

Stochastic Human Exposure and Dose Simulation Model
For Multimedia, Multipathway Chemicals

SHEDS-Multimedia Model version 3 Technical Manual

DRAFT

June 14, 2007

Disclaimer: This report is currently undergoing EPA review and should not be considered final.

Prepared by: Valerie Zartarian¹, Graham Glen², Luther Smith², Jianping Xue¹

¹ US Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory

² Alion Science and Technology

Acknowledgments

We would like to thank Dana Vogel, Steve Nako, David Hrды, and Phillip Villanueva of the U.S. Environmental Protection Agency, Office of Pollution Prevention and Toxic Substances, Office of Pesticide Programs and Miles Okino of the U.S. Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory for providing comments on the draft technical manual.

TABLE OF CONTENTS

Acknowledgments.....	2
TABLE OF CONTENTS.....	3
LIST OF FIGURES	6
LIST OF TABLES	7
ACRONYMS AND ABBREVIATIONS	8
INTRODUCTION	9
Purpose of SHEDS-Multimedia	9
SHEDS-Multimedia Model Development History.....	10
SHEDS-Multimedia Version 3 Overview	11
Features and Advantages of SHEDS-Multimedia Version 3	14
QUALITY ASSURANCE (QA).....	16
Development QA.....	16
One-person, one-day simulation for code verification	16
MODEL STRUCTURE AND APPROACH	18
Defining Exposure Scenarios and Simulation Information.....	18
Creating a Simulated Individual.....	19
Age Groups	19
Assigning ages to individuals	20
Assigning other person-level variables	20
Generating Longitudinal Activity Patterns for a Simulated Individual.....	21
Selecting Activity Diaries	21
Simulating an Individual’s Contact with Scenario-Relevant Media	22
Contact Media.....	22
Determination of contact events	23
Generating Concentration Time Series for the Contact Media	25
Chemical usage patterns (applications).....	25
Media concentration generation methods	31
Generating Exposure Time Series for Individuals	35
SHEDS equations for exposure by pathway.....	35
New Exposure.....	37
SHEDS Exposure Adjustments.....	42

Calculation of Dose Time Series for the Individual	45
Extract Daily Statistics from Exposure or Dose Time Series	46
Generating Population Variability Estimates of Exposure or Dose	46
MODEL INPUTS	48
Stochastic Sampling	49
Probability vectors	49
Continuous Probability Distributions	50
Sampling frequency	56
General Issues.....	57
Units.....	57
Formats	57
Demo File and multiple distributions for one variable	58
Simulation period.....	58
Chemical Usage-Related Inputs	59
Application-specific inputs	59
Co-Occurrence Factors	61
Activity Diary-Related Inputs	61
CHAD Diaries.....	61
Re-extracting data from CHAD for use in SHEDS	62
Pre-processing the CHAD files.....	62
Contact Probability-related Inputs.....	66
Concentration-Related Inputs	67
Decay-dispersion method.....	67
Post-application method.....	68
Time Series method	68
Exposure-Related Inputs	69
Physical exposure factors.....	69
Activity-related exposure factors	70
Transfer-related exposure factors	72
Dose Factors	73
OUTPUTS	74
Individual Time Series	74
Population Variability Distributions.....	75
SENSITIVITY ANALYSES	78
Pearson Correlation	79

