Evaluation of LRT Models to Estimate Single Source Impacts on Secondary Pollutants as Part of the IWAQM Phase 3 Process

Bret A. Anderson

USDA Forest Service
Air Resource Management Program
Fort Collins, CO

Western Meteorological, Emissions, and Air Quality Modeling Workshop
Boulder, CO
July 8, 2013

Outline

- Introduction on long range transport (LRT) models and their role in regulatory air modeling
- Background on EPA evaluation program
  - Evaluation paradigm
  - Statistical frameworks
  - Candidate model platforma
- Review of results from European Tracer Experiments (others in progress)

HISTORICAL USEPA LRT MODEL EVALUATION EFFORTS

Evaluation Efforts at a Glance

- 1986 8-model study
- 1990 Rocky Mountain Acid Deposition Model Assessment Program
- 1993 Interagency Workgroup on Air Quality Modeling Phase I Evaluation
- 1998 Interagency Workgroup on Air Quality Modeling Phase II Evaluation

1986 USEPA Short-Term Long Range Transport Model Evaluation Project

- Eight long range transport models evaluated in the project for consideration as Appendix A LRT model:
  - MESOPUFF (Environmental Research and Technology, Inc.)
  - MESOPLUME (Environmental Research and Technology, Inc.)
  - MSPUFF (ND Dept. of Health)
  - MESOPUFF II (Environmental Research and Technology, Inc.)
  - MTDDIS (Rockwell International, Inc.)
  - ARRPA (Tennessee Valley Authority)
  - RADM (Dames and Moore, Inc.)
  - RTM-II (Systems Applications, Inc.)
- Evaluated against two mesoscale tracer experiments:
  - Oklahoma (1980)
  - Savannah River Laboratory Kr^{85}(1976)
- Evaluated using graphical and statistical methods recommended by American Meteorological Society (Fox, 1981)

1986 USEPA Short-Term Long Range Transport Model Evaluation Project (2)

- Models Evaluated Across Multiple Data Organization Strategies
  - Space/Time Paired
  - Paired in Space
  - Unpaired in Time
  - Unpaired in Space/Time, etc.

Rocky Mountain Acid Deposition Model Assessment Project

- Acid Rain Mountain Mesoscale Model (ARM3) (SAI) developed for Western Acid Deposition Task Force.
- Compared against 7 other models from 1986 LRT study
  - Evaluated using same approach as 1986
    - Oklahoma and SRL data sets
    - AMS statistics for various data pairing strategies
    - Model scoring system weighted each tracer experiment and data pairing combination equally. Best performing model in each tracer/data combination awarded four points, three points for second, two for third, etc.
      - MESOPUFF-II performed best for unpaired data combination for each tracer experiment
      - ARM3 performed best for both tracer experiments for space/time data pairing.
      - Final overall scoring: ARM3 – 21, MESOPUFF II - 20
  - Model evaluation approach exposes fundamental issue – need for defining performance objectives according to nature of regulatory applications and defining an objective scoring scheme reflecting these performance objectives.

Interagency Workgroup on Air Quality Modeling (IWAQM)

- Phase I Evaluation – “off-the-shelf” models, ARM3 and MESOPUFF-II evaluated, coding errors discovered in ARM3, leaving MESOPUFF II only model available
- Phase II Evaluation – CALPUFF/CALMET and CITPUFF/NUATMOS evaluated.
  - Trajectory evaluation using CALMET and NUATMOS using observations and “hybrid” fields based upon observation blending with 80-km MM4 data.
  - Statistical evaluation using ASTM/Irwin methods for evaluation (Oklahoma and SRL datasets).
    - CALMET/MM4 combination produced more accurate trajectory statistics than NUATMOS/MM4 combination

Lessons Learned from Prior Evaluation Efforts

- No USEPA recommended methodology for evaluation of air quality models. No consistent approach between efforts in 1980's and 1990's.
- Evaluation methodology used all published AMS metrics and data organizational strategies. This did not take into consideration regulatory use of LRT models, weighting schemes not most appropriate for particular methods LRT models are used for.
- High sensitivity of LRT models to meteorological inputs. Need for more objective meteorological performance evaluation measures.
- No data sets available to evaluation chemical transformation mechanisms of LRT models

