emissions are highly variable with meteorology and day of the week, and are often available on a monthly or even daily basis that reflects variability in temperature, traffic patterns, and other factors. In addition, more detailed temporal data from electric generating units (EGUs), biogenic sources, windblown dust, sea salt, lighting, wildfires,

[19] The definition of "ozone season day" emissions specifies weekday emissions and is intended to reflect the periods of emissions associated with monitored ambient ozone exceedances.

ED_002666_00014243-00039

prescribed fires, agricultural burning, or other sources should be used when available to enhance the technical quality of the modeled attainment demonstration.

For regional haze, the use of annual emissions inventories is appropriate given the nature of the regional haze problem. However, the same considerations apply for the inventory's temporal aspects as they do in the cases above; more detailed temporal data should be used as the inventory starting point when it is available for important sources of the pollutants included in the inventory.

2.7.5 What Spatial Coverage Is Needed for Base Year Inventories?

Modeling inventories must cover all areas of the modeling domain, which will include areas outside of a nonattainment area and Class I areas and most often includes areas in neighboring states or tribes. This subsection addresses areas within and outside the nonattainment area separately.

2.7.5.1 Areas Within the Nonattainment Area

As mentioned above, planning inventories (which are a CAA and regulatory requirement for ozone and $PM_{2.5}$ SIPs) include emissions within the nonattainment area. The EPA expects that modeling inventories will be consistent with those planning inventories; however, some exceptions may exist. Where possible, the planning inventories (which are typically annual or seasonal) can be a sum (for annual data) or average (for ozone or $PM_{2.5}$ season day data) of day-specific or hour-specific data used for modeling. In some cases, however, this approach may not be sufficient for modeling purposes. For example, greater spatial and temporal detail are needed for on-road mobile modeling inventories as compared to the base year (planning) inventory for the nonattainment area. For the planning inventory, one goal is to allow for the repeatability of the approach in order to create average, seasonal, or annual inventories for use in rule requirements, such as reasonable further progress or conformity demonstrations. That goal is not necessarily compatible with the modeling need for spatial and temporal detail. In cases where some differences between planning and modeling inventories are unavoidable, air agencies should attempt to promote consistency where feasible. More information on the temporal basis of inventories is provided in section 4.5 of the EI SIP Guidance.

2.7.5.2 Areas Outside of the Nonattainment Area

For air quality modeling, emissions inventories are often needed for areas outside the nonattainment area. Developing modeling inventories for these areas should be relatively straightforward and can be included as part of the SIP planning inventories. For areas outside of

ED_002666_00014243-00040

the air agency's jurisdiction, however, it can be more difficult for air agencies to develop the inventories themselves. In these cases, the EPA encourages participation in regional modeling efforts, which are designed to allow sharing of data and help promote consistent approaches across state boundaries. When those organizations lack the needed inventories, the EPA recommends using the EPA's publicly available emissions, as described in the next section.

2.7.5.3 Areas Outside of the United States

In some cases, the modeling domain may need to extend into areas of Canada and/or Mexico. In this case, emissions modelers will need to include emissions from these countries in their modeling domain. Both Canada and Mexico have emission inventory programs that provide inventories that can be used in modeling, and the EPA includes the latest available inventories from these countries in its national emissions modeling platform. States are encouraged to use the latest available data for these countries. While the EPA is one source for that data, states may work directly with other countries or work with their EPA Regional office to try to get inventories that are more current or otherwise improved over what is available directly from the EPA.

2.7.6 How Can the National Emission Inventory Be Used by Air Agencies?

The EPA recommends that state, local, and tribal agencies start with available inventories suitable for air quality modeling of the selected time period(s). If no such inventories are available from a prior state or tribal effort, from regional modeling groups, or from some other source, air agencies may choose to derive an inventory suitable for use with models from EPA-provided information. Two sources of information are available, as described below. In both cases, the EPA expects that air agencies will only use such data after additional review to ensure the appropriate quality of the data for SIP planning purposes.

First, the National Emission Inventory (NEI) provides criteria air pollutant and hazardous air pollutant emissions as annual emissions totals for most (if not all) sources of emissions relevant for ozone, $PM_{2.5}$, and regional haze implementation purposes. Air agencies may use the most recent public release of the NEI as a starting point for inventory development. The data and documentation are available from the NEI Website.

Additionally, the EPA provides emissions modeling "platforms" that are based on the NEI and can be a starting point for air agencies. The modeling platform data provide additional information not available from the NEI, such as daily EGU and fire emissions, monthly nonroad emissions, biogenic emissions and/or input to biogenic emissions models, and inputs to mobile source models. Furthermore, these modeling platforms usually include future-year emissions

ED_002666_00014243-00041

projections. Both base and future-year data are provided in a format used by emissions models and can, therefore, be more readily adopted by air agencies that need to process emissions for use in air quality models. The data and documentation are available from the latest platform release at the Air Emissions Modeling Website.

The detail and accuracy of the EPA emissions inventories and platforms may or may not be sufficient for use in any particular SIP. Air agencies should review the emissions and associated documentation and make improvements where deficiencies exist. Such improvements will often be necessary to demonstrate appropriate model performance, as described in section 3. Additional degrees of review are needed for areas closer to the nonattainment area. For example, EPA data may be sufficient for estimating emissions in states that are hundreds of kilometers upwind or downwind of the nonattainment area, but may not be sufficient for estimating emissions closer to the nonattainment area because those emissions can have a greater impact on the final results of the attainment demonstration.

2.7.7 How Should Base Year Emissions Be Created?

Emission inventory data from at least five general categories are needed to support air quality modeling: stationary point-source emissions, stationary area-source emissions (also called non-point), emissions for on-road mobile sources, emissions for nonroad mobile sources (including railroad and marine vessels), and biogenic/geogenic emissions. A sixth category of wildland fires also exists and includes day-specific emissions from wildfires and prescribed burns. These fire emissions are included as an "event" source in the NEI. Emissions inventories are expected to include "all" sources of emissions; if an air agency wishes to exclude a particular source (e.g., because it is a very small contributor), then the air agency should document the reasons such omissions are acceptable. In addition, some "traditional" nonroad mobile sources (e.g., airports and rail yards) can also be included in inventories as a point source. For example, in the NEI, airports and rail yards are included as point sources and railroad and marine vessel emissions as part of the nonpoint data category.

The EI SIP Guidance (U.S. EPA, 2017b) addresses numerous details about emissions inventory development. Readers are encouraged to refer to information in that document as a resource for building base year emissions inventories. That guidance contains additional information on the definitions of the data categories (point sources, nonpoint sources, etc.), provides more information about emissions estimation for each data category, includes a process for prioritization of inventory development, and includes quality assurance recommendations. All of this information is highly relevant to inventories developed for air quality modeling. In the remainder of this section, we provide a basic description of the data categories and their base

ED_002666_00014243-00042

year emissions and additional information needed for their use in modeling applications where appropriate.

2.7.7.1 Point Sources

Point sources are emissions sources that are attributed to specific facilities and emissions release points within those facilities. The Air Emissions Reporting Rule (AERR) provides information on how to define which sources to include as specific point sources in the inventory. This definition is the same as the New Source Review program major source definition, with a lower potential-to-emit threshold for nonattainment areas. More information on the relationship between SIP inventories and the AERR and how that affects SIP inventories is available in the EI SIP Guidance. For modeling inventories, data developers may include sources with emissions smaller than the required reporting threshold, and should be careful to prevent double counting with nonpoint source categories that may overlap with some of these smaller sources. This is especially a concern when compiling point and nonpoint inventories from different sources, because a nonpoint inventory is compiled using certain assumptions about what an associated point source inventory includes. This issue is described as "nonpoint reconciliation" and is discussed further in the EI SIP Guidance.

Point source inventories for modeling should be compiled at a minimum by country, state/tribe, county, facility, unit, process or source category code (SCC) and by "release point" or stack (see references below to point-source inventory development). The point source data should include information on the location of the sources (e.g., latitude/longitude coordinates); their stack parameters (e.g., stack diameter and height, exit gas temperature and velocity); and operating schedules (e.g., monthly, day-of-week, and diurnal). Information on existing control device types, measures and associated emissions reductions are also useful to identify sources that might be further controlled and to prepare future year inventories.

For air quality modeling purposes, the precision of the latitude/longitude values can be very important, as it will determine the modeled grid cell of the source. It is recommended to use coordinates reported to one-ten-thousandth of a degree (four digits after the decimal point) to give a precision of approximately 30 feet. In addition, metadata about the latitude/longitude should ideally be collected, such as the horizontal reference datum code used (e.g., WGS84). The EPA provides a data standard for latitude/longitude which can be considered when collecting, storing, and reporting latitude and longitude data values (see here). When working with other states to share point source data, it may be useful to use the same horizontal reference datum across states, such as the World Geodetic System of 1984 (WGS84). Other datum codes listed on the EPA Data Element Registry Services include the North American Datum of 1927 (NAD27) and the North American Datum of 1983 (NAD83).

ED_002666_00014243-00043

For modeling needs, many point sources will also be post-processed for the estimation of plume rise of the emissions vertically through layers of the atmosphere as modeled by the air quality model. This plume rise depends on the stack parameters (listed above) and on meteorological conditions. Plume rise can be calculated by emissions models or as part of air quality models (either approach is acceptable and the results can be equivalent). In this latter case, inventory developers are responsible only for identifying for which sources' plume rise will be computed using tools available in emissions models.

Also, for modeling needs, hourly emissions data are usually a starting point for inventories of EGUs. In most cases, such data are readily available from Continuous Emissions Monitoring Systems (CEMS) installed as part of EPA trading programs[20] and include hourly emissions of NO_X, SO_2, CO_2, and heat input. The heat input data can be used to estimate hourly emissions of other pollutants (such of primary $PM_{2.5}$ emissions) using facility-specific emission factors or other emission factors. The EPA's modeling platforms also include these hourly data.

2.7.7.2 Nonpoint Sources

Nonpoint sources are also called "area sources." These sources collectively represent sources of emissions that have not been inventoried as specific point sources. The individual sources within the nonpoint emissions county totals are typically too small, numerous, or difficult to inventory using the methods for point sources. The nonpoint emissions data should be compiled by country, state/tribe, county, and SCC, and sometimes by emission type (e.g., hoteling and cruising for commercial marine emissions in the NEI) and Geographic Information Systems (GIS) shape (for locomotive and commercial marine emissions in the NEI). The specification of nonpoint emissions at a more detailed resolution than county is typically not necessary for modeling inventories, but can provide a more detailed spatial representation for areas where such an approach is needed. Spatial allocation of emissions with spatial surrogates (see section 2.7.8.3) can be equally effective.

2.7.7.3 On-Road Mobile

Emissions from on-road vehicles are the result of several emission processes, including the combustion of fuel while vehicles are starting, idling, and moving, evaporation of fuel from the fuel system and during refueling, and from brake wear and tire wear.

[20] CEM data are provided by the EPA's Clean Air Markets Division in a format suitable for use in emissions processing at the Air Markets Program Data Website. From this page, click the "Prepackaged data" tab and scroll to the bottom to download the data in a form that can be input into SMOKE.

ED_002666_00014243-00044

Mobile source inventory modelers should use the latest version of the Motor Vehicle Emission Simulator (MOVES) for on-road mobile source emissions. They should also subscribe to the MOVES email list to be notified of any bug fixes or updates to the MOVES model. For on-road emissions in California, the most recent EPA-approved version of the EMission FACtors model (EMFAC) should be used. In addition, the Sparse Matrix Operator Kernel Emissions (SMOKE) emissions modeling system includes tools that support the use of MOVES emission factors with SMOKE, and thereby allows modelers to account for the sensitivity of on-road emissions to temperature. SMOKE uses county-specific inputs and gridded hourly temperature and other meteorological information, which is typically based on the same meteorological data used for air quality modeling.

Emissions modelers have a choice in MOVES between "inventory" mode and "emission factors" mode (which creates emission factors that can be further processed using SMOKE-MOVES). Both approaches are valid for supporting air quality modeling for SIPs. More information on the different ways of running MOVES is provided in the MOVES documentation and guidance.

Modelers should consider using a refined approach (such as the SMOKE-MOVES system, day-specific MOVES runs, or some other approach) for use in emissions modeling needed for modeled attainment demonstrations. The SMOKE-MOVES approach first generates emissions rates from MOVES by process (running, start, vapor venting, refueling, etc.), vehicle type, road type, temperature, speed, and hour of day, according to specified fuel parameters and other inputs. Then, subsequent steps apply the appropriate MOVES emissions rates along with vehicle miles traveled (VMT) and vehicle population data for the counties and grid cells in the modeling domain for each hour of the modeling episode.

The MOVES model allows modelers to override default database settings to obtain a locality-specific on-road inventory, and EPA strongly encourages data developers to develop and use MOVES input databases that are specific to the states and counties within the modeling domain. The MOVES Technical Guidance[21] discusses when local data should be used in place of model default data, as well as potential sources of local data.

For some input parameters, there is overlap with the inputs needed for other models, such as the nonroad mobile models. For example, meteorology and fuels information is needed to drive both the mobile and non-road emissions models. Efforts should be made to use the same source of data across multiple categories when the inputs are shared by multiple models. Not doing so calls into question the validity of one or both approaches. Air agencies should explain

[21] The most current version of this document is available on the MOVES website.

ED_002666_00014243-00045

the use of shared inputs, or reasons for not using them, as part of the documentation prepared with their modeled attainment demonstration.

On-road emissions and VMT should be compiled at least at the country, state/tribe, county, and SCC level. Vehicle population data by county and vehicle type should also be compiled as well as idling hours, if available. For some inputs such as VMT, modelers may optionally compile and use data for individual road segments (called "links"). Link approaches require starting and ending coordinates for each road link. Link-based modeling is typically performed for fine scale (i.e., 4-km or finer) modeling studies. Travel Demand Models (TDMs) can be a source of data needed to develop link-based approaches. In regional scale modeling, county-level on-road emissions can be allocated onto road links using spatial surrogates that are generated based on link data and mapped into the modeling grid cells using appropriate weights.

2.7.7.4 Nonroad Mobile Equipment

Nonroad mobile equipment emissions result from the use of fuel in a diverse collection of vehicles and equipment, and the EPA's NONROAD emissions model estimates these emissions. The NONROAD model was originally a separate model, but has now been incorporated into the MOVES model (starting with MOVES2014a). Nonroad emissions should be compiled as country, state/tribe, county and SCC totals.

For the MOVES2014a version of MOVES and later versions, the MOVES model should be used to calculate nonroad source emissions. MOVES2014a (and newer versions) includes features not previously available in MOVES that simplify processing of emissions output and includes updated fuel input files that result in small changes in emission results. In addition, the prior tools available for nonroad emissions estimations (NONROAD2008 and NMIM2008) may no longer work with current operating systems, and the EPA cannot continue to provide technical support for these models. Therefore, the EPA recommends that for modeling nonroad emissions, MOVES2014a or newer versions be used for all new SIP development, although state and local agencies that have already completed significant work with MOVES2014, NONROAD2008, or NMIM2008 can continue to do so to allow for timely submission of the SIP. As a general matter, air agencies should review and possibly update and customize the MOVES model inputs to give better emissions estimates for their state.

Unlike models for onroad emissions, the EPA does not specifically "approve" nonroad models. However, use of alternative models can be justified as part of a SIP submission. For example, California has previously developed and used a state-specific approach for estimating its nonroad source emissions for SIP purposes, called OFFROAD2007. Although this model has now been replaced by category-specific methods for many categories, it is still the default approach for categories not listed with newer methods, as described on the OFFROAD website. Any use

ED_002666_00014243-00046

of such an alternative approach is subject to review by the EPA as part of the review of the SIP, and thus states are encouraged to coordinate with Regional offices on any use of alternative models. The MOVES nonroad capabilities include more than 80 basic and 260 specific types of nonroad equipment and further stratifies equipment types by horsepower rating. Fuel types include gasoline, diesel, compressed natural gas (CNG), and liquified petroleum gas (LPG). The nonroad capabilities of the MOVES model estimates emissions for all criteria pollutant and precursors from both exhaust and non-exhaust processes (diurnal, refueling spillage, vapor displacement, hot soak, running loss, take permeation, hose permeation, and crankcase emissions). MOVES supports many SIP-related inventory development needs, including support for partial counties and seasonal emissions.

Air agencies are encouraged to replace default model inputs with more representative data. Common input adjustments include equipment population, geographic allocations, and local growth rates. If agencies make changes to default model values, the agency should submit the input files to the EPA as well as a description of why the defaults were changed. As mentioned in the previous discussion of on-road mobile emissions, for some input parameters, there is overlap with the inputs needed for other models, such as the on-road mobile models (e.g., meteorology and fuels). Efforts should be made to use the same source of data across multiple categories when the inputs are shared. Not doing so calls into question the validity of one or both approaches. Air agencies should explain the use of shared inputs, or reasons for not using them, as part of the documentation provided with their inventories.

2.7.7.5 Nonroad Mobile: Commercial Marine Vessels, Locomotives, and Aircraft

Emissions from commercial marine vessels, locomotives, and aircraft are estimated in other ways as described in detail in the EI SIP Guidance. These sources can be particularly important for urban areas and in some Class I areas (especially in coastal areas). The EI SIP Guidance contains information about methods available to estimate emissions from these sources.

Commercial marine vessel emissions can be a particular challenge for nonattainment areas that include ports or are adjacent to major shipping lanes. While the level of emissions from these vessels can be outside of the control of state air agencies (coming under national rather than state jurisdiction), it is still important to accurately represent these emissions when they are significant contributors in the vicinity of a nonattainment area or Class I area. Vessels are divided into the largest ocean-going “category 3” (C3) vessels and the smaller “category 1” (C1) and “category 2” (C2) vessels such as tug boats, ferries, support vessels, fishing vessels, and others. Commercial marine emissions include both in-port emissions and underway emissions, with most (~90% of nationwide emissions) emissions associated with underway activity. For the planning inventories covered by the EI SIP Guidance, air agencies need to include only those

ED_002666_00014243-00047

emissions within state waters. For the modeling inventories, however, air agencies may also need to include emissions in Federal waters when the modeling domains extend more than a few miles offshore. This is especially important for coastal areas where offshore commercial marine emissions may be an important contributor to ozone and/or PM in the nonattainment area and/or visibility impairment in Class I areas.

The current EPA methods for estimating commercial marine emissions use a bottom-up approaches for C1, C2, and C3 vessels. The method uses activity data from Entrance and Clearance Waterbourne Commerce (both from the Army Corps of Engineers) and from a 2007 EPA census of C1 and C2 vessel activities. The most recent documentation for EPA methods is provided in the 2014 NEI v2 Technical Support Document, section 4.19.

Uncertainties in estimating commercial marine emissions with EPA's current methods remain. Thus, the EPA strongly encourages air agencies preparing modeling inventories to seek out, develop, and use local-specific information about their ports and waterway vessels to better compute emissions whenever possible, and particularly when the commercial marine sector is a major factor in emissions conditions that lead to nonattainment and/or visibility impairment in Class I areas.

2.7.7.5 Biogenic Emissions

Biogenic sources are a subset of natural emissions sources that may be an important component of the emissions inventory. Vegetation (i.e., forests and crops) is the predominant biogenic source of VOC and is typically the only source that is included in a biogenic VOC emissions inventory. Additionally, microbial activity in the soil contributes to biogenic NO emissions.

Biogenic emissions from vegetation and soils are computed using a model that utilizes spatial information on vegetation and land use and environmental conditions of temperature and solar radiation. The model inputs are typically horizontally allocated (gridded) data, and the outputs are gridded biogenic emissions which can then be speciated and utilized as input to chemical transport grid models. Several models exist, and are described more fully in section 2.7.9.2 and the EI SIP Guidance.

Modeled biogenic vegetation emissions are significant for understanding ozone formation, particularly in the Eastern U.S., and the dependence on meteorology makes using case-specific meteorology very important. In addition, the impact and model characterization of biogenic monoterpene and sequiterpene emissions on secondary aerosol formation is important in many areas of the U.S., most notably the Southeast. Therefore, biogenic emissions should be

ED_002666_00014243-00048

generated using hourly meteorology reflective of the time period being modeled to best capture important relationships between temperature, solar radiation, and biogenic emissions.

2.7.7.6 Geogenic and Other Natural Sources

Geogenic emissions are primarily the result of oil or natural gas seeps, soil wind erosion, and sea salt spray. In addition, lightning may also be an important contributor to natural NO emissions in an inventory area. Volcanoes and fumaroles (i.e., vapor or gas vents in a volcanic region) can be additional sources of geogenic emissions. SIP developers should consider these sources as part of their conceptual description. If geogenic sources are identified as potentially contributing to air quality, such emissions should be included in the modeling.

Oil or Natural Gas Seeps

The U.S. Geological Survey Bureau of Ocean Energy Management (BOEM) has documented emissions from oil and gas seeps in the Gulf of Mexico. The University of California Santa Barbara also provides inventory information about California- area seeps. This is an area of ongoing research.

Windblown Dust

Wind erosion may contribute substantially to PM emissions in an area. Emissions from wind erosion is an area of active research and development, and multiple approaches for estimating wind-blown dust emissions are available. For example, the EPA's CMAQ model includes a module that dynamically estimates hourly natural emissions of fine and coarse dust particles due to wind action over arid and agricultural land. The windblown dust approach is based on the work of J.S. Fu et al. (2016). Other examples of windblown dust approaches include Draxler, Ginoux and Stein (2010), Mansell et al. (2006), and Sundram et al. (2004).

Air agencies using any "in-line" approach for windblown dust emissions should evaluate the results of the approach to determine if it provides a reasonable representation of dust emissions in their modeling domains. Air quality models can have an option to output in-line emissions for evaluation, so that modelers can assess and quality assure the emission estimates directly.

Sea Salt and Other Ocean Emissions

The interaction of sea-salt particles in coastal environments with air chemistry can impact concentrations of $PM_{2.5}$ and gas-phase species in the atmosphere. Various emissions estimation methods are available to account for sea salt emissions. For example, the CMAQ model has been updated to account for these emissions, so that emission inventory developers are able to

ED_002666_00014243-00049

rely primarily on the air quality model to internally calculate sea-salt emissions. Additional information for CMAQ is available in Kelly et al., (2010) and Gantt et al., (2015). For CAMx, additional information on sea salt simulations is available in Athanasopoulou et al. (2008). When an in-line approach is used, emissions should be output from the model so that modelers can assess and quality assure the emission estimates directly. When an in-line approach is not used, modelers should consider the possible need for including sea salt emissions data as an input to the air quality model.

In addition to particulate impacts, ozone in coastal areas can be affected by ocean emissions of halogen compounds including chlorine, bromine, and iodine. Many approaches could be taken to represent those emissions. For example, Yarwood et al. (2014) describes an updated chemical mechanism to include halogen chemistry and a CAMx preprocessor that was used to estimate emissions of halogen precursors in coastal areas, and Sarwar (2015) describes a similar halogen chemistry implementation that has been used in CMAQ. Numerous other references are available in the peer reviewed literature that describe halogen chemistry impacts on ozone formation and modeling of that process.

Lightning

Another natural source of emissions is lightning, which produces NO that forms NO_2 in the presence of ozone or in a photochemically reactive atmosphere. The NO and NO_2 are formed in the mid-upper troposphere and can impact the vertical distribution of reactive nitrogen as well as ozone. Because lightning is not a direct source of NO_2, accounting for this source category is more important for air quality modeling purposes than for SIP planning inventory purposes.

Various methods for estimating lightning emissions could be used. For example, the EPA's CMAQ model includes a parameterized approach to estimate NO_X emissions and the vertical distribution that is generated from lightning. The peer reviewed literature contains both a description of the approach (Allen, 2012) and a model evaluation (Appel, 2010). In addition, Koo et al. (2010) describe estimating annual total lightning NO emissions for North America and then spatially and temporally allocating those emissions to model grid cells using convective precipitation activity as a surrogate.

Volcanoes and Fumaroles

Basic information on volcanic gas is available from the USGS. The USGS also has a monitoring program focused on predicting eruptions. Although this program does not include information on pollutant concentrations, the monitoring program website provides potentially useful maps of volcanos using the "Find a U.S. Volcano" pull-down menu near the top of the site. Finally, the publications related to volcanoes search allows for searching publications for pollutant

ED_002666_00014243-00050

keywords such as CO_2, SO_2, gas, hydrogen sulfide, and volcanic gas. Most of these emissions-related articles focus on information in California, Washington, Hawaii, and the Yellowstone area.

2.7.7.7 Wildland and Cropland Fires

Wildland fires are generally defined as any non-structure fire that occurs in wildland (an area in which human activity and development are essentially non-existent, except for roads, railroads, power lines, and similar transportation facilities and structures, if any, are widely scattered). Cropland fires are any non-structure fire that occurs on croplands. The Exceptional Events Rule[22] (which applies to NAAQS implementation) and the Regional Haze Rule define two types of fire that can occur on wildland:

- *Prescribed fire* means any fire intentionally ignited by management actions in accordance with applicable laws, policies, and regulations to meet specific land or resource management objectives.
- *Wildfire* means any fire started by an unplanned ignition caused by lightning; volcanoes; other acts of nature; unauthorized activity; or accidental, human-caused actions, or a prescribed fire that has developed into a wildfire. A wildfire that predominantly occurs on wildland is a natural event.

Inventories used for ozone, $PM_{2.5}$, and regional haze modeling purposes should include wildland (wild and prescribed) and cropland fire emissions. Wildland and cropland fires can emit large amounts of primary $PM_{2.5}$ as well as precursors that can react in the atmosphere to form ozone and secondary $PM_{2.5}$ downwind potentially impacting both urban and Class I areas. Thus, wildfire and prescribed fire emissions should be included in the nonattainment area inventories when and where these emissions occur within the modeling domain during time periods being modeled. Cropland fires can also be inventoried when and where these emissions occurred or as a nonpoint source.

Detailed information about how the EPA develops wildland and cropland (Pouliot et al, 2016) fire inventories can be found in the latest version of an NEI Technical Support Document, such as that provided with the 2014 NEI. The EPA's approach for wildland fires relies on a combination of satellite detection of fires merged with on-the-ground observational data where available. The EPA encourages the use of ground-based observations and local fuel information whenever possible, as these factors can have a large impact on the emissions. One option for obtaining fire date and location information includes the latest version of the

[22] *See* 81 FR 68216, October 3, 2016.

Satellite Mapping Automated Reanalysis Tool for Fire Incident Reconciliation (SMARTFIRE) system (Du et al., 2013; Sullivan, et al., 2008). Detailed information on the SMARTFIRE system is available in Raffuse et al., 2007. Additional references and information are available as part of the 2014 NEI Technical Support Document. Other sources of fire location, size, duration, and emissions may be available and could be used instead of the SMARTFIRE approach. Examples of other approaches include the WRAP FETS (JFSP, 2013) and the Fire Inventory from NCAR (FINN).

Important aspects of wildland fire that impact resulting emissions estimates include the location of the fire, timing of the fire (start and end date/time), fire type (wild, prescribed, or cropland), fuel type, fuel loading, area burned, and fuel moisture. Daily emissions are allocated to hour of the day using profiles based on fire type. VOC and primary $PM_{2.5}$ emissions are speciated to specific compounds based on profiles generated from laboratory and field testing. The activity information noted is not just critically important for estimating emissions, it is also important toward estimating the heat flux of the fire and subsequent plume rise in the photochemical model. The estimated plume rise can have a large impact on how far downwind smoke can be transported (to a nonattainment area and/or Class I area). Combustion phase can also be an important consideration when modeling wildland fires in terms of emissions, plume rise, and vertical allocation of emissions. Fires that are largely smoldering (such as peat fires) can have different emission factors and have a different profile of VOC and primarily emitted $PM_{2.5}$ emissions.

Since fire emissions can be a large source of emissions, the treatment of fires should be discussed with the relevant EPA Regional office(s) as part of the modeling protocol development process.

2.7.8 What Other Data Are Needed to Support Emissions Modeling?

For modeling purposes, emission inventories need to be converted through emissions modeling from their original resolution (e.g., annual emissions by point source or count) to air quality model input files. Air quality models generally require emissions inputs to be specified by model grid cell, hour, and model chemical species. Modeling inventories also identify which point sources should have their plume rise computed. This section describes the ancillary data that modelers should collect and prepare to allow emissions models to convert the emission inventory data into air quality model inputs.

2.7.8.1 Temporal Allocation

Ancillary data for temporal allocation are necessary for stationary point, stationary area, and all mobile sources. To facilitate temporal allocation of the emissions, factors called "temporal

ED_002666_00014243-00052

profiles" should be created to convert annual emissions to specific months (i.e., monthly profiles), average-day emissions to a specific day of the week (i.e., weekly profiles), and daily emissions to hours of the day (i.e., hourly profiles). Additionally, a cross-reference file is needed to assign the temporal profiles to the inventory records by SCC, facility, geographic area such as state or county, or some other inventory characteristics. Where available, the operating information that may be available from the point-source inventory should be used to create inventory-specific temporal factors. The EPA provides a starting point for the temporal profiles and cross-reference files as part of the latest modeling platforms (previously described), available at the Air Emissions Modeling website.

The EPA has developed a SMOKE utility program (Gentpro) that estimates temporal profiles for residential wood combustion (RWC), agricultural NH_3 from animals, and other generic (user-defined) area sources by relating meteorology to air emission fluxes. Gentpro reads in hourly gridded meteorological data (temperature and wind) from the Meteorology-Chemistry Interface Processor (MCIP)[23] and a gridded spatial surrogate to produce temporal profiles and cross-reference data. Annual MCIP input data are required to calculate temporal profiles with Gentpro and the spatial surrogates are needed to identify the grid cells for computing county-averaged meteorology variables. The Gentpro program is described in the latest SMOKE User's Manual.

For RWC, Gentpro allocates emissions to specific days from annual totals based on a regression equation that was developed by relating daily minimum temperatures and chemical tracers of RWC. This approach results in more RWC emissions assigned to periods of colder temperatures using a user-defined temperature threshold that restricts RWC emissions to days with morning lows below a specific value. Air agencies that use EPA's emissions data for this category should note that this approach has been used to temporally allocate the RWC emissions.

For agricultural NH_3 from animals (i.e., from concentrated animal feeding operations, or CAFOs), Gentpro allocates annual, monthly or daily emissions to hourly emissions based on empirical equations derived from investigations of emissions from animal waste decomposition as a function of temperature and wind speed. The resolution of emissions input to Gentpro for this sector depends on the available data. Air agencies that use EPA's emissions data for this category should note that this approach has been used to temporally allocate agricultural NH_3 emissions from CAFOs.

[23] MCIP is a meteorology postprocessing program. It is typically used to convert WRF meteorological model output data to CMAQ input format. The latest version of MCIP can be found here.

ED_002666_00014243-00053

For point sources, hourly CEMS data are recommended for use in base year emissions modeling rather than temporal profiles, where such CEMS data are available, and with the following caveats. CEMS data are collected as part of emissions compliance and/or trading programs. Because these programs require complete hourly data, alternative emissions values are assigned to hours when CEMS devices are not functioning. These alternative values can be extremely high values (compared to actual emissions rates), in order to motivate CEMS operators to provide a complete data record to the extent possible. As a result, CEMS data may have some anomalies and should be reviewed prior to use in a modeling study. The CEMS data fields include an emissions method code, which identifies the hours with alternative values. Furthermore, CEMS data may be available for only part of a year; therefore, for annual modeling, users should be careful to ensure CEMS data are used for the appropriate periods and total annual emissions are temporally allocated to fill in the other hours of the year.

For wildland fire emissions, the EPA recommends using daily, location-specific fire data such as the NEI "events" sources. These sources can have a large impact on ambient levels of ozone and $PM_{2.5}$, and on regional haze. Although emissions from these sources are largely uncontrollable, the inclusion of all emissions sources (including biogenic and uncontrollable sources) is needed to accurately predict pollutant concentrations and derive relative response factors for both PM and ozone.

As a general matter, the temporal allocation approach should remain consistent between the base year modeling and the future year modeling. Consistency is important when using the RRF approach to make future year projections of ozone, $PM_{2.5}$ and/or visibility, because the results of that approach can be affected by changes in when emissions are modeled to occur. In general, this means that the temporal allocation will be the same between base and future years, except when an intentional change is made to reflect a real expected change in the temporal distribution. An example of such a change includes intentionally reducing emission peaks from units during incidents of poor air quality. In addition, modeling can be specially done for the current year with an emissions case which uses "typical" emissions temporal allocations for major sources, and this case may be different from the current year case, where actual source schedules are used to evaluate model performance.

For future-year modeling there are two choices for representing EGU emissions derived from CEMs data and wildfire emissions. Modelers can use day-specific emissions for both the base and future years, or a "typical year" temporal allocation (averaging) approach can be used to create representative emissions. If a typical year approach is used, it should include similar daily temporal variability as could be expected for any given year (in most cases, the daily temporal variability also needs to be consistent with the base year meteorology). Air agencies should

ED_002666_00014243-00054

discuss the temporal approach to fires (see also section 2.7.11) and point sources with the appropriate EPA Regional office as part of the modeling protocol development process.

2.7.8.2 Chemical Speciation

The emissions models also need information about the chemical species of the VOC, NOx, and $PM_{2.5}$ emissions for all emissions sources. These data are used to disaggregate the total VOC, NOx, and $PM_{2.5}$ emissions to the chemical species expected by the air quality model and are called speciation "factors" or "profiles."

The EPA provides a starting point for the VOC, NOx, and $PM_{2.5}$ speciation data, which are available with the EPA modeling platforms (previously described). The VOC and $PM_{2.5}$ speciation profiles are based on the SPECIATE database, which is available on the SPECIATE website. In addition to the speciation profiles, a speciation cross-reference file assigns speciation profiles. The default EPA speciation cross-reference assigns profiles to sources using SCCs. These speciation profiles in many cases are highly uncertain and should be assessed and improved where possible for emissions source categories that are critical to model performance.

For large or critical VOC and $PM_{2.5}$ sources in the modeling domain, air agencies should consider determining the individual chemical compounds contributing to the total VOC and $PM_{2.5}$. If collected, this information should then be used to compile speciation profiles for the critical facilities or source categories. These speciation profiles should be assigned to the facilities by updating the speciation cross-reference file to support a facility-specific or facility-SCC-specific assignment.

For VOC, low reactivity chemical components including methane have traditionally been calculated by applying a factor to the VOC emissions to calculate total organic gases (TOG). These TOG emissions are then split during emissions processing into the VOC model species that include methane. In limited situations, air agencies may want to investigate the extent that methane emissions and low volatility compounds play a role in ozone episodes and/or secondary $PM_{2.5}$ formation. Additional efforts going beyond the use of the default VOC-to-TOG profiles from the speciation profiles may develop over time.

In some limited cases, it may be of interest to include methane emissions inputs to air quality models to achieve sufficient model performance for a modeled attainment demonstration. Approaches can include incorporating methane emissions from other sources into emissions inventories used for modeled air quality. For example, the MOVES model estimates methane emissions directly, so there would be no need to use VOC speciation to estimate those emissions. Air agencies considering methane as a contributing factor should consult with their

ED_002666_00014243-00055

EPA Regional office to discuss the best sources for such data as this inventory development area is undergoing many changes at this time. Furthermore, more specific VOC-to-TOG factors and speciation profiles may be useful, particularly for analysis related to areas of very high methane emissions. Additionally, the EPA encourages the use of HAP emissions data when available to augment VOC speciation; therefore, we suggest using the inventory HAPs where criteria air pollutants (CAPs) and HAPs were provided by the same data source. When possible, the EPA's modeling platforms generally process the following HAPs rather than relying on VOC speciation: benzene, acetaldehyde, formaldehyde and methanol. These 4 HAPs are collectively known as "BAFM" and are removed from VOC speciation calculations for the remaining VOC HAPs needed for the air quality models. Ethanol can also be added to this "VOC integration" process where ethanol emissions are available. It is entirely up to the user to determine the "integration" status of specific sources in each inventory. A simple test for possible integration of VOC HAPs is the following. For each inventory unique source (i.e., a stack or nonpoint/mobile FIPS+SCC): 1) confirm that VOC HAPs and VOC were reported by the same data provider, and 2) compute VOC HAPs to ensure they are non-zero and less than inventory VOC. A detailed methodology is provided in section 3.2.1 of the 2011 emissions modeling platform technical support document.

County-specific speciation profile combinations are also available for cases where user-defined mixtures of two or more speciation profiles are needed. The "GSPRO_COMBO" feature in SMOKE can be applied by state and county and time period (e.g., month) and is currently used for on-road and nonroad mobile and gasoline-related stationary sources where the emissions sources use fuels with varying ethanol content (e.g., E0 and E10 profiles). Ethanol content varies spatially by state and county via nonattainment areas or local programs, temporally (e.g., during ozone season months), and by modeling year because future years have more ethanol. Certain source categories will require more careful review of speciation assignments. For example, the oil and gas sector (point, area and nonroad) inventories are generally constructed for specific drilling areas (i.e., drilling basins) which may include multiple counties with unique production gas characteristics. Knowledge of the type of gas (e.g., dry, wet, coal bed methane) or oil produced, flared and vented in these basin inventories is helpful for developing basin specific speciation assignments.

NO_x emission factors and, therefore, NO_x inventories are developed on a NO_2 weight basis. For air quality modeling, emissions processors must split NO_x to NO and NO_2 and in some cases HONO. Although emissions models allow using multiple speciation profiles for NO_x, historically only one split factor has been used of 90% NO and 10% NO_2. For mobile sources, the MOVES model calculates NO and NO_2 internally based on a more nuanced selection of factors and uses an NO_2 factor of 0.008 for onroad sources based on Sarwar (2008).

One additional NO_x speciation data resource is stack test data of NO/NO_2 ratios for use in AERMOD modeling for the 1-hr NO_2 standard. These data are available here.

2.7.8.3 Spatial Allocation

For all source sectors that are compiled at a county resolution, the emissions models also need information about allocating the countywide emissions to individual grid cells that intersect the county boundaries. Such sectors include stationary area, nonroad mobile, and (non-link) on-road mobile sources. The spatial allocation process assigns fractions of county-total emissions to the model grid cells intersecting the county based on a "surrogate" data type (e.g., population, housing data). The appropriate types of surrogate data to use for each SCC in the inventories should be identified for this processing step. Spatial surrogates can be created using GIS to calculate the fraction of countywide emissions to allocate to each grid cell based on the surrogate type. These calculations can also be made using the EPA's Surrogate Tool, which is based on the Multimedia Integrated Modeling System (MIMS) Spatial Allocator. In addition, all SCCs needing spatial surrogates should be assigned a surrogate in a cross-reference file. Point sources do not need spatial surrogates, since the emissions models assign the grid location based on the latitude and longitude of the point sources. The EPA provides spatial surrogates and cross-references as part of its latest modeling platform. The same spatial surrogate data are normally used in both the base and future year modeling. However, emissions developers can choose to alter the spatial surrogate data based on predicted changes in land use patterns, population growth, and demographics, although the impacts and utility of such approaches are not well characterized,[24] so their use needs to be well documented and explained.

2.7.8.4 Other Ancillary Inputs

On-road emissions for fine-scale model grids (e.g., 4-km grid cells or smaller) may be based on a link-based approach as mentioned in section 2.7.7.3. The VMT and speed data needed for a link-based approach can be provided using a Travel Demand Model (TDM). These models require their own sets of inputs, which depend on the specific TDM used. The MOVES Technical Guidance provides general guidance on the development of MOVES inputs. Details on using TDMs for link-based on-road mobile emissions are available from the EIIP document "Use of Locality-Specific Transportation Data for the Development of Source Emission Inventories." An example of the use of TDMs can be found in "Use of Travel Demand Model Data to Improve Inventories in Philadelphia."

[24] At the time this document was written, tools to readily predict future-year land use patterns are not readily available in a form for use in emissions modeling.

ED_002666_00014243-00057

Emissions models have other input files that must be created. Previously, we noted that for point sources, emissions developers will have to identify which sources should be treated as elevated sources by an air quality model. To do this, data developers will select criteria such as stack height or a screening plume height value to allow an emissions model to flag point sources as elevated sources. In the SMOKE modeling system for example, the Elevpoint program can be used for this purpose.

The emissions models have a large number of files and settings which work together in fairly complex ways; therefore, users should be careful to select the appropriate files needed for the emissions model, and to prepare all input files in a way that will support using the emissions model for the specific air quality model application. A possible shortcut can be starting with the emissions modeling scripts included with the EPA's modeling platforms.

2.7.9 How Are Inventory Data Converted into Air Quality Model Input?

Emissions models are used to convert inventory data to inputs for air quality modeling. As described in section 2.7.8, as part of the emissions modeling process, additional ancillary data are needed to augment the raw emissions inventories. The emissions data for each of the six emissions sectors (point, area, on-road mobile, nonroad mobile, biogenics, and wildland fires) are temporally allocated, chemically speciated, and spatially allocated by the emissions model. The resulting hourly, gridded, and speciated emissions from all sectors are then combined before being used by an air quality model. In this section, we will provide information on several emissions models and summarize some key issues with the application of emissions models.

2.7.9.1 Emissions Models

Several emissions models are available for use in developing SIPs. While no single model has been specifically created for all situations, each model is generally capable of performing the necessary emissions processing steps including temporal, chemical, and spatial allocation. Users of such models are responsible for ensuring that the emissions processing steps transform the emission inventories as intended and are not changing the emissions in any unexpected way. Each model has different capabilities, limitations, and nuances. Therefore, when choosing an emissions model, it is worthwhile to discuss the choice with the developers of these systems and/or with EPA to determine which model is best suited for a particular application. Note that there are a number of programs that process emissions for individual source sectors. These are discussed below in the sector-specific subsections. Currently, SMOKE is the primary emissions

ED_002666_00014243-00058

model used to develop emissions data for input into chemical transport grid models. Other emissions models can also be used, such as the Emissions Preprocessing System (EPS).

SMOKE supports processing of criteria, mercury, and toxics inventories for stationary point, stationary area, mobile, and biogenic emissions. This emissions model can create input files for the CMAQ and CAM_X air quality models. Applications of SMOKE have been presented at several of the International Emissions Inventory Workshops (Houyoux, 2000; Strum, 2003, Zubrow, 2012, Baek, 2010, Adelman, 2008, Baek, 2007a, Baek, 2007b, Baek, 2006, Eyth, 2006, Pouliot, 2005). SMOKE is available for UNIX and Linux operating systems, but is not recommended for use on PCs running Windows. It does not require third party software. The SMOKE software and User's Guide are available on the SMOKE website.

Utilities for creating speciation profiles, biogenic land use, or spatial surrogates are not included within SMOKE. However, the Speciation Tool can create speciation profile inputs for SMOKE (Eyth, 2006), and biogenic land use and spatial surrogates can be built using the MIMS Spatial Allocator Tool along with the related Surrogate Tools.

2.7.9.2 Biogenic Emissions Models

There are several available biogenic emissions models that can be used to develop biogenic emissions for modeled attainment demonstrations and regional haze modeling. One such model is the Biogenic Emissions Inventory System, version 3.6 (BEIS3.6) (Bash et al, 2016). This model can be run both as part of a CMAQ model run and alternatively as a module of the SMOKE system. The CMAQ approach is described here. The latest documentation on the SMOKE approach to run BEIS is available here.

Another biogenic model commonly used to support air quality model applications is the Model of Emissions of Gases and Aerosols from Nature (MEGAN). The MEGAN model also uses meteorological data including temperature and solar radiation as input to generate biogenic emissions. MEGAN is distributed with vegetation/land-use data that includes all of North America. Other biogenic emissions models have been used to support air quality modeling, such as BEIGIS, which is a GIS-based biogenic emission model developed by the California Air Resources Board.

The BEIS modeling system uses land-use data from the Biogenic Emissions Land Use Database, version 4.1 (BELD4). BELD4 data provides data on the 230 vegetation classes at 1-km resolution over most of North America. These land use data can be created with the MIMS Spatial Allocator. Based on inputs from the emissions developer, the Spatial Allocator will select the 1-km grid cells from the BELD data that intersect the modeling domain and (if necessary)

aggregate to the grid resolution of the modeling domain. Air agencies can alternatively rely on the EPA's 12-km gridded data for modeling domains that overlap with the EPA national domain, provided with the EPA's latest modeling platform on the Emissions Modeling Clearinghouse.

In situations where air agencies consider biogenic emissions important for ozone, PM, or regional haze, the data used in these biogenic emission models could be updated to provide a better representation of vegetation and vegetation specific emission factors. For future-year modeling, land use data is typically held constant (same as the base case). Emissions developers can choose to change their land cover data based on predicted changes in land use patterns, however, the impact and utility of such approaches is not well characterized, so their use needs to be well documented and explained. Where fundamental changes are made to landuse/vegetation data or emission factors, those changes should be noted in the modeling protocol and/or final documentation package along with references to the source of updated/alternative data.

The meteorological data used as input to any biogenic emission model should be consistent with the data used for the air quality model. Biogenic emissions models are generally set up to use gridded hourly meteorological data, so inventory developers should be able to convert their meteorological inputs from air quality modeling inputs for use in their selected biogenics emissions model. Surface (or canopy level) temperature and solar radiation data can be obtained from modeled meteorological data, such as the WRF Model. Alternatively, solar radiation data could be obtained from satellite products.

Biogenic emission models evolve over time, so agencies using these models should evaluate the suitability of one model over the other as part of a modeling analysis. Once a choice has been made for modeling purposes, emissions should also be summed across the NAA for inclusion in the base year inventory for the NAA and/or the attainment projected inventory for the NAA.

