Uncertainty of Transport and Dispersion Models

Steven R. Hanna
Harvard School of Public Health
shanna@hsph.harvard.edu

OFCM Session on Model Uncertainty
GMU, Fairfax, VA
14 July 2010

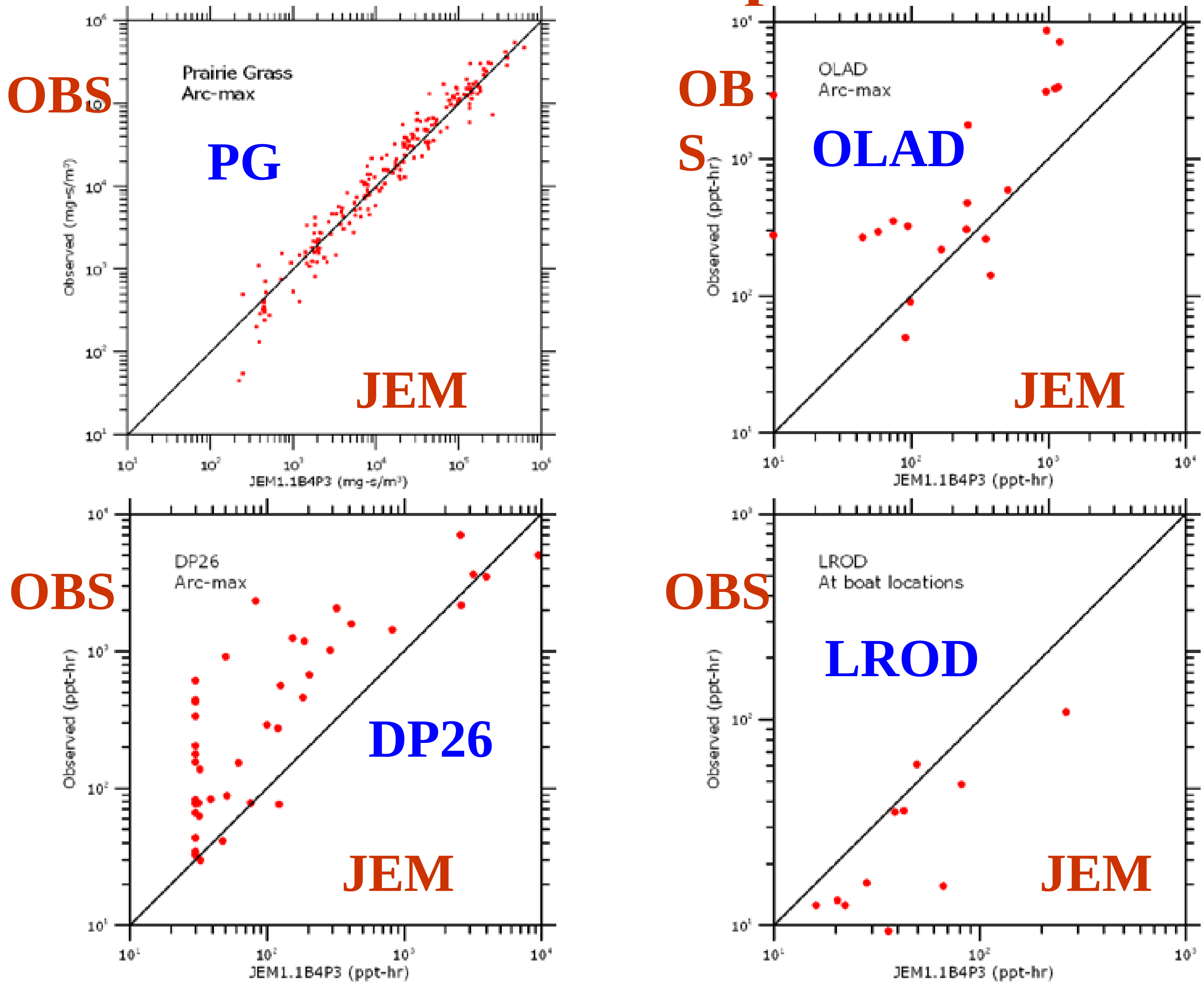
What is a model and its outputs?

