# PC-Based Supercomputing for Uncertainty and Sensitivity Analysis of Models

Emerging Technologies

Justin Babendreier, Rajbir Parmar, Kurt Wolfe (USEPA); Simone Uter, Monica McKendrick (NAFEO)
Office of Research and Development, National Exposure Research Laboratory, Ecosystems Research Division, Athens, Georgia

## Uncertainty Analysis

*…describing potential differences between model predictions and nature.*

**Uncertainty (UA):** *due to lack of knowledge and data.*
**Analysis→** *given uncertainty in both models and their inputs, quantify/qualify uncertainty in model output(s).*

**Unified Uncertainty Theory: $U ?=? P + M^2C$**

**Parameter** input uncertainty & variability

**Model** models in same group (e.g., air models) give different predictions

**Conceptual** for a given model, can implement reality differently

**Modeler** manifests as model selection and implementation; people use the same models differently


## Sensitivity Analysis & Parameter Estimation

*…discovering relationships between model predictions and unit changes in input variables.*

**Sensitivity (SA):** *finding the subset of input variables that are most responsible for variation in model output.*
**Analysis→** *relate importance of uncertainty in inputs to uncertainty in model output(s).*

**Parameter Estimation (PE):** *use measured output(s) to back-calculate best estimates of (some) model inputs.*

### Input Space Assessment Techniques

**Local** works intensely around a specific set of input values (i.e., the local condition)

**Screening** quick and simplistic, ranks input variables and ignores interactions between variables

**Global** quantifies scale & shape of the I/O relationship; all input ranges; assesses parameter interaction

**Key Words:** Model, Uncertainty, Sensitivity, Parameter Estimation, Multimedia, Supercomputing

## UA/SA/PE – How Is It Done?

- Many techniques and methods available, improving constantly.
- Current knowledge and execution capabilities usually limited to a select few, out of reach from most model developers and model users.
- An "**embarrassingly parallel**" computational problem; solutions involve running a model over and over with slightly different inputs.
- Many EPA models written for Windows, but most supercomputing solutions today require mainframes or Linux-based PC clusters.

## The UA/SA/PE Runtime Problem

- As model complexity, time & space grid density, or types of uncertainty and sensitivity analyzed increases, computational burden (runtime) typically increases geometrically.
- Greatest reason UA/SA/PE techniques not widely applied to EPA models is lack of Windows based computer processing capacity.
- General trend → typical to see PC-based model developers increase model complexity over time, offsetting concurrent gains in CPU speed.
- Depending on the EPA model/application, need 100's to 10's of millions of model simulations.


NERL/ERD's 270 GHz SuperMUSE

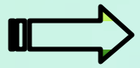
(total 180 PCs)

## Runtime Problem Solution for PC-Models → SuperMUSE

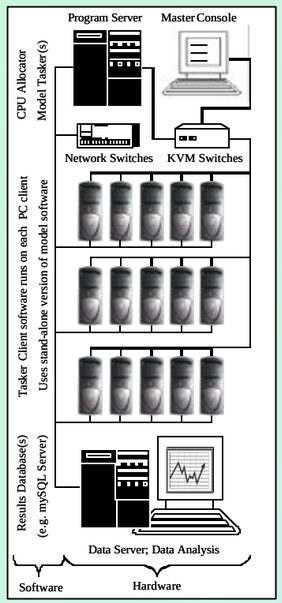
**Supercomputer for Model Uncertainty and Sensitivity Evaluation**

### Clustering to Increase Computational Capacity


Speed of 1 PC




Speed of many




## Conceptual Layout of SuperMUSE Hardware & Software Solution

**….tools to parallelize stand-alone PC-based models**

**Model Tasker**
*Model Dependent-- Create an MT for each modeling system*
Parallelizes a stand-alone PC-Model's System User Interface

**CPU Allocator**
*Model Independent*
Oversees Multiple Active Model Taskers

At launch, MT's register with the CPU Allocator.

1. TaskerClient $TC_1$ announces availability, one of many client PCs.
2. If no active MT, idles. This $TC_1$ call was assigned to $MT_1$.
3. $TC_1$ requests job from $MT_1$.
4. Do job X = a single task line in any MT (e.g., single model run)
5. Send job X warnings, errors, and results to Data Server's project file area identified by MT.
6. $TC_1$ says done with job X.

**Tasker Client** ($TC_k$; k = 1…$n_c$)
*Model independent.*
Executes DOS commands in batch files delivered by MT.

## Why Facilitate Use of Model UA/SA/PE?

- Communicate prediction uncertainty to decision makers.
- Identify critical gaps in knowledge and data.
- Increasing technical focus for regulatory-driven litigation.
- We are called upon to establish validity, trustworthiness, and relevance in model predictions. *(Chen and Beck, 1999)*

## Beneficial Impacts of PC-Based SuperMUSE*ing*

- ✓ SuperMUSE is scalable to individual user (or program & regional office) needs; clustering from 2 to 1000+ PCs.
- ✓ Supports Windows or Linux based modeling systems.
- ✓ Can handle PC models with 10's to 1000's of variables.
- ✓ Solves "**embarrassingly parallel**" computing problems.
- ✓ A local solution → empowers model developers and users.
- ✓ Autonomy from supercomputing centers, removes barriers.
- ✓ Simple, inexpensive, can be built/operated by PC novices.
- ✓ Ideal for debugging models and performing UA/SA/PE.
- ✓ For an average model runtime of 2 minutes, ERD's SuperMUSE can run over 4 million simulations/month.

## Collaborations

- Office of Solid Waste, Hazardous Waste Risk Assessments
- Drs. Beck and Osidele, UGA; global sensitivity analyses
- Dr. Hill, USGS; inverse problem software technologies
- Multi-agency workgroup DoE, DoD, NRC, USDA, NOAA

Partnering to Protect Human Health and the Environment