EXISTING USEPA MODEL EVALUATION GUIDANCE

Relevant Model Evaluation Guidance

- Fox, D.G., 1981: Judging Air Quality Model Performance: a summary of the AMS workshop on dispersion model performance, Woods Hole, MA, 8 – 11 September 1980. *Bull. Am. Met. Soc..*, **62**, 599-609.
- Interim Procedures for Evaluating Air Quality Models (Revised) (EPA-450/4-84-023)
- Protocol for Determining the Best Performing Model (EPA-454/R-92-025)
- Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, PM2.5, and Regional Haze (EPA-454/B-07-002)

Paradigm Question

- Carhart et al. (1989) noted that there was no recommended USEPA modeling guidance specific LRT models. 1986 EPA study applied all model metrics across all data organization strategies in the evaluation of the 8 models included in that study.
  - Are LRT models dispersion models or chemistry models?
    - Concerned about peak values
    - Concerned about accuracy of transport
    - Concerned about transformation and removal mechanisms
- Should they be evaluated in the same manner as near-field dispersion models, photochemical grid models, or some other way?

1984 USEPA Interim Evaluation Procedures Document

- The USEPA interim procedures document suggests:
  - definition of performance objectives relative to the nature of the regulatory applications of the model;
  - compilation of data sets and performance measures that will be used for each performance objective;
  - objective scheme for assigning weights to each performance measure and data set combination; and
  - an objective scheme for scoring the performance of any models relative to one another.

Defining Performance Objectives – Start with Regulatory Use of Model

- Current Regulatory Uses:
  - PSD Class I NAAQS and increment analyses
  - Visibility and deposition for Air Quality Related Values analysis for PSD
- Future Uses:
  - Single source O_3 NAAQS analyses
  - Single source $PM_{2.5}$ NAAQS analyses

Regulatory Niche for LRT Models

- Section 165(d) of the Clean Air Act requires suspected adverse impacts on federally protected Class I areas be determined under the federal major new source review program called Prevention of Significant Deterioration of Air Quality (PSD) program
- Many Class I areas are located areas are located more than 50 km from source under review.
- EPA near field regulatory models (ISC, AERMOD, etc.) not applicable beyond 50 km because steady-state wind field assumption not applicable beyond these distances
- LRT models used to assess PSD increment, visibility impacts from secondary aerosols, and acid deposition in federally protected Class I areas

Regulatory Background

- Interagency Workgroup on Air Quality Models (IWAQM) in 1991 in response to emerging need to assess air pollutant impacts in federal Class I areas.
- In 1998, EPA published IWAQM Phase 2 report recommending CALPUFF for regulatory LRT model applications. Phase 2 report provided recommended settings for CALPUFF model control options.
- In 2003, EPA promulgated the CALPUFF modeling system as its “preferred” model for LRT model applications. IWAQM Phase 2 report becomes de-facto “recommendations for regulatory use” for regulatory CALPUFF applications.

Regulatory Background

- In 2008-2009, EPA, US Fish and Wildlife Service, National Park Service, and US Forest Service updated Phase 2 guidance. LRT model evaluation program initiated by EPA as part of IWAQM effort.
- IWAQM Phase 3 – evaluation of possible model platforms for development/adaptation for single source, full photochemistry model applications

IWAQM Phase 3 LRT Model Evaluation Project Goals

- Develop meteorological, inert tracer, and chemistry databases for evaluation of long range transport models.
- Develop a consistent and objective method for evaluating long range transport (LRT) models used by the EPA and FLM's.
- Promote the best scientific application of models based upon lessons learned from evaluations and reflect this in EPA modeling guidance.
- Evaluate new models as part of IWAQM Phase 3 process.