- The model can be a single dispersion model (e.g., AERMOD or SCIPUFF) or a linked model system (e.g., linked with a met model such as WRF)
- Most models predict the “ensemble mean” of a concentration averaged in time and/or space. The averaging method must be prescribed (i.e. define the “output”).
- Some models (e.g., SCIPUFF) also predict the variance of the concentration, as well as an assumed pdf.
- Uncertainty is due to model errors, instrument uncertainty, and random turbulent uncertainty

Definition of Uncertainty

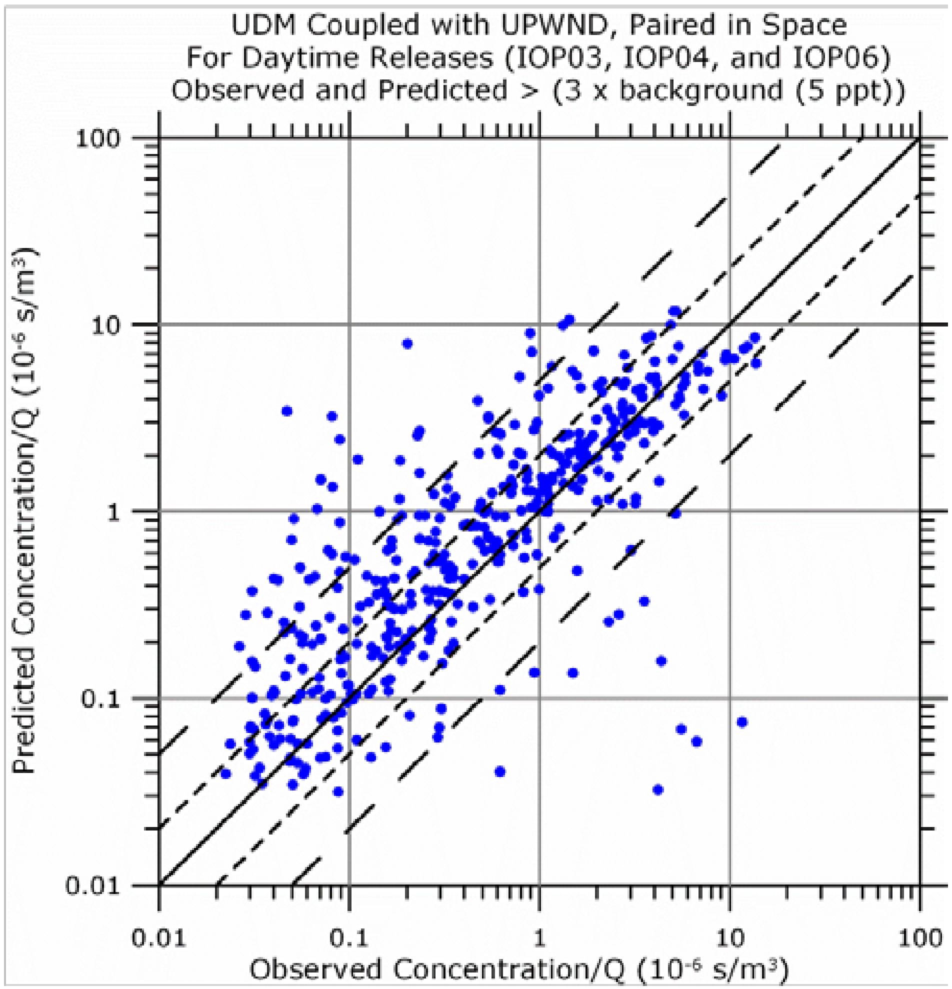
- The word “uncertainty” is best defined in statistical terms, e.g. the root mean square of the difference between model predictions and observations
- “Uncertainty” could be the pdf that is directly predicted by the model (e.g., SCIPUFF)
- Explain (in clear terms) to the decision-maker.

