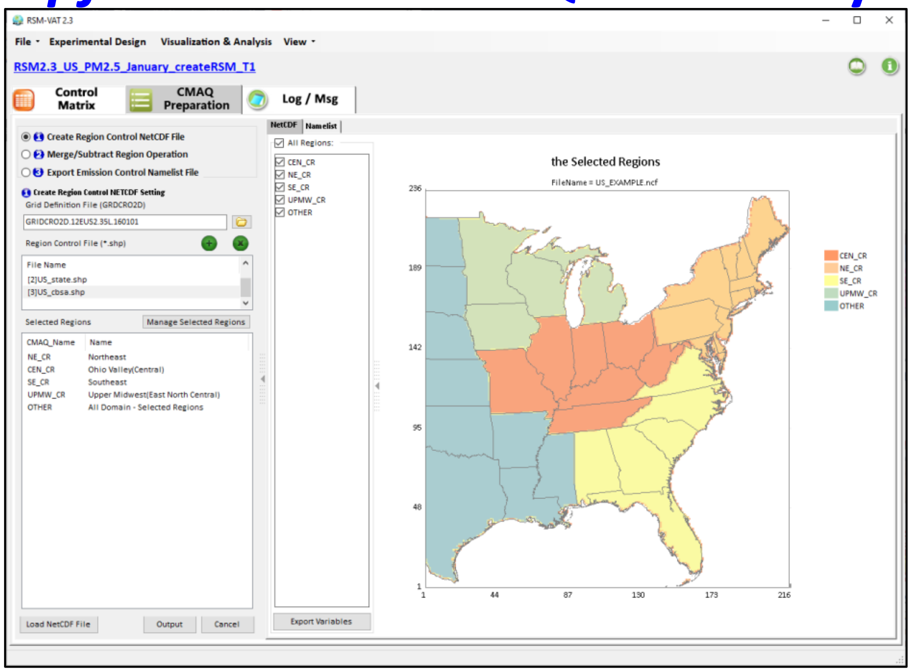
Development and Evolution of Response Surface Model (RSM)

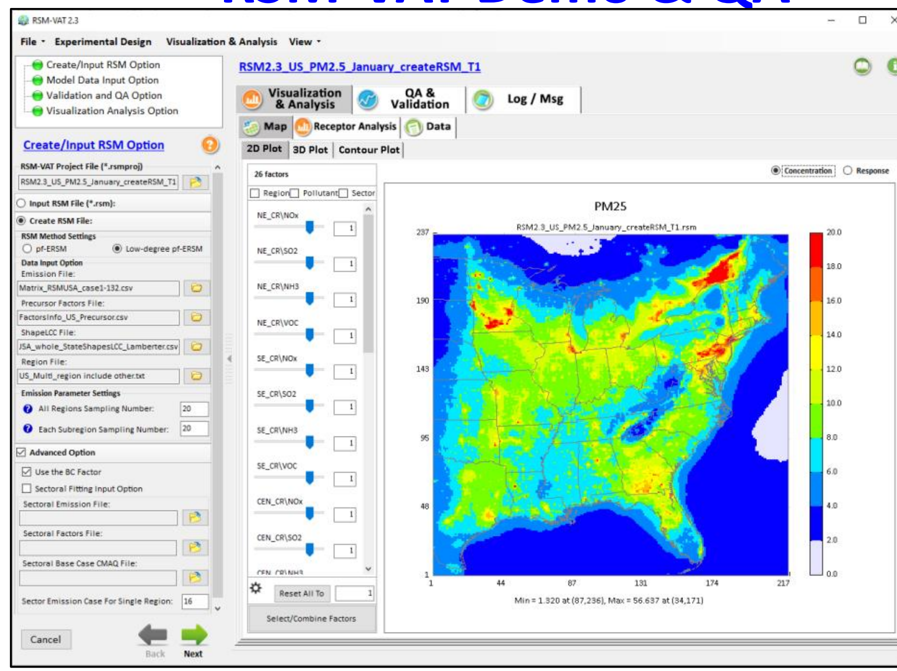
and

RSM-VAT Demo (EUS pilot case)

"pf-RSM" EUS CMAQ Case study



RSM-VAT Demo & QA



Carey Jang and James Kelly, OAQPS/AQAD/AQMG,
08/12/2020

Outline

1. *Development and Evolution of RSM*

- **Traditional *RSM***
- **Extended Multi-Region RSM (*E-RSM*)**
- **Polynomial Function RSM (*pf-RSM*)**
- **"Deep Machine Learning" "Indicator-based" *pf-RSM* (*DeepRSM*) (under development)**

2. *RSM-VAT Tutorial/Demo & QA*

- **CMAQ "Eastern US" 12-km pilot case based on *pf-RSM***
- **QA results: "Out-of-Sample" validation**

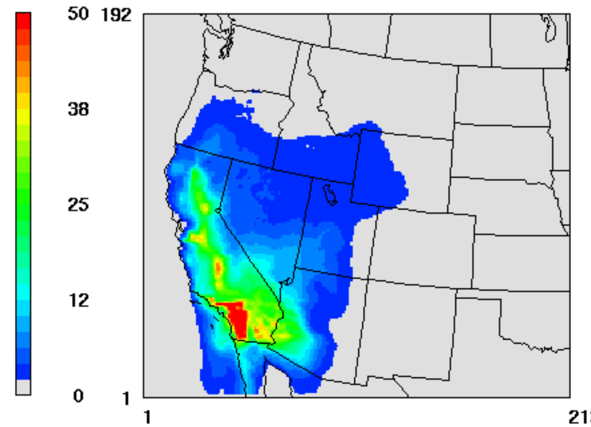
The "RSM-VAT 2.4" package ("RSM-VAT 2.4 Setup.exe" & "Data.exe": run both to install) is available at the Google Drive (or EPA's "ScienceFTP2" site upon request): [*https://drive.google.com/open?id=1Xl3VqtlRXeBt_FrfHpuCZYumxjqMR9hL*](https://drive.google.com/open?id=1Xl3VqtlRXeBt_FrfHpuCZYumxjqMR9hL)

Reduced Form Modeling Techniques:
Source Contribution & Control/Response Modeling

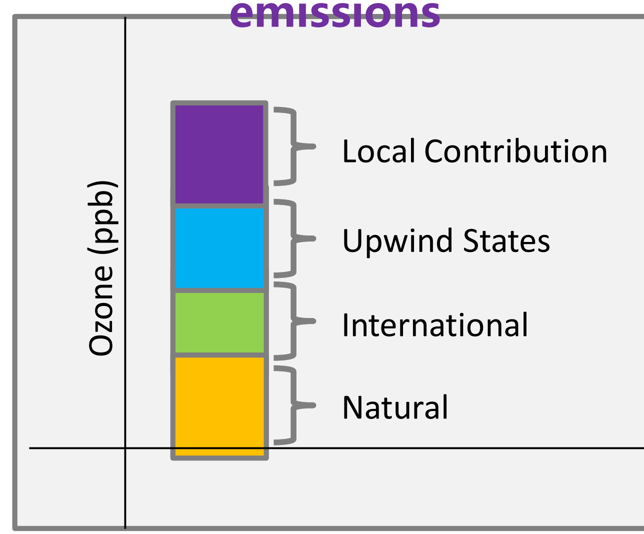
- **"Source Contribution" Assessment**
  - **Source Apportionment Modeling (Chem. approximation)**
    - **CAMx's OSAT/PSAT/APCA, CMAQ's ISAM &TSSA, etc.**
  - **Source/Receptor Sensitivity Modeling (Math. approximation)**
    - **Decoupled Direct Method (DDM/HDDM), Adjoint, etc.**
- **"Control Response" Assessment of emissions reduction (Control efficacy)**
  - **Zero-Out (Brute Force) Modeling**
  - **Response Surface Modeling (RSM): Provide "real-time" AQ response of emissions control**
    - **Previous USA applications: REMSAD (PM2.5):2008, CAMx (O3):2009, CMAQ (EGU RSM): 2011 (Math./Stat. approximation)**
    - **Evolving into Extended RSM (*ERSM*):2015, Polynomial Functions RSM (*pf-RSM*):2018, Deep Machine Learning/Indicator RSM (*DeepRSM*):2020 (Math. & Chem. approximation)**

Source Contribution Analysis for Policy Assessment

- Transport rules or Transport SIPS:
  - Assess "significant contribution" of sources in upwind states to downwind states
- NAAQS Area designations:
  - Provide analysis for States to assess impacts from boundaries
- Exceptional events demonstrations:
  - Assess impacts of special events such as fire
- Inform control strategy development:
  - O3, PM2.5, Reg haze SIPs & RIAs
- Single source Impact assessments:
  - PSD/NSR permitting
- Sector impact assessments:
  - EGU, Transportation, Industrial Boiler, etc.
- Characterizing international contribution:
  - 179b demonstrations



U.S. EPA Source apportionment modeling for assessing source contributions to PM2.5 & O3 from emissions



Ozone Source Contribution (example, not to scale)

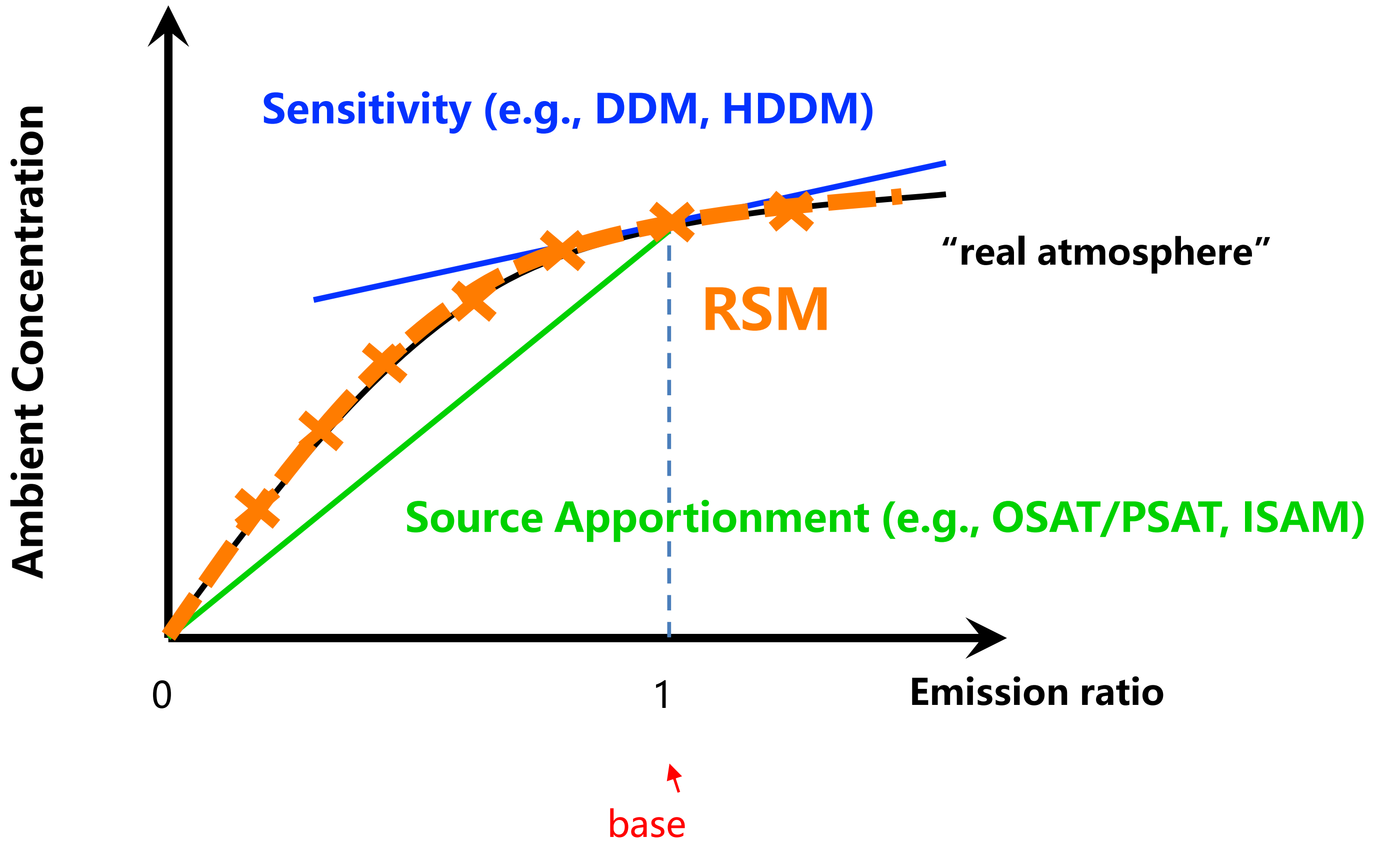
(Courtesy of Dr. Kirk Baker)

RSM: Real-time estimate of air quality responses

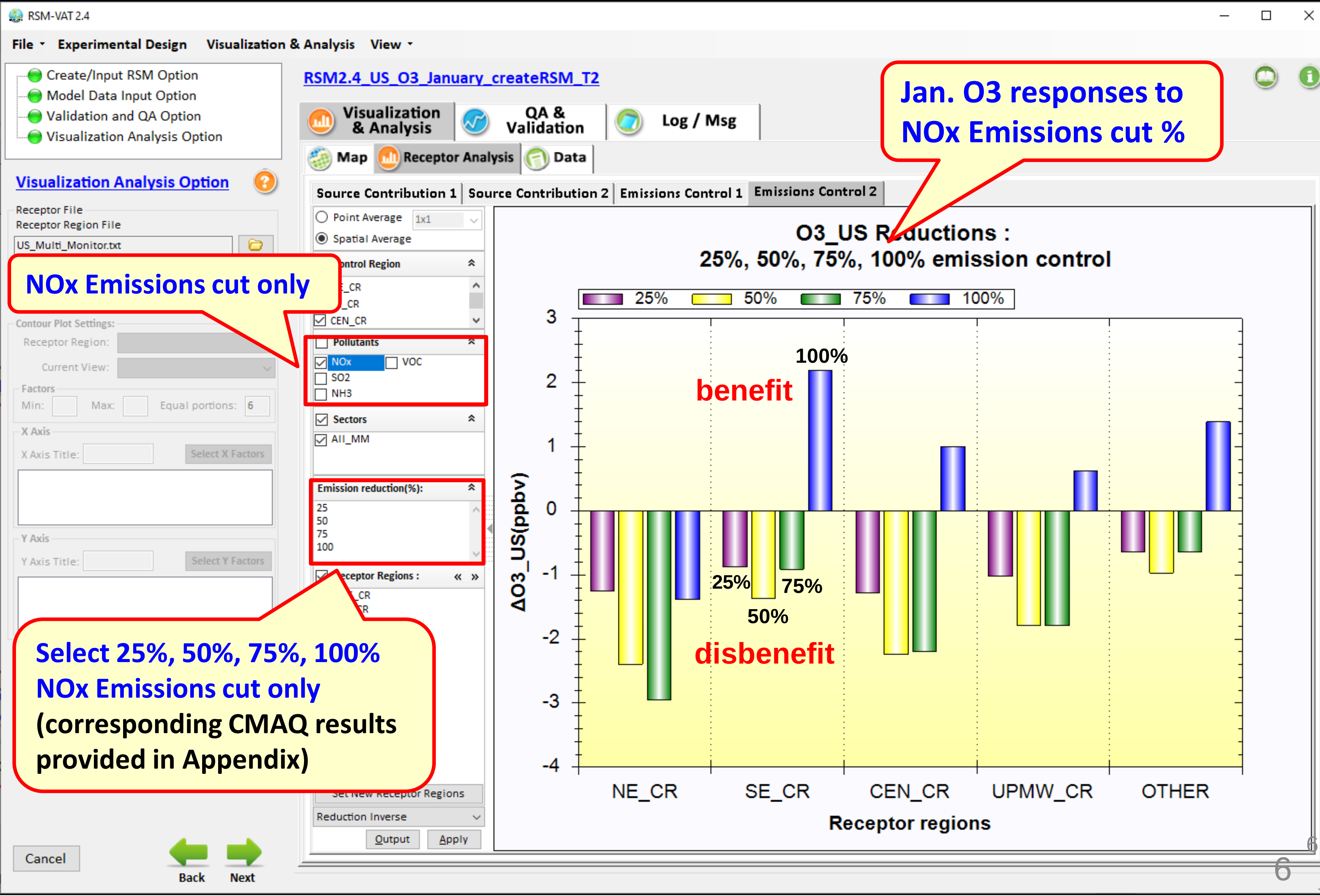
Response Surface Model (RSM):

-- A reduced form meta-model built upon multiple "brute force" AQ model simulations to explore "real-time" response relationship between emission changes and AQ (O3 & PM2.5 species)

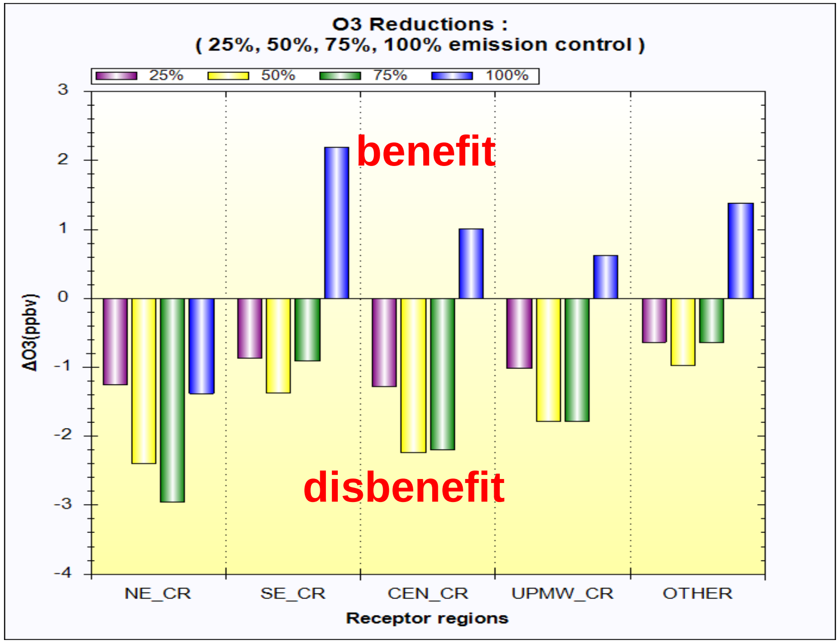
-- Major advantages: provide "real-time" & improved "non-linear" relationship of emissions control and AQ response



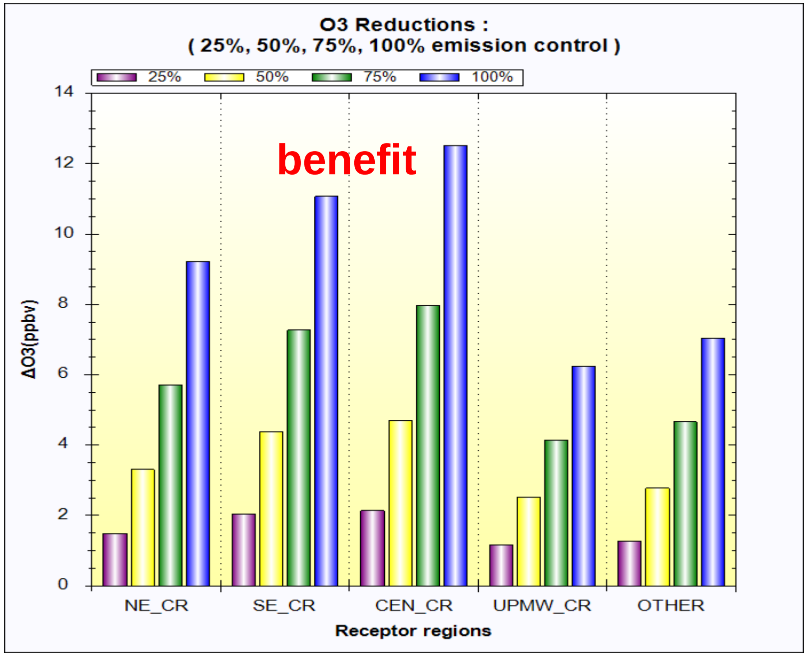
RSM example: "O3" & "PM_SO4" non-linear chemistry



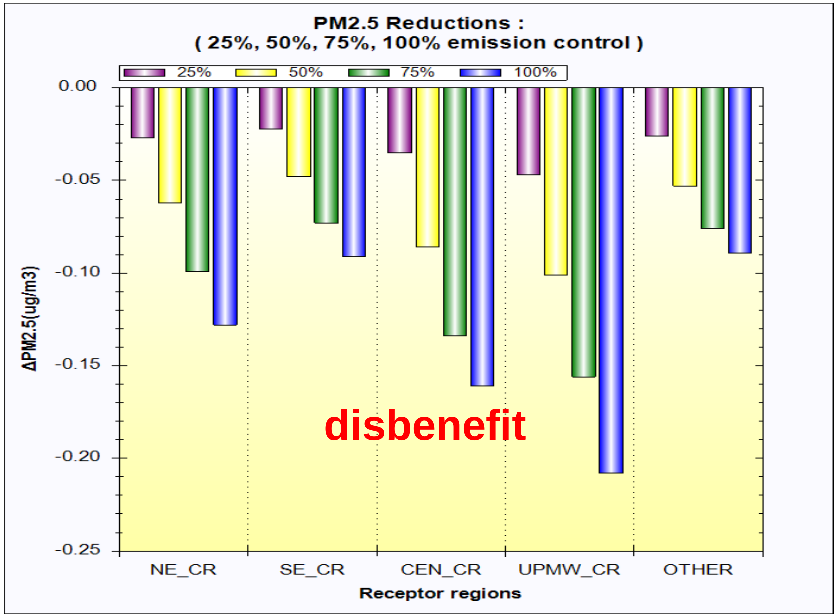
Jan. O3 responses to NOx cut



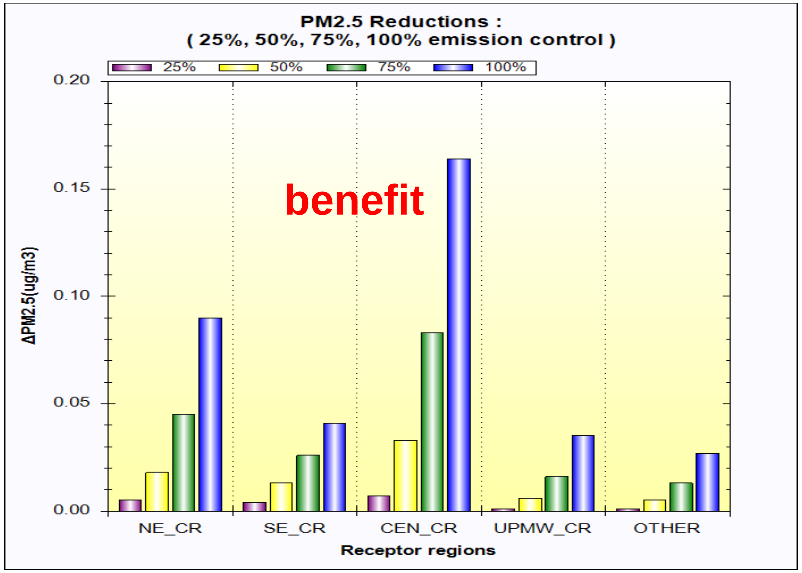
July O3 responses to NOx cut



Jan. PM_SO4 responses to NOx cut



July PM_SO4 responses to NOx cut

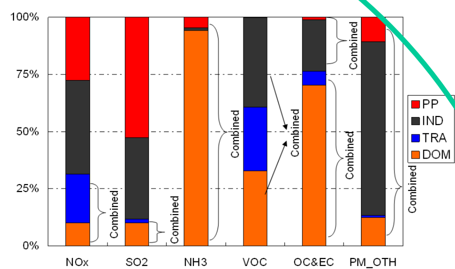


"4S" Steps in RSM (Traditional RSM)

Real-time AQ Response Surface

1.Selection

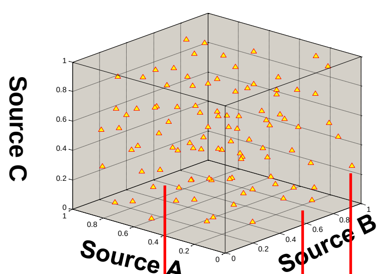
"Region/Pollutant /Sector" Factors



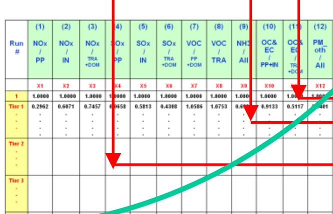
- Source A- Local NOx
- Source B- Regional VOC
- Source C- Area NOx
- …

2.Sampling

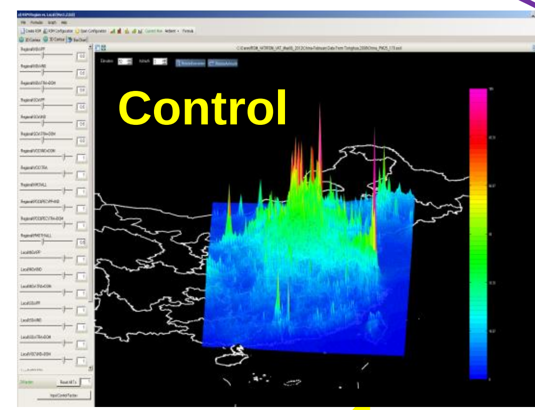
Control Matrix



- Latin Hypercube Sample (LHS)
- Hammersley quasi-random Sequence Sample (HSS)



Experimental Design

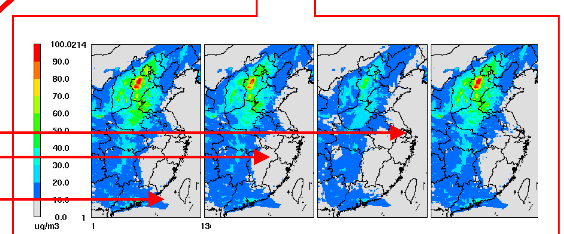


Create RSM

4.Statistic

- Empirical Best Linear Unbiased Predictors (EBLUPs)

$$\vec{Y}(x_0)=\vec{Y}_0=\sum_{j=1}^{d} f_j(x)\vec{\beta}_j+Z(x)=f_0^T\vec{\beta}+\gamma_0^T R^{-1}\left(Y^n-F\vec{\beta}\right)$$



3.Simulation

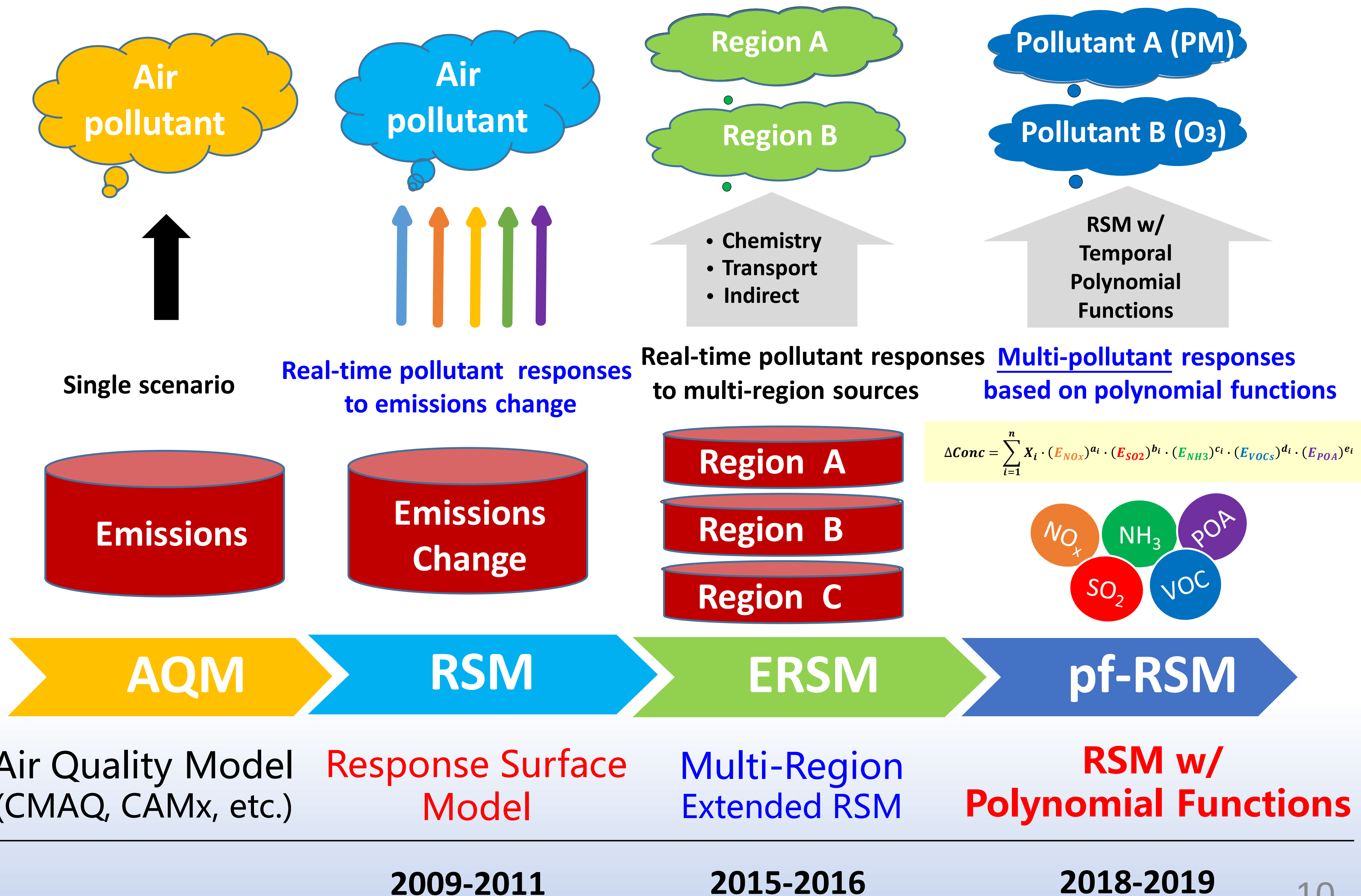
- Multi-case CMAQ simulations

Resource-intensive

Motivation for New RSM Development:

How to reduce # of model simulations while maintaining RSM accuracy?

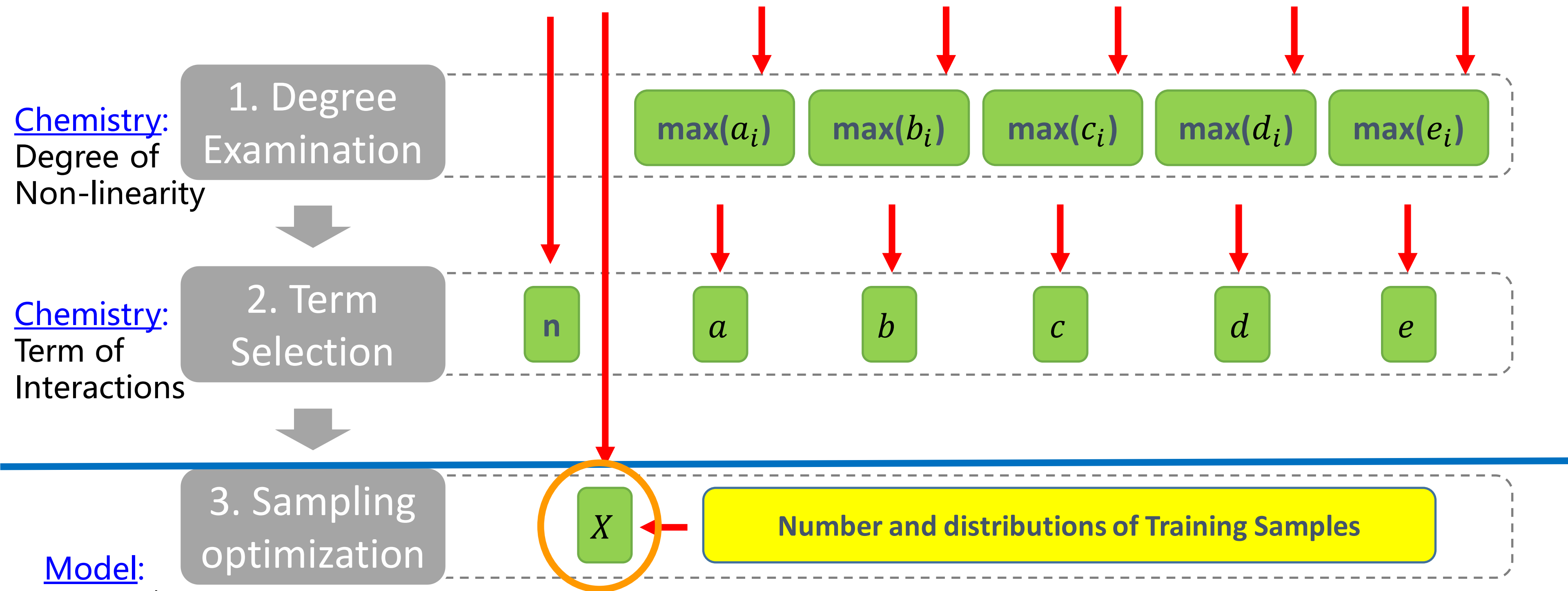
New RSM Development: *"pf-RSM"* & *"DeepRSM"*



"*pf-RSM*" Methodology

Polynomial Functions

pf-RSM: $\Delta Conc = \sum_{i=1}^{n} X_i \cdot (E_{NOx})^{a_i} \cdot (E_{SO2})^{b_i} \cdot (E_{NH3})^{c_i} \cdot (E_{VOCs})^{d_i} \cdot (E_{POA})^{e_i}$

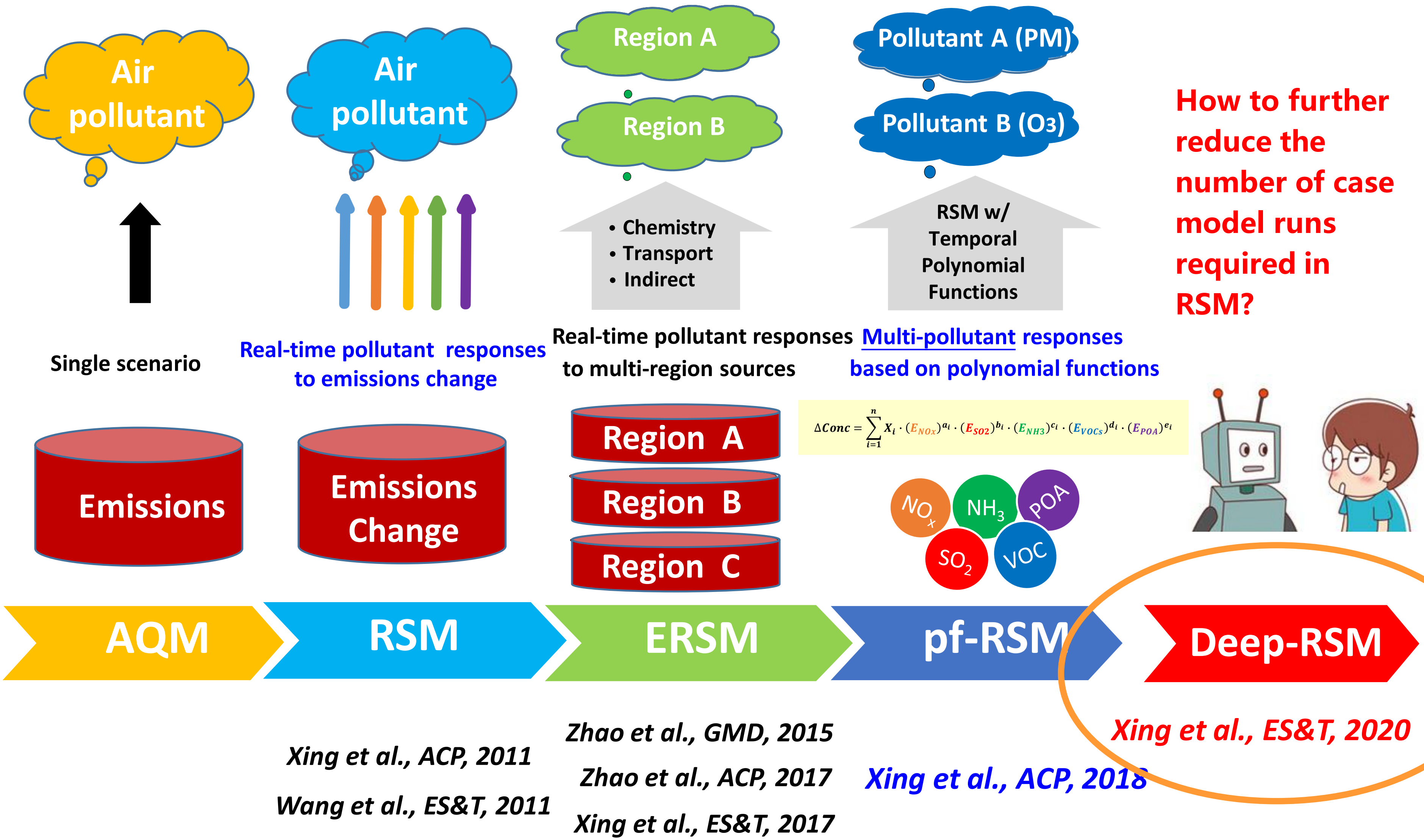


Only 3rd step involved in training samples (i.e., model runs), and thus significantly reduces # of model runs required for accuracy (20~30 runs per control region)

"pf-RSM" : Response Surface Model based on Polynomial Functions (Temporal & multi-region RSM)

Xing et. al., 2018, Atmos. Chem. & Phys.

DeepRSM: Machine Learning w/ Indicator-RSM



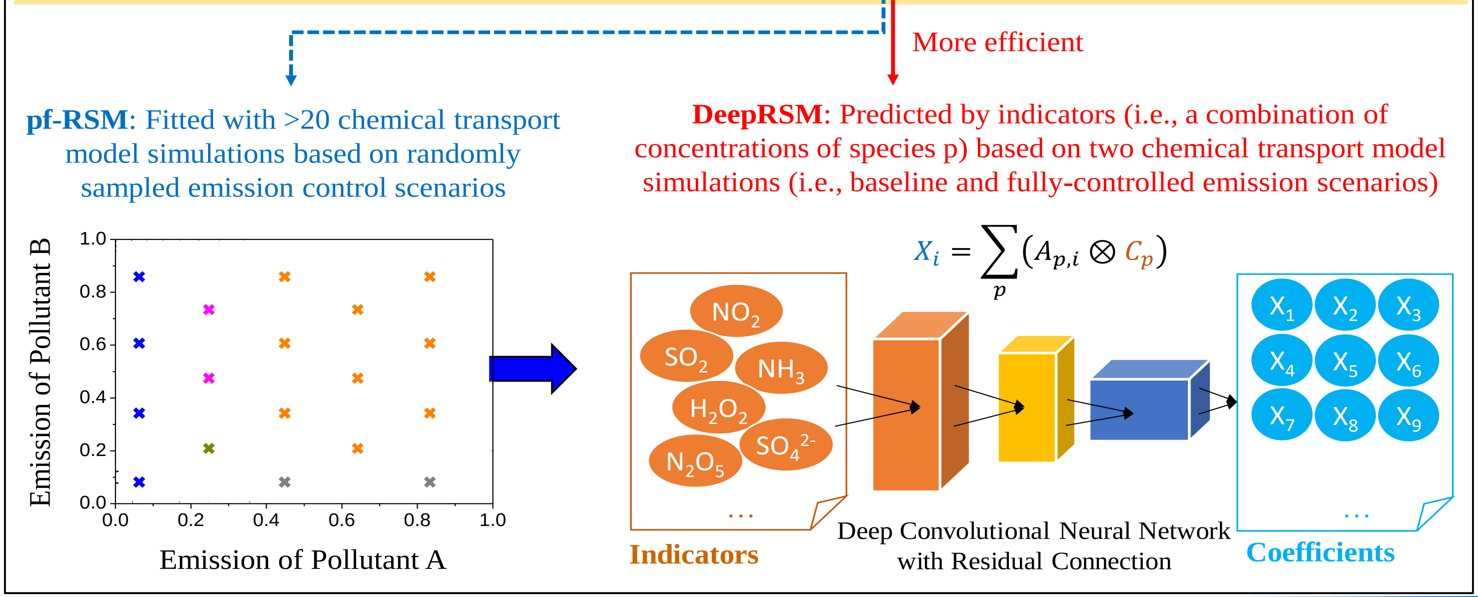
Machine Learning "DeepRSM"

DeepRSM: "Deep Learning" "Indicator-based" pf-RSM

Use 18 Species indicators and deep learning CNN to predict PM2.5 & O3 response to emissions control

Response function of $PM_{2.5}$ and O_3 concentrations to precursor emission change:

$$\Delta Conc = \sum_{i=1}^{n} X_i \cdot (E_{NOx})^{a_i} \cdot (E_{SO2})^{b_i} \cdot (E_{NH3})^{c_i} \cdot (E_{VOCs})^{d_i}$$



# of model runs needed for maintaining RSM accuracy in predicting control responses per selected control region

Traditional RSM	"pf-RSM"	"DeepRSM"
>100s	20~30	2

Xing et. al., ES&T, 2020 (Further details please see Appendix A)

Outline

1. *Development and Evolution of RSM*

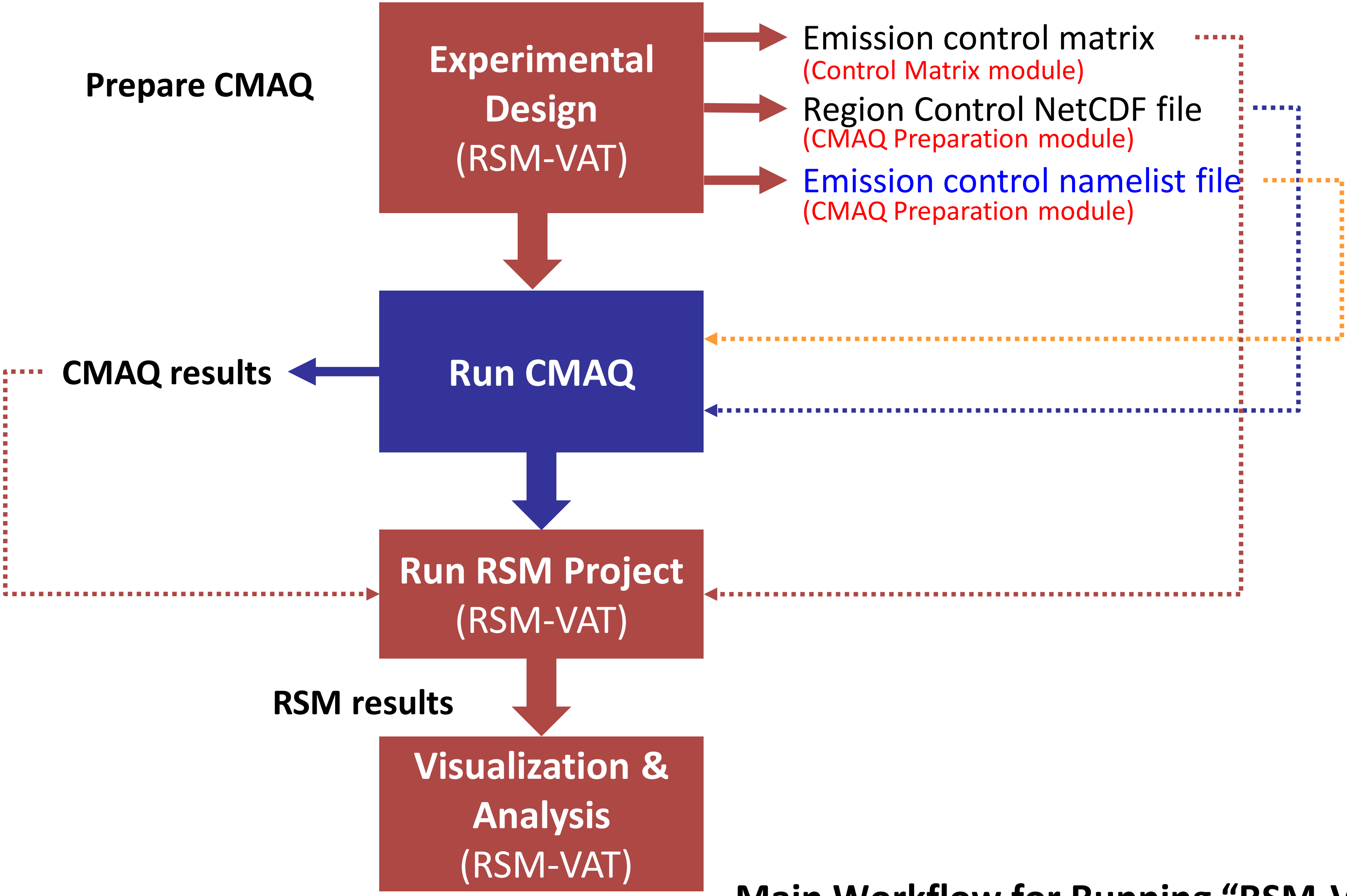
- **Traditional *RSM***
- **Extended Multi-Region RSM (*E-RSM*)**
- **Polynomial Function RSM (*pf-RSM*)**
- **"Deep Machine Learning" Indicator-based *pf-RSM* (*DeepRSM*)**

2. RSM-VAT Tutorial & Demo

- **CMAQ "EUS" 12-km pilot case based on *pf-RSM***
- **QA results: "Out-of-Sample" validation**

