Probabilistic Risk Assessment

LTC Stephen Lewandowski, PhD, CPH

11/10/22

Lesson Objectives

- Explain rational for the use of probabilistic risk assessment
- Describe the tiered approach for risk assessment
- Understand the basics of Monte Carlo sampling
- Discuss the advantages and disadvantages of using probabilistic assessments compared to deterministic assessments

Probabilistic Risk Assessment

- "An analytical methodology used to incorporate information regarding uncertainty and/or variability into analyses to provide insight regarding the degree of certainty of a risk estimate and how the risk estimate varies among different members of an exposed population…"

A group of techniques that incorporate uncertainty and variability into risk assessments

EPA, 2014 Risk Assessment Forum White Paper

Variability and Uncertainty Review

- **Variability**: the inherent natural variation, diversity and heterogeneity across time, space or individuals within a population or lifestage
- **Uncertainty**: imperfect knowledge or a lack of precise knowledge of the physical world, either for specific values of interest or in the description of the system

Probabilistic Approach - Motivation

- Risk assessors, risk managers and others, particularly within the scientific and research divisions, have recognized that **more sophisticated statistical and mathematical approaches** could be utilized to **enhance the quality and accuracy of Agency risk assessments**
- Various stakeholders, inside and outside of the Agency, have called for a more comprehensive characterization of risks, including uncertainties, to **improve the protection of sensitive or vulnerable populations and lifestages**

EPA, 2014 Risk Assessment Forum White Paper

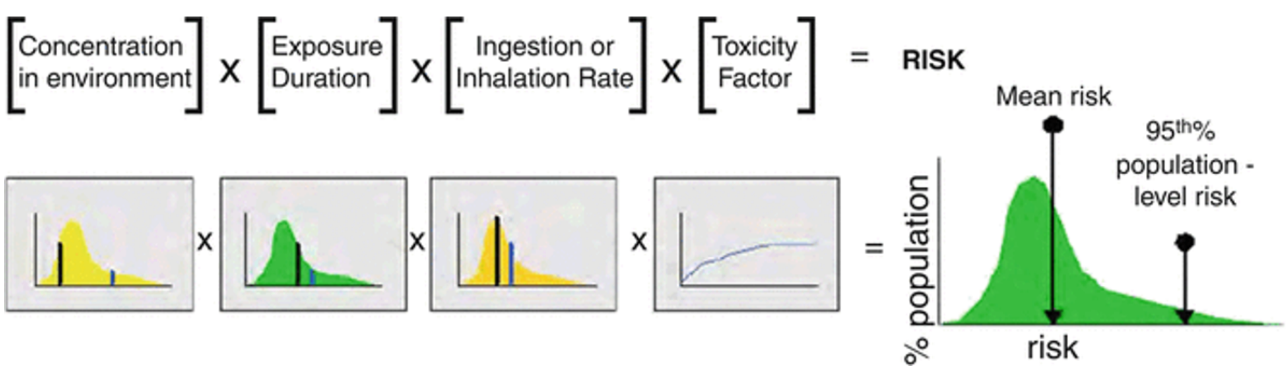
Deterministic: Point Estimate of Exposure Dose

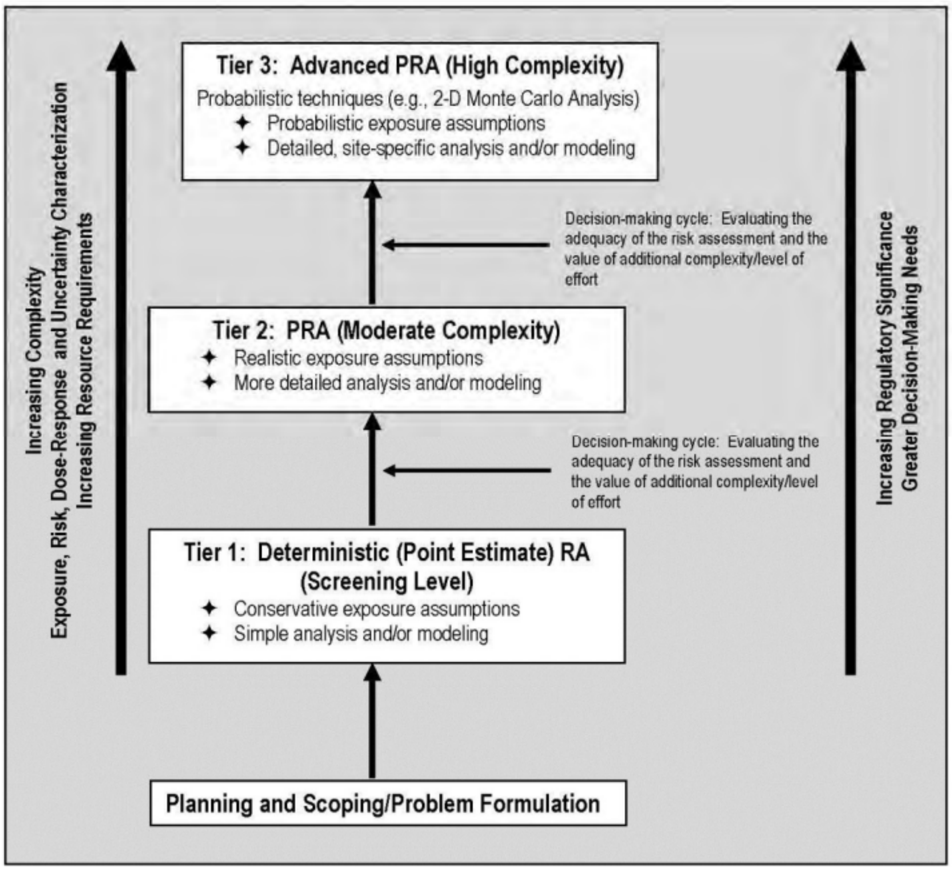
- Deterministic risk assessments express health risks as a single numerical estimate of risk
  - Assuming reasonable maximum exposure
    * Compounds unrealistically high estimates
    * Difficult to know/communicate the level of conservatism
  - Assuming average exposure
    * May present unacceptable risks
  - Mostly qualitative assessment of uncertainty and variability