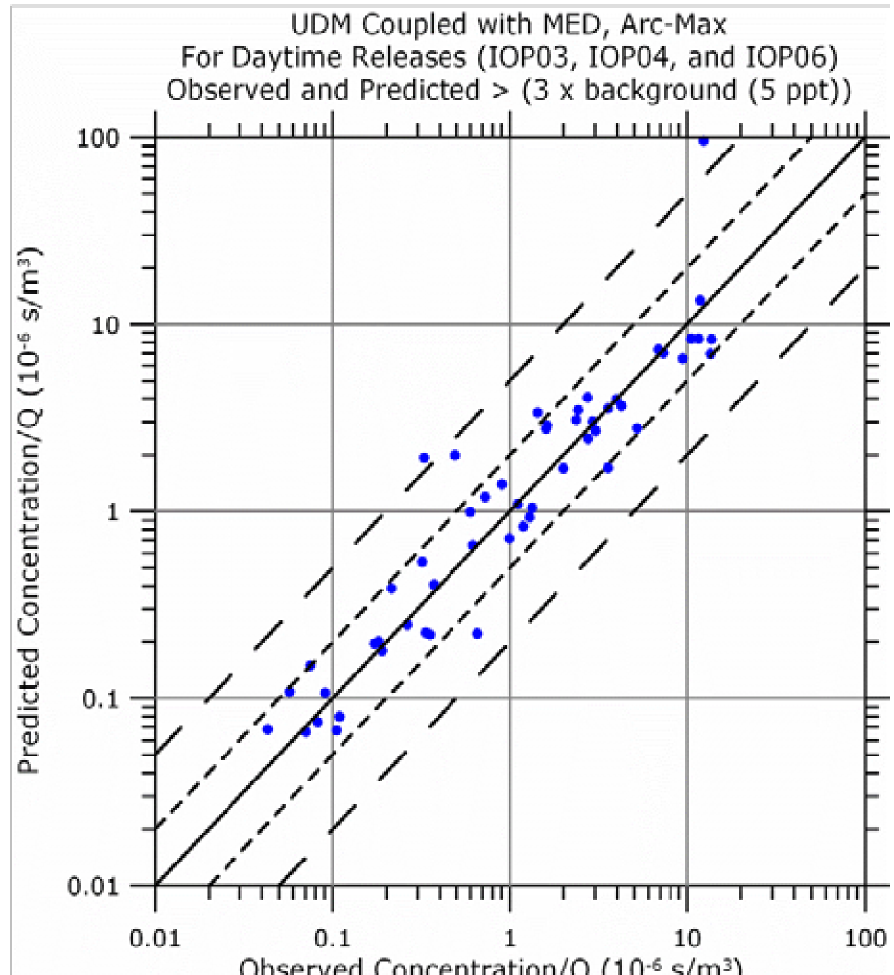
Rural PG, DP26, OLAD, & LROD Scatter Plots for JEM Dispersion Model



Note different scales

Differences in scatter plots for pred vs obs for paired in space (left) and arc-max (right)
For HPAC/UDM for JU2003 field obs (three daytime IOPs)





Rules of Thumb Based on Experience and Field Experiments

- Experts' experiences suggest "factor of two" uncertainty in dispersion model predictions in best scenarios. This was suggested by Frank Pasquill 40 years ago and is still valid.
- Uncertainty increases to "factor of 5 or 10" for poorly defined scenarios or complex terrain and/or met conditions
- In-plume σ_C/C is about unity on the plume centerline for one-hour sampling times and is larger (factor of 5 to 10) on plume edges.
- The model uncertainty is no better than the uncertainty of its inputs.

Sources of Model Uncertainty

- Natural stochastic (turbulent) variations
- Errors or non-representativeness of input data (e.g., wind speed observations by anemometers and by radiosondes)
- Physics assumptions in the model technical document are incorrect or inadequate or inappropriate for the application
- Model parameters (e.g., scaling constants) are uncertain
- Coding/software errors
- The users guide is unclear about which input data to use and what switches to set, causing different users to get different results
- The model is best suited (tuned) for certain simple scenarios where field data were available. “Gaps” in knowledge may exist for complex scenarios.

