

Model Evaluation

Report of Model Evaluation Workgroup

R/S/L 2005

Clint Bowman

Model Evaluation Workgroup

- R/S/L modelers
- Northwest modelers
 - NW-Airquest (realtime air quality modeling)
 - NW Regional Modeling Consortium (MM5)
 - ClearSky (ag burning)
 - BlueSky-RAINS (silvacultural burning)
- EPA RTP

Model Evaluation

- Metrics
- Effects of Error
- Limits to Accuracy
- Accuracy \Leftrightarrow Utility
- Representativeness

Model Evaluation (cont)

- Episodes???
- Communication of Uncertainty
- Standards
- Acceptance

Metrics

- What metric(s) of the air quality modeling needs to be evaluated?
 - cumulative distribution of concentrations (permits)
 - correspondence in time and space (episodes)
 - accuracy of diurnal, seasonal, decadal behavior

Effects of Error

- How does the accuracy of the meteorological fields affect the accuracy of the air quality model?
 - need to be able to determine the accuracy of the meteorological fields
 - similarly for the concentration fields

Limits to Accuracy

- Are there limits to the expected accuracy of various types of models imposed by:
 - Sub-grid scale processes
 - Unmodeled physical processes
 - Unmodeled chemical processes
 - Unknown inventory
 - Unknown activity
 - Unknown geophysical parameters

More Limits to Accuracy

- Random fluctuations
- Errors in initial or boundary conditions

Accuracy \Leftrightarrow Utility

- How accurate does an estimate need be to provide adequate information to be useful?

A Question of Representativeness

- How representative are the observations?
 - How can we determine and document?
 - meta-data
 - Establish repositories (observations and predictions)

Episode vs Continuous Evaluation

- Episode Evaluation
 - Appropriate metrics
 - Confidence in results
- Long-term Evaluation
 - Similar concerns
- Domain Size
 - Evaluation points

Communication of Uncertainty

- Sources of Uncertainty
- Levels (sizeof) Uncertainty
- Performance Indicators
- Decision Making

The Role of Standardized Evaluation

- ASTM
 - Standard datasets
- Metrics
- Revision
- Updates

Procedures for Model Acceptance

- Define testing methodology
 - Test dataset
- Define acceptable level of performance