Overview
of Major Advances by EOHSI:
The MENTOR SYSTEM - Its Tools and Applications

Peer-Review of US EPA - National Exposure Research Laboratory’s
Human Exposure and Dose Simulation University Partnership (HEADSUP) Program
Washington, DC - July 9-10, 2002

by
Paul J. Lioy and Panos G. Georgopoulos

Exposure Measurement and Assessment Division
Environmental and Occupational Health Sciences Institute (EOHSI)
170 Frelinghuysen Road, Piscataway, NJ 08854
Overview

WHAT IS the **M**odeling **E**nvironment for **T**otal **R**isk - **M**ENTOR?

- **MENTOR** is an evolving open “modeling support system” which can also be described as an expandable computational toolbox.

- **MENTOR** is intended to facilitate consistent multiscale source-to-dose modeling exposures to contaminants for individuals and populations.

- **MENTOR** uses a combination of **existing and new approaches**
  - to enhance evaluation and application of existing models
  - to understand environmental and biological processes
MENTOR SYSTEM: Source-to-Dose Exposure Analyses
MENTOR is being constructed to develop, apply and evaluate state-of-the-art methods and computational tools for a wide range of environmental applications, that utilize existing models when available or utilize new models to “fill gaps” in the source-to-dose sequence:

» Example: A “proof of concept” source-to-dose dose analysis of exposure to fine particles and gaseous photochemical pollutants (Philadelphia, PA, 1999)

» Analyses incorporate many new tools (for spatiotemporal interpolation, uncertainty characterization and reduction, population dosimetry, etc.) that allow existing EPA models (for different scales) to work together
OVERVIEW: MENTOR SYSTEM

Computational tools are being developed or adapted to:

- Include refined or new models for environmental, microenvironmental, and biological processes and activities
- Facilitate application and evaluation of models in the source-to-dose sequence (either “stand-alone” or in MENTOR framework)

New methods are being designed for integrated prognostic and diagnostic analyses through:

- “data/model fusion” MENTOR (primarily Bayesian)
- Inverse problem solution (classical and Bayesian)
- Pattern recognition in data and model outcomes (classical and Bayesian)
- Systematic model reduction
- Efficient sensitivity/uncertainty analysis

These new methods are especially promising for “..omic” analysis and computational toxicology source-to-dose applications
Structure and Components of MENTOR

DISTRIBUTED KERNEL WITH SHARED ROUTINES

- Problem definition front end with GIS options
- Hypertext help libraries
- Environmental and biological process models
- Data modules libraries
- Tools for integrated probabilistic & diagnostic analyses
- Software extension libraries
- Internet serving tools libraries

DATA WAREHOUSING, MANAGEMENT, AND MINING ENGINES

- Data warehousing, management, and mining engines

ENVIRONMENTAL INFORMATION SYSTEM TOOLBOXES

- Environmental information system toolboxes

- Macroenvironmental models
- Ecological food-web models
- Local multimedia environmental models
- Microenvironmental models
- Activity pattern/exposure event models
- Biological fate and transport models
- Dose-response models

- Multiscale domain data import/export
- Integrated sensitivity/uncertainty analysis
- Variability characterization
- Bayesian model/data assimilation
- Solution of inverse and ill-posed problems
- Systematic model reduction
- Operational/diagnostic model evaluation
OVERVIEW: PROGRESS

System Design

- Started in Year 01 and there has been steady progress
  - **Computational tool development**
  - **Libraries of models, databases etc.**
  - **Exposure Information System (EXIS) development**
  - **Techniques in uncertainty analysis**
  - **Graphics and hypertext interfaces**

- **Milestones**
  - **Support for EPA models**
  - **Improvement of models for source to dose applications**
  - **Inclusion of computational tools for uncertainty, and probabilistic analysis**
  - **Successful testing of modules and software components**

- Applications of the MENTOR SYSTEM to national issues and emerging problems
- Strong Scientific Collaborations on applications with EPA, and LBNL

Challenges

- Getting the components of the MENTOR to work as a System
- Achieving user friendly computational environment for scientific applications
THE MENTOR MODELING TEAM 2002