Spearman Correlation	79
Stepwise Regression	80
The Percentile Scaling Approach	82
UNCERTAINTY ANALYSES	83
The Bootstrap Approach for Input Uncertainty Distributions	84
Uncertainty inputs and outputs	85
Sensitivity Analysis on Results from a Two-Stage Monte Carlo Model Run	87
REFERENCES	90
APPENDIX A: Terminology Used in SHEDS-Multimedia	A-1
APPENDIX B: The SHEDS Code Modules and their Functions	B-1
List of SAS Code Modules and Their Calling Sequence	B-4
APPENDIX C: Description of Residential Code Input Files.....	C-1
List of Input Files	C-1
Detailed Description of User-Specified Input Files	C-2
Variables file.....	C-2
Distributions file	C-4
Correlations file	C-6
Contactmedia file	C-6
ModelDates file.....	C-7
UserDates file	C-7
AppMedia file.....	C-8
Diarydetails file.....	C-9
Brief Description of Other Input Files.....	C-10
bwsa file	C-10
diarycohorts file	C-10
diaryevents file.....	C-10
diarylevels file.....	C-11
diarylocs file	C-11
diarymets file	C-11
diaryQARules file	C-12
diaryquest file	C-12
pop2000 file	C-12
APPENDIX D: Tables for the One-Person, One-Day Simulation for Code Verification.....	D-1
APPENDIX E: Selected SHEDS Variable Nomenclature Used in the Technical Manual.....	E-1

LIST OF FIGURES

Figure 1. Hypothetical exposure profile for an individual.....	11
Figure 2. Monte Carlo sampling in SHEDS.	12
Figure 3. Overview of SHEDS methodology.	14
Figure 4. SHEDS approach for simulating one-year activity patterns.....	22
Figure 5. Choosing application dates for a single treatment type.....	28
Figure 6. Co-occurrence of treatment types, choosing application dates for treatment type 2.....	29
Figure 7. Example time profile of dermal exposure from surface residues.....	35
Figure 8. Example time profile of GI tract exposure from surface residues.....	36
Figure 9. Example cumulative density function (CDF) plot from a SHEDS variability run.....	76
Figure 10. Hypothetical uncertainty analysis results: CDFs for low-, medium-, and high-dose populations, as ranked by their medians.....	86
Figure 11. Hypothetical uncertainty analysis results: CDFs for the 5th, 50th, and 95th percentiles from all simulated populations.....	86
Figure B-1. Conceptual overview of SHEDS-Multimedia version 3 module linkages.....	B-2

LIST OF TABLES

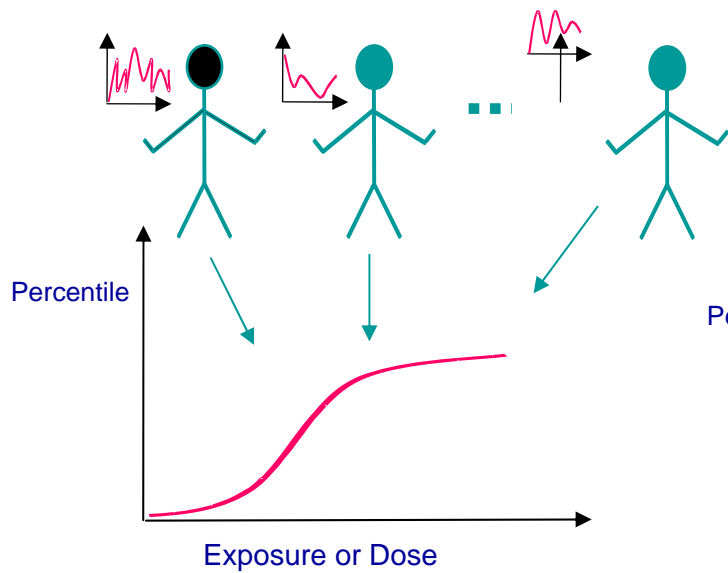
Table 1. Default list of contact media in SHEDS-multimedia version 3.	23
Table 2. Micro environments (micros) in SHEDS.	24
Table 3. Chemical application types in SHEDS.	26
Table 4. Example input file for concentration time series.	34
Table 5. Exposure pathways in SHEDS.	36
Table 6. Summary of SHEDS user-specified input variables (not all apply to all model runs).	48
Table 7. List of user-supplied probability vectors.	50
Table 8. SHEDS input probability distribution types.	51
Table 9. List of continuous variables in SHEDS.	55
Table 10. Rules for diary pre-processing in SHEDS.	63
Table 11. Variables that may be used in diary QA rules.	63
Table 12. Contact probability variables.	66
Table 13. Output files for SHEDS-multimedia version 3 variability run	77
Table 14. Hypothetical stepwise regression and percentile scaling sensitivity analysis results.	81
Table 15. Hypothetical Spearman correlation uncertainty analysis results.	88
Table 16. Hypothetical stepwise regression uncertainty analysis results.	89
Table C-1. Column definitions for the variable file.	C-3
Table C-2. Definitions for variables from the distributions and variables files.	C-3
Table C-3. Column definitions for the distributions file.	C-4
Table C-4. Column definitions for the correlations file.	C-6
Table C-5. Column definitions for the contactmedia file.	C-6
Table C-6. Default contact media in the contactmedia file.	C-7
Table C-7. Column definitions for the modeldates file.	C-7
Table C-8. Column definitions for the userdates file.	C-8
Table C-9. SAS variables in both appmedia files.	C-8
Table C-10. Contact probability definitions.	C-9
Table C-11. Column definitions for the diaryevents file.	C-10
Table C-12. Example diarylevels input file.	C-11
Table C-13. Current microenvironments used by SHEDS.	C-11
Table D-1. Inputs to exposure for one-day calculation.	D-2
Table D-2. Verification equations for SHEDS Q/A.	D-9
Table D-3. Output variables from the exposure calculations.	D-12

