9.0 3MRA Version 1.0 Uncertainty Analysis and Sensitivity Analysis Plan

This Section describes the specific efforts needed to complete uncertainty and sensitivity analyses (UA/SA) of the 3MRA Version 1.0 modeling system. Furthermore, it summarizes overall efforts undertaken to date by OSW and ORD to conduct a comprehensive model evaluation of 3MRA in the context of the model evaluation terminology introduced in Section 2.

Overview of 3MRA Modeling System Documentation

In communicating a formal plan for UA/SA of the Multimedia, Multireceptor, and Multipathway modeling system (3MRA), supporting information was first developed to place in context its technical underpinning. In addition to documentation provided in Volumes I, II, and III, and the *3MRA Modeling System: Technology Design and Users Guide*, the following elements, defining the context for uncertainty and sensitivity analysis for 3MRA Version 1.0, were discussed in detail in Sections 1 through 8 of this report (Volume IV), covering:

- An overview of the national assessment problem statement under consideration for the application of 3MRA Version 1.0 the in risk assessment of hazardous contaminant disposal (and treatment) in land-based waste management units (Section 1).

- A delineation of the various taxonomies of probabilistic assessment, uncertainty and sensitivity analyses that introduces a comprehensive suite of methodologies available for "evaluation" of environmental exposure and risk assessment modeling systems (Section 2).

- A brief review of the underlying science methodology of 3MRA, and its constituent components and datasets (Section 3).

- A brief review of the technological solution for the 3MRA science methodology, formulated as the 3MRA Version 1.0 modeling system software set, along with the dimensionality of the overall problem statement posed by the national assessment strategy under consideration (Section 4).

- An overview of issues of "embarrassing parallelism" associated with performing uncertainty and sensitivity analyses on complex environmental models such as 3MRA, along with discussion of the hardware solution provided by

SuperMUSE, a 225 GHz PC-based, Windows-based, Supercomputer for Model Uncertainty and Sensitivity Evaluation (Section 5).

- A discussion of the software elements supporting the SuperMUSE hardware design approach, along with key enhancements needed for 3MRA model evaluation, embodied in the pending release of 3MRA Version 1.x (Section 6).

- Representing a typical application of 3MRA Version 1.0 and the SuperMUSE hardware and software toolset, an evaluation of Benzene disposal in various land-based waste management units (Section 7).

- Finally, 3MRA Version 1.0 model inputs and model outputs were defined and summarized. (Section 8).

9.1 Review of Model Evaluation Tasking Concepts

Reiterating the discussion in Section 2.3 of the 3MRA modeling system's "synthesis" and "analysis" stages, verification, validation and predictive uncertainty serve as the ultimate focal points of the overall model evaluation program being undertaken for 3MRA. Quality assurance in 3MRA technology design is the end goal of this synthesis and analysis effort, culminating in a product of overall model validation for the specific task defined (i.e., the national assessment strategy). Specific model evaluation tasking can be conceptually broken-down along lines of "compositional validity", an internal measurement of the model, and "performance validity", an external measurement of the model, made in the context of its intended use for a specific purpose.

As an example of how UA and SA interrelate in this procedure, sensitivity analysis tasking performed at various levels (e.g., unit or module level, and system-level) is used to evaluate aspects of both the modeling system's composition, and its performance. Compositional uncertainty is used to assess the state of compositional validity, and is embedded in the process conducted to date to develop the conceptual model for 3MRA (and its modules); for example, in the delineation of minimum and maximum values that model inputs are allowed to take-on, and in specification (or internalization) of correlated model inputs.

9.1.1 Composition

In distinguishing the attributes of composition versus performance, the first of two stages in developing the model application, model synthesis, is completed by the act of code verification, at which point the model's compositional validity is established (Beck *et al.*, 1997). Model evaluation then proceeds, involving aspects of model input specification for the task at hand, calibration of input data, which can be infused in a variety of ways during model input specification, and performance validation. Collectively, uncertainty and sensitivity analyses; model verification, comparison, and validation; and peer review tasks comprise a formal program of model evaluation for 3MRA. This evaluation program is intended to establish the requisite level of confidence needed for the 3MRA's use for the given, prospective task of prediction (i.e., the national assessment strategy).

