Higher Level Probabilistic Approaches - Applying Monte Carlo Modeling to EPA Risk Evaluations

Case Example: 1,4-Dioxane Supplemental Risk Evaluation

Franklyn Hall, B.S., Chemical Engineer,

Risk Assessment Branch 1

Existing Chemical Risk Assessment Division

Office of Pollution Prevention & Toxics

Office of Chemical Safety and Pollution Prevention

U.S. Environmental Protection Agency

General Outline of this Presentation

1. Background on the incorporation of Monte Carlo methods in EPA Risk Evaluations.
2. 1,4-Dioxane Supplemental Risk Evaluation peer review process, focusing on feedback received on Monte Carlo methods and the associate modifications made by EPA.
3. Improvements in Monte Carlo methods from the peer review process and key takeaways for the use of Monte Carlo in EPA Risk Evaluations.

Probabilistic Modeling

- A type of stochastic modeling that allows for random variation of multiple inputs where the random variation may be probabilistically influenced by historical data.
- Monte Carlo methods fall under the umbrella of probabilistic modeling. The model is run multiple times, and each run uses different input values and generates different output values.
- For the 1,4-Dioxane Supplemental Risk Evaluation, EPA implemented Monte Carlo methods in a Microsoft Excel-based model using a Monte Carlo add-in tool called @Risk.

Uncertainty and Variability

- Uncertainty refers to lack of knowledge about specific factors, parameters, or models.
  - **Parameter uncertainty** (measurement errors, sampling errors, systematic errors)
  - **Model uncertainty** (uncertainty due to necessary simplification of real-world processes, use of inappropriate surrogate variables)
  - **Scenario uncertainty** (descriptive errors, aggregation errors, errors in professional judgement)
- Variability refers to observed differences attributable to *true heterogeneity* or diversity in a population or exposure parameter.
  - **Spatial variability** refers to differences that may occur because of location.
  - **Temporal variability** refers to variations over time, whether long- or short-term.
  - **Inter- and intra- site variability** refers to differences between sites and variability in the day-to-day activities at any particular site.

Benefits of Monte Carlo Methods

- Monte Carlo modeling offers various benefits to EPA’s Risk Evaluations.
  - Incorporation of variability using distributions of input parameters.
  - Reduces uncertainty because the model outputs a range of results that are more likely to reflect the true distribution.
  - Elimination of potential bias from using only monitoring data, such as:
    - If the monitoring data are very site-specific, or
    - If the monitoring data only capture a portion of the process.
    - Incorporates different inputs for engineering controls which may not be captured by monitoring data or deterministic approach (e.g., dust capture efficiencies, ventilation rates, etc.).
- These benefits increase the confidence in the resulting release and/or exposure estimates.
- However, the results of Monte Carlo modeling are only as good as the data available to define the input parameters.

Implementation of the Monte Carlo Method

1. Define probability distributions for input parameters.
2. Generate a set of input values by randomly drawing a sample from each probability distribution.
3. Execute the deterministic model calculations.
4. Save the output results.
5. Repeat steps 2 through 4 through the appropriate number of iterations.
6. Aggregate the saved output results and calculate statistics.

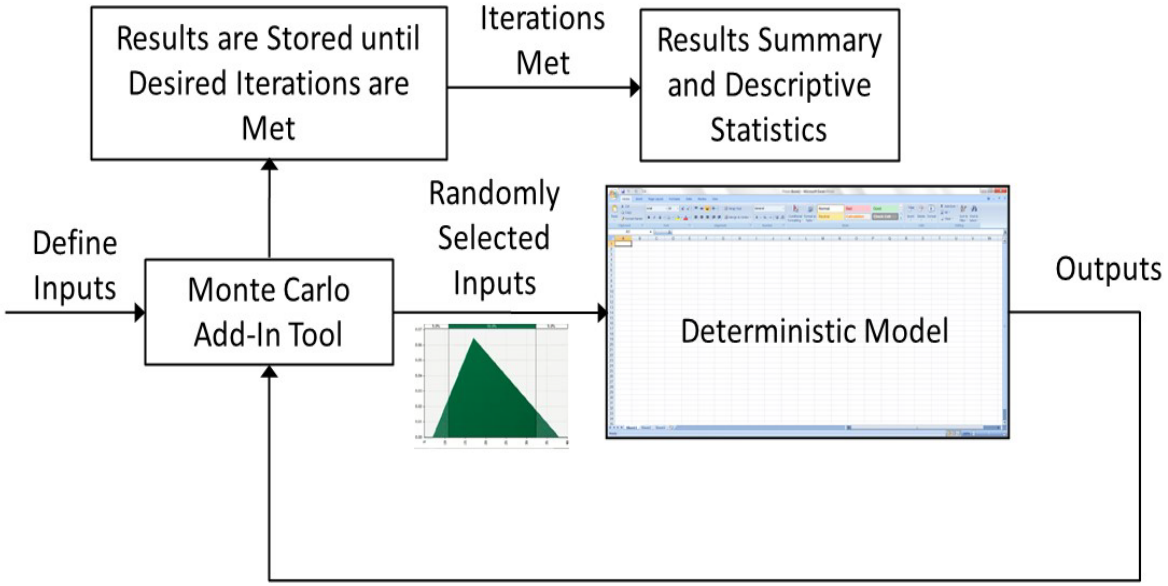


Figure 1. Flowchart of a Monte Carlo Method Implemented in a Microsoft Excel-Based Model Using a Monte Carlo Add-In Tool