Tiered Approach for Risk Assessment

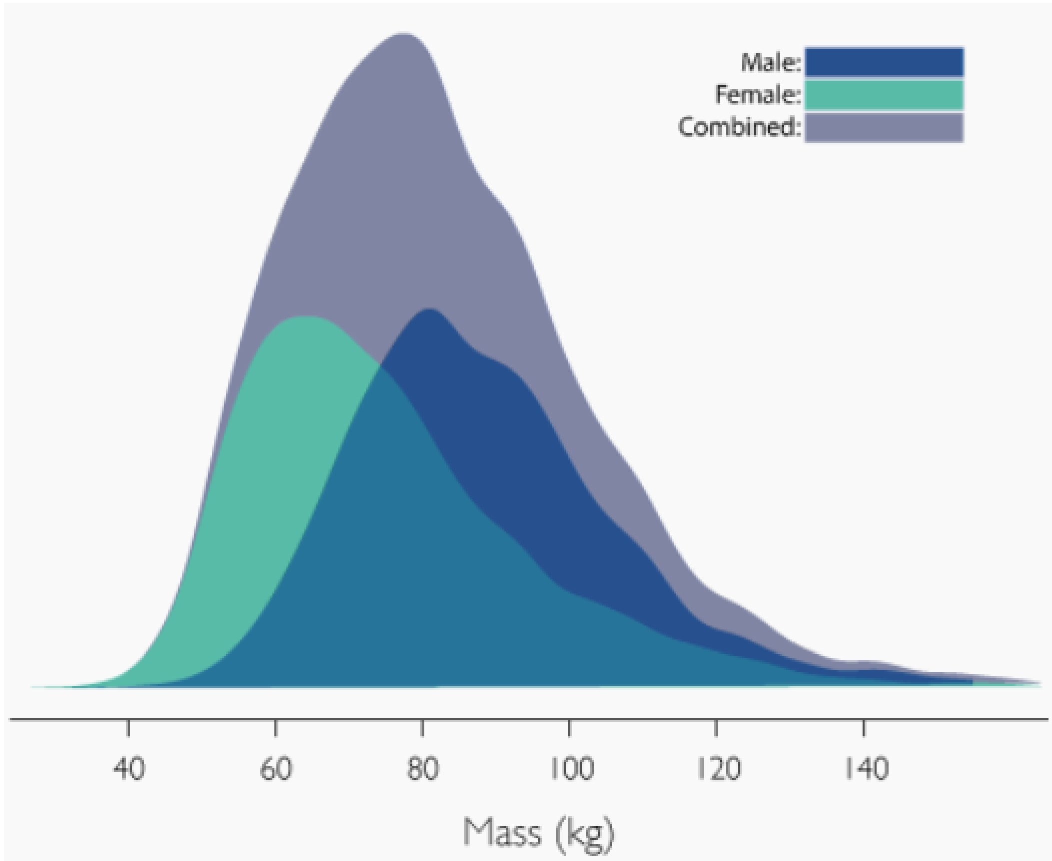
Assessments that are high in complexity and regulatory significance benefit from the application of probabilistic techniques

Probabilistic Approach




Tier 3: Advanced PRA (High Complexity)
Probabilistic techniques (e.g., 2-D Monte Carlo Analysis)
Probabilistic exposure assumptions
Detailed, site-specific analysis and/or modeling
Decision-making cycle: Evaluating the adequacy of the risk assessment and the value of additional complexity/level of effort
Tier 2: PRA (Moderate Complexity)
Realistic exposure assumptions
More detailed analysis and/or modeling
Decision-making cycle: Evaluating the adequacy of the risk assessment and the value of additional complexity/level of effort
Tier 1: Deterministic (Point Estimate) RA (Screening Level)
Conservative exposure assumptions
Simple analysis and/or modeling
Planning and Scoping/Problem Formulation
Increasing Complexity
Exposure, Risk, Dose-Response and Uncertainty Characterization
Increasing Resource Requirements
Increasing Regulatory Significance
Greater Decision-Making Needs

Example Parameter Variability



Source: PSU Open Design Lab (NHANES 2010) Adults 20-80 http://tools.openlab.psu.edu/tools/nhanes.htm

Probabilistic Example Scenario: AIRBORNE!

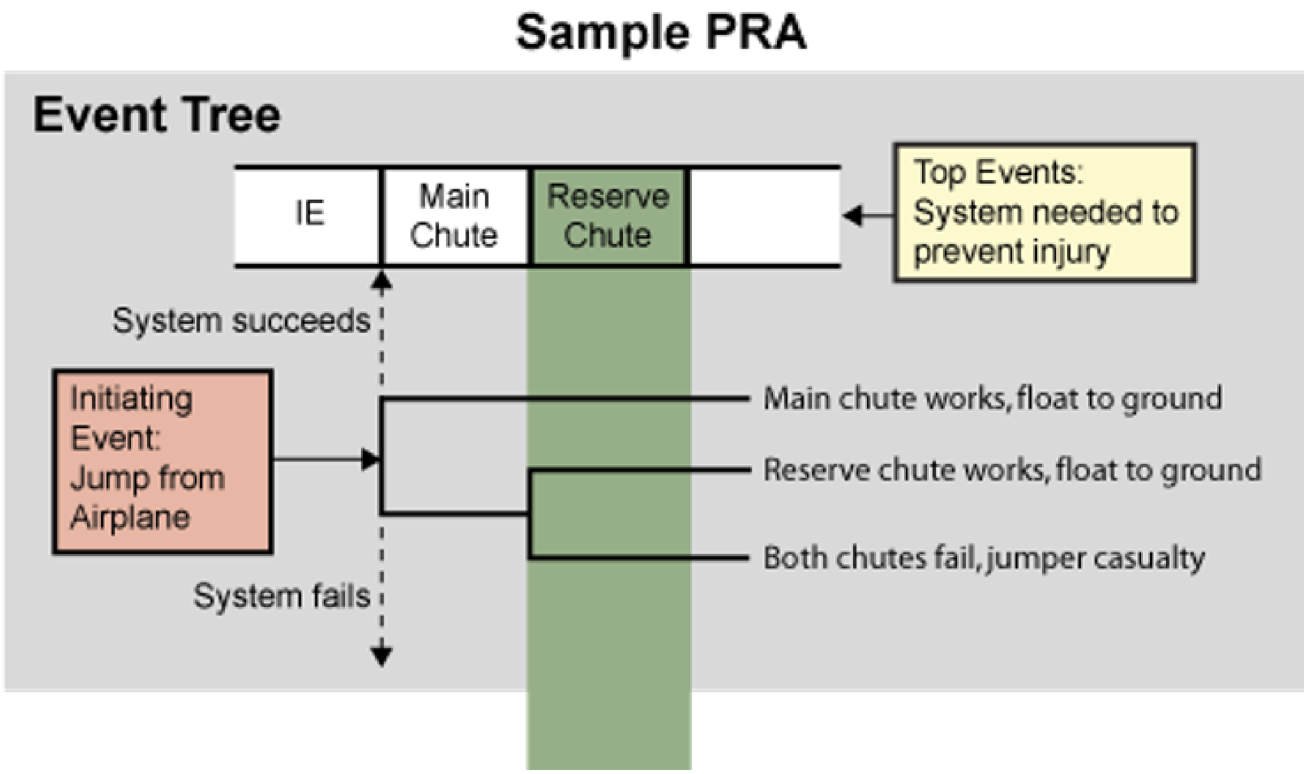
https://www.nrc.gov/about-nrc/regulatory/risk-informed/pra.html

Probabilistic Example Scenario: AIRBORNE!