Quantitative descriptions of compositional validity of a model can be tied to perspectives of model input uncertainty, and is best suited overall to a process of group peer-review, which together form a basis of integrating expert judgment and historical experience (Beck *et al.*, 1997).

9.1.2 Performance

Output uncertainty analysis being undertaken for 3MRA serves a vital role in characterizing performance of the model for the application under consideration. As a performance validation measure for predictive risk analysis, our US/SA plan builds upon the works of Young, Hornberger, Spear, Chen, and Beck (Section 2.8) in applying the notion of a model having maximum relevance to the performance of a specific task, thereby broadening the discussion of model validation into one of quality assurance in environmental forecasting.

Here, the concept of maximum relevancy is interpreted as an assessment of key (sensitive) and redundant (insensitive) model inputs, an identification process irrespective of output uncertainty, placed in the context of a binary classification of behavior and non-behavior in the model output space. One may also utilize an intermediate concept of important (or moderately sensitive) inputs, representing a gradation of statistical significance, between significance levels assigned in determining the key and redundant inputs. Notably, the underlying sensitivity analysis technique (i.e., RSA) provides opportunities to both qualify input uncertainty importance, and to assess model performance in forecasting.

9.2 3MRA Model Evaluation Tasking Completed To Date

While many model evaluation tasks have been executed, several tasks remain to be completed (i.e., the model comparison study, UA, system-level SA, and $SA_{p;}$ see Section 2.3.5) in evaluating 3MRA Version 1.0/1.x for the national scale assessment strategy. Embodied throughout this Volume Set I-IV and outlined in Figure 2-3, one can place associated tasking completed to date in terms of composition and (constituent) performance validation procedures for model evaluation as discussed in Section 2.3. Such activities completed to date to support the application of 3MRA version 1.0 for the national assessment strategy include:

- Compositional validity supported through the extensive, iterative peer-review process of the science methodology design (i.e., the science plan; Marin *et al.*, 1999, 2003).

- Compositional validity supported through the extensive, iterative peer-review process of the technology design and implementation embodied by 3MRA Version 1.0.

- Compositional and performance validation is further supported by:
    - *3MRA Modeling System: Technology Design and Users Guide*.
    - Modeling system and module-specific description and documentation. (Volume I).
    - 3MRA national, regional, and site-based data collection (Volume II).
    - 3MRA chemical and meteorological data collection (Volume II).
    - Documentation of module verification and validation studies (Volume III).
    - Initial system-level model comparison efforts (Volume III).
    - Extensive system-level verification and precision testing (Volume IV).
    - System-level testing of specific chemical scenarios (Volume IV).
    - The final peer-review for 3MRA Version 1.0, and enhancements in 3MRA Version 1.x, both under current consideration by EPA's Science Advisory Board.

- Development of suitable computational capacity and reliable supporting software infrastructure needed (3MRA Version 1.x) to perform the millions of individual model runs required to conduct random sampling-based uncertainty and sensitivity analysis experiments for 3MRA Version 1.0 for multiple chemicals and source types.

The above activities address the verification and validation needs of both individual constituent 3MRA model components, and the integration of its components at the system-level. Module level performance analysis is used in part, for example, to further assert compositional validity at the system-level, or in other words, to assess the integrity of the system's constituent hypotheses. Furthermore, the computational scheme delivered via SuperMUSE facilitates the ability to perform the extensive sampling-based experimentation that will be needed to conduct supporting uncertainty and sensitivity analyses of 3MRA to address a variety of assessment questions posed by the decision-maker.