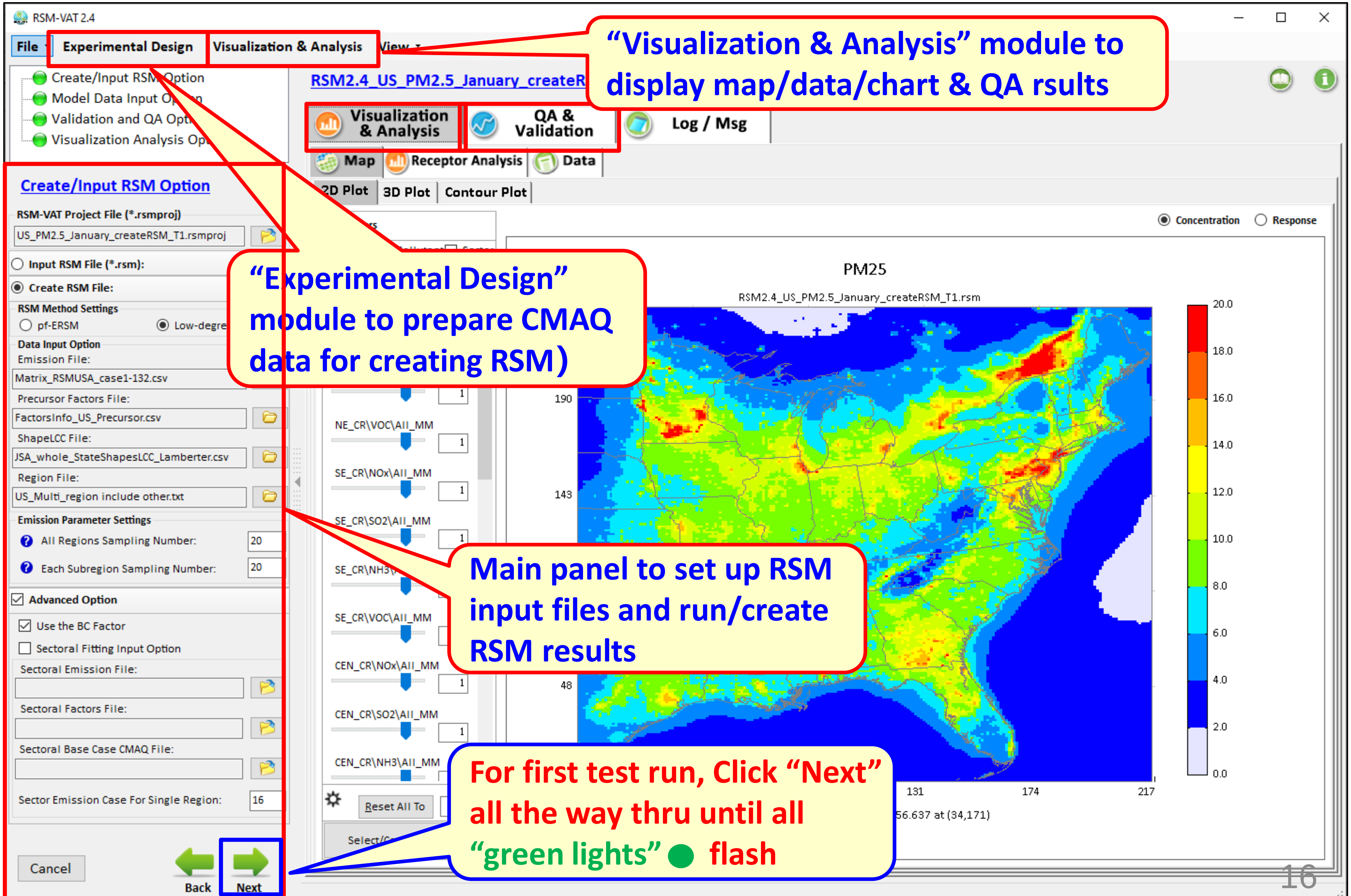
The "RSM-VAT 2.4" package ("RSM-VAT 2.4 Setup.exe" & "Data.exe": run both to install) is available at the Google Drive (or EPA's "ScienceFTP2" site upon request): *https://drive.google.com/open?id=1Xl3VqtlRXeBt_FrfHpuCZYumxjqMR9hL*

Workflow to Create RSM in *"RSM-VAT"*



Main Workflow for Running "RSM-VAT"

RSM-VAT: Main Page & Key Modules



"Visualization & Analysis" module to display map/data/chart & QA rsults

"Experimental Design" module to prepare CMAQ data for creating RSM)

Main panel to set up RSM input files and run/create RSM results

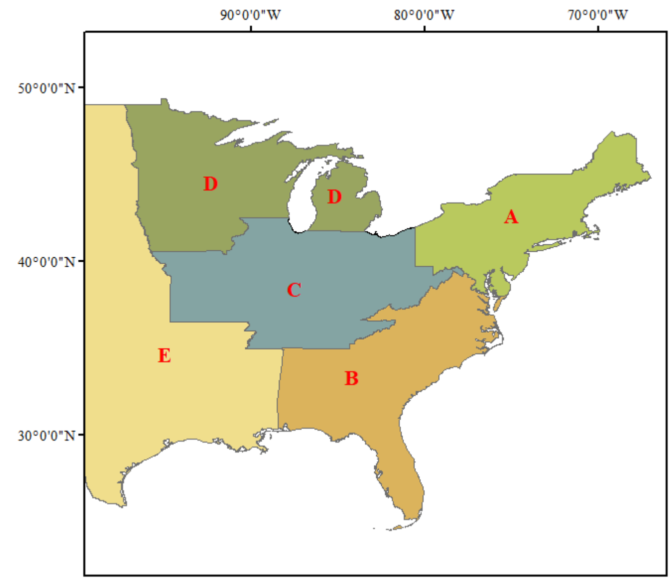
For first test run, Click "Next" all the way thru until all "green lights" ● flash

Experimental Design: CMAQ EUS Case using *"pf-RSM"*

- 5 regions (CR: climate region)

A: Northeast (NE_CR); B: Southeast (SE_CR); C: Central (CEN_CR); D: Upper Midwest (UPMW_CR); E: Other

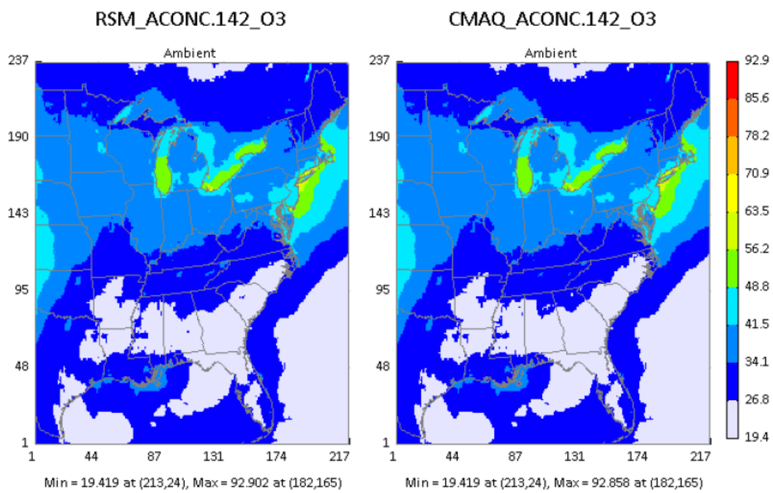
- 5 Pollutants
  - 4 non-linear pollutant (NOx, NH_3, SO_2, VOCs) factors + 1 Primary PM factor
- 1 Sector
  - All manmade (All_MM)
- 26 factors = 5 Pollutants * 5 regions *1 Sector + 1 BC
- 147 CMAQ simulation cases (12-km, January & July)



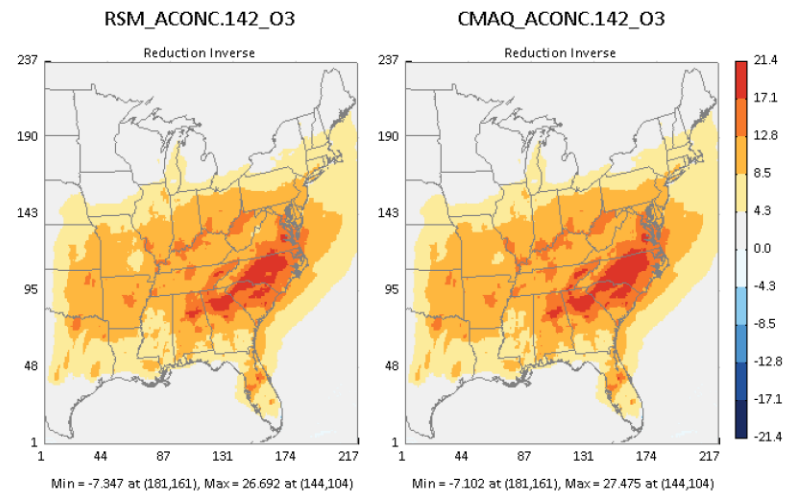
- 1 base case
- 120 cases for non-linear factors = 20 cases * (all region + 5 regions), created by Hammersley quasi-random Sequence Sample (HSS) method (0 ~ 1.2)
- 10 cases for linear Primary PM factors = 5 regions * 2 (0 and 0.5)
- 1 Clean Boundary Conditions
- **15 cases for "out-of-sample" validation** ***(QA for pf-RSM)***

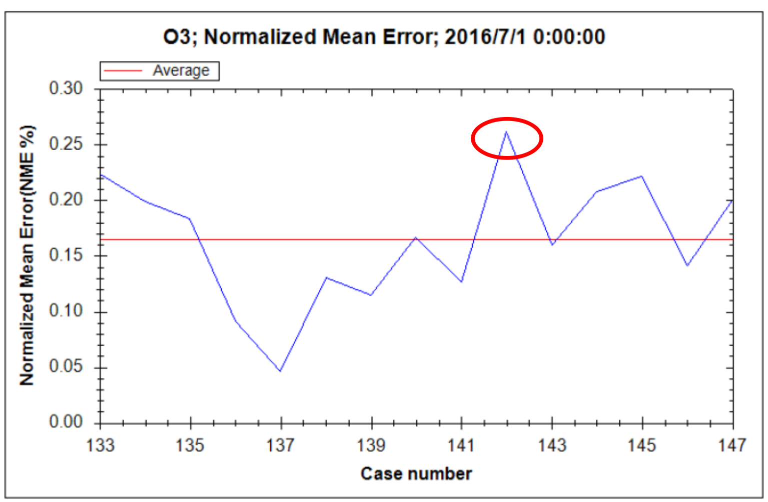
"Out-of-Sample" Validation: O_3 Conc. & Response (July, O_3-8hr avg.): Worst #1

O_3 Conc: RSM vs. CMAQ

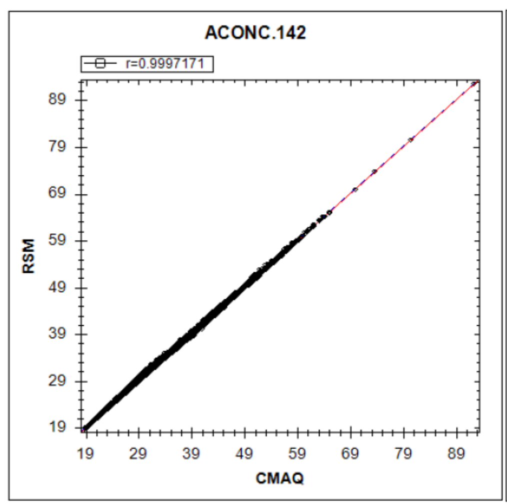


O_3 Response: RSM vs. CMAQ

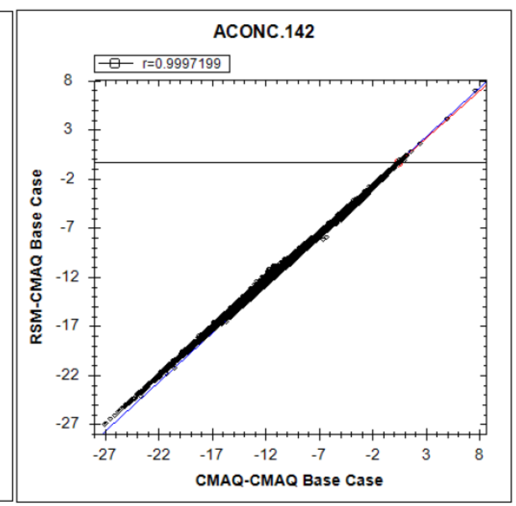




All "Out-of-Sample" runs: Normalized Mean Error



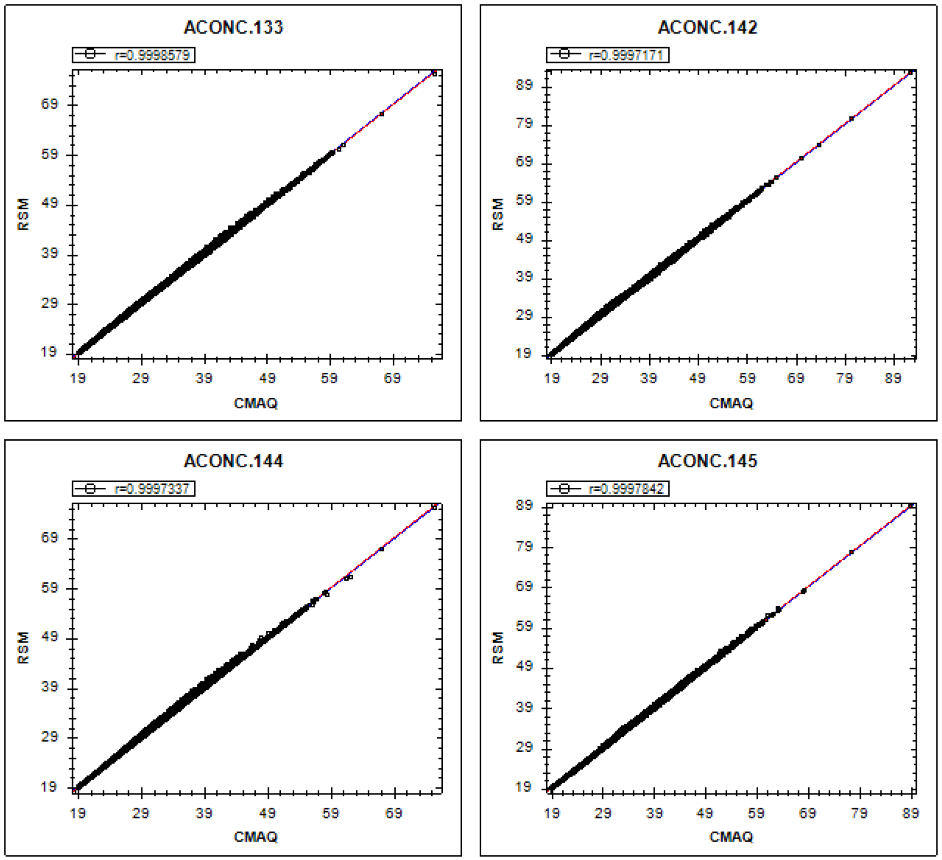
Predicted O_3 Concentration (RSM vs. CMAQ)



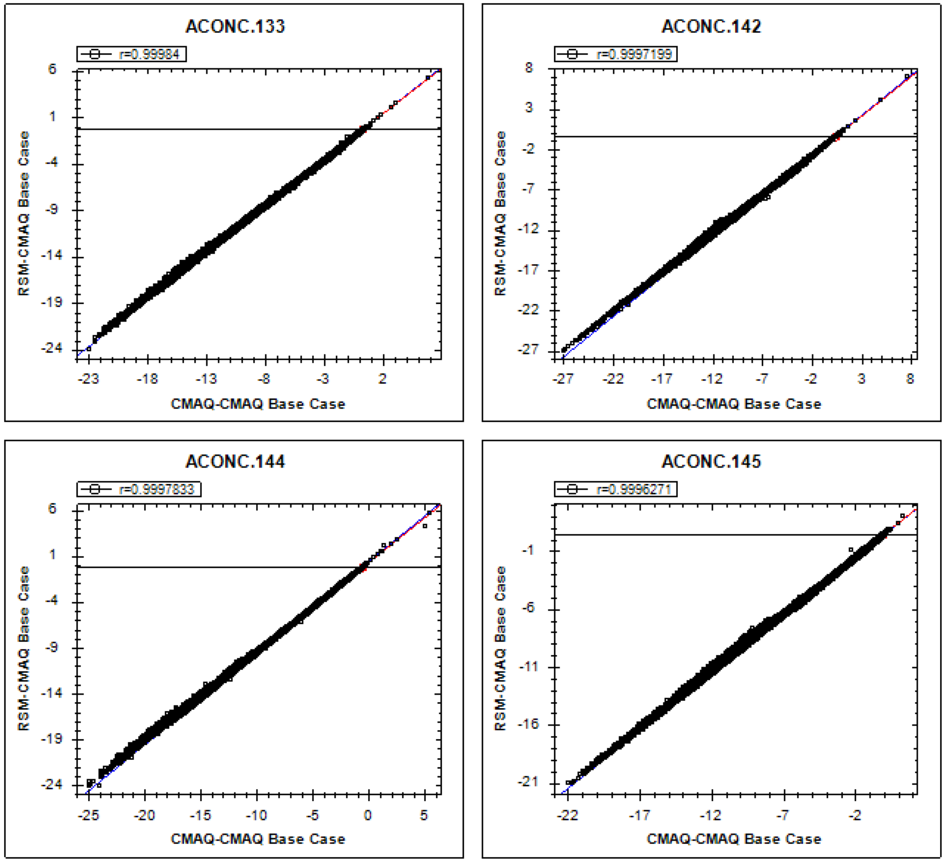
Predicted O_3 Response (RSM vs. CMAQ)

"Out-of-Sample" Validation: O_3 Conc. & Response (July, O_3-8hr avg. , US): Worst 4 cases

Predicted O_3 Conc. (RSM vs. CMAQ)



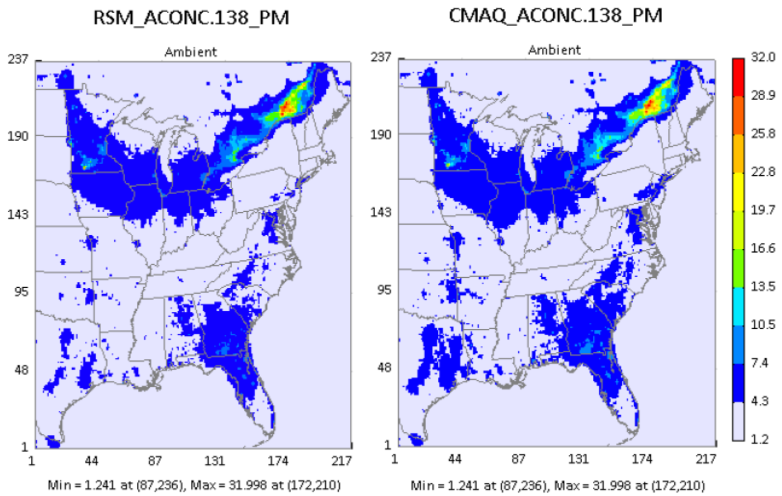
Predicted O_3 Responses (O_3) (RSM-Base vs. CMAQ-Base)



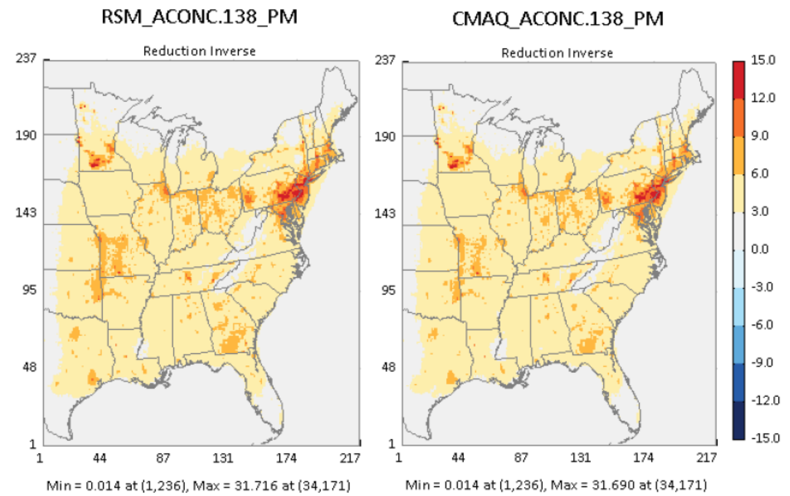
Please see "Appendix B" for additional "out-of-sample" validation slides

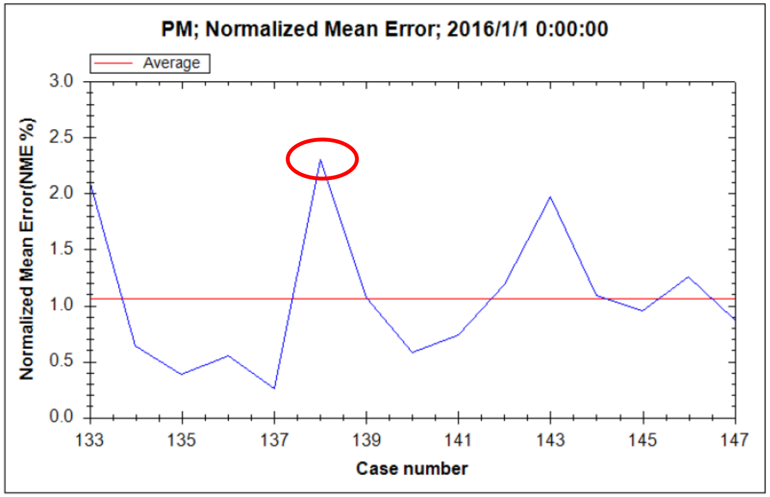
"Out-of-Sample" Validation: $PM_{2.5}$ Conc. & Response (Jan., $PM_{2.5}$ monthly avg.): Worst Case #1

$PM_{2.5}$ Conc: RSM vs. CMAQ

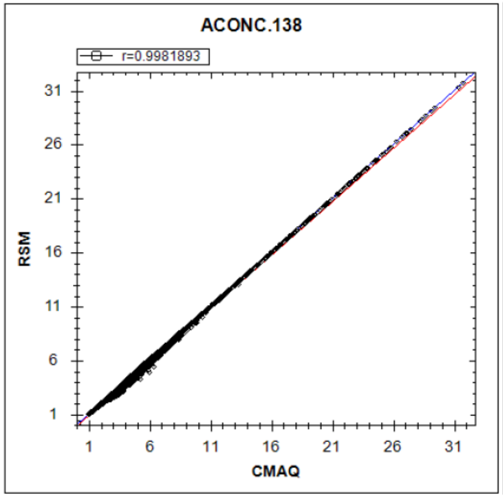


$PM_{2.5}$ Response: RSM vs. CMAQ

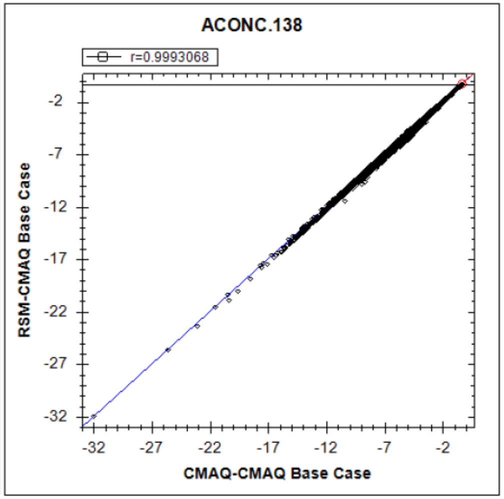




All "Out-of-Sample" runs: Normalized Mean Error



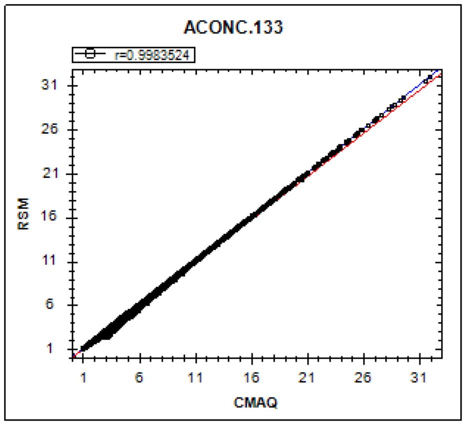
Predicted $PM_{2.5}$ Concentration (RSM vs. CMAQ)

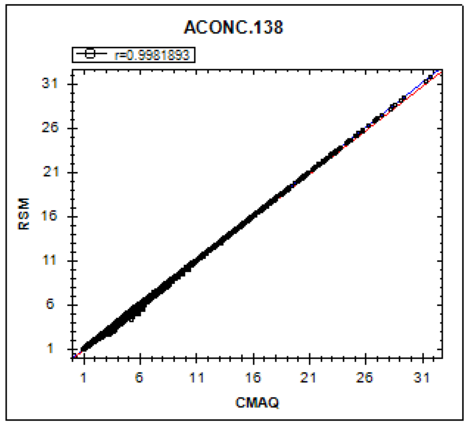


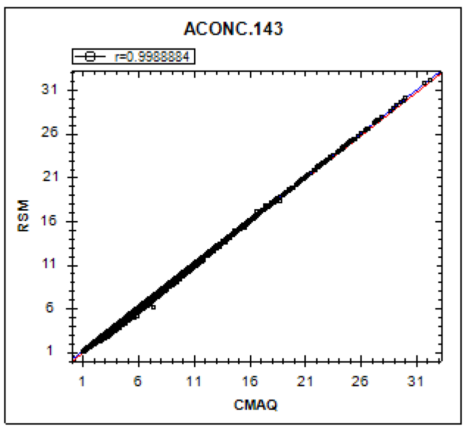
Predicted $PM_{.5}$ Response (RSM vs. CMAQ)

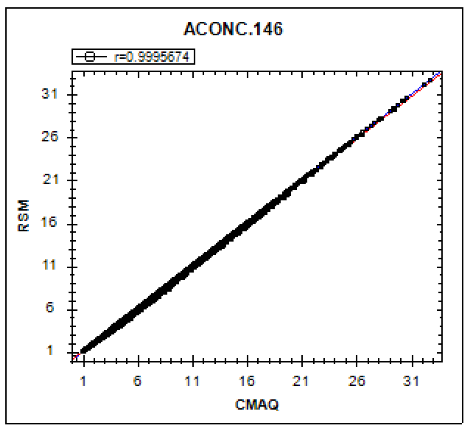
"Out-of-Sample" Validation: PM2.5 Conc. & Response (Jan., $PM_{2.5}$ monthly avg., US): Worst 4 cases

Predicted $PM_{2.5}$ Conc.
(RSM vs. CMAQ)

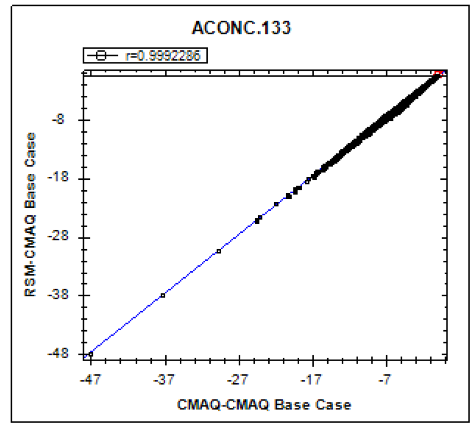


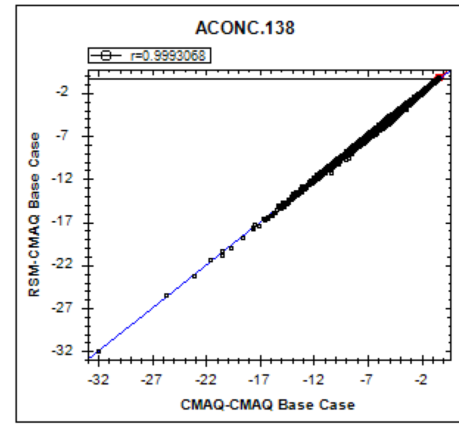


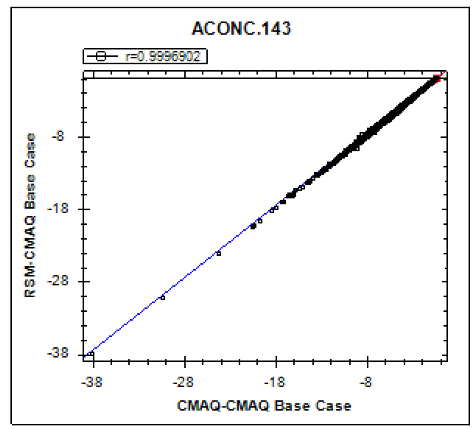


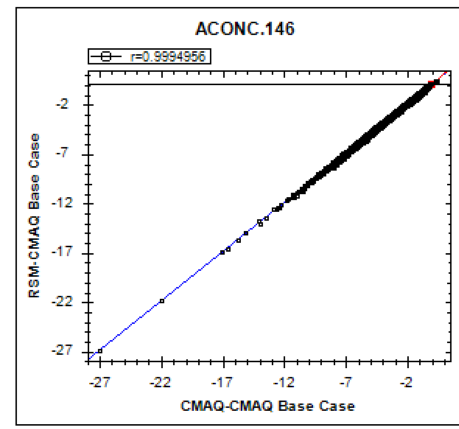


Predicted $PM_{2.5}$ Responses
(RSM-Base vs. CMAQ-Base)





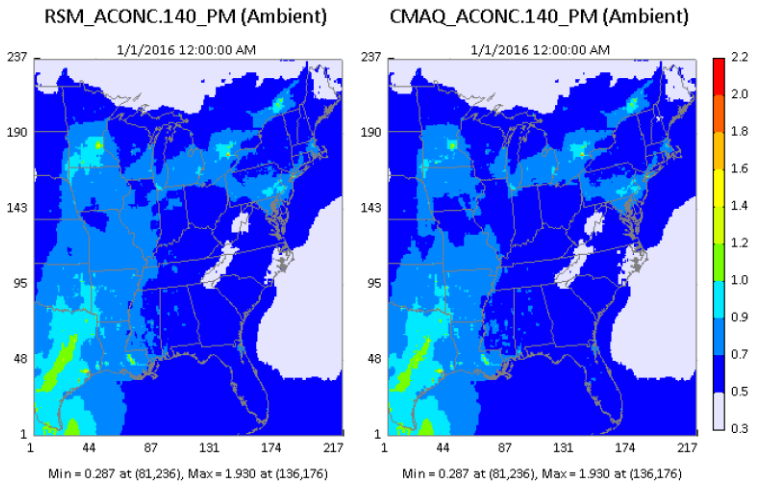




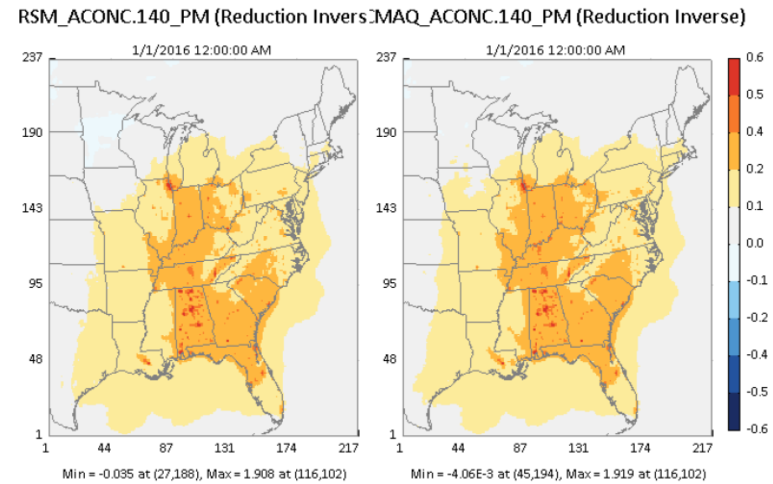
Please see "Appendix B" for additional "out-of-sample" validation slides

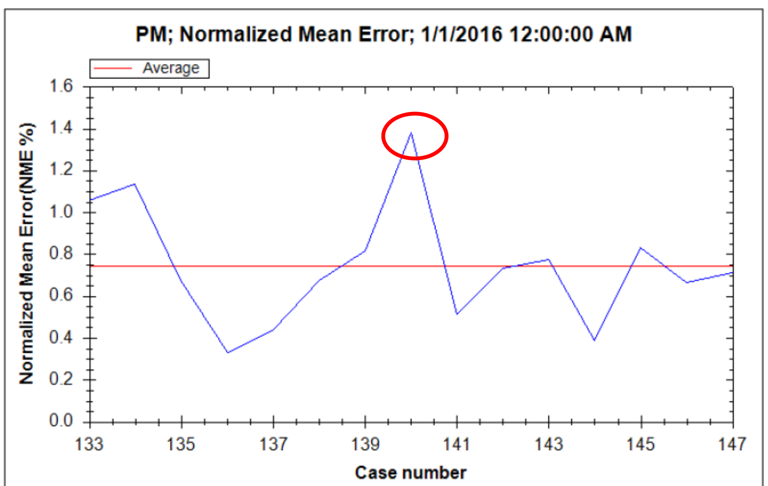
"Out-of-Sample" Validation: PM_SO_4 Conc. & Responses (Jan., PM_SO_4 monthly avg., US): Worst #1

$PM_{2.5}$ Conc: RSM vs. CMAQ

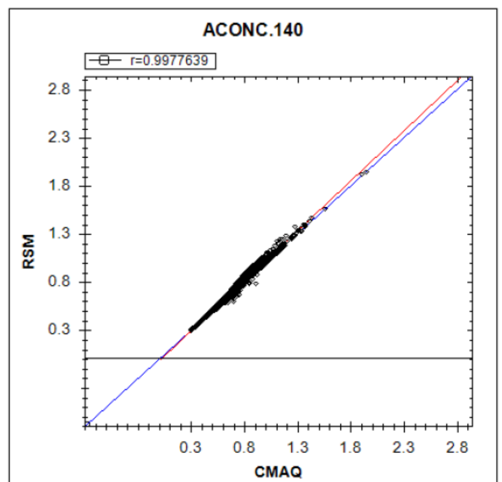


$PM_{2.5}$ Response: RSM vs. CMAQ

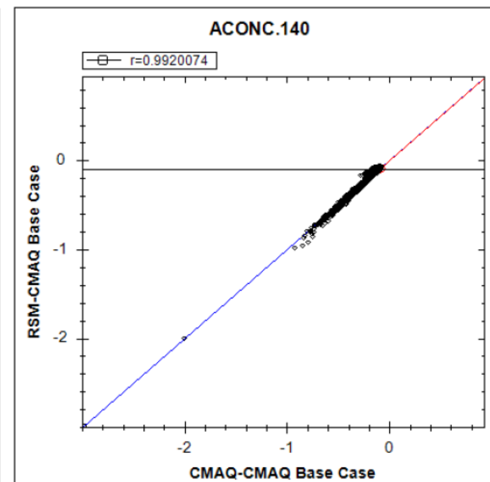




All "Out-of-Sample" runs: Normalized Mean Error



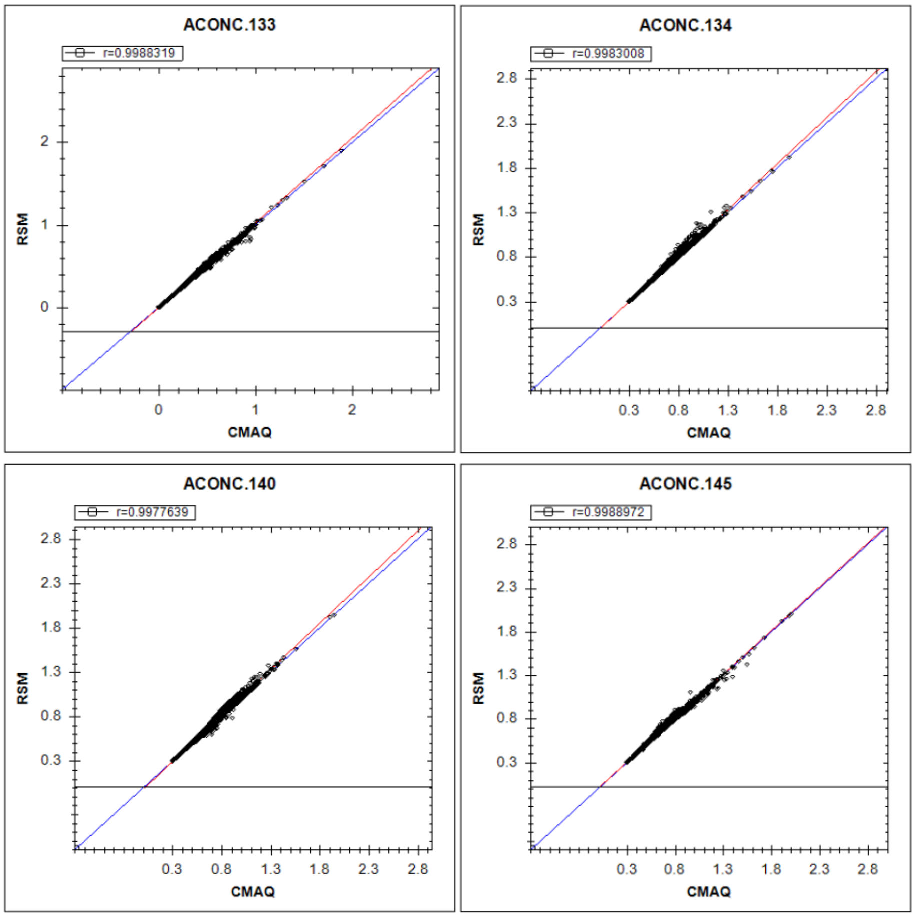
Predicted $PM_{2.5}$ Concentration (RSM vs. CMAQ)



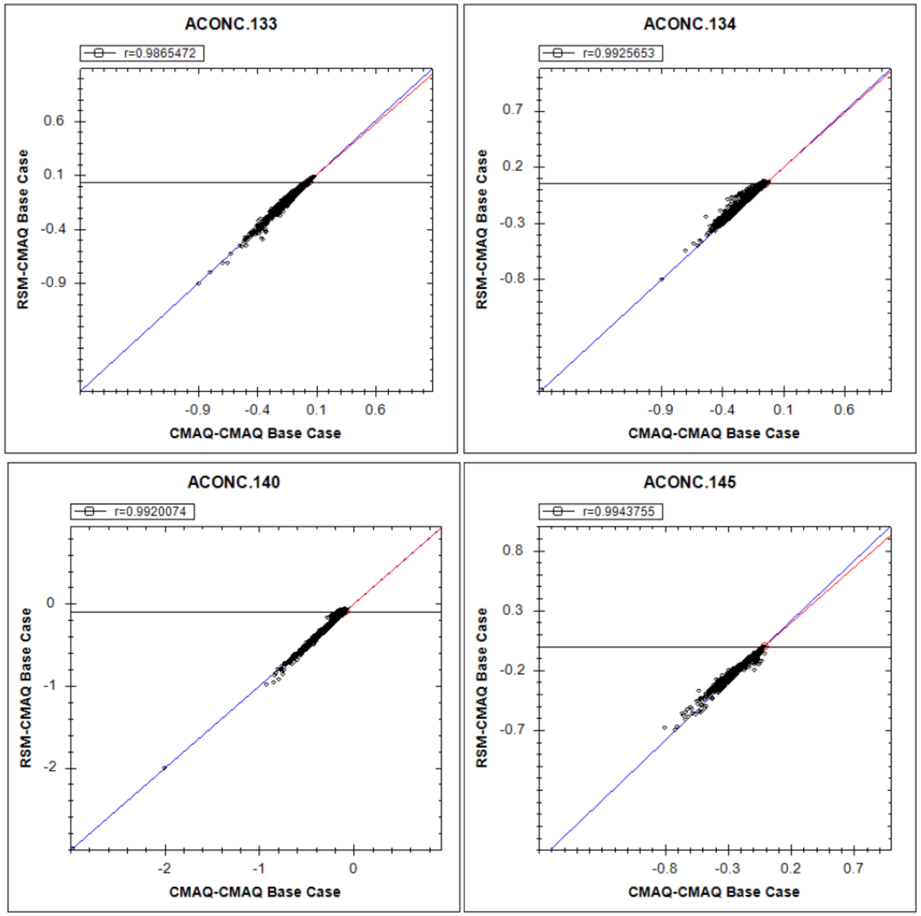
Predicted $PM_{.5}$ Response (RSM vs. CMAQ)

"Out-of-Sample" Validation: PM_SO_4 Concentration & Response (Jan., PM_SO_4 monthly avg.): Worst 4 cases

Predicted $PM_{2.5}$ Conc.
(RSM vs. CMAQ)



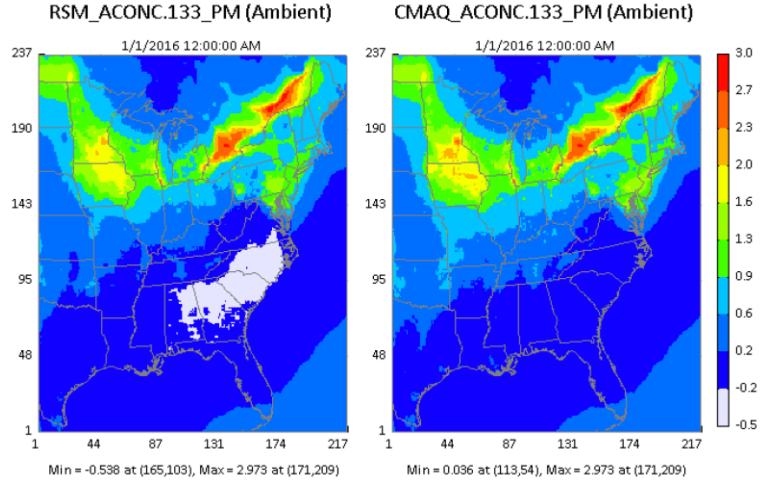
Predicted $PM_{2.5}$ Responses
(RSM-Base vs. CMAQ-Base)



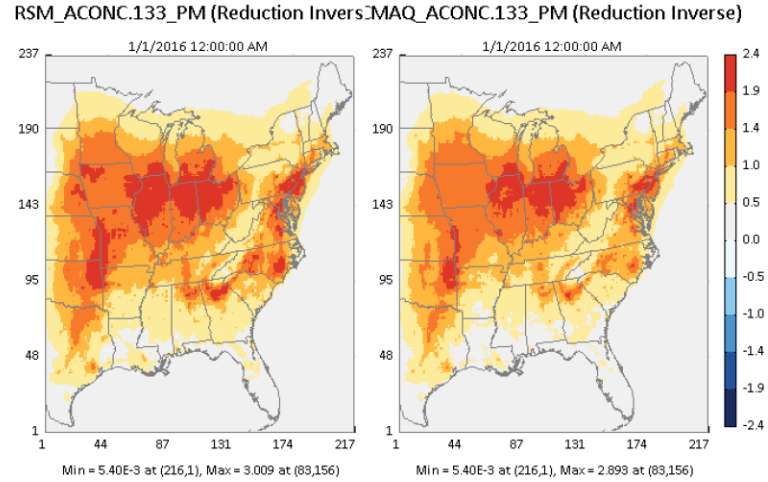
Please see "Appendix B" for additional "out-of-sample" validation slides

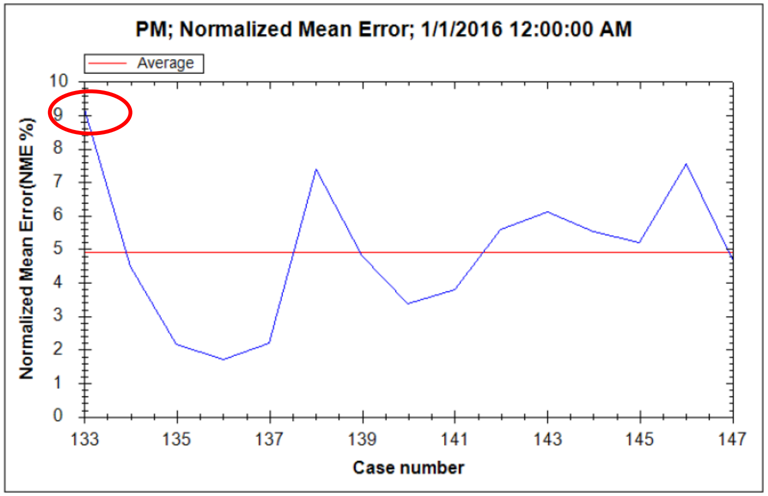
"Out-of-Sample Validation": PM_NO$_3$ Conc. & Responses (Jan., PM_NO$_3$ monthly avg., US): Worst #1

PM$_{2.5}$ Conc: RSM vs. CMAQ

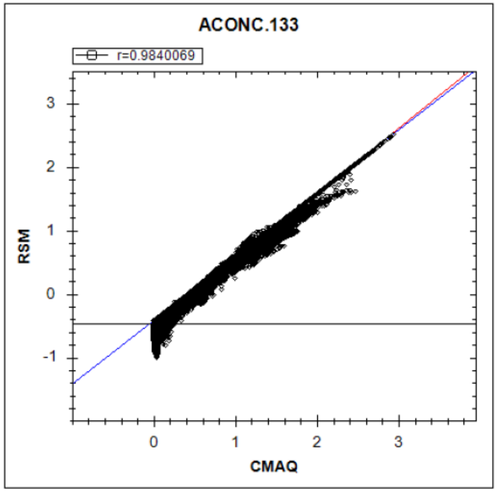


PM$_{2.5}$ Response: RSM vs. CMAQ

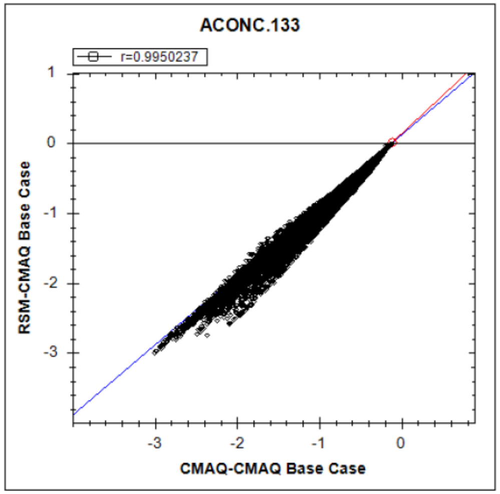




All "Out-of-Sample" runs: Normalized Mean Error



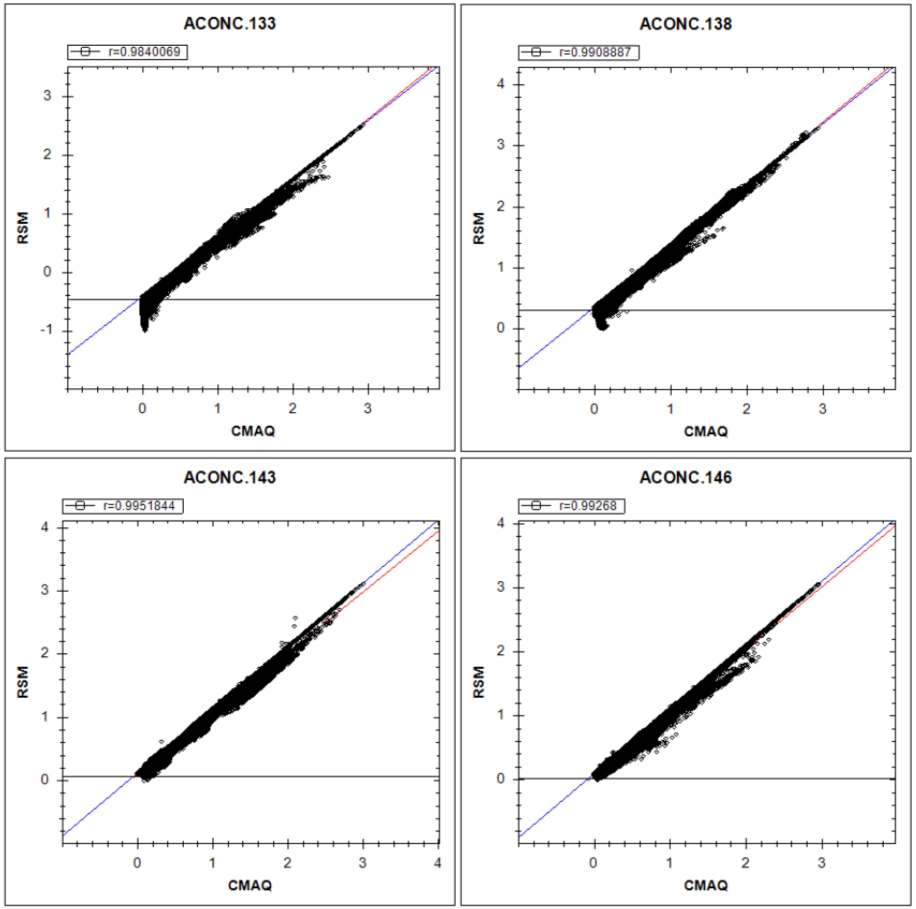
Predicted PM$_{2.5}$ Concentration (RSM vs. CMAQ)



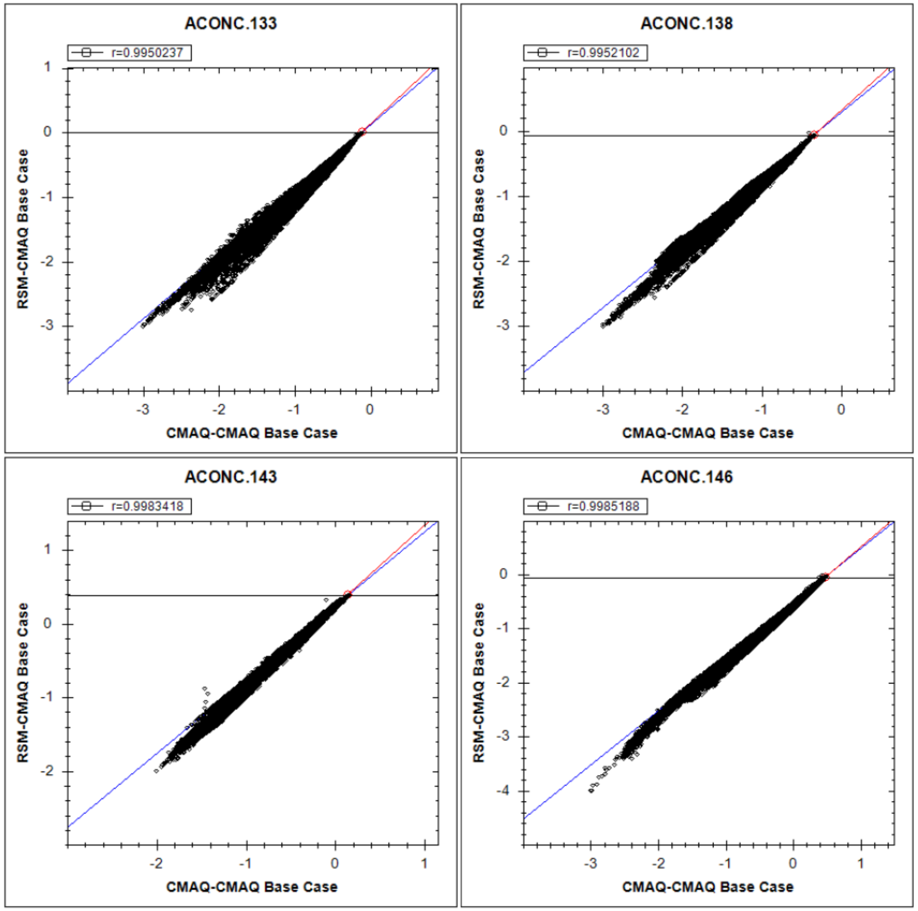
Predicted PM$_{.5}$ Response (RSM vs. CMAQ)

"Out-of-Sample" Validation: PM_NO_3 Concentration & Response (Jan., PM_NO_3 monthly avg.): Worst 4 cases

Predicted $PM_{2.5}$ Conc.
(RSM vs. CMAQ)



Predicted $PM_{2.5}$ Responses
(RSM-Base vs. CMAQ-Base)



Please see "Appendix B" for additional "out-of-sample" validation slides

"RSM-VAT" Demo

Main GUI:

1. Launch/open PM2.5 Jan. project
2. Main GUI -> File/Analysis/View
3. Menu: VA/QA/Log&Msg
4. Traffic Light: Each module & input files
5. Help: Online User's Manual

Module -> Experimental Design module

1. "Control Matrix" creation
2. "CMAQ Preparation"
   - Create Region (default CR, State & CBSA)
   - Merge/Substract
   - Export Control Namelist File

V/A module:

1. MAP -> 2D/3D/Contour
2. Conc./Response toggle
3. Radio button: Control %: Set & Reset
4. NOx/SO2/NH3 & Primary PM
5. Output to SMAT/BenMAP/CSV & color legend
6. Chart: EC 1 &2: NOx: 25/50/75/100%
7. Region/Pollutant/Sector checked
8. Chart: SC 1: NOx, 20%/40%/60%/80%/100%
9. SC2: NOx & Primary PM

QA/Validation module:

1. OOSV: RSM vs. CMAQ
   - "Conc." toggle "to Response"
2. OOSV: Comparison Plot
   - "Ambient" to "Reduction"
   - Show Difference Plot
   - Adjust legend/scale
3. "CMAQ Viewer": O3 & PM2.5 species

O3 July case:

1. Launch/open O3 July project
2. V/A->MAP: Concs & Response toggle
3. V/A-> Chart: NOx: 25/50/75/100%
4. QA/Validation:
   - OOSA: RSM vs. CMAQ
   - OOSA: Comparison Plot

Discussion & Next Steps

Next Steps:

- Implement *"DeepRSM"* under RSM-VAT
- Conduct *"DeepRSM"* U.S. pilot case studies under 2016 or 2017 modeling platform (CMAQ & CAMx?)