2.7.10 Other Emissions Modeling Issues

In addition to the emissions modeling process and tools described above, there are several other important issues to consider. In the remainder of this section, we briefly address each of these issues.

ED_002666_00014243-00060

2.7.10.1 Elevated Point Sources

Point sources need to be assigned to an appropriate model layer[25] (the vertical modeling dimension). Depending on the air quality model that is being used, different emissions modeling steps can be taken. Models such as CMAQ and CAM_X expect input emissions files separately for surface level emissions versus the elevated point-source emissions.

2.7.10.2 Treatment of 'Atypical' Sources of Emissions

In the NEI, there are several sources of emissions that are reported by only one state or a small number of states. These sources typically have very low emissions, so whether they are included in the modeling inventory typically has little impact on air quality modeling. Examples of these sources include domestic and wild animal waste emissions (primarily NH_3), human perspiration and cigarette smoke, and swimming pools. In some inventories, large catastrophic/accidental releases are reported (e.g., car accidents, tire fires). It is generally considered inappropriate to include these sources in emissions processing for air quality models. However, some source categories may be important or unique to a particular area. Before removing seemingly unimportant source categories, the relative importance (e.g., as a percentage of emissions in the area) of each category should be evaluated.

Air agencies should also review and consider the draft EI SIP Guidance and the Ozone Implementation Rule regarding EGU emissions during High Electric Demand Day (HEDD) periods. Such emissions can be crucial for having a complete understanding of relevant emissions sources within a modeling domain. Where these emissions are important for assessing a nonattainment problem, air agencies should consider the potential impact of HEDD emissions when selecting time periods (e.g., episodes, season, year) to model.

2.7.10.3 Transport and Meteorology-based Reductions in Fugitive Dust Emissions

Fugitive dust emissions are another large source of primary $PM_{2.5}$ emissions. However, past modeling experience has shown modeled $PM_{2.5}$ concentration over-predictions from these emissions (Pouliot et al, 2010). Therefore, the EPA recommends reducing area-source fugitive dust emissions prior to air quality modeling to account for "capture" by the terrain (such as deposition on vegetation and buildings), changes in the emissions potential (such as frozen ground), and removal by meteorology (such as precipitation). These two components should be accounted for when adjusting the inventory PM emissions from PM sources such as paved and unpaved roads, construction, mining, agricultural activities, and industrial dust and other low-

[25] Point sources generally comprise most of the elevated emissions (above layer 1), although other sources, such as fires, ships, and aircraft may also have emissions assigned to elevated layers.

ED_002666_00014243-00061

level fugitive dust source categories. The "captured" portion of the reduction can be based on gridded land use data such as vegetation and building characteristics. The EPA has developed a methodology within SMOKE to adjust the fugitive dust emissions. The "meteorological" component of the reduction requires intermediate SMOKE processing scripts to merge hourly, gridded precipitation data and snow cover values to zero-out all fugitive dust emissions where appropriate. To avoid double-counting the precipitation-based reduction, the user should make sure that the inventory fugitive dust inventory does not already include a "MET-adjusted" reduction prior to these manipulations. This combined transport and met-based reduction approach is documented in Pouliot et al., 2010.

2.7.10.4 Quality Assurance

Quality assurance (QA) is a key component of any successful emissions modeling effort. A brief synopsis of appropriate QA approaches for emissions modeling is available in section 2.20 of the SMOKE manual. The purpose of QA for emissions modeling is to ensure that the inventories are correctly processed using the information the modeler intended. (It is assumed here that the inventory itself has already been QA'd and erroneous values and/or outliers removed through inventory QA procedures, as described in the EI SIP Guidance.)

Emissions modeling QA includes such activities as:

- Reviewing log files for errors and warnings and addressing problems;
- Comparing emissions between each of the processing stages (e.g., data import, speciation, temporal allocation) to ensure mass is conserved;
- Checking that the correct speciation, temporal allocation, and spatial allocation factors were applied; and
- Reviewing the modeling-specific parameters such as stack parameters.

It is also useful to compare air quality modeling-ready emissions summaries to previous emissions modeling efforts to verify expected emissions changes by source category(s) and geographic area(s) of interest. In addition, the process of running emissions inventories through emissions models and air quality models often provides insights into the emission inventories themselves. These insights can lead to inventory changes that improve the quality of inventories for additional modeling iterations.

In general, this guidance also encourages the use of graphical analysis and GIS for improved QA of emissions data and processing. A commonly used analysis tool for model-ready emissions data is the EPA-sponsored Visualization Environment for Rich Data Interpretation (VERDI) tool. The VERDI software package provides for efficient, flexible and modular capabilities of

ED_002666_00014243-00062

overlaying meteorology, emissions and air quality modeling data with GIS shapefiles. More information on VERDI is available on the VERDI website.

2.7.11 How Are Emissions Estimated for Future Years?

Emissions estimates for future years are called "emissions projections." The inventory developed for modeled attainment demonstrations is called the *attainment projected inventory for modeling.*[26] These emissions projections include emissions changes (due to increased or decreased activities), facility and/or unit-level shutdowns, and emissions controls (which can be due to, among other things, regulations, settlements or fuel-switching that reduce emissions in specific ways in the future). The goal in making projections is to obtain a reasonable estimate of future-year emissions that accounts for the key variables that will affect future emissions. Each air agency is encouraged to incorporate in its analysis the variables that have historically been shown to drive the economy and emissions within the modeling domain, as well as the changes in growth patterns and regulations that are expected to take place between the time of the base year and future years.

Complete guidance on emissions projections is provided in the EI SIP Guidance, with subsections on EGUs, non-EGU stationary sources, on-road mobile sources, nonroad mobile equipment, other nonroad mobile sources, and a list of other projection resources. The EI SIP Guidance includes a lengthy discussion about considerations that should be made for models or other tools that forecast EGU emissions. As noted in that guidance, "the states are advised to make sure that the tools meet the criteria laid out in this document because EPA will use these criteria to assess any EGU projection information submitted as part of SIPs." In addition, several additional issues are listed here for consideration for modeling inventories.

Emissions models (e.g., SMOKE, EPS) provide the capability to create future-year inventories using base year inventories and projection information. Emissions modelers will need to convert the projection data into specific formats for input to these models. Inventory developers should determine which emissions model will be used to perform the calculations and make sure that the type(s) of information needed by the model is being collected.

In the context of modeling for ozone and $PM_{2.5}$ purposes, for fires, it is recommended that the base year fires be held constant for use in the future year modeling. This approach allows for

[26] The name applies even if it turns out that the modeling shows that the area is not able to attain the standard by the latest applicable attainment date. Note that projected inventories are also used for regional haze modeling, however, the name "attainment projected inventory for modeling" applies specifically to attainment demonstrations.

ED_002666_00014243-00063

the impact of these fires to be minimized using the RRF approach, while still including realistic fire patterns that are relevant to ozone and $PM_{2.5}$ formation necessary for modeled attainment demonstrations. States have flexibility in how they estimate base year fires, but it is recommended that states should hold these constant in the future modeling case. Additional specific discussion of future-year fires as they relate to $PM_{2.5}$ SIPs is available in section IV.D.3.b of the preamble to the $PM_{2.5}$ SIP Requirements Rule. For regional haze purposes, air agencies should consult sections IV.E and IV.G of the 2017 Regional Haze Rule, 82 FR 3078 (Jan. 10, 2017), as well as their Regional EPA office as appropriate, to develop an approach for the treatment of emissions from fires.

Ancillary input files for inputs to emissions models, such as spatial surrogates, temporal allocation factors and especially speciation profiles may also need to be adjusted to reflect conditions expected in the future. The EPA modeling platforms, for example, consider the impact on speciation of fuels used in the future, which are expected to change over time due to regulations.

Once a future-year inventory and other data have been created, it should undergo the same steps as for the base-year modeling, including temporal allocation, speciation, spatial allocation, elevated source selection, special on-road mobile processing, any other custom processing steps (e.g. applying meteorology-based information to fugitive dust, agricultural NH_3 and residential wood combustion), and QA. Except where intentional, every attempt should be made to use consistent approaches between the future year and the base year for all of these modeling steps. Inconsistencies in approaches between the future-year modeling and the base-year modeling can lead to artificial differences in air quality modeling results that can affect conclusions.[27] Therefore, it is critical to avoid such differences whenever possible. If inconsistent base and future year approaches are necessary, they should be well documented and explained.

2.8 Initial and Lateral Boundary Conditions

Air quality models require chemical time and space boundary conditions in order to solve for concentrations within the modeled domain and time period. The time boundary or *initial condition* (IC) is provided at the beginning of the first simulated time step and the space boundary or *lateral boundary condition* (LBC) is provided at the outer edge of the modeling domain. As with other model inputs, the establishment of the IC and LBC data should be guided

[27] This is especially important due to the use of the "relative" change in modeled concentrations as part of the modeled attainment test. An inconsistency in base and future year emissions may lead to a very large modeled percent change in ozone and/or $PM_{2.5}$ concentrations.

ED_002666_00014243-00064

by the conceptual model of the air quality issue being simulated. For instance, more attention will be needed in developing accurate LBC for cases in which ambient data analyses (or previous modeling scenarios) have indicated the potential for long-range transport of pollutant material, compared to cases in which the air quality problem has been identified as largely local in nature. Numerous studies have highlighted the importance of accurately representing LBC in optimizing model performance and identifying the most effective path to attaining the NAAQS (Jacob *et al.*, 1999; Tang *et al.*, 2009; Hogrefe *et al.*, 2011).

Given limitations in available ambient data, it is impossible to exactly specify the complex three-dimensional chemical characteristics of the IC or LBC. As a result, two basic approaches to establishing these fields have been used within the modeling community. The first and preferred approach is to characterize the IC/LBC data as accurately as possible given the existing state of knowledge about ambient conditions at those times and in those locations. The second approach is to attempt to minimize the impacts of the IC and LBC by configuring the domain and modeling period such that the air quality concentrations in the locations and times of interest are largely unaffected by the IC/LBC inputs.

The impacts of the IC are most easily minimized when they are reasonably accurate. In most cases, it will be possible to minimize the impacts of the IC data via the use of a model "spin up" period. A spin up period is defined as some number of modeling days that precede the time period of actual interest. The model outputs for this spin up period are discarded from any analytical post-processing and evaluation, as they are potentially affected by the uncertain initial conditions. The length of the needed model spin-up period will vary as a function of: the domain size, prevailing meteorological conditions, and the chemical lifetimes of the pollutants being simulated. It is a simple modeling exercise to test the sensitivity of the key time periods to the initial conditions. Previous sensitivity modeling has determined that the IC are typically "washed out" of the system in as few as 3-5 days for a local-scale modeling analysis; approximately 10-20 days for a regional-scale analysis; and on the order of 6 months or more for a hemispheric-scale analysis (Berge *et al.*, 2001; U.S. EPA 2010; Anenberg et al., 2011). More recent work shows that 2 weeks is generally sufficient to reduce IC influence to under 1%, with more realistic IC decreasing the spin-up time (Hogrefe et al., 2017). Upper air (elevated layers in the model) takes longer to minimize the influence of IC, and as a result, some locations may require longer spin-up). If there is any question as to the appropriate length of the spin-up period, then sensitivity modeling tests should be conducted. Otherwise, the above guidelines should be sufficient for most ozone and $PM_{2.5}$ attainment demonstrations and regional haze assessments.

ED_002666_00014243-00065

In many cases, it will not be feasible to build a large enough domain to minimize the impacts from emissions outside the domain. For these cases, it will be necessary to downscale the requisite species concentration data from well-performing, coarser-grid, larger domain modeling runs. (In rare cases, it may be possible to assign LBC from detailed field study species data, but most domains will not have complete chemical profiles in space and time along the domain edges.). These larger "parent" grids may be generated by the same regional scale model at coarser resolution; or they can be generated via global or hemispheric models. In the last decade, several global/hemispheric air quality models have been used to generate regional LBC inputs. A sample of models used for LBC include: GEOS-Chem (Bey et al., 2001), the AM3 atmospheric component of the GFDL global coupled model (Donner, et al., 2011), the Model for Ozone And Related Tracers (MOZART, Emmons et al, 2010), the WRF-Chem model (Zhang, et al., 2012a), and the hemispheric CMAQ model (Byun and Schere, 2006; Fu et al., 2012). The EPA does not recommend any one particular global/hemispheric model. Each regional modeling exercise should individually assess whether the chosen global/hemispheric or coarse regional model accurately captures lateral boundary conditions.

For cases in which the LBC will be extracted from larger-scale regional/hemispheric modeling, the EPA recommends two efforts to optimize the LBC characterizations. First, modelers should try to minimize any disconnects in the downscaling process. Wherever possible, one should attempt to utilize the same species, chemical mechanism, layer depths, and meteorological inputs in the regional/hemispheric model as are being used in the finer-scale attainment demonstration modeling (Lam and Fu, 2009). This will minimize the number of uncertain translations needed to convert coarse-scale outputs to finer-scale LBC inputs. In particular, careful consideration of the stratosphere in the two modeling systems will be needed to ensure that regional/hemispheric ozone that is stratospheric in origin is not excessively mixed down to the surface in the finer-scale modeling due to disconnects between the two systems.[28] Second, there should be a comprehensive evaluation of the regional/hemispheric model performance (Bey *et al*., 2001). While there may not be a detailed data set available for model evaluation along the domain lateral boundaries in most cases, consideration of model performance across the domain will provide a first-order indication of the ability of the coarser-scale model to characterize air quality at the boundaries. Additionally, it may also be possible to infer

[28] In establishing the layer structure for both global and regional models, it is advisable to avoid selecting upper model layers that span both the upper troposphere and lower stratosphere. It has been seen that "hybrid" layers can allow for inappropriate diffusive mixing between those two generally decoupled sections of the atmosphere. In cases where stratospheric intrusions are not expected to influence ozone in the area, it may be appropriate to artificially cap stratospheric ozone concentrations to avoid ozone performance issues. This should be a "last resort" fix, only applied after detailed assessments of the vertical diffusion and vertical advection schemes within the model.

ED_002666_00014243-00066

information about the quality of the coarse-model estimates via comparisons with satellite data (Liu *et al.*, 2006, Henderson *et al.*, 2013).

In some situations, it may be possible to construct a modeling platform in which the LBC will not substantially influence the air quality issue under consideration. For instance, areas for which there is no evidence of pollutant transport from the far edges of the model domain to the local area may be able to demonstrate that the determination of LBC data will not impact the eventual modeling demonstration. As with the IC data, this conclusion can be proven via simple sensitivity tests in which the LBC are modified from a base state and the results between the two runs are compared (e.g., Appel *et al*, 2007). For situations in which alterations to the LBC data will not impact key findings, it is acceptable to use time-invariant profiles for LBC. Because there are many publicly available sources of modeled boundary conditions, this approach should only be used only when there is good reason.

Due to the potential impact of LBCs on model response to emissions controls, it is important to accurately characterize the relative influence of LBC versus local emissions. For example, if the impacts of LBC on local air quality are overestimated, it will be more difficult than it otherwise should be to show reductions in future design values from local controls. Conversely, if the modeling system underestimates the role of inflow air quality, local emissions reductions may appear more effective in the model than they will be in reality.

The selection of future year LBCs also has potentially important ramifications on the relative response factors derived from the modeling for the various control scenarios. If LBC are known to impact local air quality, then modelers may need to consider how these contributions will change between the base year and the attainment year (Adelman *et al.*, 2011). Ambient trends analyses for 1990-2010 have indicated that free tropospheric ozone levels above the western U.S. have increased by approximately 0.5 ppb/year over those two decades with increases seen at rural surface monitors in the west (Cooper et al., 2012; Cooper et al., 2010; Lin et al., 2017). More recent analysis focusing on 2000-2016 shows mixed trends in the free troposphere with only one high altitude surface monitor with a significant increase in its median or 95th percentile concentration. The most recent trends may suggest decreasing future ozone contributions from LBC due to international controls (e.g., van der A et al., 2017). It is left to the discretion of the air agency to determine whether varying future LBC are needed as part of an attainment demonstration. Updated future boundary conditions may be needed when emissions changes are expected outside the domain that will meaningfully change the LBC and impact the area(s) of interest within the domain. This decision should consider both the length of the projection period and the potential for significant air quality changes outside the modeling domain. Additionally, careful consideration of upstream emissions will be needed as

ED_002666_00014243-00067

part of any additional future-year regional/hemispheric modeling conducted to develop future LBC. If updated future year LBC are used, those data should be fully described and justified in the modeling protocol.

ED_002666_00014243-00068

3.0 Evaluating Model Performance

3.1 Overview of Model Performance Evaluation

The results of a model performance evaluation (MPE) should be considered prior to using modeling to support an attainment demonstration or regional haze assessment. The objective of the MPE is to demonstrate that the base case model can simulate observed pollution concentrations during historical pollution episodes, and to develop confidence that the model can reliably predict how future pollution levels will change in response to changes in emissions. A particular concern in photochemical models is that compensating errors in the model can cause the model to reproduce observed pollution concentrations in the base case while incorrectly representing the emissions, dispersion and chemistry processes that control pollution formation. Models that achieve good performance through compensating errors may not be reliable for predicting how pollution levels will respond to future emissions reductions. However, if an operational evaluation is conducted over a large enough spatial and temporal scale to incorporate a variety of meteorology and emissions conditions, good model performance is less likely the result of compensating error. Thus, a key goal of the MPE is to demonstrate that the model is getting good results for the right reason (Russell and Dennis, 2000) and to show that the model is able to capture pollution concentrations over the range of conditions that are relevant for the regulatory application being undertaken. In addition, the MPE can provide air agencies with tools to identify and fix causes of poor model performance such as problems in the meteorological, emissions, or boundary condition input files.

Dennis et al. (2010) describe an MPE framework that includes four different approaches that can be used to evaluate air quality models:

1. Operational evaluation techniques include statistical and graphical analyses aimed at determining whether the modeled simulated variables are comparable to measurements.
2. Diagnostic evaluation focuses on process-oriented analyses that determine whether the individual processes and components of the model system are working correctly, both independently and in combination.
3. Dynamic evaluation assesses the ability of the air quality model to predict changes in air quality given changes in source emissions or meteorology, the principal forces that drive the air quality model.
4. Probabilistic evaluation attempts to assess the level of confidence in the model predictions through techniques such as ensemble model simulations.

ED_002666_00014243-00069

The goal of the operational evaluation is to determine how well the model replicates observed concentrations of $PM_{2.5}$ components, ozone and their precursors. Diagnostic and dynamic evaluations are used to determine whether the model accurately represents the processes that determine pollutant concentrations, including emissions, dispersion, chemical reactions, and deposition (i.e., to determine if the model is getting good results for the right reason). For cases in which the model performs poorly in the operational evaluation, the diagnostic and dynamic evaluation can provide insights into the causes of the poor performance and to improve model performance. Ultimately, the goal of these evaluation methods is to assess the degree of confidence in the use of the model as a tool to inform the planning process, and more specifically, to determine how reliably the model predicts the response of ozone and/or $PM_{2.5}$ to changes in emissions. The modeled attainment and regional haze tests use models to predict the response of ozone and/or $PM_{2.5}$ to changes in emissions and then applies the resulting relative response factors (RRFs) to observed (rather than modeled) ambient data. Thus, while historically, most of the effort has focused on the operational evaluation, the relative test makes the diagnostic and dynamic evaluation also important.

At a minimum, a model used for air quality planning purposes should include a complete operational MPE using all available ambient monitoring data for the base case model simulation period. Section 3.2 describes the types of metrics and plots that are typically developed as part of an operational MPE, and section 3.3 describes ambient data that is available through continuously operating monitoring networks. If available, monitoring data from special field studies and other research data may also be used in the operational MPE.

Where practical, the MPE should also include some level of diagnostic evaluation. The use of sensitivity studies, process analysis, indicator ratios and source apportionment approaches for diagnostic evaluation, including dynamic and probabilistic evaluations are described in section 3.4. Given that air agencies might have limited resources and time to perform diagnostic and dynamic evaluation, the use of these methods may be limited in scope in a typical regulatory modeling application. However, more comprehensive diagnostic testing of the model can be a valuable component of the weight of evidence analysis used to support a SIP model attainment demonstration.

From an operational standpoint, the 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 et al., 2017) summarize photochemical model performance for applications published in the peer-reviewed literature between 2006 and 2012. These reviews may serve as a resource for identifying typical model performance for state of the science modeling applications. It is clear that there is no single definitive test for evaluating model performance. All of the tests

ED_002666_00014243-00070

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, the 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. The 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 (i.e., tests that provide insight into the accuracy of the model's relative response to emissions reductions should be given more weight). 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.

3.2 Operational Evaluation

An operational evaluation is used to assess how accurately the model predicts observed concentrations. The underlying rationale is that if we are able to correctly characterize historical concentrations of ozone, PM and their precursors for a variety of meteorological conditions, this gives us some confidence that we can correctly characterize future concentrations under alternative emissions levels. Typically, this type of evaluation is comprised of statistical assessments of modeled versus observed data. Operational evaluations are generally accompanied by graphical and other qualitative descriptions of the model's ability to replicate historical air quality patterns. An operational evaluation provides a benchmark for model performance and can therefore identify model limitations and uncertainties that may require further model development/improvement and/or diagnostic evaluation

The robustness of an operational evaluation is directly related to the amount and quality of the ambient data available for comparison. An operational evaluation for ozone would ideally include co-located measurements of ozone precursors NOx and VOC and vertical profile measurements that can be used to determine the extent of vertical mixing of pollutants and the concentration of ozone and precursors above the boundary layer.

An operational evaluation for $PM_{2.5}$ and regional haze is similar to that for ozone; however, an important difference is that $PM_{2.5}$ consists of many component species and is typically measured with a 24-hour averaging time. The chemical components of $PM_{2.5}$ should be evaluated individually. In fact, in undertaking an operational assessment, it is more important to evaluate the components of $PM_{2.5}$ than to evaluate total $PM_{2.5}$. Apparent "good performance" for total $PM_{2.5}$ does not indicate whether the model is predicting the proper mix of components, which is important for the modeled attainment test or regional haze

ED_002666_00014243-00071

assessment. If performance of the major components is good, then it follows that performance for total $PM_{2.5}$ will also be good. In addition to measurements of VOC and NOx, which are precursors for organic aerosols and nitrate, a $PM_{2.5}$ model evaluation should also include measurements of gaseous precursors SO_2 and (where available) NH_3. Due to the influence of initial conditions, model predictions from the spin-up days should be excluded from the analysis of model performance.

3.2.1 Metrics

Operational evaluations quantify model performance through a variety of statistical metrics. Recommended metrics are described below and shown in Table 3.1. It is important to include multiple statistical measures in any operational evaluation to fully characterize model performance. At a minimum, we recommend evaluating: Mean Observed, Mean Model, Mean Bias, Mean Error and/or Root Mean Square Error, Normalized Mean Bias and/or Fractional Bias, Normalized Mean Error and/or Fractional Error, and the correlation coefficient. The equations for calculating each of these metrics are given in Table 3.1.

Mean Observed: The time-averaged mean observed value (in ppb or $\mu g/m^3$).

Mean Model: The time-averaged mean predicted value (in ppb or $\mu g/m^3$) paired in time and space with the observations.

Mean Bias (MB): This performance statistic averages the model/observation residual paired in time and space. A value of zero would indicate that the model over-predictions and model under predictions exactly cancel each other out. This metric is reported in the unit of measure (ppb or $\mu g/m^3$).

Mean (Gross) Error (ME/MGE): This performance statistic averages the absolute value of the model/observation residual paired in time and space. A value of zero would indicate that the model exactly matches the observed values at all points in space/time. This metric is reported in the unit of measure (ppb or $\mu g/m^3$).

Root Mean Square Error (RMSE): This performance statistic (in units of ppb or ug/m^3) is a measure of the average distance between predicted and observed values. It is calculated as the standard deviation of the difference between modeled and observed values.

Normalized Mean Bias (NMB): This statistic (given in units of percent) is used to normalize MB to the average observed value. NMB values range from -100% to +infinity. Consequently, negative and positive bias values using this metric are not symmetrical around 0.

Normalized Mean Error (NME): This performance statistic (given in units of percent) is used to normalize the mean error relative to the average observation. This statistic averages the

ED_002666_00014243-00072

absolute value of the difference (model - observed) over the sum of observed values. NME values range from 0 to +infinity.

<u>(Mean) Fractional Bias</u> (MFB/FB): Fractional bias is determined by normalizing the MB by the average of observed and modeled concentrations. The range of FB is -200% to +200%. The fractional bias for cases with factors of 2 under- and over-prediction are -67 and + 67 percent, respectively (as opposed to -50 and +100 percent, when using normalized bias). Fractional bias equally weights positive and negative bias estimates (underestimates and overestimates are symmetrical around 0).

<u>(Mean) Fractional Error</u> (MFE/FE): Fractional error is determined by normalizing the ME by the average of observed and modeled concentrations. The range of values for FE is 0 to 200%. It is similar to the fractional bias except the absolute value of the difference is used so that the error is always positive.

<u>Correlation Coefficient (r) and Coefficient of Determination (R^2)</u>: These performance statistic measures the degree to which two variables are linearly related. A correlation coefficient of 1 indicates a perfect linear relationship; whereas a correlation coefficient of 0 means that there is no linear relationship between the variables. A correlation coefficient of less than 0 indicate anti-correlation. The Coefficient of Determination is always positive.

Table 3.1: Definitions of recommended statistical metrics

Abbreviation	Term	Definition
MB	mean bias	$\frac{1}{\ }\Big)\ (\ \ -\ \ )$
ME	mean error	$\frac{1}{\ }\Big)\ \lvert\ \ -\ \ \rvert$
RMSE	root mean squared error	$\Big)\overline{\frac{\Sigma(\ \ -\ \ )^2}{\ }}$
FB	fractional bias	$100\% \times \frac{2}{\ }\Big)\ \frac{(\ \ -\ \ )}{(\ \ +\ \ )}$
FE	fractional error	$100\% \times \frac{2}{\ }\Big)\ \frac{\lvert\ \ -\ \ \rvert}{(\ \ +\ \ )}$
NMB	normalized mean bias	$100\% \times \frac{\Sigma(\ \ -\ \ )}{\Sigma}$

ED_002666_00014243-00073

NME	normalized mean error	$100\% \times \frac{\sum \lvert \quad - \quad \rvert}{\sum}$
r	Correlation coefficient	$\left(\frac{\sum_1 (\quad - \quad) \times (\quad - \quad)}{\sqrt{\sum_1 (\quad - \quad)^2 \sum_1 (\quad - \quad)^2}}\right)$
r^2	coefficient of determination	$\left(\frac{\sum_1 (\quad - \quad) \times (\quad - \quad)}{\sqrt{\sum_1 (\quad - \quad)^2 \sum_1 (\quad - \quad)^2}}\right)^2$

M_i = modeled concentration i. O_i = observed concentration i. N = number of paired obs/model concentrations. = mean modeled concentration. = mean observed concentration.

Observations and model predictions should initially be paired in time and space, matching the temporal resolution of the raw observation data. When matching in time, modelers should pay special attention to making sure that monitor and model data are both in local standard time. The units of time associated with model and observed concentrations can be 24-hour (usually for PM and its species), hourly (usually for species with continuous measurements, like ozone), or sometimes weekly (such as for CASTNet filter pack measurements). Temporal averaging is appropriate when used to match a relevant regulatory metric such as maximum 8-hr daily average ozone. As appropriate, air agencies should then aggregate the raw statistical results into meaningful groups of sub-regions or sub-periods. For example, examining model performance within and near the non-attainment area(s) of interest may be more important than performance in other parts of the modeling domain. For larger areas with more spatial variation, it may also be informative to evaluate sub-regions within the non-attainment area. In addition to any spatially aggregated statistics, model performance should also be evaluated at individual monitors within the nonattainment area. Similarly, priority may be placed on examination of the days that are potentially used in the attainment test (i.e., base period days with 8-hour ozone > 60 ppb). That is not to say that model performance evaluations should ignore performance on lower ozone days or in areas outside of the nonattainment areas. Model performance on lower concentration days and in areas outside of the nonattainment area are still important, but it is appropriate to give more attention to the model outputs that most directly impact the outcome of the attainment test. Unlike attainment demonstrations, model performance on days with low concentrations may be especially important for regional haze analyses. Some IMPROVE sites routinely measure extremely low PM species concentrations (even on the most impaired days).

In terms of pairing model predictions with monitored observations, the EPA recommends that the grid cell value in which the monitor resides be used for the calculations. It would also be acceptable to consider bi-linear interpolation of model predictions to specific monitoring

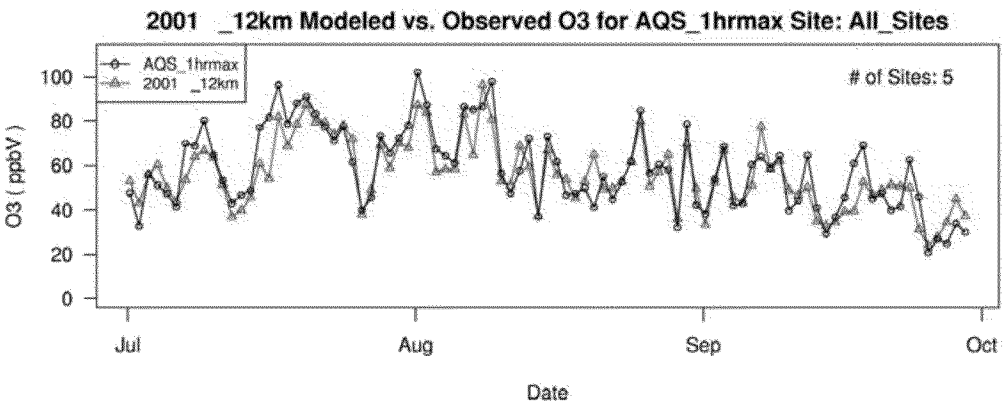
locations.[29] Air agencies should recognize that, even in the case of perfect model performance, model-observed residuals are unlikely to result in exact matches due to measurement uncertainty and differences in the spatial extent represented by model predictions, which are volume averages, and the observations, which are point values.

3.2.2 Plots

In addition to statistical summaries, graphical displays of data allow for a fuller characterization of model performance. Therefore, plots play a key role in any model performance evaluation. Below are examples of some types of plots which have been useful in past evaluations.

Time series plots of model and predicted hourly and/or daily concentrations (e.g., maximum 1-hour average, maximum 8-hour average, or 24-hour average) for each monitoring location in the nonattainment area, as well as key sites outside of the nonattainment area. These plots are especially important for examining performance at individual monitors and show how well the model captures temporal variations in pollutant concentrations at individual locations. Time series plots can indicate if there are particular times of day or days/episodes when the model performs especially poorly.

Figure 3.2.1 Example time series plot of daily 1-hour maximum ozone concentrations



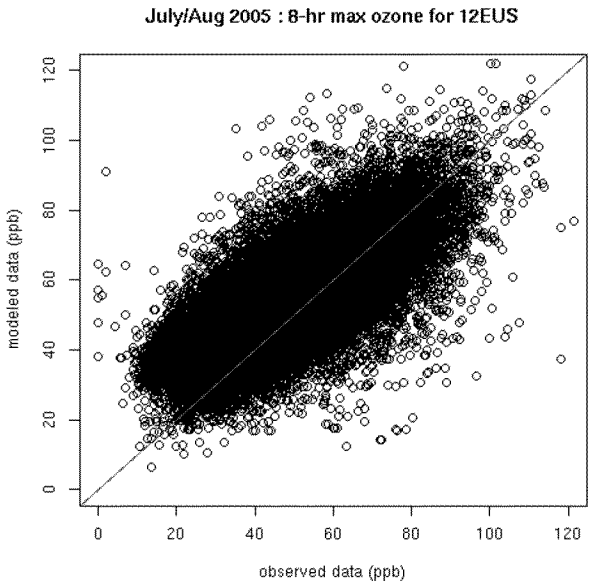
Scatter plots of predicted and observed concentrations at each site. These plots are useful for indicating if there is a particular part of the distribution of observations that is poorly

[29]In certain instances, air agencies may also want to conduct performance evaluations using a "near the monitor" grid cell array (i.e., the best match within a 3X3 grid cell array centered on the monitor). A "near the monitor" analysis may be useful when strong ozone gradients are observed, such as in the presence of a sea breeze or in strongly oxidant limited conditions.

ED_002666_00014243-00075

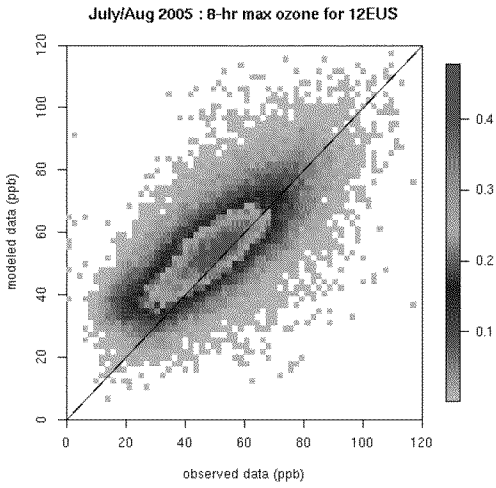
represented by the model. It may also be useful to develop separate plots for individual time periods or key sub-regions.

Figure 3.2.2 Example scatter plot of 8-hour daily maximum ozone concentrations



Density plots are a variation on scatter plots which are especially useful with large amounts of data when individual points overlap on a scatter plot. Density plots color code each point on the scatter plot based on the number of points falling in that location.

Figure 3.2.3 Example density plot of 8-hour daily maximum ozone concentrations



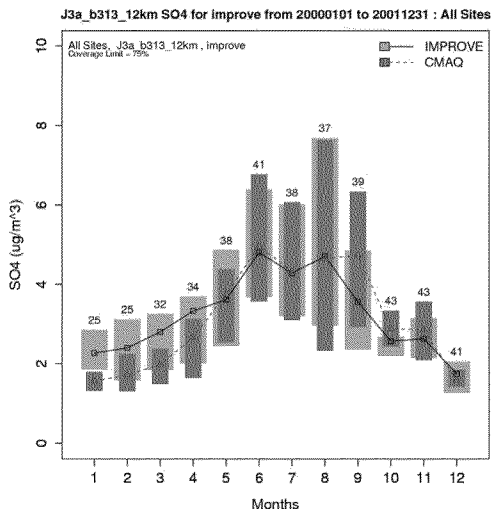
Box plots can be developed for model performance evaluation. These types of plots show the distribution of observations, model estimates, or performance metrics, which can be grouped

ED_002666_00014243-00076

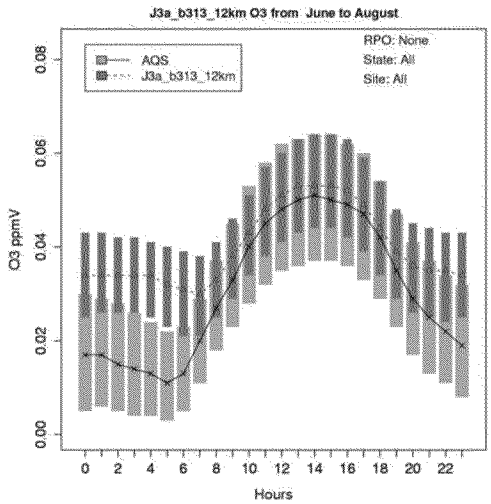
in a variety of ways, most commonly by observed ozone concentration, month, or hour of the day. Box plots can show several quantities: the 25% to 75% percentiles, the median values, and outliers. The monthly box plot can be used to quickly visualize model performance across the entire year, highlighting the seasonal change in model performance. The hourly or "diurnal" box plot is constructed using hourly data, and shows how the model predictions compare against observations throughout an entire day.

Figure 3.2.4 Example box plots of (a) monthly average sulfate concentrations, (b) daily diurnal average ozone concentrations, and (c) seasonal ozone bias binned by concentration range

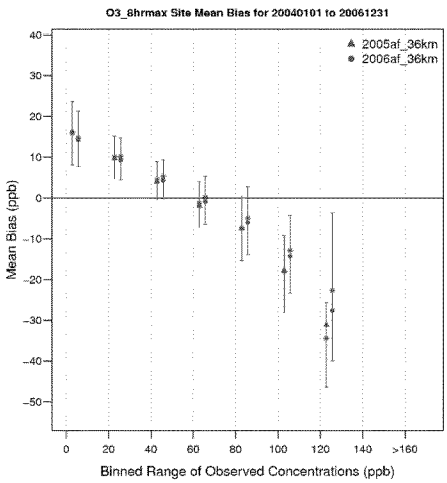
(a)



(b)

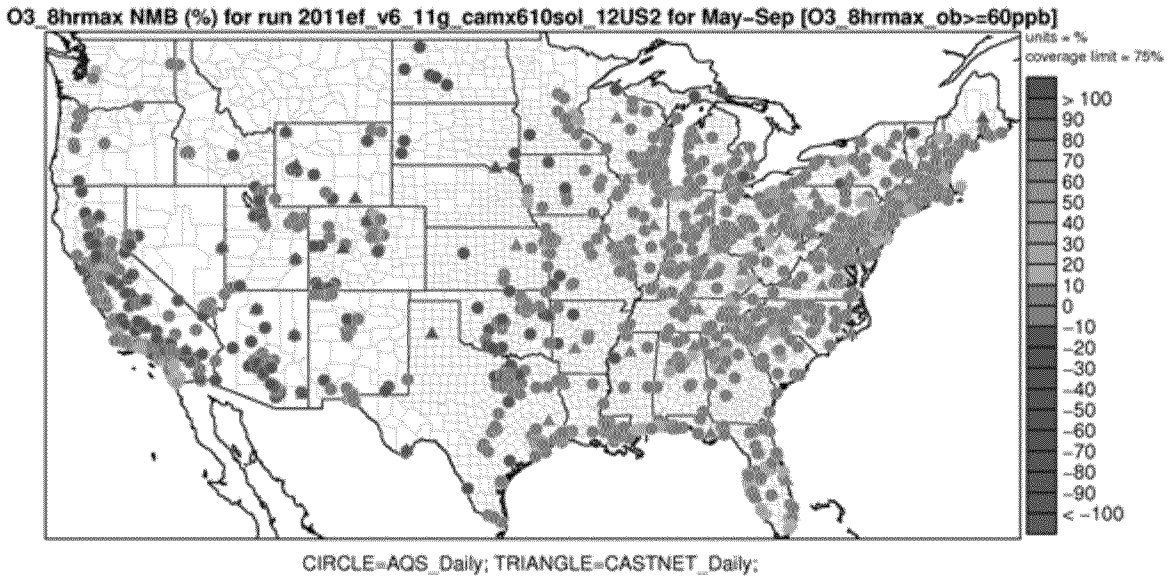


(c)



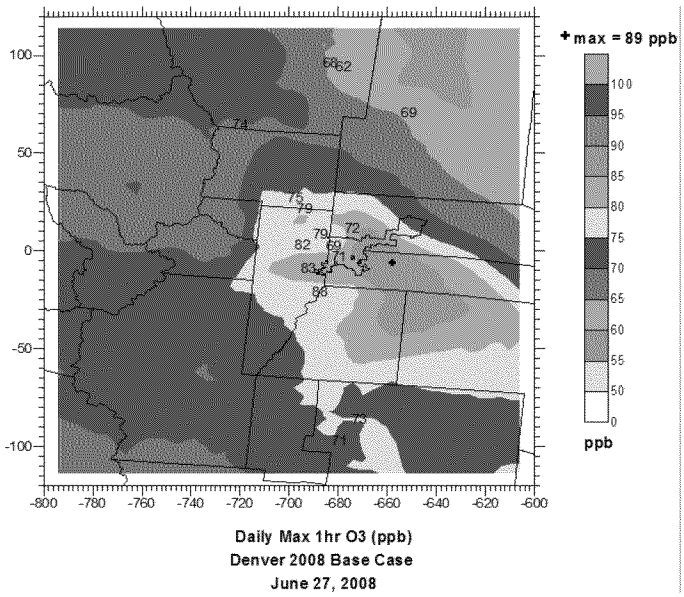
Spatial plots of model performance at monitor locations provide an overall picture of the geographic patterns in model performance. Any performance metric can be plotted in this manner, but it is most common to include spatial plots of MB, ME, NMB, and NME. Spatial plots of correlation may also be useful. These plots are most useful when constructed on a monthly or seasonal basis. For ozone these plots may be constructed to include all days or only "high ozone days." For regional haze, these plots may be constructed to include only the 20% most impaired days or 20% clearest days.

Figure 3.2.5 Example of spatial plot of NMB at monitor locations for days with O_3 above 60 ppb during the 2011 ozone summer season (May-September).



Daily gridded tile plots of predicted concentrations across the modeling domain with the actual observations as an overlay. Plots should be prepared for relevant averaging time (i.e., 1-hour maxima, daily 8-hour maxima, 24-hour average, etc.). These plots can reveal locations where the model performs poorly. Superimposing observed concentrations on the predicted isopleths reveals useful information on the spatial alignment of predicted and observed plumes.

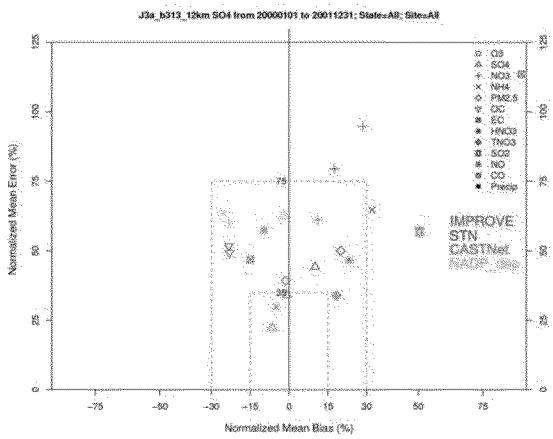
Figure 3.2.6 Example gridded tile plot of daily maximum 1-hour ozone concentrations (with overlaid observations)



Animations of predicted hourly concentrations for all episode days or for certain periods of interest. Animations are useful for examining the timing and location of pollutant formation. Animations may also reveal transport patterns (especially when looking at aloft layers). Animations can also be used to qualitatively compare model outputs with the conceptual model of particular pollution episodes. These plots may also reveal whether an important feature was simulated but was spatially or temporally displaced.

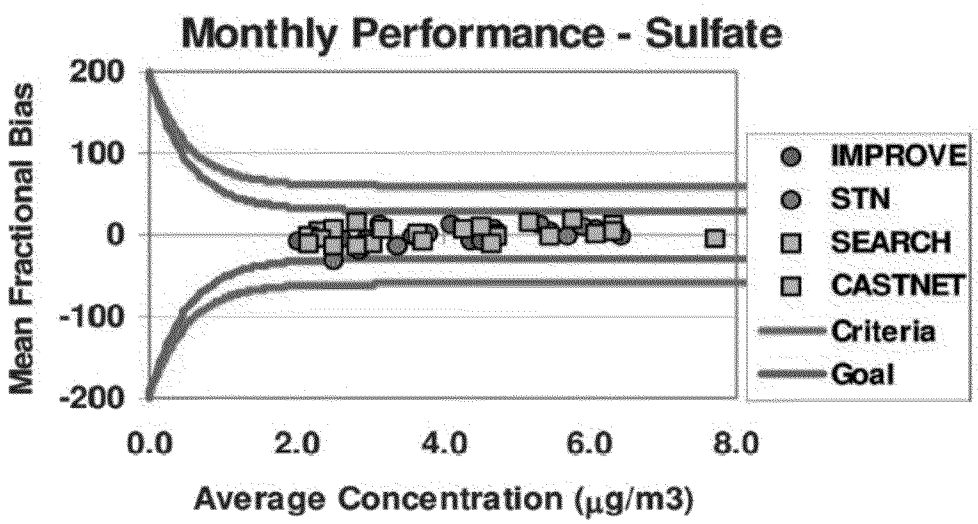
The soccer plot (Tesche et al, 2006) is so named because the dotted lines illustrating performance goals resemble a soccer goal. The plot is a convenient way to visualize model performance of both bias and error on a single plot. As bias and error approach zero, the points are plotted closer to or within the "goal," represented by the dashed boxes.

Figure 3.2.7 Example soccer plot of sulfate normalized mean bias and error.



The bugle plot (Boylan and Russell, 2006; Tesche et al., 2006), named for the shape formed by the criteria and goal lines, is another plot available for model performance evaluation. The bugle plots are shaped as such because the goal and criteria lines are adjusted based on the average concentration of the observed species. As the average concentration becomes smaller, the criteria and goal lines become larger to adjust for the model's ability to predict at low concentrations.

Figure 3.2.8 Example bugle plot of sulfate mean fractional bias as a function of modeled concentration.



Source: Morris et al, 2005

3.2.3 Ozone Model Performance

It is recommended that, at a minimum, statistical performance metrics be calculated for hourly ozone and 8-hour ozone maxima for each day of the model simulation used to support the attainment demonstration. The EPA recommends that the metrics be calculated both for all pairs of modeled and observed data and for pairs in which the hourly or 8-hour daily max values are above 60 ppb. Another acceptable approach would be to estimate metrics binned by observed ozone concentration. This would allow for a thorough evaluation of the model tendencies to estimate ambient concentrations in different observed ranges of interest. Wherever possible, these types of performance measures should also be calculated for ozone precursors and related gas phase oxidants.