9.3 UA/SA 3MRA Model Evaluation Tasking To Be Completed

With the sum total effort embodied in the previously described tasking completed to date, as outlined in Section 9.2, the remaining portions of the UA/SA plan, discussed in the following materials, coupled with the ongoing 3MRA-TRIM model comparison and validation study underway, as described in Volume III, represent the tasking needed to complete a sufficiently comprehensive picture of overall compositional and performance validations for 3MRA. This is, of course, rooted in the 3MRA application for the intended purpose of conducting national scale risk assessments of land-based disposal (and treatment) of hazardous waste constituents. The supporting UA/SA plan is formulated within a probabilistic, sampling-based analysis using Monte Carlo Simulation.

9.3.1 Five Components of UA/SA

Ignoring for now the inter-dimensionality of individual model inputs x_i, which can vary from site to site, and referencing the terminology introduced in Section 1.3.1 (Helton and Davis, 2000, 2002, 2003) and extended in Section 8 for the national assessment strategy, described by Helton and Davis, UA and SA involves five major components:

1) Definition of uncertainty distributions D_1, D_2, …,D_{n_x} that describe the total uncertainty in the input factors, if any; where these distributions may be site-specific, region-specific, or national-based, and where each distribution itself may be further described as a family of distributions representing empirical uncertainty in the specification of a specific D_i (i.e., 2nd-order analysis for ISE).

2) Generation of the $\mathbf{x}_k$ vectors from the input distributions; i.e., through random sampling;

3) Generation of output samples $\mathbf{y}(\mathbf{x}_k)$; i.e., through n_s Monte Carlo simulations;

4) Displays of the uncertainty in $\mathbf{y}$ from the analysis outcomes of $\mathbf{y}(\mathbf{x}_k)$; i.e., total uncertainty analysis via ELP2 (e.g., using CDFs); and

5) Exploration of the mapping $[\mathbf{x}_k, \mathbf{y}(\mathbf{x}_k)]$ to determine the effects of elements of $\mathbf{x}$ on $\mathbf{y}$; i.e., sensitivity analysis.

Where:

- The values n_x and n_y represent the number of elements of **x** and **y**, respectively,
- The value n_s is the number of samples or iterations,
- The value n_{ss} is the total number of site-source combinations in the 3MRA site-based database,
- The value n_{Cw} is the number of wastestream concentrations considered, and
- This construction is repeated $n_{ss}*n_s*n_{Cw}$ times for each chemical under consideration.

Summary of the UA/SA Plan

Referencing terminology introduced in Section 2, the remainder of the 3MRA UA/SA plan to be completed is comprised of three major tasks:

- Performance uncertainty analysis (UA_p):
    - Basically entails propagation of input uncertainty through the modeling system, while also addressing output sampling error (OSE) associated with computational limitations of the sampling-based MCS strategy.
    - This will initially be performed using a pseudo 2nd-order analysis to address empirical input uncertainty and OSE, less ISE (Section 2.6.6).

  - Depending on outcomes of the sensitivity analyses (SA), a limited full 2nd-order analysis (Section 2.6.6) could possibly be undertaken, to the degree feasible. Such an analysis might further address uncertainty in the empirical distribution specifications associated with sample measurement error (SME) and input sampling error (ISE) for extremely sensitive model inputs, provided suitable information could be made available to form the analysis.

  - Model error (ME) cannot be formally quantified due to an overall lack of information available that would make such an effort meaningful at this point in time.

- System-level sensitivity analysis (SA):

  - Basically exploring the mapping [$\mathbf{x}_k$, $\mathbf{y}(\mathbf{x}_k)$] through use of several analytical techniques, identifying key, important, and redundant model inputs.

  - SA to be conducted for this purpose will enhance both compositional and performance validation aspects for the modeling system.

  - The latter (i.e., an aspect of performance uncertainty analysis) is reflected upon as a qualification of the importance of accurately quantifying input uncertainty in support of the final UA_p.

  - The former (i.e., compositional validation) represents additional activity supporting system-level modeling system verification (through identification of unexpected model output behavior over the allowable ranges of inputs; e.g., discontinuities, linearity, non-monotonicity etc.).