Potential Applications & Projects

- Hemispheric CMAQ applications
- Inter-comparisons with APCA (O_3/PM) & DDM/HDDM (O_3)

Potential Issue:

- No current year funding for RSM development; Need new funding to support *"DeepRSM"* pilot project under new GSA contract starting Oct. 1st, 2020

Appendix A:

"DeepRSM" Supplemental information

Appendix B:

RSM CMAQ EUS Case "QA":

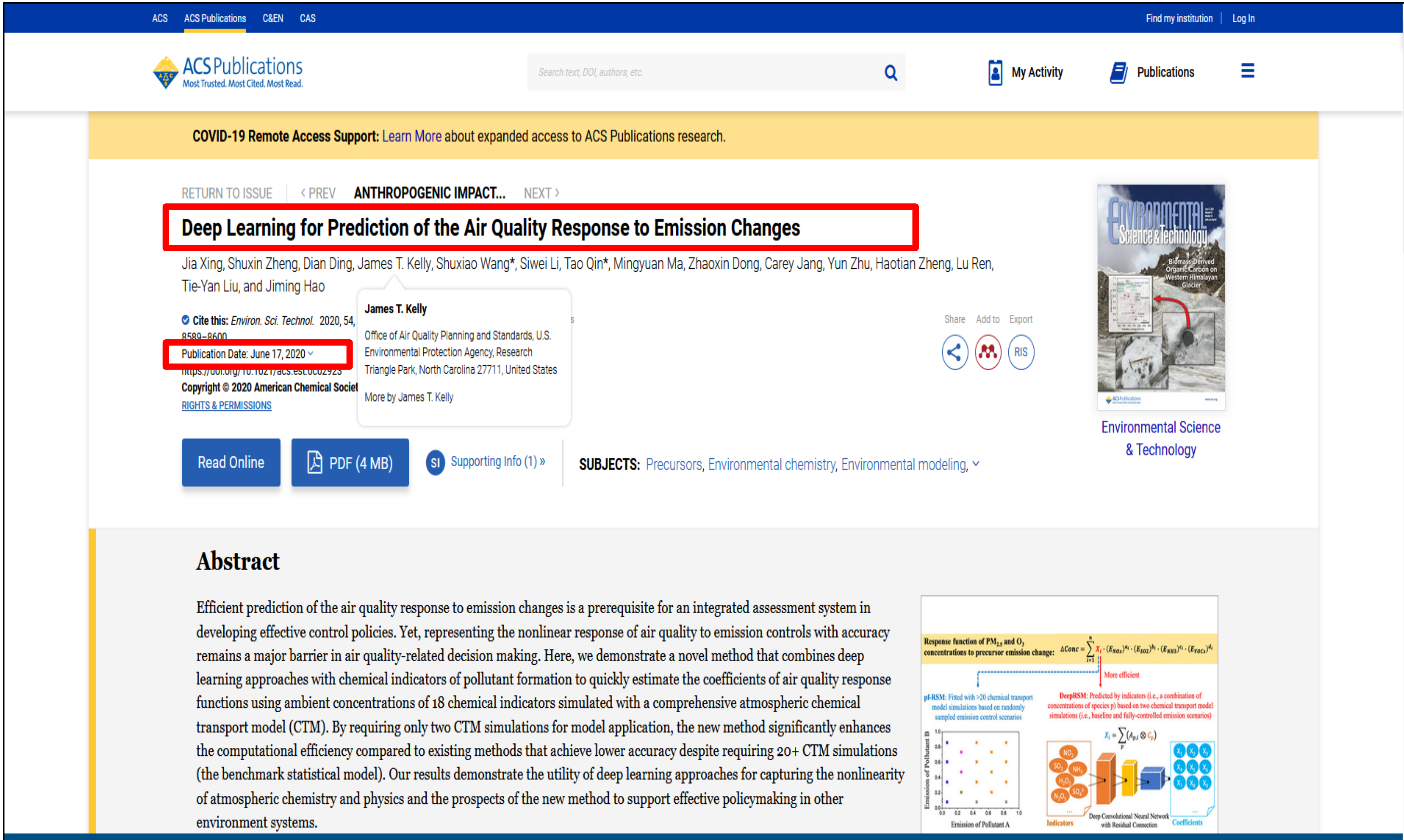
pf-RSM "Out-of-Sample" Validation

Appendix C:

RSM-VAT Demo & Tutorial:

EUS 12-km Case using "pf-RSM"

DeepRSM Publication (2020)


ACS ACS Publications C&EN CAS Find my institution | Log In

ACS Publications Most Trusted. Most Cited. Most Read.

My Activity Publications

COVID-19 Remote Access Support: Learn More about expanded access to ACS Publications research.

RETURN TO ISSUE ‹ PREV ANTHROPOGENIC IMPACT... NEXT ›

Deep Learning for Prediction of the Air Quality Response to Emission Changes

Jia Xing, Shuxin Zheng, Dian Ding, James T. Kelly, Shuxiao Wang*, Siwei Li, Tao Qin*, Mingyuan Ma, Zhaoxin Dong, Carey Jang, Yun Zhu, Haotian Zheng, Lu Ren, Tie-Yan Liu, and Jiming Hao

James T. Kelly
Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711, United States
More by James T. Kelly

Cite this: *Environ. Sci. Technol.* 2020, 54, 8589–8600

Publication Date: June 17, 2020

https://doi.org/10.1021/acs.est.0c02923

Copyright © 2020 American Chemical Societ

RIGHTS & PERMISSIONS

Share Add to Export

Environmental Science & Technology

Read Online PDF (4 MB) Supporting Info (1) »

SUBJECTS: Precursors, Environmental chemistry, Environmental modeling,

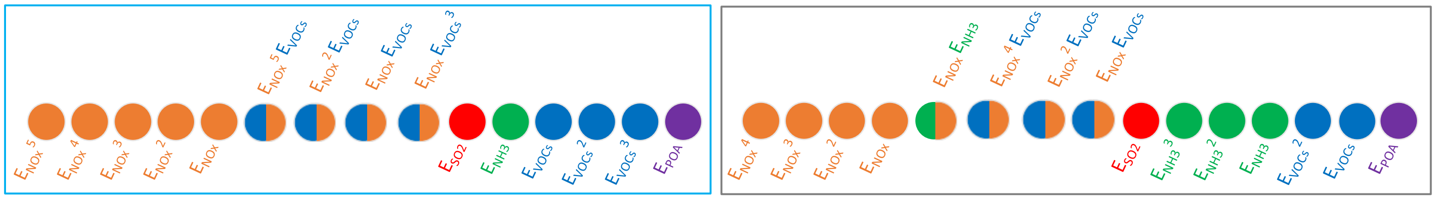
Abstract

Efficient prediction of the air quality response to emission changes is a prerequisite for an integrated assessment system in developing effective control policies. Yet, representing the nonlinear response of air quality to emission controls with accuracy remains a major barrier in air quality-related decision making. Here, we demonstrate a novel method that combines deep learning approaches with chemical indicators of pollutant formation to quickly estimate the coefficients of air quality response functions using ambient concentrations of 18 chemical indicators simulated with a comprehensive atmospheric chemical transport model (CTM). By requiring only two CTM simulations for model application, the new method significantly enhances the computational efficiency compared to existing methods that achieve lower accuracy despite requiring 20+ CTM simulations (the benchmark statistical model). Our results demonstrate the utility of deep learning approaches for capturing the nonlinearity of atmospheric chemistry and physics and the prospects of the new method to support effective policymaking in other environment systems.

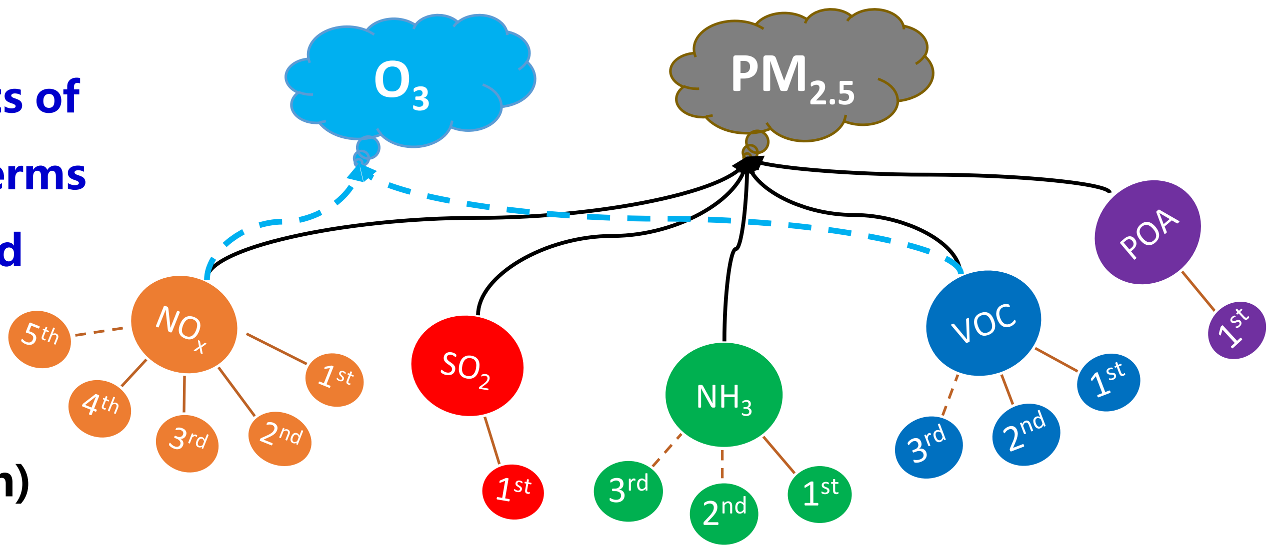
Xing et. al., ES&T, 2020 https://pubs.acs.org/doi/10.1021/acs.est.0c02923

RSM with polynomial functions: *pf-RSM*

$$\Delta Conc = \sum_{i=1}^{n} X_i \cdot (E_{NOx})^{a_i} \cdot (E_{SO2})^{b_i} \cdot (E_{NH3})^{c_i} \cdot (E_{VOCs})^{d_i} \cdot (E_{POA})^{e_i}$$



14 coefficients of polynomial terms can be derived from 20~30 samples (per control region)

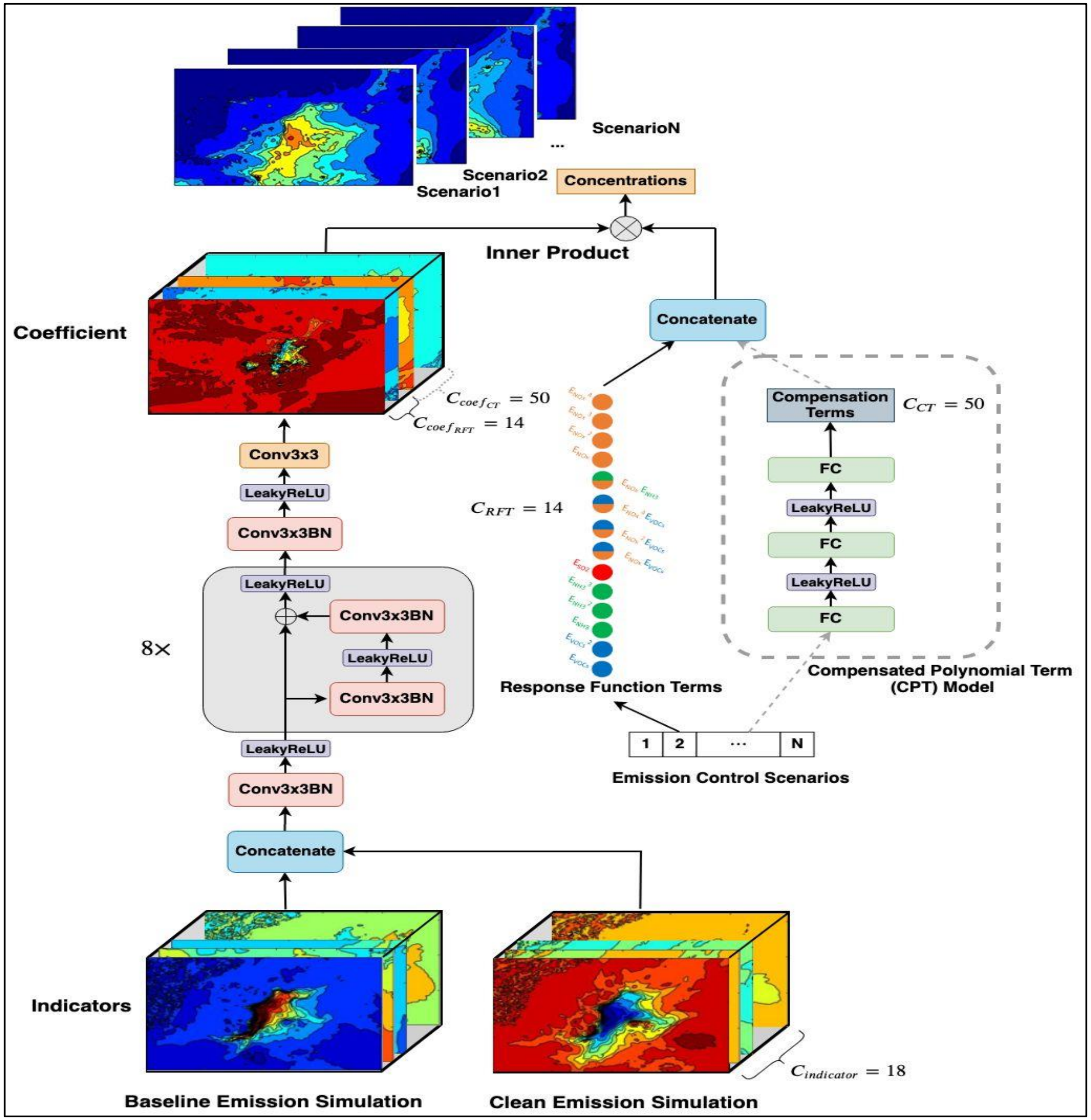


High order for NOx and VOC

Linear behavior for SO_2 and POA

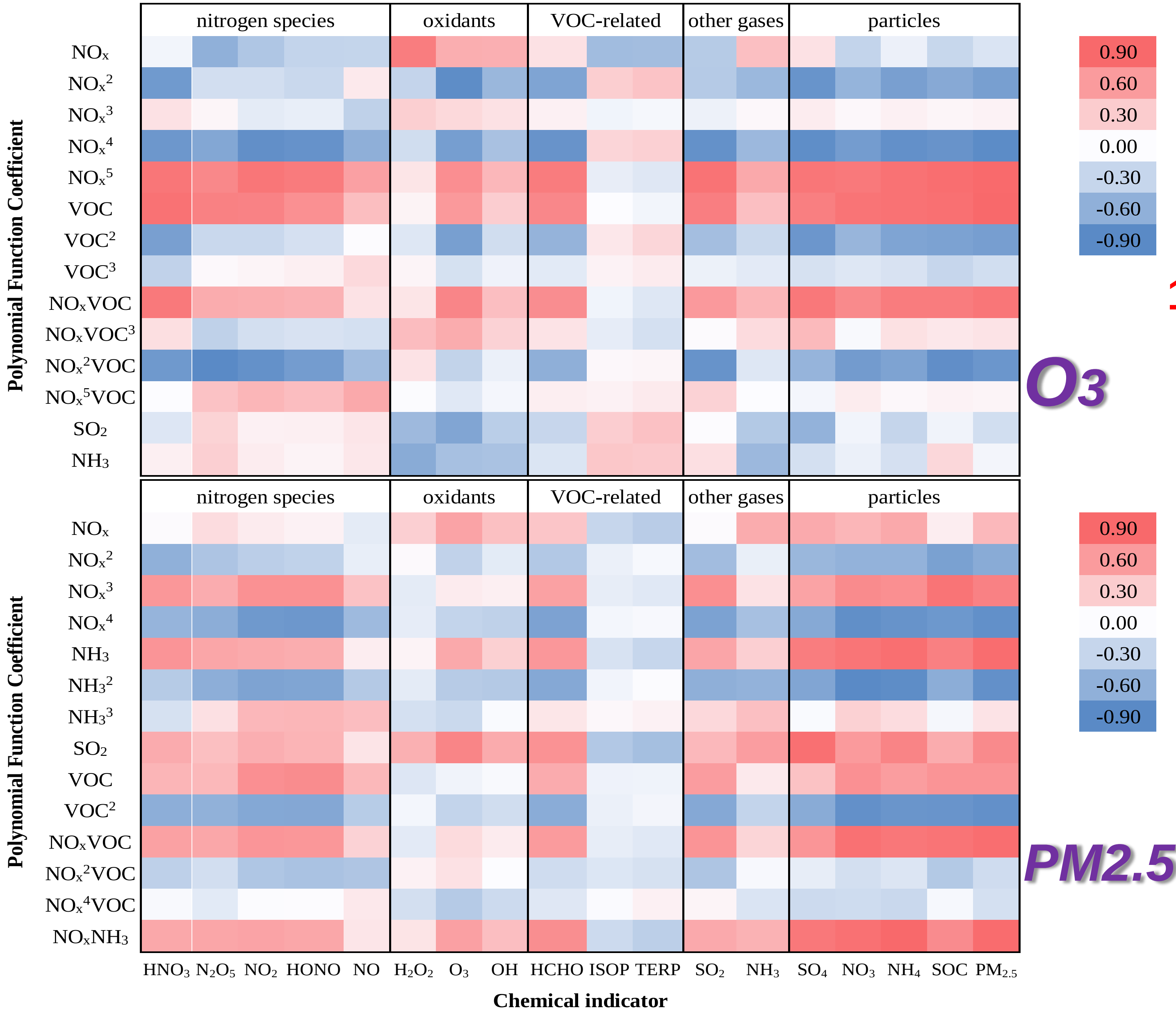
Xing et al., ACP, 2018

DeepRSM: Machine Learning with Indicator-RSM



The CNN architecture of the DeepRSM with optimized polynomial structure for predicting the air quality response functions. ("Conv3x3BN": 3x3 convolution followed by batch normalization; "LeakyReLU": leaky rectified linear unit activation function; "FC": fully connected layer)

“Indicator-based” DeepRSM



14 coefficients of terms in in responding to 18 chemical indicators for O_3 and $PM_{2.5}$ (example)

Appendix B:

QA for RSM CMAQ EUS Case:

pf-RSM "Out-of-Sample" Validation

1. O_3 July 8-hr avg.
2. PM2.5 July monthly avg.
3. PM2.5 Jan. monthly avg.
4. PM_SO4 Jan. monthly avg.
5. PM_NO3 Jan. monthly avg.

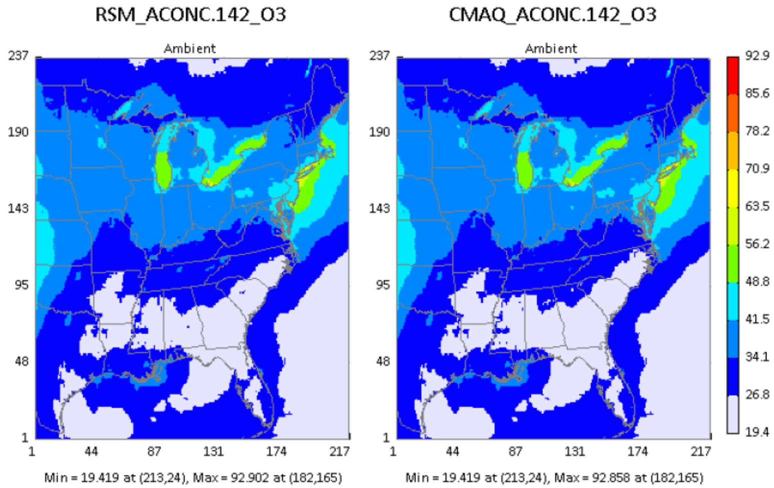
UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

33

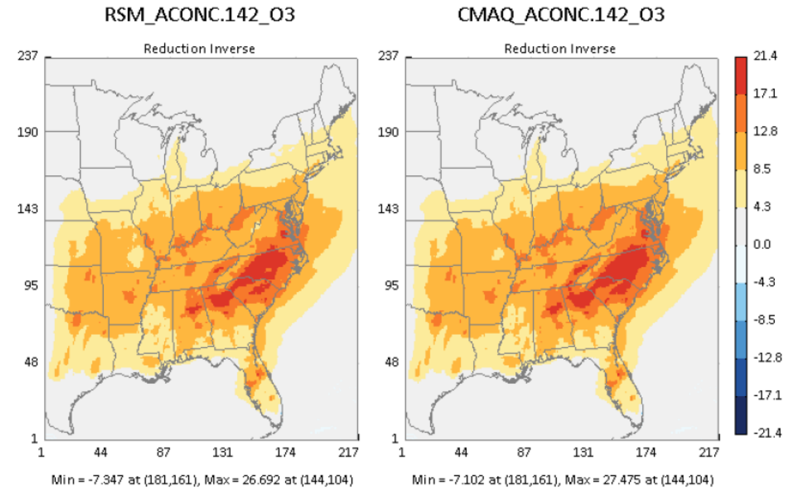
July, O_3-8hr avg. , US

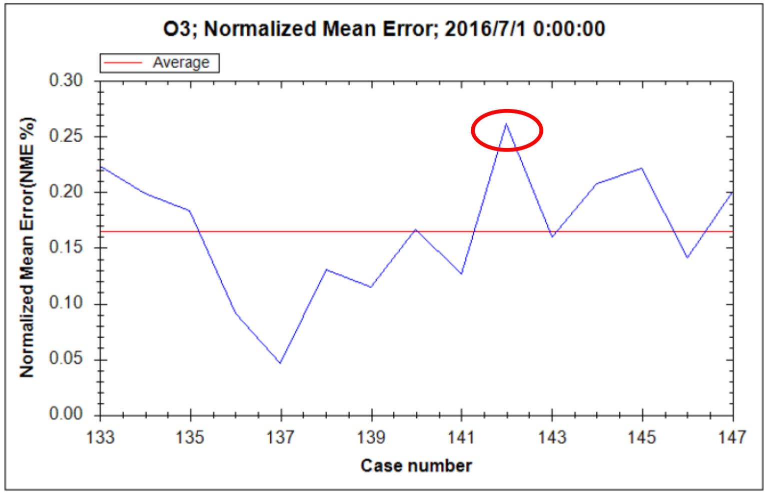
Out-of-Sample Validation: Predicted O_3 Concentrations & Responses (July, O_3-8hr avg. , US) Worst 1

O_3 Conc: RSM vs. CMAQ

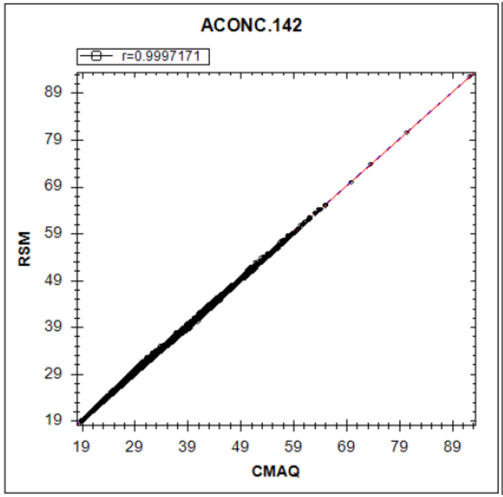


O_3 Response: RSM vs. CMAQ

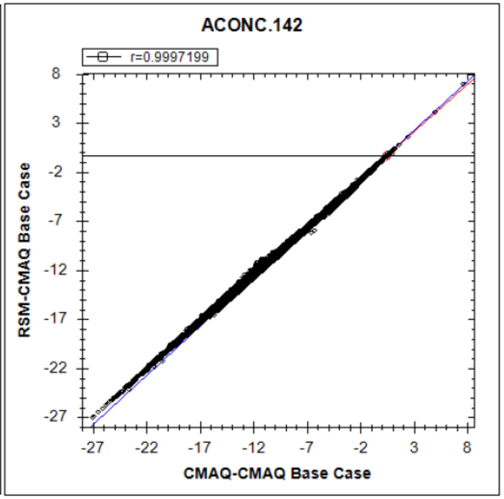




All "Out-of-Sample" runs: Normalized Mean Error



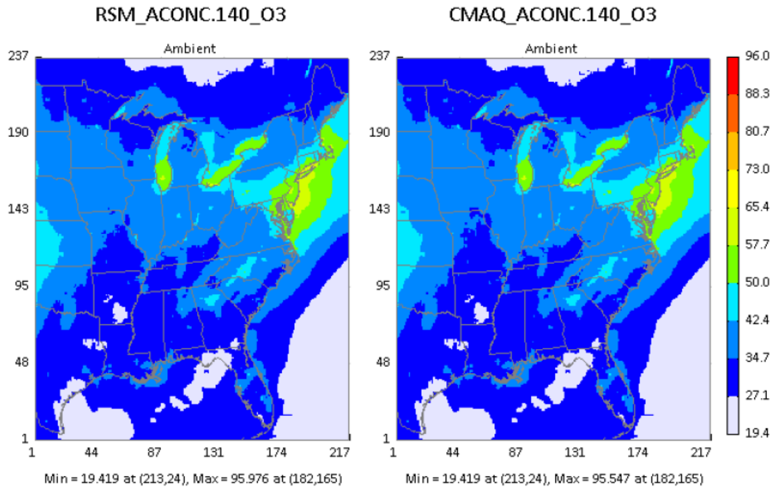
Predicted O_3 Concentration (RSM vs. CMAQ)



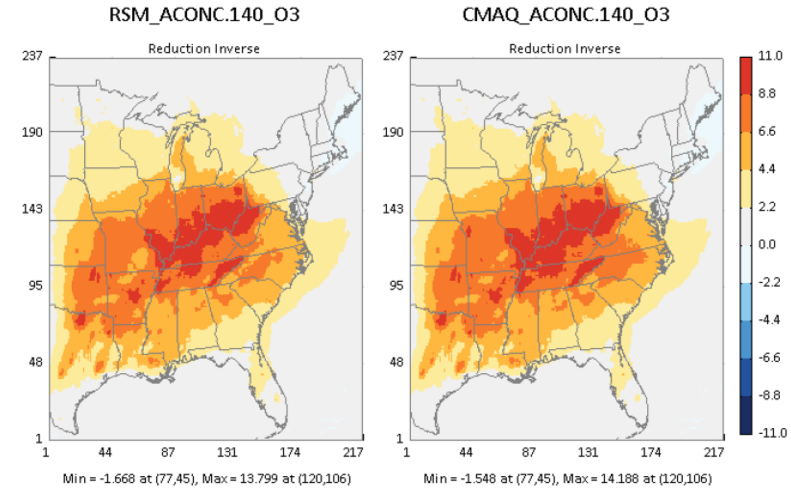
Predicted O_3 Response (RSM vs. CMAQ)

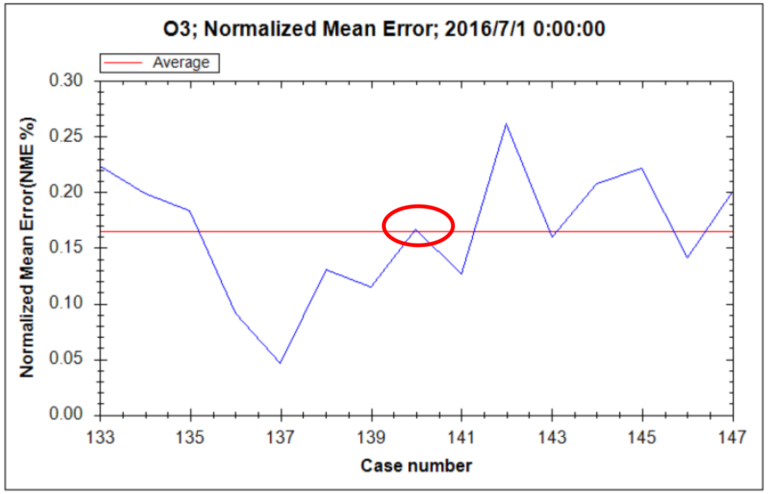
Out-of-Sample Validation: Predicted O_3 Concentrations & Responses (July, O_3-8hr avg. , US) Typical 1

O_3 Conc: RSM vs. CMAQ

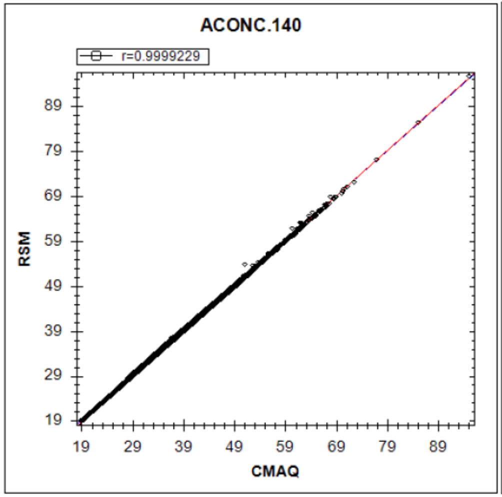


O_3 Response: RSM vs. CMAQ

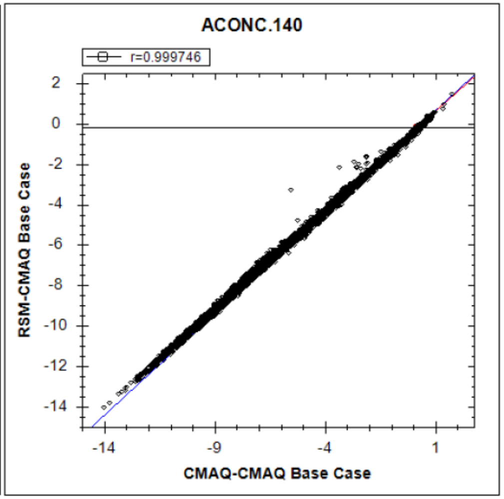




All "Out-of-Sample" runs: Normalized Mean Error



Predicted O_3 Concentration (RSM vs. CMAQ)

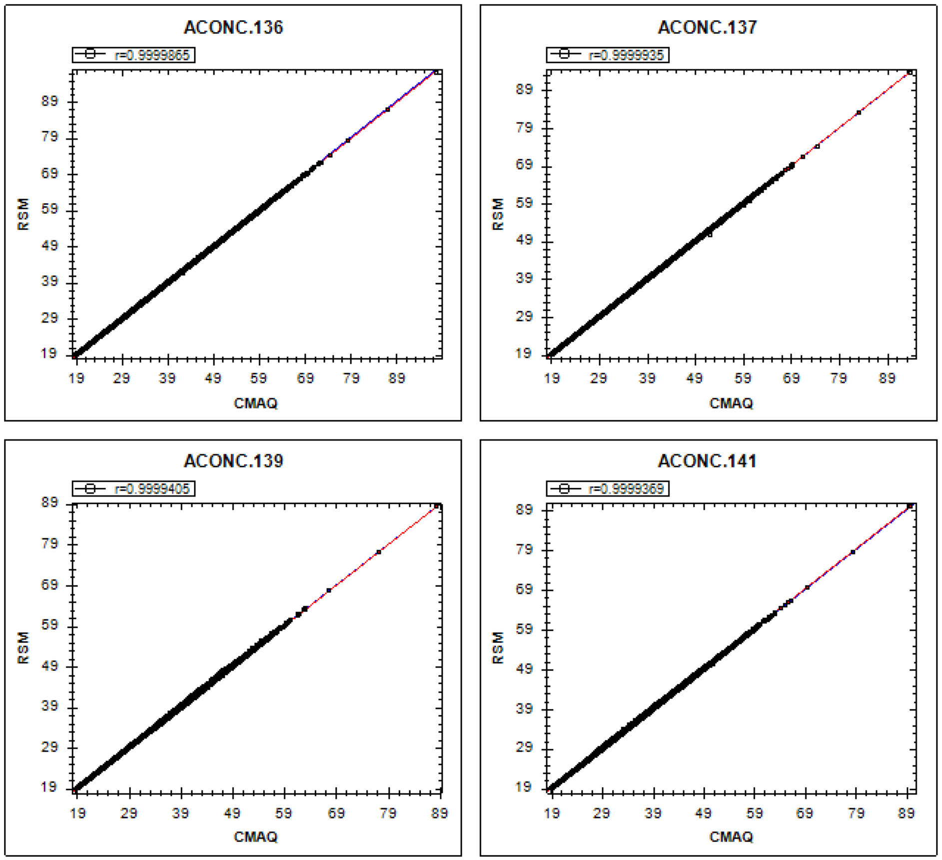


Predicted O_3 Response (RSM vs. CMAQ)

Out-of-Sample Validation: Concentration Comparison (July, O_3-8hr avg. , US) Best 4

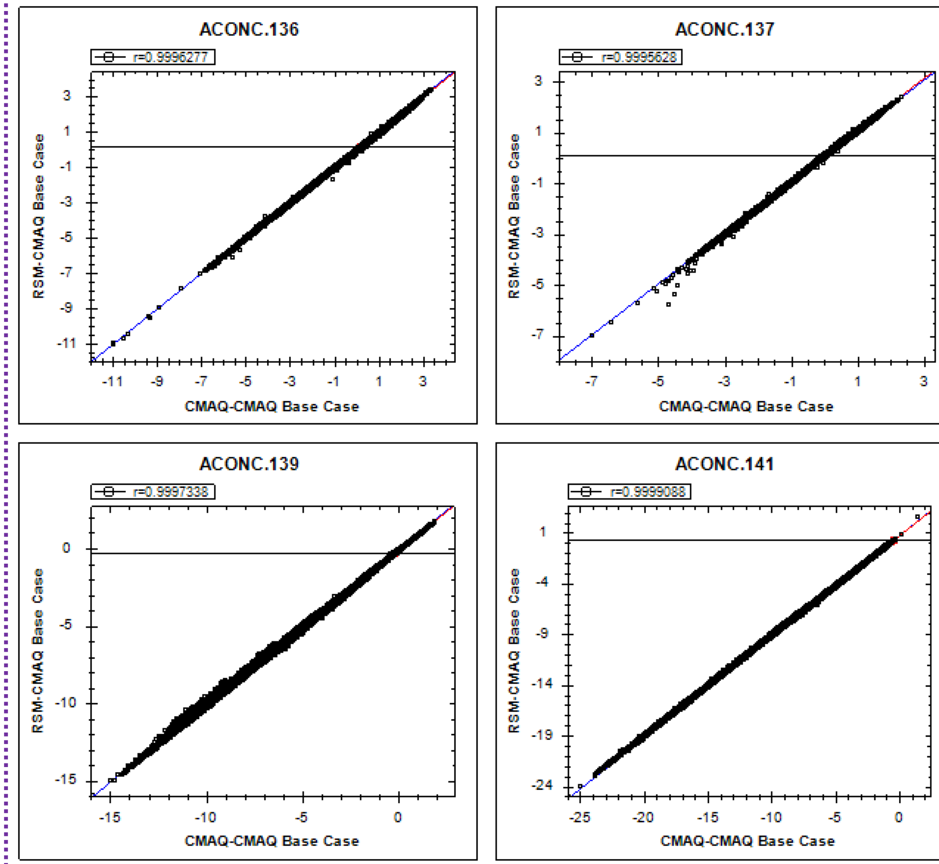
Predicted O_3 Conc.