Ozone precursors and related gas-phase oxidants (partial list) are listed below.

- Total Volatile Organic Compounds (VOCs)
- Speciated VOC

- Nitric acid (HNO_3)
- Nitric oxide (NO)
- Nitrogen dioxide (NO_2)
- Peroxyacetyl nitrate (PAN)
- NOy (NOx=NO+NO_2 and other oxidized nitrogen compounds)
- Carbon monoxide (CO)

3.2.4 $PM_{2.5}$ Model Performance

Performance metrics for PM and regional haze are similar to those calculated for ozone. Metrics should be estimated for components of $PM_{2.5}$ and PM precursors. Since modeling for $PM_{2.5}$ and regional haze will likely require modeling different times of year, season-specific statistics and graphical displays are helpful for evaluating and diagnosing model performance. Statistics and graphics can be averaged for various time scales depending on the model application. For example, statistical metrics and scatterplots can show daily averaged ambient-modeled pairs or annual average pairs for attainment demonstration of the daily and annual $PM_{2.5}$ standards respectively. In addition, the EPA recommends that daily pairs be grouped by season or month and statistics be given for different times of the year since different processes dominated $PM_{2.5}$ formation and fate in different seasons. Note that in this case, pairings should be performed on 24-hour average concentrations and statistics may then be reported for subgroups (seasons, regions etc.) of this data. This is different than pairing monthly or seasonally averaged data since statistics and plots tend to look "better" as the averaging time increases from daily to monthly to quarterly to annual. As such, daily pairs should always be examined to ensure a detailed look at model performance on the time scale of the measurements (24-hour average). Finally, air agencies may also want to subset the 20% clearest and most impaired visibility days at IMPROVE sites or the "high" modeled $PM_{2.5}$ days (e.g. the 10% highest modeled days in each quarter) to examine model performance on the days which will potentially be used to evaluate air quality goals.

Because $PM_{2.5}$ is a mixture, a meaningful performance evaluation should include an assessment of how well the model is able to predict individual chemical components that constitute $PM_{2.5}$. Components of $PM_{2.5}$ of interest include:

- Sulfate ion (SO_4)
- Nitrate ion (NO_3)
- Ammonium ion (NH_4)
- Elemental Carbon (EC)

ED_002666_00014243-00081

- Organic Carbon (OC) and/or Organic Carbon Mass[30]
- Crustal (weighted average of the most abundant trace elements in ambient air)
- Sea salt constituents (Na and Cl)[31]
- Other mass including Non-Carbon Organic Matter (NCOM) and metal oxides not captured by the crustal component

There are a number of gas-phase species that can be used to evaluate $PM_{2.5}$. Their presence may affect the formation of secondary PM and the response of secondary PM components to emissions strategies. Also, the model's performance in predicting certain gaseous species may provide diagnostic clues which help explain poor performance. Gas phase species of most interest for performance evaluation due to being a direct precursor to secondarily formed $PM_{2.5}$ are listed below:

- Nitric acid (HNO_3)
- Ammonia (NH_3)
- Sulfur dioxide (SO_2)
- Volatile Organic Compounds (VOCs)

Other related gas-phase oxidants (partial list) that indirectly influence $PM_{2.5}$ formation and destruction are listed below.

- Ozone (O_3)
- Nitric oxide (NO)
- Nitrogen dioxide (NO_2)
- Peroxyacetyl nitrate (PAN)
- NOy (sum of NOx and other oxidized compounds)
- Carbon monoxide (CO)
- Hydrogen peroxide (H_2O_2)

Some of the species listed above are currently available from either existing monitoring network systems and/or special studies. However, it is important to note that many of the species listed above are not routinely measured (e.g., H_2O_2, HNO_3, and NH_3) or measurements are not co-located with speciated PM measurements. Model performance can best be assessed using extensive databases, such as those obtained in major field studies. Where possible, model

[30] For predicted/observed comparisons, organic carbon is preferred over organic carbon mass to allow for transparency between model predictions and observed measurements. Ambient measurements will generally only account for the carbon portion of organic carbon. Organic carbon mass may include hydrogen, oxygen, and other components.

[31] White (2008) recommends the use of $1.8xCl^-$ to characterize ambient sea salt concentrations at coastal IMPROVE sites due to uncertainty in the XRF Na measurements.

ED_002666_00014243-00082

time periods can be prioritized to correspond to periods when field study or special study data was collected. However, it is recognized that many of the precursor species will not be available for every model application.

3.3 Ambient Measurement Networks

Provided below is an overview of some of the various ambient air monitoring networks currently available. Network methods and procedures are subject to change annually due to systematic review and/or updates to the existing monitoring network/program. Please note, there are other available monitoring networks which are not mentioned here and more details on these networks and measurements can be obtained from other sources, if necessary.

3.3.1 AQS

The Air Quality System (AQS) is a repository of ambient air pollution data and related meteorological data collected by EPA, state, local and tribal air pollution control agencies from tens of thousands of monitors (it is not an actual monitoring network). The monitoring data in AQS are the result of the various CAA requirements to provide a national database of ambient air pollution data. This information management system contains data on over 1,000 pollutants from 1957 through the present day. AQS contains all the routine hourly gaseous pollutant data collected from State and Local Air Monitoring Stations (SLAMS) sites. SLAMS is a dynamic network of monitors for state and local directed monitoring objectives (e.g., control strategy development) which consists of thousands of monitoring stations. The SLAMS network includes stations classified as NCore, PAMS, and Speciation, and stations formerly categorized as NAMS, but does not include Special Purpose Monitors (SPM) and other monitors used for non-regulatory or industrial monitoring purposes. AQS also contains data from the IMPROVE monitoring network, described separately in section 3.3.3 of this document.

The AQS database includes criteria pollutant data (SO_2, NO_2, O_3, CO, PM_{10}, $PM_{2.5}$, and Pb) and speciation data of particulate matter (SO_4, NO_3, NH_4, EC, OC, and crustal material), air toxic data, photochemical assessment data, and meteorological data. The data are measured and reported on an hourly or daily average basis. The AQS system continues to expand to include more ambient air pollutants. An overview of the AQS can be found here. For querying the database or viewing the data one can use the EPA's AirData website, which is a collection of user-friendly visualization tools for air quality analysts. The tools generate maps, graphs, and data tables dynamically. An AQS User's Guide is also available.

ED_002666_00014243-00083

3.3.2 NCore

The term "NCore monitoring network" (for national core) refers to a set of monitoring sites in major urban areas that generally started operating in 2011. These sites combine several advanced measurement systems for particles, pollutant gases and meteorology. Specifically, chemical measurements at these sites includes $PM_{2.5}$ speciation, continuous mass, 24-hr filter-based mass, ozone, CO, SO_2, NO, NOy. More information about this network is available here.

3.3.3 IMPROVE

The Interagency Monitoring of PROtected Visual Environments (IMPROVE) 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 U.S., mostly at national parks, national wilderness areas, and national wildlife refuges. As of 2018, there were approximately 160 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 additional IMPROVE sites are "IMPROVE protocol" sites, which are generally located in rural areas throughout the U.S., although there are also a handful of urban sites in the U.S. These protocol sites provide additional spatial information across the country, being generally located in areas where there are few Class I areas. The protocol sites use the IMPROVE monitoring samplers and collection routines. In addition to IMPROVE data that is available in AQS, the IMPROVE program provides summary datasets that contains information and pre-calculated data needed for Regional Haze Rule analyses. This includes daily average and annual data for the 20% most impaired and 20% clearest visibility days. The IMPROVE documentation and data can be found here. IMPROVE data advisories are posted here.

3.3.4 CASTNet

Established in 1987, the Clean Air Status and Trends Network (CASTNet) is a dry deposition monitoring network where PM data are collected and reported as weekly average data (U.S. EPA, 2012a). In addition, this network measures and reports hourly ozone concentrations. CASTNet provides atmospheric data on the dry deposition component of total acid deposition, ground-level ozone and other forms of atmospheric pollution. The data (except for ozone) are collected in filter packs that sample the ambient air continuously during the week. As of 2018, CASTNet is comprised of 95 monitoring stations across the U.S. The longest data records are primarily at eastern U.S. sites. More information can be obtained through the CASTNet website.

ED_002666_00014243-00084

3.3.5 CSN

The Chemical Speciation Network (formerly known as STN: The Speciation Trends Network) began operation in 1999 to provide nationally consistent speciated $PM_{2.5}$ data for the assessment of trends at representative sites in urban areas in the U.S. The CSN was established by regulation and is a companion network to the mass-based Federal Reference Method (FRM) network implemented in support of the $PM_{2.5}$ NAAQS. As part of a routine monitoring program, the CSN quantifies mass concentrations and $PM_{2.5}$ constituents, including numerous trace elements, ions (sulfate, nitrate, sodium, potassium, and ammonium), elemental carbon, and organic carbon. As of 2018, there were 52 trends sites in the CSN nationally. CSN trends sites are largely static urban monitoring stations with protocols for sampling methods that are dedicated to characterizing aerosol mass components in urban areas of the U.S. to discern long-term trends and provide an accountability mechanism to assess the effectiveness of control programs. In addition, in 2018, there were approximately 100 supplemental speciation sites that are also part of the CSN. The CSN data at trends sites are collected 1 in every 3 days, whereas supplemental sites collect data either 1 in every 3 days or 1 in every 6 days. Comprehensive information on the CSN and related speciation monitoring can be found here.

3.3.6 NADP

Initiated in the late 1970s, the National Acid Deposition Program (NADP) monitoring network began as a cooperative program between federal and state agencies, universities, electric utilities, and other industries to determine geographical patterns and trends in precipitation chemistry in the U.S. NADP collects and reports wet deposition measurements as weekly average data (NADP, 2013). The network is now known as NADP/NTN (National Trends Network), with nearly 200 sites in operation. The NADP analyzes the constituents important in precipitation chemistry, including those affecting rainfall acidity and those that may have ecological effects. The NTN measures sulfate, nitrate, hydrogen ion (measure of acidity), ammonia, chloride, and base cations (calcium, magnesium, potassium). Detailed information regarding the NADP/NTN monitoring network can be found here.

The Ammonia Monitoring Network (AMoN) is a NADP network that provides additional ambient and deposition ammonia data. The network began as part of a special study in 2007, was approved as an NADP network in 2010, and as of 2016 included approximately 50 monitoring sites across the U.S. AMoN monitors collect 2-week average data. More information can be found here.

ED_002666_00014243-00085

3.3.7 SEARCH

The South Eastern Aerosol Research and Characterization (SEARCH) monitoring network was established in 1998 and is a coordinated effort between the public and private sector to characterize the chemical and physical composition as well as the geographical distribution and long-term trends of $PM_{2.5}$ in the Southeastern U.S. (Edgerton, 2005; Hansen, 2003) The SEARCH network officially shut down in December 2016. However, SEARCH data are still useful for evaluating model performance for any simulations that included the Southeastern U.S. between 1998 and 2016. SEARCH data were collected and reported on an hourly/daily basis. A total of eight sites were operated at one time by SEARCH, although only six operated during SEARCH's final years: Birmingham, Alabama (urban), Centreville, Alabama (rural), Gulfport, Mississippi (urban), Jefferson Street, Atlanta, Georgia (urban), Oak Grove, Mississippi (rural) (not operational as of 2010), Yorkville, Georgia (rural) (not operational as of 2009), suburban Pensacola, Florida (suburban), and downtown Pensacola, Florida (urban). Historical SEARCH data from 1992-2016 can be accessed here.

3.3.8 Speciated $PM_{2.5}$ Data Summary

Speciated $PM_{2.5}$ data play an especially important role in performance evaluation, as the deterministic models predict exact chemical components which can be compared to some of the corresponding measured analytes. There are known positive and negative sampling artifacts as well as analysis artifacts associated with FRM and speciation monitors (Frank, 2006). Due to the complex nature of the measurements, a basic understanding of the various $PM_{2.5}$ and speciation monitoring technology and measurements are needed before undertaking a model evaluation.

Table 3.2: Known measurement artifacts for $PM_{2.5}$ and speciated PM for different monitoring networks.

Network	Measurement	Notes	Sampling Time
FRM	$PM_{2.5}$	- Negative artifacts: ammonium nitrate, organic carbon - Positive artifacts: organic carbon, water	24-hour average 1/3 days or 1/6 days (some daily sites)
	$PM_{2.5}$	- Negative artifacts: ammonium nitrate, organic carbon	

ED_002666_00014243-00086

Network	Measurement	Notes	Sampling Time
CSN (formally STN)		- Positive artifacts: organic carbon, water -Note that positive and negative artifacts for $PM_{2.5}$ may be different from artifacts for individual PM components because the speciated $PM_{2.5}$ samples are collected on different types of filters than are used for total $PM_{2.5}$ gravimetric measurements -CSN $PM_{2.5}$ mass may be different from FRM measurements due to differences in handling procedures	24-hour average 1/3 days or 1/6 days
	SO_4		
	NO_3		
	NH_4		
	OC	- Blank correction is needed to account for positive artifact[32] - Before 2007, OC/EC split was operationally defined using Thermal Optical Transmittance (TOT) method (NIOSH). CSN monitors switched to using the Thermal Optical Reflectance (TOR) method (consistent with	

[32] CSN OC measurements reported in AQS currently have artifact correction applied, although ambient data provided in SMAT-CE does already include a blank correction. The same artifact correction techniques are applied to measurements from CSN and IMPROVE networks. The CSN began implementing IMPROVE-like samplers and the IMPROVE analysis method in 2007. The implementation was fully completed in late 2009. Work by Malm et al (2011) suggests that CSN OC from 2005 and 2006 (before the switch in samplers) likely had a positive multiplicative bias of ~1.2 and a positive additive bias between 1.0 and 1.9 $\mu g/m^3$ depending on the time of year. We recommend using artifact corrected measurements where available and adjusting any uncorrected AQS CSN OC data by subtracting 1.4 $\mu g/m^3$ for measurement before the switch to IMPROVE-like measurements and by subtracting 0.12 $\mu g/m^3$ from measurements after the switch based on Frank (2012).

ED_002666_00014243-00087

Network	Measurement	Notes	Sampling Time
CSN (formerly STN) (con't.)		IMPROVE monitors) between 2007and 2009.[33]	
	EC	- Before 2007, OC/EC split was operationally defined using TOT method (NIOSH). CSN monitors switched to using the TOR method (consistent with IMPROVE monitors) between 2007 and 2009(see footnote 23)	
CASTNet	SO_4		weekly average
	NO_3	Split between NO_3 and HNO_3 may not be reliable in warm weather due to conversion from NO_3 to HNO_3. Recommend comparing total nitrate (NO_3 + HNO_3) between model and observations.	
	NH_4		
	SO_2		
	HNO_3	Split between NO_3 and HNO_3 may not be reliable in warm weather due to conversion from NO_3 to HNO_3. Recommend comparing total nitrate (NO_3 + HNO_3) between model and observations.	
	$PM_{2.5}$	- Negative artifacts: ammonium nitrate, organic carbon - Positive artifacts: organic carbon, water	

[33] The transition from TOT to TOR for CSN monitors occurred in shifts: 56 sites switched in May 2007, 63 sites switched in April 2009, and 78 sites switched in October 2009.

ED_002666_00014243-00088

Network	Measurement	Notes	Sampling Time
IMPROVE		-Note that positive and negative artifacts for $PM_{2.5}$ may be different from artifacts for individual PM components because the speciated $PM_{2.5}$ samples are collected on different types of filters than are used for total $PM_{2.5}$ gravimetric measurements	24-hour average 1/3 days
	SO_4 and S	-IMPROVE monitors measure both elemental S (by XRF) and ionic SO_4^- by ion chromatography. Generally, either measurement is reliable and may be used for model evaluation purposes although some caveats should be noted.[34]	
	NO_3		
	NH_4	-Only available at a limited number of sites for a limited number of years	
IMPROVE (con't.)	Organic Carbon	- OC/EC split is operationally defined using Thermal Optical Reflectance (TOR) analysis. - Blank corrections based on a median measured positive OC artifact from back-up filters at six sites are applied on a month-specific basis to all data before they are reported (Chow et al, 2010; Watson et al., 2009). Frank (2012) recommended changing blank	

[34] Hyslop et al (2012) report that XRF elemental sulfur measurements are somewhat affected by changes in XRF methodologies over time. Also, prior to 2000, filters used to collect sulfate for ion chromatography became clogged during high sulfate episodes. Changes in filter collection techniques eliminated this problem after 2000. In addition, recent literature by Tolocka and Turpin (2012) suggests that differences between XRF sulfur and ion chromatography sulfate are due to the presence of organic sulfur in aerosol (which would be measured by XRF but not IC). If this is the case, then the IC sulfate measurement is more appropriate for comparison with modeled sulfate which generally does not include formation of organic sulfur compounds.

ED_002666_00014243-00089

Network	Measurement	Notes	Sampling Time
		correction procedures for both IMPROVE and STN OC measurements. These recommendations have not yet been implemented.	
	Elemental Carbon	- OC/EC split is operationally defined using Thermal Optical Reflectance (TOR) analysis.	
	Sodium	XRF analysis is not as sensitive to sodium as other methods and generally under-reports sodium concentrations.	
	Silicon	XRF silicon measurements are unreliable at high sulfur conditions (when S/Si > about 7)	
	Aluminum	Elemental concentrations above the MDL can go undetected. IMPROVE has published an alternate soil calculation which estimates Al based on Si concentrations.[35]	

In general, the speciated $PM_{2.5}$ data can be directly compared to the model outputs. Of note, care should be taken in examining issues with blank corrections for organic carbon (especially when measured OC is low) as well as the EC/OC split. Prior to 2006, the CSN and IMPROVE networks used different analysis techniques when estimating EC and OC (CSN monitor locations switched to a similar method to IMPROVE between 2007 and 2009, depending on the monitor). The EC/OC analysis method will generally have a larger impact on EC concentrations than OC concentrations. The difference in OC/EC splits between the two techniques has been investigated in depth in the literature and by IMPROVE staff (Chow, 2001, Chow, 2010, White, 2007). It may not be appropriate to compare EC concentrations between networks for years prior to 2009 (although the data can still be used to evaluate model performance).

Total $PM_{2.5}$ mass measured gravimetrically at CSN and IMPROVE monitors is known to have several measurement artifacts (both positive and negative) (Frank, 2006, Simon et al., 2011,

[35] Due to difficulties in detecting Al, IMPROVE sometimes uses an alternative SOIL calculation which does not include Al: SOIL = 3.48*Si + 1.63*Ca + 2.42*Fe + 1.94*Ti (http://vista.cira.colostate.edu/improve/publications/SOPs/ucdavis_sops/sop351_V2.pdf).

ED_002666_00014243-00090

Malm et al., 2011). Consequently, summing the mass of all $PM_{2.5}$ species components is not likely to give the same answer as the total measured $PM_{2.5}$ mass from a Teflon filter. Therefore, model evaluations should concentrate on comparisons with speciated $PM_{2.5}$ components at these sites, rather than total $PM_{2.5}$ mass.

FRM monitors do not measure all of the $PM_{2.5}$ in the air, and the speciation samplers don't measure $PM_{2.5}$ in the same way as the FRM monitors (Frank, 2006). Due to these positive and negative artifacts, the FRM data may not be directly comparable with model outputs without adjustment. The models predict PM species as they might be measured in the ambient air, but the FRM measurements may not always measure what is in the air (due to known positive and negative artifacts). As part of the $PM_{2.5}$ attainment test, we recommend default methodologies to adjust speciation (CSN and IMPROVE) data to better approximate the FRM data (see sections 4.4 and 4.5). A similar adjustment could be made to the model output data to allow the direct comparison to FRM measurements. However, as noted earlier, there may be limited value of information gained by evaluating daily average total measured $PM_{2.5}$ mass (as compared to $PM_{2.5}$ components). More meaningful information may be gleaned from continuous $PM_{2.5}$ ambient data.

Another issue to consider when comparing modeled and observed PM concentrations is the size definitions for PM. IMPROVE and CSN ambient monitors measure $PM_{2.5}$ using inlets which screen out any particles with aerodynamic diameters greater than 2.5 µm. Air quality models use varying schemes to characterize particulate matter size distributions. The two most common modeling schemes are the sectional treatment and the modal treatment. The default configuration for CAMx includes a sectional aerosol treatment while the default configuration for CMAQ includes a modal aerosol treatment, although other versions of these models exist with alternate aerosol treatments. Sectional aerosol modules split particles into bins defined using discrete size cut-points based on aerodynamic diameter (e.g., particles with aerodynamic diameters between 1 µm and 2.5 µm). One common sectional set-up includes 2 bins, one containing particles with aerodynamic diameters < 2.5 µm (fine PM) and one containing particles with aerodynamic diameters > 2.5 µm (coarse particles). Other models have split the fine PM into more than two bins (Nolte et al., 2008; Pandis et al., 1993). With sectional models, the comparison to measured $PM_{2.5}$ is straightforward so long as one of the PM bins has a cutoff at an aerodynamic diameter of 2.5 µm (i.e., add up PM mass in all bins less than 2.5 µm). The modal approach treats PM as occurring in distinct "modes," each one having a lognormal distribution. Each PM mode is characterized by the geometric mean diameter and the geometric standard deviation of the diameter (the geometric mean and standard deviation can be tracked for particle number, surface area, mass or all three). With modal aerosol treatment, $PM_{2.5}$ can be estimated in several ways. One method is to simply sum the mass in the modes that generally contain fine PM. Another method would be to use the geometric mean and

ED_002666_00014243-00091

standard deviation to calculate the fraction of PM mass in each mode which is associated with particles less than 2.5 µm. However, there are several complicating factors with taking the second approach. First, the EPA's NEIy and many other inventories define $PM_{2.5}$ as being entirely less than 2.5 µm in diameter. If these primary emissions are injected into a modal model (generally into a fine PM mode), they are instantaneously mixed with existing aerosol mass in that mode. The tail of the fine PM mode will include particles greater than 2.5 µm. Therefore, a fraction of the PM that was operationally defined as primarily emitted $PM_{2.5}$ (diameter < 2.5 µm) would be incorrectly distributed into the > 2.5 µm size range based on the mode distribution. If primary emissions were quantified based on modes versus a discrete cut-point, then this would not be a problem. Secondly, some modal models have been shown to over-predict the peak-concentration diameter and distribution width of the fine PM mode (Kelly et al., 2011; Kelly et al., 2010). This would lead to under-predictions of the $PM_{2.5}$ fraction by increasing the mass fraction in the > 2.5 micron diameter range. Air agencies should consider these complications when deciding whether to calculate the fraction of PM mass in each mode below the 2.5 µm cut-point or to simply sum mass over the fine PM mode(s).

Finally, although we recommend weighing performance of speciated PM components more heavily than total PM, some additional insight can be gained by looking at Federal Equivalent Method (FEM) continuous $PM_{2.5}$ measurements[36]. Since speciated measurements are generally derived from filters, these measurements are usually taken as 24-hour averages. The FEM continuous measurements provide a diurnal profile of $PM_{2.5}$ concentrations that can be compared against hourly model outputs to determine whether the model is replicating the measured temporal pattern.

3.4 Diagnostic Evaluation

The goal of a diagnostic model evaluation is to determine whether the model correctly represents the physical and chemical processes that control the ambient concentrations of precursors and secondary pollutants. It is possible that a model can incorrectly represent the relative amount of primary emissions versus secondary production or that a model can incorrectly represent the importance of dispersion, deposition or chemical reaction as a sink. Thus, the ability of a model to accurately reproduce the observed concentration of a particular pollutant does not guarantee that the model correctly simulated the processes that determine the concentration of the pollutant. The goal of a diagnostic evaluation is to investigate the processes that determine the ambient concentrations of ozone, $PM_{2.5}$ and their precursors, and to develop confidence that the model can accurately predict how these concentrations will

[36] Monitoring of atmospheric air quality for purposes of determining compliance with the NAAQS generally requires the use of either Federal reference methods (FRM) or equivalent methods (FEM), as specified in Section 2.1 of Appendix C to 40 CFR part 58.

ED_002666_00014243-00092

change in the future (e.g., Arnold et al., 2006). It is likely that research grade measurements needed to perform a diagnostic evaluation will not be available in many SIP modeling studies, however, it can still be valuable to perform diagnostic analyses and compare the results with previous modeling research studies.

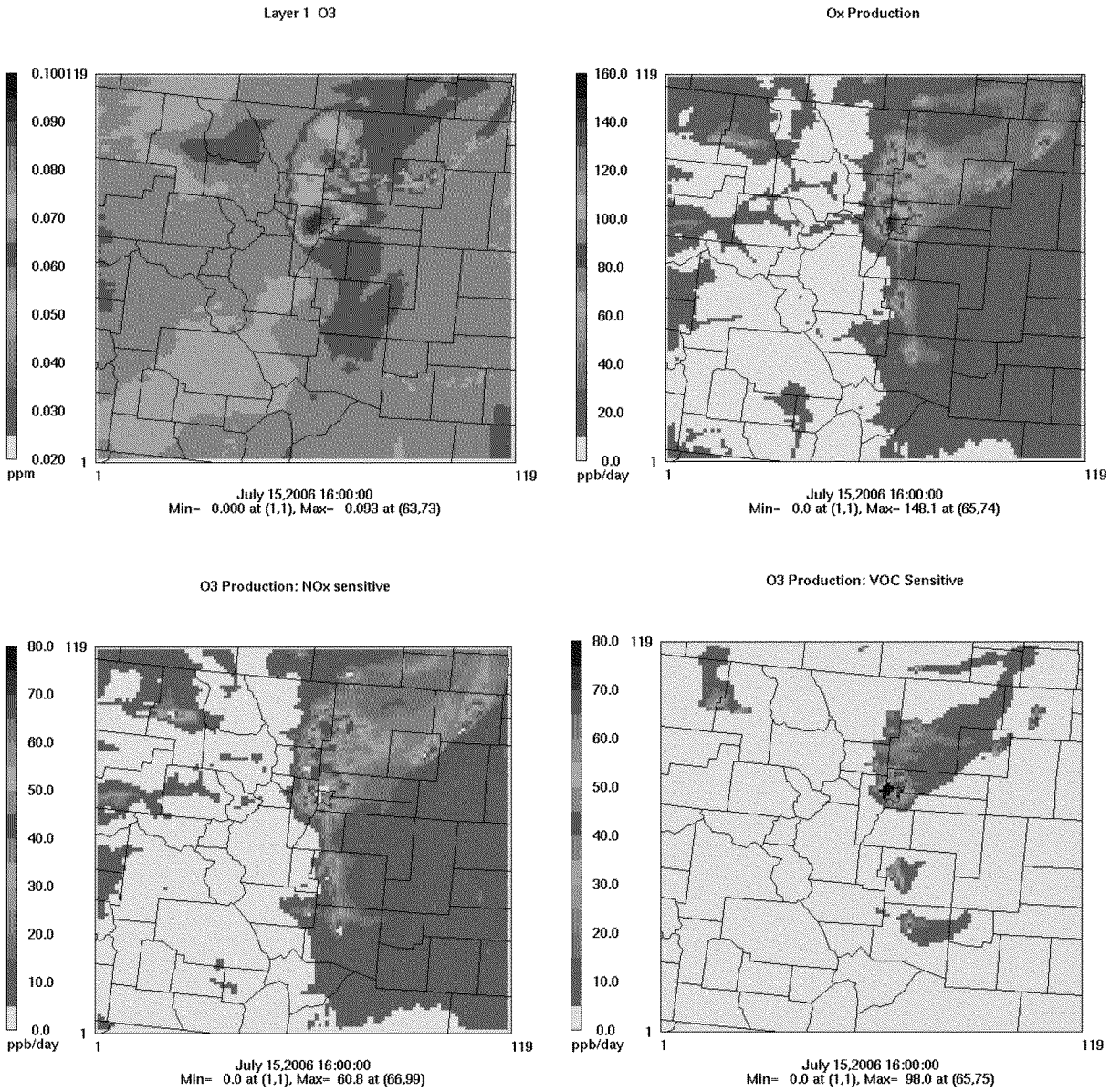
Several different tools and analysis methods are available for use in diagnostic model evaluation including process analysis, indicator ratios, source apportionment methods, model sensitivity studies, and retrospective studies.

3.4.1 Process Analysis

Both the CMAQ and CAMx models include a process analysis tool (Jeffries and Tonnesen, 1994) that provides additional diagnostic information on the chemical reactions and physical processes (e.g., emissions, deposition, advection, vertical mixing, and net change from chemistry) that determine species concentrations. Chemical reaction data can be output either as the hourly integrated reaction rate (IRR) for each reaction or as a summary of key reactions in the chemical process analysis (CPA). Physical processes can be output as hourly concentration changes in the integrated process rate (IPR) file. Detailed analysis off IRR and IPR outputs have been demonstrated in research studies (Jang et al., 1995; Vizuete et al., 2008; Henderson et al., 2011), but significant time and effort are required to complete these analyses. The CPA output provides a more limited set of chemical process information that is more accessible and that can be visualized using the same tools used to analyze model concentration output files. Minimal computational cost is required to produce the CPA files, and they are in the same format as the model concentration files and can be viewed using the same visualization tools. Below are examples of these plots comparing modeled ozone concentration, daily total Ox (Ox = O3 + O3P + O1D + NO2 +2*NO3 + PNA + PAN + 3*N2O5) production, and net ozone production in NOx limited and oxidant limited chemical regimes. The CPA tool can also provide information on the chemical reactions of VOC, NOy and radical species. The need for a process analysis evaluation and the specific processes investigated should be determined on a case-by-case basis.

ED_002666_00014243-00093

Figure 3.4.1 Chemical process analysis plots showing modeled ozone concentration, daily total Ox production, and net ozone production under VOC sensitive and NOx sensitive conditions.



3.4.2 Indicator Ratios

One way to evaluate the response of the model is to examine predicted and observed ratios of indicator species. Indicator species techniques have been developed for both ozone and secondary PM species (in particular nitrate) (Sillman, 1995, 1998; Ansari and Pandis, 1998;

Blanchard et al., 2000; Tonnesen and Dennis, 2000a and 2000b). If ratios of observed indicator species are very high or very low, they provide a sense of whether further ozone or secondary $PM_{2.5}$ production at the monitored location is likely to be limited by availability of NOx or VOC (or NH_3 in the case of particulate nitrate). Agreement between paired observed and predicted ratios suggests a model may correctly predict the sensitivity of ozone or secondary $PM_{2.5}$ at the monitored locations to emission control strategies. Thus, the use of indicator species aims to validate the modeled chemical regime and may help to build confidence in the RRFs predicted from the specific model application.

For ozone, measurements of certain "indicator species ratios" (e.g., comparisons between modeled and observed ratios: O_3/NO_y, O_3/HNO_3) are a potentially useful way to assess whether local ozone formation is VOC or NO_x limited at any particular point in space and time, as well as help reveal whether the model is correctly predicting the sensitivity of ozone to VOC and/or NO_x controls (Millford et al., 1994; Sillman, 1995, 1997, 1998, 2002; Lu and Chang, 1998). Tonnesen and Dennis (2000a and 2000b) distinguished between short lived indicators of instantaneous rates of production of ozone (PO_3) or odd oxygen (PO_x) versus long lived indicators of peak ozone concentration and found that the ratio of production of H_2O_2 to production of HNO_3 ($PH_2O_2/PHNO_3$) is strongly associated with PO_x sensitivity to VOC and NO_x. Tonnesen and Dennis also found that the ratio of formaldehyde to NO_2 was useful as an indicator of ozone sensitivity, and this ratio can be evaluated using satellite data (Martin et al., 2004; Duncan et al., 2010). For PM, such a comparison may reveal whether the model is predicting sensitivity of secondary components of $PM_{2.5}$ to changes in SO_2, NH_3, VOC and/or NO_x controls correctly (Ansari and Pandis, 1998; Blanchard et al., 2000; Pun and Seigneur, 1999, 2001). If a model accurately predicts observed indicator species relationships, then one can conclude with additional confidence that the predicted change in future year ozone or PM may be accurate.

Although ratios of indicator species have the potential to be quite useful, they should be used with caution since different studies have identified different ranges of values to characterize VOC and NOx limited regimes in different locations. Table 3.3 summarizes the range of values determined for 10 indicator species or ratios in various studies. In cases where the NOx and VOC limited ranges are broad or overlap, air agencies should attempt to use values from studies in the local area. Air agencies should be cautious with respect to the use of indicator species when the range of observed values does not provide a clear direction of control. When this occurs, agreement between predictions and observations does not necessarily imply that the response to controls as predicted by the model is correct (especially for secondary particulate matter to changes in precursors). If a model predicts observed values of indicator species relationships such that observed and predicted value fall within the same range, this provides

some confidence that the predicted change in ozone or secondary particulate matter may be accurate.

A second precaution is that this method often requires more types of measurements than are commonly available. In some cases, it may be difficult to calculate meaningful indicator ratios from relatively imprecise measurements from routine monitoring networks.

Table 3.3: Ranges of indicator species/ratios that have been reported for VOC and NOx limited regimes.

indicator	median VOC-sensitive	transition	median NO_x-sensitive
O_3/NO_y	1.6 - 8.8[a,b,d,f-m]	6 – 15[c,d,m,n,q]	8.8 – 26[a,b,d,f-m]
O_3/NO_z	5.8 – 16[a,d,f,j-m]	7 – 100[c-f,n,q]	10.1 – 35[a,d,f,j-m]
O_3/NO_x	< 15	15 – 60[n,o,q,t]	> 40
O_3/HNO_3	7 - 33.9[a,d,j,k,m]	12 – 25[c,d]	17.2 - 62.9[a,d,j,k,m]
H_2O_2/HNO_3	0.3 - 1.5[d,j-m]	0.2 - 2.4[c,e,h,n-q]	0.4 - 3.6[d,j-m]
H_2O_2/NO_z	0.2 - 1.1[a,j]	0.2 - 0.25[c]	0.6 - 1.3[a,j]
H_2O_2/NO_y	0.2 - 0.4[a,j]	0.12 - 0.17[c]	0.6 - 1.0[a,j]
$HCHO/NO_2$	< 1	1-2[n,o,u]	> 2
NO_z/NO_y	< .5	0.5-0.6[f]	> 0.6
NOy	7.7 - 12.2[j]	5 -2 0[c,e,f,n,r,s]	3.2 - 5.7[j]

[a]Castell et al., 2009; [b]Xie et al., 2011; [c]Torres-Jardon and Keener, 2006; [d]Sillman and He, 2002; [e]Sillman, 1995; [f] Lu and Chang, 1998; [g]Stein et al, 2005; [h]Sillman et al., 1997; [i]Martilli et al., 2002; [j]Jimenez and Baldasano, 2004; [k]Torres-Jardon et al, 2009; [l]Chock et al, 1999; [m]Sillman et al., 2003; [n]Zhang et al, 2009; [o]Tonnesen and Dennis, 2000a; [p]Hammer et al., 2002; [q]Liang et al., 2006; [r]Milford et al, 1994; [s]Vogel et al, 1995; [t]Tonnesen and Dennis, 2000b; [u]Martin et al., 2004

Much of the work done to date with indicator species has focused on peak concentrations of ozone. Limited research has been conducted on development of indicator ratios for secondary PM. This includes work by Blanchard et al., (2000) who examined how indicator species might be used to assess whether particulate nitrate concentrations are limited by NOx or by ammonia

ED_002666_00014243-00096

emissions using mechanisms which incorporate reactions dealing with secondary PM. These authors identify several indicators which appear potentially useful for determining limiting precursors for secondary nitrate PM. The indicators include:

(1) excess NH_3 defined in Equation 3-1 (values greater than zero indicate NOx limited regimes and values less than 0 indicate ammonia limited regimes);

$$_{i} = [\mathrm{NH3(g)}] + [\mathrm{NH4+(a)}] - 2\,x\,[\mathrm{SO42-(a)}] - [\mathrm{NO3-(a)}] - [\mathrm{HNO3(g)}] - [\mathrm{HCl(g)}] + 2\,x\,[\mathrm{Ca2+}] + 2\,x\,[\mathrm{Mg2+}] + [\mathrm{Na+}] + [\mathrm{K+}] + [\mathrm{Cl-}]$$

[3-1][37]

(2) The ratio of particulate to total nitrate. Values above 0.9 at high temperatures indicate NOx limited regimes (values below 0.3 to 0.5 were indicative of ammonia limited regimes).

Ansari and Pandis (1998) also have suggested the use of the gas ratio (GS), defined in equation 3-2, as an indicator of PM formation regimes.

$$= \frac{_{i}}{_{i}}$$

[3-2][38]

In summary, major challenges in evaluating indicator species include variability in the characteristic ratios for VOC and NOx limited regimes in different locations and the availability of ambient data. Despite these caveats, comparing predicted and observed relationships of indicator species provides a means of assessing a model's ability to accurately characterize the sensitivity of predicted ozone and predicted secondary components of $PM_{2.5}$ to changes in precursors.

3.4.3 Source Apportionment

Source apportionment methods and model sensitivity studies are two complementary approaches that can be used to identify emissions sources that contribute to high pollutant concentrations. While these tools do not directly diagnose the model's ability to replicate the real world, they are useful in their ability to further characterize the sources that are most important to high pollutant concentrations within the model. The CMAQ and CAMx models both include options to use source apportionment tools to identify sources that contribute to

[37]Concentrations are in µmol/m^3, (g) represent gas-phase and (a) represents aerosol phase.
[38] NH_3^F is free ammonia (total ammonia – 2*sulfate). HNO_3^T is total nitrate (gas + PM).

ED_002666_00014243-00097

the production of ozone (Ramboll, 2018; Kwok et al, 2014) and some components of $PM_{2.5}$, including sulfate, nitrate, ammonium and primary emissions of $PM_{2.5}$ (Ramboll, 2018; Kwok et al., 2013). Source apportionment tools employ a set of model tracers to attribute pollutant concentrations to user selected emissions source sectors and geographic areas. This approach can be useful both as a diagnostic tool to identify sources that contribute to ozone and $PM_{2.5}$ and also as an aid in regulatory analysis to identify emissions control measures that will be effective for attaining air quality standards. An advantage of source apportionment approaches is that tracers attribute the mass contribution from a source to a receptor and, therefore, this approach is less susceptible to the effects of non-linear chemistry, which can complicate the interpretation of model sensitivity simulations.

3.4.4 Sensitivity Studies

Model sensitivity simulations can be used to identify model inputs (IC/LBC, meteorological data, chemical reactions, and emissions) that have the greatest impact on model-predicted concentrations. The objective of sensitivity tests is to identify whether the inputs have a large enough influence that errors in these inputs could be a driving influence on the model predictions. In "brute-force" sensitivity simulations, the model input data is modified directly (e.g., increasing or decreasing emissions in the model input file). Brute-force simulations are relatively easy to implement but are resource intensive when many different inputs must be evaluated. The Direct Decoupled Method (DDM) is an approach that can be used to evaluate the first order sensitivity to multiple model inputs in a single model simulation (e.g., Cohan et al., 2005; Napelenok et al., 2006). Because of non-linearity in chemical reactions, first-order sensitivity is not always sufficient, and high-order DDM (HDDM) can be used to evaluate second order sensitivities to inputs that are subject to non-linear chemistry (Zhang et al., 2012b).

In cases where the model performs poorly in the base case, sensitivity simulation may be useful to identify potential causes for poor model performance. However, caution should be used in modifying inputs to achieve good model performance because arbitrarily "tuning" of model input data to fit the observations can introduce compensating errors. For example, if a model underestimates observed ozone because the meteorological data is overly dispersive, the model performance could be improved by increasing emissions or LBC concentrations, but this model would not be reliable for predicting future concentrations because its good performance in the base case was achieved through compensating errors in the input. An appropriate use of sensitivity approaches is to identify each of the processes or inputs that has a large effect on model predictions, and then perform additional analysis of those inputs to determine if they are accurately represented in the base case model.

ED_002666_00014243-00098

3.4.5 Dynamic Model Evaluation

A second method for assessing a model's performance in predicting the sensitivity of ozone or $PM_{2.5}$ species to changes in emissions is to perform a retrospective dynamic analysis. This involves comparing model predicted historical trends with observed trends. Retrospective analyses provide potentially useful means for diagnosing why a strategy did or did not work as expected. They also provide an important opportunity to evaluate model performance in a way that is closely related to how models are used to support an attainment demonstration.

A dynamic evaluation uses various analyses to assess the accuracy of the model in characterizing the sensitivity of ozone and/or $PM_{2.5}$ to changes in emissions. Although dynamic evaluation is always recommended, the measurements and resources needed for a dynamic evaluation may not be readily available to air agencies performing SIP modeling. Additional challenges for retrospective analyses include identifying appropriate case studies for which responses to emissions changes are large enough to be distinguished in ambient data, developing or obtaining equivalent and appropriate emissions for multiple historical years, and disentangling the confounding influences of year to year changes in meteorology and emissions.

A retrospective analysis is intended to examine the ability of the model to respond to emissions changes by comparing recent trends in observed ozone or $PM_{2.5}$ concentrations to the model-predicted trend over the same time period. The approach is a direct assessment of what is most important in an attainment demonstration: does the model accurately predict changes in air quality as a result of changes in emissions? As part of a retrospective analysis, the model is run for (a) current episodes or time periods and (b) episodes in one or more historical time periods using the emissions and meteorological inputs appropriate for each time period modeled.

While retrospective analyses may be useful, it may be difficult to obtain meteorological and emissions inputs for the historical time period(s) that are calculated using techniques and assumptions that are consistent with the calculation of these same inputs for the current time period. Using inconsistent inputs will confound the interpretation of the predicted trend. One method for ensuring consistent methodology for emissions estimates is to backcast emissions from the most recent year to the earliest year. The model should respond in a predictable way if the emissions changes are large enough. Hence, modeling an emissions change of only a few percent would not be viable. However, if NOx or SO_2 emissions have been reduced by a large percentage (e.g., 30% or more) over a relatively short time period, then that may be a good time period to test the response of the model.

Because differences in meteorology between years can confound the apparent change in pollutants, a complete retrospective analysis could include multiple sensitivity runs using both

ED_002666_00014243-00099

year-specific meteorology and constant meteorology. The year-specific meteorology would allow the modeler to compare modeled and observed trends most accurately, and the constant meteorology would help inform the analysis on what portion of the change in pollutant concentrations were due to emissions versus meteorological differences.

Retrospective analyses are one of the few tools that can be used to determine if the model is responding "adequately" or "correctly" to control measures. Recent efforts to perform this type of analysis have been undertaken by Zhou et al (2013), Gilliland et al. (2008), Godowitch et al. (2010), Pierce et al. (2010) and Foley et al. (2014). These published analyses may be used to better understand the ability of certain models to properly capture relationships between ambient pollution levels and emissions.

Another dynamic evaluation approach is to look at operational performance under varying conditions (e.g., by day of the week, by season, and by region) (Marr and Harley, 2002; Murphy et al, 2007; Harley et al, 2005). The mix of pollutants vary by day of the week and from area to area; when a model shows good operational performance across these different chemical environments, it supports the assertion that it will respond appropriately to changes in emissions. Two studies have performed model evaluations that focused on the model's ability to capture weekend/weekday changes (CRC, 2011; Pierce et al, 2010).

3.5 Evaluation Tools

There is available software and tools that can be used to process ambient and model data to create statistical output and plots for model performance evaluations. One example is the Atmospheric Model Evaluation Tool (AMET) (Appel, 2011). AMET is built on several open-source software packages and uses the MySQL database software to store paired observation and model output data. There are modules in AMET for evaluating both meteorological and air quality model output. AMET provides a useful set of scripts for creating various types of plots and statistics.

ED_002666_00014243-00100

4.0 Assessing Modeled Attainment for Ozone and $PM_{2.5}$

A *model attainment demonstration* usually consists of two key components: a) analyses which estimate whether a set of simulated emissions reductions corresponding to a control program scenario will result in ambient concentrations that meet the NAAQS and b) an identified set of control measures which will result in the required emissions reductions. This guidance focuses on the first component of a model attainment demonstration; that is, the completion and interpretation of analyses to estimate the amount of emission reduction needed to attain the NAAQS.

Air agencies should determine whether a control program scenario will provide sufficient emission reductions to demonstrate attainment of the NAAQS using the *modeled attainment test*. The modeled attainment test is a technical procedure in which an air quality model is used to simulate base year and future air pollutant concentrations for the purpose of demonstrating attainment of the relevant NAAQS. The recommended test uses model estimates in a "relative" rather than "absolute" sense to estimate future year design values. As explained in more detail in subsequent sections, the fractional changes in air pollutant concentrations between the model future year and model base year are calculated for all valid monitors. These ratios are called *relative response factors* (RRF). Future ozone and/or $PM_{2.5}$ design values are estimated at existing monitoring sites by multiplying the modeled relative response factor for each monitor by the monitor-specific base year design value. The resulting estimates of future concentrations are then compared to the NAAQS. If the future estimates of ozone and/or $PM_{2.5}$ design values do not exceed the NAAQS, then this provides evidence that attainment will be reached.

In addition to the modeled attainment test, air agencies should also consider performing a set of corroboratory analyses to further assess whether a proposed set of emission reductions is likely to lead to attainment of the NAAQS. As discussed later in this section (see section 6), the modeled attainment test and these supplemental analyses should be aggregated into a *weight of evidence determination* to evaluate whether the selected emissions reductions will yield attainment. The concept of relative response factors also applies to modeling for the purpose of setting RPGs for regional haze, the subject of section 5.