ACRONYMS AND ABBREVIATIONS

ADD – average daily absorbed dose
CCA –Chromated Copper Arsenate
CDFs – cumulative density functions
CHAD – Consolidated Human Activity Database
EOHSI – Environmental Occupational Health Sciences Institute
EPA – United States Environmental Protection Agency
EPOC – post-exercise oxygen consumption
FIFRA – Federal Insecticide, Fungicide, Rodenticide Act
g – gram
GI – gastrointestinal
GM – geometric mean
GSD – geometric standard deviation
GUI – graphical user interface
LBNL – Lawrence Berkeley National Laboratory
METS – metabolic equivalents
NERL – National Exposure Research Laboratory
NHANES – National Health and Nutrition Examination Survey
OPP – Office of Pesticide Programs
ORD –Office of Research and Development
PBPK – Physically-based pharmacokinetic
PDFs – probability density functions
PK- pharmacokinetic
SA – surface area
SAP – Scientific Advisory Panel
SHEDS – Stochastic Human Exposure and Dose Simulation
UPA – University Partnership Agreement
ug (in SAS printout or variable names) – microgram
yr – year

Variability

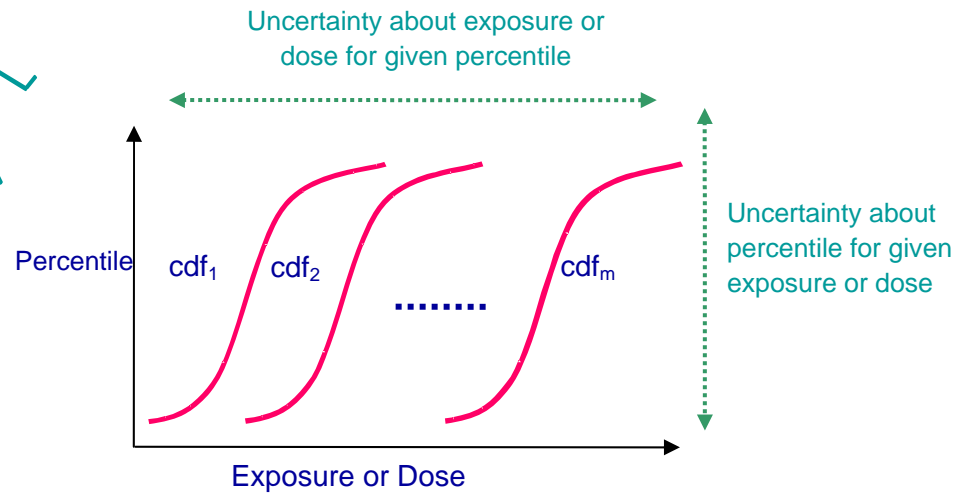
- temporal, spatial, or inter-individual differences in the value of an input