EOHSI (UMDNJ & Rutgers)
- Faculty: Panos Georgopoulos, Paul Lioy, Amit Roy, Vikram Vyas, A. Chandrasekar (visiting), Eric Vowinkel (adjunct), Charlie Weschler (adjunct)
- Senior Research Associates: Ming Ouyang, Qing Sun, Sheng-wei Wang
- Research Specialists: Srinivas Bandi, Linda Everett, Samir Goel, Hao-Chen Tan
- GRAs: Suhrid Balakrishnan, Pay-ling Chu, Christos Efstatiou, Wei Li, Paromita Hore, Eric Jayjock, Pamela Shade, Simriti Tanwar, Yu-ching Yang

Rutgers University, Department of Statistics
- Faculty: Bill Strawderman

Princeton University, Departments of Chemistry and Applied Mathematics
- Faculty: Hersh Rabitz; Research Associate: Genyuan Li

Harvard School of Public Health
- Faculty: Petros Koutrakis; Deborah Bennet

TRJ, Inc.
- Ted Johnson, Tom Long

INTERACTIONS/LINKS/COLLABORATIONS
- Amherst, UNC, Vanderbilt, MCNC, STI, ISPRA (Italy), U. Athens & U. Crete (Greece)

- EPA Collaborators: Gary Foley, Linda Sheldon, Haluk ÖzKaynak, Jerry Blancato, Janet Burke, Tom McCurdy, Ram Vedanthan, Lester Grant, William Petersen, Steven Perry, Roger Thompson, Alan Huber

- LBNL Collaborators: Tom McKone, Wayne Ott, Randy Maddalena, Michael Sohn, Bill Riley
Important Achievements

MENTOR is now an operational system that is under continual development and refinement

MENTOR has been successful in supporting and enhancing the application and performance of EPA models

MENTOR has been successfully applied to understand source to dose relationships for important environmental health issues

- Problems specified in the year 02 and 03 workplan for collaboration with EPA and/or members from the LBNL – UPA – e.g. PM and Pesticides
- National Issues that were identified and collaboratively developed between EOHSI and EPA during year 03-04.
  - Arsenic in drinking water
  - Ozone exposure profiles and the new standard
  - WTC plume reconstruction and exposure for WTC epidemiology
Presentation Summary: EOHSI’s “Top Ten plus 1”

1. Development of exposure assessment for \( O_3 \) and PM and tools for evaluating emission control strategies; Integration of MENTOR System with Models-3/CMAQ and linkage to SHEDS

2. Incorporation of receptor/activity variability and pollutant dynamics within biological dosimetry model:
   1. Dosimetry model for hygroscopic PM
   2. Population dosimetry model

3. “Proof-of-concept” demonstration of population source-to-dose assessment for \( PM_{2.5} \) and Ozone
   
   *Case study: Philadelphia, 2 weeks in July 1999*

4. New database for evaluating and refining Pesticides-SHEDS within the MENTOR system, and analysis of field results by a “beta” user

5. Application of MENTOR with integrated individual based model for source-to-dose (environmental/microenvironmental/biological) multimedia/multipathway arsenic exposures
Presentation Summary: EOHSI’s “Top 10 plus 1” cont’d.


7. User-oriented implementation of HDMR; multiple applications (atmospheric chemistry; groundwater transport; PBPK modeling)

8. User-oriented implementation of Bayesian MCMC; application to PBPKM

7. First combined application of SRSM and MCMC to a complex environmental transport model (the finite element FACT groundwater model)

8. Tools for mining databases from large scale exposure studies, and diverse data sources

Initiation of the development of a reconstruction of the WTC exposure profiles from 9-11-2001 through 12-14-2001 (The “plus 1”)
1.a On-Going Integration of MENTOR/SHEDS with CMAQ/Models-3; Development of Compatible Components

“Seamless” integration of Data and Models
First applications are with ambient Particulate Matter and Ozone
1.b PM2.5 Predictions of CMAQ for 10:00 AM, July 19, 1999, EDT
1.c Ozone and PM2.5 Potential Outdoor Exposure Predictions based on CMAQ for July 18, 1999
2.a Incorporation of Receptor/Activity Variability and Physicochemical Pollutant Dynamics in Biological Dosimetry

The new biological dosimetry modules in MENTOR are implemented using a flexible design, and interactive with CHAD (the Consolidated Human Activities Database):

- physiological variability due to age, gender, weight, etc.
- continuous temporal variability due to physical activity (metabolic expenditure)

Both respiratory PM deposition and multimedia/multipathway PBPK models have been implemented with flexible design.