4.1 Overview of Modeled Attainment Test

There are four reasons why the recommended modeled attainment test is based on application of the model in a relative sense. First and foremost, this approach has the effect of anchoring the future concentrations to a "real" measured ambient value, which is important given model bias and error in the base year simulation(s). It is reasoned that factors causing bias (either under or over-predictions) in the base case will also affect the future case. While good model

ED_002666_00014243-00101

performance remains a prerequisite for use of a model for regulatory analyses, problems posed by imperfect model performance on individual days are expected to be reduced when using the relative approach. An internal EPA analysis (U.S. EPA, 2014a) considered whether daily ratios of model future/current maximum daily 8-hour ozone averages (MDA8) varied strongly as a function of site-specific base case model performance. This analysis was completed using a national model simulation that projected 2020 ozone concentrations from a 2007 base case. The analysis determined that when modeled, MDA8 ozone bias was relatively small (e.g., less than +/- 20 ppb), the average response ratios were not a strong function of the model MDA8 bias. This provides confidence that the model can detect the air quality response in the midst of reasonable levels of absolute bias and error. Second, the relative modeled attainment test allows for a future projection of 3-year average ozone and $PM_{2.5}$ design values or 5-year average visibility values without explicitly having to model a 3-year or 5-year period. Third, because $PM_{2.5}$ concentrations typically consist of a diverse mix of primary and secondary components, the modeled attainment test based on RRFs decreases the possibility of choosing ineffective control strategies based on inaccurate model estimations of $PM_{2.5}$ composition, by assessing component specific RRFs. Fourth, and finally, there is evidence, based on a retrospective modeling analysis of ozone changes between 2002 and 2005 (Foley et al, 2014) and 2002 and 2010 (Hogrefe et al, 2014), that using the model in a relative sense provides better estimates of future ozone design values than using the absolute future year simulation. The correlation between model-projected and actual design values, as well as the accuracy of model projections of attaining a 75 ppb NAAQS, appear to be slightly improved when a relative modeled attainment test is utilized.

Equation (4.1) describes the recommended modeled attainment test in its simplest form, which directly applies for ozone, as applied for monitoring site *i*:

$$(\mathrm{DVF})_i = (\qquad)\ x\,(\qquad) \qquad [4.1]$$

where DVF_i is the estimated design value for the future year in which attainment is required at monitoring site *i*; RRF_i is the relative response factor at monitoring site *i*; and DVB_i is the *base design value* monitored at site *i*. Each of these terms is discussed in more detail below. For $PM_{2.5}$ and visibility, the modeled attainment test is more complicated in that each PM component (defined in a certain way) has its own RRF.

The EPA has developed the Software for Modeled Attainment Test-Community Edition (SMAT-CE) tool to enable completion of the modeled attainment tests for $PM_{2.5}$ and ozone, as well as for calculating changes in visibility in Class I areas. SMAT replaced the previous EPA attainment test software called MATS (Abt, 2014). SMAT can be found on the EPA's Support Center for Regulatory Atmospheric Modeling (SCRAM) website.

ED_002666_00014243-00102

4.1.1 Establishing the Base Design Value

The base design value for each monitoring site ( ) is the anchor point for estimating future year projected concentrations. Because the modeling is being used in a relative sense to determine how the modeled emissions changes will affect air quality design values in an area, it is important to match the base design value as closely as possible to the base year for which future/base ratios will be assessed. That is, if one is estimating the expected air quality change from emissions reductions between 2016 and 2030, it is important to establish an air quality anchor that represents 2016 emissions levels as closely as possible. However, it is well-established that inter-annual variability in meteorological conditions often leads to year to year differences in design values, even with static emissions levels (U.S. EPA, 2012). In addition, there is also year to year variability in emissions due to economic factors (such as recessions) and compliance with regulations. We, therefore, recommend using the base design value approach described below.

For the modeled attainment tests for ozone and $PM_{2.5}$, we recommend using the average of the three design value periods, which include the base emissions inventory year. This average is expected to best represent the air quality resulting from base year emissions with consideration of meteorological and emissions variability. For example, if the modeled attainment test uses base year emissions from 2014, then the base design value (DVB) would be calculated as shown in equation 4.2 for each site *i*:

$$(\quad) = \frac{((\quad 2014) + (\quad 2015) + (\quad 2016))}{i} \qquad [4.2]^{39}$$

This has the desired effect of weighting the projected ozone or $PM_{2.5}$ base design values towards the middle year of a five-year period (2012-2016); in this example, for a 2014 base emissions inventory year. Additionally, an average of three design values will be more stable (less year to year variability) than any single design value period. An EPA analysis of design values data at 761 ozone monitors over the period between 2002 and 2011 concluded that the median standard deviation of individual 3-year ozone design values was 5.3 ppb, whereas the standard deviation of the average of multiple 3-year ozone design values was only 4.0 ppb.

In cases in which there are less than 5 years of valid monitoring data at a site, we recommend that base design values be calculated only when there are at least 3 years of consecutive valid data (i.e., at least one complete design value). If a location has less than 3 consecutive years of

[39] For nomenclature purposes, the 3-year design value is referred to by the last year of the averaging period. That is, DV 2014 represents air quality data averaged over 2012, 2013, and 2014 for ozone and $PM_{2.5}$.

ED_002666_00014243-00103

valid data, then that site should not ordinarily be used in the modeled attainment test.[40] An example calculation for DVB is shown in Table 4.1 and includes the number of significant digits to include for ozone and each form of the $PM_{2.5}$ standard.

Table 4.1: Example illustrating the calculation of base design values.

	2014 DV	2015 DV	2016 DV	DVB
O3 NAAQS (ppb)[41]	76	76	78	76.7
Annual $PM_{2.5}$ NAAQS (µg/m^3)	12.9	13.1	12.6	12.87
24-hour $PM_{2.5}$ NAAQS (µg/m^3)	32	33	32	32.3

The measured 3-year average design values use the rounding/truncation rules established in 40 CFR part 50 Appendix U (8-hour ozone) and Appendix N ($PM_{2.5}$). The resultant 5-year weighted average DVB should carry one digit to the right of the decimal point for ozone and 24-hour $PM_{2.5}$ and two digits to the right of the decimal for annual $PM_{2.5}$.

The EPA also recognizes that the ambient data record may be modified under certain circumstances. For example, air agencies can request that the EPA agree to exclude event-influenced data by demonstrating that these data have been influenced by "exceptional" events and otherwise meet the criteria in the Exceptional Events Rule (81 FR 68216, October 3, 2016). Once an air agency submits an exceptional events demonstration and the EPA concurs with an air agency's request, the event-influenced data are officially noted and removed from AQS user reports (unless the AQS user specifically indicates that they should be included) and not used to calculate design values. In other cases, data may not qualify for exclusion under the Exceptional Events Rule but Appendix W to Part 51 may provide guidance on considering data which may be influenced by events that are not typical or expected to recur in the future (e.g., wildfires, construction, roadway repairs, or unusual agricultural activities) for certain modeling analyses. Potential event-influenced data may affect future year projections that are part of the modeled attainment demonstration. If potential event-influenced data from the historical

[40] Monitoring sites with less than 3 years of valid data cannot ordinarily be used to determine attainment status. Therefore, in most cases, those sites should not be used in the attainment demonstration or for regulatory analyses. However, there may be cases where a monitoring site with incomplete data indicates an area of high ozone or PM concentration. Further examination of the monitoring data and modeled response (RRF) in the area is recommended as part of either the monitor-based attainment test or an unmonitored area analysis, as appropriate.

[41] The current ozone NAAQS is defined in units of ppm. However, for ease of use, this document will use ppb units.

ED_002666_00014243-00104

record may affect regulatory outcomes, we encourage air agencies to consult with their EPA Regional office to determine how best to handle this situation.

In practice, the choice of the base design value can be critical to the determination of the estimated future year design values (Cohan, 2006) and careful consideration should be given to the selection of the years used in the calculation of base design values. There is no single methodology that can derive a "correct" base design value, but the 5-year weighted average value establishes a relatively stable value that is weighted towards the emissions and meteorological modeling year. In most cases, this provides an appropriate anchor point on which to base the future year design value calculations. Alternate, equally plausible, calculations of base design values may be considered with appropriate justification. Alternate base design values may also be considered as part of the corroborating analyses that comprise the aggregate weight of evidence determination. For instance, one may want to consider establishing the base design value on the highest of the relevant design values to ensure that the emissions measures will yield attainment even in periods with meteorological conditions especially conducive to poor air quality. In addition, in some cases, large year to year emissions changes or highly unusual meteorological conditions may make certain periods/years unsuitable for consideration in the base year period. Consideration of adjustments to the base design value(s) can be made, with appropriate justification, on a case-by-case basis and in consultation with the relevant EPA Regional office.

4.1.2 Calculation of Relative Response Factors

The relative response factor for each monitoring site ( ) is the fractional change in air quality concentrations that is simulated due to emissions changes between a base and a future year emissions scenario. The specific RRF calculation techniques vary by pollutant and are described in more detail in the remainder of section 4 below. The most important consideration associated with the RRF calculation is determining which model values are most representative of the expected air quality change for a given location. This requires consideration of all of the varying changes in hourly pollutant concentrations between the base and future cases and a determination of the most appropriate summary (average) value to apply to the base design value. As with the selection of a base design value, there may be plausible alternative means of calculating the RRFs that can differ from the approaches recommended below (e.g., Kulkarni et al., 2014, Porter et al., 2014). Where justified, alternate RRF calculation techniques can be included as corroborative analyses as part of a weight of evidence determination.

4.2 Modeled Attainment Test for the Primary Ozone Standard

The 8-hour ozone design value is currently calculated as the 3-year average of the annual fourth highest daily maximum 8-hour average concentration for a specific monitor. The standard is

ED_002666_00014243-00105

considered to be attained if the observed design value does not exceed the level of the ozone NAAQS[42]. Similarly, the modeled attainment test is considered satisfied when the $(\mathrm{DVF})_i$ is less than or equal to the current NAAQS level. The recommended test is described below.

4.2.1 Model Values to Use in the RRF Calculation

Given the formulation of the modeled attainment test, it is important to identify which model days are best suited for determining the expected air quality response resulting from a set of emissions changes. On any given modeled day, meteorological conditions may or may not be conducive to high concentrations at a monitor. If ozone predicted near a monitor on any particular day is much less than the design value, the model predictions for that day could be unresponsive to controls (e.g., the location could be upwind from most of the emissions in the nonattainment area on that day). If these days are included in the RRF calculation, they will likely lead to a higher RRF (closer to one) and a potential overestimation of the future design value.

Internal EPA analyses (U.S. EPA, 2014a) have shown that model response to decreasing emissions is generally most stable when the base ozone predictions are highest. The greater model response at higher concentrations is likely due to more "controllable" ozone at higher concentrations. In addition, the high model days are selected individually for each monitor. Meteorological conditions that lead to high ozone may differ between monitors in an area. For example, monitors that are typically "upwind" or "downwind" of an area may experience high ozone on different days, depending on the prevailing wind direction. In most urban areas, on days with high ozone concentrations, there is a relatively high percentage of locally generated ozone compared to days with low base case concentrations (U.S. EPA 2014b). Days with low ozone concentrations are more likely to have a high percentage of ozone due to background and boundary conditions. Since the form of the standard is also focused on the highest days of an ozone season (i.e., the fourth highest MDA8), the RRF calculation should also focus on days when the model predicts the highest ozone concentrations.

As the ozone NAAQS has been lowered, a key question is whether the attainment test should use an average RRF based on all modeled days that are greater than the NAAQS, as in the previous guidance approach. In some areas of the country, this could lead to RRF calculations that might include upwards of 100 days per ozone season. In these cases, including too many

[42] At the time of publication of this guidance, the ozone NAAQS is 0.070 ppm per 40 CFR 50.19. Design values for the 0.070 ppm NAAQS are truncated to the 3rd decimal digit. Therefore, 0.0709 ppm (70.9 ppb) is considered attainment and 0.0710 (71.0 ppb) is considered nonattainment. The same rounding/truncation procedures should be applied in the modeled attainment test. A detailed description of the methodology to calculate ambient ozone design values is provided in Appendix U to Part 50.

ED_002666_00014243-00106

days in the RRF calculation may lead to higher (closer to 1) RRFs. Using a reasonable number of the highest modeled days at each monitor is most likely to represent the response of the observed design value at a monitor. Since the design value is based on the seasonal 4th high observed values, we recommend selecting a set of modeled days that are likely to encompass a range of values that are somewhat higher than and somewhat lower than the 4th high value. We, therefore, recommend calculating the RRF based on the highest 10 modeled days in the simulated period (at each monitoring site). We believe this balances the desire to have enough days in the RRF to generate a robust calculation, but not so many days that the RRF does not represent days with concentrations near the observed design values. In addition, use of the highest 10 days (without an initial threshold) allows the attainment test to be easily adapted to any level of the NAAQS with little or no change in the methodology. In support of the revised recommendation, a recent assessment of modeled ozone changes between two ozone seasons (Foley et al., 2014) suggests that the use of the highest 10 days in the mean RRF calculation yields a slightly better estimate of the actual observed ozone change than the previous guidance approach.

The EPA recommends that the ozone RRF calculations be based on the 10 highest days in the base year modeling at the monitor location, as long as the base MDA8 value is greater than or equal to 60 ppb for that day. In cases for which the base model simulation does not have 10 days with MDA8 values greater than or equal to 60 ppb at a site, then EPA recommends using all days where MDA8 >= 60 ppb, as long as there are at least 5 days that meet the minimum threshold criteria. If there are less than 5 days with MDA8 >= 60 ppb, EPA recommends that air agencies do not calculate a DVF for the site.[43] If the modeled attainment test is based on less than 10 days, then other assessments in the aggregate weight of evidence demonstration will need to be more rigorous to counter the reduced robustness of the modeled attainment test.

In addition, as discussed by Vizuete et al. (2011), there can be days in which the modeled source-receptor relationships may not yield a representative response for a particular cell or array of grid cells. This can result from small inconsistencies between the model representation of transport patterns (or chemical formation) and what actually occurred on that day. In other situations, perhaps the high modeled ozone results from an atypical condition (e.g., in conjunction with a wildfire smoke plume) that is not representative of conditions that led to the area's design value. If there is compelling evidence to determine that a particular day, while being among the 10 highest MDA8 values at a location, is not representative of the expected source-receptor relationship at that location; then that day can be considered for removal from

[43] A scenario in which the 5th highest base model MDA8 values is less than 60 ppb either indicates an extremely low ozone site or a site with poor model performance. In either case, projections of DVF values are likely not reliable.

ED_002666_00014243-00107

the RRF calculation.[44] If a day or days are removed, we recommended adding the next-highest modeled day or days from outside the top 10 to maintain the use of 10 days in the RRF calculation. Air agencies should document the evidence that argues for the exclusion of any otherwise-eligible days. As noted earlier, poor model performance on individual days can lead to ratios of future to base ozone that may be biased on those days. For example, underprediction of ozone concentrations may lead to ratios that are unresponsive to emissions controls. Air agencies may want to examine the day- and site-specific model performance for days that are part of the RRF calculation to make sure that they can be appropriately used in calculating the expected response.

4.2.2 Grid Cells to Use in the RRF Calculation

It is recommended that ozone RRF calculations consider model response in grid cells immediately surrounding the monitoring site along with the grid cell in which the monitor is located. There are two primary reasons why we believe it is appropriate to include predictions from grid cells near a monitor rather than just the cell containing the monitor. First, limitations in the inputs and model physics can affect model precision at the grid cell level. Allowing some leeway in the precision of the predicted location of daily maximum ozone concentrations can help assure that possibly artificial, fine scale variations do not inadvertently impact an assessment of modeled ozone response. Second, some ozone monitors and important emission sources may be located very close to the border of a grid cell. Considering multiple cells near a monitor, rather than the single cell containing the monitor, diminishes the likelihood of inappropriate results which may occur from the artificial geometry of the superimposed grid system.

Based on the above considerations, it is recommended that the RRF be based on a 3 x 3 array of cells centered on the location of the grid cell containing the monitor. This is a change from the 2007 guidance (U.S. EPA, 2007), which pre-identified 15 km as an appropriate distance for grid cell proximity to a monitor and, therefore, recommended the selection of RRF grid cells to be contingent upon the chosen model grid resolution (i.e., a finer grid resolution considered a larger array of grid cells to maintain the pre-identified 15 km distance). The revised approach allows for small spatial differences in the response to be considered in the RRF calculation and presumes that the chosen model grid resolution is able to appropriately capture model response at the location of the monitoring site. Additionally, as described in section 4.2.3, the guidance now recommends that the grid cell with the highest base year MDA8 ozone value in the 3 x 3 array be used for both the base and future components of the RRF calculation, as opposed to the 2007 guidance approach, which recommended selecting the peak value in the

[44] Days that have been removed from the official record as exceptional events should be considered for removal from the RRF calculations.

ED_002666_00014243-00108

base and future cases even if they were in different grid cells. The rationale behind this change is that the goal of the modeling is to assess the ozone change at a specific location (i.e., a grid cell representative of a monitor location), independent of how the ozone peak may shift across an area. As the monitor location does not change between the base and future, it stands to reason that the future-year value should come from the same grid cell as the base case.

The selection of grid cells to use in the RRF calculation is especially important in "oxidant limited" areas (i.e., areas where incremental NOx decreases may lead to ozone increases) because these areas are likely to have grid cells that respond to VOC controls, with NOx sensitive grid cells nearby (perhaps even in the same 3 x 3 array). The 3 x 3 array methodology could lead to unrealistically small or large RRFs in these locations, depending on the specific case. Internal EPA analyses with a 12 km grid resolution have shown that on average, model-estimated future design values vary by less than 1 ppb when a 1x1 approach is used as opposed to the recommended array, although higher variability can occur (U.S. EPA, 2014a).

While the use of a 3 x 3 array of cells is the recommended approach for the modeled attainment test, air agencies are encouraged to consider the actual spatial scales of the monitors at which the modeled attainment test is being applied. Most ozone monitors are generally sited to be representative of an urban or neighborhood scale. In cases in which the spatial representativeness of a monitoring location is much smaller or larger than the area covered by the 3 x 3 array of cells, air agencies may consider assessing site-specific model response over an alternative grid cell array as part of corroborative analyses that inform the aggregate weight of evidence determination. Additionally, there may be cases where certain cells along the periphery of the 3 x 3 array have different modeled responses than what would be expected at the monitor location at the center of array due to a specific local topographic or geographical feature (e.g., a large water body or a significant elevation change). A potential example of this situation would be an array where several cells are over water and where the meteorological conditions and relevant emissions sources differ substantially from the land-based monitor location. Again, in these types of cases and with appropriate justification, air agencies could consider removing the unrepresentative cells from the calculation.

4.2.3 Sample Modeled Attainment Test Calculation

Once the appropriate RRF days are determined at each site, the calculation is straightforward. RRFs are not calculated on a daily basis, but instead are based on the mean 8-hour daily maximum values on all eligible calculation days. At the last step in the process, the final DVF should be calculated in accordance with the data handling conventions associated with the NAAQS (e.g., truncated to the third decimal in ppm units as per current Appendix U). An example is provided below for a 2016 base year and 2030 future year simulation based on an

ED_002666_00014243-00109

annual model simulation. This hypothetical modeled attainment test considers an area with three ozone monitors.

Step 1: Compute site-specific base year design values (DVBs) from observed data at all eligible sites in the area by calculating the average of the design value periods which include the base year inventory year. All averaged values should be rounded to the tenth of a ppb. See Table 4.2 for the calculation in this sample case. The base design value is the average of the 3-year design values for 2016, 2017, and 2018.

Table 4.2: Ambient ozone base design values in example calculation.

Monitoring site	2014 DV (ppb)	2015 DV (ppb)	2016 DV (ppb)	Base design value (ppb)
A	84	79	73	78.7
B	N/A	67	83	75.0
C	86	76	80	80.7

Step 2: Compute site-specific relative response factors (RRFs) from modeled data at all eligible sites (i.e., A, B, and C). This is done by first identifying the 10 days with the highest daily 8-hour maximum ozone concentrations greater than 60 ppb (or subset of days if there are 5-9 days greater than 60 ppb) in the base year. For each of these days, identify the grid cell with the highest base year MDA8 ozone in the 3 x 3 cell array surrounding the monitor cell location and use that value in the calculation of a multi-day average. 10-day averages (where possible) should be calculated for both the base and future simulations. The future case average should be based on MDA8 values paired to the same cell as used in the base average. Figure 4.1 below shows an abbreviated subset of the RRF calculation for 3 days at a single sample site. Model estimates of MDA8 (in ppb) should be calculated to at least four places to the right of the decimal, with the last digit truncated.

ED_002666_00014243-00110

Figure 4.1. Example demonstration of daily selections of 8-hour maxima to be used in RRF calculation.[45]

Base

6/21/2014		
60.24	68.15	66.20
73.41	**76.26**	71.27
67.62	74.02	73.24

6/26/2014		
76.20	75.35	77.41
78.51	76.79	75.92
76.66	75.95	77.21

7/17/2014		
50.22	58.55	56.22
53.47	56.22	51.28
57.76	**59.10**	53.29

Future

6/21/2014		
56.89	59.22	58.05
63.69	**68.46**	63.39
61.86	67.16	66.80

Day 1 (6/21) is straightforward. The daily maximum 8-hour average is > 60 ppb and is in the same cell in both the base and future case.

6/26/2014		
<u>69.84</u>	67.71	68.14
68.30	66.27	67.95
64.28	65.15	66.58

For day 2 (6/26), the peak location shifts from the base to another cell in the array in the future (underlined cell). However, the center left cell is used for base and future values (paired in space).

7/17/2014		
46.80	49.27	48.55
51.91	50.41	50.34
53.62	**54.10**	50.62

Day 3 (7/17) would not be eligible for the RRF calculation because there are no base values > 60 ppb.

Step 3: Once the MDA8 values are selected for each appropriate day for each monitoring site, they should be averaged together over the days while maintaining at least four places to the right of the decimal. The RRF is the ratio of the average future MDA8 values to the average base MDA8 values. Table 4.3 shows an example RRF calculation, using the results from days 1 and 2 from Figure 4.1, plus the other eight (top 10) days in which the daily model predictions were deemed appropriate for RRF calculation. RRF values should be estimated to four decimal places (rounded to the 4th decimal digit).

[45] Figure 4.1 depicts the modeled ozone in each cell with only two digits to the right of the decimal to make the formatting simpler. The actual cell values used in the calculation should carry at least 4 digits to the right of the decimal.

ED_002666_00014243-00111

Table 4.3. Relative response factor (RRF) calculation in example calculation.

Date	Base (ppb)	Future (ppb)
6/21/2014	76.2632	68.4626
6/26/2014	78.5156	68.3089
6/27/2014	75.2235	65.4333
6/28/2014	70.0509	62.7618
7/11/2014	76.3311	68.1865
7/12/2014	81.6739	72.6237
7/13/2014	78.3132	64.5255
8/02/2014	73.6956	69.9670
8/03/2014	75.8379	65.0507
8/21/2014	71.4430	59.3647
Average	75.7348	66.4685
RRF_i	**0.8776**	

The last step of the process is to multiply all of the site-specific RRFs by the corresponding site-specific DVBs and compare the results against the current level of the ozone NAAQS. In the final step, as noted in footnote 34, the design values are truncated to the integer ppb unit. Table 4.4 shows those results, assuming that the sample RRF shown above is for site C in the example. The final DVF for this sample site would be 70 ppb or 0.070 ppm. This would be considered a passing modeled attainment test for an ozone NAAQS of 0.070 ppm.

Table 4.4. Calculation of DVF_i for the three sites in the sample exercise.

Monitoring site	Base design value (ppb)	RRF	Future design value, pre-truncation (ppb)	Final Future design value (ppb)
A	78.7	0.8933	70.3	70
B	75.0	0.9138	68.5	68
C	80.7	0.8776	70.8	70

Note: The "final" future year design value represents the truncated value to compare to the NAAQS. However, it is recommended to retain the tenths (of a ppb) digit in all documentation, so that small ozone changes can be gleaned from control strategy and sensitivity analyses.

4.3 Modeled Attainment Test for the Secondary Ozone Standard

At the time this guidance document was written, the secondary ozone NAAQS is the same as the primary NAAQS. However, in the past, EPA has considered setting a secondary ozone NAAQS that is based on a cumulative measure of ozone concentrations, as opposed to the 3-year average of the annual fourth-highest daily maxima. If EPA were to establish a distinct secondary standard, it is possible that air agencies would need to separately show that the planned emissions reductions would be sufficient to meet both the primary and secondary

ED_002666_00014243-00112

NAAQS. If a new secondary standard is subsequently developed, EPA will provide supplemental guidance on how to assess attainment of that new standard at the time of its promulgation.

4.4 What is the Modeled Attainment Tests for the Annual Average $PM_{2.5}$ NAAQS?

Because ambient $PM_{2.5}$ often consists of multiple PM species, the modeled attainment test for $PM_{2.5}$ utilizes both $PM_{2.5}$ and individual $PM_{2.5}$ component species. A separate RRF and a separate concentration representing the base period is calculated for each $PM_{2.5}$ species. In order to perform the recommended modeled attainment test, air agencies should divide observed mass concentrations of $PM_{2.5}$ into 8 components (plus passive mass):

- Sulfates
- Nitrates
- Ammonium
- Organic carbon
- Elemental carbon
- Particle bound water
- Salt
- "Other" primary inorganic particulate matter
- Passively collected mass (blank mass)

To apply the attainment test, air agencies must first have outputs from air quality model simulations for a base year and a future year emissions scenario. We recommend a modeled attainment test which has four basic steps, as described in this section.

Step 1. Compute *observed* quarterly mean $PM_{2.5}$ and quarterly mean composition for each monitor.

Derive base year period quarterly mean concentrations[46] for each of the major components of

[46]Concentrations should be calculated based on calendar quarters for two reasons. First, the NAAQS is calculated as an annual average of 3 calendar years, so it would be inconsistent to average the data by season (where the winter season would span two calendar years). Second, the monitored data used to calculate design values is averaged on a calendar quarter basis before calculating annual averages.

ED_002666_00014243-00113

$PM_{2.5}$.[47] This is done by multiplying the monitored quarterly mean concentration of FRM derived $PM_{2.5}$ (40 CFR part 53) by the monitored fractional composition of $PM_{2.5}$ species for each calendar quarter (e.g., 20% sulfate x 15.0 µg/m^3 $PM_{2.5}$ = 3.0 µg/m^3 sulfate).

Step 2. Using air quality model results, derive component-specific relative response factors (RRF) at each monitor for each calendar quarter.

Use the air quality model predictions to calculate the quarterly mean concentration for each $PM_{2.5}$ component for the base year and future year scenario.[48] Take the ratio of future to base year predictions for each component. The result is a component-specific relative response factor (RRF).

The RRF for component j at a site i is given by the following expression:

$$(RRF)_{ij} = [C_{ij,\ projected}]/[C_{ij,\ base\ year}] \qquad [4.3]$$

where $C_{ij,\ base\ year}$ is the quarterly mean concentration predicted at or near the monitoring site[49] with emissions characteristic of the period used to calculate the base design value for annual $PM_{2.5}$.

$C_{ij,\ projected}$ is the future year quarterly mean concentration predicted at or near the monitoring site.

(e.g., given model predicted mean third quarter base year period sulfate of 10.12 µg/m^3 and future year mean third quarter concentration of 8.23 µg/m^{3,} then the third quarter RRF for sulfate is 0.8132).

---

[47] The mean species concentrations should be representative of the base period. The default approach to calculate mean species concentrations in SMAT is to use the 3-year average, centered about the base modeling year. These values are used to calculate species fractions by quarter. This differs from the 5-year weighted average design values recommended for base year $PM_{2.5}$ values. However, it is expected that there will be less variation in the species fractions compared to the absolute measured $PM_{2.5}$ concentrations. Alternative species fractions calculations can be considered, as long as they are representative of the expected species fractions during base year period.

[48] The calculations assume that there is one full year of model outputs.

[49] For greater (coarser) than 12km grid resolution, use the single grid cell. For 12km or less grid resolution, use the mean of the 3 x 3 array of the grid cells surrounding the monitor.

ED_002666_00014243-00114

Step 3. Apply the component specific RRFs to observed air quality to obtain projected quarterly species estimates.

For each quarter, multiply the base year period quarterly mean component concentration (step 1) times the component-specific RRF obtained in step 2. This leads to an estimated future quarterly mean concentration for each component (e.g., 3.0 $\mu g/m^3$ sulfate x 0.8132 = future sulfate of 2.44 $\mu g/m^3$).

Step 4. Calculate a future year annual average $PM_{2.5}$ estimate.

Sum the quarterly mean $PM_{2.5}$ components to get quarterly mean $PM_{2.5}$ values. Then average the quarterly mean $PM_{2.5}$ concentrations to get a future year annual average $PM_{2.5}$ estimate for each FRM site.

Compare the projected average annual arithmetic mean $PM_{2.5}$ concentration obtained in Step 4 with 12.0 $\mu g/m^3$. If the values at all FRM sites are $\leq$ 12.0 $\mu g/m^3$, the test is passed.

4.4.1 Ambient $PM_{2.5}$ Data Used in the Annual $PM_{2.5}$ Attainment Test

$PM_{2.5}$ data collected at FRM and FEM sites are used for nonattainment designations. Therefore, we recommend using FRM and FEM data as the base data for projecting future $PM_{2.5}$ concentrations. As can be seen from the list of steps, the modeled attainment test is critically dependent on the availability of species component mass at FRM sites. This raises several issues. First, the majority of FRM sites do not have co-located chemical speciation network (CSN[50]) samplers.[51] And second, the FRM filter measurements and $PM_{2.5}$ speciation measurements do not always measure the same mass (Frank, 2006). There are numerous known issues with positive and negative sampling artifacts. These issues are addressed below.

4.4.2 FRM Monitors that Do Not Have Speciation Data

Species concentration data and/or species fractions are needed in order to apply the $PM_{2.5}$ attainment test. There are approximately 1200 FRM and FEM $PM_{2.5}$ monitoring sites, but only ~200 urban speciation monitoring sites. This makes it difficult to apply the attainment test at

[50] References to CSN monitors in this document refers to the overall speciation network which includes both trends sites and SLAMS sites.
[51] There are ~1200 FRM measurement sites and ~200-250 urban speciation sites (trends and SLAMS).

ED_002666_00014243-00115

the majority of FRM/FEM sites. There are several possible ways to estimate species concentrations at FRM monitors that lack speciation data. Among them are:

1) Use concurrent data from a nearby speciation monitor to estimate species concentrations and/or fractions at one or more FRM sites.
2) Use representative speciation data (from a different time period) collected in an area to estimates species data at FRM sites.
3) Use interpolation techniques to create spatial fields using ambient speciation data.
4) Use interpolation techniques to create spatial fields using ambient speciation data **and** gridded model outputs to adjust the species concentrations based on the modeled spatial gradients.

In general, we recommend using spatial interpolation techniques to estimate species concentrations at FRM sites that do not have speciation data (numbers 3 and 4 above). But in some cases, interpolating data from nearby sites may not be appropriate, may not be feasible, or simply may not be necessary. In the EPA's SMAT, a relatively simple interpolation technique is used to estimate species concentrations at all FRM sites in the country. The analysis uses a Voronoi Neighbor Averaging (VNA) technique (consistent with approach number 3 above). Other interpolations have been done using Kriging and other more complicated methodologies such as "Downscaler" (Berrocal, 2010a and 2010b). Air agencies are encouraged to explore techniques that are most appropriate for their area and situation. EPA's SMAT software is available for air agencies to use as a default technique.

For areas that contain one or more speciation sites, and where FRM sites exhibit very little spatial gradients, it may be appropriate to assume that the speciation site(s) is/are representative of the entire area. For areas that exhibit strong spatial gradients in $PM_{2.5}$ concentrations, air agencies should strongly consider using more sophisticated techniques to estimate species concentrations. Combining ambient data with model output concentrations (number 4 above) may help adjust concentrations in areas with strong gradients, but limited speciation data. This technique has been used extensively to create spatial fields of $PM_{2.5}$ concentrations for the purpose of generating health benefits calculations and other examinations of unmonitored areas (also see section 4.6), but it has generally not been used to generate species fractions at FRM sites. The technique is further limited by uncertainties in the representativeness of the model outputs and emissions inventories.

ED_002666_00014243-00116

4.4.3 $PM_{2.5}$ Species Calculations and Definitions

Data analyses (Frank, 2006) have noted that the FRM monitors do not measure the same components and do not retain all of the $PM_{2.5}$ that is measured by routine speciation samplers and, therefore, cannot be directly compared to speciation measurements from the Chemical Speciation Network (CSN).[52] The FRM mass measurement does not retain all ammonium nitrate and other semi-volatile materials (negative sampling artifacts) and includes particle bound water associated with sulfates, nitrates and other hygroscopic species (positive sampling artifacts). This results in FRM measured concentrations (and percent contributions to $PM_{2.5}$ mass), which may be different than the ambient levels of some measured $PM_{2.5}$ chemical constituents.

Because the FRM data are used to determine compliance with the NAAQS, it is critical to estimate the species composition as measured on the FRM filters. In addition, for the purposes of predicting changes in $PM_{2.5}$ components, constructed $PM_{2.5}$ mass should match the composition of mass retained by the FRM.

The goal is to reconstruct the measured species components so that they add up to the measured FRM mass. This concept can generally be represented by equation 4.4:

$RCFM_{FRM}$ = [Ammoniated Sulfate Mass] + [Retained Ammoniated Nitrate Mass] + [Retained Carbonaceous Mass] + [Other Primary $PM_{2.5}$] + [Other Components] [4.4]

In the above characterization, $RCFM_{FRM}$, or reconstructed fine mass, and all of the listed chemical components reflect those retained during sampling and equilibration on the FRM's Teflon filter. Sulfate and nitrate mass include associated ammonium, which may be different than assumed ammonium sulfate and ammonium nitrate compounds. Carbonaceous mass includes elemental carbon (EC), organic carbon (OC), and hundreds of different carbon compounds, which are not directly measured, and which may not match the carbon mass in speciated measurement data. Also, sulfates and nitrates include particle bound water associated with these hygroscopic aerosols. In this characterization, other primary $PM_{2.5}$ mass is intended to be a more general term that includes fine soil, and oxides that result from other PM emissions.

[52] The information in this section applies to the most common samplers in the CSN. The characteristics of the sampler and the analytical procedures used to produce chemical speciation data should be considered in determining which, if any, adjustments are appropriate to make the data useful for comparison to FRM and FEM data.

ED_002666_00014243-00117

4.4.4 Recommended Treatment of Species Data

We recommend an adjustment technique based on an FRM mass construction methodology, which results in reduced nitrates (relative to the amount measured by routine speciation networks), higher mass associated with sulfates and nitrates (reflecting water included in gravimetric FRM measurements), and a measure of organic carbonaceous mass, which is derived from the difference between measured $PM_{2.5}$ and its non-organic carbon components (Frank, 2006). This characterization of $PM_{2.5}$ mass also reflects "other" primary inorganic $PM_{2.5}$ and other minor constituents. Frank (2006) terms this approach "sulfate, adjusted nitrate, derived water, inferred carbonaceous material balance approach (SANDWICH)." The resulting characterization provides a complete mass balance. It does not have any unknown mass, which is sometimes presented as the difference between measured $PM_{2.5}$ mass and the characterized chemical components derived from routine speciation measurements. The recommended SANDWICH-based characterizations should yield more accurate assessments of future $PM_{2.5}$ concentrations as measured by FRM monitors, compared to using unadjusted CSN data.

4.4.4.1 Retained Nitrate Mass

The first step in the procedure for identifying FRM mass components is to estimate the retained nitrate mass on the FRM filters. The FRM does not capture all of the semi-volatile components of the ambient air, such as ammonium nitrate. The retained amount of nitrate ion, however, can be estimated by a simple thermodynamic model that involves 24-hr ambient nitrate speciation concentrations (as measured by a standard speciation sampler using a nylon filter preceded by a HNO_3 denuder) together with hourly ambient temperature and humidity. Atmospheric nitrates are highest during the cooler months. Retention on the FRM is also higher during the cooler months and essentially all the nitrates are lost during the summer. The retention does not appear to depend on ambient NH_3 or HNO_3 concentrations. More NO_3 is retained at low temperatures and high humidity, which varies by sampling location and time of year.

Because nitrate retention varies by site and season, a simple ammonium nitrate equilibrium model can be used to predict the amount of nitrates retained on the FRM Teflon filter. As used by Hering (Hering, 1999; Zhang, 1992),

$$\text{delta } NO_3\ (\mu g/m3) = 745.7/T_R * 1/24 * \sum_{i=1}^{24} (\mathbf{K}_i^{1/2}) \qquad [4.5]$$

where delta NO_3 is the amount of volatilized nitrate mass, T_R is the reference temperature for the sampled air volume in degrees Kelvin, and $\mathbf{K}_i$ is the dissociation constant for ammonium nitrate evaluated at the ambient temperature and relative humidity for hour i. The nitrate loss can be predicted for each day, based on hourly temperature and relative humidity data. It must be subtracted from the NO_3 measured at speciation sites to estimate the FRM retained nitrate. The SMAT software is populated with a speciation ambient data file with daily average retained nitrate pre-calculated for all CSN and IMPROVE sites. The analysis uses National Weather Service temperature and relative humidity data from the closest station. Other sources of meteorological data may also be appropriate. Further details on the nitrate loss calculations can be found in Frank (2006).

4.4.4.2 Ammonium Associated with Sulfates and Retained Nitrates

There are several ways to estimate ammonium mass for use in attainment test. The most direct way is to use measured ammonium concentrations from the CSN network (IMPROVE does not measure ammonium ion). A second, more indirect method is to estimate the ammonium associated with nitrate, and to calculate the ammonium associated with sulfate and the degree of neutralization of sulfate (DON), and then use the resulting information to calculate ammonium mass. Due to uncertainties associated with the ammonium measurements and the lack of ammonium measurements in rural areas, this indirect method is recommended as the default attainment test methodology.

To determine the mass associated with nitrates, we assume retained nitrate is all ammonium nitrate. Thus, using the ratio of the molecular weight of ammonium to that of nitrate, the ammonium associated with nitrates ($NH4_{NO3}$) can be derived directly from the estimated retained nitrate ($NO3_{FRM}$) as:

$$NH4_{NO3} = 0.29 * NO3_{FRM} \qquad [4.6]$$

The difference between total FRM NH_4 (amount associated with nitrates and sulfates), termed $NH4_{FRM}$, and the measured NH_4, termed NH_{4CSN}, is needed to determine the ammoniated form of sulfates as described by equation 4.4. A measurement study by Collett (Collett, 2004) shows that NH_4 may not be completely retained during collection on nylon filters preceded by a nitric acid denuder, such as are used at CSN sites. At several FRM study sites, the CSN NH_4 which is adjusted for evaporated NH_4NO_3 tends to more closely correspond to the measured NH_4 from the FRM Teflon filter. However, for other sites, the measured CSN NH_4 appears to agree with FRM NH_4.

ED_002666_00014243-00119

The available information suggests that using measured ammonium (assuming none is volatilized) may be the best assumption. We, therefore, recommend using unadjusted ammonium, but further analysis of this issue is warranted.

To calculate the ammonium associated with sulfate, we recommend using the degree of neutralization of sulfate (DON). This is defined as the ammonium associated with sulfate (i.e., total ammonium less the ammonium associated with nitrate), divided by sulfate. This reflects the degree to which sulfate has been neutralized by ammonium. The ambient data is such that nearly all of the ammonium data is from urban sites (CSN), but the sulfate and nitrate data is from both urban (CSN) and rural (IMPROVE) sites. This may lead to an overestimation of ammonium concentration in rural areas when ammonium is directly interpolated. Therefore, we recommend using calculated DON with SO_4 and $NO3_{FRM}$, and interpolating to each FRM site to get quarterly average concentrations. The interpolated DON and species concentrations are used to calculate $NH4_{FRM}$ using the following equation:

$$NH4_{FRM} = DON * SO4 + 0.29*NO3_{FRM} \quad [4.7]$$

The indirect calculation of ammonium mass from interpolated fields tends to smooth out the gradients in mass. This is deemed to be beneficial, due to the uncertainty in the measurements.

4.4.4.3 Particle Bound Water

Because ammoniated sulfate and ammonium nitrate are hygroscopic, the retained sulfate and nitrate mass will include water. Particle bound water (PBW) can be estimated using an equilibrium model. The EPA developed an approach using the Aerosol Inorganic Model (AIM) (Clegg, 1998). PBW is derived from quarterly average FRM concentrations of sulfate, ammonium, and nitrate (as described above). Estimated hydronium ion, H+, needed to achieve ionic balance is derived from the latter values. The model enables the distribution of water and ions to be calculated between liquid, solid and vapor phases for specific temperature and relative humidity conditions. Typical FRM filter equilibration conditions of 35% RH and 22 deg C (295 deg K) temperature are used to estimate water concentrations at those conditions.

Application of AIM at the specified FRM filter equilibration conditions show that PBW is much more dependent on sulfate concentration compared to nitrate and that the relationship varies somewhat by season. There is proportionally less estimated PBW water for wintertime aerosol which has higher NO_3 and lower SO_4. The PBW concentrations are also sensitive to the degree

of neutralization of the sulfate particles (determined by the relative concentration of $NH4_{FRM}$).

We recommend calculating PBW as a component of $PM_{2.5}$ mass. The AIM model (or other equilibrium models) can be used, or a regression equation can be developed to simplify the process. For computational convenience, EPA derived a polynomial regression equation fit to the calculated water mass from AIM and the three input values to AIM (sulfate, nitrate and ammonium). The polynomial equation is used in the SMAT software to calculate PBW. See the SMAT User's Guide for more details on the default PBW equation.

4.4.4.4 Salt

In the $PM_{2.5}$ attainment test methodology, salt is represented as a species which accounts primarily for sea salt at sites near the oceans. The salt species is estimated using the CSN measured Cl^+ ion concentration multiplied by 1.8. The multiplier accounts for the sodium contribution to the species. In areas distant from oceans, measured Cl^+ may represent PM mass that is not associated with salt water. For instance, there may be chloride emissions resulting from various industrial processes. Therefore, the use of measured salt in the attainment test calculations should be evaluated on an individual area/monitor basis.

Measured salt is included by default in the SMAT species ambient data file. The inclusion of modeled salt (and hence an RRF for salt) is optional and can be included if there are accurate modeled salt concentrations. If modeled salt is not included in the SMAT input files, an RRF of 1.0 used in the SMAT calculations.

4.4.4.5 Other Primary PM2.5

The terms "crustal," "fine soil," "major metal oxides," "inorganic particulates," and "other PM," are sometimes used interchangeably. For $PM_{2.5}$ design value calculations we will refer to this material as "other primary $PM_{2.5}$" (OPP). For regional haze calculations, we will continue to refer to this material as "fine soil."

For the purposes of estimating OPP for the $PM_{2.5}$ attainment test, all measured non-carbon mass that is not organic in nature (not associated with sulfate and/or nitrate) should be counted. As with the other $PM_{2.5}$ components measured on the FRM filter, there is uncertainty associated with this estimate. The "crustal" or "fine soil" definition from IMPROVE can be used

ED_002666_00014243-00121

to estimate OPP[53] or an alternative formula can be defined which better estimates the urban nature of the FRM measurements. The IMPROVE definition of "fine soil" accounts for the typical crustal components (and attached mass) that would be expected in remote Class I areas. Fine soil is represented as five elements (Al, Si, Fe, Ti, and Ca) with coefficients to represent various oxides and material, which may be attached to or associated with the major elements. In urban areas, inorganic PM, which is not elemental carbon or associated with sulfate or nitrate, may come from many sources such as re-suspended dust or industrial sources (stack or fugitives). It is generally "crustal" in nature (dominated by silicon), but urban PM is more likely to contain heavy metals and industrial components.

Although the composition of inorganic PM may differ between urban and rural (remote) areas, an alternative equation, similar to the IMPROVE fine soil equation, can be used to estimate OPP for the $PM_{2.5}$ attainment test. The alternative equation is suggested by (Frank, 2006) and uses only four elements. The equation removes aluminum (and accounts for associated mass by increasing the coefficient for Si), due to the fact that aluminum is often missing from the speciation measurements. This allows for more complete data.

The alternative "fine soil" equation is as follows:

$$\text{Other primary } PM_{2.5} \text{ mass} = 3.73 \times [Si] + 1.63 \times [Ca] + 2.42 \times [Fe] + 1.94 \times [Ti] \qquad [4.8]$$

4.4.4.6 Blank mass

The other quantifiable components of $PM_{2.5}$ mass include passively collected mass, represented by a field blank concentration (U.S. EPA, 2002). This appears to constitute a contamination of the filter resulting from handling or contact with the FRM cassette. This value is deemed to be an important constituent of $PM_{2.5}$ mass (it is assumed to not be dependent on pollutant emissions). Based on recently collected ambient QA data, we recommend using a default nominal blank mass value of 0.2 $\mu g/m^3$ (see the EPA QA website for more details). This value represents an approximate nationwide network median value. This value can be modified based on local FRM blank mass measurements. The blank mass is assumed to remain constant through time (RRF=1.0).