Dependence on Scenario

Figure from NOAA/Hazmat – El Cajon train derailment

Two Approaches to Predicting Uncertainty

- **Internal -** The uncertainty (i.e., the pdf of C) is internally predicted by the model (e.g., HPAC/SCIPUFF), which includes formulas for internal plume fluctuations and meandering.

- **External -** The model does not directly predict the uncertainty. Instead the uncertainty is assumed to be caused by variations in inputs and model parameters and is estimated separately, through multiple model runs (ensembles), sensitivity studies, etc.

Internal - HPAC/SCIPUFF accounts for small-scale concentration variations, plus uncertainties in observed or modeled meteorology

- SCIPUFF automatically solves for the small-scale concentration variance and produces a PDF (e.g., 95 % range) of concentrations
- Large scale velocity variance and distance scale (Distance scale now has default of 100 km) for mesoscale wind variations
- SCIPUFF could ultimately use mesoscale met model variance and distance scale predictions

•

External methods for estimating uncertainty, ordered by complexity

- Full Monte Carlo probabilistic method (allows all inputs and model parameters to be simultaneously varied and correlations determined, but takes a lot of time, and may produce too much uncertainty)
- Ensemble method (a subset of the MC method with a few model runs hopefully sufficient to capture "spread")
- One-at-a-time (OAT) sensitivity studies (not good for nonlinear systems)

JEM IV&V Acceptance Criteria for Arc-Max Concentrations

	Rural	Urban
|FB|	< ~30%	< ~67%
NMSE	< ~3	< ~6
FAC2	> ~50%	> ~30%

- Based on the experience gained from many previous model evaluation exercises primarily using *rural research-grade* field data
- Urban criteria relaxed roughly by a factor of 2 compared to rural
  - FB indicates relative systematic bias
  - NMSE indicates relative random scatter

Performance Measures for Four Rural (i.e., Non-Urban) Field Data Sets

	FB	NMSE	FAC2
PG	0.124	0.34	0.886
DP26	0.407	2.33	0.357
OLAD	1.225	8.05	0.333
LROD	-0.570	1.27	0.727

- **GREEN** means rural acceptance criteria are met, i.e., |FB| < 30%, NMSE < 3, and FAC2 > 50%
- **RED** means acceptance criteria are not met

Ensemble Method for Estimating Uncertainties

- Widely used in weather forecasting; has had some tests with air quality models
- Make multiple (10 or 20 or 30) runs of the model system using different models, different initial fields, different model algorithms, etc. The median of the predictions is the best "forecast".
- It is hoped that the "spread" of the predictions is a measure of the uncertainty. This is checked with observations over some time period and domain.