Probabilistic Example: Nuclear Power Plant Operations (NRC)

Monte Carlo Simulation

- Mathematical technique used to estimate the possible outcomes of an uncertain event



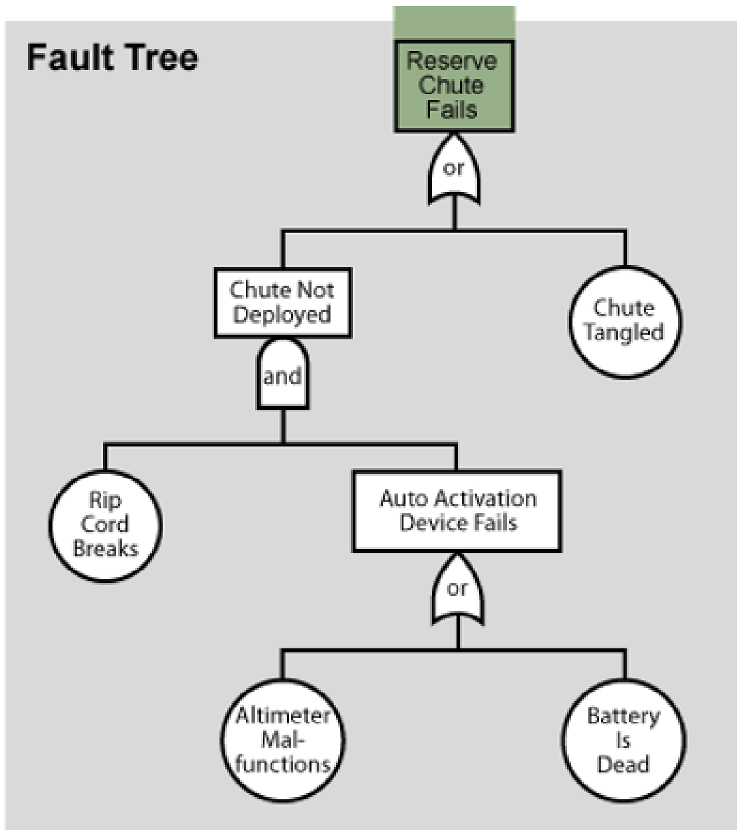
- Predicts a set of outcomes based on an estimated range of values versus a set of fixed input values
- Yields a range of possible outcomes with the probability of each result occurring
- Sensitivity analysis

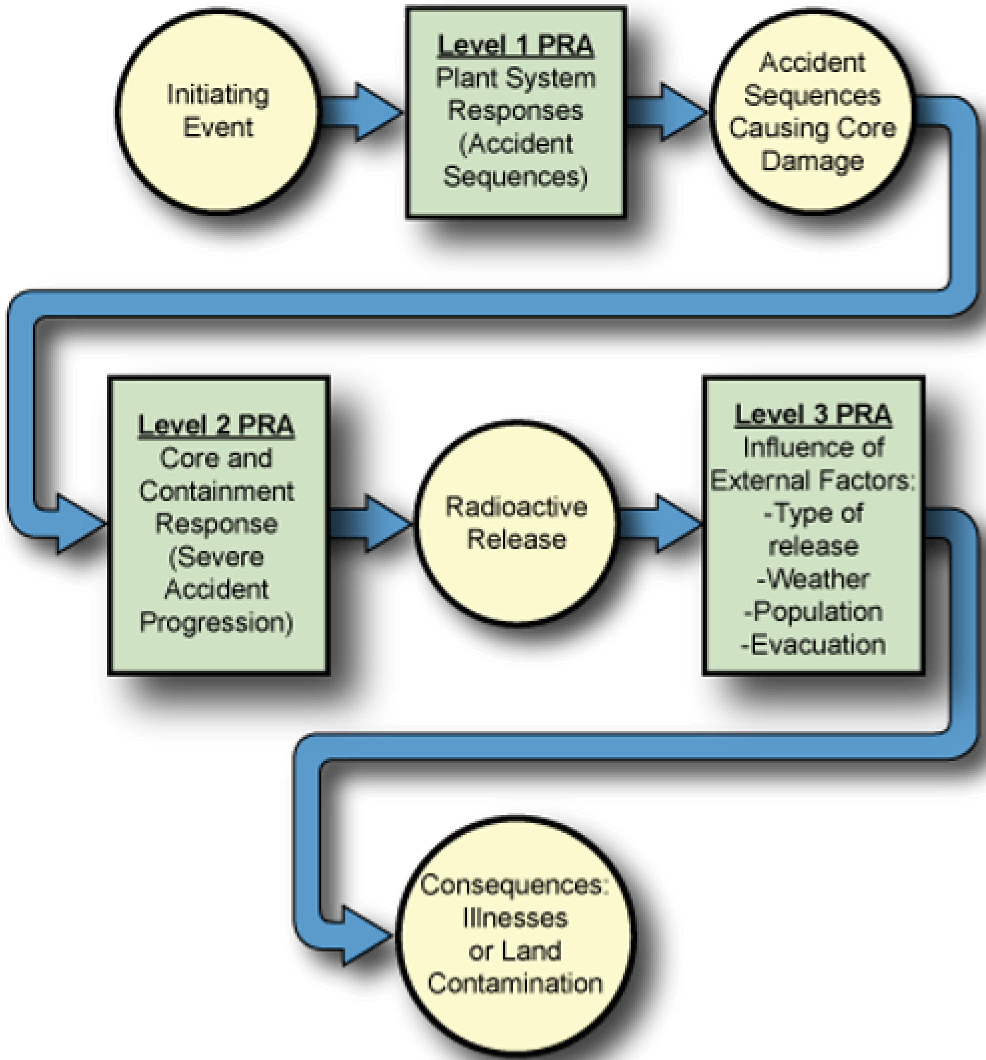
Video: What is Monte Carlo Simulation?