[53] IMPROVE estimates fine soil as: $2.2 \times [Al] + 2.49 \times [Si] + 1.63 \times [Ca] + 2.42 \times [Fe] + 1.94 \times [Ti]$

4.4.4.7 Calculation of Carbonaceous Mass

Organic carbon mass is typically estimated from blank corrected speciation data, where organic carbonaceous mass is first estimated by multiplying the organic carbon concentrations by 1.4 or alternative factors to account for the oxygen, hydrogen and other elements associated with ambient carbon particles.

There are many uncertainties in estimating carbonaceous mass from carbon measurements (Turpin, 2001; Chow, 2004). Uncertainties include differences in carbon measurement protocol between urban and rural monitoring locations; a relatively "bumpy" surface of urban carbon concentrations as derived from urban and rural organic carbon measurements; and lack of carbon measurements at all FRM locations. We, therefore, recommend an alternative approach to estimate the organic carbon contribution to $PM_{2.5}$ mass.

The recommended attainment test approach estimates organic carbon by *mass balance*. Precisely measured FRM $PM_{2.5}$ mass (U.S. EPA, 2003) is compared to the sum of its non-organic carbon components. The latter are sulfates, ammonium, nitrates, estimated particle bound water, elemental carbon, estimated other primary $PM_{2.5}$ material, salt, plus 0.2 µg/m3 blank mass as discussed earlier.

This approach estimates retained organic carbon FRM mass and explicitly accounts for the following important and difficult to estimate carbon mass properties: 1) regional and urban-rural differences in the mix of carbonaceous aerosols (i.e., the amount of oxygen, hydrogen, etc); 2) retained water associated with hygroscopic carbon compounds (Saxena, 1996; Yua, 2004*)*; 3) volatile carbonaceous material measured by speciation samplers, but not retained in FRM mass; and 4) uncertainties associated with blank corrections of measured organic carbon.

Organic Carbon Mass by mass balance (**OCM_{mb}**) is defined as,

$$\mathbf{OCM_{mb}} = PM_{2.5} - \{ [SO4] + [NO3_{FRM}] + [NH4_{FRM}] + [water] + [EC] + [OPP] + [salt] + [0.2] \} \quad [4.9]$$

In this expression, all of the above components represent the mass retained on FRM Teflon filters.

ED_002666_00014243-00123

This approach completely accounts for FRM mass[54] and OCMmb is often greater than the amount that would be derived directly from speciation measurements. Because of uncertainties in speciation measurements and their estimates from interpolated surfaces, setting a lower limit (floor) may be necessary so that the OCMmb is not unreasonably low. In some cases, the OCMmb could be lower than the measured OC or OCM. Since measured OC (without accounting for mass associated with OC) is a conservative minimum estimate of OC mass, we recommend setting the OCMmb "floor" to be equal to measured OC.

There may also be situations where an OCMmb "ceiling" is needed. In remote urban areas with relatively high FRM concentrations that may be surrounded by rural background concentrations, the OC by mass balance technique may apportion 95%+ of the $PM_{2.5}$ mass to OCM. If this is not a reasonable assumption, then a ceiling may be needed to cap the OCM as a percentage of $PM_{2.5}$ mass. Based on measured data (FRM sites with co-located speciation data), it appears that on a quarterly average basis, OCM is rarely more than 80% of total $PM_{2.5}$ mass. This may be a reasonable default ceiling, but a lower value (or in rare circumstances, a higher value) may be more appropriate in some regions of the country.

4.4.4.8 Summary of $PM_{2.5}$ Composition Calculations

The terms of equation 4.10 reflect the final estimated composition of the particles measured by the FRM (for each quarter). Quarterly average FRM mass is equal to the sum of the eight species plus blank mass.

$$PM_{2.5FRM} = \{ [OCMmb] + [EC] + [SO4] + [NO3_{FRM}] + [NH4_{FRM}] + [water] + [OPP] + [salt] + [0.2] \} \qquad [4.10]$$

The recommended order to generate the data is as follows:

1) Calculate adjusted nitrate using hourly meteorology and 24-hour average nitrate measurements.

[54] The OC by mass balance technique assumes that all other mass is accounted for and, therefore, all remaining mass is OCM. This may not always be a good assumption. The results of the technique should be carefully evaluated to ensure that OC mass is not overestimated (and, therefore, other mass components are underestimated). This may be an issue in areas that do not have nearby speciation data and have relatively large concentrations of primary $PM_{2.5}$. The OC by mass balance technique may inadvertently apportion mass to organic carbon, which may actually be EC or "other" primary $PM_{2.5}$ mass components (such as heavy metals). All available ambient data and modeling data should be used to evaluate the species apportionment results.

ED_002666_00014243-00124

2) Calculate quarterly averages for adjusted nitrate, sulfate, elemental carbon, ammonium (or degree of sulfate neutralization (DON)), OPP, salt, and measured OCM[55].

3) Quarterly average ammonium is calculated from the adjusted nitrate, sulfate, and DON values (if measured ammonium is not used directly).

4) Ammonium, sulfate, and nitrate concentrations are input into the polynomial water equation to derive particle bound water concentrations.

5) Carbon mass by difference (OMCmb) is calculated from the $PM_{2.5}$ mass, adjusted nitrate, ammonium, sulfate, water, elemental carbon, salt, other primary $PM_{2.5}$, and blank mass values. The sum of the eight species plus blank mass is equal to the FRM mass.

We illustrate application of the recommended test in example 4.4.1.

Example 4.4.1

Given: (1) Area "C" has 2 monitoring sites.

(2) Monitored FRM air quality data show the following average quarterly mean $PM_{2.5}$ concentrations based on a 5-year weighted average of observations from 2012-2016 at each site (values are in $\mu g/m^3$).

Table 4.4.1 Quarterly average $PM_{2.5}$ FRM mass concentration ($\mu g/m^3$) at 2 example sites

Site	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	17.21	16.34	20.00	14.76
2	15.39	17.98	18.23	13.76

(3) The area has 1 urban speciation site, which is co-located with FRM site 2. The speciation data is available for the entire time period, but the species concentrations for the 2013-2015 time period have been deemed to be most representative of the 2012-2016 5- year weighted average. Therefore, the species data for the average 3-year period (2013-2015) has been

[55]The measured OCM is only used to calculate the "floor" for OCMmb.

ED_002666_00014243-00125

interpolated to derive estimated species concentrations at each site. The species concentrations are matched up with FRM data for the same 2013-2015 period to derive representative species fractions. The average monitored air quality for the 3rd quarter of 2013-2015 at site 1 is as follows[56]:

Table 4.4.2 Average monitored 3rd quarter ambient data at site 1

FRM Mass (µg/m3)	Blank Mass (µg/m3)	Non-blank Mass (µg/m3)	Sulfate (µg/m3)	Nitrate (µg/m3)	Organic Carbon Mass (µg/m3)	Elemental Carbon (µg/m3)	Water (µg/m3)	Ammonium (µg/m3)	Salt (µg/m3)	OPP (µg/m3)
22.62	0 .2	22.42	8.51	1.11	5.21	0.91	2.31	3.31	0.35	0.71

The blank mass is subtracted before species fractions are calculated because the blank mass is held constant at 0.2 µg/m3 throughout the analysis. In the example, the measured FRM mass for quarter 3 is 22.62 µg/m3. The non-blank FRM mass is 22.42 µg/m3. The mass of the eight species add up to the non-blank mass.

(4) Species fractions are calculated for each quarter for each species. In the example below, a fraction of non-blank mass is calculated for each of the eight species. Blank mass remains fixed at 0.2 µg/m3.

Table 4.4.3 3rd quarter average species fractions for 2013-2015 at site 1

FRM Mass (µg/m3)	Blank Mass (µg/m3)	Non-blank Mass (µg/m3)	% Sulfate	% Nitrate	% Organic aerosol	% Elemental Carbon	% Water	% Ammonium	% Salt	% OPP
22.62	0 .2	22.42	37.96	4.95	23.24	4.06	10.30	14.76	1.56	3.17

The percentages in table 4.4.3 above are the relative composition for the 3rd quarter of 2013-2015. It is assumed that these species fractions are representative of the 2012-2016 time-period.[57]

[56] The species concentrations can be derived from co-located speciation data, nearby speciation data or from interpolated spatial fields. FRM mass is not interpolated.

[57] The default assumption is that the average of the middle 3 years of speciation data is representative of the 5-year weighted average FRM data. However, other assumptions (number of years of speciation

ED_002666_00014243-00126

(5) The weighted quarterly average FRM design values are used as the base year FRM value for each monitoring site (2012-2016). The species fractions from the 2013-2015 speciation data were used to estimate the species concentrations for the base year FRM $PM_{2.5}$ data. The percentage compositions for 2013-2015 are applied to the quarterly weighted average design values as shown in table 4.4.4. In the example below, the weighted average design value for the 3rd quarter for the site from table 4.4.1 is 20.00 µg/m3. This leads to the following concentrations of $PM_{2.5}$ species:

Table 4.4.4 Calculation of the "base case" species concentrations at site 1

Weighted Avg. FRM Mass (µg/m3)	Blank Mass (µg/m3)	Non-blank Mass (µg/m3)	Sulfate (µg/m3)	Nitrate (µg/m3)	Organic aerosol (µg/m3)	Elemental Carbon (µg/m3)	Water (µg/m3)	Ammonium (µg/m3)	Salt	OPP (µg/m3)
20.00	0.2	19.80	7.52	0.98	4.60	0.80	2.04	2.92	0.31	0.63

This procedure is repeated for each $PM_{2.5}$ site and quarter to complete the calculation of base year ambient concentrations used as the basis for future estimates of $PM_{2.5}$ mass and its components.

data) can be used. It is generally assumed that the inter-annual variability of the species fractions is small compared to the variability of species concentrations.

ED_002666_00014243-00127

(6) Modeled results show the following relative response factors (RRF) for predicted mass of 5 components of $PM_{2.5}$ for the 3rd quarter[58]:

Table 4.4.5 3rd quarter modeled RRFs for site 1

RRF Sulfate	RRF Nitrate	RRF Organic aerosol	RRF Elemental Carbon	RRF Salt	RRF OPP
0.8761	0.9432	0.9713	0.9324	1.0	1.0423

(7) The quarterly mean RRFs from table 4.4.5 are multiplied by the weighted quarterly average species concentrations from table 4.4.4 to derive future year concentrations.

From the example above, the future year 3rd quarter concentrations are:

$Sulfate_{Future}$ = 7.52 * 0.8761 = 6.58 µg/m3
$Nitrate_{Future}$= 0.98 * 0.9432 = 0.92 µg/m3
Organic carbon $mass_{Future}$= 4.60 * 0.9713 = 4.47 µg/m3
Elemental $Carbon_{Future}$= 0.80 * 0.9324 = 0.75 µg/m3
$Salt_{Future}$=0.31*1.0 = 0.31 µg/m3
OPP_{Future}= 0.63 * 1.0423 = 0.65 µg/m3

(8) The future year concentrations derived above are used to calculate the future year concentration of ammonium (if the direct ammonium RRF is not used) and particle bound water.

The future year ammonium concentrations are calculated from the sulfate, nitrate, and (base year) DON values. Assuming that the DON is unchanged from the base year[59], the ammonium is

[58] The default assumption for the annual average $PM_{2.5}$ NAAQS test is to calculate RRFs from the mean of the 3 X 3 matrix of grid cells surrounding each FRM monitoring site.
[59] Due to the uncertainty in the ammonium measurements, by default, the DON is assumed to stay constant through time. The water calculation is sensitive to the ammonium (and, therefore, the DON value) concentrations. Holding the DON constant allows for the future year ammonium and water values to be solely a function of the change in sulfate and nitrate concentrations. Otherwise, the water concentration can go up when the sulfate and nitrate concentrations go down (and vice versa). This may occur if sulfate becomes more neutralized in the future. It is a somewhat illogical outcome (although scientifically possible) and is highly dependent on an uncertain measurement (ammonium). Therefore,

calculated using the following formula:

$NH4_{future}$ = DON * $SO4_{future}$ + 0.29*$NO3_{future}$, [4.7]

In the example above, assuming the base year DON is 0.336,

$Ammonium_{Future}$= 0.336 * 6.58 + 0.29 * 0.92 = 2.48 µg/m3

The $NH4_{future}$, $SO4_{future}$, and $NO3_{future}$ concentrations can then be input into an equilibrium model (AIM or another alternative model) or through a polynomial regression equation to predict future year particle bound water concentration (based on the default equation, the future year water concentration in this case is 1.73 µg/m3).

(9) The future species concentrations at each FRM site are then summed over the eight species plus blank mass to estimate the future quarterly average $PM_{2.5}$ concentration.

In the example above, the total $PM_{2.5Future}$=
6.58 + 0.92 + 4.47 + 0.75 + 0.65 + 2.48 + 1.73 + 0.2 = 17.78 µg/m3

(10) The same calculations are completed for the other 3 quarters to get a future year $PM_{2.5}$ concentration for each quarter. The 4 quarters are then averaged to get a final future year annual average $PM_{2.5}$ concentration for each FRM site.

(11) The future year annual average concentration is compared to 12.0 µg/m^3.[60]

4.5 What Is the Recommended Modeled Attainment Test for the 24-Hour NAAQS?

Our recommended modeled attainment test for the 24-hour NAAQS for $PM_{2.5}$ is similar to the previously described test for the annual NAAQS in that it uses model predictions in a relative

---

use of a constant DON creates a more stable set of calculations. If the measured and modeled ammonium concentrations are believed to be accurate and respond in a reasonable way to emissions controls, then it would be more scientifically credible to use the model predicted change in ammonium. Otherwise, it is a reasonable assumption to keep the DON constant over time.

[60] In the final step, the future year concentration should be rounded to the tenths digit. A (rounded) value of 12.0 µg/m3 meets the NAAQS. A value of 12.1 µg/m3 or greater violates the NAAQS.

ED_002666_00014243-00129

sense to reduce site-specific *observations* (averaged over 5 years). In the test, we are interested in reducing the base year design values at each site to $\leq$ 35 µg/m^3 (the 2006 24-hour $PM_{2.5}$ NAAQS). [61]

Ideally, the modeled attainment test should reflect model results obtained for days in each season having observed $PM_{2.5}$ concentrations above the design value. Even though the 24-hour NAAQS is based on the single 98th percentile value of all days in the year, it is important to perform the test on a seasonal basis, since $PM_{2.5}$ consists of a mixture of pollutants whose composition can vary substantially from season to season. There could be a substantial amount of uncertainty associated with predictions on any single day. Thus, our test is most likely to be reliable when relative response factors (RRFs) reflect composite responses from multiple days. Therefore, we recommend modeling as many days as feasible where observed $PM_{2.5}$ is greater than 35 µg/m^3. Alternatively, the test can focus on the high end of the distribution of days in each quarter[62] (e.g., the top 10% of $PM_{2.5}$ days[63]) assuming that the high days are representative of days that violate the NAAQS. As with the annual NAAQS (and for the same reasons), the preferred approach is to develop RRFs, which are quarter specific[64].

The 24-hour $PM_{2.5}$ attainment test should be based on the same 5-year weighted average methodology that was used for the annual standard, with some slight modifications. The 24-hour design values are calculated from the 98th percentile value for each year. In the 24-hour $PM_{2.5}$ calculations, the 98th percentile value from each year is used in the final calculations. Since the 98th percentile value can come from any day in the year, all quarters and years should be carried through to near the end of the calculations. To calculate final future year design

[61] *See* 71 FR 61224 and 40 CFR 50.13.

[62] Similar to the annual $PM_{2.5}$ NAAQS, the default recommended procedures for the 24-hr NAAQS assume handling of the ambient data and model data on a calendar quarter basis. However, in some nonattainment areas, 24-hr $PM_{2.5}$ NAAQS exceedances may occur during a very limited time of the year or season. Processing the data (both ambient and model) on a sub-quarterly basis may be appropriate in certain circumstances.

[63] For most sites, the top 10% of monitored days per quarter will represent between 3 and 8 ambient measurement days.

[64] In some areas it may not be necessary to model and evaluate the 24-hour NAAQS for all quarters. For example, if observed $PM_{2.5}$ concentrations only exceed the NAAQS in the 1st and 4th quarters, and concentrations in the 2nd and 3rd quarters are very low, then it may not be necessary to model the full year. But for areas that have monitored violations (or high values that are close to the NAAQS) in all four seasons, the entire year should be evaluated.

ED_002666_00014243-00130

values, the 98th percentile day for each year is identified and then a 5-year weighted average of the 98th percentile values for each site is calculated to derive the future year design value.[65] In nonattainment areas that measure violations of the daily $PM_{2.5}$ NAAQS in multiple seasons (most often summer and winter), the highest PM days can sometimes shift between seasons when days are projected to the future. A consequence of this shift is that a base year 98th percentile winter day may have a higher future year concentration than a summer day that started with a higher concentration in the base year. Therefore, the recommended attainment test methodology has been designed so that the temporal distribution of high days in the base and future periods can shift between quarters.

In the 24-hour attainment test methodology (described below), the eight highest ambient $PM_{2.5}$ days in each quarter at each site are projected to the future (32 days per year) using species specific quarterly RRFs. After all 32 days are projected to the future, the days are re-ranked to determine the future year 98th percentile day.[66] The rank of the future year 98th percentile day is selected based on the rank of the observed 98th percentile day in the base year ambient data (which depends on the number of valid daily samples). For example, at monitoring site A, if there are 120 observations in year 1 then the 98th percentile day is the 3rd high observation day for the year. In that case, the future year 98th percentile day is selected as the 3rd high future day for the year (out of the 32 highest days).

Similar to the annual $PM_{2.5}$ attainment test, we recommend interpolation techniques for FRM monitors that do not have co-located speciation data. Because the 24-hour standard is a 98th percentile-based value, the species composition on high concentration days may be highly variable from day to day and from site to site. Therefore, while interpolation techniques may be needed, we strongly recommend collecting speciation data at all FRM sites that violate the 24-hour NAAQS. A precise estimate of the $PM_{2.5}$ components at violating sites will help reduce uncertainty in projecting the future year concentration estimates.

---

[65] Similar to the annual average $PM_{2.5}$ attainment test, design values are calculated for consecutive 3-year periods. From the 5-year base period, three design values are calculated and then averaged together to create a 5-year weighted average.

[66] The observed 98th percentile day is always between the 1st and 8th high for the year, depending on the sampling schedule. Therefore, projecting the 8 highest days in each quarter ensures that the observed 98th percentile day is always captured. Additional days could be projected, but a maximum of 32 days per year is needed to guarantee that the 98th percentile observation day at each site is projected to the future.

ED_002666_00014243-00131

We recommend a modeled attainment test for the 24-hour $PM_{2.5}$ NAAQS with the following steps.

Step 1. Identify the high observed $PM_{2.5}$ days at each FRM monitoring site for each year.

The first step in projecting the daily design value is to identify at each FRM site, the eight highest observed 24-hour $PM_{2.5}$ concentration days in each quarter for each year of the base period (up to 5 years), and identify the day rank of the observed 98th percentile value for each year based on the number of collected ambient samples (i.e., 3rd highest day of the year). This results in a data set containing 32 days of data for each year (for up to 5 years) for each site.[67]

The test should be performed for each monitoring site that meets the data completeness criteria for calculating a design value for the 24-hour NAAQS. There may not always be data available for all four quarters and all 5 years. We recommend using 8 days per quarter because the 98th percentile day for a year is always one of the 8 highest days of the year.[68] If all of the high days occur in a single quarter, then using the 8 highest days from each quarter will ensure that the actual 98th percentile day is always captured. This may result in processing more days than necessary, but effectively limits the future year design value calculations to a reasonably small number of days.

Step 2. Calculate "high days" species fractions for each quarter for each FRM monitor.

In this step, quarterly ambient species fractions on "high" days are calculated for each of the major component species of $PM_{2.5}$ (i.e., sulfate, nitrate, ammonium, elemental carbon, organic carbon mass, particle bound water, salt, and blank mass). This calculation is performed by dividing the average monitored concentrations of FRM-derived total $PM_{2.5}$ mass on the high $PM_{2.5}$ days at each site for each quarter, by the average monitored species concentration on the high $PM_{2.5}$ days for each quarter.[69] The default recommendation for identification of "high"

---

[67] The methodology does not assume that the temporal distribution of high days in the base and future periods will remain the same. We recommend projecting at least 8 ambient days per quarter from the base period to the future and then re-rank the entire set of days to find the new future year 98th percentile value (for each year).

[68] If there are 365 samples in a year, then the 98th percentile is the eighth high day.

[69] Similar to the annual average calculations, for FRM sites that do not have co-located species data, we recommend calculating the quarterly species fractions using interpolated species data. For the 24-hour species interpolations, we recommend interpolating the average of the highest 10% of monitor days in each quarter.

ED_002666_00014243-00132

days is the top 10% of days in each quarter. This results in a relatively robust calculation, which typically uses between 3 and 9 days per quarter (depending on the sampling frequency). The end result is a set of quarterly "high days" species fractions for each FRM site.

Step 3. Calculate species concentrations for each of the high ambient days.

Multiply the quarterly "high day" species fractions from step 2 by the $PM_{2.5}$ mass concentration for the 8 high days per quarter identified in step 1. This results in a set of species concentrations for each of the 32 days per year identified in step 1.

Step 4. Calculate component specific RRFs for the high days in each quarter.

For each quarter, calculate the ratio of future year to base year modeled predictions for sulfate, nitrate, elemental carbon, organic carbon, salt, and OPP for the top 10 percent of modeled $PM_{2.5}$ days based on modeled concentrations of 24-hour average $PM_{2.5}$. The result is a set of species-specific "high day" RRFs for each quarter.

The relative response factor for component j at a site i is given by the following expression:

$$(RRF)_{ij} = ([C_{j,\ projected}]/[C_{j,\ base\ year}])_i$$

where $C_{j,\ base\ year}$ is the base year mean species concentration (for the high modeled $PM_{2.5}$ days for each quarter) predicted at the (single) grid cell which contains monitoring $site_i$.

$C_{j,\ projected}$ is the future year mean species concentration (for the high modeled days for each quarter) predicted at the (single) grid cell which contains monitoring $site_i$.

For example, assume that base year predicted sulfate mass on the 10 percent highest $PM_{2.5}$ days for quarter 3 for a particular location is 20.0 $\mu g/m^3$ and the future year modeled sulfate concentration is 16.0 $\mu g/m^3$, then the RRF for sulfate for quarter 3 is 16.0/20.0 or 0.80. Due to potentially large spatial and temporal gradients, we recommend 24-hour $PM_{2.5}$ NAAQS RRFs to be calculated based on the modeled concentrations at the single grid cell where the monitor is located.

ED_002666_00014243-00133

Step 5. Apply the component specific RRFs to observed air quality by quarter.

For each of the 8 days in each quarter, multiply the daily species concentrations from step 3 by the quarterly species-specific RRFs obtained in step 4. If there is 1 modeled base year, then there will be four quarterly RRFs at each monitor. The modeled quarterly RRF for quarter 1 is multiplied by the ambient data for quarter 1 (8 days in each year) for each of the 5 years of ambient data. For example, for day A, 21.0 $\mu g/m^3$ nitrate x 0.75 = future nitrate of 15.75 $\mu g/m^3$. The same procedure is applied for the 8 high days per quarter in the other 3 quarters. This leads to an estimated future concentration for each component for each day.

Step 6. Calculate remaining future year PM2.5 species.

The future year concentrations for the remaining species are then calculated for each of the days.[70] Similar to the annual $PM_{2.5}$ attainment test, we recommend that the future year ammonium is calculated based on the future year sulfate and nitrate concentrations, using a constant value for the degree of neutralization of sulfate (from the ambient data). The future year particle bound water concentration is then calculated from an empirical formula derived from the AIM model. The inputs to the formula are the future year concentrations of sulfate, nitrate, and ammonium calculated in step 5.

Step 7. Sum the species components to get total $PM_{2.5}$ concentrations for each day.

Sum the species concentrations for each day to obtain total $PM_{2.5}$ values for the 32 days per year per site.

Step 8. Determine future year 98th percentile concentrations for each site year.

Sort the 32 days for each site for each year by total $PM_{2.5}$ concentration. For each site year, the monitored 98th percentile rank (for each year) is used to identify the 98th percentile rank for each year.[71] For example, if the base year 98th percentile value for year 1 was the 3rd high concentration, then the future year 98th percentile concentration is identified as the 3rd high future year $PM_{2.5}$ concentration (out of the 32 days).

---

[70] If salt is not explicitly modeled, then the salt RRF should be held constant. Blank mass is assumed to be constant between the base and future year.

[71] This assumes that the ambient sampling schedule (and number of samples) would remain the same between base and future years.

ED_002666_00014243-00134

Step 9. Calculate future 5-year weighted average 24-hour design values and compare to the NAAQS.

The estimated 98th percentile values for each of the 5 years are averaged over 3-year intervals to create 3-year average design values (e.g., the 98th percentile values from year 1, year 2, and year 3 are averaged. Then the 98th percentile values from year 2, year 3, and year 4 are averaged, etc.). These design values (up to 3) are then averaged to create a future year 5-year weighted average design value for each monitoring site.

The preceding steps for determining future year 24-hour $PM_{2.5}$ concentrations are applied for each FRM site. The 24-hour $PM_{2.5}$ design values are truncated after the first decimal place. This approach is consistent with the truncation and rounding procedures for the 24-hour $PM_{2.5}$ NAAQS. Any value that is greater than or equal to 35.5 $\mu g/m^3$ is rounded to 36 $\mu g/m^3$ and violates the 24-hour NAAQS. Similarly, a value of 35.4 $\mu g/m^3$ is rounded down to 34 $\mu g/m^3$ and is attaining the 24-hour NAAQS.

4.6 Local Area Analyses - Special Considerations for Primary $PM_{2.5}$

Primary PM does not undergo chemical transformation between being emitted and its arrival at a receptor location. Thus, high concentrations of primary PM can occur without relatively lengthy travel times from source to receptor. Unlike secondary PM or ozone, we would expect concentrations of primary PM to increase the closer one gets to the source(s) of emissions. Therefore, if a designated nonattainment area contains concentrated sources of primary PM, we would expect there to be relatively large spatial gradients near sources of the primary portion of the organic carbon, elemental carbon and other PM components of ambient $PM_{2.5}$ (these localized areas of high concentrations are often called "hotspots").

Note that the $PM_{2.5}$ NAAQS ambient monitoring rule language[72] indicates that some monitoring locations may not be comparable to the annual $PM_{2.5}$ NAAQS. $PM_{2.5}$ measurement data from monitors that are not representative of "area-wide" air quality, but rather of relatively unique micro-scale, or localized hot spot, or unique middle-scale impact sites, are not eligible for comparison to the annual $PM_{2.5}$ NAAQS. The ambient air quality monitor siting rules define the appropriate scales of influence for the $PM_{2.5}$ monitoring network.[73]

[72] *See* 40 CFR 58.30.
[73] *See* 40 CFR part 58, Appendix E.

ED_002666_00014243-00135

4.6.1 Analysis of Primary $PM_{2.5}$ Impacts at Monitors

The majority of FRM and FEM $PM_{2.5}$ monitors are located at "urban scale" sites. These sites tend to represent relatively large spatial scales and do not have large gradients compared to other monitors in the area. Some sites are classified as "neighborhood" scale and may be influenced by more local sources. There are also near-road sites that have specifically been sited to measure gradients that may exist near heavily travelled roads. For reasons stated above, local influences creating large spatial gradients are likely to consist mostly of primary $PM_{2.5}$ (OC, EC, and OPP). These sources may be point sources, or they may be nearby roads or other mobile or area sources.

As part of the attainment demonstration modeling, it may be necessary to evaluate the local-scale impacts of primary $PM_{2.5}$ sources for contributions to the 24-hour and/or the annual NAAQS. As stated earlier, for the purposes of determining attainment with the NAAQS, it may not always be appropriate to compare ambient data from $PM_{2.5}$ monitoring sites that are dominated by point sources to the annual NAAQS. But there are numerous cases where local source contributions may not be dominant, but are a sizable contributor to total $PM_{2.5}$ (~10-30% of total annual average $PM_{2.5}$). In these cases, a more refined analysis of the contribution of local primary $PM_{2.5}$ sources to $PM_{2.5}$ at the monitor(s) will help explain the causes of nonattainment at and near the monitor and may lead to the more efficient ways to attain the NAAQS by controlling emissions from local sources which may be important contributors to the violating area.

There are several modeling tools that can be used to evaluate contributions of local $PM_{2.5}$ sources. A grid model can be run at very high horizontal resolution (1 or 2 km or finer) or a Gaussian dispersion model can be used. Grid based models simulate chemical transformation and complex meteorological conditions, while dispersion models are generally more simplistic; being limited to a local-scale, using Gaussian approximations with little or no chemistry. Therefore, while dispersion models may not be an appropriate tool for determining secondary $PM_{2.5}$ or ozone concentrations, they work well for use in determining local primary $PM_{2.5}$ impacts. The model(s) and model setup should be evaluated to determine the most appropriate tools for a specific situation. Regardless of which type of models are used to evaluate changes in primary $PM_{2.5}$ at monitoring locations, we recommend that the model results be used in a relative manner. This is consistent with the already described attainment tests. If a grid model is used at very high resolution, the attainment test as described in sections 4.4 and 4.5 should be followed. If a Gaussian dispersion model is used, then the application of the attainment test

will vary slightly. The test will need to combine results from the chemical transport grid based modeled attainment test and the results from local-scale dispersion modeling. Because each nonattainment area has unique emissions sources and source-receptor relationships, air agencies should work closely with their EPA Regional office in developing local area analysis applications. The nature of the analyses and calculations of $PM_{2.5}$ concentrations (base and future) depend on the type(s) of local sources, measured PM gradients, the NAAQS being examined, and the model(s) used.

4.6.2 Analysis of Primary $PM_{2.5}$ Impacts at Monitors Using a Gaussian Dispersion Model

To apply a dispersion model in an attainment test, it is important to determine the local component of primary $PM_{2.5}$ at the monitor and the sources that are contributing to this component. There is no single, simple method for quantifying this contribution, but detailed analysis of ambient data and advanced data analysis techniques such as receptor modeling may help quantify the contribution. EPA provides software and information on the CMB, UNMIX, and PMF receptor models.

The simplest method for identifying the local component of $PM_{2.5}$ is to examine local ambient $PM_{2.5}$ concentrations. For this type of analysis, it is important to identify the local contributions from as small a source area as possible. This will make it easier to identify contributing sources. One aspect of this analysis is to examine monitored concentration differences between urban monitors (with the highest concentrations) and more suburban measurements. This is likely to provide a representative indication of the excess contribution from a relatively local area. "Urban excess" calculations which pair an urban monitor with an appropriate rural background monitor (U.S. EPA, 2004b) are likely to indicate "local" contributions that may be representative of a metropolitan area. In most cases, the local component will include contributions from more than one source.

Sources identified as contributing to the monitor through emissions analysis or data analysis (such as receptor modeling) can then be modeled with a dispersion model (or alternatively, a very high resolution grid model). It is common practice to run dispersion models for a limited number of major point sources or line sources in relatively small areas (out to a maximum

ED_002666_00014243-00137

distance of ~50 km).[74] Dispersion models can also be run for all sources of primary $PM_{2.5}$ in a limited area (including mobile, non-road, and area sources).

When applying a model to evaluate the impacts from emissions of primary $PM_{2.5}$, one should determine the $PM_{2.5}$ species that make up the local primary PM.[75] We recommend that the individual components of primary $PM_{2.5}$ be tracked separately. This will allow for a more thorough evaluation of the modeling results (comparing model output to speciated ambient data) and may aid in the development of primary $PM_{2.5}$ control strategies.

The majority of the primary $PM_{2.5}$ will consist of the primary portion of the organic carbon component, elemental carbon (EC), and the general category of "other primary particulate matter" (OPP). Speciated $PM_{2.5}$ measurements may be able to identify some of the more important "other" inorganic components such as heavy metals or salt. In some cases, directly emitted sulfate (and in rare cases nitrate) may also be a significant component of the local primary $PM_{2.5}$.

As part of a local-source impact analysis, an estimated concentration or fraction of primary OC may be needed. There are several methods available for estimating the primary versus secondary portion of ambient OC. Among these are the EC tracer method and receptor modeling. The EC tracer method is the most common method used to estimate secondary and primary OC concentrations (Turpin, 1995), (Strader, 1999), (Cabada, 2004), (Chu, 2005), (Saylor, 2006). The method uses measurements of OC and EC and calculated OC to EC ratios to identify periods when OC is likely to be mostly primary. This information is then used to calculate the secondary contribution to OC. Receptor models such as CMB and PMF have also been used to estimate secondary organic concentrations (Na, 2004), (Yuan, 2006).

Because each nonattainment area has unique emissions sources and source-receptor relationships, air agencies should work closely with their EPA Regional office in developing local area analysis applications.

---

[74] This guidance document only applies to $PM_{2.5}$ attainment demonstrations to satisfy requirements in CAA Section 189. Other guidance documents and procedures are applicable for SO_2, NO_2, CO, and/or Pb attainment demonstrations and/or $PM_{2.5}$ permitting applications.

[75] The $PM_{2.5}$ emissions should be broken out into individual species using the best information available. Typically, SCC specific speciation profiles are used to "speciate" the PM into individual components. Local, source specific information should be used whenever possible.

ED_002666_00014243-00138

4.6.2.1 Double Counting

Application of a local area analysis may result in double counting of local emissions sources and reductions. This can happen when a change in emissions is examined with both a dispersion model and a chemical transport model.

The potential for double counting can be evaluated by considering the grid resolution of the chemical transport model and the number of sources in the dispersion model. The finer the resolution of the grid model, the more double counting may be a problem (because the local sources will be more finely resolved). Additionally, the more sources that are run with the dispersion model, the more double counting may be a problem. If the grid cell size is relatively large (e.g., 12km) and the primary emissions are relatively small, it may be possible to assume that double counting is small enough to be ignored. But the nature of the issue should be evaluated on a case by case basis.

The simplest way to evaluate the magnitude of the problem is to run the chemical transport model with and without the sources from the dispersion model. This will indicate the maximum impact that the sources can have on the chemical transport modeling results. A very small impact means that double counting is not a problem. A large impact indicates that a more explicit analysis is needed in order to resolve double counting issues.

4.6.2.2 Analysis of Primary $PM_{2.5}$ at Monitors- Quality Assurance

As with any modeling exercise, it is important to quality assure the model inputs and outputs. In particular, we recommend focused quality assurance checks on emissions from sources flagged for a dispersion modeling analysis. Prior to applying a model, air agencies should review available information to ensure that there are no major discrepancies between modeled estimates and nearby monitored data for PM. The emission factors, activity data, and speciation profiles of the $PM_{2.5}$ emissions should also be analyzed for accuracy. If a speciation monitor exists, the speciated data from the monitor can be compared to the speciation profiles of the flagged sources. Receptor models can also be used as a QA tool. Discrepancies between receptor modeling results (which are based on analyzing ambient data) and speciated emissions may indicate a problem with the magnitude and/or the speciation of the emissions sources. If discrepancies are found, those implementing the modeling protocol (see section 2.1) should consult with the appropriate EPA Regional office to reach agreement on what, if any, adjustments should be made to the emissions estimates for the identified sources.

ED_002666_00014243-00139

It is also important to quality assure the model outputs. Modeling each species of primary $PM_{2.5}$ separately within the dispersion model should aid in this analysis, especially if the selected control strategies applied in the future year do not control each primary species to the same degree. If a speciation monitor exists, the speciated data from the monitor may also help quality assure the dispersion model results. However, the application of the dispersion model results in a relative sense will help to reduce the impact of possible over- or under-estimations by the dispersion model due to emissions, meteorology, or general selection of other model input parameters.

4.7 Estimating design values at unmonitored locations

High ozone and/or $PM_{2.5}$ concentrations can occur at (or near) existing ambient monitors or in unmonitored areas. The modeled attainment test is primarily a monitor-based test. As such, the focus of the attainment test is whether attainment can be reached at existing monitors.
An additional "unmonitored area analysis" can also be performed to examine ozone and/or $PM_{2.5}$ concentrations in unmonitored areas. The unmonitored area analysis is intended to be a means for identifying high ozone and/or $PM_{2.5}$ concentrations outside of traditionally monitored locations, particularly in nonattainment areas where modeling or other data analyses have indicated potential high concentration areas of ozone and/or $PM_{2.5}$ outside of the existing monitoring network. An unmonitored area analysis may also be more important in areas where the ozone or $PM_{2.5}$ monitoring networks just meet or minimally exceed the minimum required network in a nonattainment area.

This review can help ensure that a control strategy leads to reductions in ozone or $PM_{2.5}$ at other locations which could have base case (and future) design values exceeding the NAAQS were a monitor deployed there. In addition, analysis of concentrations in unmonitored areas may serve other purposes such as helping to site new (or move existing) monitors, examining health benefits from emissions reductions, and analyzing the impact of potential new or modified emissions sources. This document describes how an unmonitored area analysis can be performed for a particular area. See the respective ozone and $PM_{2.5}$ SIP Requirements Rules (including the preambles and response to comments documents) for more information on when an unmonitored area analysis is recommended or required.

We use the term "unmonitored area analysis" to describe an analysis that uses a combination of model output and ambient data to identify areas that might exceed the NAAQS in areas that are not currently monitored. An unmonitored area analysis for a particular nonattainment area is intended to address the potential for air quality problems within or near that nonattainment

ED_002666_00014243-00140

area. The analysis should include, at a minimum, all nonattainment counties and counties surrounding the nonattainment area, as appropriate. It is possible that unmonitored area violations may appear in counties far upwind or downwind of the local area of interest. In those cases, the distance to the nonattainment area and ability of the modeling to represent far downwind areas should be evaluated on a case-by-case basis.

The spatial resolution of the unmonitored area analysis should be consistent with the appropriate model resolution for the attainment demonstration. For example, if the attainment demonstration for ozone or $PM_{2.5}$ is conducted at 12km grid resolution, then in most cases, the unmonitored area analysis should also be performed at the same horizontal resolution.

4.7.1 Why does the unmonitored area analysis need to use both ambient data and model output?

When performing an unmonitored area analysis, it is recommended to use interpolated spatial fields of ambient data in combination with gridded modeled outputs (i.e., fused surfaces). Ambient data can be interpolated to provide a set of spatial fields. The resultant spatial fields will provide an indication of pollutant concentrations in monitored and unmonitored areas. But interpolated ambient data cannot identify unmonitored areas with higher concentrations than those measured at monitors. The interpolated concentration between monitors will generally be the same or lower than the measured concentration at the monitors (assuming that more sophisticated statistical techniques are not used). Simple interpolation techniques do not account for emissions or chemistry information that may be needed to identify high concentrations in unmonitored areas.

In addition to interpolated surfaces, gridded model output (absolute) concentrations (from a chemical transport model) can be used to examine unmonitored area concentrations. The model provides an hourly concentration output for every grid cell. The concentrations can be analyzed to determine unmonitored areas where the model predicts high values. But the absolute predictions from the model may not be entirely accurate (they may be biased high or low). But unlike the interpolated ambient data, the model explicitly accounts for emissions, chemistry, and meteorological processes over the entire domain.

Interpolated ambient data and absolute model outputs both have strengths and weaknesses. We can take advantage of the strengths of each dataset by combining the two types of data. Interpolated spatial fields of ambient data provide a strong basis for estimating accurate pollutant concentrations at and near monitor locations. And the model outputs can be used to

ED_002666_00014243-00141

adjust the interpolated spatial fields (either up or down) so that more accurate estimates can be derived in unmonitored areas.[76] A recommended way to use the modeling in the creation of fused surfaces is to adjust the ambient spatial fields using the concentration *gradients* derived from the model predictions. It is preferable to assume that the model is accurately predicting areas of generally high or low ozone or PM, rather than to assume that the absolute predictions from the model are correct. For example, in unmonitored locations where the model predicts relatively high ozone or PM concentrations, the spatial fields can be adjusted upward. In unmonitored locations where the model predicts relatively low ozone or PM concentrations, the spatial fields can be adjusted downward. In this way, it may be possible to predict unmonitored locations that may be likely to record high concentrations (were a monitor located there). We refer to this combination of interpolated spatial fields and modeled output as "gradient adjusted spatial fields."

4.7.2 Implementation of Gradient Adjusted Spatial Fields

A recommended methodology for developing gradient adjusted spatial fields is to first create base year fields. Future year estimates can then be created by applying gridded RRFs to the base year gradient adjusted spatial fields. The basic steps are as follows:

1) Interpolate base year ambient data to create a set of spatial fields.
2) Adjust the spatial fields using gridded model-based spatial gradients (base year values).
3) Apply gridded model-based RRFs to the gradient adjusted spatial fields.
4) Determine if any unmonitored areas are predicted to exceed the NAAQS in the future.

EPA's SMAT will spatially interpolate data, adjust spatial fields based on model output gradients, and multiply the fields by model calculated RRFs (steps 1-3 above) (U.S. EPA, 2018), (ABT, 2014). Air agencies can use the EPA-provided software or may wish to develop alternative techniques that may be appropriate for their area, situation, and/or purpose. Air Agencies should consult with the appropriate EPA Regional office to determine if and how an unmonitored area analysis should be conducted.

Step 1

The first step in developing an unmonitored area analysis is to interpolate ambient data. Ideally, design values should be interpolated. Either the 5-year weighted average design values

[76] The accuracy of interpolated fields can be evaluated by removing data from monitors (either in groups or one at a time) to see how well the interpolation scheme estimates known concentrations at ambient monitoring sites. This is known as a cross-validation.

ED_002666_00014243-00142

that are used in the monitor based modeled attainment test or a single design value period can be used in the development of ambient spatial fields. Care should be taken so that the interpolated fields are not unduly influenced by monitoring sites that do not have complete valid data. Since the design values can vary significantly from year to year, it is important to use a consistent set of data.

We are not recommending a single interpolation technique. EPA's SMAT uses the Voronoi Neighbor Averaging (VNA) interpolation technique. Past analyses have used a Kriging interpolation technique (U.S.EPA, 2004).

For ozone analyses, a single spatial field of ozone values is needed. But for the $PM_{2.5}$ NAAQS, the EPA recommends creating interpolated spatial fields for $PM_{2.5}$ (FRM) and for each component species of $PM_{2.5}$. For the annual $PM_{2.5}$ NAAQS, a set of quarterly average spatial fields can be created. The four quarters are averaged to get an annual average set of fields.

For the 24-hour $PM_{2.5}$ standard, a spatial field can be created using the high end of the distribution of 24-hour $PM_{2.5}$ concentrations. This is best represented by interpolating the measured high values from each quarter. For the $PM_{2.5}$ component species, we recommend interpolating the high $PM_{2.5}$ days in each quarter. This can be based on the highest monitored days in each quarter.

Step 2

The second step in the process involves the use of gridded model output to adjust the spatial fields derived from ambient data. SMAT uses the eVNA technique, where the VNA interpolation is "enhanced" by adjusting the interpolated values based on the modeled ozone or PM (or PM species) gradient. See the SMAT user's guide for more details (U.S. EPA, 2018), (Abt, 2014). Other "model adjusted" spatial fields techniques include Hierarchical-Bayesian (McMillan, 2010) and "Downscaler" (Berrocal, 2010a and 2010b).

A specific metric is needed to determine the model predicted gradient in concentrations for each of the NAAQS. For ozone, a logical metric is the 4th highest ozone prediction in each grid cell.[77] For the annual $PM_{2.5}$ NAAQS, the model predicted quarterly mean concentrations for $PM_{2.5}$ and $PM_{2.5}$ species can be used to adjust the ambient spatial fields. For the 24-hour $PM_{2.5}$ NAAQS, the gradient adjusted fields can be derived from the high end of the distribution of

[77] The metric should approximate the measured design values at monitoring sites. Depending on the days modeled, other metrics, such as the average of the top 4 modeled days may be a better proxy for the design value.

ED_002666_00014243-00143

daily averages in each quarter. This could be for the top 10% of modeled days in each quarter or for all days with 24-hour average concentration > 35 µg/m3 (or some other relatively high value).

Step 3

The next step is to create future year spatial fields by multiplying the base year gradient adjusted spatial fields by model derived gridded RRFs. The RRFs for the unmonitored area analysis are calculated in the same way as the monitored based attainment test (except that the grid cell array is not typically used in the spatial fields based analysis). The future year concentrations are equal to the base year concentration multiplied by the RRF in each grid cell. For ozone, a single spatial field is multiplied by a single set of model derived RRFs. For $PM_{2.5}$, the RRFs for each of the species, for each quarter, are multiplied by the spatial fields for each species, for each year.

4.7.3 Using the Results of the Unmonitored Area Analysis

For ozone and/or $PM_{2.5}$ attainment demonstrations, the future year gradient adjusted spatial fields can be analyzed to determine if any grid cells are predicted to remain above the NAAQS. States should consult with their EPA Regional office to determine if an unmonitored area analysis should be conducted for a particular nonattainment area and, if so, how the analysis should be conducted and how the results should be used.

ED_002666_00014243-00144

5.0 What Is the Recommended Modeling Method for Developing Air Quality Goals for Regional Haze?

This section focuses on the modeling analysis needed to set RPGs that reflect the enforceable emission limitations, compliance schedules, and other measures included in the long-term strategy of a regional haze SIP.[78] In this section, we describe the recommended modeled analysis to assess visibility conditions at the end of an implementation period (i.e., the future end-point of a multi-year planning period under the regional haze program). The visibility analysis has many similarities to the $PM_{2.5}$ attainment tests described in sections 4.4 and 4.5. The recommended visibility analysis and the attainment tests both use monitored data to define base year air quality. They divide $PM_{2.5}$ into major species, and component-specific relative response factors are multiplied by base year monitored values to estimate future concentrations.