(RSM vs. CMAQ)



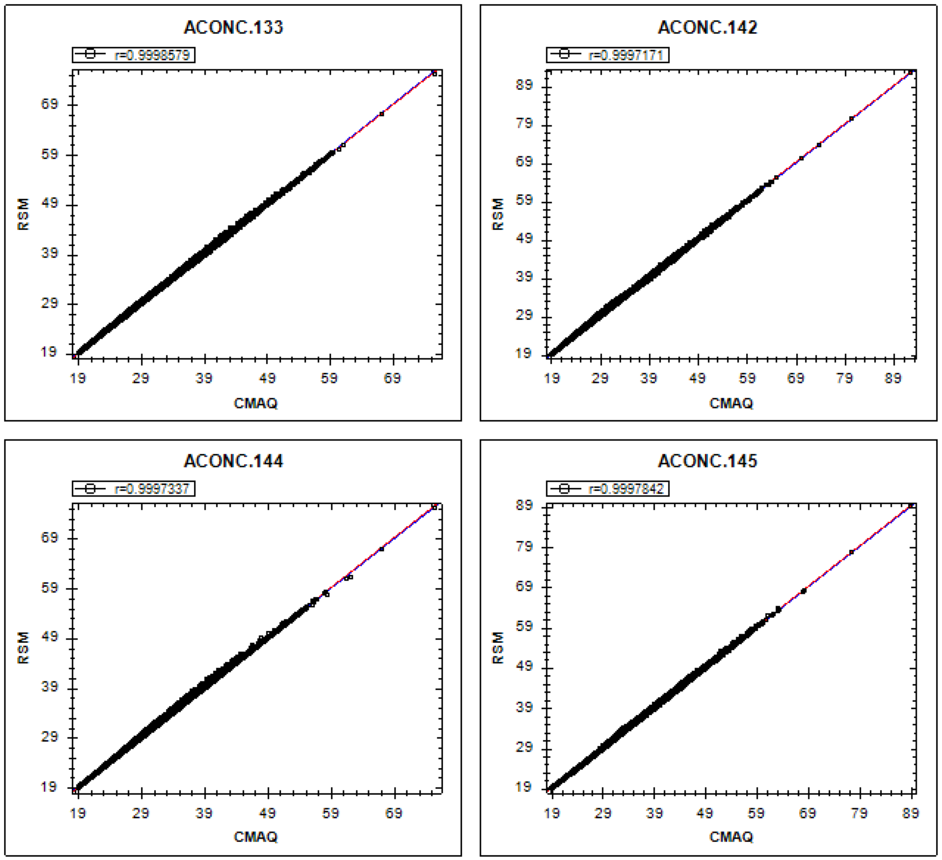
Predicted O_3 Responses (O_3)

(RSM-Base vs. CMAQ-Base)

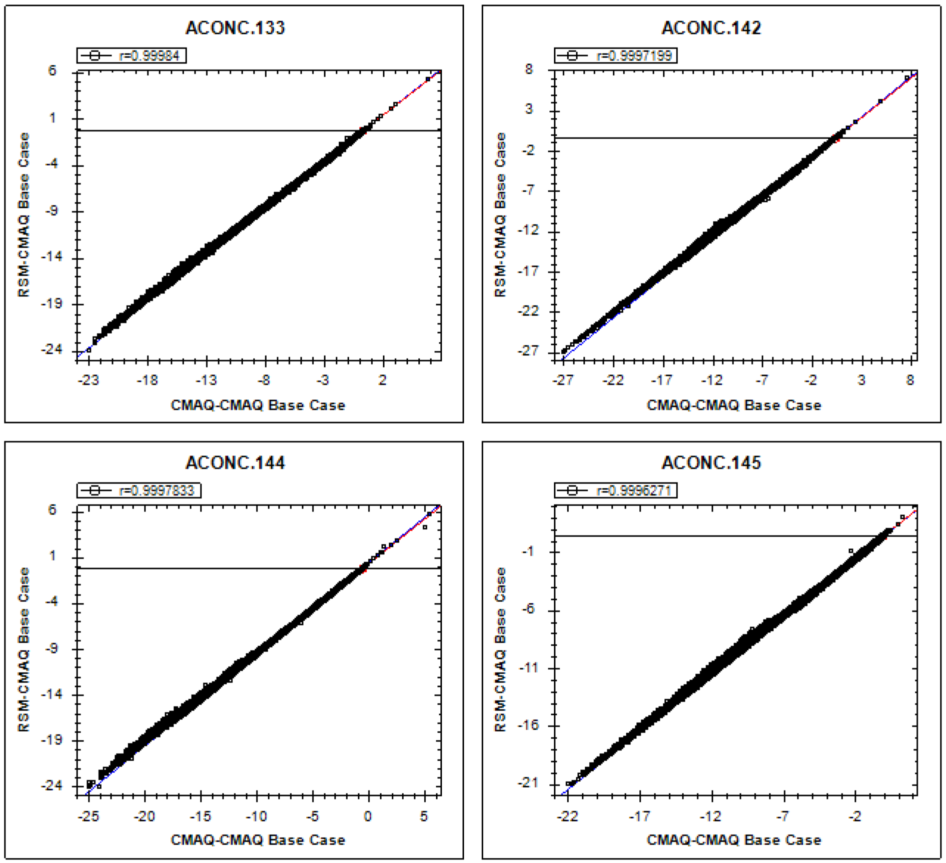


Out-of-Sample Validation: Concentration Comparison (July, O_3-8hr avg. , US) Worst 4

Predicted O_3 Conc.
(RSM vs. CMAQ)

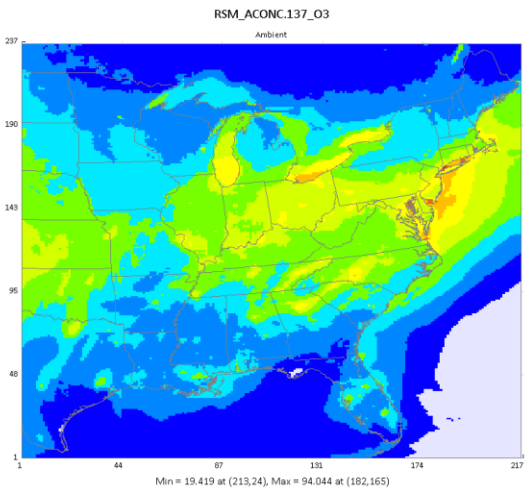


Predicted O_3 Responses (O_3)
(RSM-Base vs. CMAQ-Base)

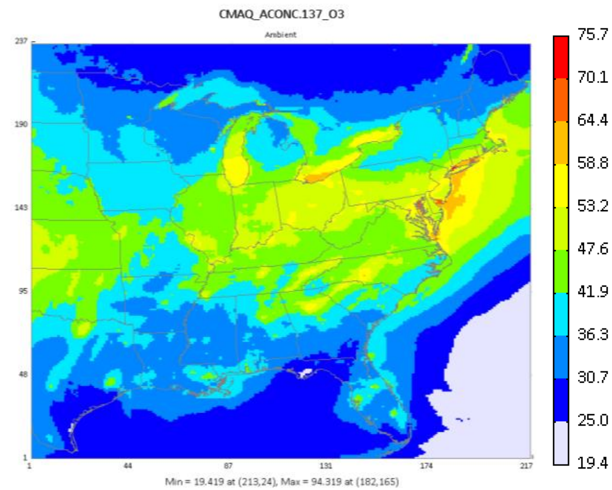


Out-of-Sample Validation: Predicted Concentrations (July, O_3-8hr avg. , US) Best 1

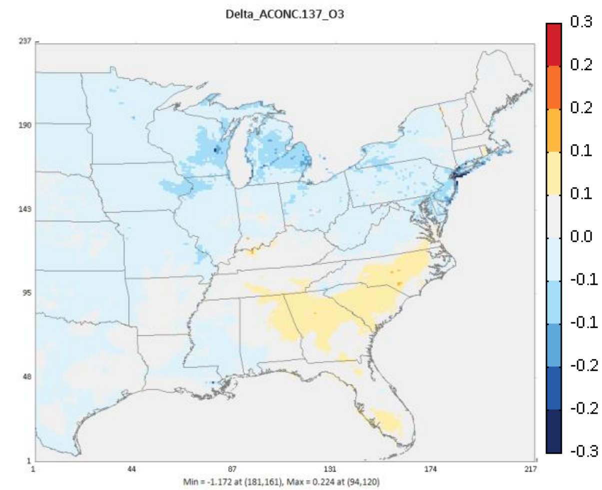
RSM



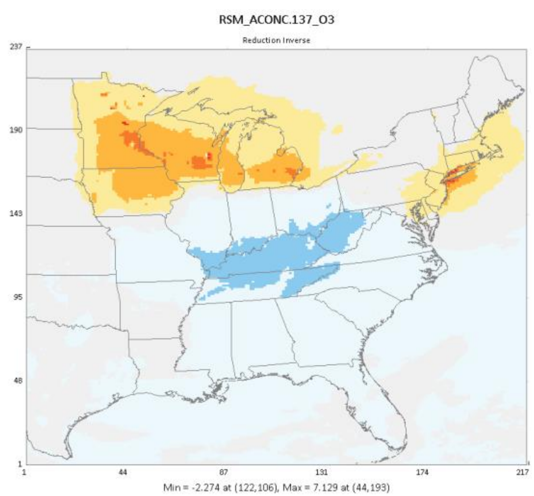
CMAQ



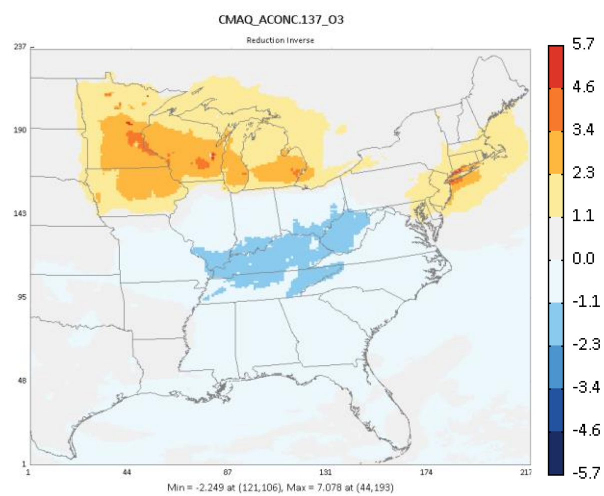
Delta



Base - RSM

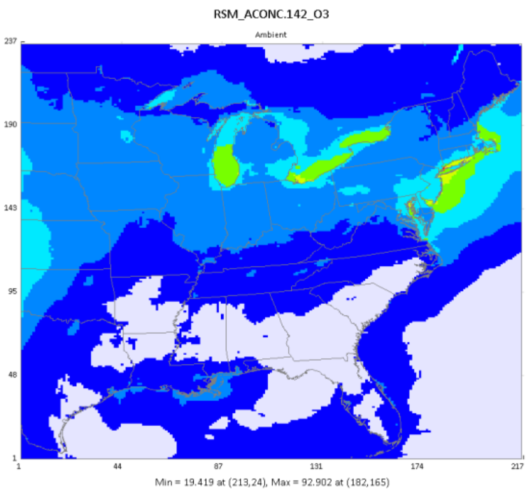


Base - CMAQ

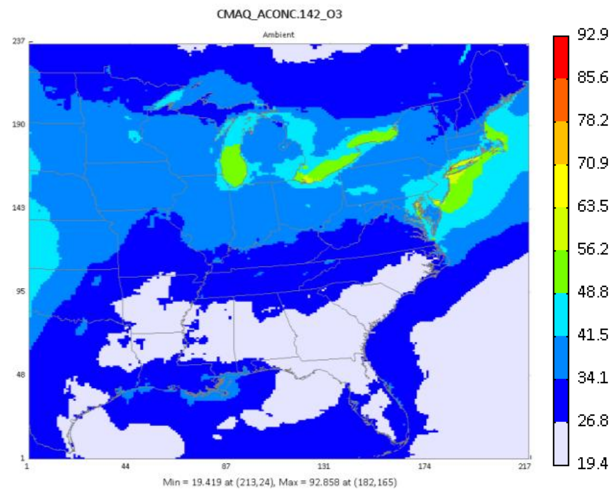


Out-of-Sample Validation: Predicted Concentrations (July, O_3-8hr avg. , US) Worst 1

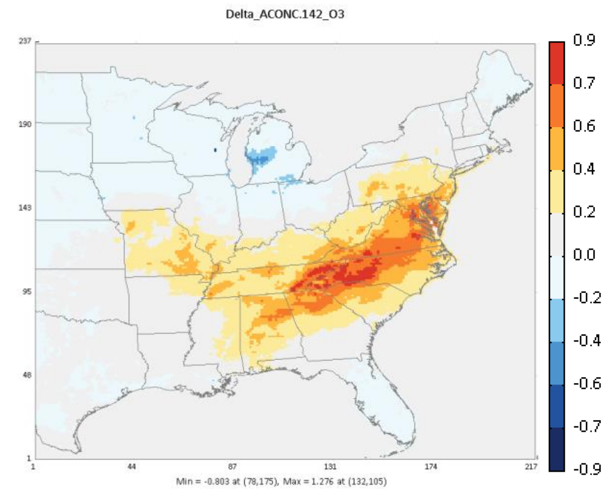
RSM



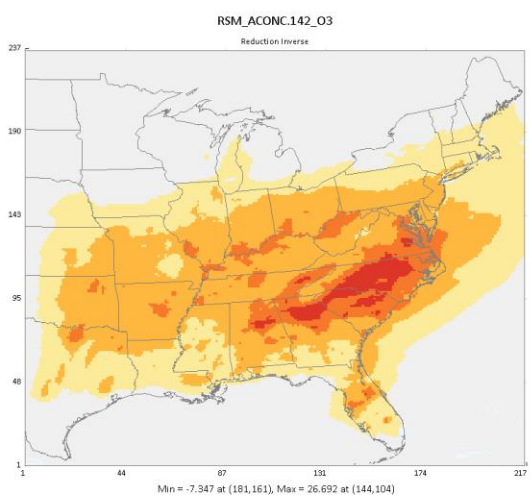
CMAQ



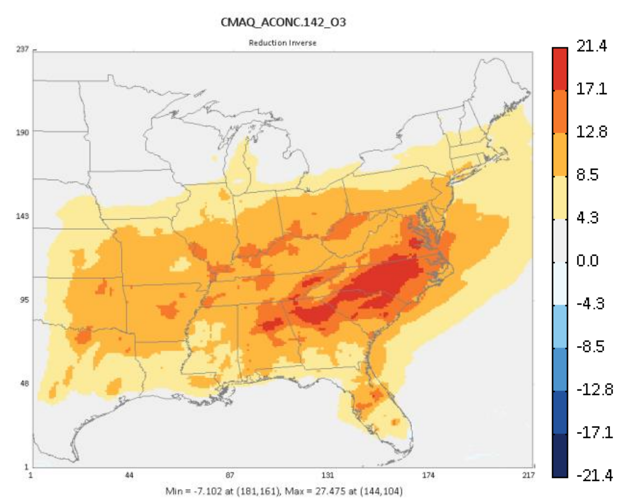
Delta



Base - RSM



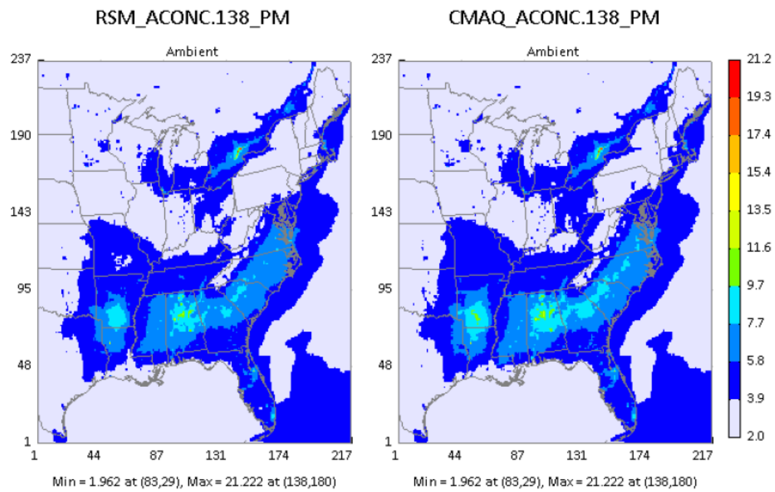
Base - CMAQ



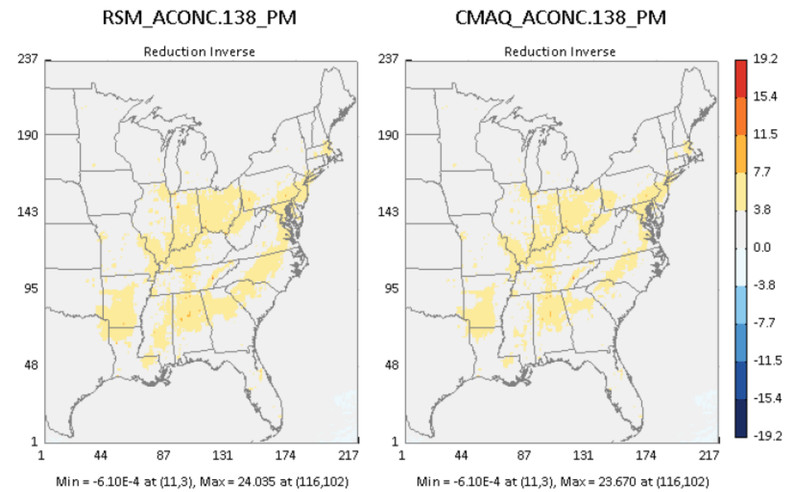
July, $PM_{2.5}$ monthly avg., US

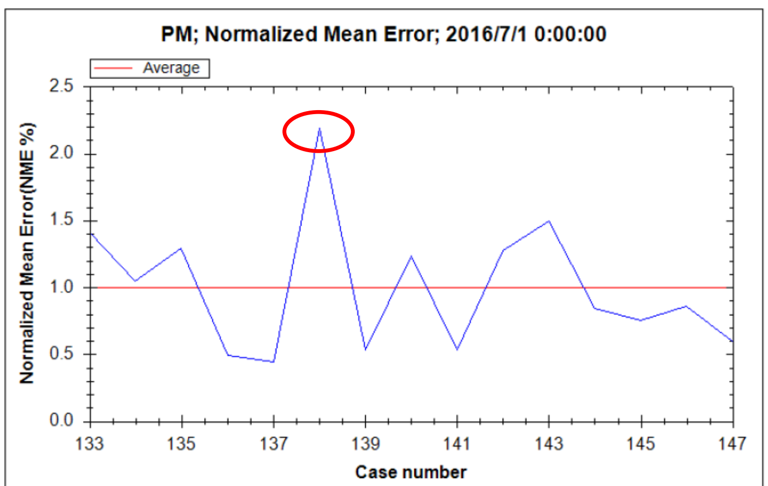
Out-of-Sample Validation: Predicted $PM_{2.5}$ Concentrations & Responses (July, $PM_{2.5}$ monthly avg., US) Worst 1

$PM_{2.5}$ Conc: RSM vs. CMAQ

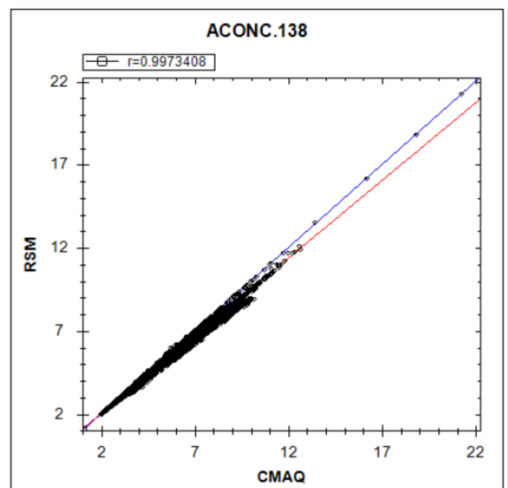


$PM_{2.5}$ Response: RSM vs. CMAQ

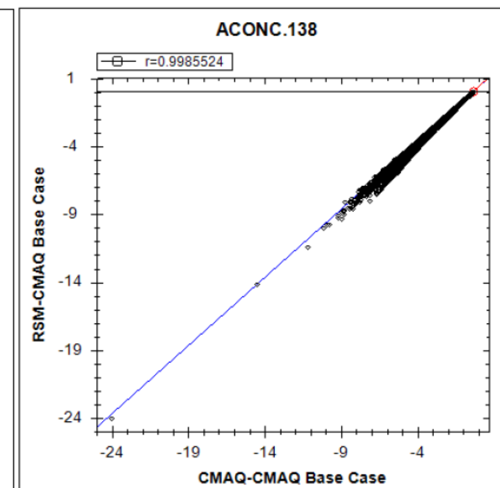




All "Out-of-Sample" runs: Normalized Mean Error



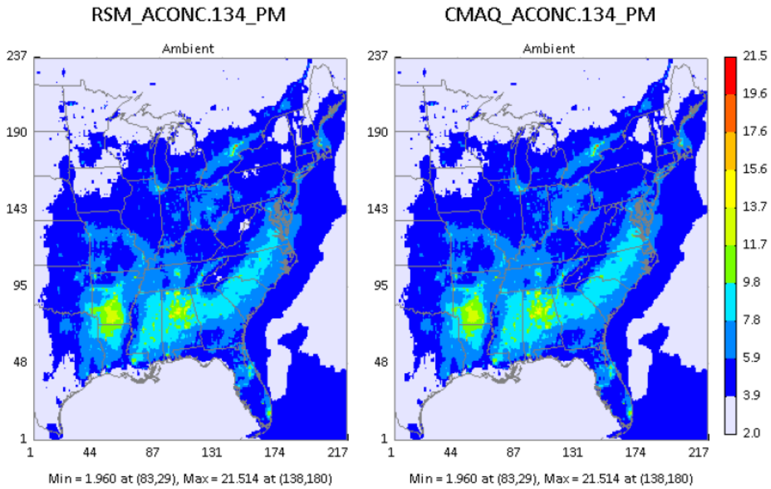
Predicted $PM_{2.5}$ Concentration (RSM vs. CMAQ)



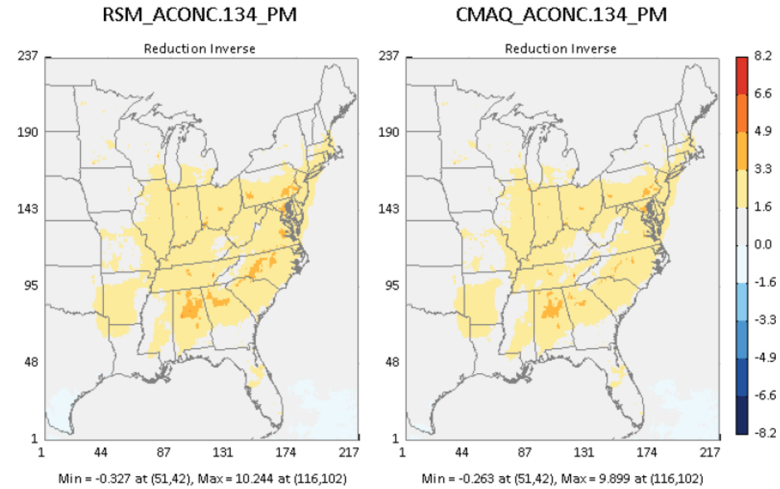
Predicted $PM_{.5}$ Response (RSM vs. CMAQ)

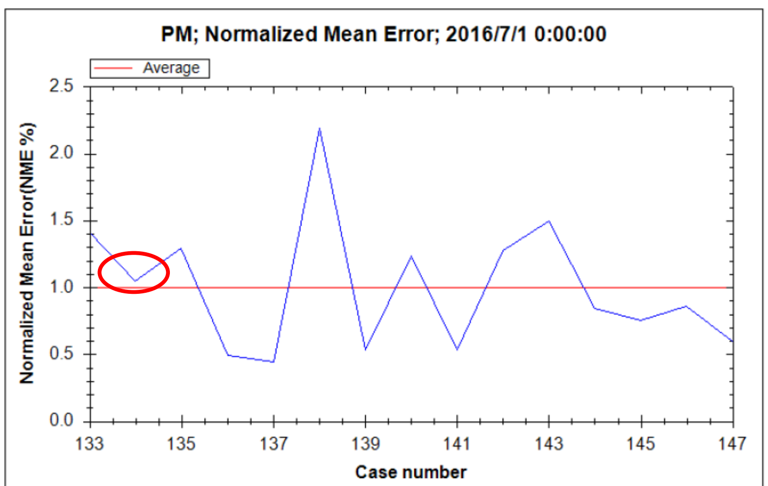
Out-of-Sample Validation: Predicted $PM_{2.5}$ Concentrations & Responses (July, $PM_{2.5}$ monthly avg., US) Typical 1

$PM_{2.5}$ Conc: RSM vs. CMAQ

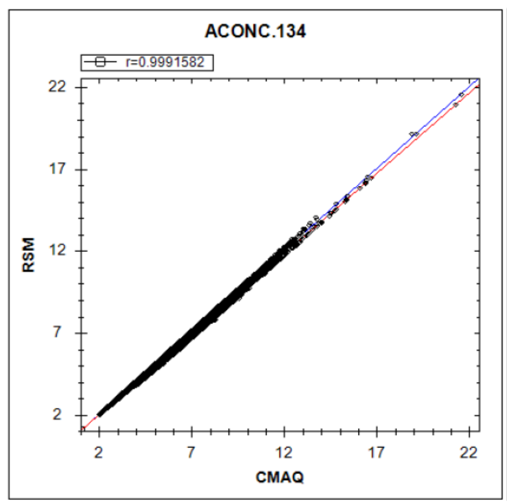


$PM_{2.5}$ Response: RSM vs. CMAQ

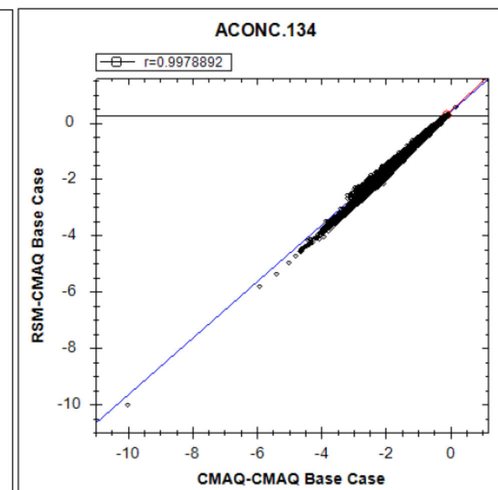




All "Out-of-Sample" runs: Normalized Mean Error



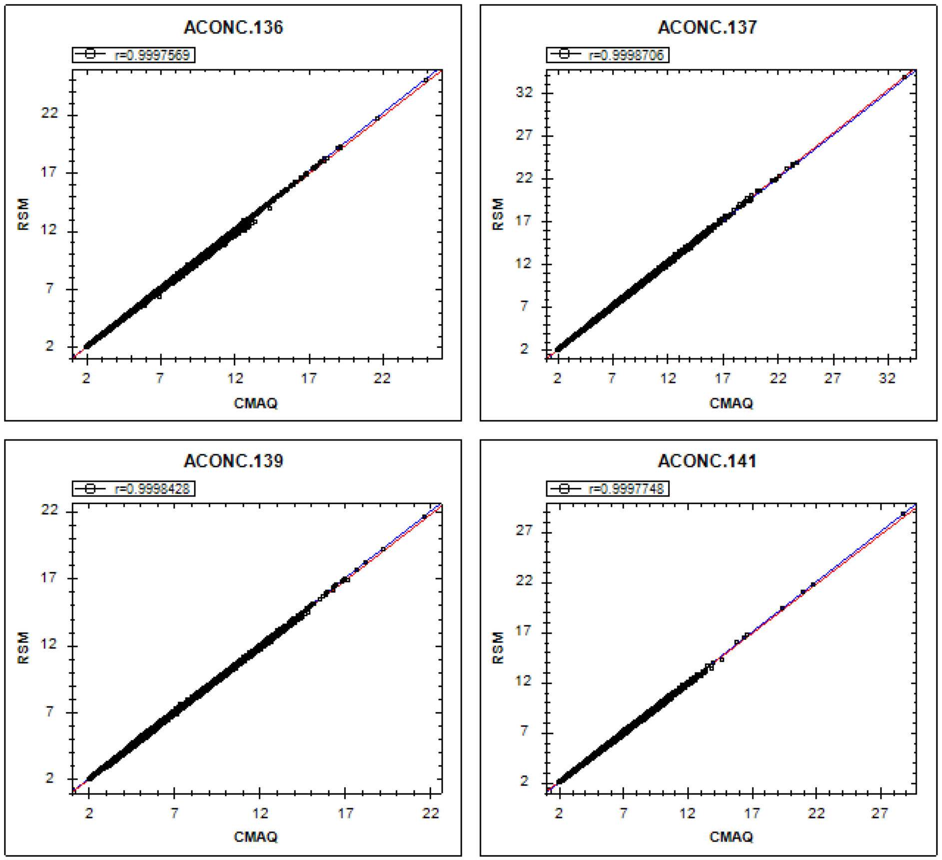
Predicted $PM_{2.5}$ Concentration (RSM vs. CMAQ)



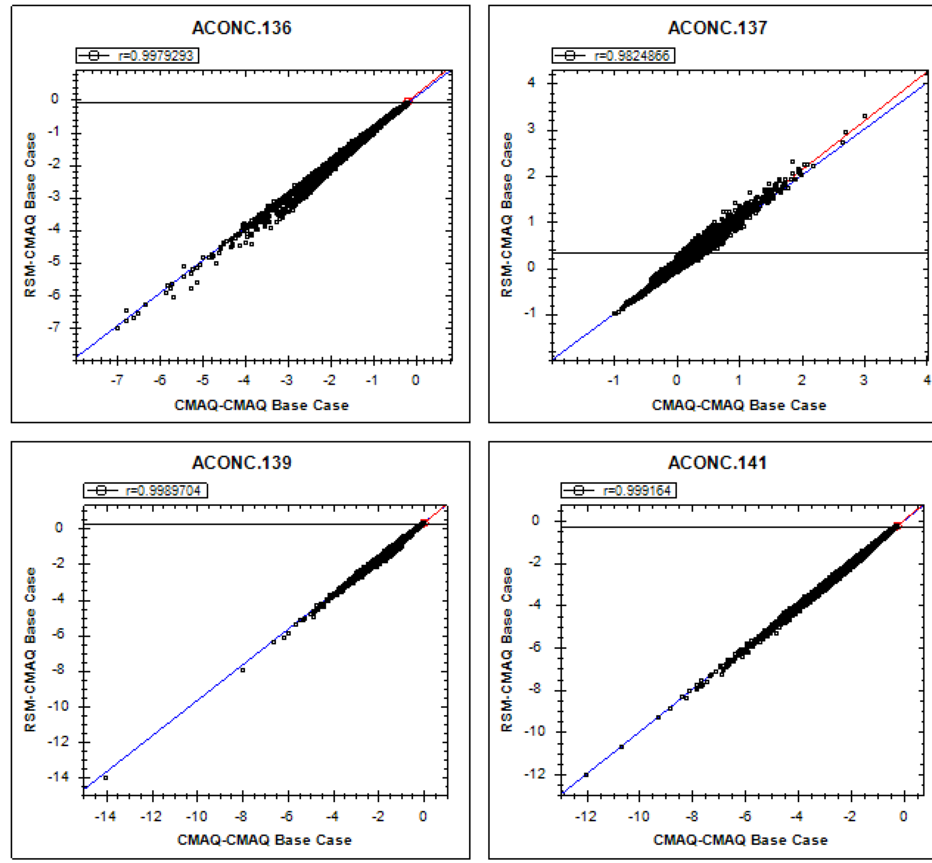
Predicted $PM_{.5}$ Response (RSM vs. CMAQ)

Out-of-Sample Validation: Concentration Comparison (July, $PM_{2.5}$ monthly avg., US) Best 4

Predicted $PM_{2.5}$ Conc.
(RSM vs. CMAQ)

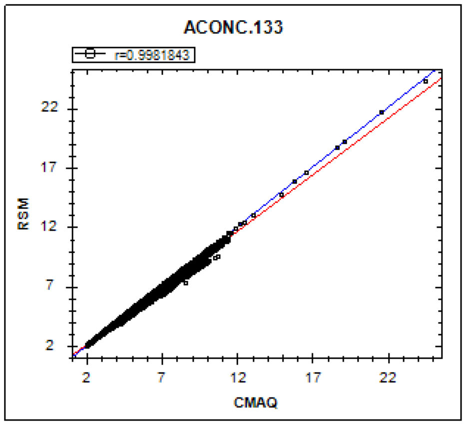


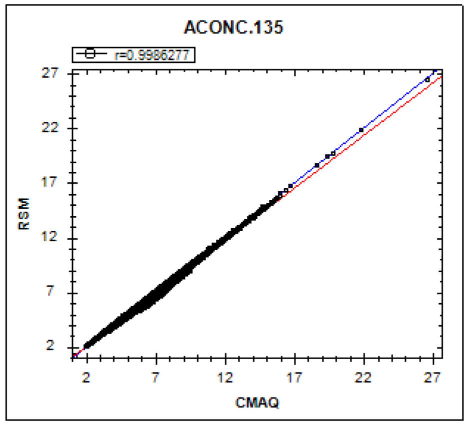
Predicted $PM_{2.5}$ Responses
(RSM-Base vs. CMAQ-Base)

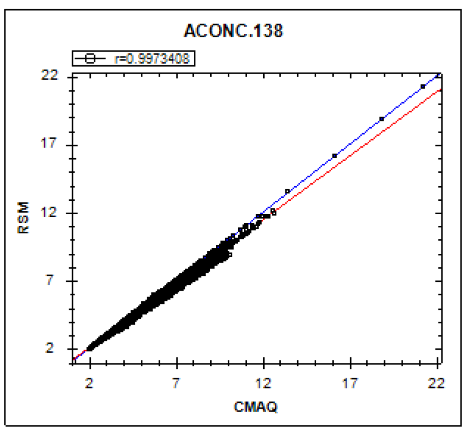


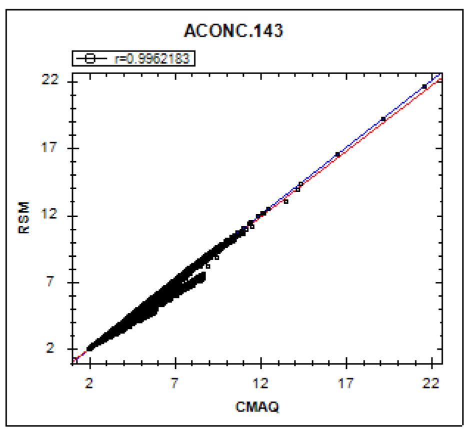
Out-of-Sample Validation: Concentration Comparison (July, $PM_{2.5}$ monthly avg., US) Worst 4

Predicted $PM_{2.5}$ Conc.
(RSM vs. CMAQ)

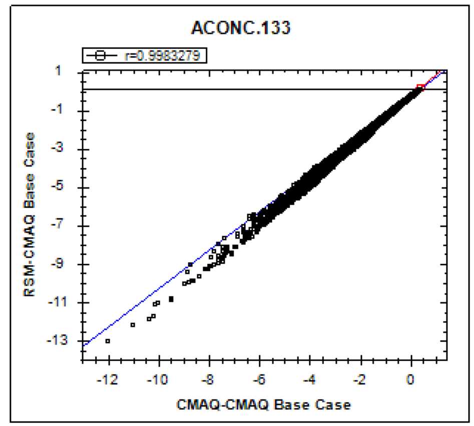


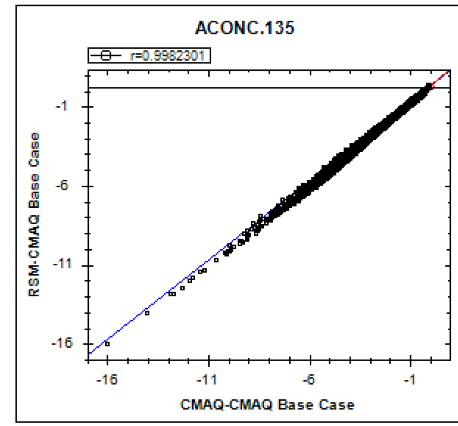


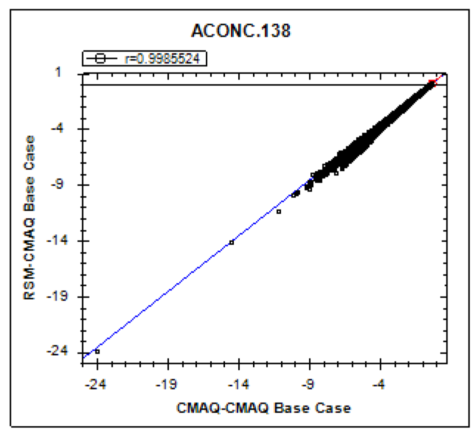


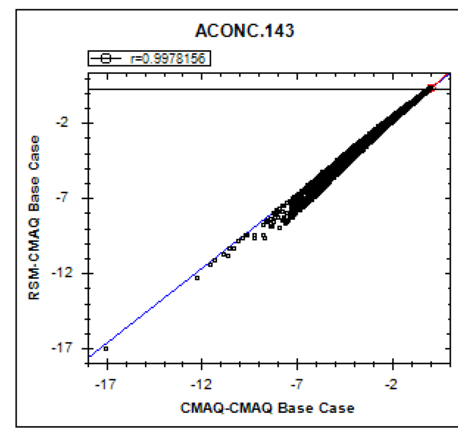


Predicted $PM_{2.5}$ Responses
(RSM-Base vs. CMAQ-Base)



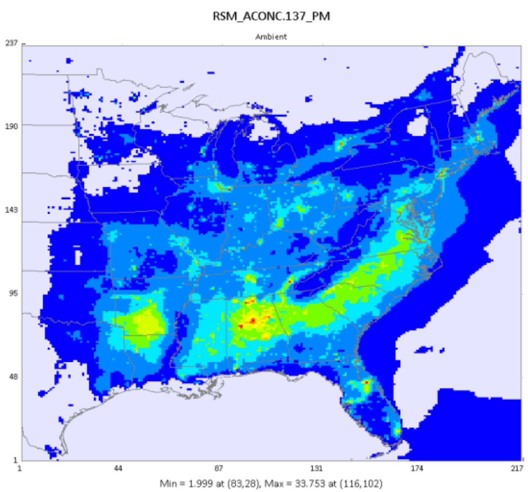




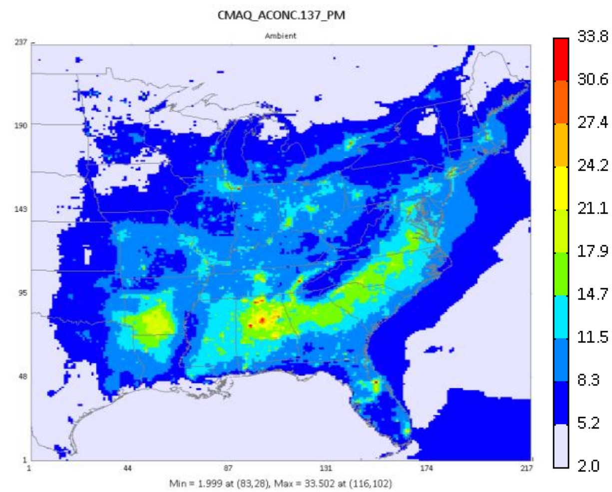


Out-of-Sample Validation: Predicted Concentrations (July, $PM_{2.5}$ monthly avg., US) Best 1

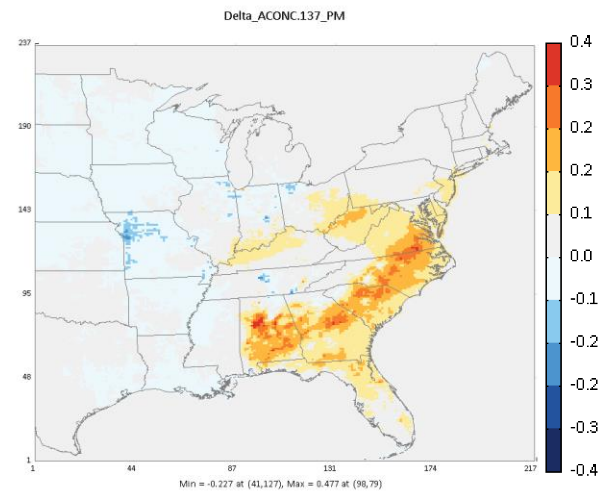
RSM



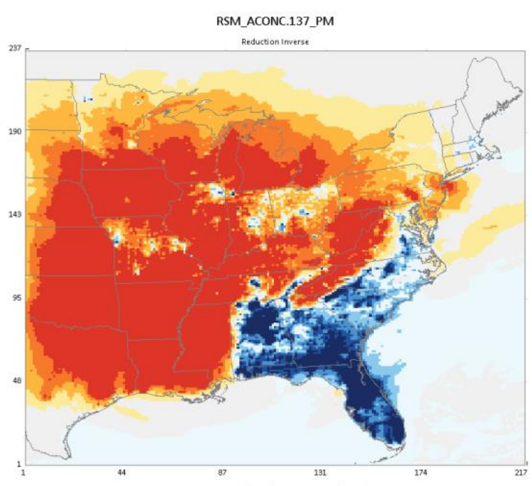
CMAQ



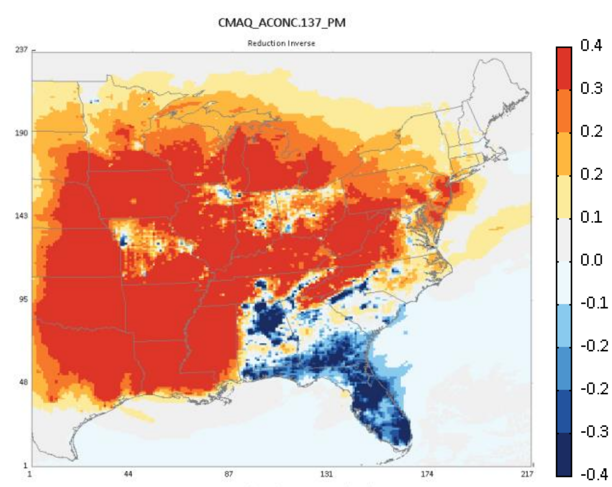
Delta



Base - RSM

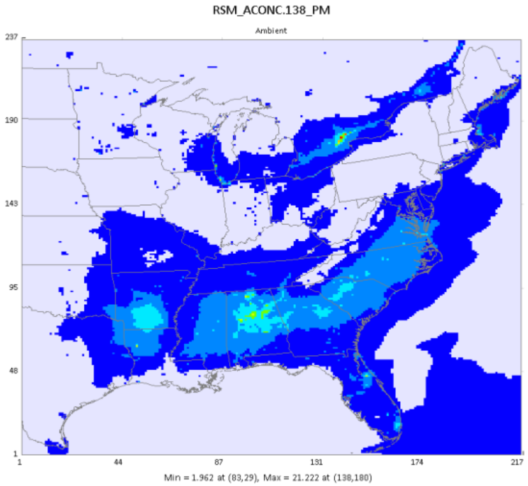


Base - CMAQ

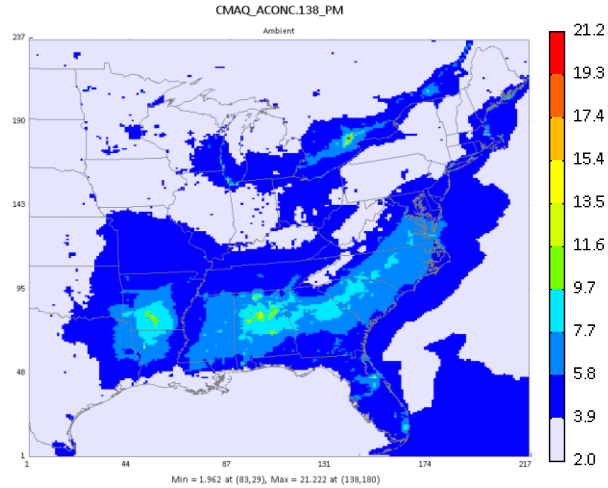


Out-of-Sample Validation: Predicted Concentrations (July, $PM_{2.5}$ monthly avg., US) Worst 1

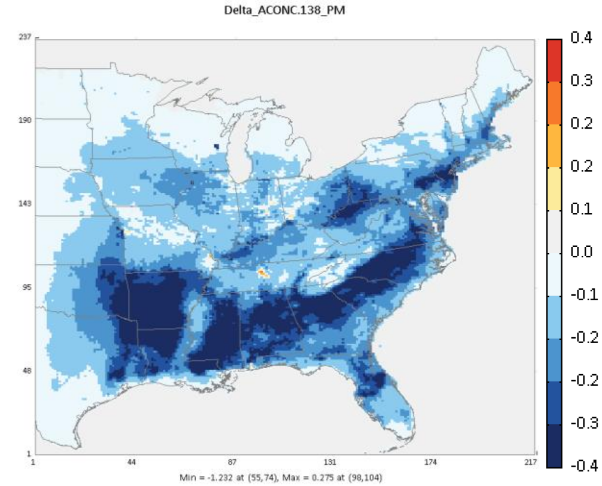
RSM



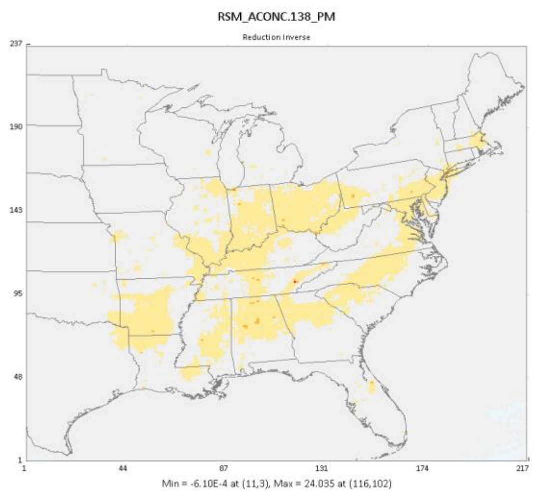
CMAQ



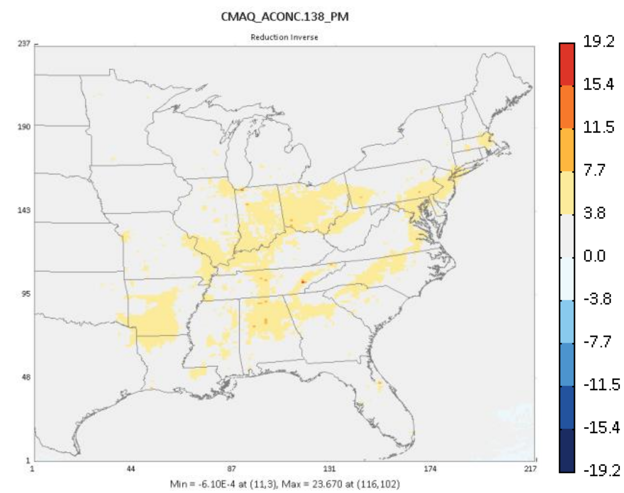
Delta



Base - RSM



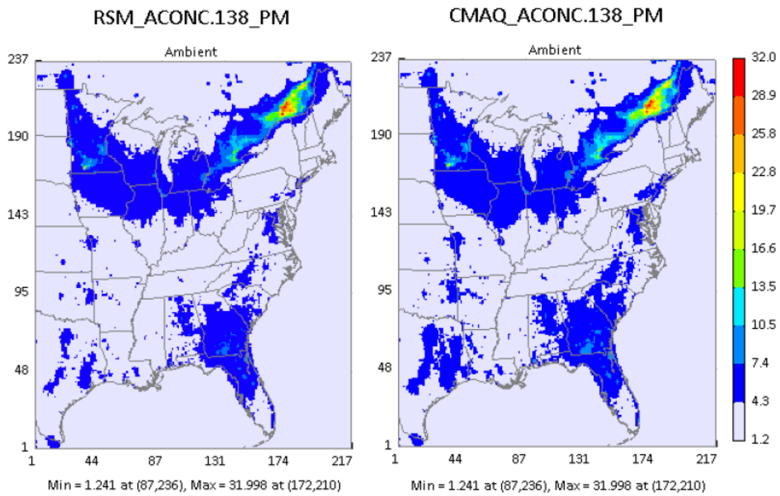
Base - CMAQ



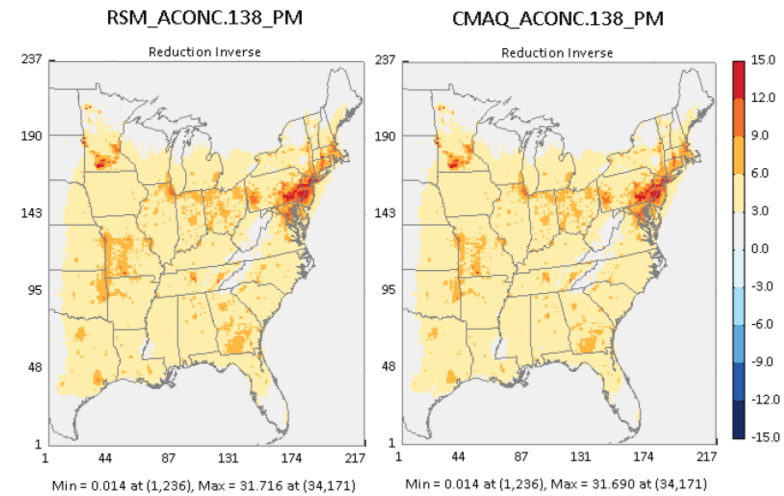
Jan., $PM_{2.5}$ monthly avg., US

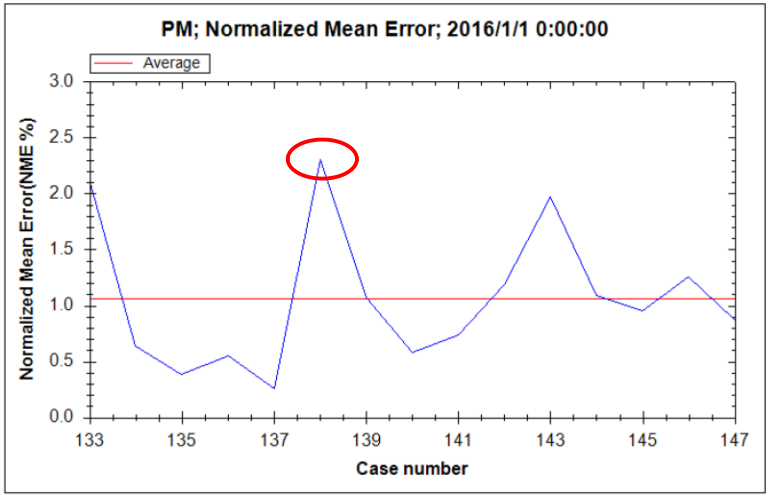
Out-of-Sample Validation: Predicted $PM_{2.5}$ Concentrations & Responses (Jan., $PM_{2.5}$ monthly avg., US) Worst 1

$PM_{2.5}$ Conc: RSM vs. CMAQ

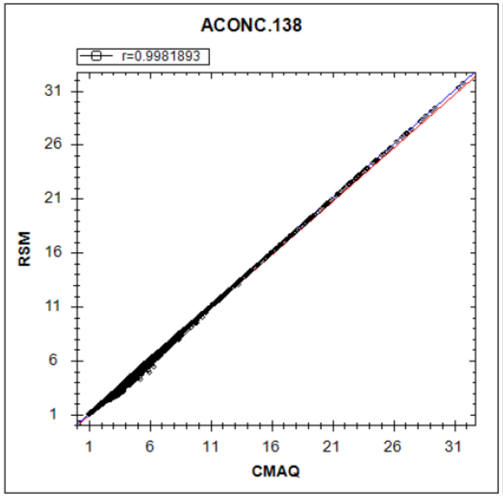


$PM_{2.5}$ Response: RSM vs. CMAQ

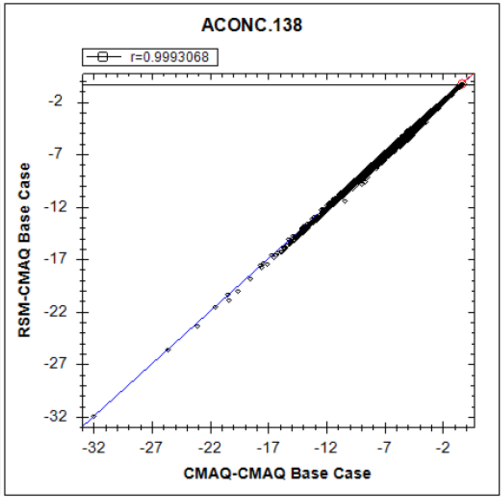




All "Out-of-Sample" runs:
Normalized Mean Error



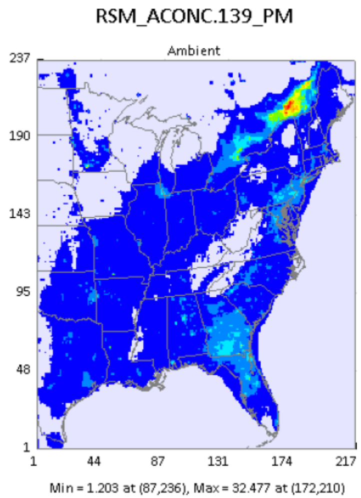
Predicted $PM_{2.5}$ Concentration
(RSM vs. CMAQ)

