A Standardized Approach for Developing Probabilistic Exposure Factor Distributions

Randy L. Maddalena¹
* Thomas E. McKone¹,²
Michael D. Sohn¹,²

1. Lawrence Berkeley National Laboratory, USA
2. University of California, Berkeley, USA
Acknowledgements

This work was supported by the U.S. Environmental Protection Agency and carried out at Lawrence Berkeley National Laboratory through the U.S. Department of Energy.

Funding was provided by the Office of Emergency and Remedial Response (OERR) through Interagency Agreement #DW-899-38393-01-0.

http://eetd.lbl.gov/ied/era/
Overview

• Distributions for probabilistic risk assessment and the need for a “standardized approach”

• Exploring and communicating the critical attributes of an input distribution

• Balancing flexibility with consistency and transparency in the distribution development process
Why a Standardized Approach

• Probabilistic risk assessments are increasingly used in Europe and America

• Regulatory decision makers need to understand and judge the adequacy of Probabilistic Risk Assessments (PRA)

• The adequacy of a PRA depends on the adequacy of the input distributions
  — The 7 critical attributes of a distribution
Critical attributes of a distribution

• Three information categories defining the **task-specific adequacy** of a distribution:

  — The level of **influence** that the input has on the PRA model outcome (i.e., sensitivity analysis)

  — The original **data** or information used to develop the input distribution

  — The ability of the **distribution** to “simulate” the original and new data
Communicating the critical attributes

Sensitivity score

Amount of information

Quality of information

Relevance of information

Cross validation performance

Visual goodness-of-fit performance

Standard goodness-of-fit performance
Step 1: Developing archetypal distributions

1. Select exposure factor $E_{F_j}$
2. Acquire data for $E_{F_j}$
3. Identify robust demographic variables (e.g., gender, race, ...)
4. Partition $E_{F_j}$ data into demographically based subsets
5. Construct archetypal distribution, one for each subset of data

Demographically based archetypal distributions
A standard approach: Step 1

Step 1: Developing archetypal distributions

1. Select exposure factor $EF_j$
2. Acquire data for $EF_j$
3. Identify robust demographic variables (e.g., gender, race, ...)
4. Partition $EF_j$ data into demographically based subsets
5. Construct archetypal distribution, one for each subset of data

Illustrative Case Studies in this presentation:
1. Body weight (BW) & 2. Exposure Duration (ED)

Demographically based archetypal distributions
A standard approach: Step 1

Step 1: Developing archetypal distributions

1. Select exposure factor $EF_j$
2. Acquire data for $EF_j$
3. Identify robust demographic variables (e.g., gender, race, ...)
4. Partition $EF_j$ data into demographically based subsets
5. Construct archetypal distribution, one for each subset of data

Raw data acquired from national surveys of the US population

Demographically based archetypal distributions
A standard approach: Step 1

Step 1: Developing archetypal distributions

- Select exposure factor $EF_j$
- Acquire data for $EF_j$
- Identify robust demographic variables (e.g., gender, race, 
  …)
- Partition $EF_j$ data into demographically based subsets
- Construct archetypal distribution, one for each subset of data

Classification and Regression Tree (CART) software used but other methods are also appropriate

Demographically based archetypal distributions
A standard approach: Step 1

Step 1: Developing archetypal distributions

1. Select exposure factor $EF_j$
2. Acquire data for $EF_j$
3. Identify robust demographic variables (e.g., gender, race, …)
4. Partition $EF_j$ data into demographically based subsets
5. Construct archetypal distribution, one for each subset of data

$BW$ data split by gender and age

$ED$ data split by tenure (i.e., rent or own)

Demographically based archetypal distributions
A standard approach: Step 1

Step 1: Developing archetypal distributions

1. Select exposure factor $EF_j$
2. Acquire data for $EF_j$
3. Identify robust demographic variables (e.g., gender, race,…)
4. Partition $EF_j$ data into demographically based subsets
5. Construct archetypal distribution, one for each subset of data
6. Exploratory data analysis and goodness of fit to empirical cumulative distribution to select & parameterize models

Demographically based archetypal distributions
A standard approach: Step 1

Step 1: Developing archetypal distributions

- Select exposure factor $EF_j$
- Acquire data for $EF_j$
- Identify robust demographic variables (e.g., gender, race,...)
- Partition $EF_j$ data into demographically based subsets
- Construct archetypal distribution, one for each subset of data

Demographically based archetypal distributions

$BW$s described by 3-parameter log normal model

$ED$s described by simple exponential model
A standard approach: Step 2

Step 2: Develop scenario specific distributions

Identify important characteristics of target population

Random samples from archetypal distributions

Select and parameterize distribution for $EF_j$ using realization of values from archetypal distributions

Final distribution of $EF_j$ for use in PRA

Demographically based archetypal distributions

Uncertainty loop for small samples

From Step 1:
**BW and Gender**

![Graph showing the distribution of reported or measured body weight (Kg) for different gender categories.](image-url)
**BW and Age**

Growth curves, U.S. Center for Disease Control (2000)

Adult Weight-for-Age data and curves constructed from U.S. National Data

---

**Male Body Weight Across All Age Categories**

Body Weight (kg)

Age (y)
Constructing \textit{BW} archetypes

Construct distributions by sampling from each age/gender categories

3-parameter Log Normal PDF typical across all age/gender categories

Distribution for 29 year old males
The site specific BW distribution

- Adult males in exposure district based on census data, $n = 3895$ adult males
- Appropriate number of BWs sampled from each archetypal distribution
- Sampled values combined to construct single distribution for site specific BW
- Repeat the process to characterize uncertainty and variability

Extreme value distribution
- mode = 75.4 ± 0.35
- scale = 13.1 ± 0.23
Tenure and the *ED* archetype

Likelihood of move for a given household estimated from exponential model using tenure-based moving rates
The site specific ED distribution

- Total occupied housing units = 1490 (renter = 362; owner = 1128)

- Ten year average annual geographic mobility rates and exponential model for each occupied home in district (renter = 11.2 ± 0.7; owner = 3.3 ± 0.32)

- Simulate total occupied housing units for 85 years

- Repeat process several times to estimate uncertainty

median = 9 to 13 y
95% ucl = 40 to 54 y
Comparison to other *ED* models

Distribution of Residential Exposure Duration

Cumulative Percentile

Exposure Duration (y)

- Distribution based on "Movers out of County"
- Johnson and Capel (1992)
- Price, Sample and Strieter (1992)
- Israeli and Nelson (1992)
- Sedman, Funk and Fountain (1998)

<table>
<thead>
<tr>
<th>mobility</th>
<th>mortality</th>
<th>current occupancy</th>
<th>tax records</th>
<th>demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Summary remarks

• Adequacy of a PRA is related to adequacy of inputs (among other things)
  —Seven-points of an adequate distribution

• Archetypal distributions provide consistency and flexibility

• Archetypal distributions provide a focus point for developing consensus and identifying data gaps
Summary remarks

• Variability, uncertainty and demographics
  —Relating “archetypal distributions” to specific population requires knowledge about the population and the area of the exposure district

• Many exposure factors lack sufficient data for this approach (i.e., soil ingestion) or require site specific information (i.e., exposure concentrations)