Initial EPA Efforts for IWAQM Phase 3

- EPA WA 4-06: Single-Source LRT Model Evaluation (7 Tasks)
  - Task 1: Work Plan
  - Task 2: Update and Finalize MMIF
  - Task 3: CALPUFF Met-Input NetCDF for Visualization
  - Task 4: Document CALPUFF Tracer Test Evaluations
  - Task 5: Critical Review of Single-Source LRT Models
  - Task 6: Proof of Concept for using PGMs for Single-Source AQ/AQRV Assessments and Compare with CALPUFF
  - Task 7: Evaluate LRT Models using Tracer Test and Plume Chemistry Field Experiments

Task 2: Update MMIF

- Subtask 2a: Enhance MMIF to support AERMOD and SCIPUFF/SCICHEM
  - For AERMOD, two options to output data for user selected location (MM5/WRF grid cell):
    - Output surface and upper-air meteorological data to run through AERMET; and
    - Output meteorological data in format to directly input into AERMOD
  - SCIPUFF/SCICHEM MEDOC format for more evaluation of this modeling system

Task 2: Update MMIF

- Subtask 2b: MMIF enhancements to support CALPUFF
  - Add option to extract MM5/WRF data based on Latitude/Longitude window rather than just (i,j) offset
  - Add support for Polar Stereographic and Mercator projections
  - Add in option to use WRF/MM5 LAI, L and z0 estimates rather than lookup table approach (i.e., pass through)
  - MMIF testing and evaluation
  - Update MMIF user's guide
    - Full documentation of MMIF critically important

Task 3: Enhance CALMET-to-NetCDF Software

- USFS has developed prototype CALMET binary output files to NetCDF/IOAPI format that can be visualized by the public available PAVE/VERDI software
  - This task would enhance this software to operate on multiple platforms and compilers and develop documentation
  - Will extend to VERDI visualization tool since PAVE no longer supported (but still useful)
  - Will allow visualization of CALPUFF meteorological inputs without expensive third part software
    - Compliments statistical evaluation using MMIFStat

Task 4: Report on CALPUFF and Other LRT Model Tracer Experiment Evaluation

- EPA/USFS performed LRT model evaluation against tracer test data for multiple LRT models including CALPUFF (CALMET & MMIF), HYSPLIT, FLEXPART, SCIPUFF (and some CAMx)
  - 1994 European Tracer Experiment (ETEX)
  - 1983 Cross-Appalachian Tracer Experiment (CAPTEX)
  - 1980 Great Plains Tracer Experiment (GR80)

Task 5: LRT Model Review (UNC)

- Review of LRT Modeling Techniques
  - Lagrangian Puff
  - Particle
  - Chemical plume model
  - Photochemical grid/hybrid model
- Literature review and assessment
  - Capability for primary and secondary pollutants
  - Data requirements
  - Publicly availability
  - Advantages and disadvantages
  - Evaluation studies

Task 6: Single-Source PGM Demonstration

- Comparison of Single-Source Estimation Techniques for Ozone and AQ/AQRV
  - Two modeling test domains
    - 2006 12 km MM5 database covering eastern Utah and western Colorado (UT-CO)
    - 2005 Four Corners Air Quality Task Force (FCAQTF) 12/4 km domain
  - Five model configurations:
    - CAMx using APCA ozone and PSAT PM source apportionment (ENVIRON/EPA)
    - CALPUFF v5.8 using MMIF (ENVIRON/EPA)
    - CALPUFF v5.8 using CALMET and August 2009 application procedures (ENVIRON/EPA)
    - CALPUFF v6.42 using MMIF
    - CALPUFF v6.42 using CALMET

Task 6: Single-Source PGM Demonstration

- Evaluation for ozone and far-field AQ and AQRVs
  - Like one would do for a PSD or NEPA application
  - Use existing and hypothetical sources
  - PSD and other (e.g., SO4 and NO3) pollutant concentrations
  - Visibility (FLAG 2010 procedures)
  - Deposition (S and full and subsets of N species)
- Optional Tasks (not funded)
  - 6-1: SCICHEM (USFS – Summer 2013)
  - 6-2: HYSPLIT
  - 6-3: Brute Force CAMx Zero-Out
  - 6-4: Brute Force CMAQ Zero-Out

Task 7: Evaluate Single-Source LRT using Field Experiments

- Subtask 7a: Evaluate CAMx using three classic inert tracer experiments and compare with CALPUFF and other LRT models
  - ETEX, CAPTEX and GP80
  - Compare with CALPUFF/MMIF and CALPUFF/CALMET
  - HYSPLIT, FLEXPART, SCIPUFF/SCICHEM, CALGRID