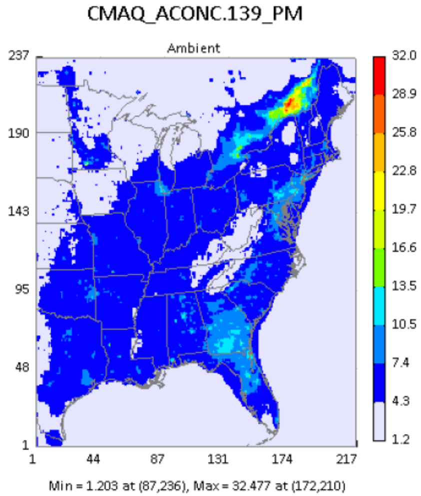


Predicted $PM_{.5}$ Response
(RSM vs. CMAQ)

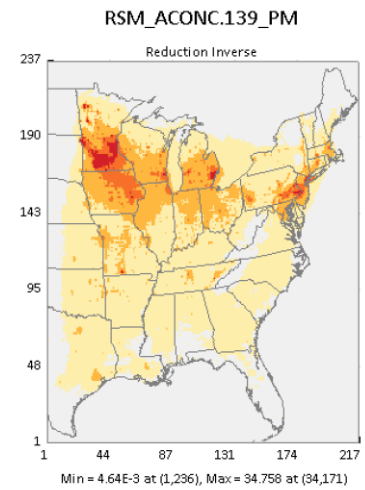
Out-of-Sample Validation: Predicted $PM_{2.5}$ Concentrations & Responses (Jan., $PM_{2.5}$ monthly avg., US) Typical 1

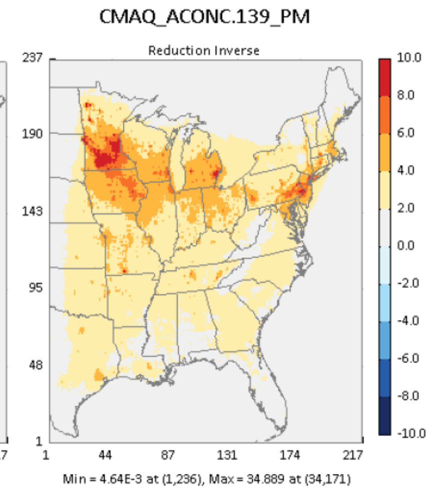
$PM_{2.5}$ Conc: RSM vs. CMAQ

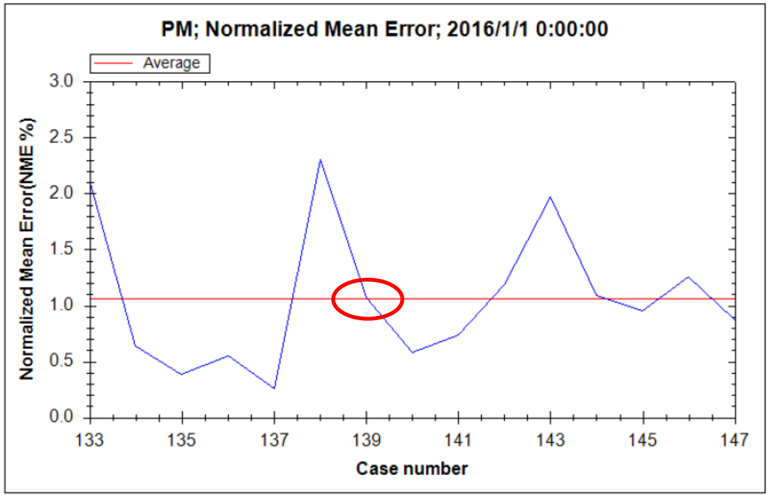




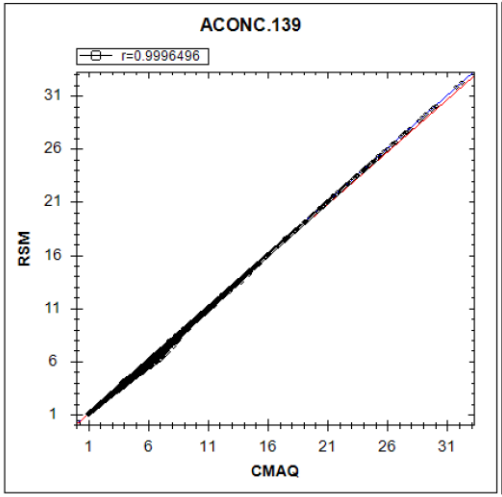
$PM_{2.5}$ Response: RSM vs. CMAQ



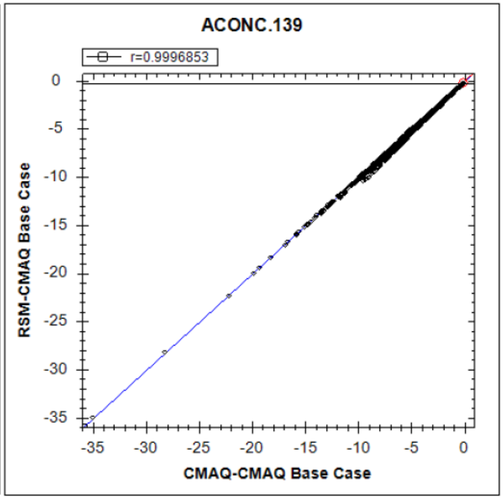




All "Out-of-Sample" runs: Normalized Mean Error



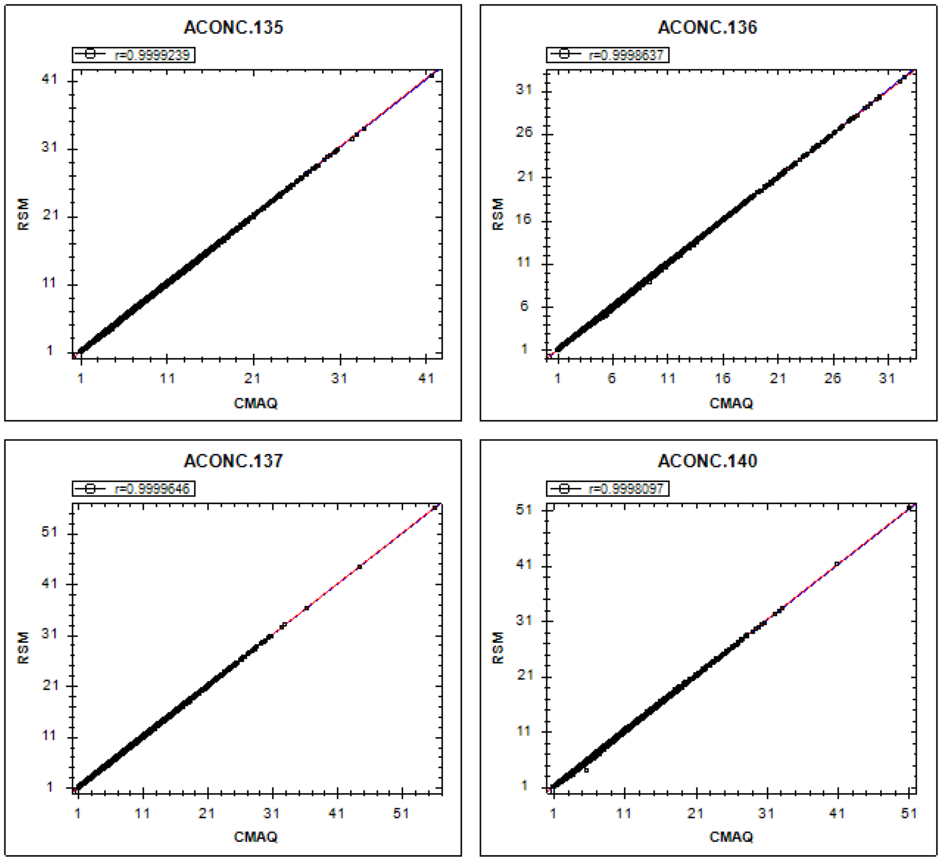
Predicted $PM_{2.5}$ Concentration (RSM vs. CMAQ)



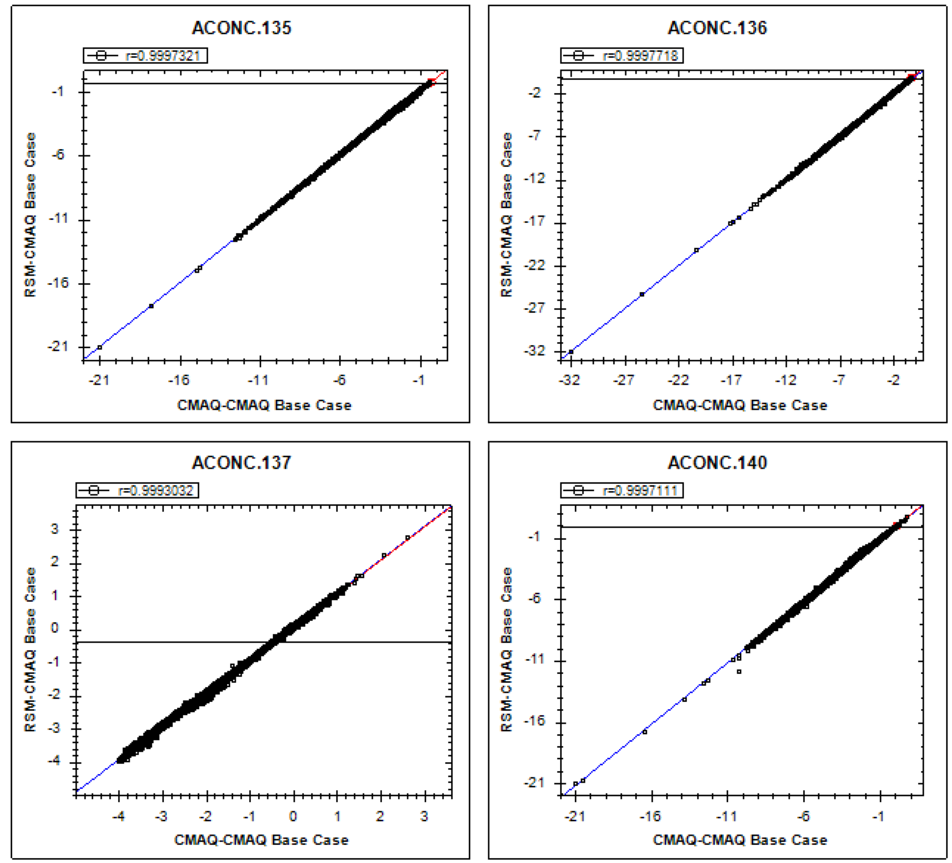
Predicted $PM_{.5}$ Response (RSM vs. CMAQ)

Out-of-Sample Validation: Concentration Comparison (Jan., $PM_{2.5}$ monthly avg., US) Best 4

Predicted $PM_{2.5}$ Conc. (RSM vs. CMAQ)

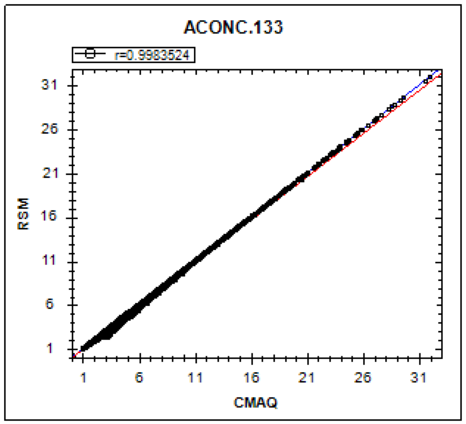


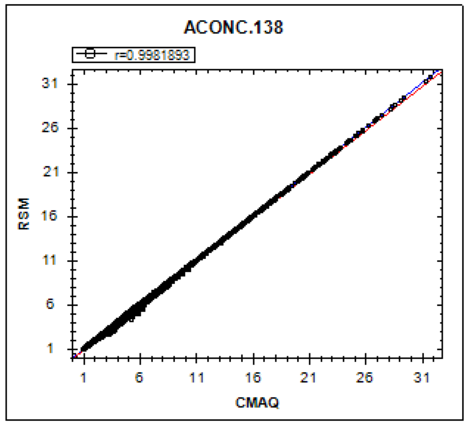
Predicted $PM_{2.5}$ Responses (RSM-Base vs. CMAQ-Base)

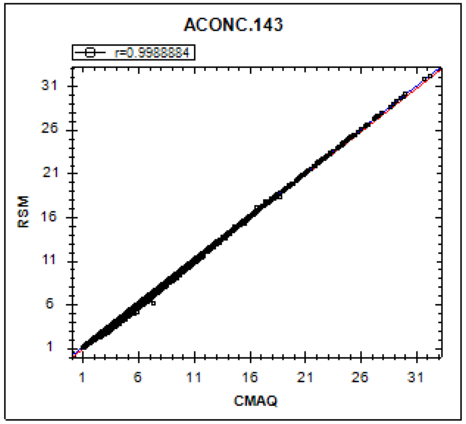


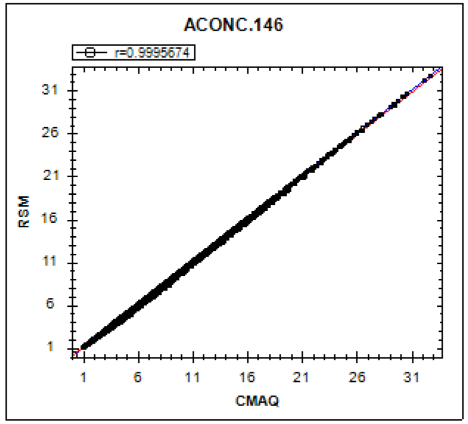
Out-of-Sample Validation: Concentration Comparison (Jan., $PM_{2.5}$ monthly avg., US) Worst 4

Predicted $PM_{2.5}$ Conc.
(RSM vs. CMAQ)

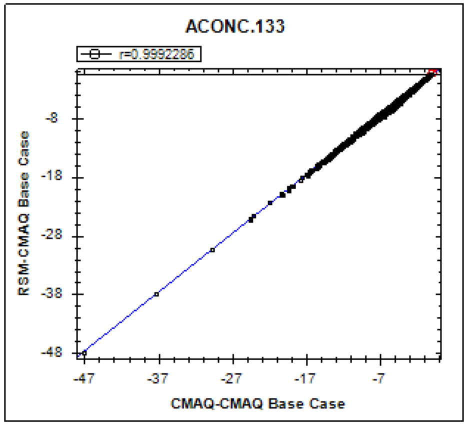


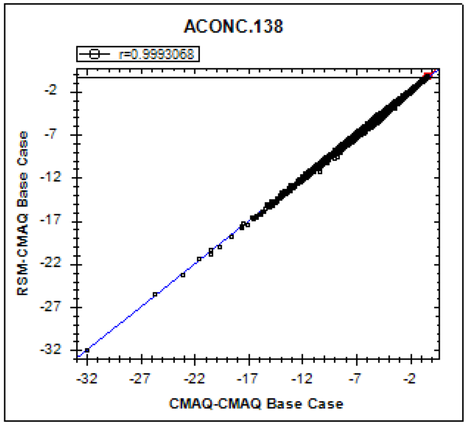


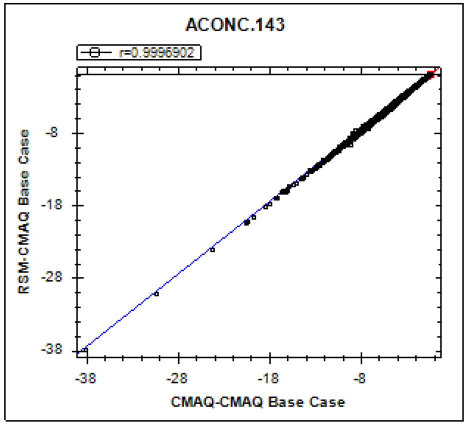


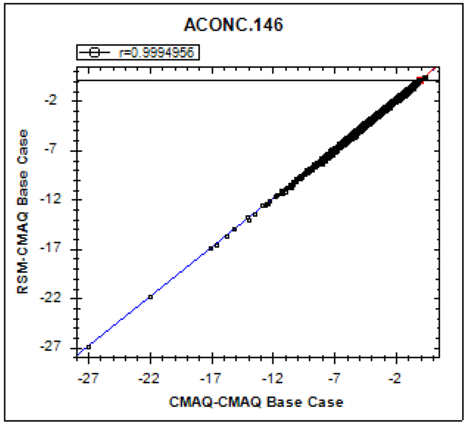


Predicted $PM_{2.5}$ Responses
(RSM-Base vs. CMAQ-Base)



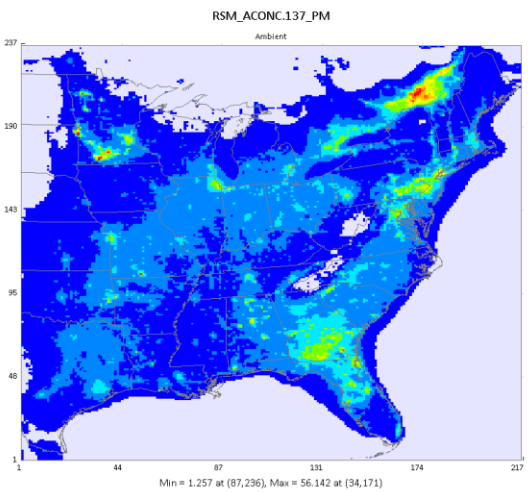




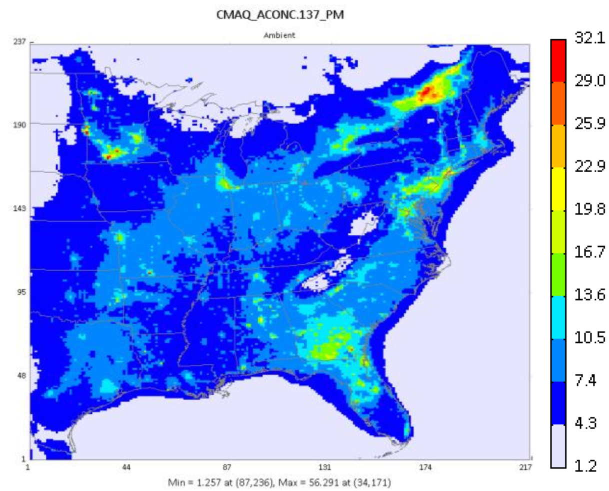


Out-of-Sample Validation: Predicted Concentrations (Jan., $PM_{2.5}$ monthly avg., US) Best 1

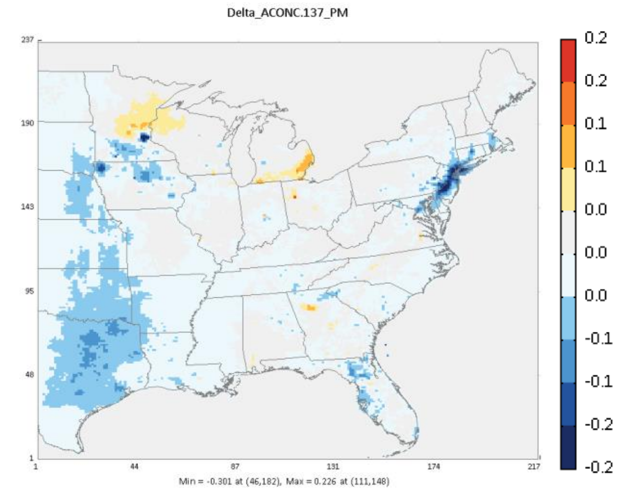
RSM



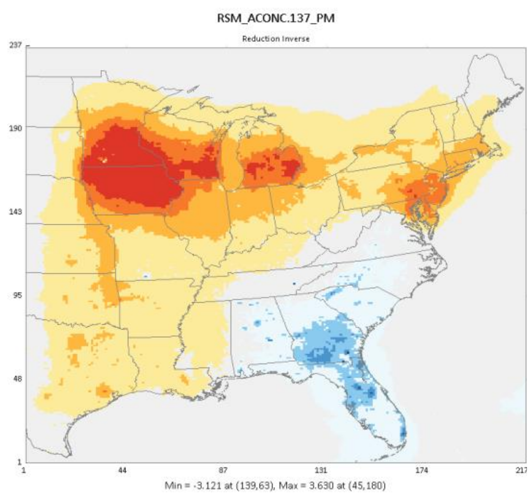
CMAQ



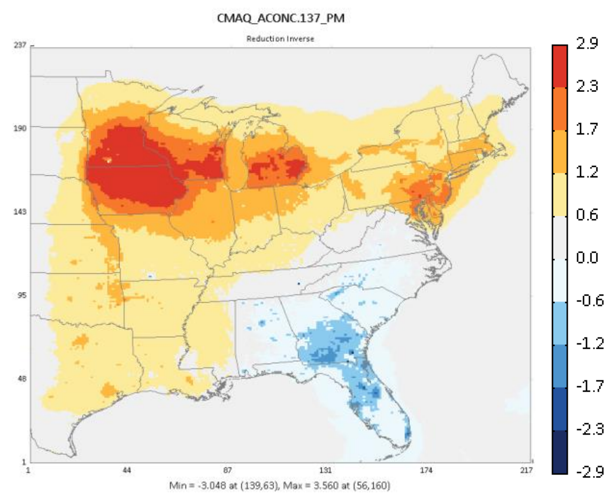
Delta



Base - RSM

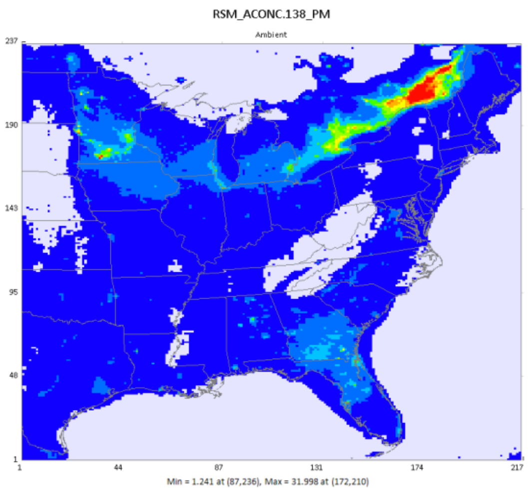


Base - CMAQ

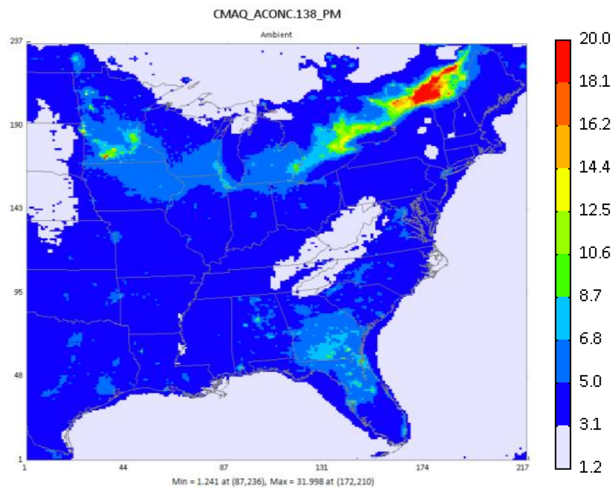


Out-of-Sample Validation: Predicted Concentrations (Jan., $PM_{2.5}$ monthly avg., US) Worst 1

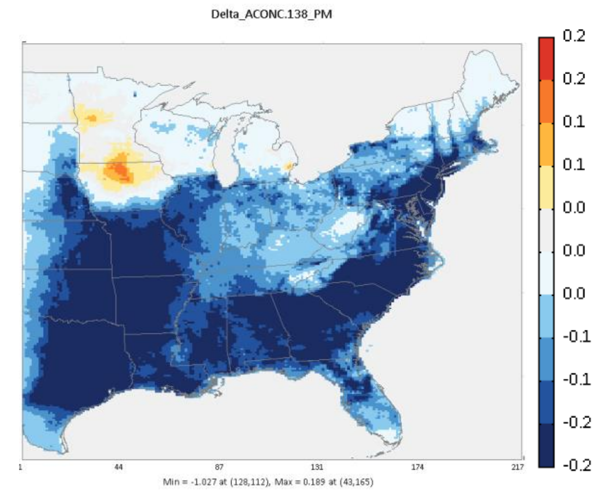
RSM



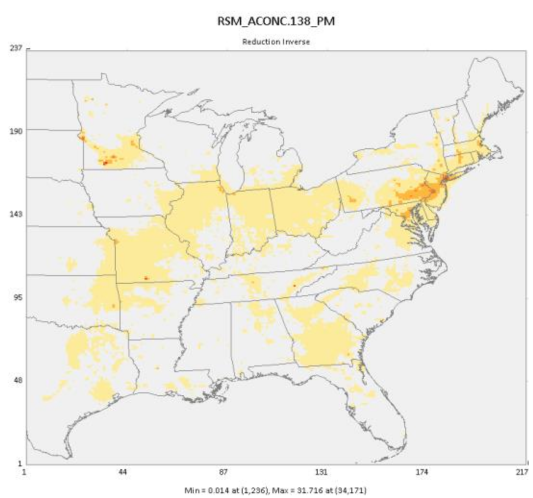
CMAQ



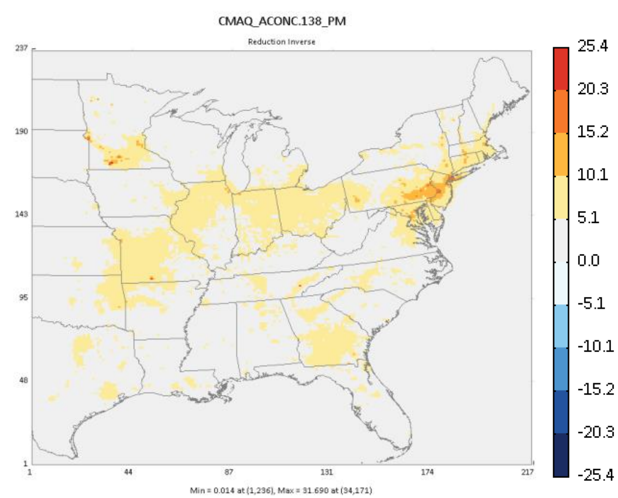
Delta



Base - RSM



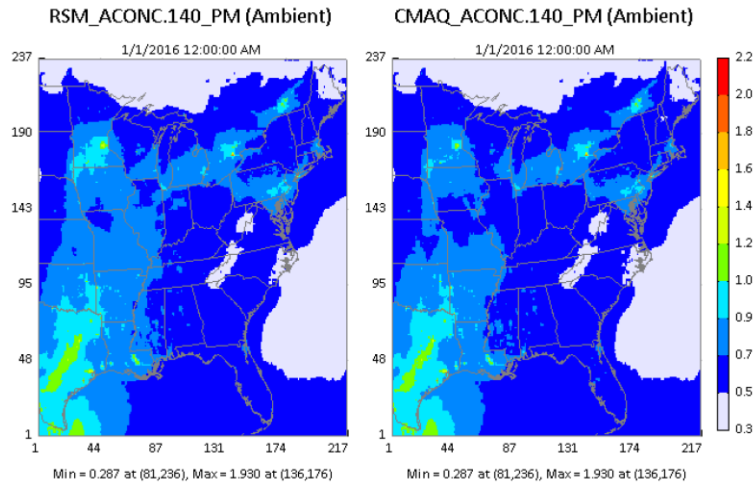
Base - CMAQ



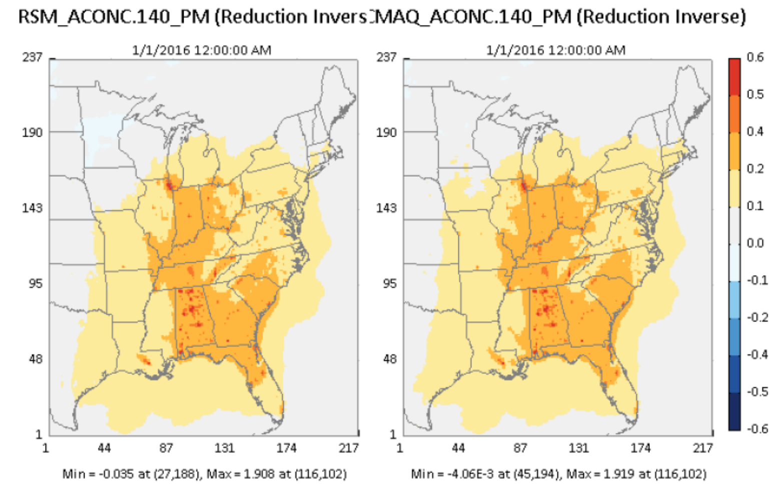
Jan., PM_SO_4 monthly avg., US

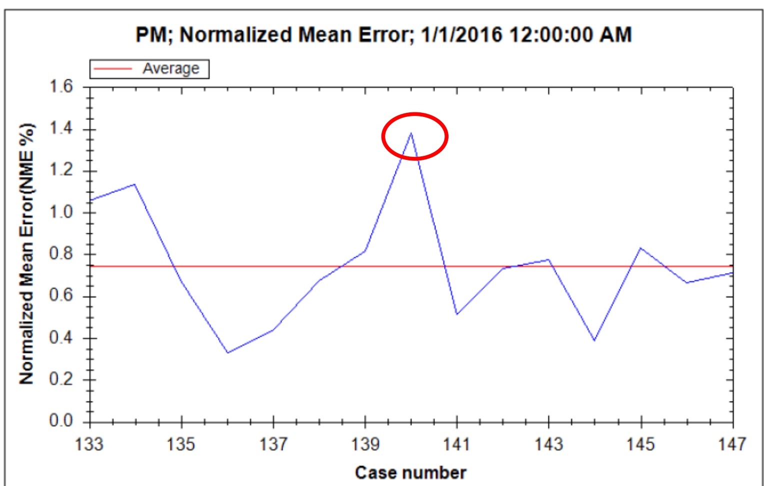
Out-of-Sample Validation: PM_SO_4 Conc. & Responses (Jan., PM_SO_4 monthly avg., US) Worst 1

$PM_{2.5}$ Conc: RSM vs. CMAQ

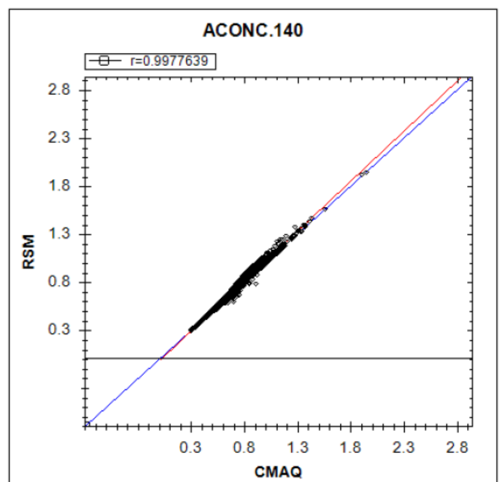


$PM_{2.5}$ Response: RSM vs. CMAQ

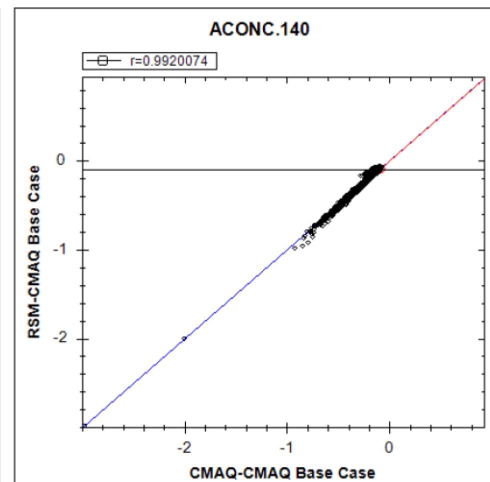




All "Out-of-Sample" runs: Normalized Mean Error



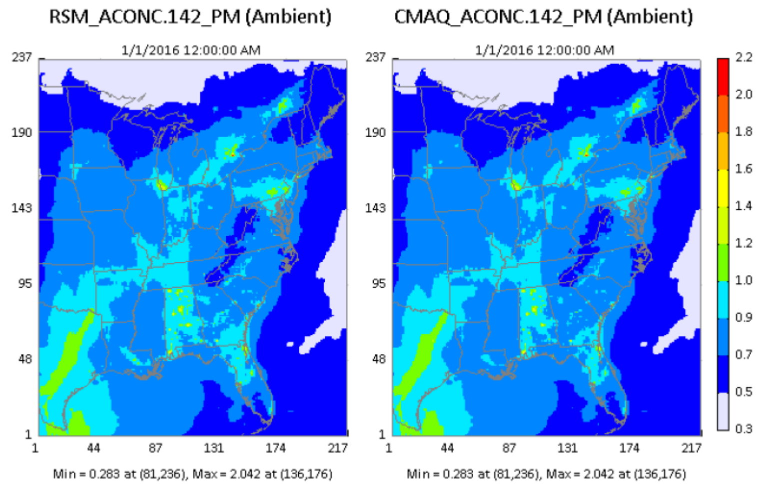
Predicted $PM_{2.5}$ Concentration (RSM vs. CMAQ)



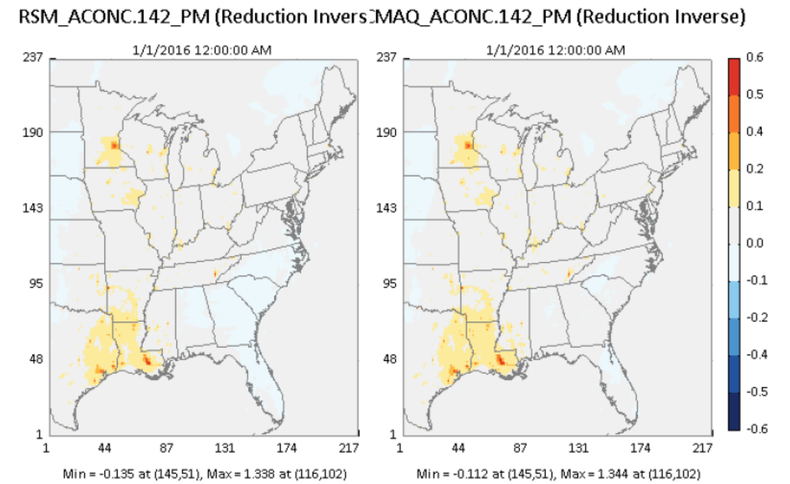
Predicted $PM_{.5}$ Response (RSM vs. CMAQ)

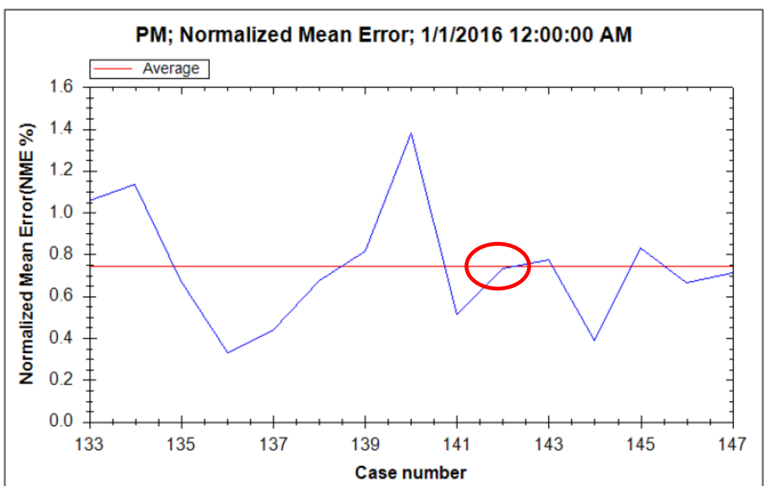
Out-of-Sample Validation: PM_SO_4 Conc. & Responses
(Jan., PM_SO_4 monthly avg., US) Typical 1

$PM_{2.5}$ Conc: RSM vs. CMAQ

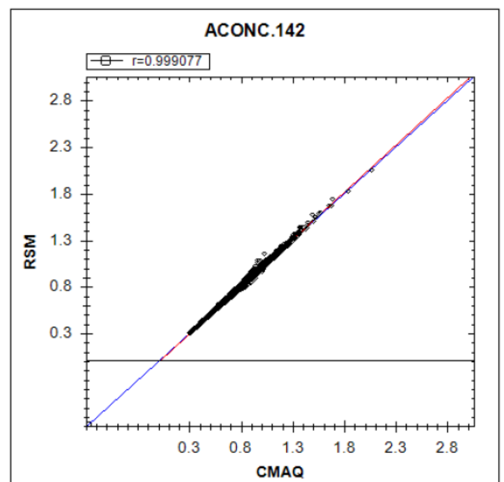


$PM_{2.5}$ Response: RSM vs. CMAQ

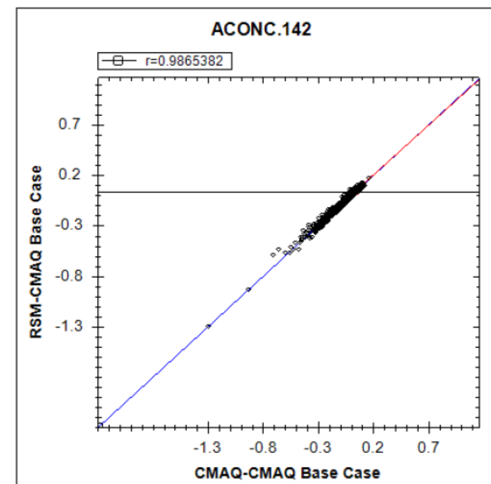




All "Out-of-Sample" runs: Normalized Mean Error



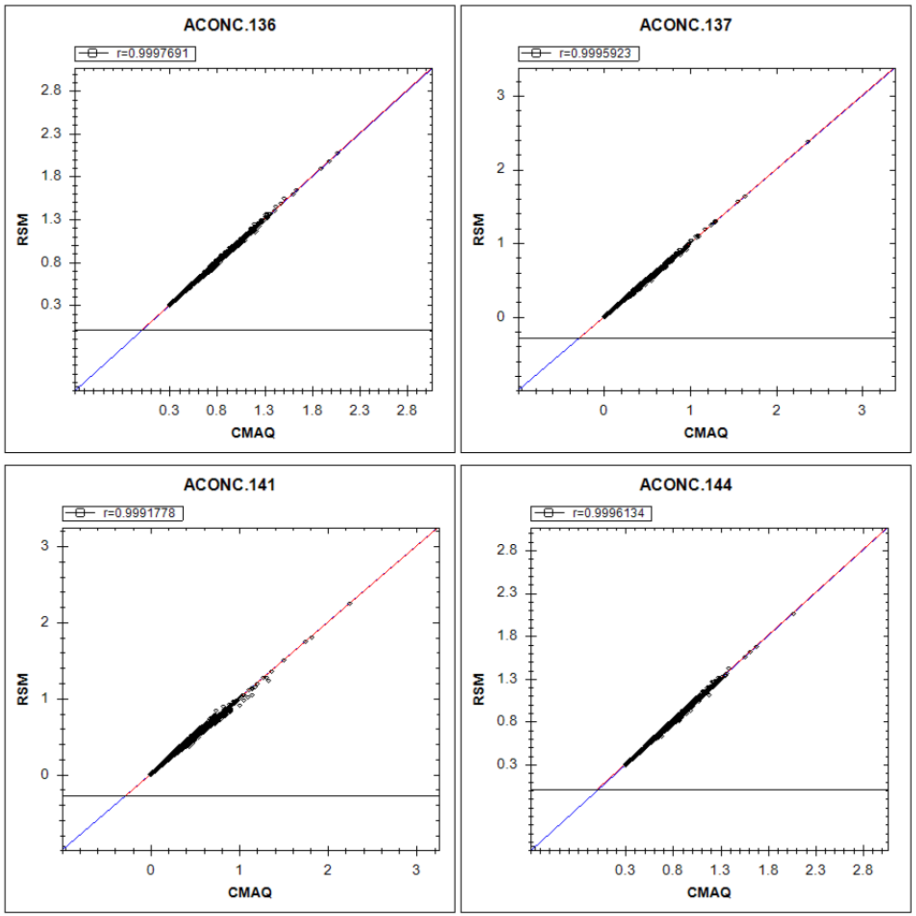
Predicted $PM_{2.5}$ Concentration (RSM vs. CMAQ)



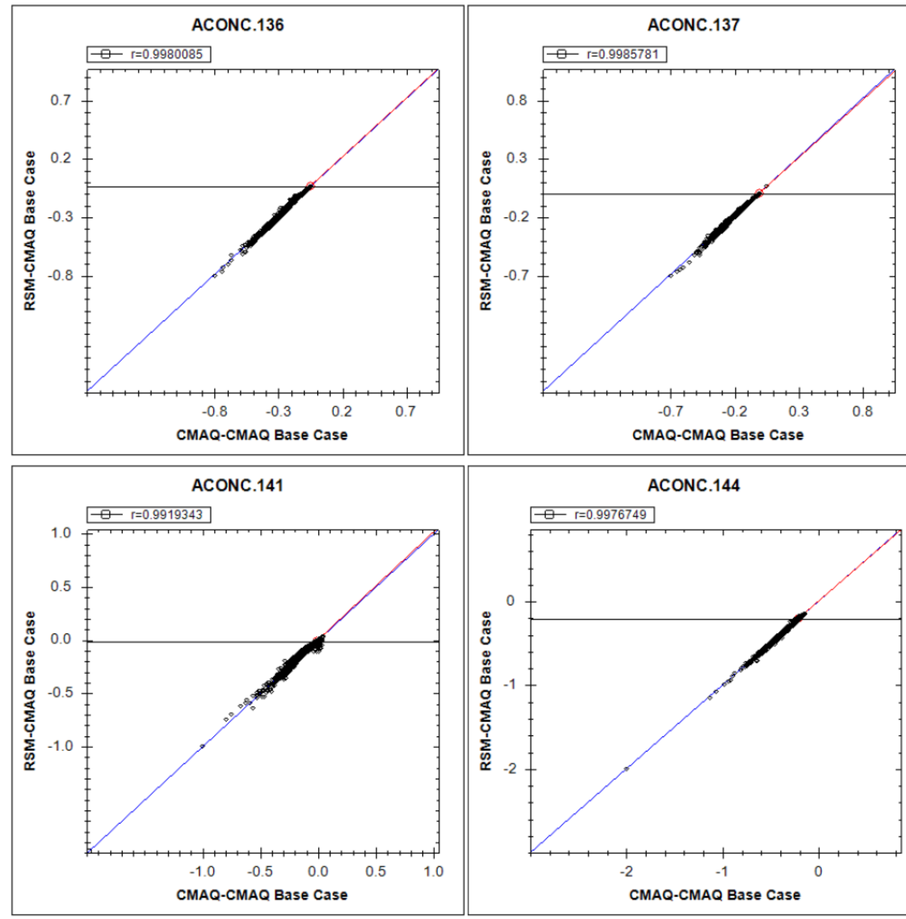
Predicted $PM_{.5}$ Response (RSM vs. CMAQ)

Out-of-Sample Validation: Concentration Comparison (Jan., PM_SO_4 monthly avg., US) Best 4

Predicted $PM_{2.5}$ Conc. (RSM vs. CMAQ)

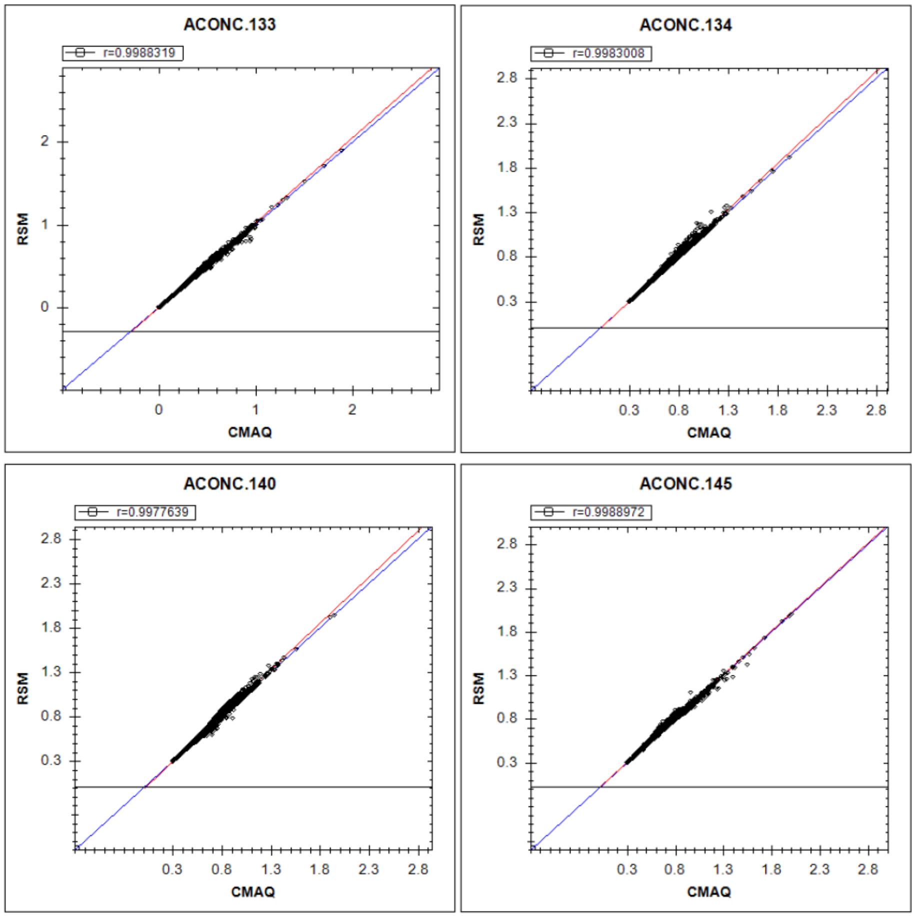


Predicted $PM_{2.5}$ Responses (RSM-Base vs. CMAQ-Base)

