Advanced Statistical Tools For Estimating Wildlife Exposure

Dwayne Moore

The Cadmus Group

Ottawa, Canada

Overview

- This presentation focuses on transfer of contaminants from prey to wildlife predators
- Two major issues:
  - Do the contaminant levels in the dataset represent the levels to which predators are exposed?
  - Is the dataset adequate to confidently estimate exposure?
- Briefly describe available techniques and present a case study

General Approach

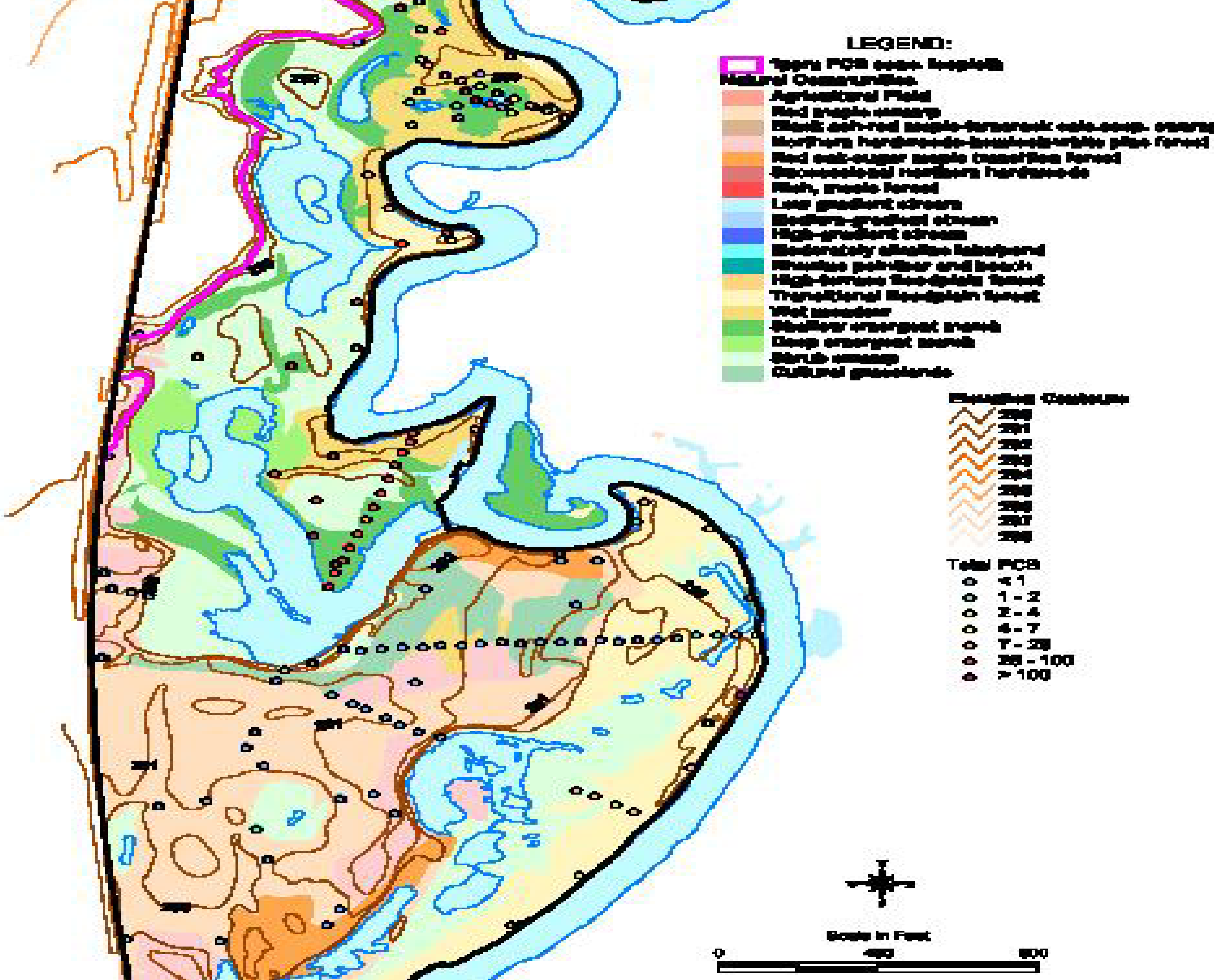
- Identify and characterize wildlife at risk
  - diet, foraging range, preferred habitats
  - life history, migration patterns, etc
- Select exposure model, identify critical input variables
- Review available data, where possible, collect data to fill critical gaps
- Parameterize exposure model, run analyses