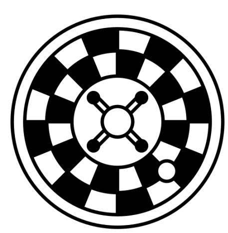
https://www.youtube.com/watch?v=7TqhmX92P6U"

Monte Carlo Simulation Applications

- Intergovernmental Panel on Climate Change: probability density function analysis of radiative forcing
- Computer graphics: Path/ray tracing renders a 3D scene by randomly tracing samples of possible light paths
- US Coast Guard: computer modeling software SAROPS calculates the probable locations of vessels during search and rescue operations


Fault Tree
Reserve Chute Fails
or
Chute Not Deployed
Chute Tangled
and
Rip Cord Breaks
Auto Activation Device Fails
or
Altimeter Mal-functions
Battery Is Dead

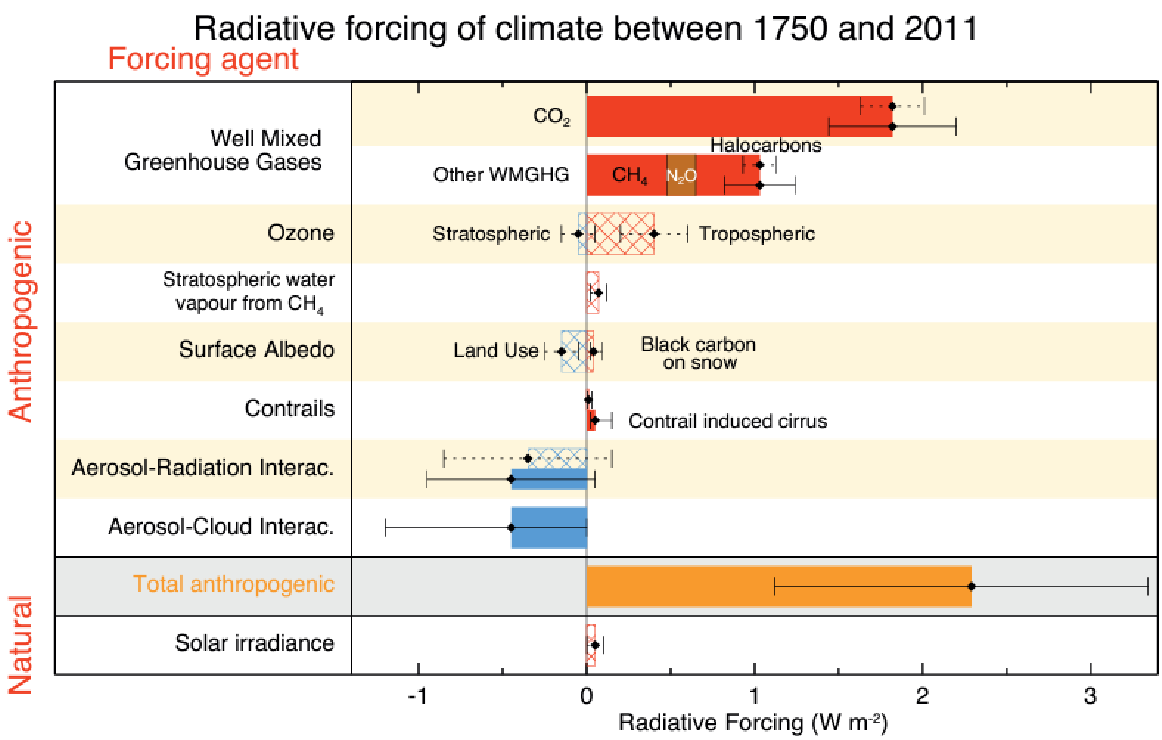

Initiating Event
Level 1 PRA
Plant System Responses (Accident Sequences)
Accident Sequences Causing Core Damage
Level 2 PRA
Core and Containment Response (Severe Accident Progression)
Radioactive Release
Level 3 PRA
Influence of External Factors:
-Type of release
-Weather
-Population
-Evacuation
Consequences: Illnesses or Land Contamination

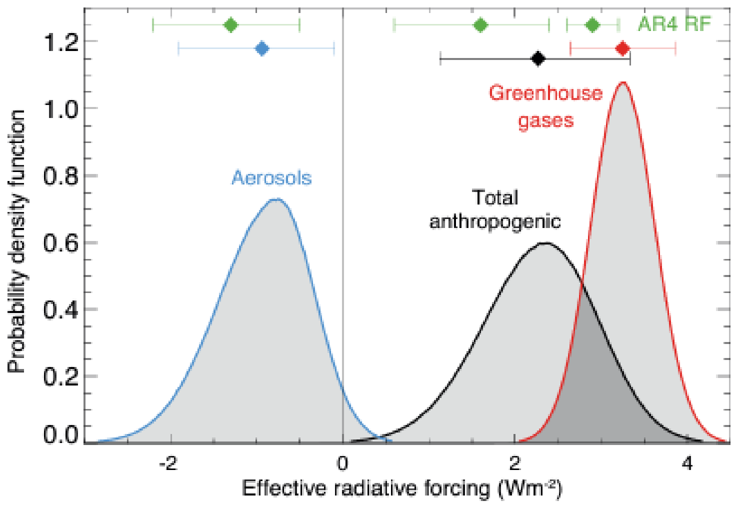


Monte Carlo Simulation: IPCC (WG1AR5)

	Evidence	Agreement	Confidence Level	Basis for Uncertainty Estimates (more certain / less certain)	Change in Under-standing Since AR4
Well-mixed greenhouse gases	Robust	High	*Very high*	Measured trends from different observed data sets and differences between radiative transfer models	No major change
Tropospheric ozone	Robust	Medium	*High*	Observed trends of ozone in the troposphere and model results for the industrial era/Differences between model estimates of RF	No major change
Stratospheric ozone	Robust	Medium	*High*	Observed trends in stratospheric and total ozone and model-ling of ozone depletion/Differences between estimates of RF	No major change
Stratospheric water vapour from CH_4	Robust	Low	*Medium*	Similarities in results of independent methods to estimate the RF/Known uncertainty in RF calculations	Elevated owing to more studies
Aerosol–radiation interactions	Robust	Medium	*High*	A large set of observations and converging independent estimates of RF/Differences between model estimates of RF	Elevated owing to more robust esti-mates from independent methods
Aerosol–cloud interactions	Medium	Low	*Low*	Variety of different observational evidence and modelling activities/ Spread in model estimates of ERF and differences between observations and model results	ERF in AR5 has a similar confidence level to RF in AR4