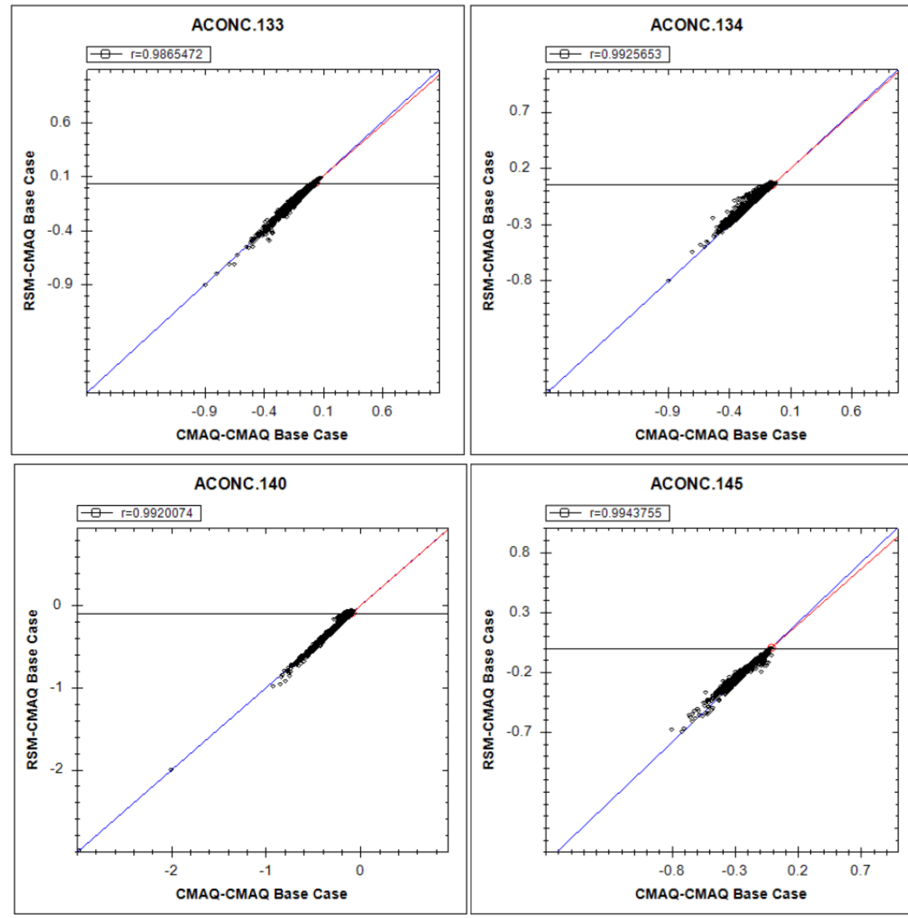


Out-of-Sample Validation: Concentration Comparison (Jan., PM_SO_4 monthly avg., US) Worst 4

Predicted $PM_{2.5}$ Conc.
(RSM vs. CMAQ)

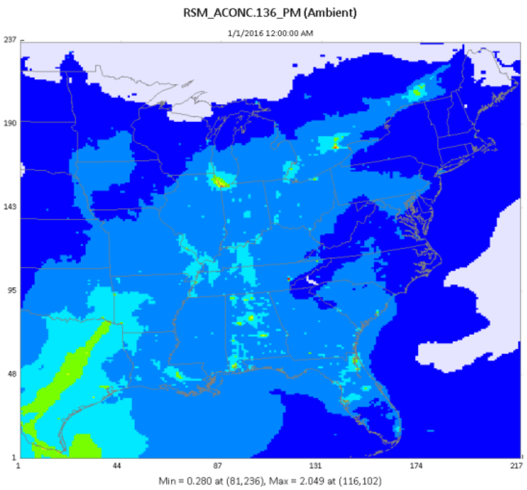


Predicted $PM_{2.5}$ Responses
(RSM-Base vs. CMAQ-Base)

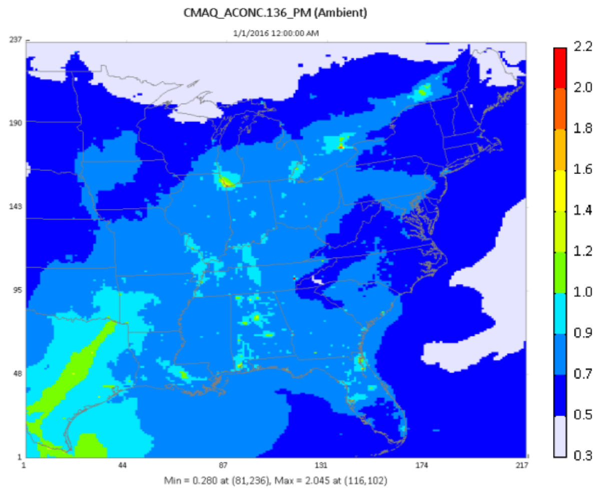


Out-of-Sample Validation: Predicted Concentrations (Jan., PM_SO_4 monthly avg., US) Best 1

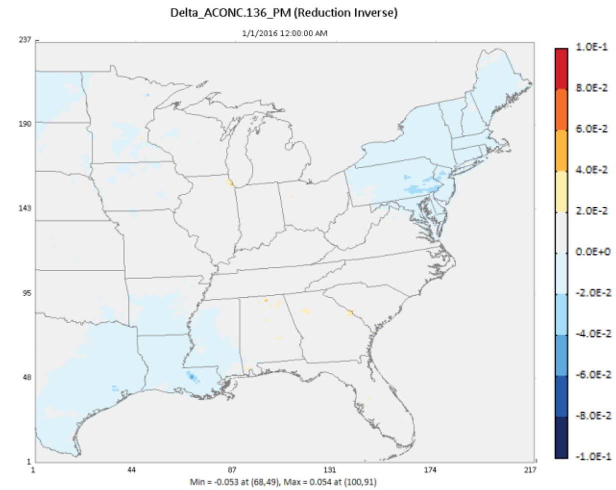
RSM



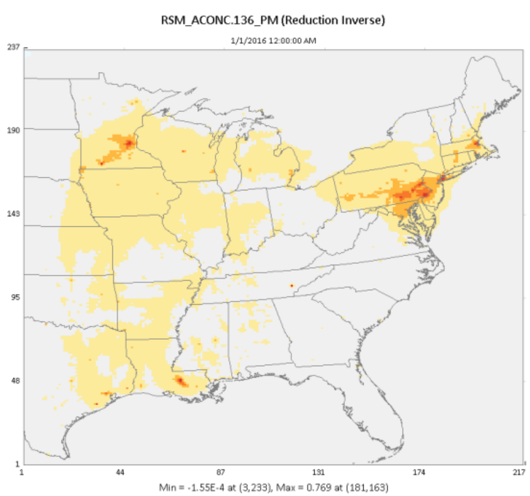
CMAQ



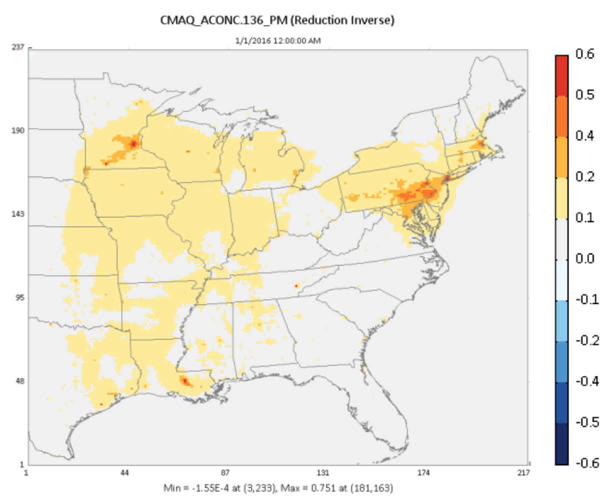
Delta



Base - RSM

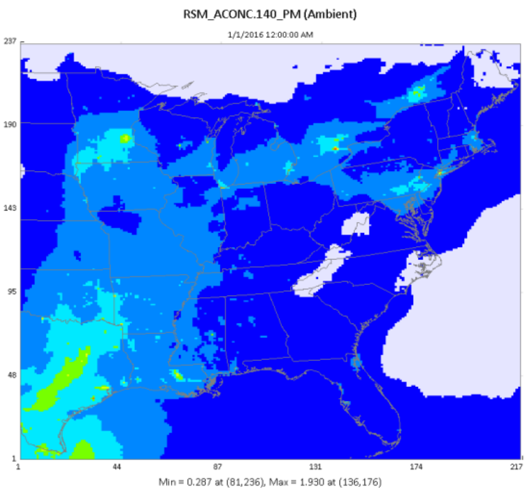


Base - CMAQ

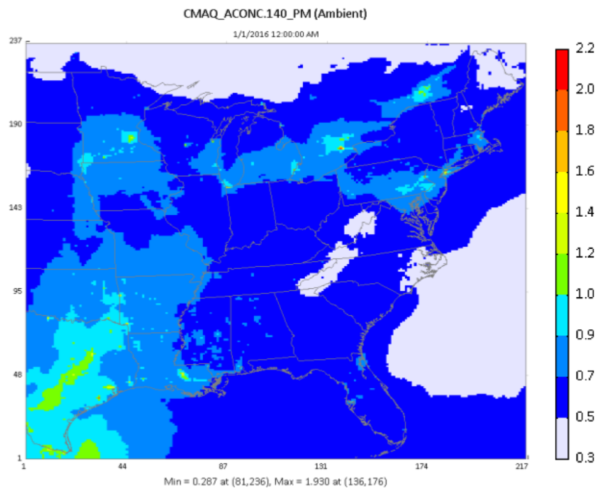


Out-of-Sample Validation: Predicted Concentrations (Jan., PM_SO_4 monthly avg., US) Worst 1

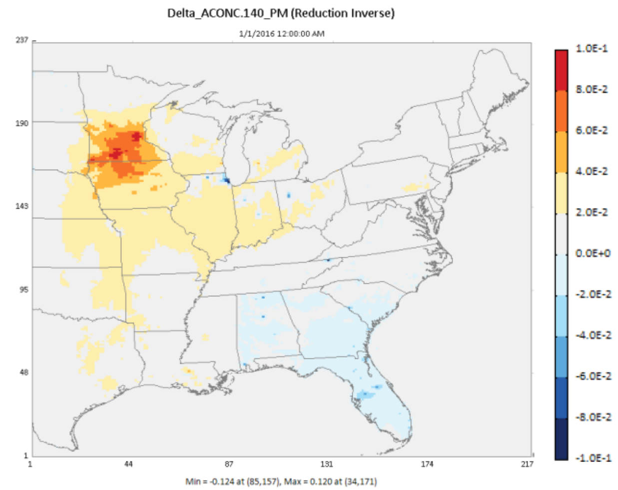
RSM



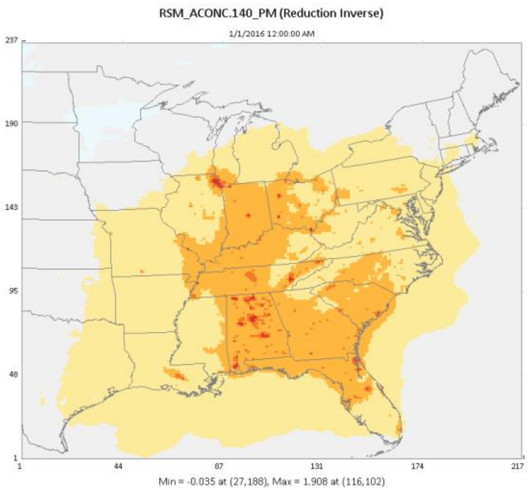
CMAQ



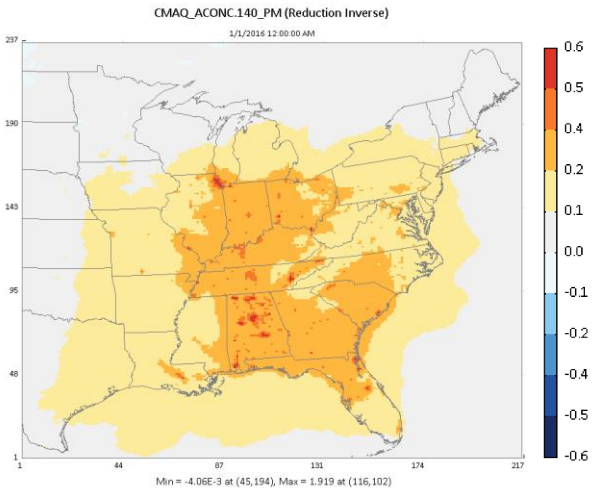
Delta



Base - RSM



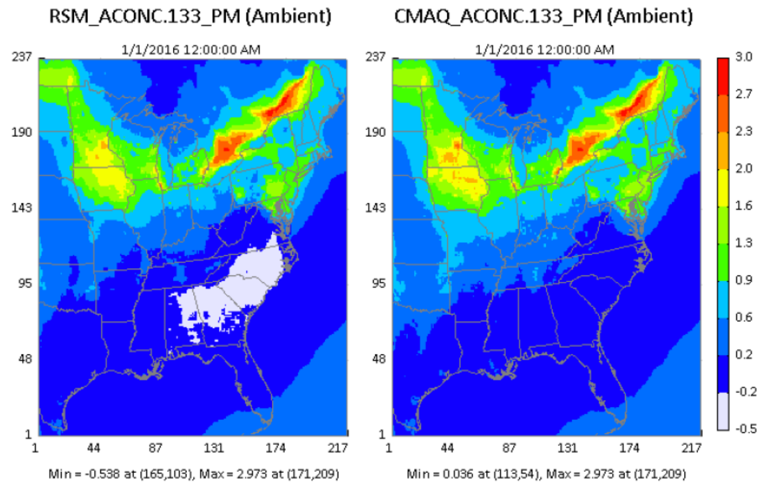
Base - CMAQ



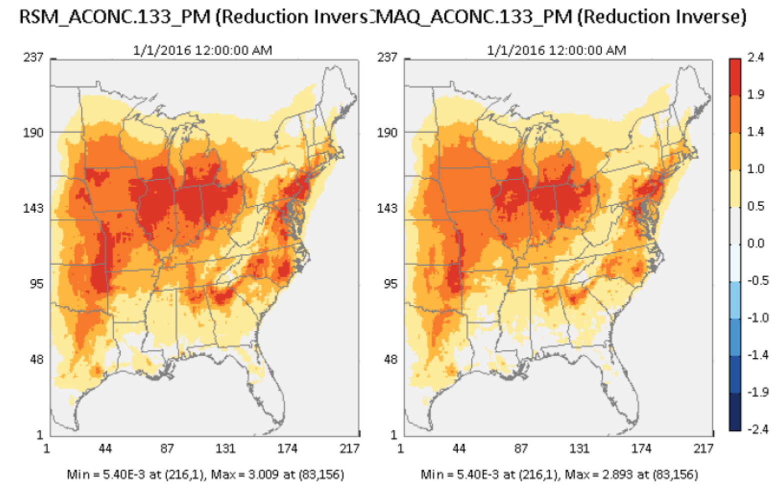
Jan., PM_NO_3 monthly avg., US

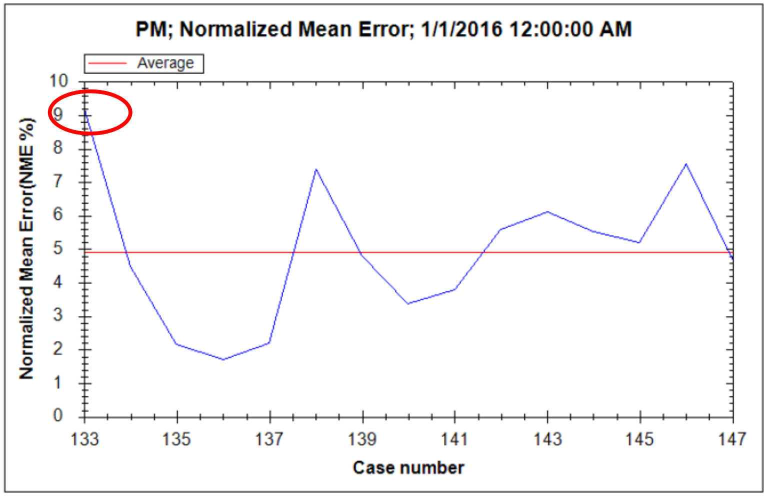
Out-of-Sample Validation: PM_NO_3 Conc. & Responses
(Jan., PM_NO_3 monthly avg., US) Worst 1

$PM_{2.5}$ Conc: RSM vs. CMAQ

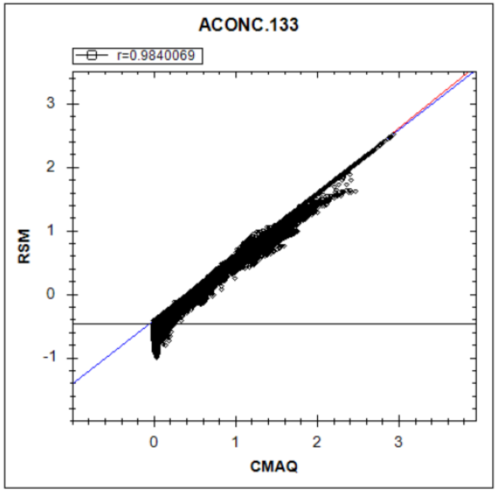


$PM_{2.5}$ Response: RSM vs. CMAQ

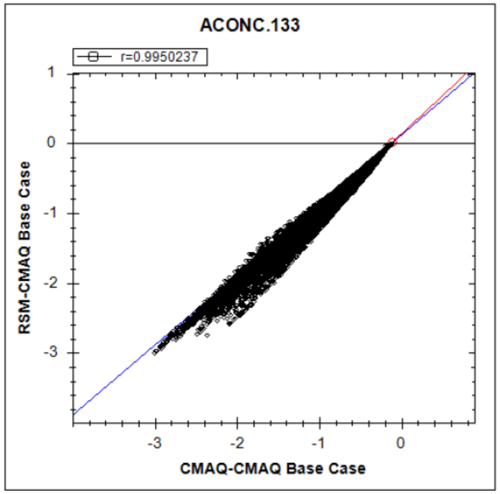




All "Out-of-Sample" runs: Normalized Mean Error



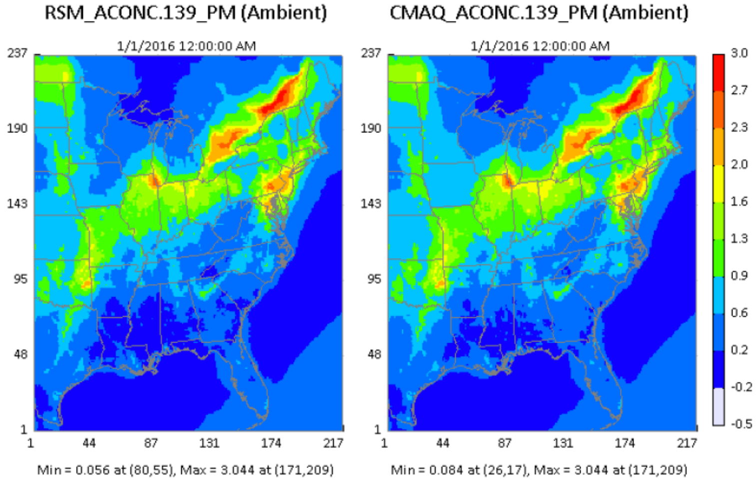
Predicted $PM_{2.5}$ Concentration (RSM vs. CMAQ)



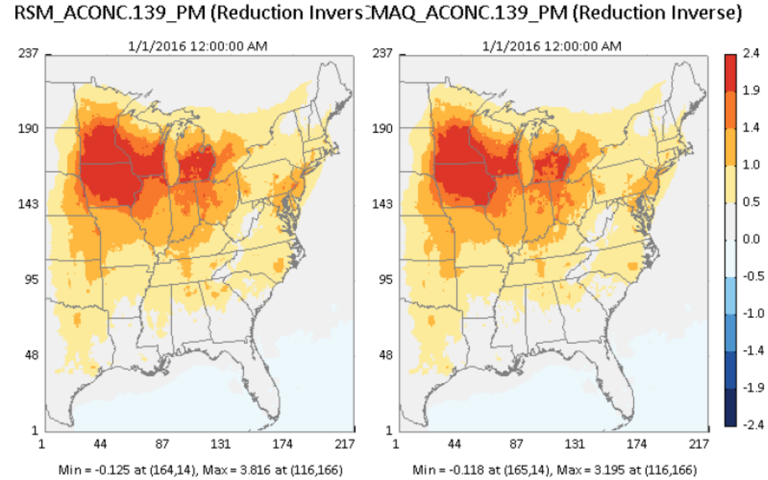
Predicted $PM_{.5}$ Response (RSM vs. CMAQ)

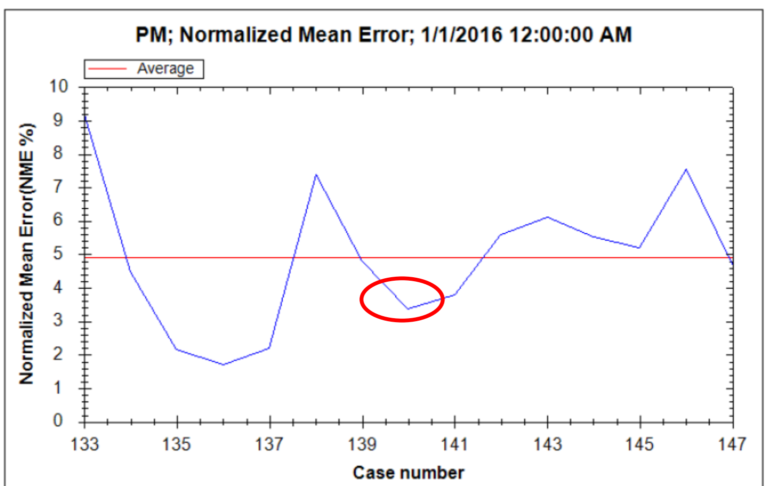
Out-of-Sample Validation: PM_NO$_3$ Conc. & Responses (Jan., PM_NO$_3$ monthly avg., US) Typical 1

PM$_{2.5}$ Conc: RSM vs. CMAQ

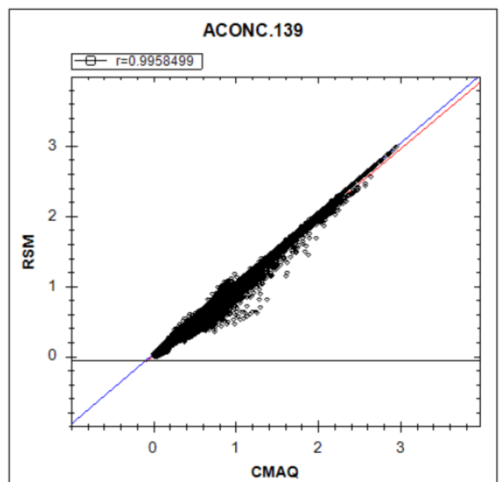


PM$_{2.5}$ Response: RSM vs. CMAQ

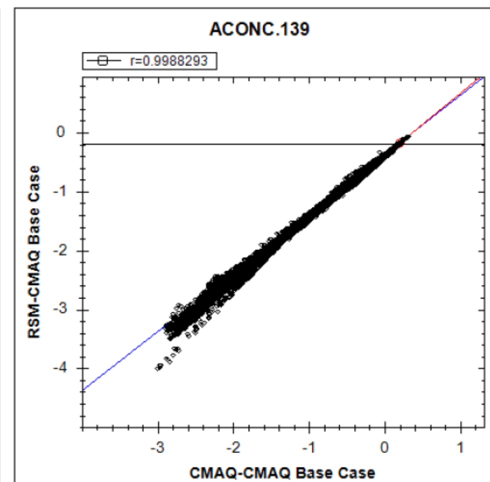




All "Out-of-Sample" runs: Normalized Mean Error



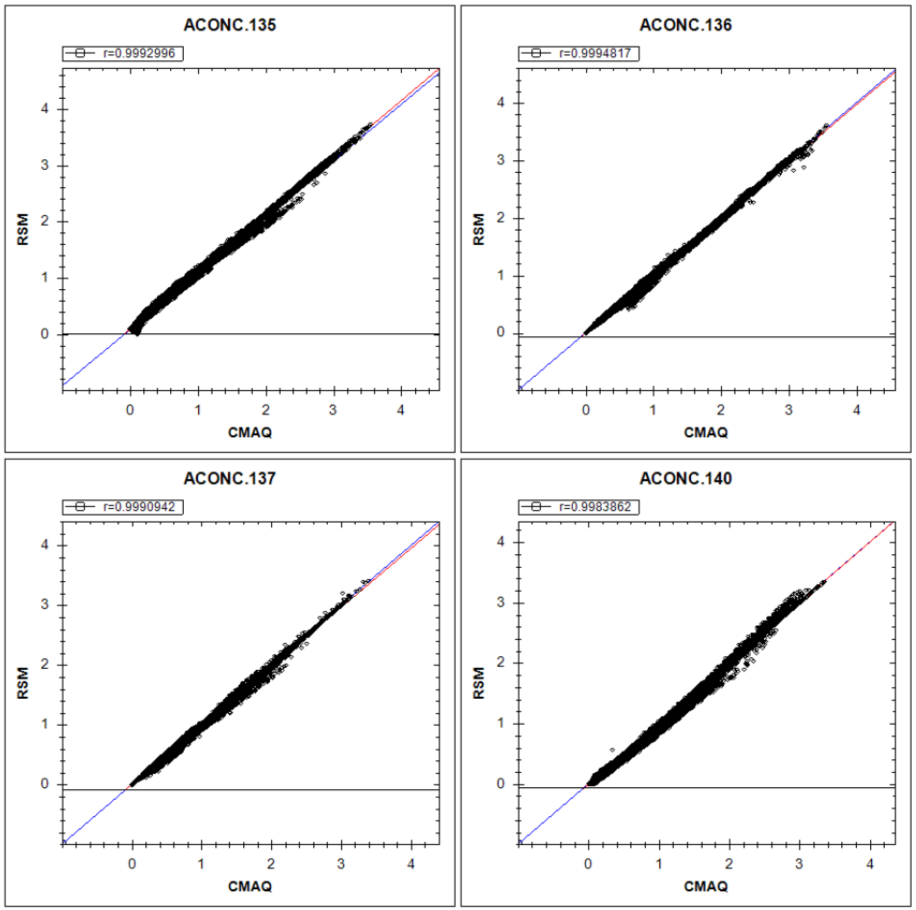
Predicted PM$_{2.5}$ Concentration (RSM vs. CMAQ)



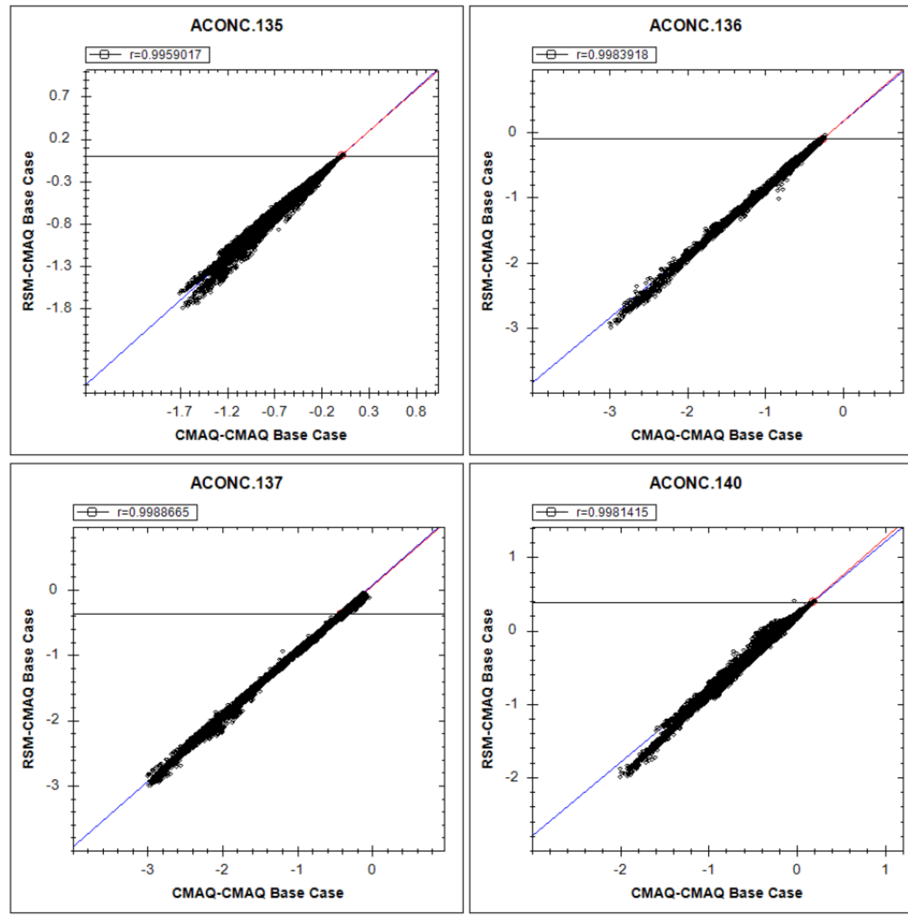
Predicted PM$_{.5}$ Response (RSM vs. CMAQ)

Out-of-Sample Validation: Concentration Comparison (Jan., PM_NO_3 monthly avg., US) Best 4

Predicted $PM_{2.5}$ Conc. (RSM vs. CMAQ)

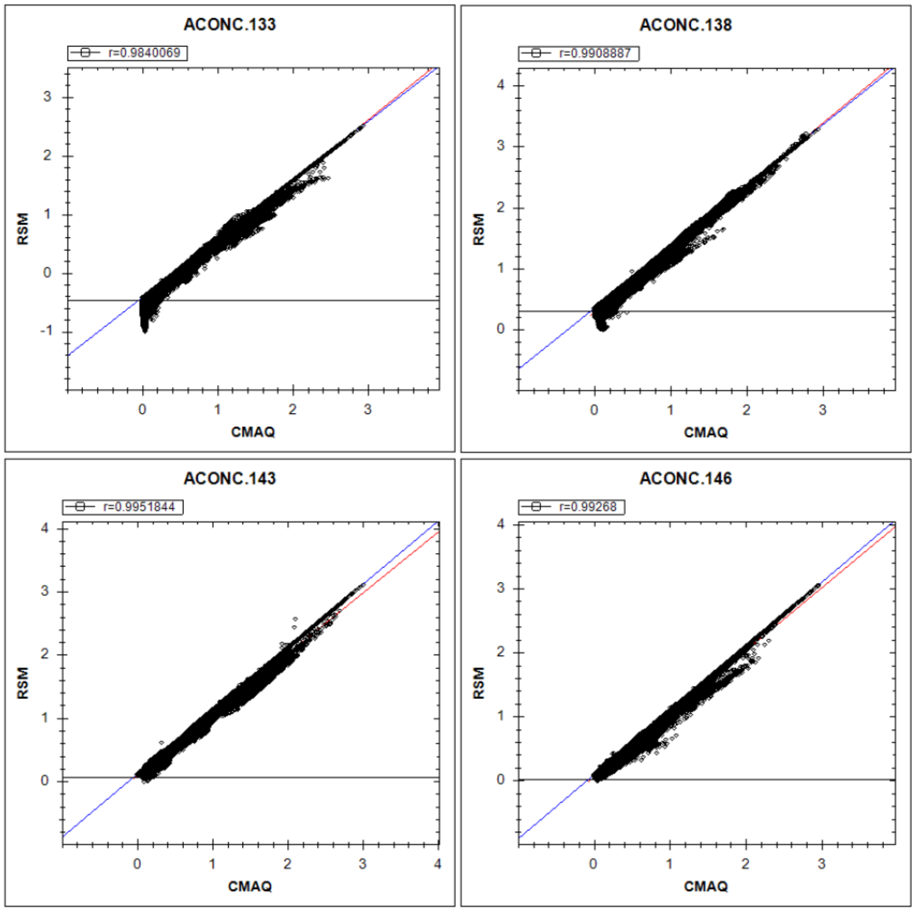


Predicted $PM_{2.5}$ Responses (RSM-Base vs. CMAQ-Base)

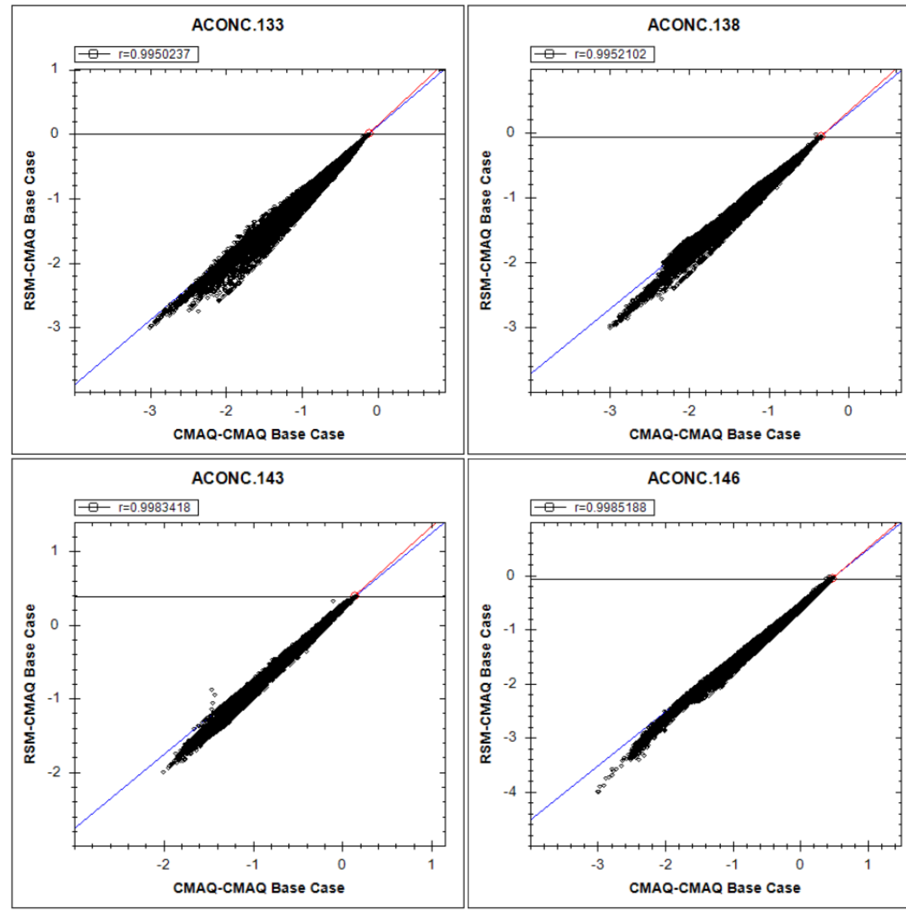


Out-of-Sample Validation: Concentration Comparison (Jan., PM_NO_3 monthly avg., US) Worst 4

Predicted $PM_{2.5}$ Conc. (RSM vs. CMAQ)

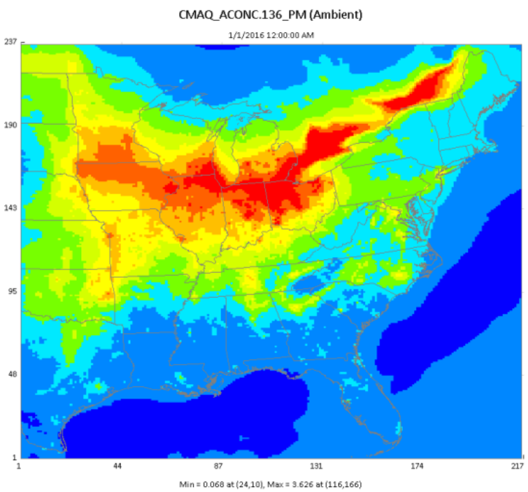


Predicted $PM_{2.5}$ Responses (RSM-Base vs. CMAQ-Base)

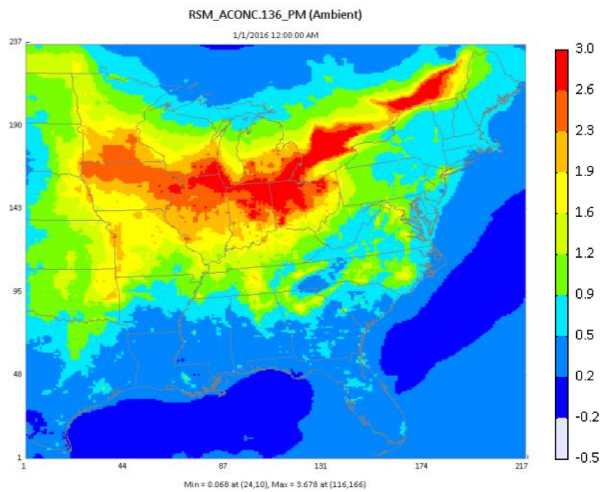


Out-of-Sample Validation: Predicted Concentrations (Jan., PM_NO_3 monthly avg., US) Best 1

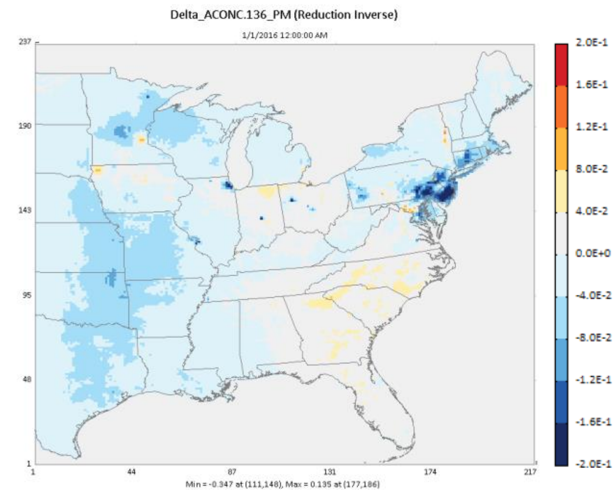
RSM



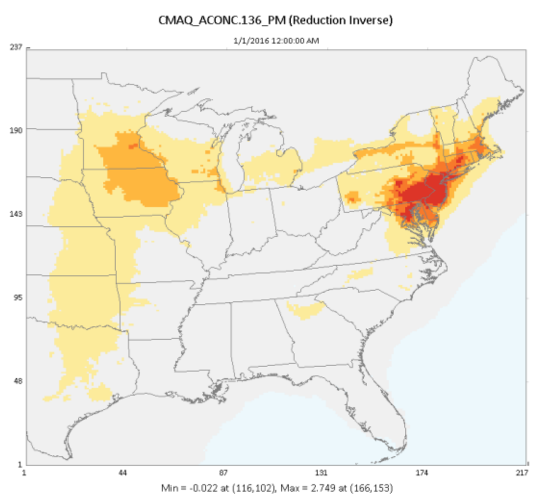
CMAQ



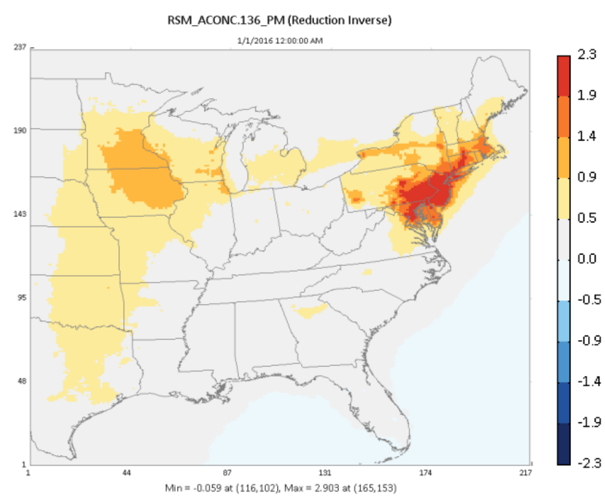
Delta



Base - RSM

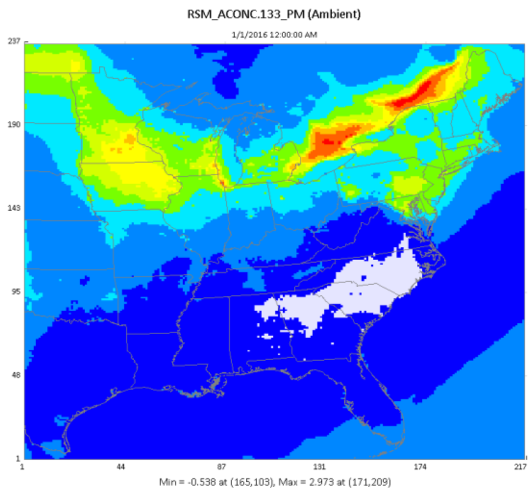


Base - CMAQ

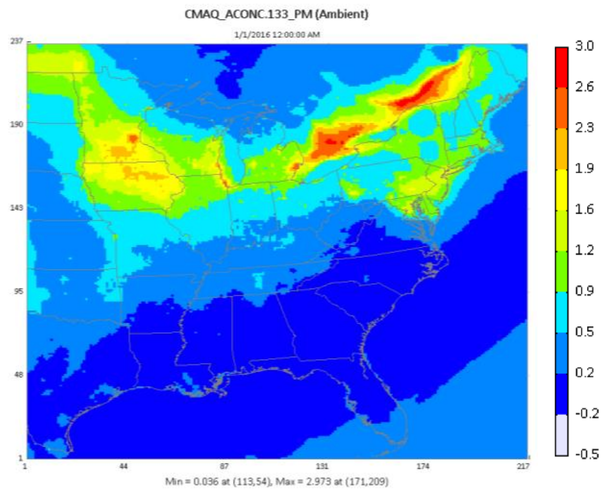


Out-of-Sample Validation: Predicted Concentrations (Jan., PM_NO_3 monthly avg., US) Worst 1

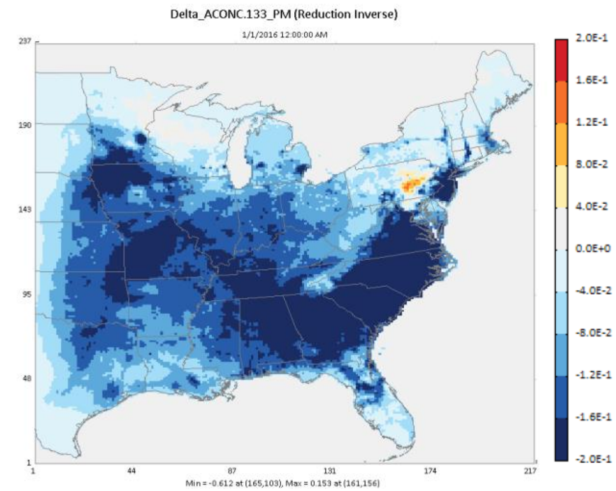
RSM



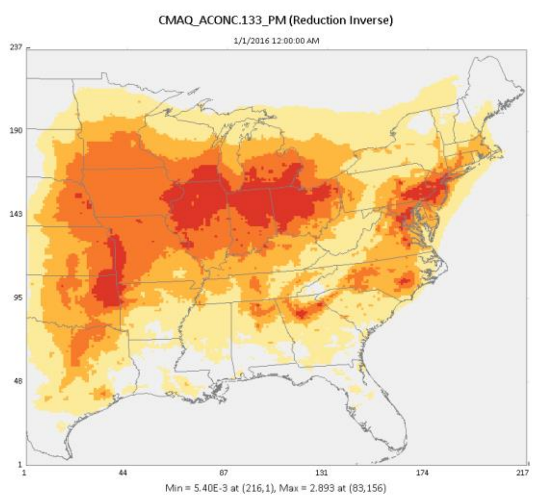
CMAQ



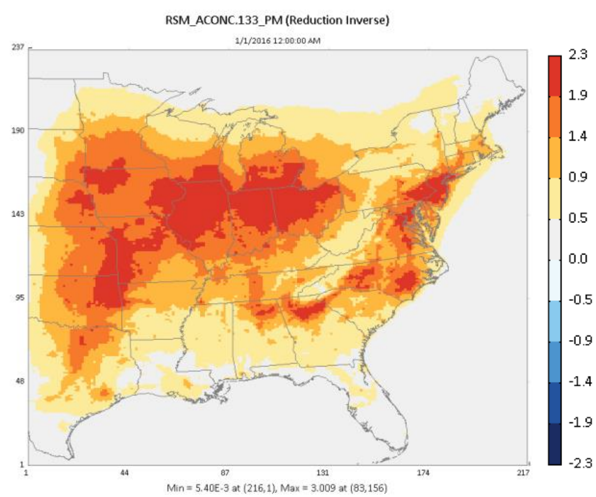
Delta



Base - RSM



Base - CMAQ



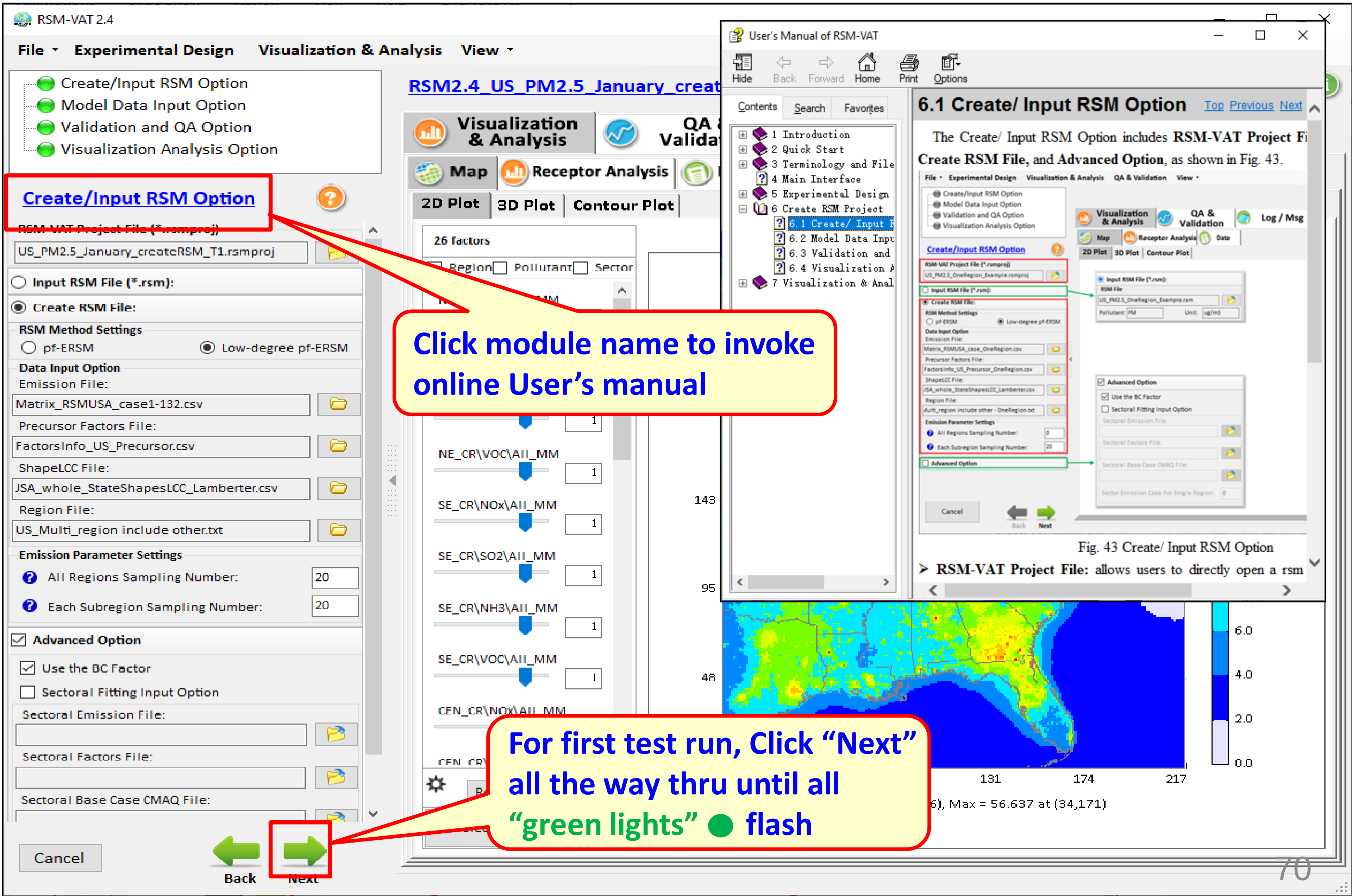
Appendix C:

RSM-VAT Demo & Tutorial:

EUS 12-km Case using "pf-RSM"

69

RSM-VAT: Launching Page & User's Manual



Click module name to invoke online User's manual

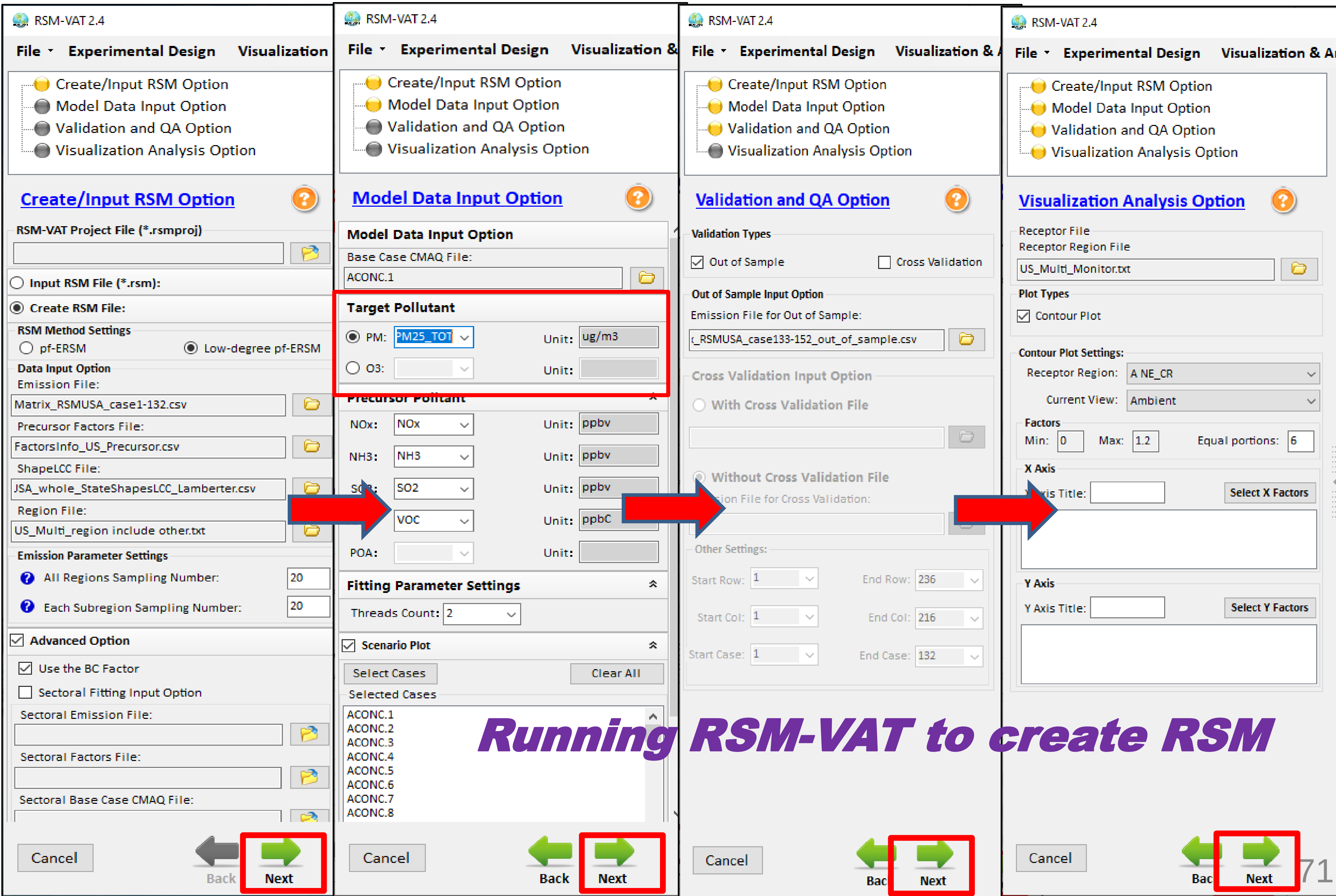
For first test run, Click "Next" all the way thru until all "green lights" ● flash

1. Create/Input RSM Option

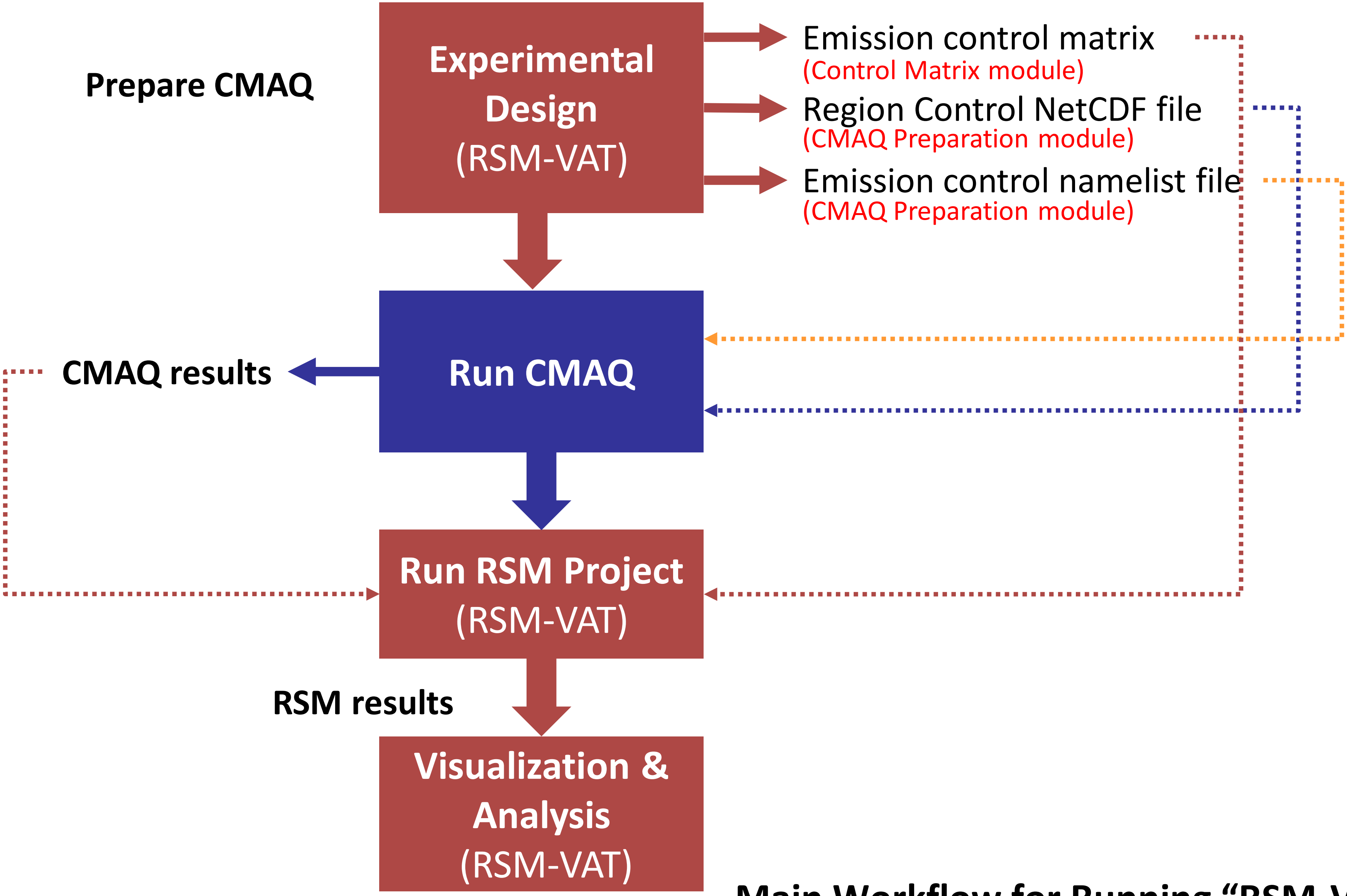
2. Model Data Input Option

3. Validation & Q/A Option

4. Vis./Analysis Option



Workflow to Create RSM in *"RSM-VAT"*

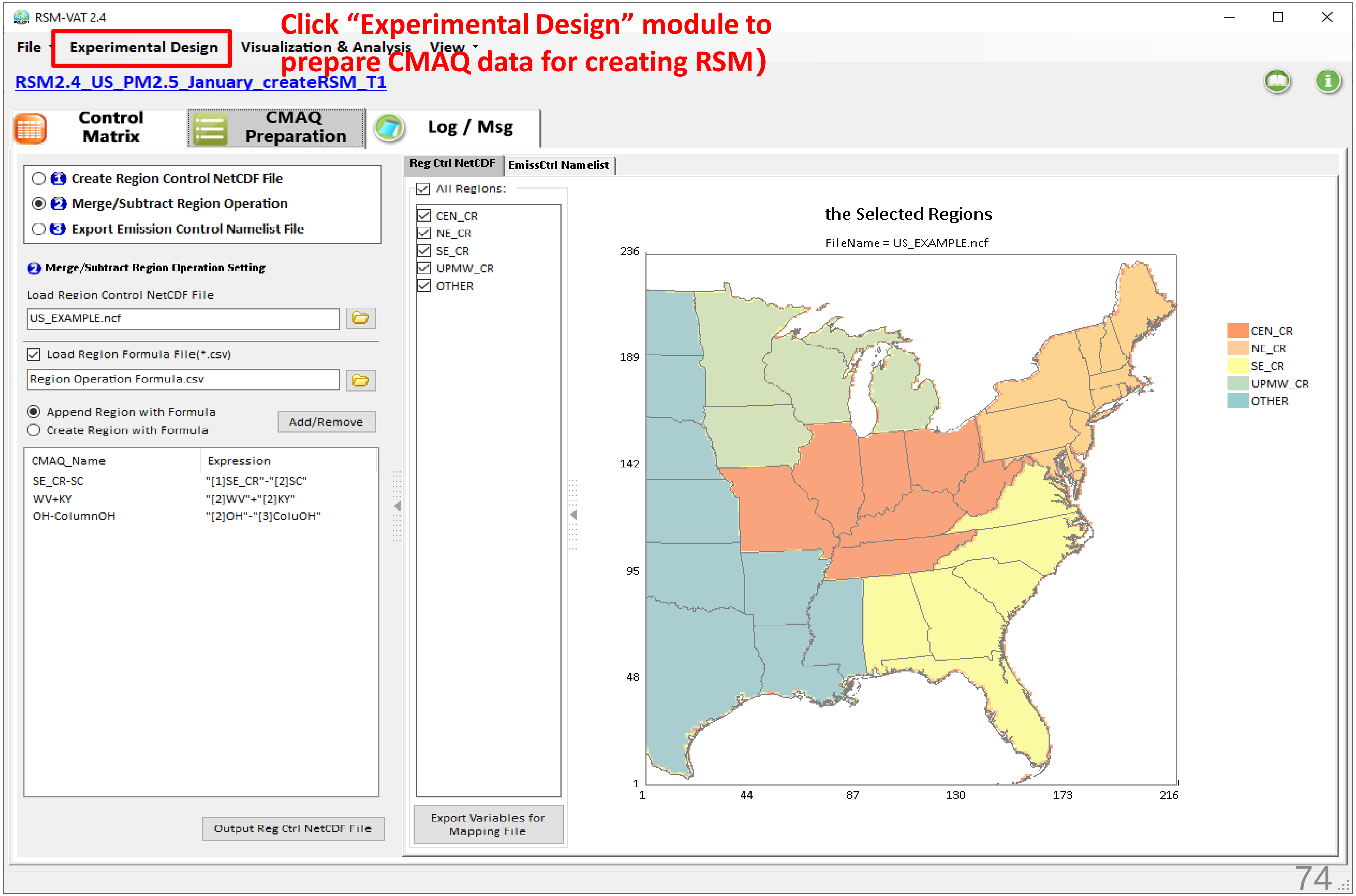


Main Workflow for Running "RSM-VAT"

Experimental Design
Running "Single Region" RSM

***<u>Note</u>**: Default RSM case is "5 Climate Regions" in RSM-VAT*

RSM-VAT: Experimental Design Module



Example Data

Default case for Experimental Design is for "5 Climate Region", but you can run "One Region" case & "Sector" case with the following setup:

- Input Data for Experimental Design and RSM runs for OneRegion: *C:\Users\...\Documents\My RSM-VAT Files\Data\Example\US_RSM – OneRegion*

Experimental Design projects (.expproj):

1. US_PM2.5_MultiRegion_Example.expproj
2. US_PM2.5_OneRegion_Example.expproj
3. US_PM2.5_Sector_OneRegion_Example.expproj

RSM projects (.rsmproj):

1. US_PM2.5_OneRegion_Example.rsmproj
2. US_PM2.5_Sector_OneRegion_Example.rsmproj

Note:

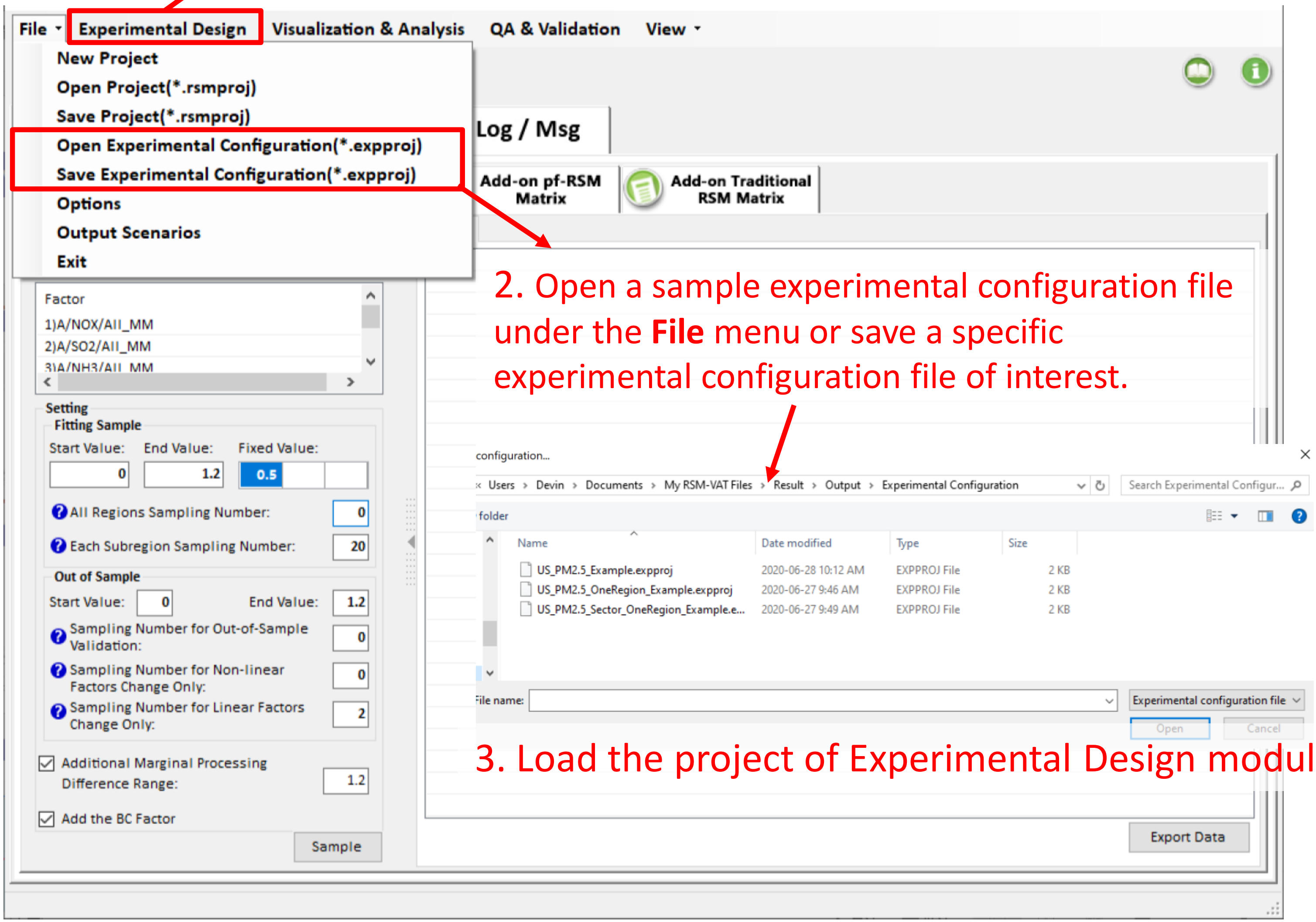
- **"US_PM2.5_MultiRegion_Example.expproj"**: the example project of experimental design module for multi-region RSM, while the corresponding input data (*C:\Users\...\Documents\My RSM-VAT Files\Data\Example\US_RSM*) and RSM project (.rsmproj) has been run and packaged in our last version of RSM, as shown here.

US_PM2.5_July.rsmproj	2020-04-20 12:16	RSMPROJ	56,064 KB
US_PM2.5_January.rsmproj	2020-04-20 12:12	RSMPROJ	56,064 KB

- **"US_PM2.5_OneRegion_Example"**: the example project of experimental design module for single-region RSM, in which all regions were merged into one region;
- **"US_PM2.5_Sector_OneRegion_Example"**: the example project of experimental design module for sectoral single-region RSM, in which the sources of NOx were divided into 'Points', 'Mobile' and 'Others'

Experimental Design

1. Click **Experimental Design** menu to switch to this module.



2. Open a sample experimental configuration file under the **File** menu or save a specific experimental configuration file of interest.

3. Load the project of Experimental Design module

Steps for Experimental Design

Option 1 (recommended and used in our example case)

Step 1: Create the emission control matrix for single-region RSM

⇩

Step 2: Create region control NetCDF file based on region control shapefile (.shp) that consisted of multiple regions

⇩

Step 3: Create the new single-region NetCDF file through merging the regions in the NetCDF file generated in Step 2 using RSM-VAT

⇩

Step 4: Export Emission Control Namelist file based on the emission control matrix and single-region NetCDF file generated in step 1 and step 3 respectively

Option 2 (not recommended because the Step 3 is time-consuming)

Step 1: Create the emission control matrix for single-region RSM

⇩

Step 2: Create single-region shapefile (.shp) that consisted of only one region using ArcGIS

⇩

Step 3: Create the single-region NetCDF file based on the single-region shapefile (.shp) generated in Step 2

⇩

Step 4: Export Emission Control Namelist file based on the emission control matrix and single-region NetCDF file generated in step 1 and step 3 respectively

Control Matrix for Single Region

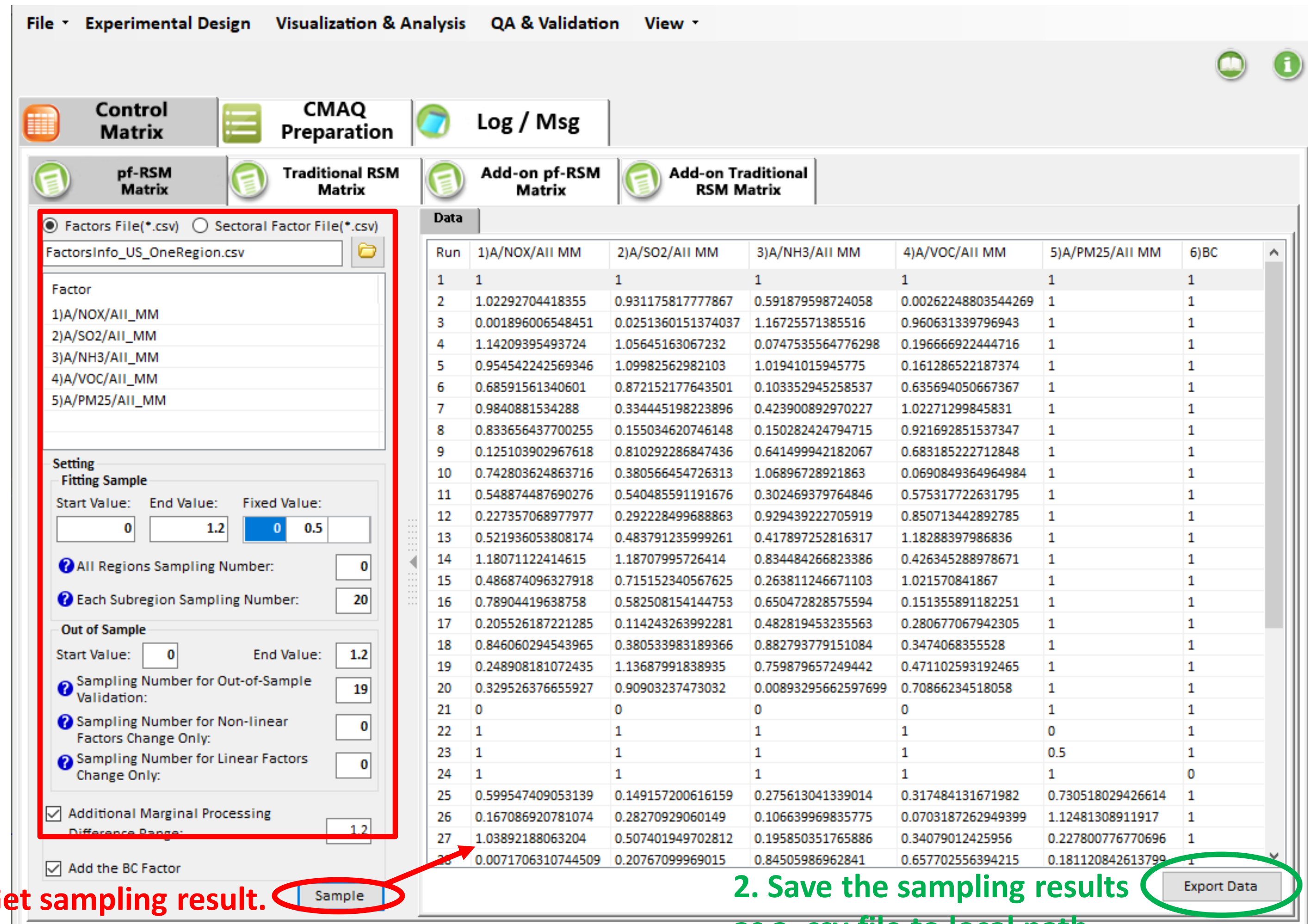
➢ Step 1: Create emission control matrix for RSM

Key Settings for single region matrix:

All Regions Sampling Number: 0

Each subregion Sampling Number: 20

Sampling Setting: see more tips by hovering the mouse over the icon ?



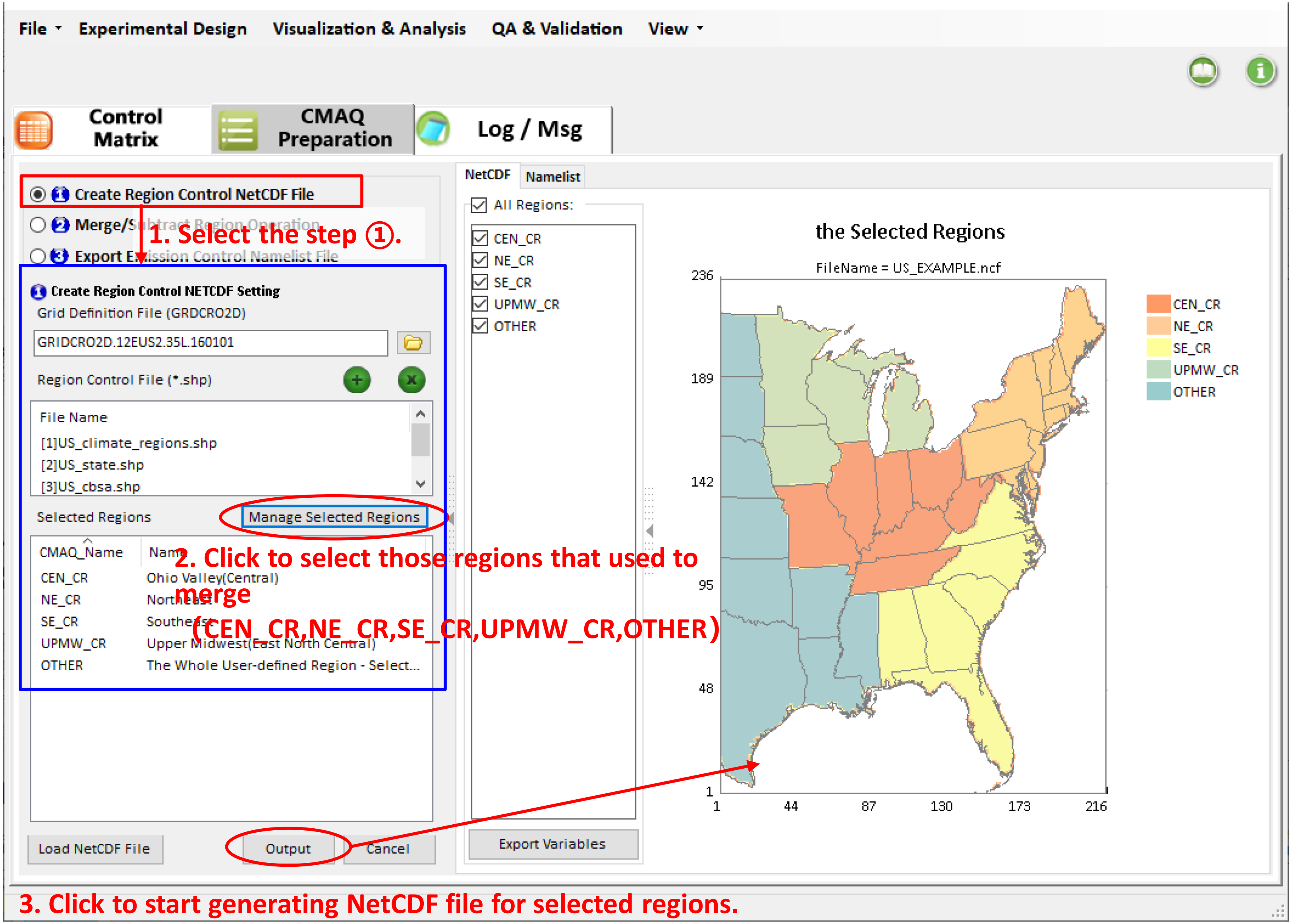
1. Get sampling result.

2. Save the sampling results as a .csv file to local path.

(\My RSM-VAT Files\Data\Example\US_RSM - OneRegion\Config Files\FactorsInfo_US_OneRegion.csv)

CMAQ Preparation for Single Region

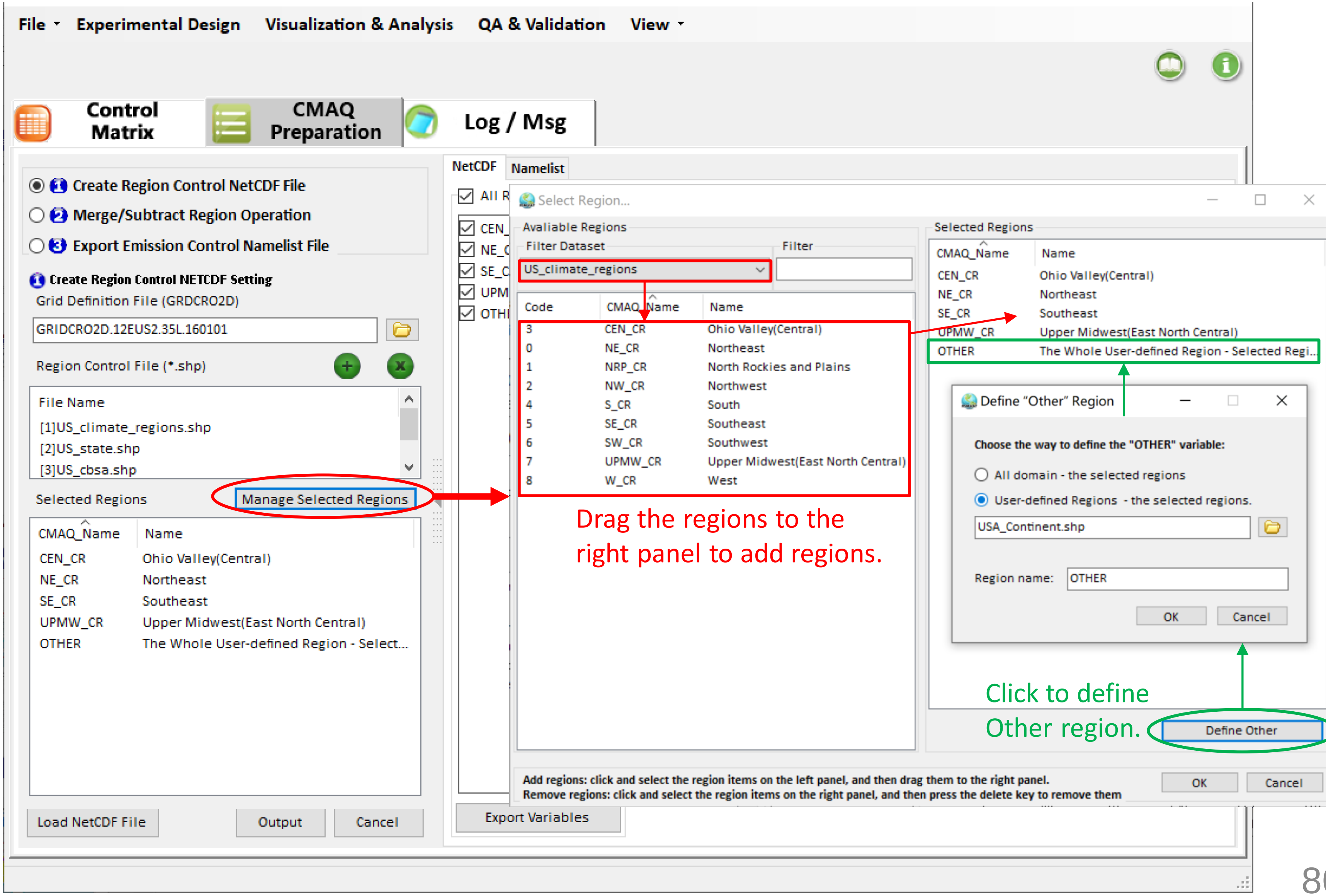
- **Step 2: Create region control NetCDF file based on region control shapefile (.shp)**



CMAQ Preparation for Single Region

Mange Selected Regions:

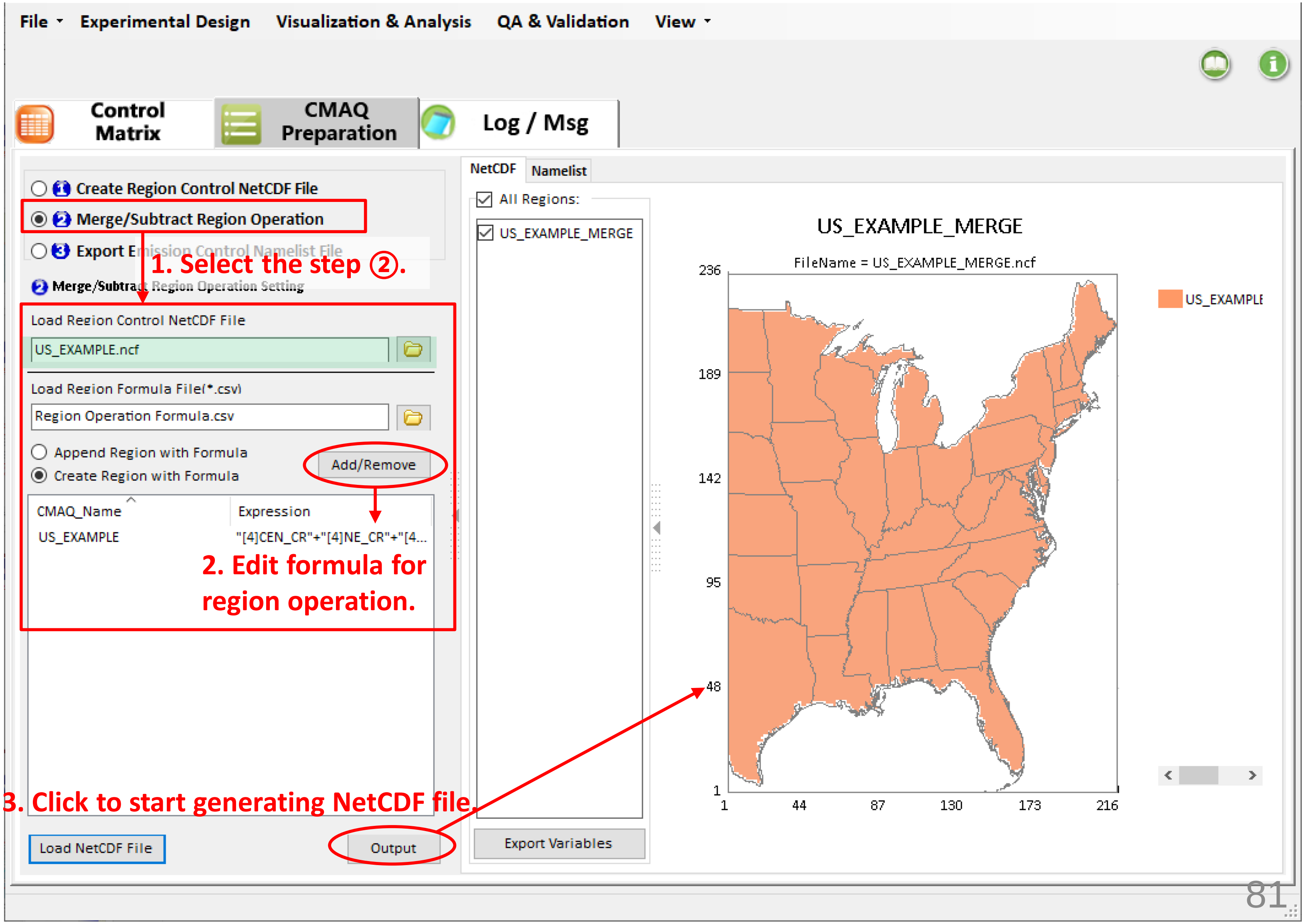
Click to select those regions that used to merge (CEN_CR,NE_CR,SE_CR,UPMW_CR,OTHER)



CMAQ Preparation for Single Region

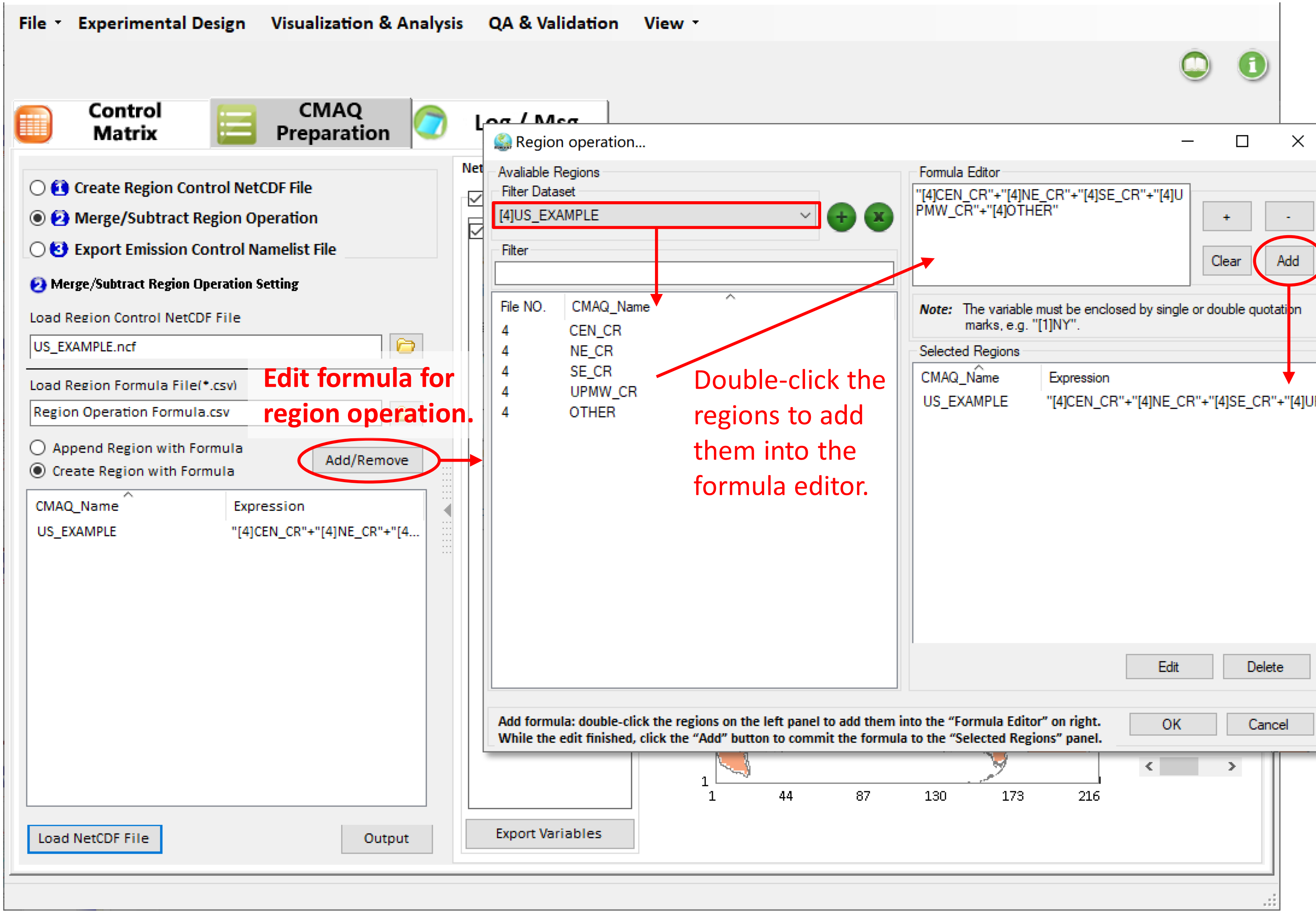
- **Step 3: Get the new NetCDF file for Single Region through merge operation on the generated NetCDF file from step ①.**

Use the NetCDF file generated from step ① as the input.



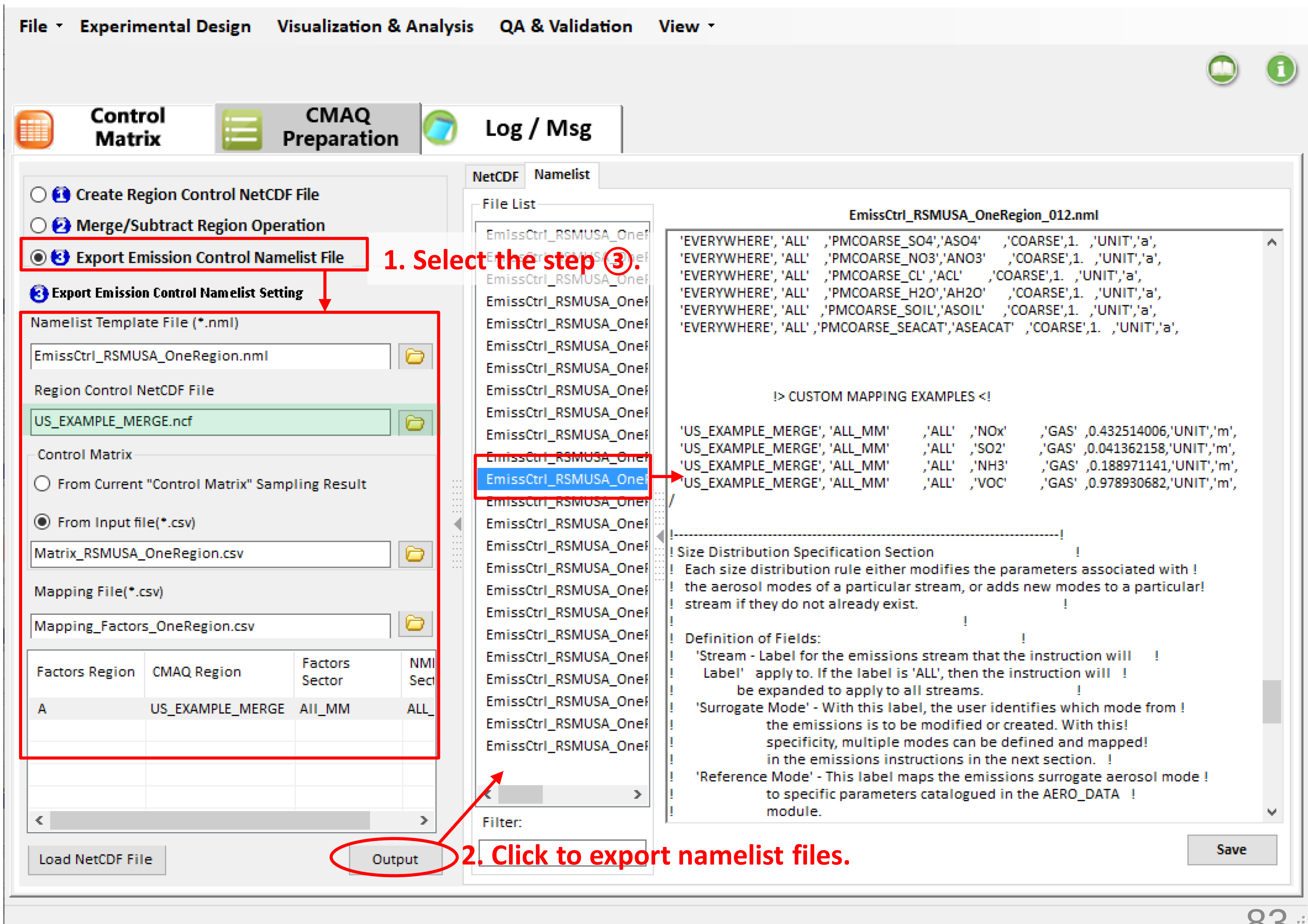
CMAQ Preparation for Single Region

Edit formula for region operation (merge regions)



CMAQ Preparation for Single Region

- **Step 4: Export Emission Control Namelist file.**



Use the NetCDF file generated from step ② as the input.

Experimental Design
(Single Region with sectors)

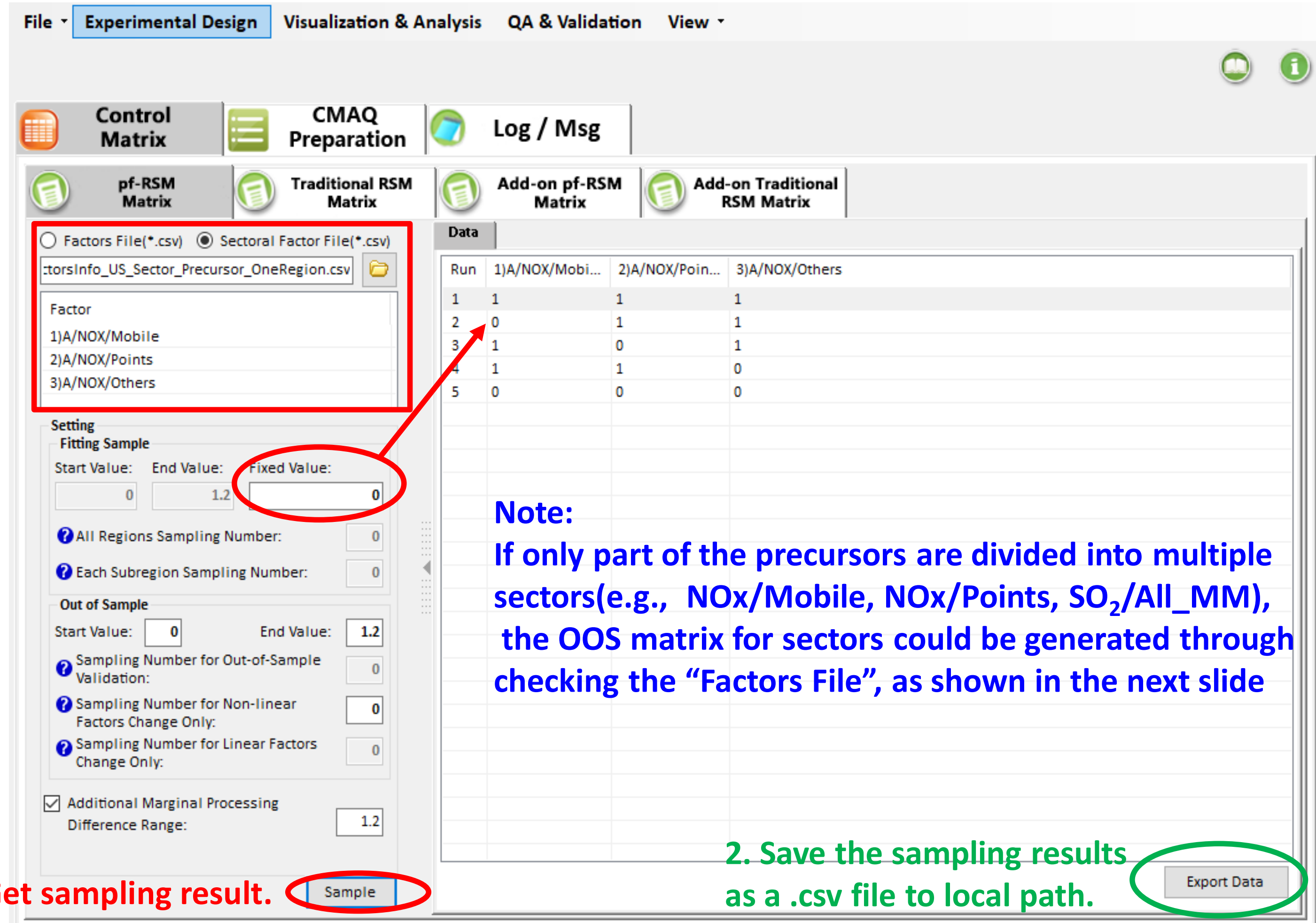
Control Matrix for Single Region with sectors

➢ **Step 1: Create sectoral emission control matrix for RSM**

Key Settings for Single Region with sectors:

Fixed Value: 0

Only the precursors that contain different sectors are needed to be listed in the sectoral file



Note:
If only part of the precursors are divided into multiple sectors(e.g., NOx/Mobile, NOx/Points, SO_2/All_MM), the OOS matrix for sectors could be generated through checking the "Factors File", as shown in the next slide

1. Get sampling result.

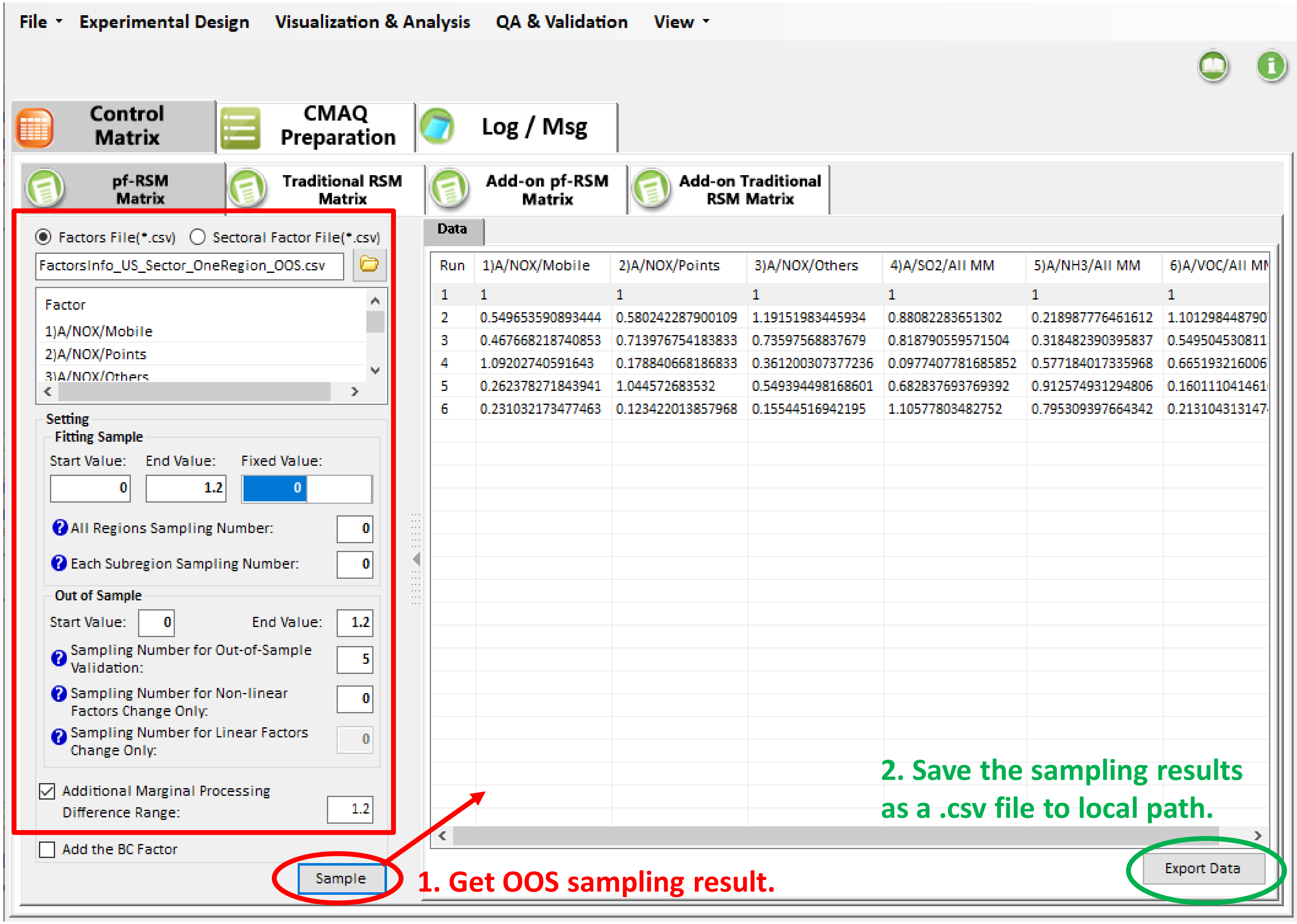
2. Save the sampling results as a .csv file to local path.