Parameterizing Exposure Models

- Representativeness of data
  - many data sets biased to more contaminated areas
  - may provide uneven coverage of wildlife foraging area
  - data need to be manipulated to account for spatial and temporal foraging patterns of wildlife focal species
- Adequacy of data
  - costly to obtain large data sets
  - need to account for variability in contaminant levels and uncertainty introduced by small data sets
  - probabilistic techniques account for variability and uncertainty

Representativeness of Data

- Various spatial weighting techniques available
  - inverse distance weighting
  - Thiessen polygons
  - kriging
- Habitat weighting techniques can be used to account for wildlife foraging patterns
- Random walk and other models can be used to account for wildlife foraging patterns over time and space


LEGEND:
Total PCB
< 1
1 - 2
2 - 4
4 - 7
25 - 100
> 100
Scale in Feet
0
800

Types of Uncertainty in an Exposure Analysis

- Variability
- Incertitude
- Model uncertainty

Probabilistic Methods For Propagating Variability and Uncertainty

- Probability bounds
- Second-order Monte Carlo

Probability Bounds

How?

– specify what you are sure about
– establish bounds on probability distributions
– select dependencies (no assumption, independence, perfect, etc.)

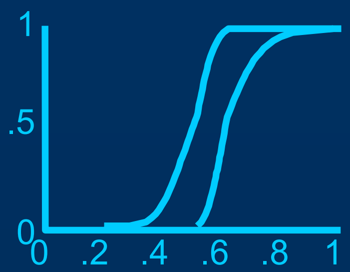
Why?

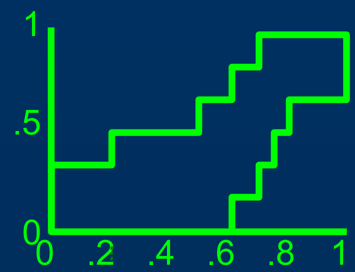
– accounts for uncertainty and variability
– puts bounds on Monte Carlo results
– bounds get narrower with better empirical information

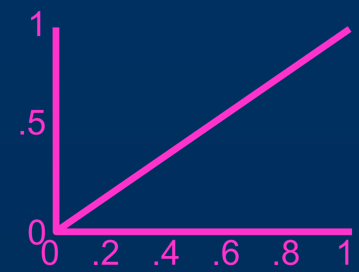
Why not?

– cannot handle second-order probabilities
– may not be able to use subtle information to tighten bounds
– optimum bounds expensive to compute when variables repeated

Probability Bounds







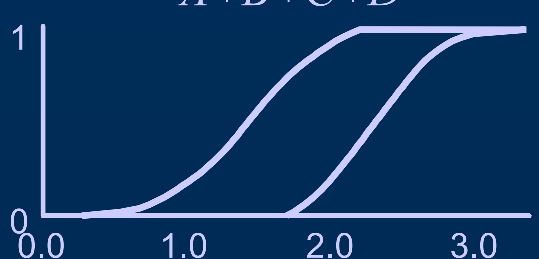
A = {lognormal, mean=[.5,.6], variance=[.001,.01]}

B = {min=0, max=.5, mode=.3}

C = {data = (.2, .5, .6, .7, .75, .8)}

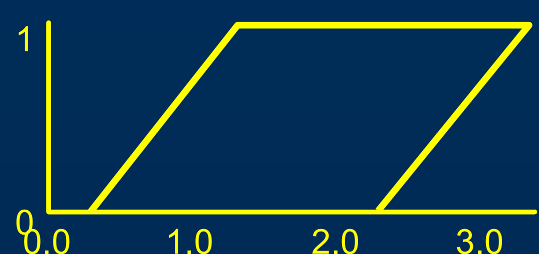
D = {shape = uniform, min=0, max=1}

$A+B+C+D$



Under independence

$A+B+C+D$



Without independence

Second-order Monte Carlo

How?