The receptor activity modules are fully integrated with the microenvironmental modules to reflect modifications due to receptor’s presence and activities.

Modules for respiratory PM deposition account dynamically for various size-specific physicochemical properties of inhaled particles (such as hygroscopicity) and changes occurring in the lungs (particle growth, etc.).

- Two peer-reviewed articles on: “non-ideal” (i.e. hygroscopic, reactive, etc.) aerosol dosimetry modeling were published in 2001 (in Aerosol Science and Technology and in Environmental Science and Technology)
2.c Example of a “Level B” Calculation: Evolution of PM Size Distribution in the Human Respiratory Tract via 1D Macromodeling

Fine Hygroscopic PM Concentration/ Size Variation Along the Conducting Airways of the Human Respiratory Tract (Persistent vs Deposition)
3.a “Proof-of-Concept” Demonstration of Integrated Population Source-to-Dose Assessment for PM$_{2.5}$

First demonstration (“proof-of-concept”) of a complete, integrated, population source-to-dose analysis for PM2.5 and Ozone

Case study

- Philadelphia, 2 weeks in July 1999*
  - The study involved the combined application of MM5, SMOKE, and CMAQ with a new census-tract-level population exposures/dose model which merged approaches and tools of SHEDS and MENTOR
  
  - The revised/expanded SHEDS code (MENTOR/PM-SHEDS in MATLAB) is interactively linked with CHAD and allows exposure and dose calculation for each activity event (flexible time implementation)
  
  - A sample of 500 individuals developed to match the demographic characteristics of each census tract is used to extract activity information from CHAD and drive the exposure and dose modules
3.c Example Results from the First Source-to-Dose Population Exposure Assessment for PM2.5 (Case Study: Philadelphia, 7/11-24/1999)

Percentile Plot for 1-Hour Aggregated Total Dose, Dose due to Outdoor Sources, and Dose due to Indoor Sources (All Days: 7/11-24/1999).

24-Hour Aggregated Total Dose, Dose due to Outdoor Sources, and Dose due to Indoor Sources on 18 July 1999
4a. Application of MENTOR/ Pesticides Sheds

Approach: Use of information and results from EOHSI/EPA - Children’s Post Pesticide Application Exposure Study (CPPAES):

1. Pesticide accumulation patterns for child accessible surfaces and objects
2. Microenvironmental data coupled with urinary metabolite excretion by children for two weeks after professional residential crack and crevice application of chlorpyrifos
4b. Study Objectives

• To examine the time course of chlorpyrifos accumulation patterns within the indoor environment (i.e. air, surfaces, objects, etc.) for two weeks following an indoor “crack and crevice” application

• To use microenvironmental data coupled with activity data as input parameters for a mechanistic application of MENTOR/Pesticides-SHEDS and estimate of the amount of chlorpyrifos absorbed by a child exposed to the pesticide

• To test the performance of MENTOR/Pesticides-SHEDS by comparing the model estimates with the actual amounts of chlorpyrifos absorbed by a child, from the measurement of chlorpyrifos metabolites in the urine of each subject
4c. Summary of Current Results and Activities for MENTOR/SHEDS - pesticides

- Instead of sampling from default distributions, child specific & activity specific daily input parameters were extracted/estimated from the CPPAES activity diaries & environmental measurements.

- Daily trends for TCP elimination & inhalation exposure were comparable between the field measurements and model simulations.

- Estimates of TCP metabolite levels in the urine by MENTOR/Pesticides - SHEDS were within the same order of magnitude as the results from the field study for each child.