Section 5.1 provides background information on the requirements of the Regional Haze Rule and explains how RPGs are used in the regional haze program once they are set. The rest of the section describes the recommended modeling analysis to evaluate future year changes in visibility impairment. This document does not address how a state should develop its long-term strategy of emissions controls for regional haze or determine the uniform rate of progress for a Class I area.

5.1 Regional Haze Rule Background

The modeling results are used to quantify the amount of progress, expressed in deciviews of visibility improvement, anticipated to be made at individual Class I areas as of the end of an implementation period. Reasonable progress goals are calculated separately for the 20% most impaired days and the 20% clearest days at each Class I area. These values are then compared to visibility conditions in the "baseline period" (2000-2004 for most areas)[79] for the same type of

[78] The main regional haze section of the guidance is related to setting reasonable progress goals. However, the guidance methods may also be applicable to other regional haze related modeling, including, but not limited to, evaluation of visibility impacts from sources and/or source sectors (including international impacts).

[79] In the context of attainment demonstrations for ozone and $PM_{2.5}$, the term "baseline period" is commonly used to refer to the base period used for modeling. Because the Regional Haze Rule defines "baseline period" specifically as 2000-2004, this section uses "baseline" and "baseline period" only when referring to 2000-2004. The term "base period" is used when referring to a more recent 5-year period used in the modeling-based projection of RPGs.

ED_002666_00014243-00145

days (i.e., most impaired and clearest) to verify that the SIP provides for an improvement in visibility on poor visibility days and allows no degradation in visibility on good visibility days. The RPGs for the most impaired days are also plotted against the uniform rate of progress that would be needed at individual Class I areas to achieve natural visibility conditions in 2064, starting at baseline visibility conditions for 2000-2004. The results of this comparison inform whether the SIP must contain an additional component demonstrating that no additional control measures are necessary to make reasonable progress.

Reasonable progress goals are a regulatory construction promulgated to implement the statutory requirements for visibility protection. Section 169A(a)(1) of the Clean Air Act states, "Congress hereby declares as a national goal the prevention of any future, and the remedying of any existing, impairment of visibility in mandatory class I Federal areas which impairment results from manmade air pollution." This section later directs EPA to promulgate regulations to assure reasonable progress towards meeting this goal. Section 169B calls for EPA to "carry out the Administrator's regulatory responsibilities under [Section 169A], including criteria for measuring 'reasonable progress' toward the national goal."

In response to these mandates, EPA promulgated the Regional Haze Rule on July 1, 1999.[80] The rule was revised on January 10, 2017.[81] The regulations promulgated at 40 CFR 51.308(f) contain the relevant requirements for SIPs for the second and subsequent implementation periods, which are the SIPs that this guidance document addresses. Under section 51.308(f)(2), every state must determine the emission limitations, compliance schedules, and other measures that are necessary to make reasonable progress at Class I areas in the state and other Class I areas that may be affected by sources in the state. The collection of these measures is referred to as the long-term strategy. Under section 51.308(f)(3), a state in which a mandatory Class I Federal area is located must establish RPGs, one each for the 20% most impaired days and the 20% clearest days (expressed in deciviews) for each such area. These RPGs reflect the visibility conditions that are projected to be achieved by the end of the applicable implementation period as a result of its own and other states' long-term strategies, as well as the implementation of other requirements of the CAA. They are interim goals that represent the incremental visibility improvement achieved in each implementation period toward Congress's goal of eliminating visibility impairment from manmade emissions sources. The long-term strategy and the RPGs are developed in consultation with other affected air agencies and

[80] *See* 64 FR 35714.

[81] *See* 82 FR 3078.

ED_002666_00014243-00146

Federal Land Managers.[82]

The Regional Haze Rule contains the following requirements that involve the RPGs once they have been set.

> The reasonable progress goals (RPGs) must provide for an improvement in visibility for the 20% most anthropogenically impaired days relative to baseline visibility conditions and ensure no degradation in visibility for the 20% clearest days relative to baseline visibility conditions.[83] As explained in the RHR, the baseline for each Class I area is the average visibility (in deciviews) for the years 2000 through 2004.[84] The visibility conditions in these years are the benchmark for both the "provide for an improvement" and "no degradation" requirements.
>
> States are also required to determine the rate of improvement in visibility needed to reach natural conditions by 2064 for the 20% most anthropogenically impaired days, given the starting point of the 2000-2004 baseline.[85] The "glidepath," or uniform rate of progress (URP), is the amount of visibility improvement needed to stay on a linear path from the baseline period to natural conditions. Every state must compare its RPG for the 20% most anthropogenically impaired days to the point on the glidepath corresponding to the end of the implementation period. The results of this comparison determine whether a SIP must contain a demonstration that no additional emission reductions are needed to make reasonable progress towards achieving natural visibility conditions.

This section does not address the process of calculating the glidepath or determining the measures for inclusion in the long-term strategy.[86]

5.2 How is Regional Haze Measured?

Regional haze is measured by an *extinction coefficient* (b_{ext}) that represents light attenuation resulting from scattering and absorption of light from ambient PM plus scattering of light due to gas molecules in the air (i.e., Rayleigh scattering). Although b_{ext} can be estimated by several

[82] *See* 40 CFR 51.308(f)(2)(ii) and 51.308(i).

[83] *See* 40 CFR 51.308(f)(3)(i).

[84] *See* 40 CFR 51.308(f)(1) and definitions in 51.301.

[85] *See* 40 CFR 51.308(f)(1).

[86] For more details on calculating the URP and regional haze policy guidance, see draft guidance for the second implementation period [note that the draft guidance document will be updated in Spring 2019].

ED_002666_00014243-00147

different methodologies, the RHR requires that it be estimated using measured ambient PM. This follows since, for a given set of meteorological conditions, visibility can be improved by reducing concentrations of ambient PM. Thus, deriving b_{ext} in this manner provides a direct link between regional haze and related pollutant concentrations.

5.2.1 IMPROVE Algorithm

The IMPROVE equation or algorithm reflects empirical relationships derived between measured mass of PM components and transmissometer measurements of b_{ext} at a subset of monitoring sites in Class I areas within the IMPROVE monitoring network. The IMPROVE program revised the IMPROVE algorithm in 2006 (Hand, 2006); (Pitchford, 2007). The revised algorithm is intended to reduce biases in light extinction estimates. The revised algorithm is as follows:

$$
\begin{aligned}
bext = {} & 2.2 \times f_s(RH) \times [\text{Small Sulfate}] + 4.8 \times f_L(RH) \times [\text{Large Sulfate}] \\
& + 2.4 \times f_s(RH) \times [\text{Small Nitrate}] + 5.1 \times f_L(RH) \times [\text{Large Nitrate}] \\
& + 2.8 \times [\text{Small Organic Mass}] + 6.1 \times [\text{Large Organic Mass}] \\
& + 10 \times [\text{Elemental Carbon}] \\
& + 1 \times [\text{Fine Soil}] \\
& + 1.7 \times f_{ss}(RH) \times [\text{Sea Salt}] \\
& + 0.6 \times [\text{Coarse Mass}] \\
& + \text{Rayleigh Scattering (site specific)} \\
& + 0.33 \times [NO_2\ (\text{ppb})] \qquad \textbf{[5.1]}
\end{aligned}
$$

The numerical coefficients on the right hand side of the equation represent the light scattering or absorption efficiency, m^2/gm of the corresponding component of particulate matter,

$f_s(rh)$, $f_L(rh)$, $f_{ss}(rh)$ are relative humidity adjustment factors applied to the light scattering efficiency (to be described in greater detail shortly), dimensionless;

SO_4 is the mass associated with sulfates, $\mu g/m^3$;

NO_3 is the mass associated with nitrates, $\mu g/m^3$;

OC is the mass associated with organic carbon, $\mu g/m^3$;

EC is the mass associated with elemental carbon, $\mu g/m^3$;

ED_002666_00014243-00148

Fine Soil is inorganic primary particulate matter (excluding primary sulfate and nitrate particles) associated with soil components with aerodynamic diameter $\leq$ 2.5 µm, µg/m^3;

CM is coarse PM with aerodynamic diameter > 2.5 µm, but $\leq$ 10 µm, µg/m^3;

$b_{rayleigh}$ is light-scattering attributable to Rayleigh scattering, Mm^{-1} (i.e., inverse "mega-meters"); and

b_{ext} is the estimated extinction coefficient, Mm^{-1}.

The total sulfate, nitrate and organic mass concentrations are each split into two fractions, representing small and large size distributions of those components. The organic mass concentration is calculated as 1.8 times the measured IMPROVE organic carbon concentration, to adjust for organic mass from elements other than carbon. Terms are included for sea salt (important for coastal locations) and for absorption by NO_2 (only used where NO_2 data are available). Site-specific Rayleigh scattering is calculated based on the elevation and annual average temperature of each IMPROVE monitoring site.

The apportionment of the total concentration of sulfate compounds into the concentrations of the small and large size fractions is accomplished using the following equations:

[Large Sulfate] = [Total Sulfate]/20 µg/m^3 x [Total Sulfate], for [Total Sulfate] < 20 µg/m^3
[Large Sulfate] = [Total Sulfate], for [Total Sulfate] $\geq$ 20 µg/m^3

[Small Sulfate] = [Total Sulfate] - [Large Sulfate]

The same equations are used to apportion total nitrate and total organic mass concentrations into the small and large size fractions.

Sea salt is calculated as 1.8 x [Chloride] or 1.8 x [Chlorine] if the chloride measurement is below detection limits, missing or invalid. The algorithm also uses three water growth adjustment terms. They are for use with the small size distribution and the large size distribution sulfate and nitrate compounds and for sea salt (f_S(RH), f_L(RH) and f_{SS}(RH) respectively).

All or nearly all states used the revised algorithm in their SIPs for the first implementation period and the IMPROVE program has stopped reporting visibility data using the original

IMPROVE equation. Therefore, we now recommend using the revised IMPROVE equation for all reasonable progress related modeling calculations.

5.2.2 Deciview Haze Index

The RHR requires that reasonable progress is to be measured in terms of changes in a *deciview haze index* or simply "deciviews" (dv). Deciviews are defined as the natural logarithm of the ratio of the extinction coefficient to Rayleigh scattering (Pitchford and Malm, 1994).

$$\text{Deciview} = 10 \ln(b_{ext}/10) \qquad \textbf{[5.2]}$$

Where the units of b_{ext} and light scattering due to Rayleigh scattering[87] (i.e., the "10" in the denominator of the logarithmic expression) are both expressed in Mm^{-1}.

A *change* in deciviews, which is how progress is tracked, is given by Equation (5.3). A one deciview change is equivalent to ~10% change in b_{ext}.

$$\Delta dv = dv_{future} - dv_{base} \qquad \textbf{[5.3]}$$

A negative number indicates a reduction in deciviews, which is an improvement in visibility.

5.2.3 Estimating Mass Associated with Components of Particulate Matter

All Class I areas either have on-site speciated $PM_{2.5}$ measurements or have been assigned a representative IMPROVE monitoring site with measurements.[88] Therefore, it is generally not necessary to interpolate measured $PM_{2.5}$ species data to a site. The existing IMPROVE database of $PM_{2.5}$ measurements should be adequate to provide data for all Class I areas.[89]

FRM data is not used in the regional haze analysis and it is not necessary for the sum of species

---

[87] Even though the "revised" IMPROVE equation uses a site-specific Rayleigh scattering value, the denominator in the deciview equation should always be 10. In this way, the deciview calculation will be consistent across all Class I areas. When the site-specific Rayleigh scattering is less than 10 (at high elevation sites), the deciview value can be negative (on very clean days), but this should not be considered a problem.

[88] See U.S. EPA 2003b, Appendix A, Table A-2.

[89] There are some IMPROVE sites that do not have enough complete data to provide baseline condition information for the 2000-2004 period. Data substitution is addressed in 40 CFR 51.308(f)(1)(i) and also discussed in (U.S. EPA 2003b).

ED_002666_00014243-00150

components to equal gravimetric measurements obtained with a Federal Reference or equivalent method for measuring $PM_{2.5}$. Therefore, for regional haze calculations, it is not necessary to use the speciated data adjustment procedures introduced in section 4.4 (the "SANDWICH" adjustments).

Hand (2006) has developed a set of default assumptions for mass associated with each of the components of PM for Equation (5.1). These are presented in Table 5.1. The components in Table 5.1 are similar to those used in the modeled attainment demonstrations for the $PM_{2.5}$ NAAQS. Notice, however, that sulfate and nitrate mass are assumed to be fully neutralized (therefore, ammonium is not needed as a separate component); organic carbon is assumed to be equal to 1.8 times **measured** OC mass; there is no water component; and there is a term for coarse PM.

Table 5.1. Default Assumptions Used to Derive Aerosol Species Concentrations from IMPROVE Data for Estimating Extinction Coefficients

(1) Species	(2) Formula	(3) Assumptions
Sulfate	**1.375 *measured SO4**	**Sulfate is assumed to be fully neutralized (ammonium sulfate)**
Nitrate	**1.29 * measured nitrate**	**Denuder efficiency is ~100% & all nitrate is from ammonium nitrate**
EC	**high + low temperature EC**	**All high temperature carbon is elemental**
OC	**1.8 * organic carbon**	**Average organic molecule is 56% carbon**
Fine Soil	**2.2*Al + 2.49*Si + 1.63*Ca + 2.42*Fe + 1.94*Ti**	**(Soil K)=0.6(Fe), FeO & Fe_2O_3 are equally abundant, a factor of 1.16 is used for MgO, Na_2O, H_2O & CO_3**
Sea Salt	**1.8*Cl-**	**Sea salt is 55% chloride ion by weight**

ED_002666_00014243-00151

$PM_{2.5}$		**-Measured gravimetrically -Represents dry ambient fine aerosol mass**
CM (coarse mass)	**(PM_{10}) - ($PM_{2.5}$)**	**Consists only of insoluble soil particles**

Throughout this section we make every attempt to be consistent (aside from the use of the revised IMPROVE algorithm) with corresponding parts of the "Guidance for Tracking Progress under the Regional Haze Program" (hereafter referred to as the "Tracking Guidance") (U.S. EPA, 2003b). It will be easier to interpret the modeling results and compare them to measured values if the modeled future year data are calculated in a manner similar to the ambient data.

5.2.4 Normalizing Trends in Regional Haze - f(rh) Factors

It is clear from equation 5.1 that relative humidity can have a substantial effect on estimated extinction coefficients, as well as the relative importance of changes in different components of PM on trends in regional haze. Because of the importance of relative humidity as a determinant of regional haze, it is necessary to normalize any apparent trend in the estimated extinction coefficient for differences in relative humidity. This enables an assessment of how much an emissions control strategy will reduce regional haze, without the confounding effects of different relative humidity levels during the base and future periods.

There are two possible approaches to normalize trends in visibility for changes in relative humidity between the past and future. The first is to assume that the same day-specific values for relative humidity which were observed in the base period calculations will occur in future years. Thus, one would use the relative humidity observations made on a specific day together with measured components of particulate matter on that day to compute the day-specific visibility extinction coefficient on that day in the future. However, the approach could lead to misleading conclusions if humidity observations were missing for some days or if the humidity observations are atypical in some way. Further, if an air agency wanted to perform visibility calculations in a number of Class I areas, it would need to obtain hourly relative humidity data for each area.

The second approach is to review relative humidity data over a long period of record to derive climatological estimates for relative humidity adjustment factors. These climatological estimates can then be used in Equation 5.1 to estimate visibility extinction coefficients. These

ED_002666_00014243-00152

estimates are more likely to reflect "typical" relative humidity at different times of year and, thus, expected future visibility extinction.

There are obvious advantages in using climatological f(rh) data for tracking progress of measured data. The measured relative humidity on the 20% most impaired days in a historical period will be different than the measured relative humidity in a future period. Therefore, the only way to normalize for relative humidity between the base and future years in the context of measured data is to use a single climatological value. However, in the context of modeling, there is a choice between climatological data or modeled (or measured) relative humidity data from a meteorological model. Since the modeled meteorological data is held constant when predicting future year air quality, the modeled relative humidity values would also be constant. This would be one way to normalize the data. But since the Tracking Guidance recommends tracking visibility using climatological f(rh) values, it will be easier to interpret the modeling results and compare them to future measured values if the data are calculated in a similar manner. Therefore, we recommend using climatological f(rh) values to calculate base period and future visibility. Appendix A (Table A-1) from the Tracking Guidance displays the relationship between relative humidity and f(rh) values. The f(rh) values were calculated from relative humidity data reported by Tang (1996). The f(rh) values are 1.00 up to 36% relative humidity and grow to a maximum value of 7.40 at a relative humidity of 95%.[90] These values were calculated using 10 years of meteorological data.

For calculations with the revised IMPROVE equation, three separate f(rh) curves are needed (small, large, and sea salt). These can be found in (Hand, 2006). The monthly f(rh) values for both the old and revised IMPROVE equations have been incorporated into pre-calculated daily extinction and deciview data that can be found on the IMPROVE website. The use of pre-calculated f(rh) and visibility data makes it easier for air agencies and regional organizations to calculate visibility changes and will also provide for more consistency.

5.3 How Do I Apply a Model to Calculate Changes in Visibility?

The purpose of a regional haze modeling assessment is to project future visibility conditions representing the RPGs, which are used to satisfy the requirements described in section 5.1. In

[90]For the regional haze calculations, the f(rh) factors have been capped at the f(rh) value associated with 95% relative humidity. Relative humidity value above 95% use the same value.

ED_002666_00014243-00153

the procedures below, the observed base period visibility data should be linked to the base modeling year. Similar to the ozone and $PM_{2.5}$ attainment tests, the 5-year ambient data period should be centered about the base modeling year. For example, for a base modeling year of 2014, the ambient IMPROVE data should be from the 2012-2016 period.[91] However, unlike the ozone and $PM_{2.5}$ attainment tests, the calculation is a 5-year mean where each year counts equally (unlike the 5-year weighted average values for the ozone and $PM_{2.5}$ attainment test). This is consistent with the regional haze rule ambient data calculations for the 20% most impaired days and 20% clearest days. The analysis we recommend has 6 steps.

1) For each Class I area, estimate anthropogenic impairment on each day using observed speciated $PM_{2.5}$ and PM_{10} data plus climatological f(rh) and Rayleigh scattering data for each of the 5 years comprising the base period (this estimation step is not addressed in this document), and rank the days on this indicator. This ranking will determine the 20% most anthropogenically impaired days and the 20% clearest days.
2) For each of the 5 years comprising the base period, calculate the mean deciviews for the 20% most anthropogenically impaired days and 20% clearest days. For each Class I area, calculate the 5-year mean deciviews for most impaired and clearest days from the 5 year-specific values.
3) Use an air quality model to simulate base period emissions and future emissions. Use the resulting information to develop RRFs for each component of PM identified on the right-hand side of Equation (5.1).
4) Multiply the RRFs by the measured species concentration data during the base period (for the measured 20% most impaired days and 20% clearest days). This results in daily future year species concentrations data.
5) Using the results in Step 4 and the IMPROVE algorithm (equation 5.1), calculate the future daily extinction coefficients for the 20% most impaired days and 20% clearest days in each of the 5 base years.
6) Calculate daily deciview values (from total daily extinction) and then compute the future average mean deciviews for the 20% most impaired days and 20% clearest days for each

[91] The *baseline period* for the regional haze program continues to be 2000-2004, and the uniform rate of progress is calculated using that historical data. However, the modeled reasonable progress visibility projections should use ambient data from a 5-year *base period* that corresponds to the modeled base year meteorological and emissions data. In addition, ambient data trends for the 5-year base period should be examined to make sure the base modeling year emissions are representative of the period. In some cases, an alternative base ambient data period may be more representative of a base modeling year.

ED_002666_00014243-00154

year. Average the 5 years together to get the final future year mean deciview values for the 20% most impaired days and 20% clearest days.

We describe each of these steps more fully below. The methodology follows the same basic procedures outlined in the Tracking Guidance, however, based on the revised regional haze rule, a draft methodology (for steps 1 and 2) is available for calculating the most anthropogenically impaired and clearest days (U.S. EPA, 2016d).[92] We conclude this subsection with an example illustrating the recommended modeled uniform rate of progress analysis.

Step 1. Using monitored data, rank base period visibility for each day with PM_{10}, $PM_{2.5}$ and speciated $PM_{2.5}$ measurements within a Class I area.

Ranking should be performed separately for each of the 5 years comprising the base period.[93] The deciview (dv) should serve as the basis for ranking. Day-specific observations for mass associated with SO_4, NO_3, OC, EC, soil, and CM, as defined in Table 5.1, should be used to calculate b_{ext} for each day. The appropriate month- and area-specific climatological relative humidity adjustment factor(s) (*f(rh)*) should be used. Total b_{ext} for all components should be converted to deciviews for each day to get a daily deciview value.

Step 2. Calculate the average base period deciviews for the 20% most anthropogenically impaired days and 20% clearest days.

For each of the 5 years in the base period, order all days considered in Step 1 from worst (highest deciview value) to best (lowest deciview value) visibility. For each year, note the 20% most impaired days and 20% clearest days. Calculate the arithmetic mean deciview value for the identified 20% most impaired days and 20% clearest days in each year. Average the resulting 5 yearly mean deciview values reflecting the most impaired days. This represents the value subject to improvement requirement (i.e., the value must decrease). Average the 5-yearly mean deciview values reflecting mean visibility on the clearest days. This is the value subject to the requirement of no degradation (i.e., the value cannot increase).

---

[92] A final recommendation is expected in Fall 2018.

[93] Pre-calculated and ranked anthropogenic impairment, extinction, and deciview calculations for all Class I areas is available on the IMPROVE website. This data can be used to satisfy steps 1 and 2. States may also choose to use an alternative methodology to calculate anthropogenic impairment for Class I areas, which may differ from the default EPA recommended methodology. States should consult with their EPA Regional office to discuss appropriate methodologies for alternative impairment calculations.

ED_002666_00014243-00155

Step 3. Estimate RRFs for each component of $PM_{2.5}$ and for CM.

The RRFs should be determined using results from air quality modeling of base year and future year emissions. These model simulations should be performed for a large number of days (presumably at least a full year; see section 2.3.1 for more details). The (temporal) arithmetic mean concentration for each $PM_{2.5}$ component (and coarse mass) computed near the Class I monitoring site from the future year modeling is divided by the corresponding arithmetic mean concentration for each component from the base year modeling. The resulting ratios are the component-specific RRF's. A separate set of RRF values are calculated for the 20% most impaired days and 20% clearest days identified in step 2. The RRFs are calculated using the model outputs corresponding to the *monitored* 20% most impaired days and 20% clearest days at each Class I area. This will likely be a different set of days at each monitor.

Step 4. Using the RRFs and ambient data, calculate future year daily concentration data for the clearest and most impaired days.

Multiply the RRFs derived in Step 3 by measured daily concentration data for each component of $PM_{2.5}$ and CM to get future daily estimates of species concentrations for $PM_{2.5}$ components and CM on the 20% most impaired days and 20% clearest days. These multiplications produce future concentration estimates for SO_4, NO_3, OC, EC, soil, sea salt, and CM for each of the previously selected 20% most impaired days and 20% clearest days. This calculation is performed for the 5-year period using an RRF for each PM component (a separate set of RRFs for the 20% most impaired days and 20% clearest days).

Step 5. Use the information developed in Step 4 to compute future year daily b_{ext} values for the 20% most impaired days and 20% clearest days.

Use the future year concentration data calculated in step 4 to calculate future year daily b_{ext} values for each PM component for each of the 20% most impaired days and 20% clearest days for the 5-year period. This is accomplished by applying the revised IMPROVE visibility algorithm (equation 5.1).

Step 6. Use the daily total b_{ext} values from step 5 to calculate future mean deciview values for the 20% most impaired days and 20% clearest days .

The total daily b_{ext} for each day is converted to deciviews. This gives a future year daily deciview value for each of the 20% most impaired days and 20% clearest days.

Next, compute the arithmetic mean future deciview value for the 20% most anthropogenically

impaired days and 20% clearest days for each year. This leads to five future estimated mean deciview values for the worst and five future estimated mean deciview values for the best visibility days. Compute the arithmetic mean of the five mean values for deciviews on the 20% most anthropogenically impaired days, and the arithmetic mean of the five mean deciview values estimated for the 20% clearest days. The resulting 5-year average mean values for deciviews on 20% most impaired days and 20% clearest days will be compared to the base period mean deciview values calculated in step 2. If the resulting change in deciviews is a negative number (future - base), this represents an improvement in visibility. The future year visibility value for the 20% most impaired days will also be compared to the glidepath to determine if the Class I area is expected to be on, above, or below the glidepath in the future implementation period.

Example 5.1

We use example 5.1 to illustrate the modeled visibility assessment. For ease of presentation, we assume there are only 10 speciated samples for PM in each of the 5 years comprising the base period. Since sampling occurs every third day, a typical sample size is ~120 days per year. We illustrate the calculations for the first base year and then furnish information regarding mean deciviews for the other 4 base years to illustrate subsequent steps in the test.

Given:

Ten days have measured components of PM in a Class I area during the first year of a 5-year base period. Table 5.2 below shows the measurements (in µg/m^3) for each of the 10 days. Table 5.2 also shows the date of each measurement and the corresponding climatological relative humidity adjustment factor (made up for this example) for the appropriate month and area.

Table 5.2 Example observed IMPROVE data at a Class I area

Day	Date	f_S(RH)	f_L(RH)	F_{SS}(RH)	Ammonium Sulfate (µg/m3)	$SO4_L$ (µg/m3)	$SO4_S$ (µg/m3)	Ammonium Nitrate (µg/m3)	$NO3_L$ (µg/m3)	$NO3_S$ (µg/m3)
1	1/7	3.82	2.75	3.91	2.004	0.201	1.804	1.050	0.055	0.995
2	2/24	3.29	2.46	3.49	2.290	0.262	2.028	0.453	0.010	0.442
3	4/15	3.73	2.67	3.68	2.617	0.342	2.274	1.206	0.073	1.134
4	6/20	3.82	2.73	3.82	4.580	1.049	3.531	0.528	0.014	0.514
5	7/14	4.29	3.00	4.20	2.742	0.376	2.366	0.416	0.009	0.407
6	8/10	4.35	3.04	4.27	3.512	0.617	2.896	0.161	0.001	0.160
7	8/28	4.35	3.04	4.27	2.303	0.265	2.038	0.224	0.003	0.221
8	9/18	4.59	3.17	4.44	1.759	0.155	1.604	0.409	0.008	0.400
9	10/21	4.11	2.92	4.13	1.660	0.138	1.522	0.159	0.001	0.158
10	12/14	4.21	2.97	4.20	1.975	0.195	1.780	0.854	0.036	0.818

ED_002666_00014243-00157

Day	Date	Organic Mass (µg/m3)	OM_L (µg/m3)	OM_S (µg/m3)	EC (µg/m3)	Soil (µg/m3)	Sea Salt (µg/m3)	CM (µg/m3)	Rayleigh Mm^{-1}
1	1/7	3.024	0.457	2.567	0.345	0.106	0.214	1.334	12
2	2/24	1.073	0.058	1.016	0.150	0.058	0.150	2.004	12
3	4/15	2.409	0.290	2.119	0.241	0.589	0.357	5.928	12
4	6/20	2.482	0.308	2.174	0.282	0.367	0.132	3.549	12
5	7/14	3.608	0.651	2.957	0.297	0.343	0.157	6.252	12
6	8/10	1.768	0.156	1.612	0.268	0.134	0.026	1.989	12
7	8/28	1.772	0.157	1.615	0.162	0.190	0.165	2.521	12
8	9/18	1.049	0.055	0.994	0.123	0.112	0.725	3.880	12
9	10/21	1.267	0.080	1.186	0.094	0.218	0.086	1.996	12
10	12/14	1.454	0.106	1.348	0.248	0.170	0.046	2.081	12

Using the IMPROVE concentration data for the Class I area and modeled RRFs, calculate the future year extinction and deciview values for the 20% most impaired days and the 20% clearest days for the end of the implementation period addressed in the SIP.

Step 1. Using monitored data, rank base period visibility for each day with PM_{10}, $PM_{2.5}$, and speciated $PM_{2.5}$ measurements within a Class I area.

First, estimate the extinction coefficient for each day that has the needed PM measurements. This is done using the information in table 5.2 with Equation (5.1). For day 1 in year 1, the current extinction coefficient is:

b_{ext} = (2.2*3.82*1.804) + (4.8*2.75*0.201) + (2.4*3.82*0.995) + (5.1*2.75*0.055) + (2.8*2.567) + (6.1*0.457) + (10*0.345) + (1*0.106) + (1.7*3.91*0.214) + (0.6*+ 1.334) + 12

b_{ext} = 55.461 Mm^{-1}

Then convert extinction into deciviews:

dv = 10 * ln (55.461/10) = 17.131 dv

Base year extinction coefficients and deciviews for the remaining 9 days with monitored data in year 1 are calculated in a similar manner. The days are then ranked. The day with the highest impairment is given a rank of "1." The results of these calculations are displayed in the table 5.3 below. Based on these rankings, days 4 and 5 comprise the 20% most impaired. Days 2 and 8 comprise the 20% clearest days.[94]

[94] Note that this is an abbreviated example. The actual IMPROVE sampling schedule collects data every third day. Therefore, with a complete set of data, there will be ~121 samples per year at each site. Both the 20% most impaired days and 20% clearest days will be comprised of an average of ~24 sample days.

Table 5.3 Example observed IMPROVE extinction, deciviews, and rank at a Class I area

Day	Date	Rayleigh (Mm^{-1})	SO4 (Mm^{-1})	NO3 (Mm^{-1})	OM (Mm^{-1})	EC (Mm^{-1})	Soil (Mm^{-1})	CM (Mm^{-1})	Sea Salt (Mm^{-1})	B_{ext} [base period] (Mm^{-1})	Deciviews	Rank
1	1/7	12	17.809	9.897	9.975	3.447	0.106	0.801	1.425	55.461	17.131	5
2	2/24	12	17.777	3.621	3.195	1.500	0.058	1.203	0.891	40.244	13.924	9
3	4/15	12	23.049	11.139	7.702	2.411	0.589	3.557	2.231	62.678	18.354	3
4	6/20	12	43.421	4.905	7.967	2.815	0.367	2.129	0.857	74.460	20.077	1
5	7/14	12	27.743	4.327	12.249	2.969	0.343	3.751	1.118	64.499	18.641	2
6	8/10	12	36.711	1.687	5.467	2.680	0.134	1.194	0.187	60.061	17.928	4
7	8/28	12	23.369	2.351	5.481	1.624	0.190	1.512	1.199	47.726	15.629	7
8	9/18	12	18.556	4.546	3.117	1.234	0.112	2.328	5.473	47.366	15.553	8
9	10/21	12	15.695	1.577	3.811	0.943	0.218	1.198	0.607	36.048	12.823	10
10	12/14	12	19.267	8.817	4.418	2.475	0.170	1.248	0.328	48.723	15.836	6

Step 2. Calculate the average base year deciviews for the 20% most impaired days and 20% clearest days.

For year 1, mean deciview value for 20% most impaired days = (20.077 + 18.641) / 2 = 19.359 dv,
and mean deciview value for the 20% clearest days = (12.823 + 13.924) / 2 = 13.374 dv

Mean deciview values for the 20% most impaired and 20% clearest days for years 2-5 are provided in table 5.4.

Table 5.4 Example mean worst days visibility and best days visibility at a Class I area

Year	Mean $dv_{current}$ 20% Most Anthropogenically Impaired Days	Mean $dv_{current}$ 20% Clearest Days
1	19.35	13.37
2	20.32	12.54
3	19.21	11.78
4	18.93	12.32
5	18.34	12.91

ED_002666_00014243-00159

The average mean base period visibility for the 20% most impaired and for the 20% clearest days is obtained by taking the arithmetic mean of these days for the 5 years. Thus, the 5-year mean of the average visibility for the 20% most impaired days is given by

$dv_{base\ period}$ = (19.35 + 20.32 + 19.21 + 18.93 + 18.34) / 5 = **19.23 dv**

The 5-year mean of the average visibility for the 20% clearest days is

$dv_{base\ period}$ = (13.37 + 12.54 + 11.78 + 12.32 + 12.91) / 5 = **12.58 dv**

Step 3. Apply a model to develop component specific RRF's for SO4, NO3, OC, EC, Soil and CM.

Tables 5.5 and 5.6 show the procedure for calculating component-specific relative response factors using an air quality model. The example shows the recommended calculation of RRFs for the 20% most impaired days in the year modeled (the abbreviated example assumes that the 20% most impaired days is represented by only 5 modeled days). The same calculation is repeated for the clearest days.

Table 5.5 Base year modeled species concentrations on the 20% most impaired days (in µg/m3)

Modeled Output	SO_4	NO_3	OC	EC	Soil	CM
Day 1	3.13	1.02	3.56	0.87	0.32	5.43
Day 2	5.21	3.11	4.67	1.23	0.34	4.68
Day 3	4.56	0.59	6.32	1.45	0.87	3.58
Day 4	5.51	1.67	4.56	0.54	1.07	4.67
Day 5	3.99	0.57	4.12	0.69	1.01	5.21
Mean Base Concentration	**4.48**	**1.39**	**4.65**	**0.96**	**0.72**	**4.71**

ED_002666_00014243-00160

Table 5.6 Future year modeled species concentrations on 20% most anthropogenically impaired days (in µg/m3)

Modeled Output	SO_4	NO_3	OC	EC	Soil	CM
Day 1	2.45	0.91	3.54	0.82	0.36	5.51
Day 2	3.98	2.98	4.62	1.15	0.37	4.81
Day 3	3.56	0.50	6.30	1.32	0.92	3.76
Day 4	4.12	1.42	4.51	0.46	1.10	4.76
Day 5	2.87	0.39	4.07	0.61	1.08	5.34
Mean Future Concentration	**3.40**	**1.24**	**4.61**	**0.87**	**0.77**	**4.84**

Table 5.7 Component specific RRFs for the 20% most impaired days

Most Impaired days RRF	SO_4	NO_3	OC	EC	Soil	CM
RRF (mean future/mean base)	0.758	0.891	0.992	0.912	1.061	1.026

Step 4. Multiply the relative response factors times the measured daily species concentration data during the base period to compute future daily species concentrations.

In year 1, we previously identified days 4 and 5 as those included in the 20% most impaired days (i.e., see Step 1). Similarly, days 2 and 8 are the 20% clearest days (see Table 5.3). In this step, we estimate future concentrations for components of $PM_{2.5}$ and for CM for these two sets of days. This is done using information shown in tables presented in Steps 1 and 3 as well as the best days RRFs given in table 5.8 below:

Table 5.8 Component specific RRFs for the 20% clearest days

Clearest days RRF	SO_4	NO_3	OC	EC	Soil	CM
RRF (mean future/mean base)	0.780	0.752	0.956	0.768	1.022	1.043

ED_002666_00014243-00161

Most Anthropogenically Impaired Days

Day 4: $[SO_4]_{future} = (RRF)_{SO4}\ [SO_4]_{base\ period} = (0.758)\ [4.580] = 3.472\ \mu g/m^3$
$[NO_3]_{future} = (0.891)\ [0.528] = 0.470\ \mu g/m^3$
$[OC]_{future} = (0.992)\ [2.482] = 2.462\ \mu g/m^3$
$[EC]_{future} = (0.912)\ [0.282] = 0.257\ \mu g/m^3$
$[Soil]_{future} = (1.061)\ [0.367] = 0.389\ \mu g/m^3$
$[CM]_{future} = (1.026)\ [3.549] = 3.641\ \mu g/m^3$
$[Salt]_{future} = (1.00)\ [0.132] = 0.132\ \mu g/m^3$

Day 5: $[SO_4]_{future} = (0.758)\ [2.742] = 2.079\ \mu g/m^3$
$[NO_3]_{future} = (0.891)\ [0.416] = 0.371\ \mu g/m^3$
$[OC]_{future} = (0.992)\ [3.608] = 3.578\ \mu g/m^3$
$[EC]_{future} = (0.912)\ [0.297] = 0.271\ \mu g/m^3$
$[Soil]_{future} = (1.061)\ [0.343] = 0.364\ \mu g/m^3$
$[CM]_{future} = (1.026)\ [6.252] = 6.414\ \mu g/m^3$
$[Salt]_{future} = (1.00)\ [0.157] = 0.157\ \mu g/m^3$

Clearest Days

Day 2: $[SO_4]_{future} = (0.780)\ [2.290] = 1.786\ \mu g/m^3$
$[NO_3]_{future} = (0.752)\ [0.453] = 0.341\ \mu g/m^3$
$[OC]_{future} = (0.956)\ [1.073] = 1.026\ \mu g/m^3$
$[EC]_{future} = (0.768)\ [0.150] = 0.115\ \mu g/m^3$
$[Soil]_{future} = (1.022)\ [0.058] = 0.059\ \mu g/m^3$
$[CM]_{future} = (1.043)\ [2.004] = 2.090\ \mu g/m^3$
$[Salt]_{future} = (1.00)\ [0.150] = 0.150\ \mu g/m^3$

Day 8: $[SO_4]_{future} = (0.780)\ [1.759] = 1.372\ \mu g/m^3$
$[NO_3]_{future} = (0.752)\ [0.409] = 0.308\ \mu g/m^3$
$[OC]_{future} = (0.956)\ [1.049] = 1.003\ \mu g/m^3$
$[EC]_{future} = (0.768)\ [0.123] = 0.094\ \mu g/m^3$
$[Soil]_{future} = (1.022)\ [0.112] = 0.114\ \mu g/m^3$
$[CM]_{future} = (1.043)\ [3.880] = 4.047\ \mu g/m^3$
$[Salt]_{future} = (1.00)\ [0.725] = 0.725\ \mu g/m^3$

Similar calculations (using the same model derived component specific RRFs) are performed for each of the most impaired days and clearest days in each of the other 4 years in the base

period.[95]

Step 5. Using the results in Step 4, calculate the future year extinction coefficients for each of the 20% most impaired days and each of the 20% clearest days in each of the 5 base years.

Using future PM components obtained in Step 4, we can estimate future daily total b_{ext}.

For year 1

Most Impaired Days

Day 4: b_{ext} = (2.2*3.82*2.869) + (4.8*2.73*0.603) + (2.4*3.82*0.459) + (5.1*2.73*0.011) + (2.8*2.159) + (6.1*0.303) + (10*0.257) + (1*0.389) + (1.7*3.91*0.132) + (0.6*3.641) + 12 = 62.288 Mm^{-1}

Day 5: b_{ext} =(2.2*4.29*1.863) + (4.8*3.00*0.216) + (2.4*4.29*0.364) + (5.1*3.00*0.007) + (2.8*2.938) + (6.1*0.640) + (10*0.271) + (1*0.364) + (1.7*4.20*0.157) + (0.6*6.414) + 12 = 56.722 Mm^{-1}

Clearest Days

Day 2: b_{ext} = (2.2*3.29*1.627) + (4.8*2.46*0.160) + (2.4*3.29*0.335) + (5.1*2.46*0.006) + (2.8*0.973) + (6.1*0.053) + (10*0.115) + (1*0.059) + (1.7*3.49*0.150) + (0.6*2.090) + 12 = 34.786 Mm^{-1}

Day 8: b_{ext} = (2.2*4.59*1.278) + (4.8*3.17*0.094) + (2.4*4.59*0.303) + (5.1*3.17*0.005) + (2.8*953) + (6.1*0.050) + (10*0.094) + (1*0.114) + (1.7*4.44*0.725) + (0.6*4.047) + 12 = 41.682Mm^{-1}

Step 6. Use the daily total b_{ext} values from step 5 to calculate future mean deciview values for the 20% most impaired days and the 20% clearest days.

Calculate daily deciview values (from total daily extinction) and then compute the future mean

[95] Unless multiple meteorological years are being modeled, the same mean RRFs (for each component) are applied to the concentrations on all of the most impaired days for each of the 5 years of data. The mean RRFs are derived from the modeled days. A separate set of RRFs are derived (in the same manner) for the clearest days.

ED_002666_00014243-00163

deciviews for the 20% most impaired days and the 20% clearest days for each year. Then average the 5 years together to get the final future mean deciview value for each set of days.

For Year 1

Most Impaired Days

Day 4: 62.288 Mm^{-1} = 10* ln (62.288/10) = 18.291 dv
Day 5: 56.722 Mm^{-1} = 10* ln (56.722/10) = 17.355 dv

Future mean visibility on most anthropogenically impaired days = (18.291 + 17.355) / 2 = 17.82 dv

Clearest Days

Day 2: 34.786 Mm^{-1} = 10* ln (34.786/10) = 12.466 dv
Day 8: 41.682 Mm^{-1} = 10* ln (41.682/10) = 14.274 dv

Future mean visibility on clearest days = (12.466 + 14.274) / 2 = 13.37 dv

Similar calculations are performed for previously selected 20% most impaired days and the 20% clearest days in each of years 2-5. Assume these other calculations yield the following estimates for future mean dv on these sets of days.

Table 5.9 Future year mean deciviews on the 20% most impaired days and the 20% clearest days for each of the 5 years

Year	Future Mean dv on 20% Most Impaired Days	Future Mean dv on 20% Clearest Days
1	17.82	13.37
2	18.17	12.45
3	16.78	12.91
4	18.65	11.52
5	18.21	12.31

ED_002666_00014243-00164

Using results in table 5.9, we see that the estimated future average mean visibility for the most impaired days is
dv_{future} = (17.82 + 18.17 + 16.78 + 18.65 + 18.21) / 5 = **17.92 dv**

The estimated future average mean visibility for the clearest days is
dv_{future} = (13.37 + 12.45 + 12.91 + 11.52 + 12.31) / 5 = **12.51 dv**

The most impaired days result generated in step 6 can then be compared to the point on the Class I area-specific glidepath for the end of the implementation period. The most impaired days result is also compared to the baseline most impaired days to ensure that there has been improvement relative to the baseline period. The future year clearest days result is compared to the baseline (2000-2004) clearest days to ensure that visibility on the best days is not forecast to degrade relative to the same period.

This information is used to determine the predicted improvement in visibility that will result from implementation of an emissions strategy. It cannot itself determine the necessary content of the long-term strategy. See the Regional Haze Rule and related guidance documents for more information on the reasonable progress requirements.

5.4 How Do I Select Appropriate Inputs For Developing RPGs?

In section 5.3, we described the recommended regional modeling analysis to develop RPGs and to use them to evaluate progress relative to the glidepath and the baseline period. An important part of the analysis requires using component-specific RRFs, obtained with models, to estimate future concentrations of these components and, subsequently, future visibility. In this subsection, we address more details concerning the calculation of visibility RRFs. Additionally, there are several assumptions inherent in the recommended RPG analysis; we identify these assumptions and comment on their underlying rationale. More specifically, we address seven issues:

1. How to estimate base period air quality in a Class I area without monitored data;
2. How to handle a base year without data or with a small sample size;
3. Use of the same days to estimate changes in visibility for the most anthropogenically impaired days and a different set of same days to estimate changes in visibility for the clearest days;
4. Which predictions to use to derive RRFs;

ED_002666_00014243-00165

5. How many and what kind of days to use to develop RRF values;
6. RRFs when modeled species concentration are very low; and
7. Alternative RRF calculations.

Estimating baseline and base period visibility in a Class I area without monitors

There are 156 officially designated Class I areas subject to the Regional Haze Rule and 110 IMPROVE sites in or near Class I areas. Therefore, 46 Class I areas do not have co-located IMPROVE monitors. EPA's Tracking Guidance recommends using IMPROVE data from a nearby site to represent the visibility at each Class I area that does not have ambient data. Table A-2 in Appendix A of the Tracking Guidance lists the recommended monthly f(rh) values[96] for each Class I area as well as the representative site for each Class area. The representative IMPROVE site data will be used to track regional haze progress for the Class I areas. Therefore, it follows that visibility improvement for tracking purposes should be predicted at the location of the IMPROVE monitor, not at the actual Class I area. For the purposes of deriving ambient data for modeling, we recommend following the same representative site assignments contained in the Tracking Guidance. In this way, the needed visibility values can be derived for each Class I area from the network of 110 IMPROVE sites.[97] Similarly, the modeling results should be extracted for the location of the representative monitor, not the actual location of the Class I area.[98]

Considering a year in the base period with little or no monitored particulate matter or missing data

The Tracking Guidance recommends calculating base period visibility values for sites with at least 3 out of 5 complete years of data. It further contains recommendations for determining if a year has complete data. In general, a site should have 50% data completeness in all quarters and meet a 75% completeness criteria for the full year. There should be no periods with more than 30 consecutive days without data. The Tracking Guidance assumes that all IMPROVE sites

[96] The f(rh) values in the Tracking Guidance were developed for the "original" IMPROVE algorithm and should not be used to calculate visibility using the "revised" IMPROVE algorithm.
[97] Bering Sea Wilderness (Alaska) is the only Class I area that has no IMPROVE monitor **and** no representative IMPROVE site. On-site IMPROVE or representative IMPROVE data can be found for the other 155 Class I areas.
[98] There may be other reasons to calculate visibility improvement at the actual Class I area. The SMAT provides options for calculating visibility changes at both the Class I area grid cell center and at the representative IMPROVE monitoring site. It may be informative to compare model response at the representative IMPROVE site versus the grid cell center of the Class I area. Large differences in the model response may indicate that the "representative" site may not adequately represent visibility at the Class I area or across the entire Class I area.