- let parameters of input distributions be distributions too
- nest Monte Carlo analyses
- either summarize with distribution of distributions,
- or condense output into a single distribution

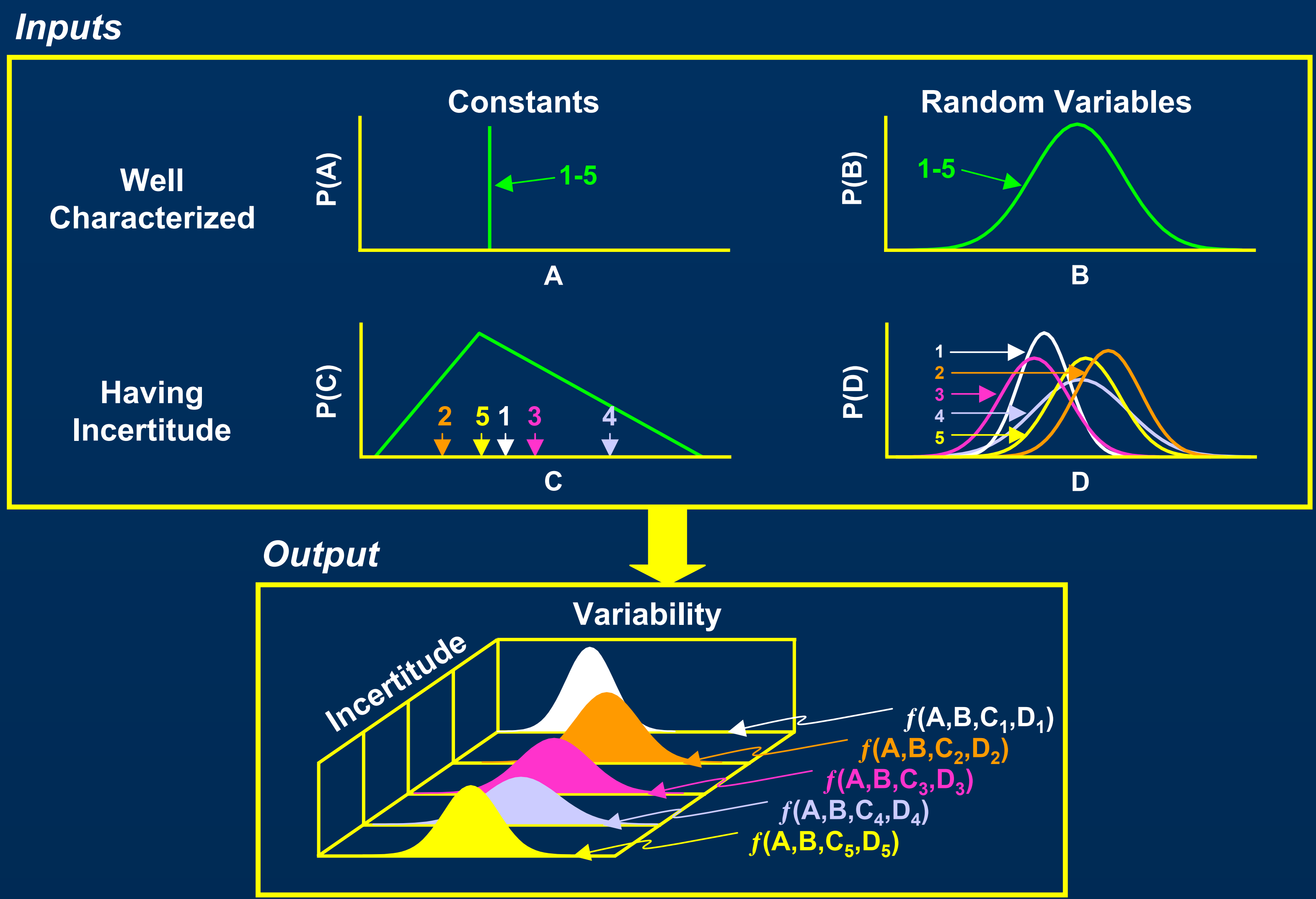
Why?

- acknowledges and accounts for the full extent of uncertainty

Why not?