Monte Carlo Simulation: IPCC (WG1AR5)



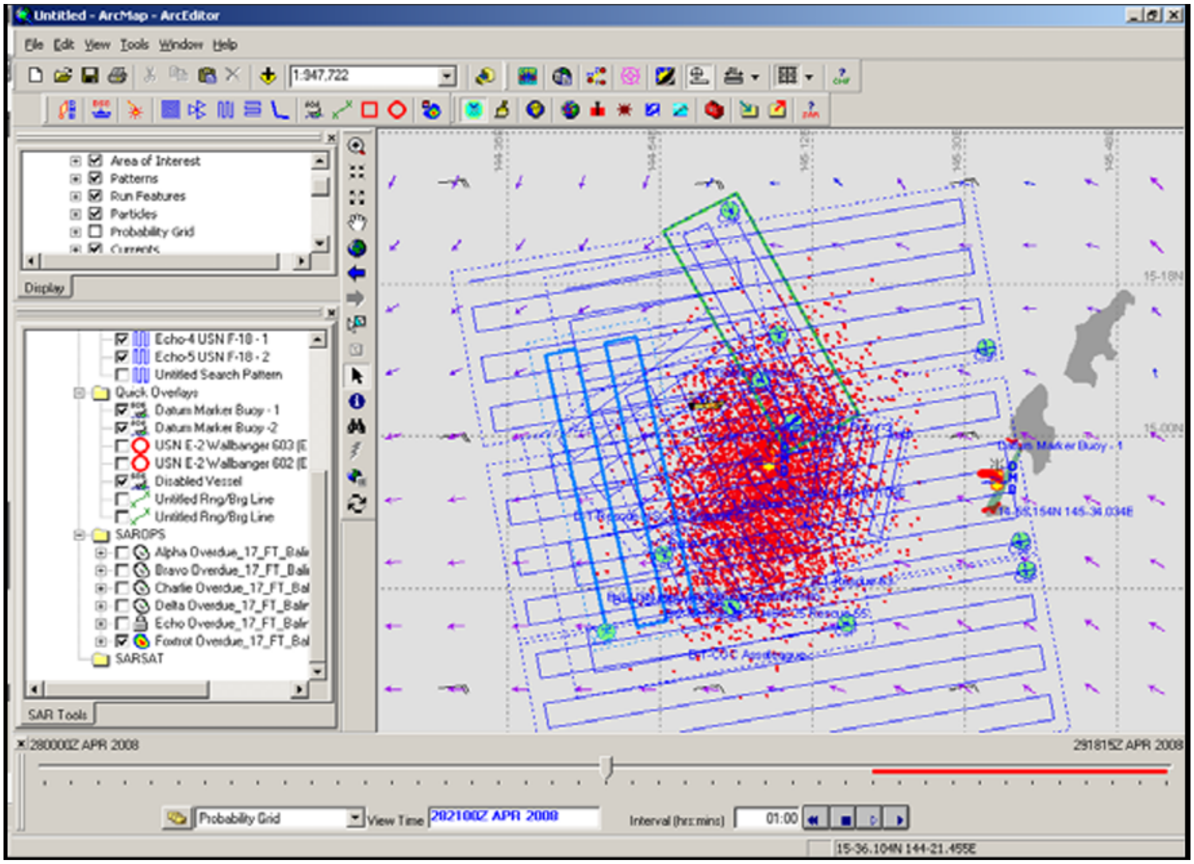


Monte Carlo Simulation: IPCC (WG1AR5)

Monte Carlo Simulation: SAROPS

- Search and Rescue Optimal Planning System (SAROPS)
- Software used by the U.S. Coast Guard for Maritime Search Planning
- SAROPS is a Monte Carlo based system that uses thousands of simulated particles generated by user inputs in a wizard based Graphical User Interface
  - Handle multiple scenarios and search object types
  - Model pre-distress motion and hazards
  - Account for the affects of previous searches

Monte Carlo Simulation: SAROPS Screen



Monte Carlo Simulation: Step 1

Set up the predictive model, identifying both the dependent variable to be predicted and the independent variables (also known as the input, risk or predictor variables) that will drive the prediction.

$$\text{Daily Intake} = \underbrace{\frac{\text{Concentration} * \text{Contact Rate}}{\text{Body Weight}}}_{\text{Intake (mg/kg/day)}} * \underbrace{\left(\frac{\text{Frequency} * \text{Duration}}{\text{Averaging Time}}\right)}_{\text{Time Weighted Exposure Factor}}$$

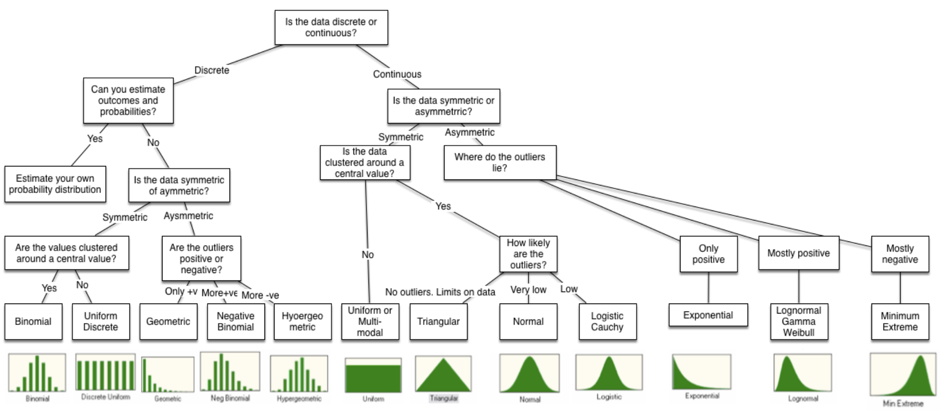
https://www.ibm.com/cloud/learn/monte-carlo-simulation