Task 7: Evaluate Single-Source LRT using Field Experiments

- Subtask 7b: Evaluate Against Atmospheric Chemical Plume Observations
  - TVA Cumberland Plume 1999 SOS
  - 2002 CCOS Aircraft Measurements
  - Three models:
    - CAMx, CALPUFF/MMIF and CALPUFF/CALMET
  - Optional Tasks
    - 7b-1: SCICHEM
    - 7b-2: HYSPLIT
    - 7b-3: CALPUFF v6.42

Testing Advection and Diffusion of Models

MESOSCALE TRACER SIMULATIONS

Evaluation Framework

- Evaluation of LRT models within their defined regulatory niche requires an evaluation of three independent components of the AQ model system
  - Meteorological component
  - Advection and diffusion component
  - Chemical transformation

Evaluation Paradigm

- Evaluation procedures follow logic of Chang et al (2003) regarding multi-model evaluations
  - Inherent amount of uncertainty due to differences in technical formulations between various modeling systems
  - Use common meteorological platform with minimal diagnostic adjustments to reduce uncertainty
    - This is a challenge when models such as SCIPUFF and CALPUFF use diagnostic wind models as primary source of 3-D meteorological data
      - Use MM5SCIPUFF developed by Penn State and MMIF (CALPUFF) developed by EPA to couple MM5 directly to these models
  - Model control options mostly default "out-of-the-box" configuration
    - CALPUFF configured for turbulence dispersion and puff-splitting similar to SCIPUFF, which is a deviation from its default configuration

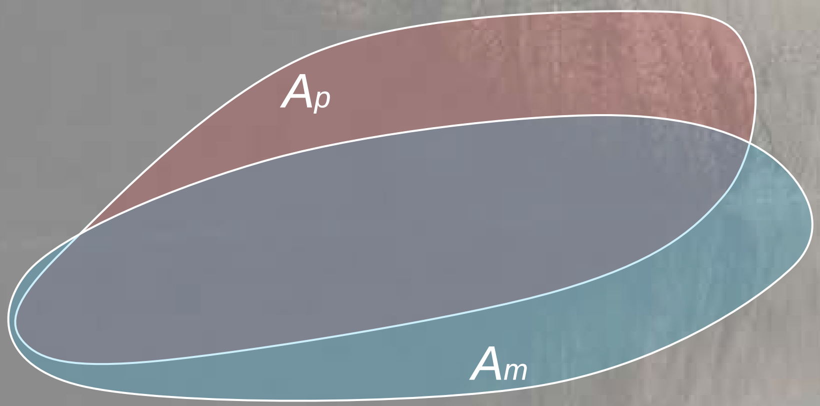
Data Organizational Issues

- The 1984 Interim Evaluation Procedures document identified a key concept in developing an evaluation protocol – data organization.
- Compilation of data sets and performance measures that will be used for each performance objective; each performance objective is tied to regulatory use of a model.
-

Performance Metrics

- Performance metrics focus upon models' ability to predict

Performance Metrics



$$FMS = \frac{A_M \cap A_P}{A_M \cup A_P}$$

- Data organization and performance metrics focus upon model's ability to predict paired in both space and time. Considered most stringent method, but most consistent with LRT regulatory applications.
- Example:
  - FMS is defined as the ratio between the intersection of measured (***Am***) and predicted areas (***Ap***) above a significant concentration level and their union. FMS expressed as a percentage corresponding to the degree of overlap.
  - The more that the predicted and measured tracer clouds overlap one another, the greater the FMS values are.