- Sensitivity analysis based performance validation (SA_p)

  - To be evaluated using 3MRA Version 1.x. Basically an assessment of a "prior" validity through the execution of a regional sensitivity analysis (RSA) procedure, realized as an assessment of the model's maximum relevancy in predicting model behavior for various population percentiles.

Representative Chemicals To Be Evaluated

Of the 43 chemicals currently available within the 3MRA modeling system chemical properties database, evaluation of 7 chemicals will be conducted as part of this plan. These 7 cover a wide range of chemical behaviors *in situ* (e.g., VOCs, SVOCs, and metals). The chemicals proposed for this evaluation are:

1. Benzene,
2. PCE,
3. 2,3,7,8-TCDD,
4. Benzo(a)pyrene,
5. Arsenic,
6. Nickel, and
7. Divalent Mercury.

These chemicals were selected as a broadly representative set, spanning various chemical properties exhibited among the existing organic chemicals and metals currently parameterized in the 3MRA chemical properties database. As a point of recommendation from ORD, a similar performance uncertainty analysis (UA_p), as described above, should be executed prior to the use of 3MRA model predictions as a basis for any regulatory-based decision-making for other chemicals under investigation by OSW.

9.3.2 Overview of Greatest Questions of Concern

As a precursor to model evaluation tasking detailed here for 3MRA, the following example questions posed by OSW cover major areas of greatest interest in the national scale risk assessment strategy. These questions were used to focus the overall development of 3MRA Version 1.0, 3MRA Version 1.x, and the UA/SA plan described herein, and will be addressed in various ways during the execution of this UA/SA plan.

Example Questions Concerning Population and Subpopulation Risk

- What is the driving pathway[s] for a particular chemical/protection group?
- Are there pathways that never drive risk/exit level?
- What is the contribution to risk/exit level by pathway?
- Who are the driving receptors?
- What is the contribution to risk/exit level by receptor?
- What subpopulations are not protected at the exit level determined by the assigned protection criteria (e.g., 90% population protection at 95% of sites)?
- What are the sensitive parameters at the system level that could help target data collection needs?
- What is the number of Monte Carlo iterations needed for convergence, what is the run-time?
- What is the impact on the results/run-time using fewer iterations than are needed for convergence (i.e., completely minimized OSE)?
- What is the % of total population protected at a given risk criteria (i.e., if 95% of receptors are protected at 95% of sites, what is the total % of the national population protected across all sites?)?
- What is the geographical distribution of unprotected sites?
- What is the individual risk (at a given "protective" C_w) to a subsistence receptor (fisher and farmer)?

9.3.3 Pseudo2nd-Order Uncertainty Analysis (UA_p)

The initial performance uncertainty analysis UA_p is straightforward, involving the execution of the modeling system, via 3MRA Version 1.x, for the chemicals of interest across all site-WMU combinations (419), for all five C_w's. Per chemical, this represents 2095 individual modeling system runs per national realization, or 14,665 model runs for the 7 chemicals selected for the analysis.

As discussed in Section 2.5.3, convergence in output sampling is expected to require on the order of millions to tens of millions of model runs for all 7 chemicals. For example, 3457 national realizations would be needed, or just over 7,240,000 individual 3MRA model runs per chemical if the rather restrictive approach outlined by Morgan and Henrion (1990) in equations 2.2, 2.3, and 2.4 is used to establish the 91st probability percentile CDF as representing a conservative limit on meeting a 90th probability percentile strategy, with 95% confidence for a given population percentile of interest (see Section 2.6.6 and Section 7).

Alternatively, as an example from data shown in Section 7, using equation 2.1 in Section 2.5.3 it was estimated that for the initial Benzene analysis conducted on C_w5, roughly 165 to 644 national realizations would be needed to arrive at 95% confidence in specifying a 1.0 to 0.5 unit percentage confidence interval, respectively, on the 50th probability percentile estimate of % sites protected for the 90% population percentile. As opposed to relying on the above approaches alone, the pseudo 2nd-order uncertainty analysis of empirical input uncertainty will be evaluated to monitor convergence in estimates of probability percentiles > 50% (i.e., 90%, 95%, etc.). Collectively, these approaches will be used to quantify, as a function of the number of model runs completed, convergence behavior in risk estimates.