Monte Carlo Simulation: Step 2

- Specify probability distributions of the independent variables
- Use historical data and/or the analyst's subjective judgment to define a range of likely values and assign probability weights for each.
- Probability distribution: mathematical function that gives the probabilities of occurrence of different possible outcomes for an experiment
  - Type of distribution is useful when you need to know which outcomes are most likely, the spread of potential values, and the likelihood of different results
  - Selection of the appropriate distribution depends on the presence or absence of symmetry of the data set with respect to the mean value

https://www.ibm.com/cloud/learn/monte-carlo-simulation

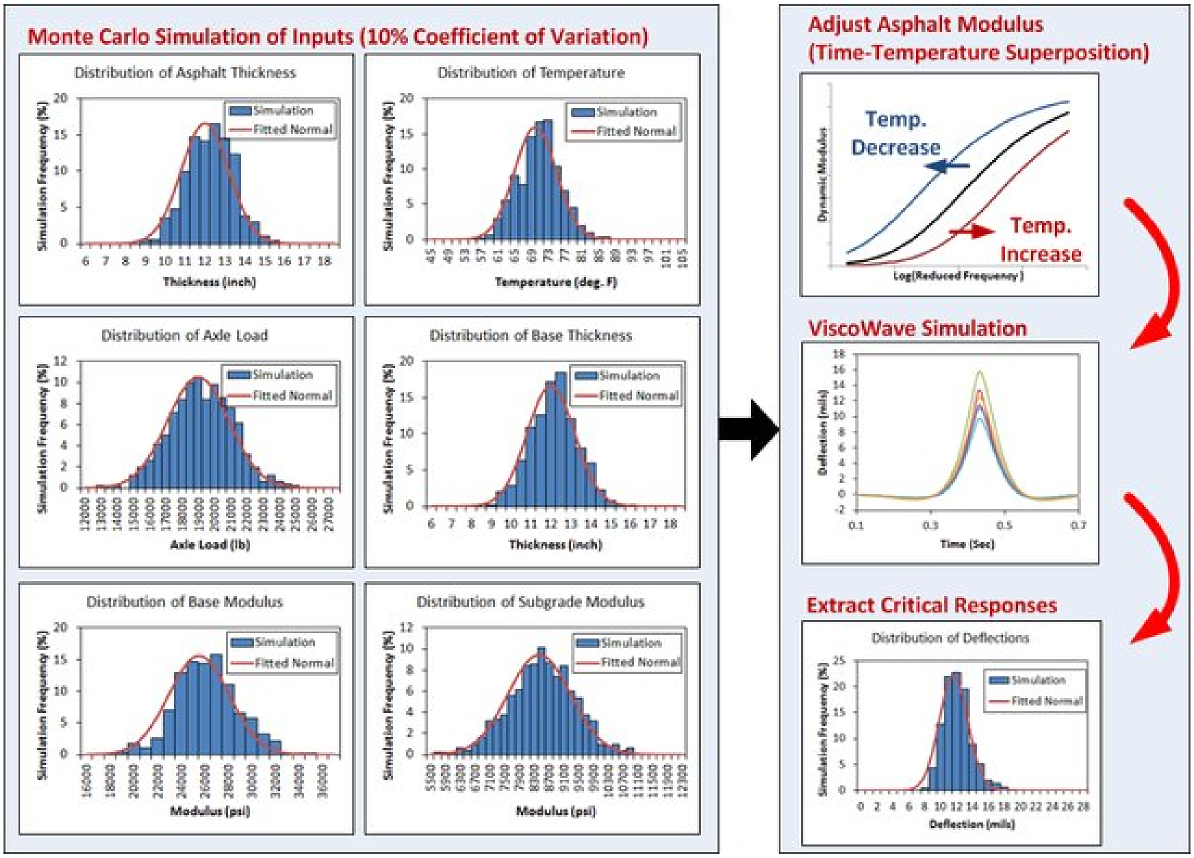
Monte Carlo Simulation: Step 2 (Distributions)



Monte Carlo Simulation: Step 3

- Run simulations repeatedly, generating random values of the independent variables
- Do this until enough results are gathered to make up a representative sample of the near infinite number of possible combinations

Monte Carlo Simulation: Step 3



https://www.researchgate.net/figure/Monte-Carlo-simulation-procedure-using-ViscoWave_fig7_311982513

Advantages and Disadvantages of Probabilistic Assessment

- **Advantages**
  - Flexibility for risk managers
  - Quantifies uncertainty and variability
  - Range of risk opposed to a single point estimate
  - Disadvantages may be offset by more informed decisions
- **Disadvantages**
  - Time
  - Resources
  - Greater technical expertise (analysts, communicators, and decision makers)

  - May require more data
  - Communicating results

Simulation Programs

- Commercial Software
  - Argo [BAH. Open Source]
  - @Risk [Palisade, \~\$1,900]
  - Crystal Ball [Oracle, ~$1,000]
  - Riskamp [Structured Data, LLC, ~$250]
- Programming Languages
  - R
  - Python

ARGO Simulation Tool

https://boozallen.github.io/argo/ https://github.com/boozallen/argo/wiki

Simulation Example: https://github.com/boozallen/argo/wiki/First-Argo-Simulation-Model

Practical Exercise: R Monte Carlo Simulation

```r
library(tidyverse)
library(mc2d)

# Define variables and distribution parameters
```

```r
c <- 0.00375 # concentration (mg/L)
  c_p1 <- 0.00375 # concentration mean
  c_p2 <- 0.001 # concentration s.d.
ir <- 1   # ingestion rate
  ir_p1 <- 0.5  # ingestion rate minimum
  ir_p2 <- 1  # ingestion rate mode
  ir_p3 <- 5  # ingestion rate maximum
ef <- 350  # exposure frequency (days)
  ef_p1 <- 200
  ef_p2 <- 350
  ef_p3 <- 365
ed <- 6  # exposure duration (days)
bw <- 70 # body mass (kg)
bw_p1 <- 70
bw_p2 <- 75
at <- 365*70  # averaging time (days)
csf <- 1.5  # cancer slope factor


## Deterministic

# Daily Intake = (C x IR x EF x ED x) / (BW x AT) Risk = CSF x (C x IR x EF x ED x) / (BW


daily_intake <-
  (c * ir * ef * ed) / (bw * at)

daily_intake
```

```
[1] 4.403131e-06
```