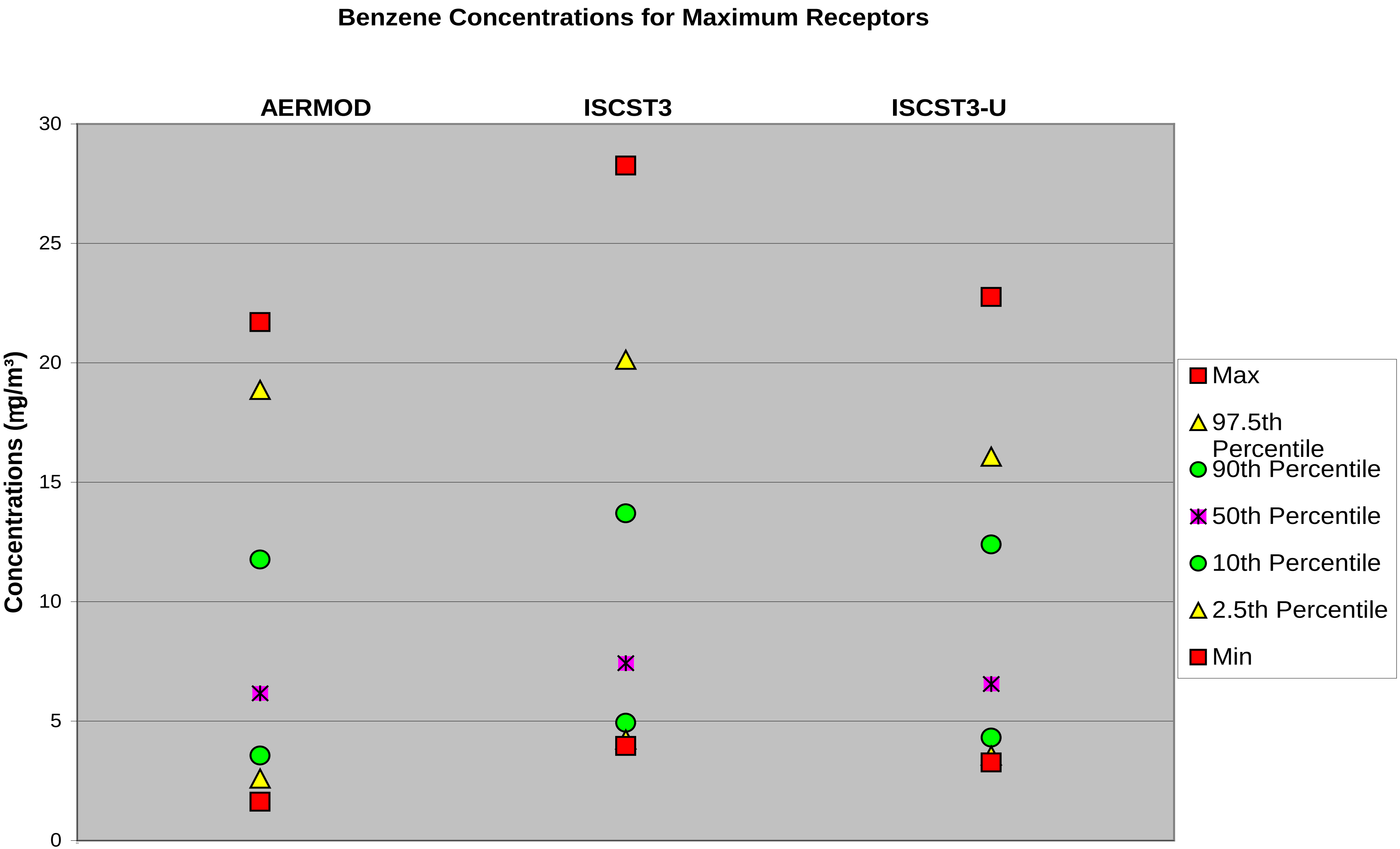
Monte Carlo Uncertainty Estimates

Assume range of variations of important inputs. Apply Monte Carlo (MC) method by resampling independently from all inputs for each MC run. Do > 100 runs. Determine statistics (e.g., ranges, means, etc.)

Gives estimates of **total uncertainty** in model outputs due to uncertainties in inputs and in model parameters. Gives estimates of **correlations** between variations in model outputs and variations in individual model inputs, allowing the inputs to be identified where uncertainties have the largest effect on the model output uncertainties.

Dispersion Model Inputs Varied in MC Houston Study with AERMOD

- Wind speed u and v components (1 m/s)
- Cloud cover (0.2)
- Daytime mixing depth (20 %)
- dT/dz (factor of two)
- σ_y and σ_z (factor of two)
- Surface roughness (AERMOD only) (factor of three)
- Bowen ratio (AERMOD only) (factor of two)



Example of Uncertainties from 100 MC Runs

Communication of Uncertainty to Decision-Makers

- Must involve decision-makers in the process (ask them what they need)
- Must answer specific questions
- Must be clear and succinct (on level of 30 second blurb on CNN)