Selecting Input Parameters for the Monte Carlo Model

- The selection of input parameters for probability distributions are largely informed by the availability of data (e.g., from GS, ESD, databases, literature, published chemical assessments).
  - If sufficient data are available, a distribution can be defined for the parameter.
  - If only a single values is known for the parameter, a distribution cannot be made.
- Input parameters may have a variety of probability distributions depending on the type of parameter and data available:
  - Uniform distribution, discrete distribution, triangular distribution, normal distribution, lognormal distribution.

Calculation Incorporating Input Parameter Distributions

- Deterministic calculation for operating hours associated with unloading containers of antifreeze:

$$OH_{cont\_unload} = \frac{(Q_{use} \times N_{jobs})}{3.79\ \frac{L}{gal} \times 1\ \frac{kg}{L} \times V_{cont} \times RATE_{fill}}$$

- The degree of stochasticity in this example is represented by the three blue input parameters for which distributions were used: number of jobs per day and container volume.
- Variability could be further incorporated if distributions were available for the other parameters, such as $RATE_{fill}$.

Parameter	Definition
OH_{cont_unload}	Duration of exposure for container unloading
Q_{use}	Daily use rate of antifreeze
N_{jobs}	Number of antifreeze jobs per day
V_{cont}	Container size
$RATE_{fill}$	Fill rate of containers

Calculation Incorporating Input Parameter Distributions (cont.)

- Deterministic calculation for the annual throughput of 1,4-dioxane in hydraulic fracturing:

$$Q_{dioxane\_site\_yr} = Q_{site\_yr} \times 3.79\ \frac{L}{gal} \times RHO_{fracturing\_fluid} \times F_{dioxane\_fracturing\_fluid}$$

- The degree of stochasticity in this example is represented by the two blue input parameters for which distributions were used: annual use rate per site of fracturing fluids containing 1,4-dioxane, and mass fraction of 1,4-dioxane in hydraulic fracturing fluid.
- Variability could be further incorporated if a distribution was available for the other parameter, $RHO_{fracturing\_fluid}$.

Parameter	Definition
$Q_{dioxane\_site\_yr}$	Annual throughput per site of 1,4-dioxane in hydraulic fracturing
Q_{site_yr}	Annual use rate per site of fracturing fluids containing 1,4-dioxane
$RHO_{fracturing\_fluid}$	Density of fracturing fluid
$F_{dioxane\_fracturing\_fluid}$	Mass fraction of 1,4-dioxane in hydraulic fracturing fluid

Calculation Incorporating Input Parameter Distributions (cont.)

- Deterministic calculation for dust releases that are captured and controlled during the use of solid laundry detergents:

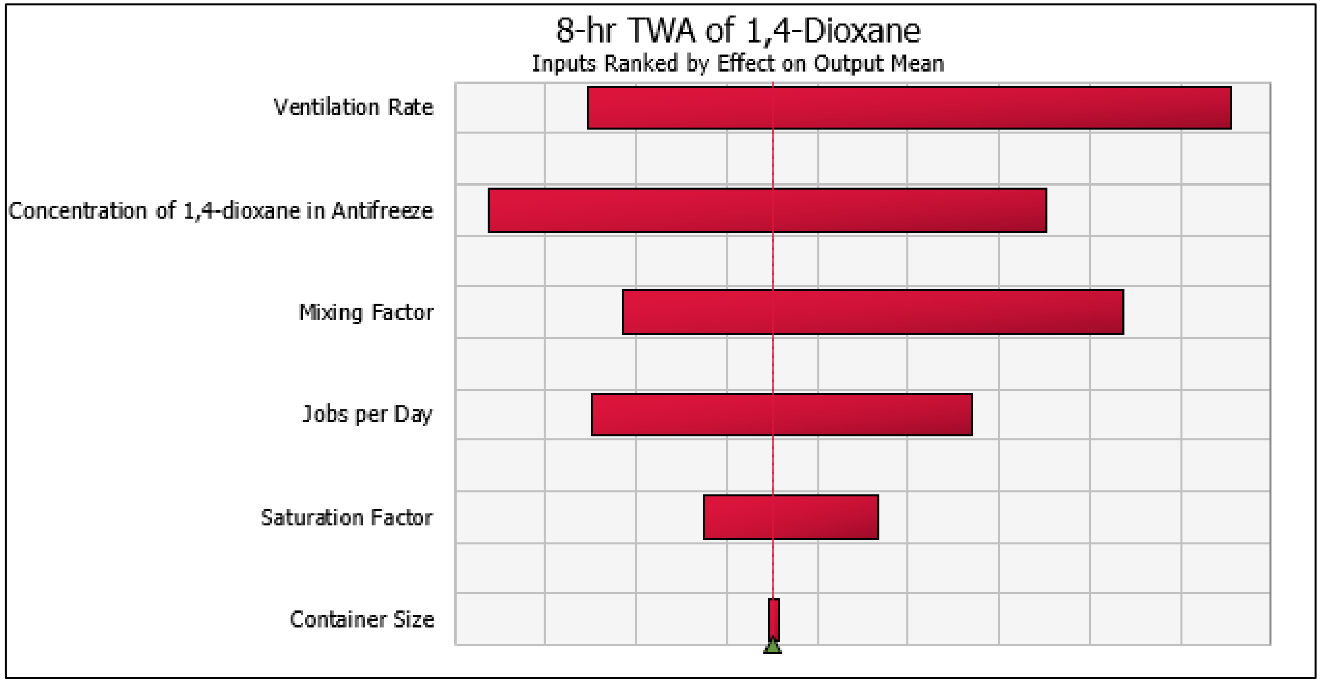
$$Release\_Daily_{dust} = Q_{facility\_day} \times F_{formulations\_dioxane} \times F_{dioxane\_laundry} \times F_{dust\_generation} \times F_{dust\_capture} \times F_{dust\_control}$$