- squared computational expense
- parameterizations can be difficult
- either results are cumbersome to interpret and explain
- or results confound ignorance with variability

Second-order Monte Carlo



Case Study

- Hypothetical case study
- Scenario
  - birds exposed to persistent chemical
  - goal: estimate distribution for chronic exposure
  - local spatial scale
  - some input variables well known, others not

Exposure Model

$$TDI = \frac{F_c\,([IR_w \; x \; C_w] + \{IR_f\,[C_{sed} \; x \; BSAF] + [C_{soil} \; x \; 0.02]\})}{BW}$$

where

TDI = Total daily intake (µg/kg bw/day)

F_c = Fraction of diet and drinking water that is contaminated (unitless)

IR_w = Intake rate for water (L/day)

IR_f = Intake rate for food (g/day)

C_w = Concentration in water (µg/L)

C_{sed} = Concentration in sediment (µg/g)

C_{soil} = Concentration in soil (µg/g)

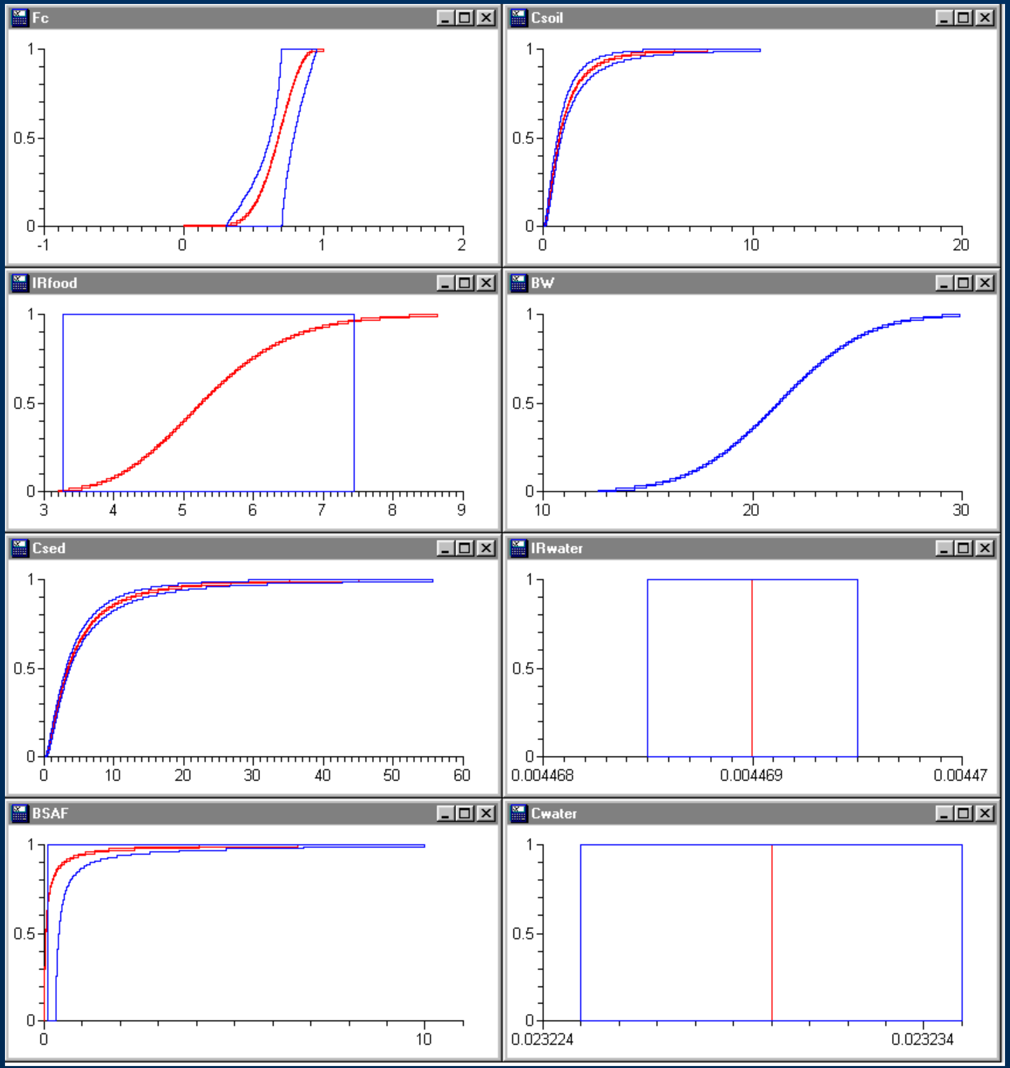
$BSAF$ = Biota-sediment accumulation factor (unitless)