- Major routes of exposure were determined to be non-dietary ingestion (Non-Dietary Ingestion > Inhalation > Dermal > Dietary).

- Collaborative EOHSI/EPA efforts revealed coding problems in the exposure & dose scenarios which were fixed during the performance evaluations using the CPPAES results.
5.a Individual Exposure to Arsenic: A New Coupled Source - to - Dose (Environmental/μEnvironmental/PBPK) Application

- EPA NET (municipal tapwater
  - GMS (private well water)
- ISC/AERMOD (outdoor air concentration)
- Location-specific monitor data (outdoor air concentration)

Microenvironmental Model

1. Choose environmental agent (TCE, PERC, and arsenic are currently available)
2. Choose attributes of exposed cohort (age range, gender, employment status)
3. Randomly select a CHAD ID that matches the attributes of exposed cohort
4. Retrieve the activity diary associated with the selected CHAD ID
5. Specify attributes of home based on attributes consistent with the selected CHAD ID
6. Map CHAD locations and activity codes to water uses and sources (showers, tubs, faucets, toilets, laundry machines)
7. Generate water use events of random durations within the CHAD events mapped to water uses
8. Retrieve outdoor air and tapwater concentrations of environmental agent as a function of time
9. Solve dynamic equations to estimate indoor air concentration time profiles of environmental agent resulting from events in activity diary
10. Call pharmacokinetic model to estimate dose

Pharmacokinetic Model

1. Map CHAD ID locations to locations in the Indoor Air Quality Model
2. Retrieve METS distribution parameters and assign a METS value to each activity event
3. Randomly select a CHAD ID
4. Retrieve food and water consumption diaries for individuals having attributes similar to that of the selected CHAD ID
5. Map CHAD location and activity codes to food and water consumption events
6. Specify food and water consumption events during these events
7. Specify extent of dermal contact during water use events based on location and activity codes
8. Solve dynamic equations to estimate inhaled, dermal and ingested doses of environmental agent

START

START
5.b Microenvironmental – PBPK Model Test: Indoors Inhalation Exposure/Dose to Arsenic (Source: Tap Water Use, Outdoor Air)

(Outdoor air: 100 pg/m³; Tap water: 50 ppb)
6.a Population Exposure to Arsenic: Source-to-Dose Applications of Multimedia MENTOR/SHEDS

Test applications:

a) Comparison of population exposures due to inhalation (from outdoor sources) and ingestion (from drinking water) in two counties with reported groundwater arsenic problems: Pima, AZ and Hunterdon, NJ

b) Multipathway exposure/dose assessment using NHEXAS V data
7.a Model Reduction and Sensitivity/Uncertainty Analysis: HDMR (High Dimensional Model Representation) Based Tools

User-defined information on original model:
- number of input variables (n)
- distributions (or ranges) of input variables
- order of HDMR
- resolution (mesh)

HDMR input processing module for sampling point generation: defines series expansion form of I/O relationship

(m x n) matrix of input values for original model
- m: number of simulations (depends on HDMR order and resolution)

HDMR-reduced model

ORIGINAL MODEL simulations

(m x p) matrix of model outputs
- p: number of model outputs

Interpolation routines

HDMR core calculation (produces output series expansion terms used to build look-up tables)

A “Fast Equivalent Operational Model” (FEOM) that can substitute the original, complex model in calculations
7.b Sensitivity/ Uncertainty Characterization of Critical Processes in the Source-To-Dose Sequence and Model Reduction

Application and testing/refinement of Stochastic Response Surface (SRSM), Automatic Differentiation (ADIFOR), and High Dimensional Model Representation (HDMR) methods for characterization of uncertainty to key parameters in exposure/dose, and systematic simplification of relevant models.