- The degree of stochasticity in this example is represented by the six blue input parameters for which distributions were used.
- As all parameters have distributions, variability has been incorporated to the highest degree.

Parameter	Definition
$Q_{facility\_day}$	Daily use rate of powder laundry detergents
$F_{formulations\_dioxane}$	Fraction of laundry detergents containing 1,4-dioxane
$F_{dioxane\_laundry}$	Mass fraction of 1,4-dioxane in laundry detergent
$F_{dust\_generation}$	Fraction of chemical lost during transfer of solid powders
$F_{dust\_capture}$	Capture efficiency for dust capture methods
$F_{dust\_control}$	Control efficiency for dust control methods

Addition of Discussion on Sensitivity

- Peer reviewers recommended to include sensitivity analysis so that parameters that are driving releases/exposures could be identified, along with potential outliers skewing the data.
- This tornado chart shows the impact of the input parameters on 8-hour TWA exposure during use of antifreeze.
  - The larger the bar, the more sensitive the daily release result is on the input value.
- The largest impact parameters were identified to ensure distributions were available. All the inputs listed in this example have distributions.



1,4 Dioxane Charge Questions for the SACC

- During the peer review meeting, EPA presented charge questions to receive targeted feedback on the risk evaluation.
- In Charge Question 1, EPA specifically requested feedback on Monte Carlo methods used in the risk evaluation, focusing on:
    - Monte Carlo method development,
    - Uncertainty and variability associated with Monte Carlo modeling,
    - Distribution shape (e.g., triangular, discrete, etc.),
    - Data used to define input parameter distributions, and
    - Application of Monte Carlo modeling to the risk evaluation.
- A key purpose for these specific charge questions was to receive peer review on the Monte Carlo method development process that could be applied to other models used in future risk evaluations.

Peer Review Recommendations on the Monte Carlo Methods

- Common recommendations:
  - Parameter distributions should be based on recent data, as opposed to old data.
  - Industry data should be used to inform parameter distributions.
  - Specific scenarios should be modeled as opposed to generic industry-wide scenarios.

Summary of Improvements from Peer Review

- EPA created a new Monte Carlo model, eliminating the use of outdated monitoring data for dish soap and detergent.
- Peer review provided additional data sources for input parameters, allowing EPA to replace older or less representative datasets.
- EPA modified the risk evaluation to add transparency in input parameter discussions and discussions of model sensitivity.
- Overall, this resulted in improvements to the models, some of which increased the Weight of Scientific Evidence conclusion for the exposure estimates.

Current and Future Monte Carlo Use in Risk Evaluations

- EPA has incorporated Monte Carlo modeling into various release and exposure scenarios throughout various chemical risk evaluations:

Release Modeling

- Textile Dyes

Exposure Modeling

- Domestic Manufacture
- Chemical Processing
- Vapor Degreasing
- Aerosol Applications
- Use of Lubricants
- Waste Handling, Disposal, Treatment, and Recycling

Release & Exposure Modeling

- Import and Repackaging
- Incorporation into Paints and Coatings
- Incorporation into Articles
- Use of Laboratory Chemicals
- Dry Cleaning and Spot Cleaning
- Hydraulic Fracturing

Key Takeaways for the Use of Monte Carlo in Risk Evaluations

- Monte Carlo methods better incorporate variability in assessments and can be used with limited datasets.
- Monte Carlo methods can provide more representative exposure estimates compared to monitoring data, where monitoring data are limited, outdated, or biased.
- However, using more generic data or assumptions to fill data gaps can add uncertainty. As highlighted in the 1,4-dioxane peer review process, having more sources of data for model input parameters adds significant value to Monte Carlo models.
- Parameter variability improves the representativeness of estimates while increasing the confidence in risk characterization.