BW = Body weight (kg)

Probability Bounds

- Several inputs have much incertitude
  - biota-sediment accumulation factor
  - food intake rate
- Several inputs have mostly variability
  - body weight
  - concentrations in soil, sediment and water

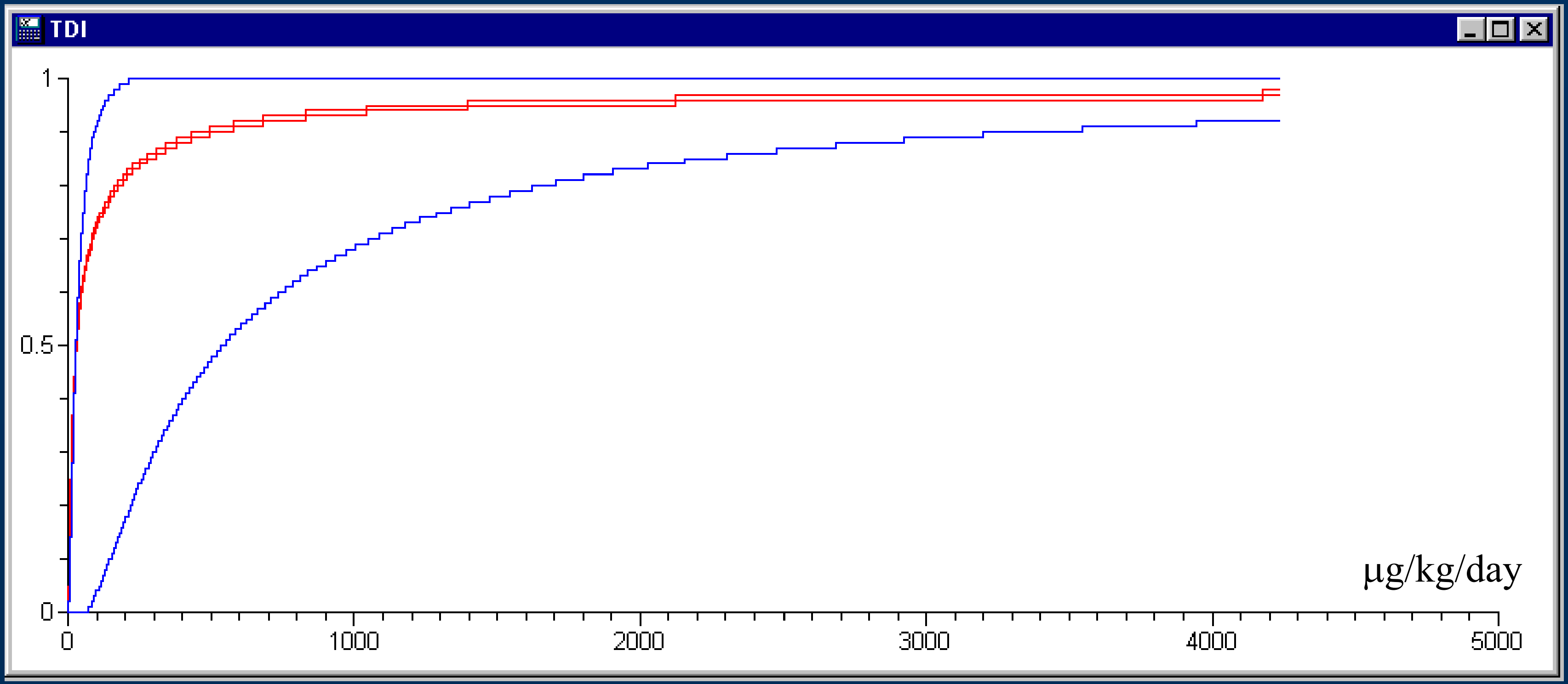
Inputs for Analysis



distributions used in Monte Carlo simulation shown in red

probability bounds shown in blue

Estimated Exposure



Estimated median	30	[27, 530]
90th percentile	450	[90, 3100]

Then What?

- Estimate risk by combining exposure distribution with effects information
- Risk curve is one line of evidence
- Also consider other lines of evidence (*e.g.*, biological surveys, *in situ* toxicity studies)
- Where remediation required, probabilistic model can be used to help determine cleanup levels