In this example HDMR was used to identify key input variable contributions to output uncertainties of the Princeton Groundwater Model, used to simulate bioremediation schemes for trace metals and radionuclides. With only 300 model runs, the HDMR methodology efficiently identified the 7 most relevant input variables (rate constants of biogeochemical reactions) among the 20 input variables, that affect uncertainty in the output (accumulated flux of U(IV) across the downstream of the remediation site).
This example shows the comparative application of the new state-of-the-art uncertainty tools in MENTOR [the High Dimensional Model Representation (HDMR), and the Stochastic Response Surface (SRS) methods], and of “classical” Monte Carlo analysis, to the propagation of uncertainties in the variables of a 3-dimensional finite element Flow and Contaminant Transport (FACT) groundwater model.

In the above figure HDMR and SRSM reproduce probability density functions of two outputs from FACT, derived from 1,000 Monte Carlo runs, using only 45 and 51 model runs, respectively.
8.a The Bayesian Approach for Analyzing Biological Problems when Mechanistic (and not only) Models are Available

“in this new century ... a significant part of the everyday practice of Statistics ... will consist of applying Bayes' formula via MCMC ...”

9.a Efficient Uncertainty Analysis of a Physiologically Based Pharmacokinetic Model Using SRSM (and SRSM/ADI FOR)

Probability densities of dose surrogates, estimated by Standard Monte Carlo, Latin Hypercube Sampling (LHS), SRSM, and the SRSM-ADI FOR.
10.a Tools for Mining and Analysis of Exposure-Related Data

Evaluation of data mining tools

- Pattern recognition in environmental and exposure databases
- Comparison of patterns in observational databases versus model predictions for diagnostic evaluation purposes
- On-going application of various alternative methods (and of specific software implementations)
  - \textit{Cluster, factor, CART, SVD analysis}

\textbf{Example}: CART characterization of the 90th percentile of \textit{blood benzene concentration} (error rate is estimated to be 4\%). The most important five attributes that separate the top 10\% group from the rest (in order of importance) are: number of cigarettes smoked per day (numerical), indoor air benzene concentration (numerical), job in contact with plastic fumes (categorical) time spent on highway (numerical), and time spent outside at home (numerical)
10.b Example of “Mining” Data from the NHEXAS Study via CART
Arsenic in Media (Food, Water, Air, Dust) and Biomarkers (Urine)
Response to a National Issue – The Attack on The WTC

- Initial discussions between Dr. Gary Foley and Dr. Paul Lioy -9/18/2001 through 10/20/2001
  - *What can be done to reconstruct exposure for epidemiology and risk assessment*
  - *What are the information available and the information needs*

- Defined our role under the UPA
  - *Collect indoor samples for analysis and use in long term exposure characterization*
  - *Develop a work plan for reconstruction of the plume for at least 9/11 though 9/26/2001*

- Expanded into EPA/EOHSI collaboration
  - *Work with CFD modeling efforts at EPA for assessments through end of Fires- 12-14-2001*
  - *Work with EPA Wind Tunnel Laboratory to develop plume characteristics for specific days of highest concern for exposure - Post day 1*
Initially Exposed Population on September 11, 2001
Sources of Support for the MENTOR System Efforts

**MAIN SUPPORT: US EPA NERL**
- through the Center for Exposure and Risk Modeling (CERM) at EOHSI-HEADSUP
- through other Grants and Cooperative Agreements
  - *Outdoor/Indoor PM Relationships Project*
  - *NE-OPS Modeling and Data Analysis Component (Penn State Subcontract)*

**OTHER SUPPORT: projects that “feed” results and modeling experiences into MENTOR**
- **US DOE**
  - through the Consortium for Risk Evaluation with Stakeholder participation (CRESP) (for PM emissions, resuspension, transport, fate)
- **NI EHS**
  - through the Environmental Health Sciences Center at EOHSI (WTC Exposure Assessment)
- **NJ DEP**
  - through the Ozone Research Center (ORC) at EOHSI
- **NJ DHSS/ATSDR**
  - Exposure assessment for the Toms River Cancer Cluster
- **Industrial**
  - PSE&G (Ozone PM, Air Toxics)
  - ICA (Multimedia Heavy Metals with focus on Copper; EXIS tools development)
  - Port of Authority - NY/NJ – Aircraft Emissions and exposure assessment
- **USNRL/Quantum Leap Consortium**
  - Homeland Defense Initiative