(\My RSM-VAT Files\Data\Example\US_RSM - OneRegion\Config Files\FactorsInfo_US_Sector_Precursor_OneRegion.csv)

Control Matrix for Single Region with sectors

➢ Create out-of-sample emission control matrix for sectors

Generating the OOS matrix for factors that contain both the sectoral precursor (e.g., NOx/Mobile) and All_MM precursor (e.g., SO_2/All)



1. Get OOS sampling result.

2. Save the sampling results as a .csv file to local path.

(\My RSM-VAT Files\Data\Example\US_RSM - OneRegion\Config Files\FactorsInfo_US_Sector_OneRegion_OOS.csv)

CMAQ Preparation for Single Region with sectors

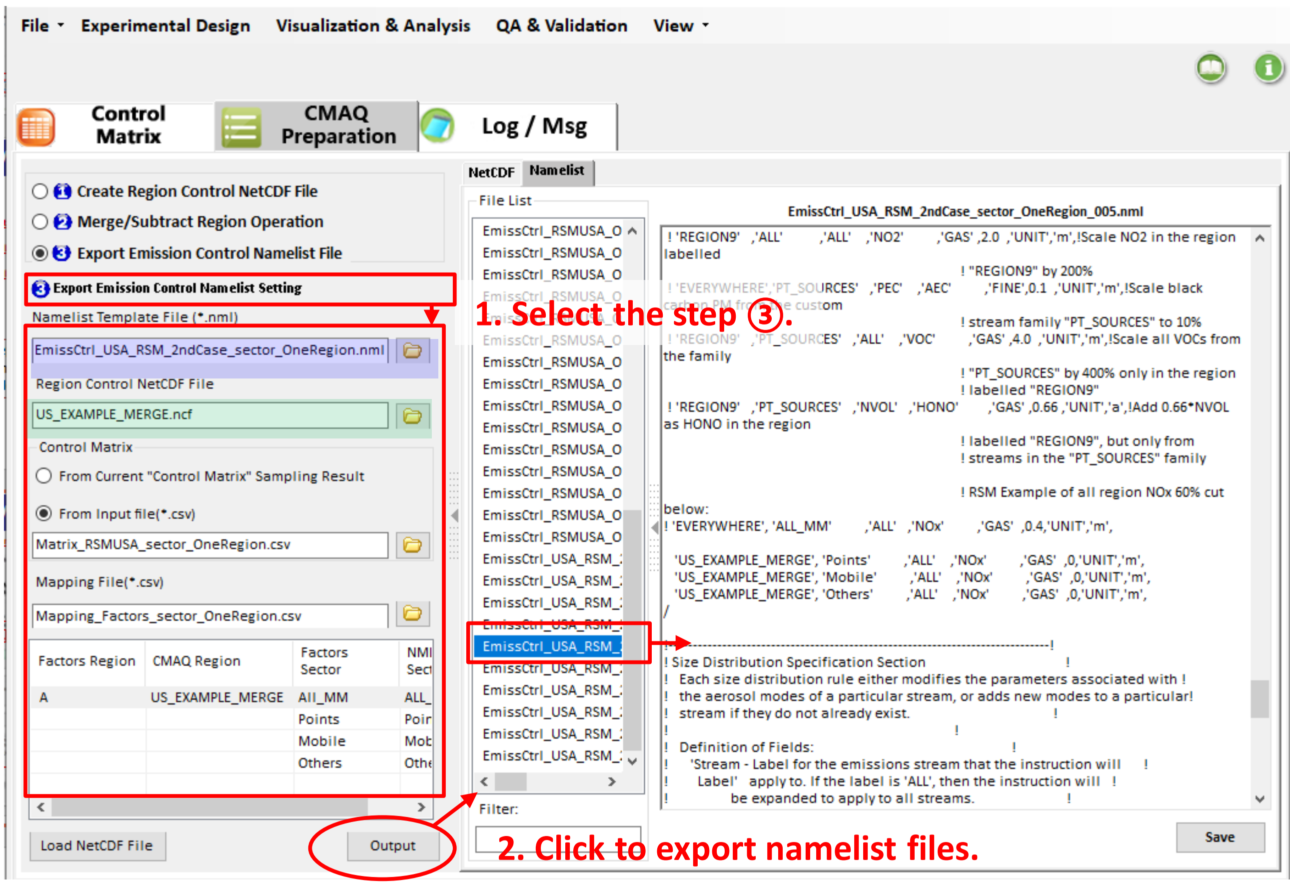
- **Step 2: Create region control NetCDF file based on region control shapefile (.shp)**
- **Step 3: Get the new NetCDF file for Single Region through merge operation on the generated NetCDF file from step ①.**

The above steps for single region with sectors are the same to "single region" (detailed in Slide 8 to 10)

- **Step 4: Export Sectoral Emission Control Namelist file.**

Use the namelist template file from EPA

Use the NetCDF file generated from step ② as the input.

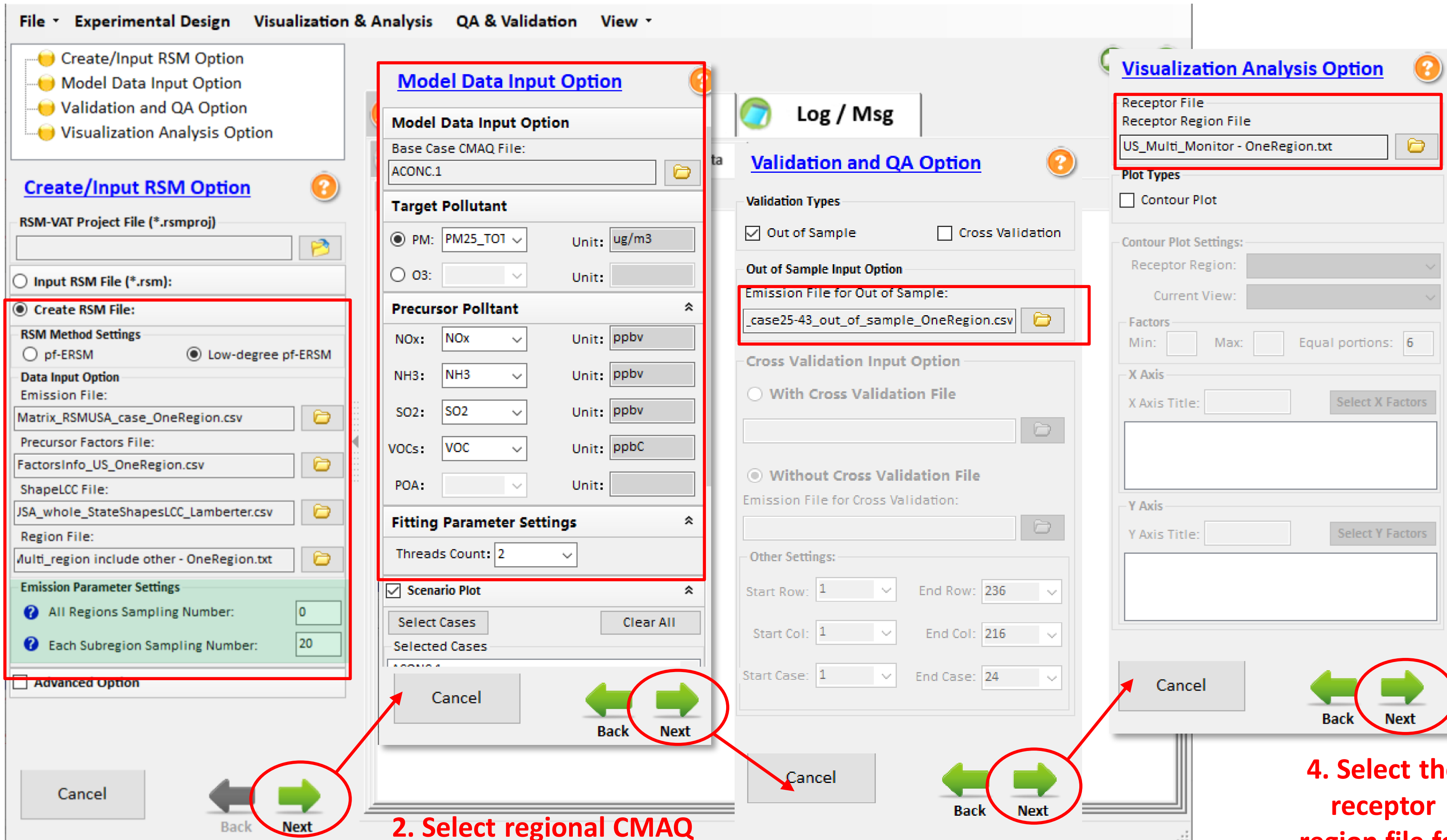


Running RSM Project with Single Region

***Note**: Default RSM case is "5 Climate Regions" in RSM-VAT*

Run RSM Project

➢ **Steps of running RSM project for single region**



1. Select input files for creating Single Region RSM

2. Select regional CMAQ files and pollutants, and set fitting parameters

3. Select the input file for out-of-sample validation

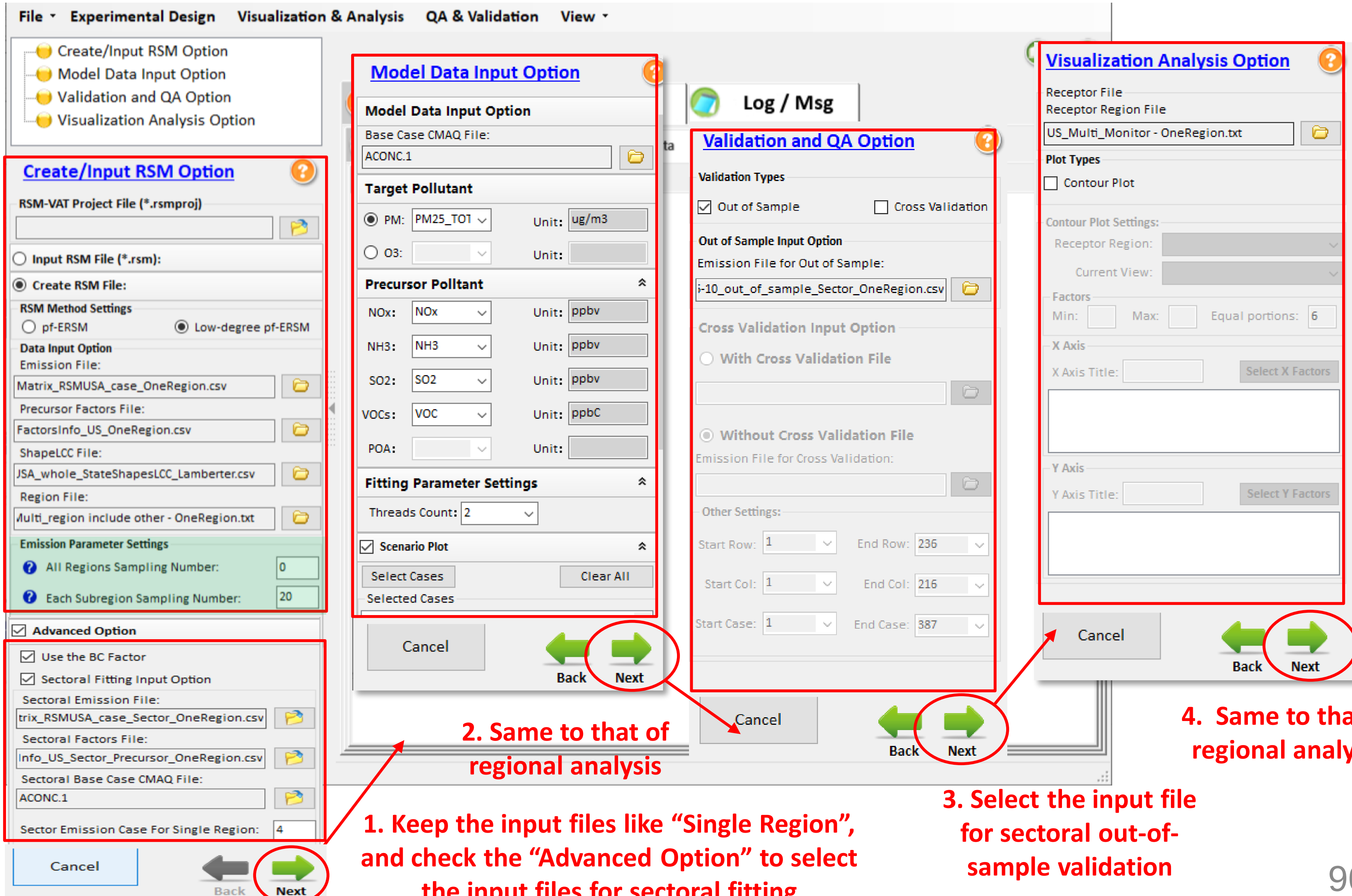
4. Select the receptor region file for visualization analysis

All Regions Sampling Number:0
Each Subregion Sampling Number:20

(C:\Users\...\Documents\My RSM-VAT Files\Data\Example\US_RSM – OneRegion)

Run RSM Project

➢ **Steps of running RSM project for single region with different sectors**



1. Keep the input files like “Single Region”, and check the “Advanced Option” to select the input files for sectoral fitting

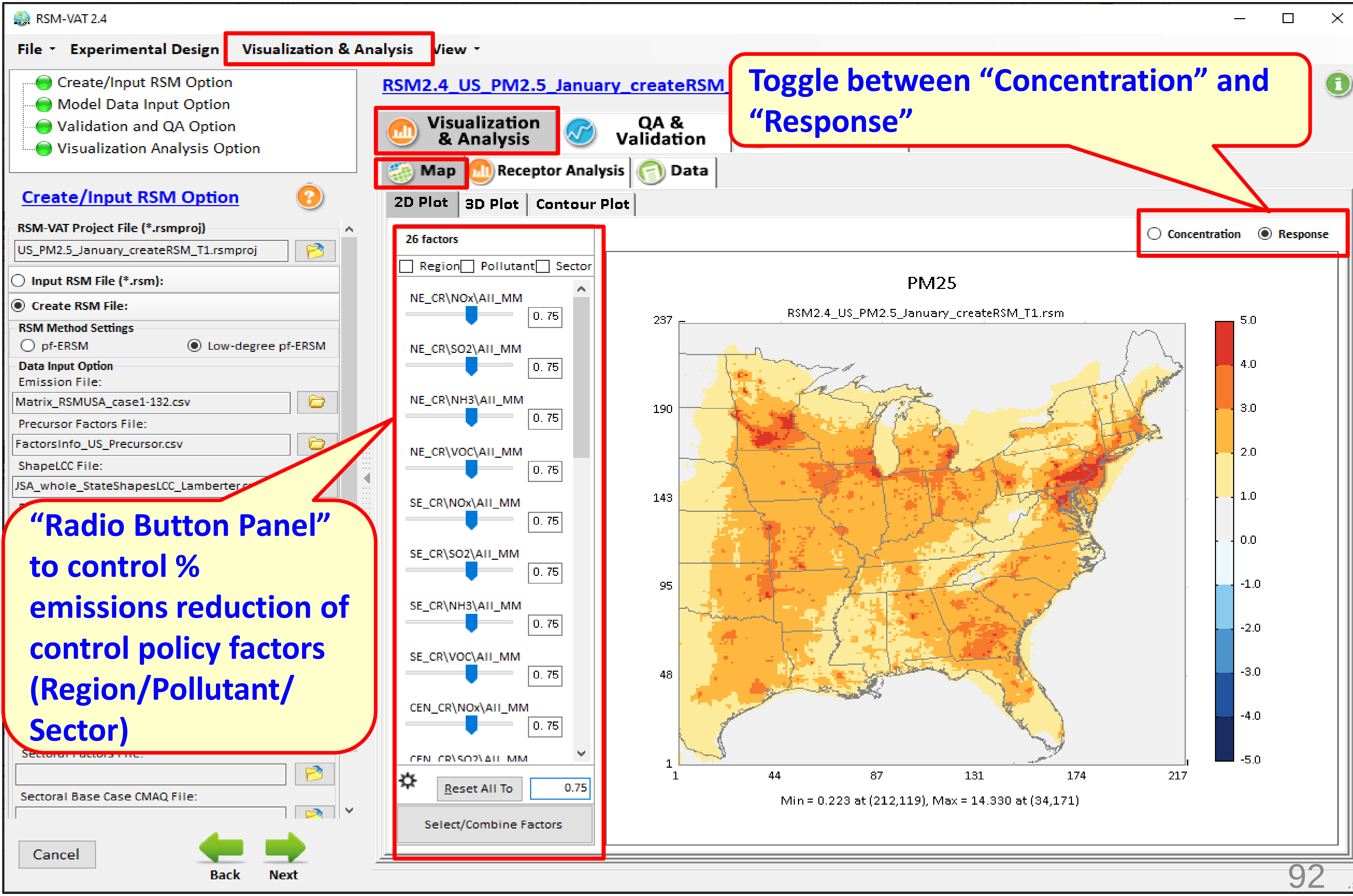
2. Same to that of regional analysis

3. Select the input file for sectoral out-of-sample validation

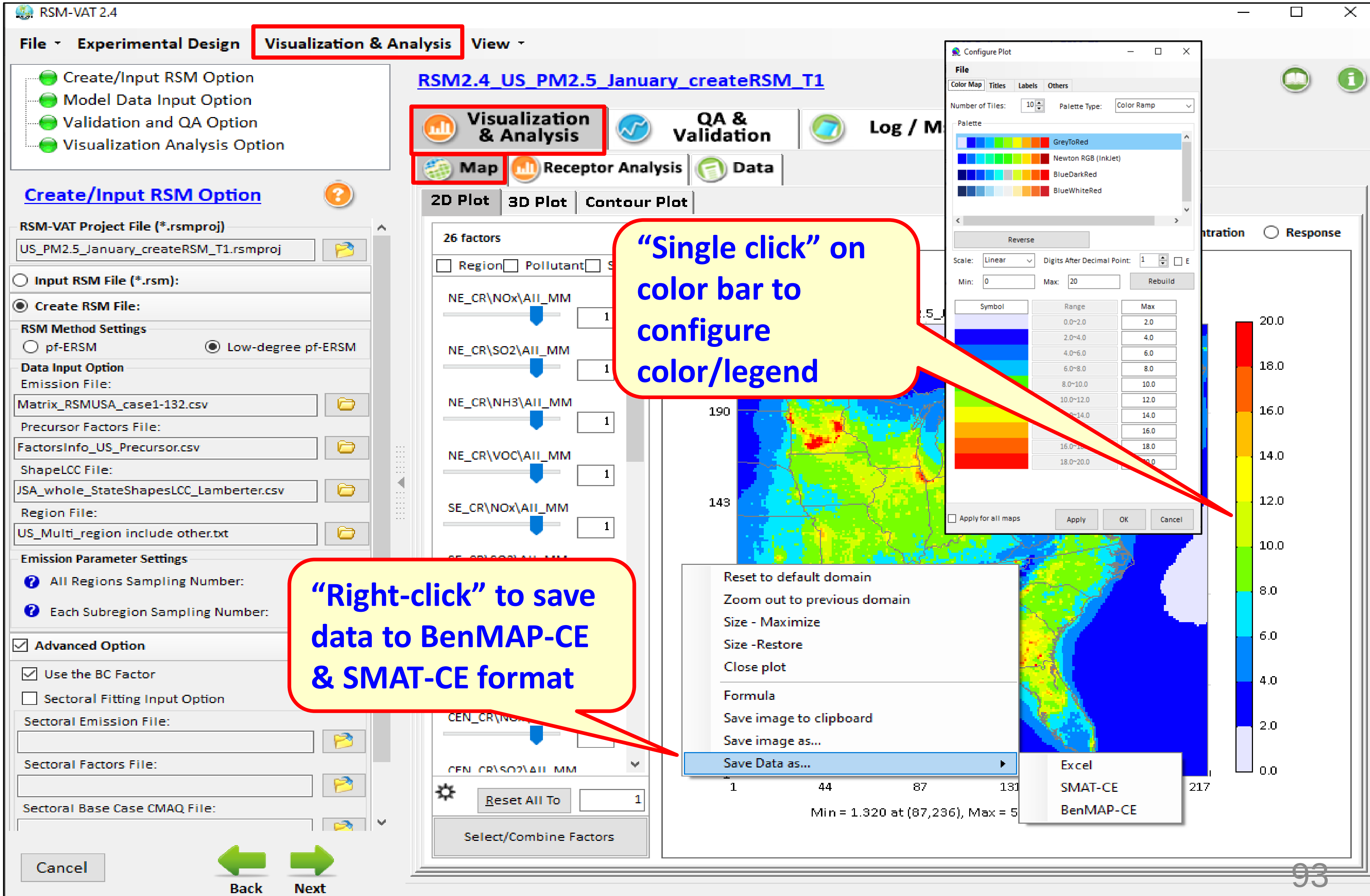
4. Same to that of regional analysis

Visualization and Analysis & QA Module

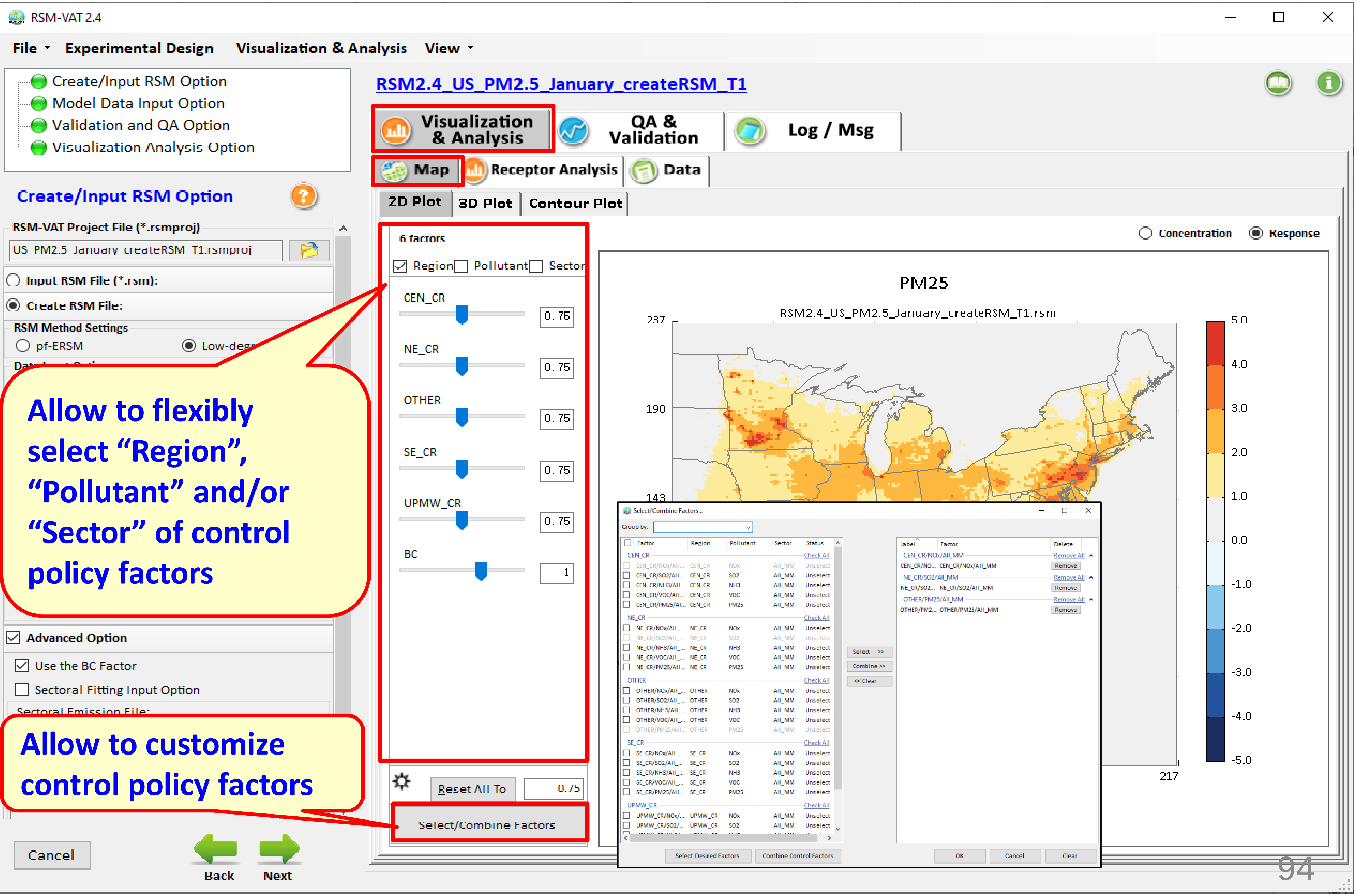
RSM-VAT: Visualization & Analysis Module



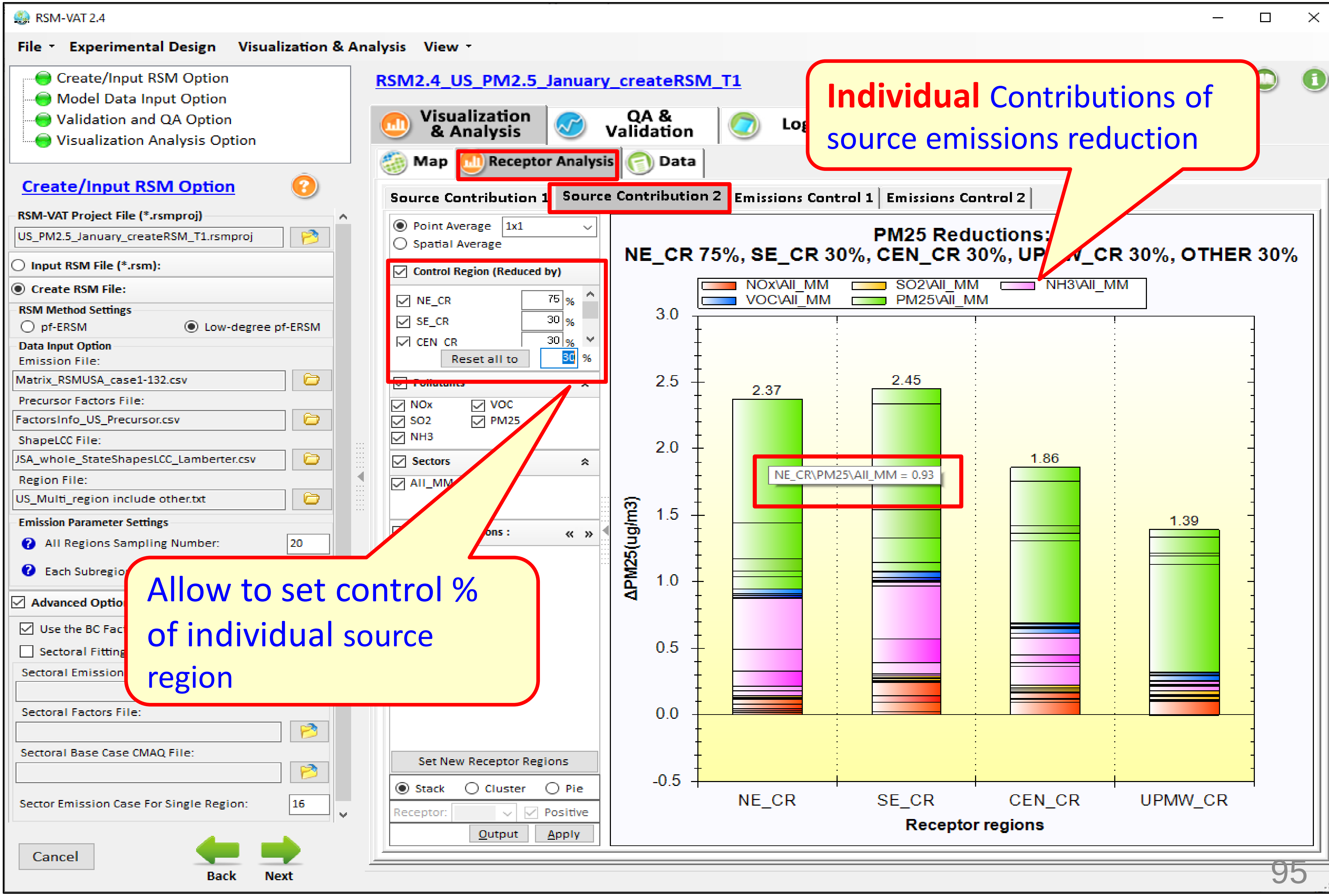
RSM-VAT: Visualization & Analysis Module



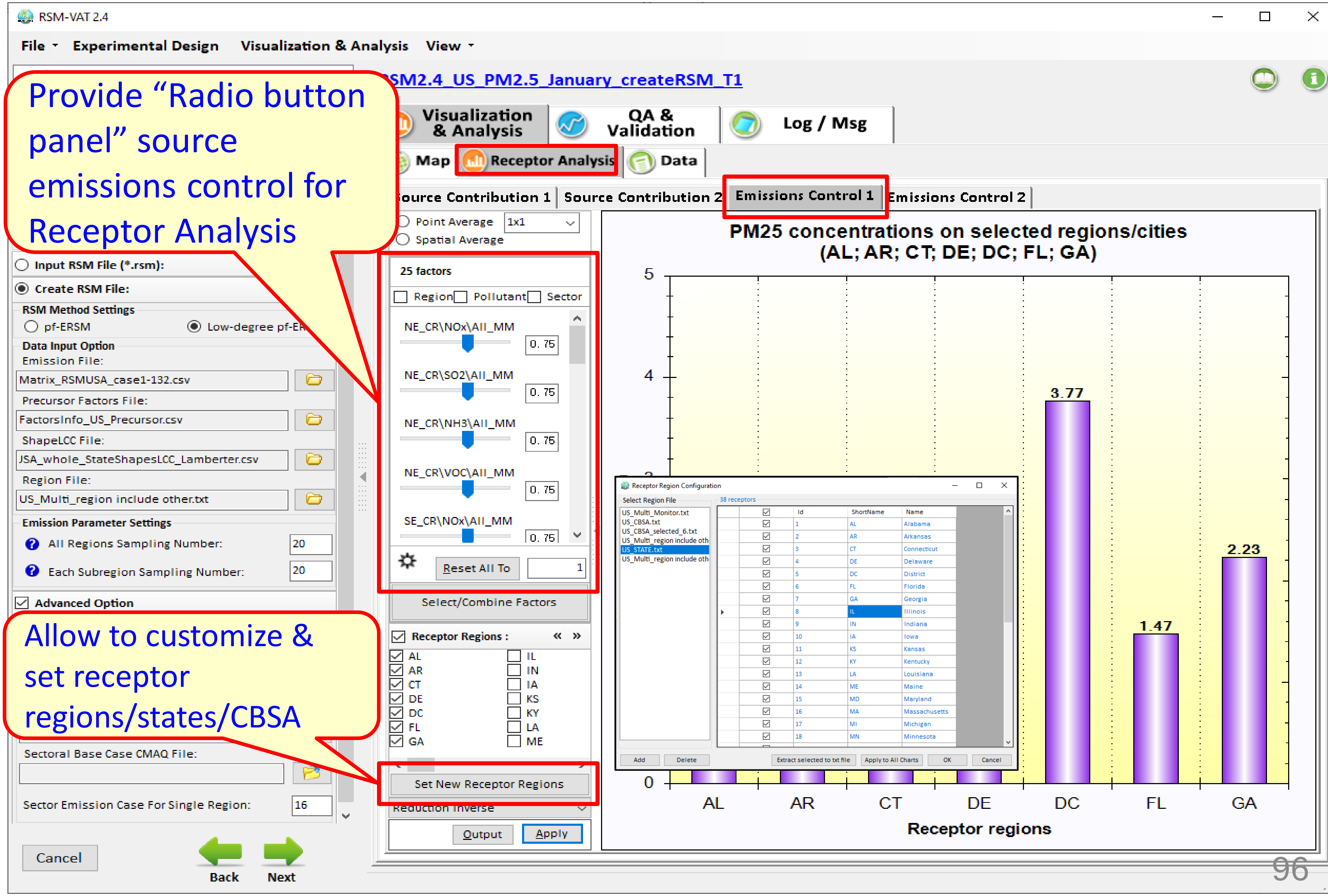
RSM-VAT: Visualization & Analysis Module



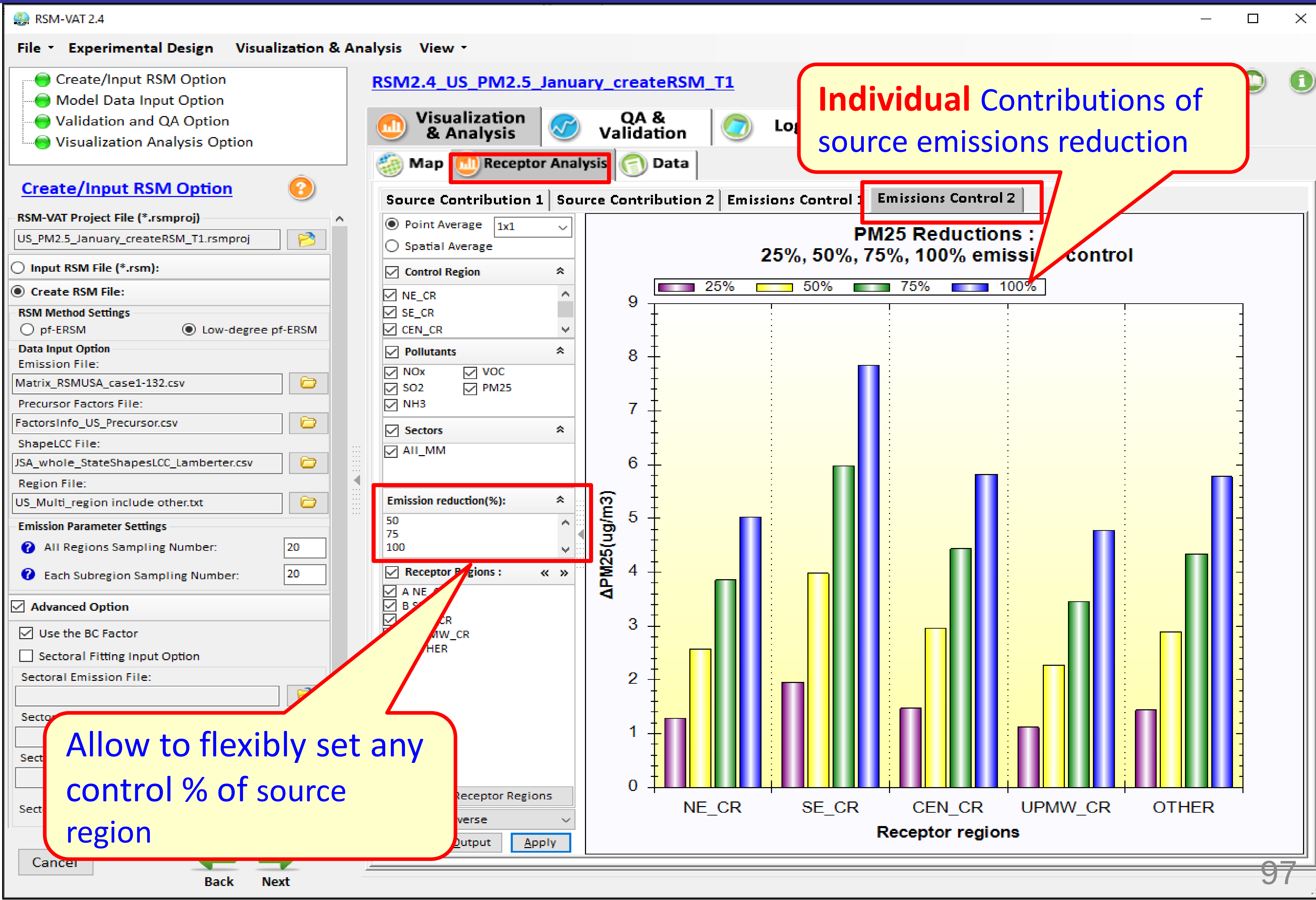
RSM-VAT: Receptor Analysis -> Source Contribution



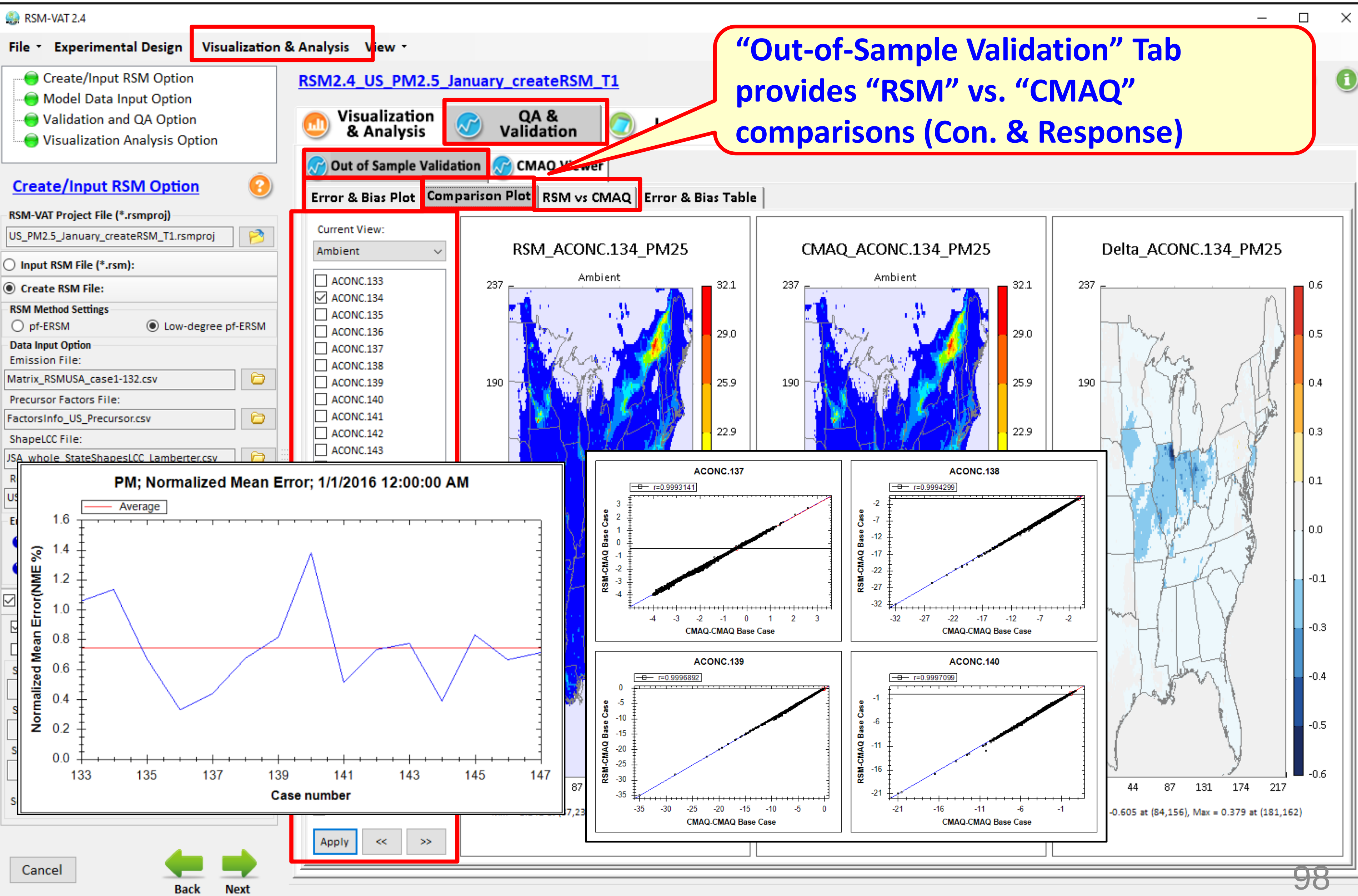
RSM-VAT: Receptor Analysis -> Emissions Control



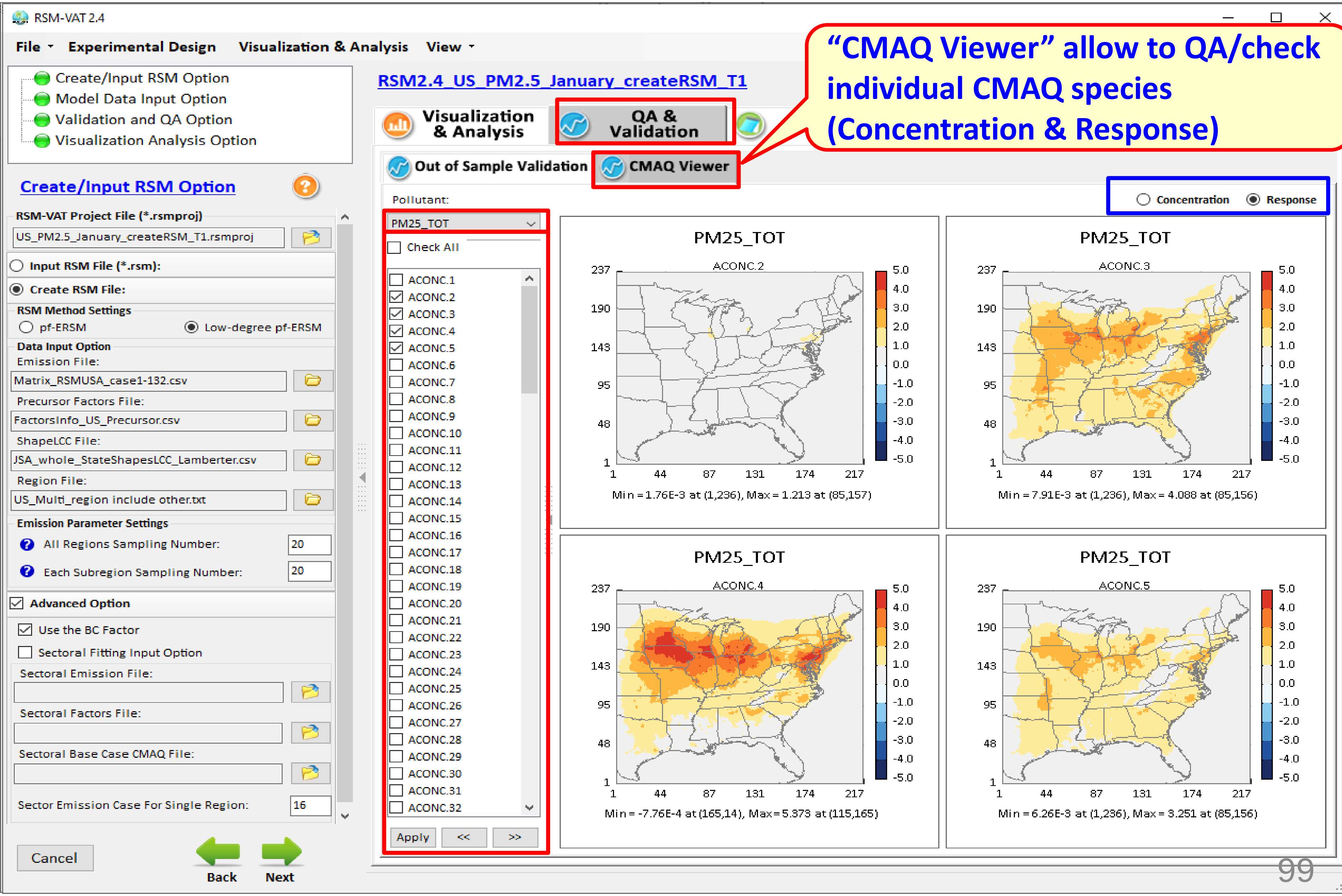
RSM-VAT: Receptor Analysis -> Emissions Control



QA & Validation: “Out-of-Sample” Validation



QA & Validation: "CMAQ Viewer" for QA



End

Carey Jang, U.S. EPA, Jang.carey@epa.gov

UNITED STATES • ENVIRONMENTAL PROTECTION AGENCY

100