Estimated UAp Runtime

In general, the approach to be used in 3MRA UAp will be to segregate ELP1 data structures by national realization, and then subsequently perform a pseudo 2nd-order uncertainty analysis, separating empirical uncertainty from variability, and describing output sampling error (see Figure 2-10). Initial analysis per chemical will likely cover 500^{+} national realizations, or 115 days of calculations on the existing SuperMUSE, with additional calculations possible depending on the desire to further minimize OSE. Depending on capability expansion realized for SuperMUSE (e.g., to a 384 node system with updated processors; Section 5), additional simulations to sufficiently minimize OSE would be relatively inexpensive to achieve. For example, at the current new PC processor speeds of 3 to 4 GHz, roughly 2500 to 3000 national realizations for the 7 chemicals could be conducted within the same 115-day time frame. The capital investment needed to realize these computational capacities is < $300K. The time, or, alternatively, computing capacity needed to minimize OSE in UA_p for all 43 chemicals in the current 3MRA database is substantial at existing SuperMUSE computing capacities. Runtime requirements will be generally linear with respect to the above analysis.

Execution Scheme and Data Management

To facilitate the intensive post-processing demands of the UA/SA work, an additional back end server was recently obtained to move national realizations off the existing back-end server as they are completed, where post-processing of ELP1 data will actually be performed. This PC offers dual 3.0 GHz processors, 4 GB RAM, and 220 GB storage space realized in 4 separate, 15,000 RPM SCSI hard drives, with RAID5 data striping to ensure data integrity. Current individual PC client hard drive capacities can manage roughly 50 realizations at a time. To maintain stability of clients throughout the calculation period, the basic simulation scheme will be to issue batches of roughly 20 to 30 national realizations at a time, using a unique, randomly selected initial seed for each batch, without replacement (see Section 4.3.4). A list of 5000 initial seed values has already been generated and will be used. ELP1 data collections of the previous batch will be conducted during modeling system simulations for the next batch.

Qualitative Analysis of Dominant Media, Pathways, and Receptors Driving Exit Levels

Using simulation data generated by the UA_p effort, analysis will also be conducted to address, in part, the questions OSW posed in Section 9.3.2. Primarily through graphical analysis, and use of: (1) fully aggregated ELP1 MySQL data structures (combining all realizations); and (2) the enhanced ELP2 Visualization tool, the driving pathways, contact media, receptors, and associated cohorts for various chemicals and protection strategies will be evaluated. This information will be used to determine exposure profiles of greatest interest for subsequent sensitivity analysis. As discussed in Sections 4.5.8 and 6.4.5, sensitivity analysis of exit levels will require use of the disaggregated ELP scheme. Judicious selection of critical profiles of interest is mandatory to feasibly handle associated data storage and post processing costs, in conducting sampling-based sensitivity analysis, which preserve the mapping between unique inputs vectors and exit levels.

9.3.4 Sensitivity Analysis (SA)

A balanced, tiered formulation of SA is planned for 3MRA. The basic approach to be undertaken for 3MRA, one of global input space assessment, entails use of sampling-based correlation and regression methods, complemented by procedures employing a tiered regional sensitivity analysis (RSA) and tree-structured density estimation (TSDE) methodology. Wastestream concentration exit levels for the dominant exposure profiles identified in the UAp work and determined through analysis of aggregated ELP1/ELP2 schemes will represent the model outputs for SA focus. This new work will be conducted using the disaggregated ELP scheme (Section 6.4.5), where exit levels will be calculated using the enhanced, disaggregated routines of the ELP2 for unique input vectors and the basic 5 C_w's.