Models Evaluated for Inert Tracers

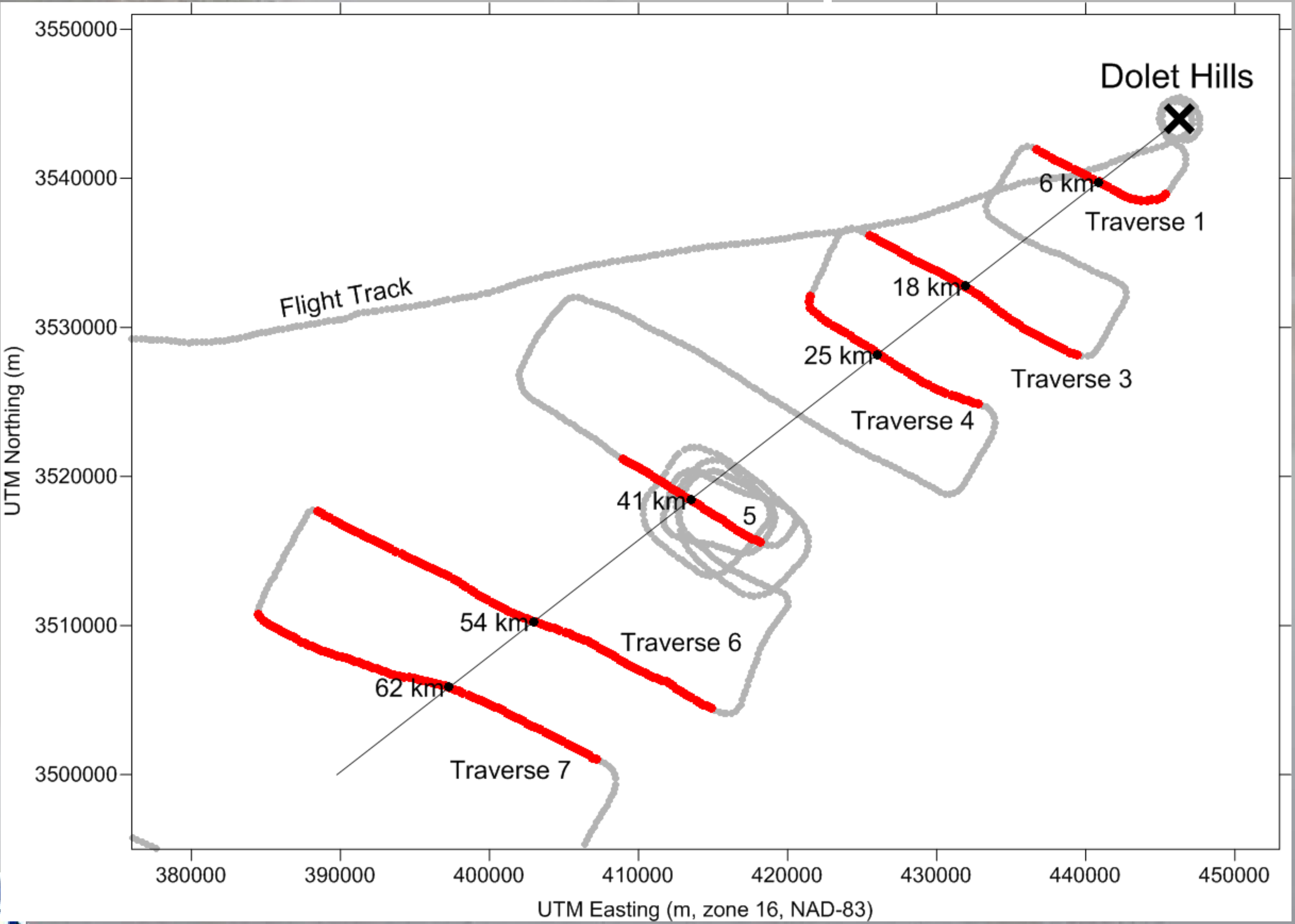
- Three Distinct Class of Models
  - Lagrangian Puff Models
  - Lagrangian Particle Models
  - Eulerian Grid Models

- CALPUFF Version 5.8 (EPA approved version)
- MM5-FLEXPART (Version 6.2)
- HYSPLIT (Version 4.8)
- SCIPUFF (Version 2.303)
- CAMx (Version 5.20)
- CALGRID (Version 2.4)

Single Source Chemistry Evaluations

- Application of LRT models for chemistry usually only involve an individual or small group of sources.
- Traditional photochemical grid model (PGM) evaluation techniques (chemistry evalaution) combined with inert tracer evaluation (advection and diffusion) are combined to examine the suitability of a model for use in single source chemistry applications.
  - The best performing chemistry model will only be as good as its ability to treat advection and diffusion appropriately.

Example: Aircraft Traverses and Model Receptors



Source: Vijayaraghavan, et al., 2010; "*Evaluation of an Advanced Reactive Puff Model using Aircraft-based Plume Measurements*"