ED_002666_00014243-00166

will have at least 3 complete years in the base period (which may or may not be true). The Tracking Guidance also contains procedures for filling in missing data.

There are several data completeness issues that may cause problems within the visibility modeling analysis. First, a site may have less than 3 years of complete data. It is possible to calculate visibility improvement based on as little as 1 year of ambient data. Additionally, the IMPROVE program has procedures to utilize substituted and/or patched data to help minimize data incompleteness. States should work with their EPA Regional office and Federal Land Managers to determine how to estimate baseline visibility for these area(s).

Another issue that is a more specific problem for modeling analyses occurs when data is missing during the meteorological time period that was modeled. It is likely that most air agencies will only be modeling a single year of meteorology and it is possible that the ambient data at one or more Class I areas is incomplete during that year. Without ambient data, it is impossible to identify the 20% most anthropogenically impaired days and 20% clearest days, which are used to calculate modeled RRFs.

Again, if this occurs, states should work with their Regional office and Federal Land Managers to determine the best way to calculate visibility on these days. Potential options are to utilize substituted and/or patched data, data from another nearby IMPROVE site, nearby data from a different ambient data network, or interpolate ambient data to the site.

Using a constant sample of days to estimate base period and future visibility

For a typical Class I area, there will be about 120 days per year having measurements needed to estimate $(dv)_{base\ period}$ with Equation (5.1). Thus, there should be about 24 most anthropogenically impaired and 24 clearest days for each of the 5 years in the base period. It is conceivable that the identity of these days could change if emissions were altered to reflect net effects of controls and growth. The recommended test described in section 4.8 assumes that the identity of the most impaired and clearest days remains unchanged. This is done primarily to avoid having to perform iterative analyses to identify future most impaired and clearest visibility days and to keep the test relatively simple and more readily understood. This assumption could cause improvement in visibility to be overestimated for the most anthropogenically impaired days and could also cause the test to overestimate the difficulty in preventing deterioration of visibility on the clearest days. However, for the reasons described below, we do not believe the effects of this assumption are substantial.

ED_002666_00014243-00167

It is unlikely that there would be any wholesale change in the identity of most anthropogenically impaired or clearest days with future versus base year emissions. Analyses performed by Meyer, et al. (1997) have shown that the predicted ranked severity of high ozone days is largely unaffected by simulated controls and growth (i.e., highest days tend to remain the highest days after the effects of growth and controls are simulated). There is no reason to expect a different outcome for other secondary pollutants. If there are differences, we would expect these to occur near the borderline between the most impaired days and more moderate days.

Because the visibility analysis relies on *mean* visibility values on 20 or more most impaired visibility days and most of these days are unlikely to change, we would expect little difference in the outcome of the analysis. Further, because of the shape of the distribution of daily extinction coefficients, the mean of the most impaired days is often more heavily influenced by the very most impaired days rather than those on the borderline between 20% most impaired and more moderate light extinction. There could be differences in some of the clearest visibility days corresponding with pre- and post-control emissions. However, because the differences in concentrations of particulate matter on such days are likely to be relatively low, differences in the computed mean visibility for clearest days are likely to be small. It should be noted that any resulting difference in the projection of RPGs for clearest days would likely be in the direction of it being more difficult to pass the no degradation test compared to the baseline 2000-2004 period, in that some of the future days selected may be higher deciview-valued than they ought to be. If our recommended procedure leads to suspected problems in the outcome of a test, an air agency may wish to perform a more rigorous version of the analysis (in which the identity of pre-control and post-control days changes) as part of additional or supplemental analyses to support the RPG.

Selecting predictions to use in deriving RRF

RRFs should be developed for each Class I area. When a Class I area contains a monitoring site, the RRF estimates should be derived using predictions which are made "near" that site. For each day, daily average surface predictions of each component of PM made near a monitor should be estimated. Similar to the annual $PM_{2.5}$ NAAQS attainment test, nearby grid cells should be averaged. These nearby estimates should be spatially averaged to estimate a spatially representative daily concentration. For 12 km or finer resolution, we recommend averaging the modeled concentrations for a 3 x 3 array of grid cells surrounding the monitor. Note that for cells larger than 12 km on a side, no spatial averaging is necessary—states should just use the

ED_002666_00014243-00168

prediction in the cell containing the monitor. Spatially representative daily concentrations obtained for each modeled day with monitored data should then be temporally averaged. This final average should be used to compute the RRF. Thus, component-specific RRF values for a Class I area with a monitor are the ratio of the temporally averaged spatial mean of nearby concentrations predicted with future emissions to that predicted with base year emissions.

Selecting days to derive RRF values

It may often happen that a planning organization or a group of states decides to model the visibility impacts of a strategy for numerous Class I areas simultaneously. As we note in section 2.3.1, this may make it advisable to simulate (at least) a full year so that RRF values for each Class I area are based on a substantial number of observed best and worst days. For the most impaired days in the chosen year, the RRF for a component of PM should be estimated as the ratio of its modeled arithmetic mean predicted value on the 20% most impaired days with future emissions to that with base year emissions. Thus, in most cases, the RRF should reflect values averaged over ~ 24 most impaired days in that year. The same procedure is followed to derive RRFs over the ~24 clearest days in the year.

Since meteorological conditions and/or emissions may be markedly different on best visibility versus worst visibility days, we recommend calculation of a separate set of RRF values for the 20% clearest days. As with most impaired days, the preferred approach is to model an entire year and select an RRF value for concentrations averaged over the 20% clearest days for each Class I area. The RRF values are the ratios of the future to base year modeled averages. The appropriate RRF values should then be used in concert with each observed best day to estimate future concentrations for each component on each identified clearest day.

RRFs when modeled species concentrations are very low

In most cases, the RRFs derived for visibility components are either less than 1.0 or slightly more than 1.0. However, large RRFs (much greater than 1.0) can occur when the mean modeled concentration of a PM component on the 20% most impaired or clearest days is very small and the concentration of that component increases in the future year case. This most often occurs with the nitrate component. On both the 20% most impaired days and 20% clearest days, the modeled nitrate concentrations at Class I areas can be very small (i.e., < 0.01 μg/m3). These low modeled concentrations can sometimes increase in future year scenarios by orders of magnitude. This occurs most often in ammonia limited areas when sulfate concentrations are reduced in the future (due to SO_2 emissions reductions) causing nitrate

ED_002666_00014243-00169

increases due to an increase in available ammonia. Large RRFs should be closely examined to identify the cause of the issue. In many cases, large RRFs are caused by poor model performance (i.e., underpredicted base case nitrate concentrations) or instability in one or more model components (i.e., the nitrate partitioning model). One solution is to cap the RRF at 1.0 or close to 1 in cases where nitrate (or any other $PM_{2.5}$ component) increases in the future by an unrealistic amount due to very low modeled concentrations. Any adjustments to the default calculations should be documented and discussed with the appropriate EPA Regional office and the Federal Land Managers.

Alternative RRF calculations

The default analysis is relatively simple in that a single mean RRF is calculated for each PM component (separate RRFs on most impaired days and clearest days). A series of tests with more complicated methods has shown that: 1) the difference between various versions of the test are usually small and 2) each of the alternative tests has limitations in its applicability (Environ, 2005). Possible variations include the use of day specific RRFs, season (or quarter) specific RRFs, or climatologically based RRFs. In some cases, these more complicated techniques may provide different answers, but sometimes not. There are specific limitations with each of these alternatives noted in the indicated reference. We have chosen to keep the single mean RRF test as the default recommendation. States are encouraged to explore other methods for estimating RRFs if it is thought that the default recommendation is too simplistic to accurately capture the change in future visibility at any particular Class I area. The SIP analysis to develop RPGs should use the most appropriate method of projecting future concentrations for the characteristics of each Class I area.[99] Alternative methods should be discussed in the modeling protocol and discussed with the appropriate EPA Regional office and Federal Land Managers.

[99] In particular, issues may arise when dealing with visibility contributions from fires, coarse mass and fine soil (from wind-blown dust), and international transport (and possibly other issues). Each of these issues should be addressed in the modeling protocol and solutions should be discussed with the appropriate EPA Regional office(s) and Federal Land Managers on a case-by-case basis.

ED_002666_00014243-00170

6.0 How Can Additional Analyses Be Used to Support an Ozone or $PM_{2.5}$ Attainment Demonstration?

By definition, models are simplistic approximations of complex phenomena. The modeling analyses used to assess whether emission reduction measures will bring an individual area into attainment for the NAAQS contain many elements that are uncertain (e.g., emission projections, meteorological inputs, science formulations, etc.). These uncertain aspects of the analyses prevent definitive assessments of future attainment status. The confidence in the representativeness of the quantitative results from a modeled attainment test should be a function of the degree to which the uncertainties in the analysis were minimized. In general, by following the recommendations contained within this guidance document, EPA hopes that the attainment demonstrations will mitigate the uncertainty as much as possible given the current state of modeling inputs, procedures, and science. However, while air quality models represent the best tools for integrating emissions and meteorological information with atmospheric chemistry, and no single additional analysis can match the expected reliability of these models' results, EPA believes that all attainment demonstrations will be strengthened by additional analyses that can supplement the modeling to enhance the assessment of whether the planned emissions reductions are likely to result in attainment.

Supplemental evidence should accompany all model attainment demonstrations. Generally, those modeling analyses that show that attainment will be reached in the future with some margin of safety will need more limited supporting material. For other attainment cases in which the projected future design value is closer to the NAAQS, more rigorous supporting analyses should be completed. There may be some areas for which the supplemental evidence is persuasive enough to support a conclusion that the area can expect to achieve timely attainment despite failing the modeled attainment test, and other areas for which the modeled attainment test demonstrates attainment, but the supplemental evidence casts significant doubt on that result. This section of the guidance will discuss some specific information and additional analyses that can be used to supplement the model projections. Of particular interest are analyses that help determine whether the modeling-based projections are likely to provide a prediction of the *air quality improvement* that is likely to occur by the attainment date. Air agencies should review these supplemental analyses, in combination with the modeling analysis, in a "weight of evidence" assessment of whether each area is likely to achieve timely attainment. Again, it should be noted that no single supplemental analysis can serve as an adequate substitute for the air quality model. However, in aggregate, supplemental

ED_002666_00014243-00171

analyses may provide information which may provide further support for the outcome of the modeled test or may indicate a different outcome than the modeled test.

In considering the use of supplemental analyses, an important factor is the time (number of years) until the attainment date. In general, modeling and related analyses are the most useful corroborative analyses for areas with attainment dates which are more than several years in the future. In contrast, ambient data and emissions trends become more important (and hence model results become less important) the closer in time the area is to its attainment date. For example, if an area is only 1 or 2 years away from its attainment date, ambient data is in most cases the best predictor of likely air quality levels in the near future. However, this may not be true if there are large emissions reductions that are expected to occur in the near-term period. In that case, modeling is needed to estimate additional ozone and/or PM reductions that are likely to occur due to the emissions controls in the short term. Similarly, if the attainment date is 5 or 10 years or more in the future, appropriate modeling will likely be the most reliable indicator of future year ozone or PM levels. Emission changes over a relatively long period are likely to overwhelm the influence of ambient data trends and meteorological variability, making ambient data and emissions trends analyses relatively less important.

6.1 What Types of Additional Analyses Should Be Completed as Part of an Ozone or $PM_{2.5}$ Attainment Demonstration?

There are three basic types of analyses that are recommended to supplement the primary modeling analysis. They are:

1) Additional modeling analyses;

2) Analyses of trends in ambient air quality and emissions;

3) Additional emissions controls/reductions.

Note that air agencies are encouraged to consult with their EPA Regional office (and Federal Land Managers, when appropriate) in advance of initiating supplemental analyses to determine which additional analyses may be most appropriate for their particular area.

ED_002666_00014243-00172

6.1.1 Modeling Analyses

The relative attainment tests described in sections 4.2, 4.4, and 4.5 are the primary modeling tools used in an attainment demonstration. The application of a chemical transport grid model on a regional or local scale is the best tool available to judge the impacts of changes in future year emissions on concentrations. In addition to this "primary" modeling analysis, there are various other models, model applications, and tools that can be used to supplement the results of the modeled attainment test. These include, but are not limited to:

- Available regional or national scale modeling applications that are suitable[100] for the local area, for example, modeling in support of EPA rulemakings or regional, multi-jurisdictional organization modeling that may be available for the appropriate future year of interest. Modeling analyses may be available that used different models and/or inputs.

- Use of other appropriate local modeling that includes the nonattainment area of interest. This may include applications using alternative models and/or inputs or research-oriented analyses.

- Use of photochemical source apportionment, DDM, and/or process analysis modeling tools to help explain why attainment is (or is not) demonstrated.

- Use of multiple air quality models / model input data sets (e.g., multiple meteorological data sets, alternative chemical mechanisms or emissions inventories, etc.). Multiple model configurations can be used to estimate sensitivity and uncertainty of future year design value predictions.
    - For results to be most relevant to the way we recommend models be applied in attainment demonstrations, it is preferable that such procedures focus on the sensitivity of estimated RRFs and resulting projected design values to the variations in inputs and/or model formulations.

- Application of the attainment test with alternative procedures compared to the default recommendations in sections 4.2, 4.4, and 4.5 of this guidance. Any alternate approaches

[100] The resolution, emissions, meteorology, and other model inputs should be evaluated for applicability to the local nonattainment area. Additionally, model performance for the local nonattainment area should be examined before determining whether the regional model results are suitable for use in the local attainment demonstration.

ED_002666_00014243-00173

should be accompanied with a technical justification to explain why the approach is appropriate for the area in question and should be discussed with the appropriate EPA Regional office.

  - Alternative RRF calculations using non-default assumptions or more detailed temporal and spatial analysis of RRFs at one or more monitoring locations. If a sound technical argument can be made for why atypically high RRFs at any particular location are not reasonable, then these types of supplemental analyses would suggest that attainment is more likely to be achieved than the default modeling analysis alone would indicate.
  - Alternate base year design values that may differ from the 5-year weighted average value. Alternative values could include different methodologies for determining base year design values or removal of potential exceptional events data (see section 4.1.1).

- Use of dispersion models to address primary $PM_{2.5}$ contributions to $PM_{2.5}$ concentrations. In areas with large spatial gradients of primary $PM_{2.5}$, dispersion models are best suited to characterizing the change in primary $PM_{2.5}$ in the future. A local area analysis may be useful as a supplemental analysis (in either monitored or unmonitored areas, as appropriate) for areas that at least partially rely on local primary PM controls to reach attainment and did not otherwise perform and submit a local area analysis of part of the attainment demonstration.

The EPA has determined that the best approach to using models to demonstrate attainment of the NAAQS is to use a model in a relative mode. However, some types of "absolute" model results may be used to assess general progress towards attainment from the baseline inventory to the projected future inventory.[101] Example metrics include:

- Percent change in total amount of ozone or $PM_{2.5}$ >= NAAQS[102] within the nonattainment area
- Percent change in number of grid cells >= NAAQS within the nonattainment area
- Percent change in grid cell-hours (days) >= NAAQS within the nonattainment area
- Percent change in maximum modeled 8-hour ozone within the nonattainment area

[101] Care should be taken in interpreting absolute metrics if the model evaluation shows a large underprediction or overprediction of ozone or $PM_{2.5}$ concentrations. An underprediction of observed concentrations will make it artificially easy to show progress towards absolute attainment levels and an overprediction will make it artificially difficult to show progress towards attainment.

[102] For each of these metrics, the appropriate comparison to the level of the NAAQS is 70 ppb for 8-hour ozone; 35 µg/m3 for 24-hour $PM_{2.5}$; and 12 µg/m3 for annual $PM_{2.5}$.

ED_002666_00014243-00174

While these metrics can be used to estimate the magnitude, frequency, and relative amount of ozone or $PM_{2.5}$ reductions from any given future emissions scenario, there are no threshold quantities of these metrics that can directly translate to an attainment determination. Generally, a large reduction in the frequency, magnitude, and relative amount of 8-hour ozone nonattainment (i.e., >= 71 ppb) or $PM_{2.5}$ nonattainment (24-hour and/or annual) is consistent with a conclusion that a proposed strategy would meet the NAAQS. In the context of a weight of evidence determination, these metrics could be used to suggest that a particular location may be "stiff" or relatively unresponsive to emissions controls, while the rest of the modeling domain/nonattainment area is projected to experience widespread reductions.

6.1.2 Analyses and Trends in Ambient Air Quality and Emissions

Generally, air quality models are regarded as the most appropriate tools for assessing the expected impacts of a change in emissions. However, it may also be possible to evaluate progress towards attainment of the ozone or $PM_{2.5}$ NAAQS based on measured historical trends of air quality and emissions. It may be possible to develop a relationship between past emissions changes and historical and current air quality. Once the relationship between past/present emissions and air quality is established, a response to the expected emissions reductions from a particular control strategy can be estimated. There are several elements to this analysis that are difficult to quantify. First, in most cases, the ambient data trends are best assessed by normalizing to account for year-to-year meteorological variations. Second, one must have an accurate accounting of the year-to-year changes in actual emissions (NOx, VOC, and/or SO_2 and NH_3) for the given area and any surrounding areas whose emissions may impact local concentrations. Third, one must have a solid conceptual model of how ozone or $PM_{2.5}$ is formed in the local area (e.g., influence of meteorology, NOx-limited, ammonia limited, transport-influenced, etc.).

If available, meteorologically adjusted ozone and $PM_{2.5}$ concentrations can be used to establish air quality trends. There are several techniques that have been used to examine the influence of meteorology on air quality. Among them are (a) statistical modeling (U.S. EPA, 2005b); (b) filtering techniques (Rao, 1995; Flaum, 1996; Milanchus, 1998; and Hogrefe, 2000), (c) using a probability distribution of meteorological severity based on climatological data (Cox and Chu, 1993, 1996), and (d) using CART analysis to identify meteorological classes and selecting days from each year so that the underlying frequency of the identified meteorological classes remains the same (Deuel and Douglas, 1996). Most of this work has examined the relationship between ozone and meteorology. Only recently have analyses examined the relationship between meteorology and $PM_{2.5}$. Additionally, compared to $PM_{2.5}$, the established relationship

ED_002666_00014243-00175

between ozone and meteorological variables is generally stronger (higher r-square values). In the case of $PM_{2.5}$, the relationship between concentration and meteorology is complicated by the fact that $PM_{2.5}$ components experience high concentrations at different times of the year and for different reasons. This makes it more difficult to meteorologically adjust $PM_{2.5}$ concentrations.

If a meteorologically adjusted trend in ozone or $PM_{2.5}$ can be estimated, then the information can be used to establish a link between emissions and air quality trends. This is not always straightforward due to the multitude of emissions precursors that may lead to high ozone and $PM_{2.5}$ concentrations. A careful analysis of (meteorologically adjusted) air quality trends and emissions trends of each of the ozone and PM precursors (as well as primary PM) is needed to fully establish relationships. Detailed emissions information as well as a solid understanding of the conceptual model of ozone or $PM_{2.5}$ formation is needed. If a trend can be established based on past emissions changes and air quality changes, then future year predicted emissions levels can be used to extrapolate future air quality.

Meteorologically adjusted trends in ozone and/or $PM_{2.5}$ attempt to adjust observed concentrations to levels which would have been observed during average meteorological conditions (Camalier, 2007). This makes it easier to observe trends primarily due to emissions changes and also allows identification of years with above average or below average meteorologically conducive conditions. Identification of "extreme" meteorological conditions, which are not likely to re-occur in a particular design value period, can be used as part of a weight of evidence analysis. For example, the presence of one or more "extreme" (high concentration) meteorological years in either the base year period or the most recent design value period may make it appear that an area is farther away from attaining the NAAQS than would otherwise be the case if more typical meteorological conditions had occurred. It is important to note that 1 or more years of meteorological conditions which are not conducive to ozone formation will have the opposite effect. An area may appear to be on track to attain the NAAQS (or close to attaining) but, in reality, may need substantial additional emissions reductions in order to attain under average or above average meteorological conditions. Meteorological adjusted concentration and/or design value calculations can help explain both of these situations. More information on meteorologically adjusted ozone trends can be found here.

A simpler (and more uncertain) way to qualitatively assess progress toward attainment is to examine recently observed air quality and emissions trends. Downward trends in observed concentrations and in emissions (past and projected) are consistent with progress towards

ED_002666_00014243-00176

attainment. Strength of the evidence produced by emissions and air quality trends is increased if an extensive monitoring network exists and if there is a strong, positive and demonstrable, correlation between past emissions reductions and current trends in ozone or $PM_{2.5}$. Until recently, EPA prepared annual air quality trends reports for all criteria pollutants. The latest trends report analyzes ambient data through 2017 and can be found here.

The appropriate weight given to trend analyses depends on several factors. Analyses that use more air quality data and apply a greater variety of trend parameters typically provide more credible results. More weight can be placed on the results if the procedure used to normalize the trend for meteorological differences explains much of the variability attributable to these differences. In addition, trend analysis is usually more reliable if the extrapolation (as applicable) does not extend very far into the future. Finally, trend analysis is most credible if the contemplated strategy is similar to a past strategy (e.g., both strategies focus on reducing sulfates for PM or NOx for ozone). For example, if a past strategy focused on reducing sulfates, but a future one envisions controlling OC, there is no guarantee that ambient OC will respond similarly to changes in past emissions.

In addition to ambient data trends, further analysis of ambient data can provide information on ozone production efficiency, evidence of transport and assist in the quality assurance of emissions inputs. In particular, photochemical assessment monitoring station (PAMS) data, special study data, research and non-routine ambient data (i.e., NOy, and VOC data) can provide insights into the ozone and/or $PM_{2.5}$ concentrations observed in the nonattainment area and can help assess the accuracy of the modeling results and control strategies that are being relied upon for attainment.

6.1.3 Additional Emissions Controls/Reductions

Models are used to predict the future expected concentrations of ozone and/or $PM_{2.5}$, based on modeled emissions changes between a base year and future year. It is expected that emission inventories are as accurate as possible and represent emissions changes (including growth and controls) that are expected to occur. However, there may be various emissions sources and/or controls that are difficult to accurately represent in the modeling analysis, and the effects of some emissions controls may be difficult to quantify. Additionally, there may be uncertainty in how emissions controls may be implemented either in a nonattainment area, or in upwind regions which may contribute ozone to the nonattainment area.

ED_002666_00014243-00177

The following are some of the types of control measures that may be appropriate to include in a weight of evidence demonstration:

- Measures that are difficult to quantify or may not be enforceable in the SIP. Examples include:
    - Energy efficiency or renewable energy (EE/RE) programs may be implemented statewide (and in fact may be enforceable). However, emissions reductions from such programs are hard to quantify and it may be difficult to assign reductions to particular utility sources. EPA has provided an "EE/RE Roadmap Manual." The roadmap includes a weight of evidence "pathway" for air agencies that want to acknowledge the emissions benefits and potential reductions from EE/RE measures, but are unable to accurately quantify the emissions reductions in their SIP.
    - Smart growth initiatives such as land use planning and transportation planning.
    - Truck stop electrification.
    - Emissions reductions from idling regulations which may not be quantified in the mobile emissions modeling.
- Voluntary measures
    - Air agencies may have implemented voluntary measures that are not enforceable and therefore are not included in the SIP. These could include measures such as "ozone action days," voluntary no burn days, telework programs, idling reduction initiatives, etc. Even though these programs may be voluntary, they can still lead to positive actions in the nonattainment area that can lead to lower emissions. Voluntary measures should be documented to the fullest extent possible, including estimates of emissions benefits, where possible.

- Regional/super-regional and/or national programs that may not have been accounted for in the attainment demonstration.
    - Federal measures which were not accounted for in the modeling because they were proposed and/or finalized after the state modeling was completed.
    - Upwind regional, state, and/or local measures which were not known or not able to be quantified in the attainment demonstration due to timing or lack of information.

6.2 Weight of Evidence Summary

A weight of evidence (WOE) determination may examine results from a diverse set of analyses, including the outcome of the primary attainment test, and attempts to summarize the results into an aggregate conclusion with respect to whether a chosen set of control measures are

ED_002666_00014243-00178

likely to result in an area attaining the NAAQS by the applicable attainment date, notwithstanding what the air quality modeling may suggest in isolation. The supplemental analyses discussed above can be part of a WOE determination, although the level of detail required in a WOE submittal will vary as a function of many elements of the model application (e.g., model performance, degree of residual nonattainment in the modeled attainment test, amount of uncertainty in the model and its inputs, etc.). Each WOE determination will be subject to area-specific conditions and data availability. Area-specific factors may also affect the types of analyses that are feasible for a specific nonattainment area, as well as the significance of each. Thus, decisions concerning which analyses to perform, and how much credence to give each, need to be made on a case-by-case basis by those implementing the modeling/analysis protocol. Air agencies are encouraged to consult with their EPA Regional office (and Federal Land Managers, when appropriate) in advance of initiating supplemental analyses to determine which additional analyses may be most appropriate for their particular area for potential use as part of a WOE evaluation.

The most useful supplemental analyses are those providing the best evidence as to how much air quality improvement can be expected as compared to the improvement projected by the air quality modeling analysis. Each analysis is weighed qualitatively, depending on: 1) the capacity of the analysis to address the adequacy of a strategy and 2) the technical credibility of the analysis. If the overall WOE produced by the combination of the primary modeling analysis and the various supplemental analyses supports the attainment hypothesis, then the air agency can demonstrate attainment of the NAAQS proposed control strategy. The end product of a WOE determination is a document which describes analyses performed, databases used, key assumptions and outcomes of each analysis, and why an air agency believes that the evidence, viewed as a whole, supports a conclusion that the area will, or will not, attain the NAAQS. In conclusion, the basic criteria required for an attainment demonstration based on weight of evidence are as follows:

1) A fully-evaluated, high-quality modeling analysis that projects future values that are close to the NAAQS.
2) A description and explanation of each of the individual supplemental analyses, preferably from multiple categories. Analyses that utilize well-established analytical procedures and are grounded with sufficient data should be weighted accordingly higher.
3) A written description as to why the full set of evidence leads to a conclusive determination regarding the future attainment status of the area that differs from the results of the modeled attainment test alone.

ED_002666_00014243-00179

7.0 What Role Should Additional and/or Supplemental Analyses Play in Regional Haze Modeling Demonstrations?

We believe additional and/or supplemental analyses can be a useful tool for air agencies when developing RPGs and applying them in the ways described in section 5. In this subsection, we note some potential supplemental analyses that may be used in this context.

7.1 Additional air quality modeling

Sensitivity tests can be performed to see if conclusions about trends in the 20% most anthropogenically impaired days and 20% clearest days are robust. One example of such an analysis is applying a model with and without a more finely resolved nested grid near source areas or near Class I areas. A purpose of this would be to see whether conclusions are affected by the degree of detail in which nearby sources are considered. A second example of an analysis would be to consider alternative future emissions (including lateral boundary conditions) and/or differing growth rate assumptions. This may be a particular concern for regional haze analyses because the emissions projection period is generally longer than for most ozone and $PM_{2.5}$ attainment demonstrations. Uncertainty in emissions and growth rates become more important as the projection period is lengthened.

If the outcomes of the several comparisons described in section 5.1 are similar using sensitivity tests, alternative models and/or alternative modeling approaches, this finding supports conclusions reached in the main analysis.

7.2 Review of trends

A review of trends generally involves a comparison, sometimes qualitative, between past trends in reconstructed visibility and estimated changes in emissions (e.g., mid '00's to mid '10's). This information could be used to confirm that additional reductions of previously controlled emissions of a component or its precursors is likely to result in the predicted visibility improvement. It may also be used to see whether certain PM components are becoming increasingly more or less important sources of light extinction.

7.3 Other models and tools

Trajectory models may be useful for identifying the types of meteorological conditions most often corresponding to observed worst and best visibility in various Class I areas. This, in turn,

ED_002666_00014243-00180

may enable air agencies to draw inferences about the areas containing sources most likely to influence visibility in a Class I area on days with "poor" and "good" visibility. Grid model-based techniques such as DDM or source apportionment may also be useful in identifying areas and emissions sources most responsible for visibility impairment on the most anthropogenically impaired days and clearest days.

7.4 Refinements to the recommended visibility analysis

A state may consider refining the recommended modeling analysis in some manner. Refinements are best made if they are based on local observations/analyses which suggest that some of the underlying assumptions in the recommended assessment may not be applicable. We list some potential refinements that could be considered. The list is intended to illustrate types of additional analyses that could be performed.

- Use an alternative light extinction equation or an area-specific version.
- The 20% most anthropogenically impaired days conceptually will not include days with high impacts from fires or dust storms because anthropogenic impairment relative to natural conditions will be low on such days. However, the particular method initially applied to identify the most anthropogenically impaired days may not implement this concept perfectly at a particular Class I area. Available speciated data and other information may be reviewed to see whether the outcome of the visibility analysis is being influenced by including one or more days with extraordinary events (e.g., a major wildfire lasting a number of days or transported dust events). If convincing arguments can be made that the event is a "natural" one, changes in the initial approach to estimate daily anthropogenic versus natural light extinction should be discussed with the appropriate EPA Regional office and Federal Land Managers.
- Daily component specific RRFs can be examined to determine if one or more days are responding in a different way compared to the majority of the clearest or most impaired days. Determining why days may be more or less responsive to emissions controls may lead to conclusions regarding the suitability of particular days to be represented in the mean response. It may be appropriate, in some cases, to re-calculate mean RRFs with suspect days removed.
- Re-rank future estimated light extinction (i.e., b_{ext} values) for all days that have current measurements and re-compute mean future most impaired and clearest days visibility (i.e., do not assume that the identity of clearest days remains the same between the base year and the future year).

ED_002666_00014243-00181

7.5 Concerns about modeling the clearest days

In some parts of the U.S., concentrations of the components of PM used in visibility calculations may be very close to background levels on the clearest days. Measurements and model estimates may be subject to more relative uncertainty (i.e., lower signal to noise ratio) on days where observed concentrations of PM are very low (and light extinction is also low). The utility of a WOE determination is heightened in such cases. An air agency should see whether a model's inability to accurately predict one or more individual components of PM has a substantial effect on the extinction coefficient calculated with Equation (5.1). If it does, and diagnostic tests are unable to resolve a performance problem, an air agency may need to address the "no degradation" requirement for the clearest days in the particular Class I area(s) without using results from a grid model.

ED_002666_00014243-00182

8.0 References

Abt, 2014. Modeled Attainment Test Software: User's Manual. MATS available at: https://www.epa.gov/scram/photochemical-modeling-tools

Adelman, Z., M. Fry, J.J. West, P. Dolwick, and C. Jang (2011): Background air quality in the United States under current and future emissions scenarios, 10th Annual CMAS conference, Chapel Hill NC, October 2011.

Allen, D.J., Pickering, K.E., Pinder, R.W., Henderson, B.H., Appel, K.W., and Prados, A., (2012), Impact of lightning-NO on eastern United States photochemistry during the summer of 2006 as determined using the CMAQ model. Atmospheric Chemistry and Physics. 12: 1737-1758.

Anenberg S.C., K. Talgo, S. Arunachalam, P. Dolwick, C. Jang, and J.J. West, (2011), Impacts of global, regional, and sectoral black carbon emission reductions on surface air quality and human mortality, Atmos. Chem. Phys., 11, 7253-7267.

Ansari, A.S., and S.N. Pandis, (1998), Response of Inorganic PM to Precursor Concentrations, *Environ. Sci. Technol.*, **32**, 2706-2714.

Appel, K. W., A. B. Gilliand, G. Sarwar, and R.C. Gilliam, (2007), Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: sensitivities impacting model performance, *Atmos. Environ.*, **41**, 9603–9615.

Appel, W., S. J. Roselle, R. C. Gilliam, and J. E. Pleim, (2010), Sensitivity of the Community Multiscale Air Quality (CMAQ) model v4.7 results for the Eastern United States to MM5 and WRF meteorological drivers, Geoscientific Model Development 3, pp. 169-188.

Appel, K.W., Gilliam, R.C., Davis, N., Zubrow, A., Howard, S.C., (2011), Overview of the atmospheric model evaluation tool (AMET) v1.1 for evaluating meteorological and air quality models. *Environmental Modeling & Software*, 26(4), 434-443.

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., 37, 1185-1198.

Arnold, J. R., and R. L. Dennis (2006), Testing CMAQ chemistry sensitivities in base case and emissions control runs at SEARCH and SOS99 surface sites in the southeastern US, *Atmos. Environ.*, *40*(26), 5027-5040.

ED_002666_00014243-00183

Arunachalam, S., Adelman, Z., Mathur, R., and Vukovich, J.M, (2002), The Impacts of Biogenic Emissions Estimates from BEIS-3 on Ozone Modeling in the Southeastern U.S., presented at the 11th Annual International Emissions Inventory Conference, Atlanta, GA, April 15-19, 2002, https://www.researchgate.net/publication/248693728_The_Impacts_of_Biogenic_Emissions_Estimates_from_BEIS-3_on_Ozone_Modeling_in_the_Southeastern_US

Arunachalam, S., A. Holland, B. Do, and M. Abraczinskas, (2006), A Quantitative Assessment of the Influence of Grid Resolution on Predictions of Future-Year Air Quality In North Carolina, USA, *Atmospheric Environment*, **40** (26), 5010-5026.

Athanasopoulou, E., Tombrou, M., Pandis, S.N., and Russell, A.G., (2008). The role of sea-salt emissions and heterogeneous chemistry in the air quality of polluted coastal areas.

Baker, K., Scheff, P., (2007), Photochemical model performance for PM2.5 sulfate, nitrate, ammonium, and precursor species SO2, HNO3, and NH3 at background monitor locations in the central and eastern United States. Atmospheric Environment 41, 6185-6195.

Bash, J.O., Baker, K.R., Beaver, M.R., 2016. Evaluation of improved land use and canopy representation in BEIS v3. 61 with biogenic VOC measurements in California. Geoscientific Model Development 9, 2191.

Berge, E., H-C. Huang, J. Chang, and T-H. Liu, (2001), A study of the importance of initial conditions for photochemical oxidant modeling, *J. Geophys. Res.*, **106**, 1347–1363.

Berrocal, V., Gelfand, A. E. and Holland, D. M. (2010a), A bivariate space-time downscaler under space and time misalignment. *The Annals of Applied Statistics* **4**, 1942-1975.

Berrocal, V., Gelfand, A. E., and Holland, D. M. (2010b), A spatio-temporal downscaler for output from numerical models. J. of Agricultural, Biological,and Environmental Statistics **15**, 176-197.

Bey, I., D.J. Jacob, R.M. Yantosca, J.A. Logan, B. Field, A.M. Fiore, Q. Li, H. Liu., L.J. Mickley, and M. Schultz, (2001), Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *J. Geophys. Res.*, **106**, 23,073-23,096.

Blanchard, C. L., (2000), Ozone process insights from field experiments- Part III: Extent of reaction and ozone formation, *Atmospheric Environment,* **34**, 2035-2043.

Blanchard, C.L., Lurmann, F.W., Roth, P.M., Jeffries, H.E. and Korc, M. (1999), The Use of Ambient Data to Corroborate Analyses of Ozone Control Strategies. Atmos. Environ. 33: 369–381.

ED_002666_00014243-00184

Blanchard, C.L. and Stoeckenius, T. (2001). Ozone Response to Precursor Controls: Comparison of Data Analysis Methods with the Predictions of Photochemical Air Quality Simulation Models. Atmos. Environ. 35: 1203–1215.

Bloomer, B. J., J. W. Stehr, C. A. Piety, R. J. Salawitch, and R. R. Dickerson, (2009), Observed relationships of ozone air pollution with temperature and emissions. Geophys. Res. Lett., 36, L09803, doi:10.1029/2009GL037308.

Boylan, J.W., Russell, A.G., (2006), PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. Atmospheric Environment 40, 4946-4959.

Byun, D., Schere, K.L., (2006), Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. Applied Mechanics Reviews 59, 51-77.

Cabada, J.C., S. N. Pandis, R. Subramanian., A. L. Robinson, A. Polidori, and B. Turpin, (2004), Estimating the Secondary Organic Aerosol Contribution to $PM_{2.5}$ Using the EC Tracer Method, *Aerosol Science and Technology*, 38, 140-155.

Camalier, L., Cox, W., Dolwick, P., (2007), The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. Atmospheric Environment 41, pp. 7127-7137.

CARB, (California Air Resources Board), 2010. Updated economic analysis of California's climate change scoping plan, (March). Available at: http://www.arb.ca.gov/cc/scopingplan/economics-sp/updated-analysis/updated_sp_analysis.pdf.

Carlton, A. and Baker, K., (2011), Photochemical modeling of the Ozark isoprene volcano: MEGAN, BEIS, and their impacts on air quality predictions. Environmental Science and Technology. 45(10): 4438-4445.

Castell, N., Stein, A.F., Mantilla, E., Salvador, R. and Millan, M. (2009), Evaluation of the Use of Photochemical Indicators to Assess Ozone-NOx-VOC Sensitivity in the Southwestern Iberian Peninsula. J. Atmos. Chem. 63: 73–91.

Chock, D.P., Chang, t.Y., Winkler, S.L., Nanac, B.I., (1999), The impact of an 8h ozone air quality standard on ROG and NO_x controls in southern California. *Atmos Environ*, 33, 2471-2485.

Chow, J. C., J. G. Watson, L.-W.A. Chen, J. Rice, and N. H. Frank (2010), Quantification of PM2.5 organic carbon sampling artifacts in U.S. networks, *Atmos. Chem. Phys.*, 10, 5223-5339.

Chow, J.C., Watson, J.G., Crow, D., Lowenthal, D.H., Merrifield, T. (2001), Comparison of IMPROVE and NIOSH carbon measurements, *Aerosol Science and Technology*, 34, 23-34.

ED_002666_00014243-00185

Chu, S.-H., (2005), Stable estimate of primary OC/EC ratios in the EC tracer method, *Atmospheric Environment*, 39, 1383–1392.

Clegg, S.L., P. Brimblecombe, and A. S. Wexler, (1998), Aerosol Inorganics Model; A Thermodynamic Model of the System H+ - NH4+ - SO42- - NO3- - H2O at Tropospheric Temperatures. *J. Phys. Chem.*, 102A, 2137-2154. http://www.aim.env.uea.ac.uk/aim/project.htm

CMAS, (Community Modeling and Analysis System), 2008, Atmospheric Model Evaluation Tool (AMET) User's Guide. Available at: https://www.cmascenter.org/help/model_docs/amet/1.1/AMET_Users_Guide_V1.1.pdf

Cohan, D. S., Hakami, A., Hu, Y. T., and Russell, A. G., (2005), Nonlinear response of ozone to emissions: Source apportionment and sensitivity analysis, Environ. Sci. Technol., 39, 6739–6748.

Cohan D.S., D. Tian, Y. Hu, and A.G. Russell, (2006), Control strategy optimization for attainment and exposure mitigation: case study for ozone in Macon, Georgia, Environmental Management, **38**, 451-462.

Collett, Jeffrey, (2004), Technical note on IMPROVE study.

Cooper, O. R., R-S Gao, D. Tarasick, T. Leblanc, and C. Sweeney, (2012), Long term ozone trends at rural ozone monitoring sites across the United States, *J. Geophys. Res.*, **117**.

Coordinating Research Council (CRC) final report for project A-69-1. Regional modeling of weekday/weekend ozone change in the midwestern US. Prepared by Environ International. May 19, 2011: http://www.crcao.com/reports/recentstudies2011/A69-1/CRCA-69-1_Final_Report_19May2011.pdf

Cox, W.M., and S. Chu, (1993), Meteorologically Adjusted Ozone Trends in Urban Areas: A Probabilistic Approach, *Atmospheric Environment,* **27B**, (4), 425-434.

Cox, W.M., and S. Chu, (1996), Assessment of Interannual Ozone Variation in Urban Areas from a Climatological Perspective, *Atmospheric Environment,* **30**, 2615-2625.

Davidson, J. P., McQueen, R. Mathur, R. Draxler, R. Wayland, N. Seaman, K. Carey, (2007). NOAA-EPA's National Air Quality Forecast Capability: Testing Expanded Capabilities Preprints, Ninth Conference on Atmospheric Chemistry American Meteorological Society, San Antonio, TX.

De Meij, A., A. Gzella, P. Thunis, C. Cuvelier, B. Bessagnet, J. F. Vinuesa, and L. Menut, (2009), The impact of MM5 and WRF meteorology over complex terrain on CHIMERE model calculations. Atmospheric Chemistry and Physics 9, pp. 6611-6632.

ED_002666_00014243-00186

Dennis, R., Fox, T., Fuentes, M., Gilliland, A., Hanna, S., Hogrefe, C., Irwin, J., Rao, S.T., Scheffe, R., Schere, K., Steyn, D.A., Venkatram, A., (2010), A framework for evaluating regional-scale numerical photochemical modeling systems, J. Environmental Fluid Mechanics, 471-489, http://dx.doi.org/10.1007/s10652-009-9163-2

Deuel, H.P., and S.G. Douglas, (1998), Episode Selection for the Integrated Analysis of Ozone, Visibility and Acid Deposition for the Southern Appalachian Mountains, Draft Technical Report Systems Applications International, Inc. (SYSAPP-98/07), prepared for the Southern Appalachians Mountains Initiative.

Donner L.J., B.L. Wyman, R.S. Hemler, L.W. Horowitz, Y. Ming, M. Zhao, J-C Golaz, P Ginoux, S.J. Lin, M.D. Schwarzkopf, J. Austin, G. Alaka, W.F. Cooke, T.L. Delworth, S.M. Freidenreich, C.T. Gordon, S.M. Griffies, I.M. Held, W.J. Hurlin, S.A. Klein, T.R. Knutson, A.R. Langenhorst, H-C Lee, Y. Lin, B.I. Magi, S.L. Malyshev, P.C.D. Milly, V. Naik, M.J. Nath, R. Pincus, J.J. Ploshay, V. Ramaswamy, C.J. Seman, E. Shevliakova, J.J. Sirutis, W.F. Stern, R.J. Stouffer, R.J. Wilson, M. Winton, A.T. Wittenberg, and F. Zeng, (2011), The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J Clim*, **24(13)**:3484–3519.

Draxler, R.R., and G.D. Hess, (1998), An overview of the HYSPLIT_4 modeling system of trajectories, dispersion, and deposition. *Aust. Meteor. Mag.*, **47**, 295-308. On-line model available at: http://ready.arl.noaa.gov/HYSPLIT.php

Draxler, R.R., P. Ginoux and A. Stein, (2010), An Empirically Derived Emissions Algorithm for Wind-Blown Dust. *J. Geo. Res.*, Vol. 115, Issue D16.27. August. (http://onlinelibrary.wiley.com/doi/10.1029/2009JD013167/abstract).

Du Y., Raffuse S.M., and Reid S.B., (2013), Technical guidance for using SmartFire2 / BlueSky Framework to develop national wildland fire emissions inventories. Draft user's guide prepared for the U.S. Environmental Protection Agency, Research Triangle Park, NC, Sonoma Technology Inc..

Duncan, B.N., et al., (2010), Application of OMI observations to a space-based indicator of NOx and VOC controls on surface ozone formation, Atmospheric Environment 44: 2213-2223.

Edgerton E.S., B.E. Hartsell, R.D. Saylor, J.J. Jansen, D.A. Hansen, and G.M. Hidy, (2005), The Southeastern Aerosol Research and Characterization Study: Part II. Filter-based measurements of fine and coarse particulate matter mass and composition, *J. Air Waste Mange. Assoc.,* **55,** 1527-1542.

ED_002666_00014243-00187

Emery, C., and E. Tai, (2001), Enhanced Meteorological Modeling and Performance Evaluation for Two Texas Ozone Episodes , prepared for the Texas Near Non-Attainment Areas through the Alamo Area Council of Governments", by ENVIRON International Corp, Novato, CA., http://www.tnrcc.state.tx.us/air/aqp/airquality_contracts.html#met01

Emery, C., Tai, E., Yarwood, G., Morris, R., (2011). Investigation into approaches to reduce excessive vertical transport over complex terrain in a regional photochemical grid model. Atmospheric Environment 45, 7341-7351

Emery, C., Liu, z., Russell, A.G., Odman, M.T., Yarwood, G., Kumar, N., (2017), Recommendations on Statistics and Benchmarks to Assess Photochemical Model Performance, Journal of the Air and Waste Management Association, 67:5, 582-598, doi:/10.1080/10962247.2016.1265027

Emmons, L. K., S. Walters, P.G. Hess, J-F Lamarque, G.G. Pfister, D. Fillmore, C. Granier, A. Guenther, D. Kinnison, T. Laepple, J. Orlando, X. Tie, G. Tyndall, C. Wiedinmyer, S.L. Baughcum, and S. Kloster,(2010), Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), *Geosci. Model Dev*., **3**, 43-67.

Environ and TCEQ, (2003), Modeling Protocol – Development of Base Case Photochemical Modeling to Address 1-hour and 8-Hour Ozone Attainment in the Dallas/Fort Worth Area, http://www.tceq.texas.gov/assets/public/implementation/air/am/docs/dfw/sip-bibliography/2003a_Environ_DFW_Modeling_Protocol_Draft_Final.pdf

ENVIRON International Corporation, (2006), Guidance for the Application of the CAMx Hybrid Photochemical Grid Model to Assess Visibility Impacts of Texas BART Sources at Class I Areas, Novato, CA. Prepared for Texas Commission on Environmental Quality (TCEQ) (http://www.tceq.state.tx.us/assets/public/implementation/air/sip/bart/BART_TX_CAMx_PSAT_Guid.pdf).