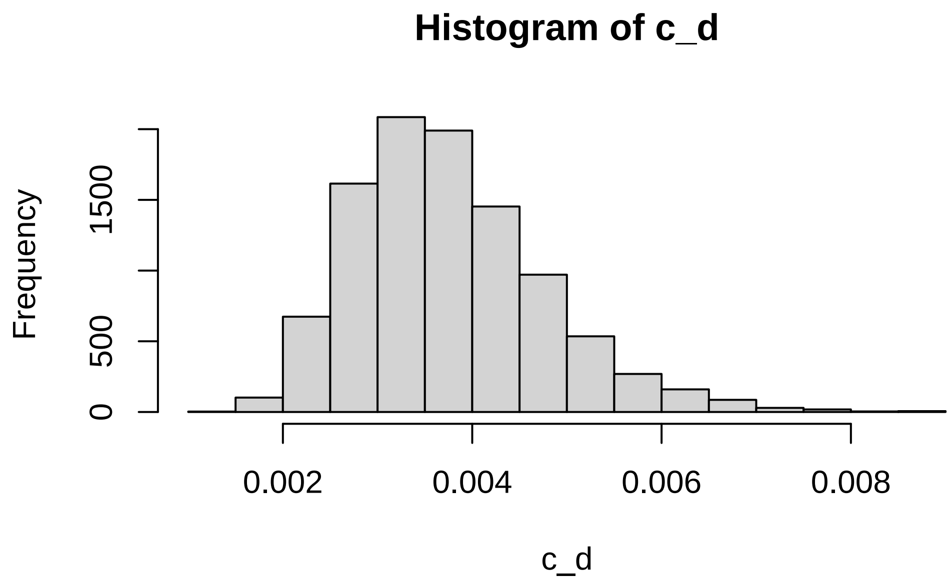
```r
risk <-
  csf * ((c * ir * ef * ed) / (bw * at))

risk
```

```
[1] 6.604697e-06
```

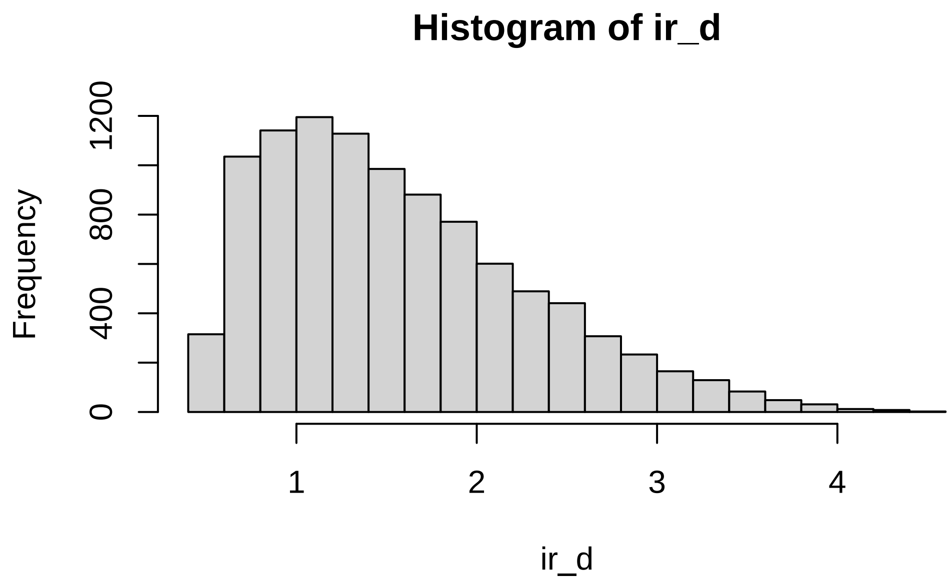
```r
## Probabilistic


# Concentration (C): Lognormal

# reparameterize for `rlnorm`
c_location <- log(c_p1^2 / sqrt(c_p2^2 + c_p1^2))
c_shape <- sqrt(log(1 + (c_p2^2 / c_p1^2)))

c_d <- rlnorm(10000, c_location, c_shape)
 c_d %>% as_tibble()
```

```
# A tibble: 10,000 x 1
     value
     <dbl>
 1 0.00602
 2 0.00299
 3 0.00265
 4 0.00320
 5 0.00480
 6 0.00350
 7 0.00379
 8 0.00491
 9 0.00256
10 0.00320
# ... with 9,990 more rows
```