As a precursor to this work, initial Site Summary input-output sampling schemes will be implemented during the UAp simulation work to assess the sensitivity of certain outputs of the human and ecological risk modules, themselves dominant drivers of exit levels (e.g., summary statistics of human hazard data: HQ_#_values; human cancer risk data: Risk_#_values; and ecological: HQMax). This screening assessment can be conducted with the standard aggregated ELP1 scheme (Section 4.5.8). The results will allow early identification of overtly or relatively insensitive inputs, which will assist in reducing subsequent overall run-by-run input extraction needs and associated data storage and post-processing costs.

Regression/Correlation Approaches

To evaluate sensitivity and uncertainty of 3MRA, several regression/correlation-based global analysis techniques will be attempted. These techniques will be based on the disaggregated, random sampling scheme (Section 6.4.5), and will rely on assumptions of near-linearity and near-monotonicity in describing the relationships between input and output vectors. Existing input distributions will be used in addition to introducing variation into existing national point estimate values. Sensitivity analysis under this area will utilize various statistical techniques (e.g., scatter plots, regression and stepwise regression analyses, correlation and partial correlation analyses; with and without use of rank transformations). Analysis will also include an evaluation of modeling system linearity and non-monotonic and nonrandom patterns. The procedures to be followed are outlined in detail in Section 2.7.4, and will follow guidance provided by Helton and Davis (2000, 2002, 2003).

Univariate RSA and Multivariate TSDE Global SA Techniques

In an effort to establish enhanced SA capabilities in global sensitivity analysis, two algorithms will be explored. These algorithms retain a unique potential to be used in tandem to define an integrated, global analysis methodology for application to 3MRA. These two algorithms are the Regional (or generalized) Sensitivity Analysis (RSA) procedure (Spear and Hornberger, 1980) and the Tree Structured Density Estimation (TSDE) procedure (Spear *et al.*, 1994). The procedures to be followed were outlined in Section 2.7.5, and are based on descriptive information on RSA and TSDE provided in the works of Chen and Beck (1999), Beck and Chen (2000), and Osidele and Beck (2001a, 2001b). An ORD contract has been underway since 2001 with Professor Beck and Dr. Osidele, of the University of Georgia, to integrate and develop the needed post-processing schemes within 3MRA Version 1.x.

The objective of the RSA procedure is to rank the importance of the uncertainties attributed to the model input factors with respect to matching the prescribed output behavior definitions, and on this basis, identify the key and redundant input parameters and related processes in the system. The aim of the TSDE procedure is to identify interactions among the input factors in the accepted *behavior* simulations derived from the RSA procedure. Thus, the dual tandem approach follows-up the univariate RSA procedure with a qualitative, multivariate analysis that ranks the relative importance of individual 3MRA modeling system inputs.

Experimental Design and Level of Effort

The same sampling experiments used in the regression/correlation effort, previously discussed, can also be utilized for the RSA and TSDE sensitivity analyses. The elegance of this approach is also extended to similar use of the same simulation experiments to conduct the sensitivity-based performance validation (SA_p) of 3MRA, discussed in Section 9.3.5 that would employ the RSA/TSDE. While the initial regression/correlation approaches can be feasibly completed within 6 to 8 months at existing SuperMUSE capacity, the RSA/TSDE sensitivity analysis work will take longer to implement. These post-processing schemes, particularly the experimental approach being developed for TSDE, which is calculation intensive, and would likely require an effort of 9 to 12 months or longer to complete results, if feasible.

9.3.5 Sensitivity-based Performance Validation (SA_p)

Prior performance validity of 3MRA Version 1.x will be assessed through the execution of a regional sensitivity analysis procedure, realized as an assessment of the model's maximum relevancy in predicting model behavior for various population percentiles. To further evaluate limitations of the univariate RSA, the multivariate TSDE approach will also be investigated in this context. The approach represents a reflection of the evaluation of the external definition of the task, or model purpose, back onto the internal composition of the model (Beck *et al.*, 1997). The overall objective is to expose underlying behavioral characteristics of 3MRA, reflecting upon the nature of its internal input-output structure.