Eyth, A., Ran, L., Partheepan, R, Yarwood, G., Jimenez, M., Rao, S. (2006): New Tools to Generate Spatial Surrogate and Speciation Profile Inputs to SMOKE, International Emission Inventory Conference, New Orleans, 2006.

Fann, N., T. Brennan, P. Dolwick, J.L. Gamble, V. Ilacqua, L. Kolb, C.G. Nolte, T.L. Spero, and L. Ziska, 2016: Ch. 3: Air Quality Impacts. The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment. U.S. Global Change Research Program, Washington, DC, 69–98. http://dx.doi.org/10.7930/J0GQ6VP6.

Fiore, A.M., V. Naik, D.V. Spracklen, A. Steiner, N. Unger, M. Prather, D. Bergmann, P.J. Cameron-Smith, I. Cionni, W.J. Collins, S. Dalsoren, V. Eyring, G.A. Folberth, P. Ginoux, L.W. Horowitz, B. Josse, J.F. Lamarque, I.A. MacKenzie, T. Nagashima, F.M. O'Connor, M. Righi, S.T. Rumbold, D.T. Shindell, R.B. Skeie, K. Sudo, S. Szopa, T. Takemura, and G. Zeng, (2012), Global air quality and climate. Chem. Soc. Rev. 41, pp. 6663-6683.

ED_002666_00014243-00188

Flaum, J. B., S.T. Rao, and I.G. Zurbenko, (1996), Moderating the influence of meteorological conditions on ambient ozone concentrations, *Journal of the Air and Waste Management Association* **46**, 35-46.

Frank, N., (2006), Retained Nitrate, Hydrated Sulfates, and Carbonaceous Mass in Federal Reference Method Fine Particulate Matter for Six Eastern U.S. Cities" *J. Air Waste Mange. Assoc.*, 56, 500-511.

Frank, N. Memo to PM NAAQS Review Docket EPA-HQ-OAR-2007-0492. Recommendations to Users of CSN and IMPROVE Speciation Data regarding Sampling Artifact Correction for PM2.5 Organic Carbon. June 14, 2012. http://www.epa.gov/ttnnaaqs/standards/pm/data/20120614Frank.pdf

Foley, K.M., Roselle, S.J., Appel, K.W., Bhave, P.V., Pleim, J.E., Otte, T.L., Mathur, R., Sarwar, G., Young, J.O., Gilliam, R.C., Nolte, C.G., Kelly, J.T., Gilliland, A.B., Bash, J.O., (2010), Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. Geoscientific Model Development 3, 205-226.

Foley. K., P. Dolwick, C. Hogrefe, H. Simon, B. Timin, and N. Possiel, (2014), Dynamic Evaluation of CMAQ Part II: Evaluation of relative response factor metrics for ozone attainment demonstrations, in review.

Fox, T., (2017), 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, PM2.5, and Regional Haze. https://www3.epa.gov/ttn/scram/guidance/clarification/20170804-Photochemical_Grid_Model_Clarification_Memo.pdf

Fu, J.S., X. Dong, K. Huang, and C. Jang: Using Hemispheric-CMAQ to provide initial and boundary conditions for regional modeling, 11th Annual CMAS conference, Chapel Hill NC, October 2012.

Fu, J.S., Dong, X., Huang, K., Tong, D., and Zhuang, G.: Model development of dust emission and heterogeneous chemistry within the Community Multiscale Air Quality modeling system and its application over East Asia, Atmospheric Chemistry and Physics, 16, 8157-8180, 2016.

Galmarini, S., Rao, S.T., Steyn, D.G., (2012), AQMEII: An International Initiative for the Evaluation of Regional-Scale Air Quality Models - Phase 1 Preface. Atmospheric Environment 53, 1-3.

Gantt, B., Kelly, J. T., and Bash, J. O., (2015), Updating sea spray aerosol emissions in the Community Multiscale Air Quality (CMAQ) model version 5.0.2, Geosci. Model Dev., 8, 3733-3746, doi:10.5194/gmd-8-3733-2015.

ED_002666_00014243-00189

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

Gilliam, C. R., J. M. Godowitch, and S. T. Rao, (2012), Improving the Horizontal Transport in the Lower Troposphere with Four Dimensional Data Assimilation. Atmospheric Environment 53, pp 186-201.

Gilliland, A.B., Hogrefe, C., Pinder, R.W., Godowitch, J.M., Foley, K.L., Rao, S.T., (2008), Dynamic evaluation of regional air quality models: assessing changes in O_3 stemming from changes in emissions and meteorology, *Atmospheric Environment*, 42, 5110-5123.

Godowitch, J. M., R. C. Gilliam, and S. T. Rao, (2011). Diagnostic Evaluation of the chemical and transport processes in a regional photochemical air quality modeling system, Atmospheric Environment 45, pp. 3977-3987.

Godowitch, J.M., Pouliot, G.A., Rao, S.T.(2010), Assessing multi-year changes in modeled and observed urban NOx concentrations from a dynamic model evaluation perspective, *Atmospheric Environment*, 44, 2894-2901.

Grell, G. A., S. E. Peckham, R. Schmitz, S. A. McKeen, G. Frost, W. C. Skamarock, and B. Eder, (2005), Fully coupled "online" chemistry within the WRF model, Atmospheric Environment 39, pp. 6957–6975.

Grell, G.A., J. Dudhia and D.R. Stauffer, (1994), A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5), NCAR/TN-398+STR, 138 pp.

Geron, C., A. Guenther and T. Pierce, (1994), An Improved Model for Estimating Emissions of Volatile Organic Compounds from Forests in the Eastern United States, *J. Geophys. Res.*, **99**, 12,773-12,791.

Guenther, A., C. Geron, T. Pierce, B. Lamb, P. Harley, and R. Fall, (2000), Natural emissions of non-methane volatile organic compounds, carbon monoxide, and oxides of nitrogen from North America, Atmospheric Environment., 34, 2205-2230.

Hammer, M.-U., Vogel, B., Vogel, H., (2002), Findings on H_2O_2/HNO_3 as an indicator of ozone sensitivity in Baden-Wurttemberg, Berlin-Brandenburg, and the Po valley based on numerical simulations. *J. Geophys Res*., 107(D22), 8190; doi:10.1029/2000JD00211.

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

Hansen, D.A., E.S. Edgerton, B.E. Hartsell, J.J.Jansen, N. Kandasamy, G.M. Hidy, and C.L. Blanchard, (2003), The Southeastern Aerosol Research and Characterization Study: Part 1 – Overview, *J. Air Waste Mange. Assoc.,* **53,** 1460-1471

Harley, R.A., Marr, L.C., Lehner, J.K. Giddings, S.N., (2005), Changes in motor vehicle emission on diurnal to decadal time scales and effects on atmospheric composition, *Environmental Science and Technology*, 39, 5356-5362, 2005.

Henderson, B.H.; Kimura, Y.; McDonald-Buller, E.; Allen, D.; Vizuete, W., (2011), Comparison of Lagrangian Process Analysis Tools for Eulerian Air Quality Models. Atmospheric Environment, 45, 5200-5211, doi:10.1016/j.atmosenv.2011.06.005

Hering, S. and G. Cass, (1999), The Magnitude of Bias in Measurement of $PM_{2.5}$ Arising from Volatilization of Particulate Nitrate from Teflon Filters. *JAWMA* 49:725-733.

Hogrefe, C., S.T. Rao, I.G. Zurbenko, and P.S. Porter, (2000), Interpreting the information in time series of ozone observations and model predictions relevant to regulatory policies in the eastern United States, *Bull. Amer. Met. Soc.,* **81**, 2083-2106.

Hogrefe C., W. Hao, E.E. Zalewsky, J-Y Ku, B. Lynn, C. Rosenzweig, M.G. Schultz, S. Rast, M.J. Newchurch, L. Wang, P.L. Kinney, and G. Sistla, (2011), An analysis of long-term regional-scale ozone simulations over the Northeastern United States: variability and trends, *Atmospheric Chemistry and Physics*, **11**, 567-582.

Hogrefe C., J. Xing, C Wei, R. Mathur, S. Roselle, S. Porter, S.T. Rao (2014), Evaluating the Impact of Emission Changes on Observed and Simulated Ozone Concentrations: Learning from the Past to Predict the Future, AWMA Annual Conference.

Hogrefe, C., Roselle, S.J., Bash, J.O., 2017. Persistence of initial conditions in continental scale air quality simulations. Atmospheric Environment 160, 36–45. https://doi.org/10.1016/j.atmosenv.2017.04.009

Hyslop, N.P., Trzepla, K., White, W.H., (2012), Reanalysis of Archived IMPROVE PM2.5 Samples Previously Analyzed over a 15-Year Period, *Environmental Science and Technology*, 46, 10106-10113.

IMPROVE, (2000), Spatial and Seasonal Patterns and Temporal Variability of Haze and its Constituents in the United States: Report III. Cooperative Institute for Research in the Atmosphere. Available at: http://vista.cira.colostate.edu/improve/publications/Reports/2000/2000.htm

IMPROVE, (2006), Revised IMPROVE Algorithm for Estimating Light Extinction from Particle Speciation Data, January 2006, http://vista.cira.colostate.edu/improve/Publications/GrayLit/gray_literature.htm

ED_002666_00014243-00191

IMPROVE, (2011), Spatial and Seasonal Patterns and Temporal Variability of Haze and its Constituents in the United States: Report V. Cooperative Institute for Research in the Atmosphere. Available at: http://vista.cira.colostate.edu/improve/publications/Reports/2011/2011.htm

Isakov, V., J.S. Irwin, and J. Ching, (2007), Using CMAQ for exposure modeling and characterizing the sub-grid variability for exposure estimates", *Journal of Applied Meteorology and Climatology*, 46(9): 1354-1371.

Jacob, D. J., Logan, J.A., and Murti, P.P., (1999), Effect of rising Asian emissions on surface ozone in the United States, *Geophys. Res. Lett.*, **26**, 2175–2178.

Jacob, D.J., D.A. Winner, (2009), Effect of climate change on air quality, Atmospheric Environment 43, pp. 51–63.

Jang, J.C., Jeffries, H.E., Tonnesen, S., (1995), Sensitivity of ozone to model grid resolution-II. Detailed process analysis for ozone chemistry, Atmospheric Environment 29, 3101-3114.

Jeffries, H.E., (1994), Process Analysis for UAM Episode 287, memo to Brock Nicholson, NC DEHNR, April 8, 1994, ftp://airsite.unc.edu/pdfs/ese_unc/jeffries/projprpt/uammodpanel287.pdf

JFSP, (2013), DEASCO3 2008 Emissions Inventory Methodology. Prepared for Join Fire Science Program Project No. 178-13, September, (https://wraptools.org/pdf/ei_methodology_20130930.pdf).

Jimenez, P., Baldansano, J.M., (2004), Ozone response to precursor controls in very complex terrains: Use of photochemical indicators to assess O_3-NO_x-VOC sensitivity in the northeastern Iberian Peninsual. *J Geophys Res*., 109, D20309, doi:10.1029/2004JD04985.

Kelly, J.T., Avise, J., Cai, C., Kaduwela, A.P., (2011), Simulating Particle Size Distributions over California and Impact on Lung Deposition Fraction. *Aerosol Science and Technology* 45, 148-162.

Kelly, J.T., Bhave, P.V., Nolte, C.G., Shankar, U., Foley, K.M., (2010), Simulating emission and chemical evolution of coarse sea-salt particles in the Community Multiscale Air Quality (CMAQ) model. *Geoscientific Model Development* 3, 257-273.

Koo, B., Chien, C.-J., Tonnesen, G., Morris, R., Johnson, J., Sakulyanontvittaya, T., Piyachaturawat, P., and Yarwood, G., (2010), Natural emissions for regional modeling of background ozone and particulate matter and impacts on emissions control strategies, *Atmos. Environ.*, 44, 2372–2382., doi:10.1016/j.atmosenv.2010.02.041.

ED_002666_00014243-00192

Kulkarni S., D. Chau, J. Avise, J. DaMassa, and A. Kaduwela, (2014), An Extended Approach to Calculate Relative Response Factors for use in the Attainment Demonstration of the Ambient Air Quality Standards for 1-hour Ozone, Submitted to Journal of the Air & Waste Management Association.

Kwok, R. H. F., Napelenok, S. L., & Baker, K. R., (2013), Implementation and evaluation of PM2.5 source contribution analysis in a photochemical model. *Atmospheric Environment*, *80*, 398-407.

Kwok, R. H. F., Baker, K. R., Napelenok, S. L., & Tonnesen, G. S., (2014), Photochemical grid model implementation of VOC, NOx, and O3 source apportionment. *Geoscientific Model Development Discussions*, *7*(5), 5791-5829.

Lake Michigan Air Directors Consortium (2009), Conceptual model of PM2.5 episodes in the Midwest, LADCO PM Data Analysis Workshop, http://www.ladco.org/re ports/pm25/post08/pm25_conceptual_model.pdf

Lam, Y-F. and J. S. Fu, (2009), A novel downscaling technique for the linkage of global and regional air quality modeling, *Atmos. Chem. Phys*. **9**, 9169–9185, 2009.

Liang, J.-Y., Jackson, B., Kaduwela, A., (2006), Evaluation of the ability of indicator species ratios to determine the sensitivity of ozone to reductions in emissions of volatile organic compounds and oxides of nitrogen in northern California. *Atmos Environ*, 40, 5156-5166.

Lin, M., Horowitz, L.W., Payton, R., Fiore, A.M., Tonnesen, G., 2017. US surface ozone trends and extremes from 1980 to 2014: quantifying the roles of rising Asian emissions, domestic controls, wildfires, and climate. *Atmospheric Chemistry and Physics* 17, 2943–2970. doi: 10.5194/acp-17-2943-2017

Liu, X, K. Chance, C.E. Sioris, T.P. Kurosu, R.J.D. Spurr, R.V. Martin, T-M Fu, J.A. Logan, D.J. Jacob, P.I. Palmer, M.J. Newchurch, I.A. Megretskaia, and R.B. Chatfield, (2006), First directly retrieved global distribution of tropospheric column ozone from GOME: Comparison with the GEOS-CHEM model, *J. Geophys. Res*., **111**.

Lu, C.H., and J.S. Chang, (1998), On the Indicator-Based Approach to Assess Ozone Sensitivities and Emissions Features, *J. Geophysical Research*, **103**, 3453-3462.

Lui, X.-H., Zhang, Y., Xing, J., Zhang, Q., Wang, K., Streets, D.G., Jang, C., Wang, W.-X., Hao, J.-M., (2010), Understanding the regional air pollution over China using CMAQ, part II. Process analysis and sensitivity of ozone and particulate matter to precursor emissions, *Atmospheric Environment*, 44, 3719-3727.

ED_002666_00014243-00193

Malm W.C., Schichtel B.A., Pitchford M.L. (2011), Uncertainties in PM2.5 gravimetric and speciation measurements and what we can learn from them. J Air Waste Manag Assoc, 61(11): 1131-49.

MANE-VU (2013), Future Modeling Platform Base Year Determination. http://www.otcair.org/MANEVU/Upload/Publication/Reports/Future%20Modeling%20Platform%20Base%20Year%20Selection%20Analysis%20-%20Oct%209%202013%20Final.pdf

Mansell, G.E., S. Lau, J. Russell and M. Omary, (2006), Fugitive Wind Blown Dust Emissions and Model Performance Evaluation Phase II. ENVIRON International Corporation and University of California at Riverside. May 5. (http://www.wrapair.org/forums/dejf/documents/WRAP_WBD_PhaseII_Final_Report_050506.pdf).

Marr, L.C.; Harley, R.A., (2002), Spectral analysis of weekday-weekend differences in ambient ozone, nitrogen oxide, and non-methan hydrocarbon time series in California, *Atmos. Environ*, 36, 2327-2335.

Martilli, A., Neftel, A., Favaro, G., Krichner, F., Sillman, S., Clappier, A., (2002), Simulation of the ozone formation in the northern part of the Po Valley. *J Geophys Res*, 107(D22), 8195.

Martin, R.V., Fiore, A.M., Donkelaar, A.V., (2004), Space-based diagnosis of surface ozone sensitivity to anthropogenic emissions. *Geophys Res Lett*, L06120, doi:10.1029/2004GL019416, 2004.

Matichuk, R., Tonneson, G., Luecken, D., Gilliam, R., Napelenok, S.L., Baker, K.R., Schwede, D., Murphy, B., Helmig, D., Lyman, S.N., Roselle, S., (2017), Evaluation of the Community Multiscale Air Quality Model for Simulating Winter Ozone Formation in the Uinta Basin. *J Geophys Res*, **122**, 13545-13572. https://doi.org/10.1002/2017JD027057

Meyer, E.L., K.W. Baldridge, S. Chu, and W.M. Cox, (1997), Choice of Episodes to Model: Considering Effects of Control Strategies on Ranked Severity of Prospective Episode Days, *Paper 97-MP112.01*, Presented at 97th Annual AWMA Meeting, Toronto, Ontario.

McGaughey G., C. Durrenberger, D. Allen, and E. McDonald-Buller, (2010), Conceptual model for ozone for the Austin area, prepared for the Capital Area Council of Governments and the Texas Commission on Environmental Quality, http://www.capcog.org/documents/airquality/cac/2010/september2010/Austin_CM_ver21.pdf.

McMillan, N., Holland, D. M., Morara, M., and Feng, J., (2010), Combining Numerical Model Output and Particulate Data Using Bayesian Space-Time Modeling *Environmetrics* **21**, 48-65.

Milanchus, M. L., S.T. Rao, and I.G. Zurbenko, (1998), Evaluating the effectiveness of ozone management efforts in the presence of meteorological variability, *Journal of the Air and Waste Management Association* **48,** 201-215.

Milford, J.B., Gao, D., Sillman, S., Blossey, P. and Russel, A.G., (1994), Total Reactive Nitrogen (NOy) as an Indicator of the Sensitivity of Ozone to Reduction in Hydrocarbons and NOx Emissions. J. Geophys. Res. D: Atmos. 99: 3533–3542.

Morris, R.E, McNally, D.E., Tesche, T.W., Tonnesen, G., Boylan, J.W., and Brewer, P., (2005), Preliminary Evaluation of the Community Multiscale Air Quality Model for 2002 over the Southeastern United States, Journal of the Air & Waste Management Association, 55:11, 1694-1708, DOI: 10.1080/10473289.2005.10464765

Morris, R.E., Koo, B., Guenther, A., Yarwood, G., McNally, D., Tesche, T.W., Tonnesen, G., Boylan, J., Brewer, P., (2006), Model sensitivity evaluation for organic carbon using two multi-pollutant air quality models that simulate regional haze in the southeastern United States, Atmospheric Environment 40, 4960-4972.

Morris and Tai, (2011), Attainment Demonstration Modeling for the Denver 8-Hour Ozone SIP, http://www.ozoneaware.org/postfiles/documentsandpresentations/modeling/documents/Denver_Model_Protocol_Draft3_080211.pdf

Murphy, J.G., Day, D.A., Cleary, P.A., Wooldridge, P.J., Millet, D.B., Goldstein, A.H., Cohen, R.C., (2007), The weekend effect within and downwind of Sacramento – Part 1: Observations of ozone, nitrogen oxides, and VOC reactivity, *Atmos. Chem. Phys.*, 7, 5327-5339, 2007.

Na, K. S., A.A. Sawant, C. Song, and D.R. Cocker, (2004), Primary and secondary carbonaceous species in the atmosphere of Western Riverside County, California, *Atmospheric Environment*, 38, 1345–1355.

NAPD, (2013), National Acid Deposition Program 2013 Annual Summary. Available at: http://nadp.sws.uiuc.edu/lib/data/2013as.pdf

Napelenok, S. L., Cohan, D. S., Hu, Y. T., and Russell, A. G., (2006), Decoupled direct 3D sensitivity analysis for particulate matter(DDM-3D/PM), Atmos. Environ., 40, 6112–6121.

Nolte, C.G., Bhave, P.V., Arnold, J.R., Dennis, R.L., Zhang, K.M., Wexler, A.S., (2008), Modeling urban and regional aerosols – Application of the CMAQ-UCD aerosol model to Tampa, a coastal urban site, *Atmospheric Environment*, 42, 3179-3191.

Otte, T.L., (2008a), The impact of nudging in the meteorological model for retrospective air quality simulations. part I: evaluation against national observation networks. J. Appl. Meteor. Climatol., 47, pp. 1853–1867.

Otte, T.L., (2008b), The impact of nudging in the meteorological model for retrospective air quality simulations. part II: evaluating collocated meteorological and air quality observations, J. Appl. Meteor. Climatology., 47, pp. 1868–1887.

Otte, T.L., J.E. Pleim, (2010), The meteorology-Chemistry interface processor (MCIP) for the CMAQ modeling system. Geoscientific Model Development, 3, pp. 243–256.

Pandis, S.N., Wexler, A.S., Seinfeld, J.H., (1993), Secondary organic aerosol formation and transport – II predicting the ambient secondary organic aerosol size distribution, *Atmospheric Environment*, 27A, 15, 2403-2416.

Pitchford, M.L., and W.C. Malm, (1994), Development and Applications of a Standard Visual Index, *Atmos. Environ.*, **28**(5),1049-1054.

Pitchford, M., W. Malm, B. Schichtel, N. Kumar, D. Lowenthal and J. Hand, (2007), Revised algorithm for estimating light extinction from IMPROVE particle speciation data, *J. Air & Waste Manage. Assoc., 57,* 1326-1336.

Pierce, T., Hogrefe, C., Rao, S.T., Porter, P.S., Ku, J.-Y., (2010), Dynamic evaluation of a regional air quality model: Assessing the emissions-induced weekly ozone cycle, *Atmospheric Environment*, 44, 3583-3596.

Pirovano, G., Balzarini, A., Bessagnet, B., Emery, C., Kallos, G., Meleux, F., Mitsakou, C., Nopmongcol, U., Riva, G.M., Yarwood, G., (2012), Investigating impacts of chemistry and transport model formulation on model performance at European scale. Atmospheric Environment 53, 93-109.

Pleim, J., Gilliam, R., Appel, W., Ran, L. (2016). Recent Advances in Modeling of the Atmospheric Boundary Layer and Land Surface in the Coupled WRF-CMAQ Model (2016), *Air Pollution Modeling and its Application*, XXIV, 391-396. https://doi.org/10.1007/978-3-319-24478-5_64

Pleim, J.E., R. Gilliam, (2009), An indirect data assimilation scheme for deep soil temperature in the Pleim–Xiu land Surface model, J. Appl. Meteor. Climatol., 48, pp. 1362–1376 Resources Conservation Commission, by ENVIRON, International Corp, Novato, CA

Porter S., S.T. Rao, C. Hogrefe, E. Grego, and R. Mathur (2014), Using Regional Air quality Models in Ozone Attainment Demonstrations, 9th International Conference on Air Quality, Garmisch-Partenkirchen Germany.

Pouliot and Pierce (2009), Integration of the Model of Emissions of Gases and Aerosols from Nature (MEGAN) into the CMAS Modeling System, Presented at 18th Annual International Emission Inventory Conference, Baltimore, MD, April 14-17, 2009, https://www.epa.gov/sites/production/files/201510/documents/pouliot.pdf.

ED_002666_00014243-00196

Pouliot, G., Simon, H., Bhave P., Tong, D., Mobley, D., Pace, T., and Pierce, T., *Assessing the Anthropogenic Fugitive Dust Emission Inventory and Temporal Allocation using an Updated Speciation of Particulate Matter*, 19th Annual International Emission Inventory Conference, San Antonio, TX, September 27-30, 2010.

Pouliot, G., Rao, V., McCarty, J.L., Soja, A., (2016), Development of the Crop Residue and Rangeland Burning in the 2014 National Emissions Inventory Using Information from Multiple Sources. Journal of the Air & Waste Management Association. 67(5), 613-622

Pun, B. and C. Seigneur, (1999), Understanding particulate matter formation in the California San Joaquin Valley: Conceptual model and data needs, Atmos. Environ., 33, 4865-4875.

Pun, B.K., and C. Seigneur, (2001), Sensitivity of Particulate Matter Nitrate Formation to Precursor Emissions in the California San Joaquin Valley, *Environ. Sci. Technol.*, **35** (14), 2979-2987.

Raffuse, S., D. Sullivan, L. Chinkin, S. Larkin, R. Solomon, A. Soja, (2007). Integration of Satellite-Detected and Incident Command Reported Wildfire Information into BlueSky, June 27, 2007. Available at: http://getbluesky.org/smartfire/docs.cfm

Ramboll-Environ, (2017). wrfcamx version 4.6 Release Notes. July 11, 2017. www.camx.com Ramboll Environ International Corporation, Novato, CA.

Ramboll, (2018). User's Guide Comprehensive Air Quality Model with Extensions (CAMx) Version 6.50. www.camx.com Ramboll Environment Health, Novato, CA.

Rao, S. T., E. Zalewsky, E., and I.G. Zurbenko, (1995), Determining temporal and spatial variations in ozone air quality, *Journal of the Air and Waste Management Association* **45**, 57-61.

Rodriguez, M.A., Barna, M.G., Moore, T., (2009), Regional Impacts of Oil and Gas Development on Ozone Formation in the Western United States. J. Air Waste Manage. Assoc. 59, 1111-1118.

Russell, A., and R. Dennis, (2000), NARSTO critical review of photochemical models and modeling, *Atmospheric Environment,* **34***,* 2283-2324

Russell, A.G., (2008), EPA Supersites Program-related emissions-based particulate matter modeling: Initial applications and advances. J. Air Waste Manage. Assoc. 58, 289-302.

Sarwar, G., S. Roselle, R. Mathur, W. Apel, R. Dennis (2008), A Comparison of CMAQ HONO predictions with observations from the Northeast Oxidant and Particle Study, *Atmospheric Environment*, 42, 5760–5770.

ED_002666_00014243-00197

Sarwar, G., Gantt, B., Schwede, D., Foley, K., Mathur, R., & Saiz-Lopez, A., (2015), Impact of enhanced ozone deposition and halogen chemistry on tropospheric ozone over the Northern Hemisphere. *Environmental Science & Technology, 49*(15), 9203-9211.

Saxena, P. and Hildemann, L, (1996), Water Soluble Organics in Atmospheric Particles: A Critical Review of the Literature and Application of thermodynamics to Identify Candidate Compounds, *J. Atmos. Chem.*, 24:57-109.

Saylor, R. D., E. S. Edgerton, and B. E. Hartsell, (2006), Linear regression techniques for use in the EC tracer method of secondary organic aerosol estimation, *Atmospheric Environment*, 40, 7546-7556.

Seigneur, C., B. Pun, P. Pai, J. F. Louis, P. Solomon, C. Emery, R. Morris, M., Zahniser, D. Worsnop, P. Koutrakis, W. White and I. Tombach, (2000), Guidance for the performance evaluation of three-dimensional air quality modeling systems for particulate matter and visibility, J. Air Waste Manage. Assoc., 50, 588-599.

Sillman, S., (1995), The Use of NOy, H_2O_2, and HNO_3 as Indicators for O3-NOx-ROG Sensitivity in Urban Locations, *J. Geophys. Res.* **100**, 14175-14188.

Sillman, S., He, D., Cardelino, C., Imhoff, R.E., (1997),The use of photochemical indicators to evaluate ozone-NO_x-hydrocarbon sensitivity; Case studies from Atlanta, New York, and Los Angeles, *J. Air Waste Manage Assoc*, 47, 642-652.

Sillman, S., (1998), Evaluating the Relation Between Ozone, NO_x and Hydrocarbons: The Method of Photochemical Indicators, EPA/600R-98/022, http://www-personal.engin.umich.edu/~sillman/publications.htm

Sillman, S., He, D., (2002), Some theoretical results concerning O_3-NO_x-VOC chemistry and NO_x-VOC indicators, *J Geophys Res,* 107,D22, 4569; doi:10.1029/2001JD001123.

Sillman, S., Vautard, R., Menut, L., Kely, D., (2003), O_3-NO_x-VOC sensitivity and NO_x-VOC indicators in Paris: Results from Models and atmospheric pollution over the Paris Area (ESQUIF) measurements. *J Geophys Res*, 108, 853, doi: 10.1029/2002JD001561.

Simon, H., Bhave, P.V., Swall, J.L., Frank, N.H., and Malm, W.C., (2011), Determining the spatial and seasonal variability in OM/OC ratios across the US using multiple regression, Atmos. Chem. Phys., 11, 2933-2949

Simon, H., Baker, K. R., Phillips, S, (2012), Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012, *Atmos Environ*, 61, 124-139.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. Duda, X.-Y. Huang, W. Wang, and J.G. Powers, (2008), A Description of the Advanced Research WRF Version 3, *NCAR Technical Note* http://www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf .

Stein, A.F., Mantilla, E., Millan, M.M., (2005), Using measured and modeled indicators to assess ozone-NO_x-VOC sensitivity in a western Mediterranean coastal environment, *Atmos Environ*, 39, 7167-7180.

Stoeckenius T., (2010), A conceptual model of winter ozone episodes in southwest Wyoming, prepared for Wyoming Department of Environmental Quality, Cheyenne WY, http://deq.wyoming.gov/media/attachments/Air%20Quality/Winter%20Ozone/Technical%20Documents/Final_03conceptModel%20Report_Environ.pdf

Strader, R., F. Lurmann, and S.N. Pandis, (1999), Evaluation of secondary organic aerosol formation in winter, *Atmospheric Environment*, 33, 4849–4863.

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: https://www3.epa.gov/ttnchie1/conference/ei17/session12/raffuse.pdf.

Sundram, I., C. Clairborn, T. Strand, B. Lamb, D. Chandler and K. Saxton, (2004), Numerical Modeling of Windblown Dust in the Pacific Northwest with Improved Meteorology and Dust Emission Models. J. Geo. Res., Vol. 109, Issue D24.27. December. (http://onlinelibrary.wiley.com/doi/10.1029/2004JD004794/full).

Tang, I., (1996), Chemical and Size Effects of Hygroscopic Aerosols on Light Scattering Coefficients, *J.Geophysical Research* 101 No.D14, pp.19,245-19,250.

Tang, Y.H., P. Lee, M. Tsidulko, H.C. Huang, J.T. McQueen, G.J. DiMego, L.K. Emmons, R.B. Pierce, A.M. Thompson, H.M. Lin, D. Kang, D. Tong, S.C. Yu, R. Mathur, J.E. Pleim, T.L. Otte, G. Pouliot, J.O. Young, K.L. Schere, P.M. Davidson, and I. Stajner, (2008), The impact of chemical lateral boundary conditions on CMAQ predictions of tropospheric ozone over the continental United States, *Environ. Fluid Mech.*, **9(1)**, 43–58.

Tesche, T. W., and D. E. McNally, (2001), Evaluation of the MM5, RAMS, and SAIMM Meteorological Model for the 6-11 September 1993 COAST and 25 August-1 September 2000 TexAQS2000 Ozone SIP Modeling Episodes, report prepared for the Business Coalition for Clean Air-Appeals Group, prepared by Alpine Geophysics, LLC, Ft. Wright, KY.

Tesche T. W., D.E. McNally, and C. Tremback, (2002), Operational evaluation of the MM5 meteorological model over the continental United States: Protocol for annual and episodic evaluation. Submitted to USEPA as part of Task Order 4TCG-68027015. July 2002.

Tesche, T.W., Morris, R., Tonnesen, G., McNally, D., Boylan, J., Brewer, P., 2006. CMAQ/CAMx annual 2002 performance evaluation over the eastern US. Atmospheric Environment 40, 4906-4919.

Tolocka, M.P., Turpin, B., (2012), Contributions of organosulfur compounds to organic aerosol mass, *Environmental Science and Technology*, 46, 7978-7983.

Tonnesen, G.S., Dennis, R.L., (2000a), Analysis of radical propagation efficiency to assess ozone sensitivity to hydrocarbons and NO_x: 2. Long lived species as indicators of ozone concentrations sensitivity, *J Geophys Res*, 105, 9227-9241, doi:10.1029/1999JD90372.

Tonnesen, G.S., Dennis, R.L., (2000b), Analysis of radical propagation efficiency to assess ozone sensitivity to hydrocarbons and NO_x: 1. Local indicators of instantaneous odd oxygen production sensitivity, *J Geophys Res*, 105, 9213-9225, doi:10.1029/1999JD900371.

Torres-Jardon, R., Garcia-Reynosos, J.A., Jazcilevich, A., Ruiz-Suarez, L.G., Keener, T.C., (2009), Assessment of the ozone-nitrogen oxide-volatile organic compound sensitivity of Mexico City through an indicator-based approach: measurements and numerical simulations comparison. *J. Air Waste Manage*, 59, 1155-1172.

Torres-Jardon, R., Keener, T.C., (2006), Evaluation of ozone-nitrogen oxides-volatile organic compound sensitivity of Cincinnati, Ohio. *J Air Waste Manage*, 56, 322-333.

Turner J., (2008), A conceptual model for ambient fine particulate matter over southeast Michigan: high concentration days, prepared for the Southeast Michigan Council of Governments, Detroit MI, https://www.michigan.gov/documents/deq/deqaqd-air-aqe-Appendix-I-HighPM-final-report_238084_7.pdf.

Turpin, B.J., and J.J Huntzicker, (1995), Identification of secondary organic aerosol episodes and quantitation of primary and secondary organic aerosol concentrations during SCAQS, *Atmospheric Environment*, 29, 3527–3544.

Turpin, B. J. and H-J Lim, (2001), Species Contributions to $PM_{2.5}$ Mass Concentrations: Revisiting Common Assumptions for Estimating Organic Mass, *Aerosol Science and Technology* 35: 602-610.

U.S. EPA, (1997), Guidance For Network Design and Optimum Site Exposure For $PM_{2.5}$ and PM_{10}, OAQPS, Research Triangle Park, NC, EPA report no. EPA-454/R-99-022.

U.S. EPA, (2002), "3 Year Quality Assurance Report- Calendar Years 1999, 2000, and 2001-$PM_{2.5}$ Performance Evaluation Program".

U.S. EPA, (2003a), "Federal Reference Method Quality Assurance Report".

U.S. EPA (2003b) Guidance for Tracking Progress Under the Regional Haze Rule, EPA-454/B-03-004 (September 2003) available at: www.epa.gov/ttn/oarpg/t1/memoranda/rh_tpurhr_gd.pdf,

ED_002666_00014243-00200

U.S. EPA (2003c) Guidance for Estimating Natural Visibility Conditions Under the Regional Haze Rule, EPA-454/B-03-005 (September 2003) available at: www.epa.gov/ttn/oarpg/t1/memoranda/rh_envcurhr_gd.pdf.

U.S. EPA, (2004), "The Particle Pollution Report: Current Understanding of Air Quality and Emissions Through 2003." EPA 454/R-04-002. http://www.epa.gov/air/airtrends/aqtrnd04/pm.html

U.S. EPA, (2005), "Evaluating Ozone Control Programs in the Eastern United States: Focus on the NOx Budget Trading Program, 2004", http://www.epa.gov/air/airtrends/2005/ozonenbp/

U.S. EPA (2007a), Guidance on the use of models and other analyses for demonstrating attainment of air quality goals for ozone, $PM_{2.5}$, and regional haze. EPA-454/B-07-002, Office of Air Quality Planning and Standards, Research Triangle Park, NC.

U.S. EPA. (2007b), "Guidance for Setting Reasonable Progress Goals Under the Regional Haze Program", June 1, 2007. Available at: http://www.epa.gov/ttn/naaqs/aqmguide/collection/cp2/20070601_wehrum_reasonable_progress_goals_reghaze.pdf

U.S. EPA (2010a), Technical Support Documentation: The Industrial Sectors Integrated Solutions (ISIS) Model and the Analysis for the National Emission Standards for Hazardous Air Pollutants and New Source Performance Standards for the Portland Cement Manufacturing Industry. Research Triangle Park, NC, August 2010.

U.S. EPA (2010b), Air quality modeling technical support document: Light duty vehicle greenhouse gas emissions standards final rule, EPA 454/R-10-003, Office of Air Quality Planning and Standards, Research Triangle Park, NC, April 2010.

U.S. EPA, (2011a), Air Quality Modeling Final Rule Technical Support Document, http://www.epa.gov/airquality/transport/pdfs/AQModeling.pdf Research Triangle Park, North Carolina.

U.S. EPA, (2011b), Air Quality Modeling Technical Support Document: Final EGU NESHAP (EPA-454/R-11-009), Research Triangle Park, North Carolina.

U.S. EPA, (2012a): Clean Air Status and Trends Network (CASTNet), 2012 annual report. Available at: http://epa.gov/castnet/javaweb/docs/annual_report_2012.pdf

U.S. EPA (2012b), Our nation's air: status and trends through 2010, EPA-454/B-07-002, Office of Air Quality Planning and Standards, Research Triangle Park, February 2012.

ED_002666_00014243-00201

U.S. EPA (2014a), EPA analyses to support revisions to the ozone attainment test, Office of Air Quality Planning and Standards, Research Triangle Park, May 2014.

U.S. EPA (2014b), Policy Assessment for the Review of the Ozone NAAQS, Final Report. August. Available at: http://www.epa.gov/ttn/naaqs/standards/ozone/data/20140829pa.pdf

U.S. EPA (2016a), Air Quality Modeling Technical Support Document for the Final Cross State Air Pollution Rule Update. August 2016. https://www.epa.gov/airmarkets/air-quality-modeling-technical-support-document-final-cross-state-air-pollution-rule

U.S. EPA (2016b), Emissions Inventory Guidance for Implementation of Ozone and Particulate Matter National Ambient Air Quality Standards (NAAQS) and Regional Haze Regulations https://www.epa.gov/air -emissions-inventories/air-emissions-inventory-guidance-implementation-ozone-and-particulate.

U.S. EPA (2016c), 2014 National Emissions Inventory Technical Support Document, https://www.epa.gov/air -emissions-inventories/2014-national-emissions-inventory-nei-documentation.

U.S. EPA (2016d), Draft Guidance on Progress Tracking Metrics, Long-term Strategies, Reasonable Progress Goals and Other Requirements for Regional Haze State Implementation Plans for the Second Implementation Period, July 2016. https://www.epa.gov/visibility/draft-guidance-second-implementation-period-regional-haze-rule

U.S. EPA (2017a), Revisions to the Guideline on Air Quality Models (Appendix W): Enhancements to the AERMOD Dispersion Modeling System and Incorporation of Approaches to Address Ozone and Fine Particulate Matter

U.S. EPA (2017b), Documentation for the EPA's Preliminary 2028 Regional Haze Modeling; Office of Air Quality Planning and Standards, Research Triangle Park NC, October 2017. https://www3.epa.gov/ttn/scram/reports/2028_Regional_Haze_ModelingTSD.pdf

U.S. EPA (2018), Software for the Modeled Attainment Test- Community Edition(SMAT-CE); Office of Air Quality Planning and Standards, Research Triangle Park NC. Available at: https://www.epa.gov/scram/photochemical-modeling-tools

van der A, R. J., Mijling, B., Ding, J., Koukouli, M. E., Liu, F., Li, Q., Mao, H. and Theys, N.: Cleaning up the air: effectiveness of air quality policy for SO2 and NOx emissions in China, *Atmospheric Chemistry and Physics*, 17(3), 1775–1789, doi:10.5194/acp-17-1775-2017, 2017.

ED_002666_00014243-00202

Vickery, J. (2004), Conceptual models of PM for North American regions. In: Particulate Matter Science for Policy Makers: A NARSTO Assessment, P.H. McMurry, M.F. Shepherd, J.S. Vickery, eds., Cambridge University Press.

Vizuete W., B.U. Kim, H. Jeffries, Y. Kimura, D.T. Allen, M.A. Kioumourtzoglou, L. Biton, B. Henderson, (2008), Modeling ozone formation from industrial emission events in Houston, Texas, Atmospheric Environment, 42 (33), pp. 7641–7650.

Vizuete, W., Biton, L., Jeffries, H.E., Couzo, E., (2010), Evaluation of Relative Response Factor Methodology for Demonstrating Attainment of Ozone in Houston, Texas. J. Air Waste Manage. Assoc. 60, 838-848.

Vizuete, W., Jeffries, H.E., Tesche, T.W., Olaguer, E.P., Couzo, E., (2011), Issues with Ozone Attainment Methodology for Houston, TX. J. Air Waste Manage. Assoc. 61, 238-253.

Vogel, B., Riemer, N., Vogel, H., Fiedler, F., (1999) Findings on NOy as an indicator for ozone sensitivity based on different numerical simulations, *J Geophys Res*, 104, 3605-3620.

Vukovich, J.M., and T. Pierce, (2002), The Implementation of BEIS3 within the SMOKE modeling framework, *11th Annual Emissions Inventory Conference of the U.S. EPA,* Session 10, https://www.epa.gov/sites/production/files/201510/documents/vukovich.pdf.

Watson, J. G., J. C. Chow, L.-W.A. Chen, N. H. Frank (2009), Methods to assess carbonaceous aerosol sampling artifacts for IMPROVE and other long-term networks, *J. Air Waste Manage. Assoc.*, 59, 898-911.

White, W. H.: IMPROVE data advisory: Shift in EC/OC split with 1 January 2005 TOR hardware upgrade, 2007.

White, W.H. (2008), Chemical markers for sea salt in IMPROVE aerosol data, *Atmospheric Environment*, 42, 261-274.

Wilczak, J. M., Djalaova, I., McKeen, S., Bianco, L., Bao, J. W., Grell, G., Peckham, S., Mathur, R., McQueen, J., Lee, P., (2009), Analysis of regional meterology and surface ozone during the TexAQS II field program and evaluation of the NMM-CMAQ and WRF-Chem air quality models. J. Geophys. Res. 114, pp. 1-22.

Wong, D.C., Pleim, J., Mathur, R., Binkowski, F., Otte, T., Gilliam, R., Pouliot, G., Xiu, A., Young, J.O., Kang, D. (2012) WRF-CMAQ Two-Way Coupled System with Aerosol Feedback: Software Development and Preliminary Results, *Geoscientific Model Development*, **5**, 299-312. www.geosci-model-dev.net/5/299/2012/ doi:10.5194/gmd-5-299-2012

Xie, Y., Elleman, R., Jobson, T., Lamb, B., (2011), Evaluation of O3-NOx-VOC sensitivities predicted with the CMAQ photochemical model using Pacific Northwest 2001 field observations, *JGR- Atmospheres*, 116, D20303, doi:_10.1029/2011JD015801.

ED_002666_00014243-00203

Yarwood, G., T. Sakulyanontvittaya, O. Nopmongcol and B. Koo, (2014), Ozone Depletion by Bromine and Iodine over the Gulf of Mexico – Final Report. Prepared for Texas Commission on Environmental Quality. Prepared by ENVIRON International Corporation, Novato, CA. November, https://www.tceq.texas.gov/assets/public/implementation/air/am/contracts/reports/pm/5821110365FY1412-20141109-environ-bromine.pdf.

Yua, J. Z., H. Yanga, H. Zhanga, and A.K.H. Lau, (2004), Size distributions of water-soluble organic carbon in ambient aerosols and its size-resolved thermal characteristics, *Atmospheric Environment* 38,1061–1071.

Yuan., Z.B., J. Z. Yu, A. K. H. Lau, P. K. K. Louie, and J. C. H. Fung, (2006), Application of positive matrix factorization in estimating aerosol secondary organic carbon in Hong Kong and its relationship with secondary sulfate, *Atmospheric Chemistry and Physics*, 6, 25-34.

Zhang, X. and P. H. McMurry, (1992), Evaporative losses of fine particulate nitrates during sampling, *Atmospheric Environment*, 26, 3305-3312.

Zhang, Y., Wen, X.-Y., Wang, K., Vijayaraghavan, K., Jacobson, M.Z., (2009), Probing into regional O3 and particulate matter pollution in the United States: 2. An examination of formation mechanisms through a process analysis technique and sensitivity study, *JGR*, 114, D22305, doi:10.1029/2009JD011900.

Zhang, Y., P. Karamchandani, T. Glotfelty, D.G. Streets, G. Grell, A. Nenes, F. Yu, and R. Bennartz: (2012a), Development and initial application of the global-through-urban weather research and forecasting model with chemistry (GU-WRF/Chem), *J. Geophys. Res.*, **117**.

Zhang, W., S. L. Capps, Y. Hu, A. Nenes, S. Napelenok, and A. G. Russell, (2012b), Development of the High-Order Decoupled Direct Method in Three Dimensions for Particulate Matter: Enabling Advanced Sensitivity Analysis in Air Quality Models. Geoscientific Model Development . Copernicus Publications, Katlenburg-Lindau, Germany, 5(2):355-368.

Zhou, W., Cohan, D.S., Napelenok, S.L., (2013), Reconciling NOx emissions reductions and ozone trends in the U.S., 2002-2006, *Atmospheric Environment*, 70, 236-244.

Zubrow, A., Eyth, A., Michaels, H., Brzezinski, D., Allen, C., Baek, B.H., (2012), SMOKE-MOVES: Description and Recent Enhancements", International Emission Inventory Conference, Tampa, Florida 2012.

ED_002666_00014243-00204

United States Environmental Protection Agency	Office of Air Quality Planning and Standards Air Quality Assessment Division Research Triangle Park, NC	Publication No. EPA-454/R-18-009 November, 2018

ED_002666_00014243-00205