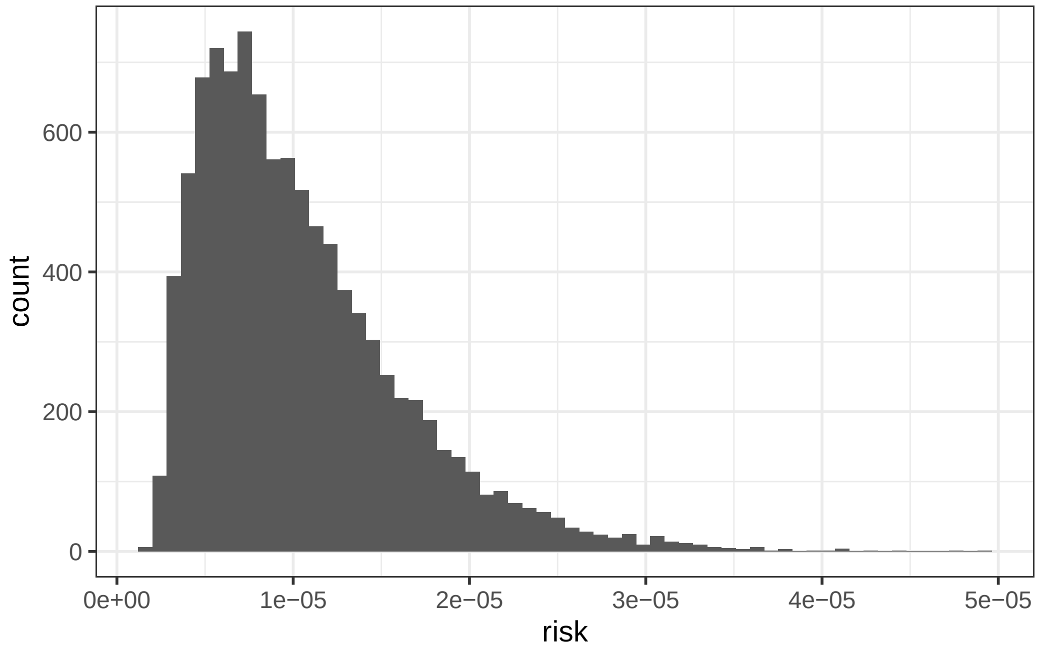
```r
hist(c_d)
```



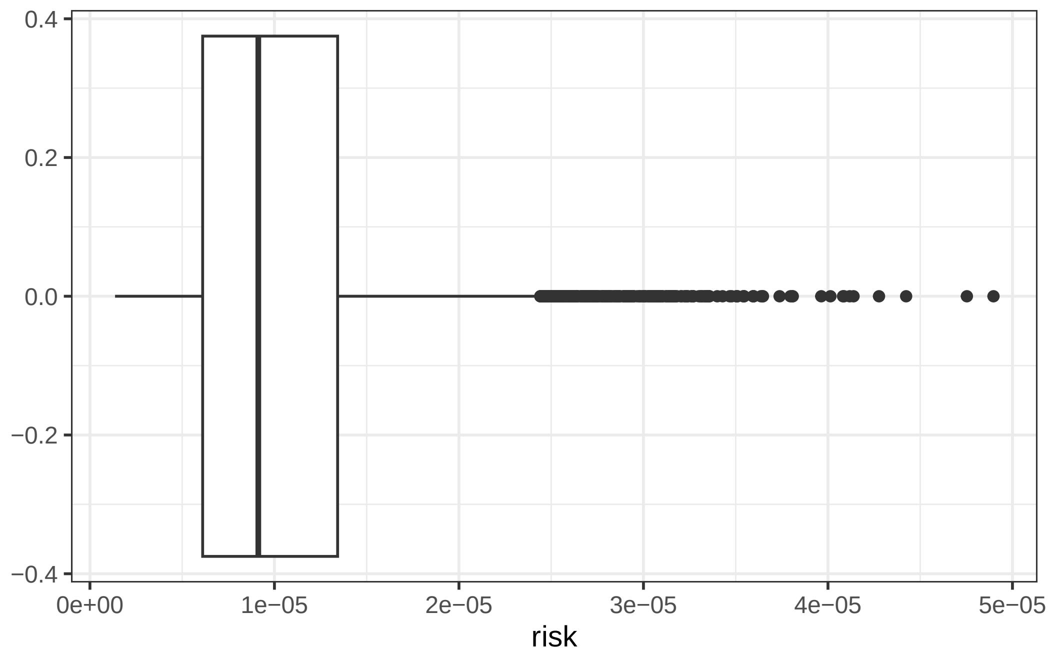

```r
# Intake Rate (IR): Beta-PERT
ir_d <- mc2d::rpert(n = 10000, min = ir_p1, mode = ir_p2, max = ir_p3, shape = 4)

# ir_d %>% as_tibble()
hist(ir_d)
```




```
# Replicate and Solve for Risk

risk_mc <-
  replicate(n = 10000, expr = {
    c_i = sample(c_d, size = 1)
    ir_i = sample(ir_d, size = 1)
    csf * ((c_i * ir_i * ef * ed) / (bw * at))
  })

risk_mc %>%
  as_tibble() %>%
  ggplot(aes(x = value)) +
  geom_histogram(bins = 60) +
  theme_bw() +
  xlab("risk")
```




```
risk_mc %>%
  as_tibble() %>%
  ggplot(aes(x = value)) +
  geom_boxplot() +
  theme_bw() +
  xlab("risk")
```




```r
summary(risk_mc)
```

```
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
1.361e-06 6.110e-06 9.125e-06 1.047e-05 1.342e-05 4.897e-05
```

EPA Case Study

Probabilistic Risk Analysis of Exposure to Polychlorinated Biphenyls via Consumption of Fish From a Contaminated Sediment Site

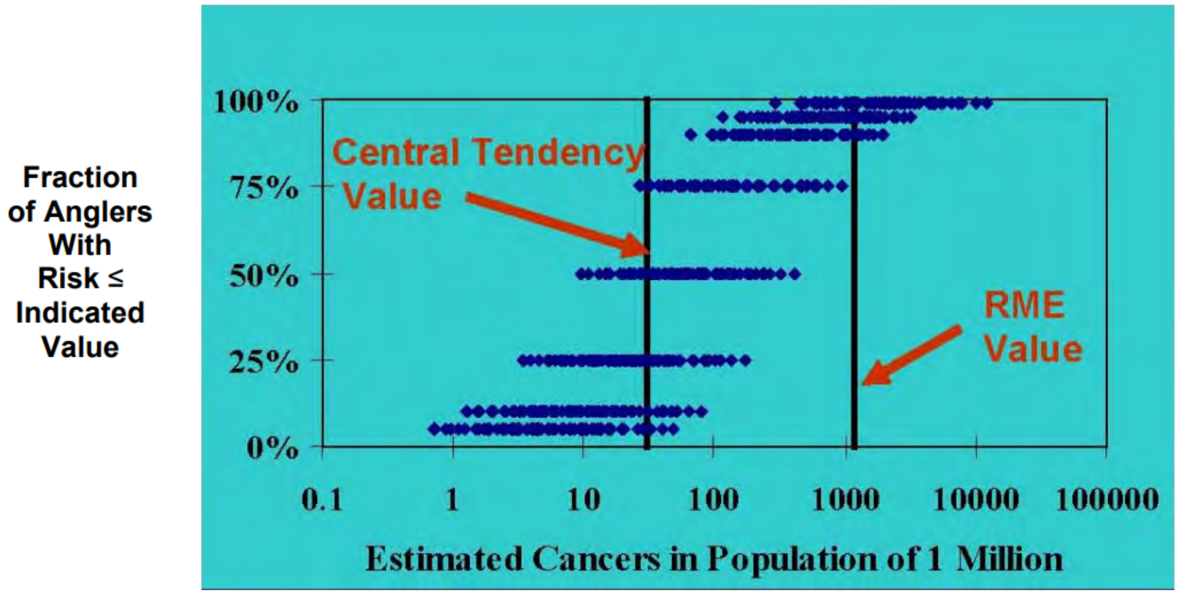
- EPA Region 2 conducted a preliminary deterministic HHRA at the Hudson River PCBs Superfund site
- Consumption of fish provided the highest exposure among relevant exposure pathways
- Cancer risks and noncancer health hazards exceeded regulatory benchmarks

EPA Case Study: Analysis

- Monte Carlo analysis of the variability of exposure as a function of the variability of individual exposure factors

- Mathematical models of the environmental fate, transport and bioaccumulation of PCBs in the Hudson River previously developed were used to forecast changes in PCB concentration over time
- Monte Carlo results consistent with deterministic results

EPA Case Study: Comparison



EPA Case Study: Comparison