The methodology for establishing "prior performance" validation of model behavior under novel conditions was outlined in Sections 2.3 and 2.8, and has been further explicated in a simplified example of a national application for multimedia modeling of hazardous waste disposal (Chen and Beck, 1999 – see Appendix B; Beck and Chen, 2000). These latter references serve as guides in this research endeavor to provide a novel performance validation of 3MRA Version 1.0/1.x. The analysis to be undertaken will focus on identification of key and redundant parameters in predicting protection of various population percentiles in the context of the national risk assessment strategy (e.g., 99%, 90%, 50%).

9.3.6 ISE-Based 2nd-Order Uncertainty Analysis

Referenced to the definitions offered in Section 2 for uncertainty, the three main types of empirical errors that can be introduced during input formulation and model execution are:

- Probability distribution function type assignment for each random variable (i.e., is a given model input distributed uniformly, normally, or lognormally, etc.?)
- Random variable distribution parameterization (i.e., description of distribution parameters; e.g. mean, variance, etc.; i.e., ISE/SME), and
- Model description (i.e., 1-D versus 3-D, etc.)

Indicated previously, there is no feasible ability at this stage of knowledge of complex, very high order modeling systems (VHOMs) such as 3MRA, to quantify model structure error.

Background discussion of the separation of variability from empirical uncertainty was presented previously in Section 2 and Section 3. The need to attempt a full 2nd-order dimensional analysis of variability and empirical uncertainty (Section 2.6.6) will depend on the outcome of sensitivity analysis results above. Further investigation into SME and ISE could possibly be conducted on a limited basis, utilizing some form of this approach, given available data that is currently lacking. The tiered approach outlined in this UA/SA plan leads to a more optimal path for identification of sensitive inputs for further evaluation over time. Even within this tiered approach, an additional tiered evaluation strategy of the efficacy of employing 2nd-order analysis for ISE or SME would be undertaken subsequent to sensitivity analysis.

Tiered Evaluation Approach for Investigating Higher Dimensional Analyses

Higher-dimensional analyses for ISE and SME would logically start with a qualitative assessment to weight expert knowledge relative to the likelihood of their importance in existing parameterizations of input distributions. This, for example, could be sufficient in characterizing the need to forgo higher-dimensional analysis. Highly sensitive model inputs potentially subject to significant levels of empirical ISE and SME uncertainty, or those without additional inference regarding importance, could be subjected to extended distribution analysis and reformulation.

Depending on the computational requirements of additional 2nd-order analysis that might be conducted for ISE and SME, several approaches could be utilized in the experimental design, with increasing complexity in level of effort:

1. For selected parameters, a screening-level perturbation analysis could be conducted by creation of several alternative distributions for each model input of interest, and an associated "family of databases" created. These then would be subjected to additional simulation with the existing toolset available in 3MRA Version 1.x. This would be accomplished by constructing several alternative representations of national and regional databases, and could also take advantage of a manually implemented LHS-type procedure in database family construction.

2. A more formal development of 2nd-order analysis through creation of a Distribution Statistics Processor could be made available over time to accomplish the same solution approach, while also being capable of completely random sampling designs.

3. Alternatively, a higher-dimensional analysis could be constructed through use of the n-stage iterator combined with LHS within the FRAMES 3MRA Version 2.0 modeling system (see Figure 1-1).

Practical Limitations in Investigating Higher Dimensional Analyses

It is emphasized that the complexity of 3MRA and resulting computational burden imposed by a higher dimensional analysis for ISE and SME are quite significant, as demonstrated for the simplest case of a pseudo 2nd-order analysis. It is further emphasized that there are no known applications of such an approach for model orders and application scales of the magnitude embodied by 3MRA for the national assessment. The likelihood of providing useful, accurate information in this approach in the near-term would necessarily rely on some form of variance reduction technique, such as LHS, to be practically achieved for multiple chemicals. The solution approach most feasible in the near-term would be use of some form of LHS coupled with a screening-level based perturbation analysis to ascertain uncertainty in probabilities of protection for various percentiles of population protection. Such an approach would, of course, need to be informed by knowledge of linearity and non-monotonicity of 3MRA model behavior to establish credibility.

In summary, while the generic, underlying science methodology of 3MRA (Marin *et al.*, 1999, 2003) has implied the feasibility of use of higher-dimensional analysis to separate ISE empirical uncertainty from variability, in its general formulation for national scale assessments, it is a non-trivial exercise, extremely computationally demanding, and is currently not readily accomplished, nor feasible based on available input data.

ORD will certainly be engaged in the future in evaluating the efficacy of these approaches for such national assessment strategies over time. NRC (1994), EPA (1996a), and this author offer caution in any attempts to establish such higher dimensional analysis as a prerequisite for regulatory-based modeling, thereby limiting regulators in their ability to make decisions through professional judgment on how best to protect human health and the environment in the interim. More important in the interim is practical recognition of the underlying quality of the constituent elements of the modeling system, UA/SA results, and the always-present limitations of the 3MRA modeling system, or any model for that matter, in "generating truth" with respect to the question(s) posed. 3MRA is a model, and as such, it is not perfect, nor will it ever be perfected.

9.4 Long-Term Research Supporting Future Model Evaluation Capabilities

As described in Sections 2.7.3, 2.7.6, 6.5, and 6.8, the intent of the National Exposure Research Laboratory's Ecosystems Research Division (ERD) over-time is to further evaluate the efficacy of: simpler SA screening methods; fully quantitative, variance-based SA global methods; and alternative stochastic sampling methods. These efforts would be constructed within the specific context of the model evaluation problem, with tasking to be undertaken to serve longer-term needs in investigating uncertainty and sensitivity of 3MRA and other EPA models.

To further evaluate the sensitivity and uncertainty of 3MRA, and to facilitate similar investigation of other high order models and modeling systems, it is anticipated that one or two screening level techniques will eventually be developed for inclusion into 3MRA Version 2.0. These will likely include application of Morris' OAT design (1st and 2nd order) and Andres' Iterated Fractional Factorial Design (IFFD) (Campolongo *et al.*, 2000b). Incorporation and implementation of higher-order, global, variance-based techniques (e.g., FAST, Sobol's Method) will also be undertaken over-time by ORD (Chan *et al.*, 2000b). As an example of ERD's current emphasis on local sensitivity analysis and parameter estimation techniques, a project is currently underway to both merge and modernize existing inverse problem formulation technologies (i.e., UCODE, PEST), under the auspices of object oriented software design for suitable inclusion into FRAMES and other modeling frameworks. This project is being conducted jointly by USGS and ERD, and will result in creation of a new technology – the Joint Parameter IdenTification and Evaluation of Reliability (JUPITER).

ERD is strategically aligned in its long-term research goal to develop a widely available, easily accessed suite of tools and computing capabilities that will allow sophisticated parameter estimation, sensitivity, and uncertainty methods to be applied to a wide range of environmental modeling problems, by a wide range of researchers and stakeholders. ERD, in general, seeks to advance these capabilities throughout the environmental modeling community in an effort to improve the quality assurance in environmental models, applications, and the decisions made based thereon.

Initiated by researchers at ERD, a pivotal development in the last two years for the family of Federal Agencies involved in various environmental modeling applications was the creation of the Multi-Agency Multimedia Modeling MOU, and an associated Workgroup for Uncertainty Assessment and Parameter Estimation (PE) for Multimedia Environmental Models. Two key MOU Workgroup activities of note to date are: (1) the development of a common Applications Programming Interface (API) for UA/SA/PE tools across Agencies; and (2) the infusion of national and international expertise in UA/SA/PE into Federal Research Programs. Co-organized by ERD, an upcoming International Workshop on Uncertainty, Sensitivity, and Parameter Estimation for Multimedia Models will be held in August 2003. A prospectus of the workshop is provided in Appendix E to assist the SAB in further understanding the increasing dedication at ORD to infuse research and development expertise into the critically important topical areas of UA, SA, and PE, and in general, model evaluation.