1-stage Monte Carlo

Uncertainty

- Lack of knowledge about
 - ◆ true value of a fixed but unknown quantity
 - ◆ true population distribution for variability



2-stage Monte Carlo

Figure 2. Monte Carlo sampling in SHEDS.

In a two-stage run (an “uncertainty run”), series of variability runs are performed, with the input distributions modified between each variability run to represent uncertainty in the input parameters of the variability runs. This allows for the construction of confidence intervals for various points on the variability distribution (see Uncertainty Analyses section for details). SHEDS may also be run in a special mode designed for sensitivity analysis, which estimates the influences of various input variables on exposure and dose statistics (see Sensitivity Analyses section for more details). Sensitivity and uncertainty runs are more computer resource intensive than variability runs, and may take several days to complete.

SHEDS-Multimedia version 3 uses the clock to set random number seeds and does not have the option of reproducing a prior run exactly; this is planned for SHEDS-Multimedia version 4. The ability to reproduce random number streams is essential to the operation of Sobol’s method of sensitivity analysis, which is also planned for version 4. Reproducible random number streams will be available even when sensitivity analysis is not being used. If two or more runs are made with the same input distributions, run settings, and random number streams, then the results will be identical.

The fundamental modeling unit in SHEDS is the individual. Since each individual is generated as a representative random sample, the rules for determining the characteristics and the exposure for each individual are simply repeated. One individual differs from another only because there is stochastic variation when sampling from distributions. For each individual in a SHEDS run, the following general steps are applied:

1. Randomly select the age, gender, and other demographic properties of interest, given the distribution of the target population.
2. Generate a longitudinal activity diary (using 8 CHAD diaries in SHEDS version 3), which indicates the sequence and duration of activities and locations for that person.
3. Generate concentration time series for each potential contact medium (e.g., indoor air, indoor smooth surfaces, indoor textured surfaces, indoor dust, outdoor air, outdoor lawn, outdoor vegetable garden, outdoor soil, and pets). These concentrations may depend on the usage patterns for household pesticides or other chemicals.
4. Simulate the contacts between the individual and the affected media. These depend on the diary activity/location information and user-specified contact probabilities.
5. Calculate pathway-specific exposure time series for the individual (with exposure equations given starting on page 35), using the results of the prior two steps and user-specified distributions for exposure factors.
6. Generate an approximation for the dose time series, if desired, using the simple built-in pharmacokinetic model in SHEDS (see section starting on page 45).
7. Export exposure time series for use in a PBPK model, or extract desired metrics or summary statistics from the exposure or dose time series.

To obtain population estimates SHEDS-Multimedia repeats this process for an individual many times using Monte Carlo simulation. Once the model run is complete, the user may view summary tables and graphs of the results. SHEDS includes 5 possible categories of variables related to exposure and dose (new exposure, running exposure, absorption, dose, and elimination), each divided into 7 pathways (residues to surfaces on hands, residues to surfaces on body, residues to surfaces in GI tract, dust or soil to hands, dust or soil to body, dust or soil to GI tract, air to lungs). Thus, there are 35 basic output variables, not counting sums across pathways. Each of these is reported on each diary event, with typically 12,000-20,000 diary events per year, per individual. When output from just one individual is examined, it is usually plotted as a time series (for any user-specified number of days up to the simulation length). When collectively examining output from a population, usually each person is summarized by a single value (often some average exposure or dose over the simulation period), and the population distribution of this statistic is generated. An overview of the SHEDS methodology is given in Figure 3.

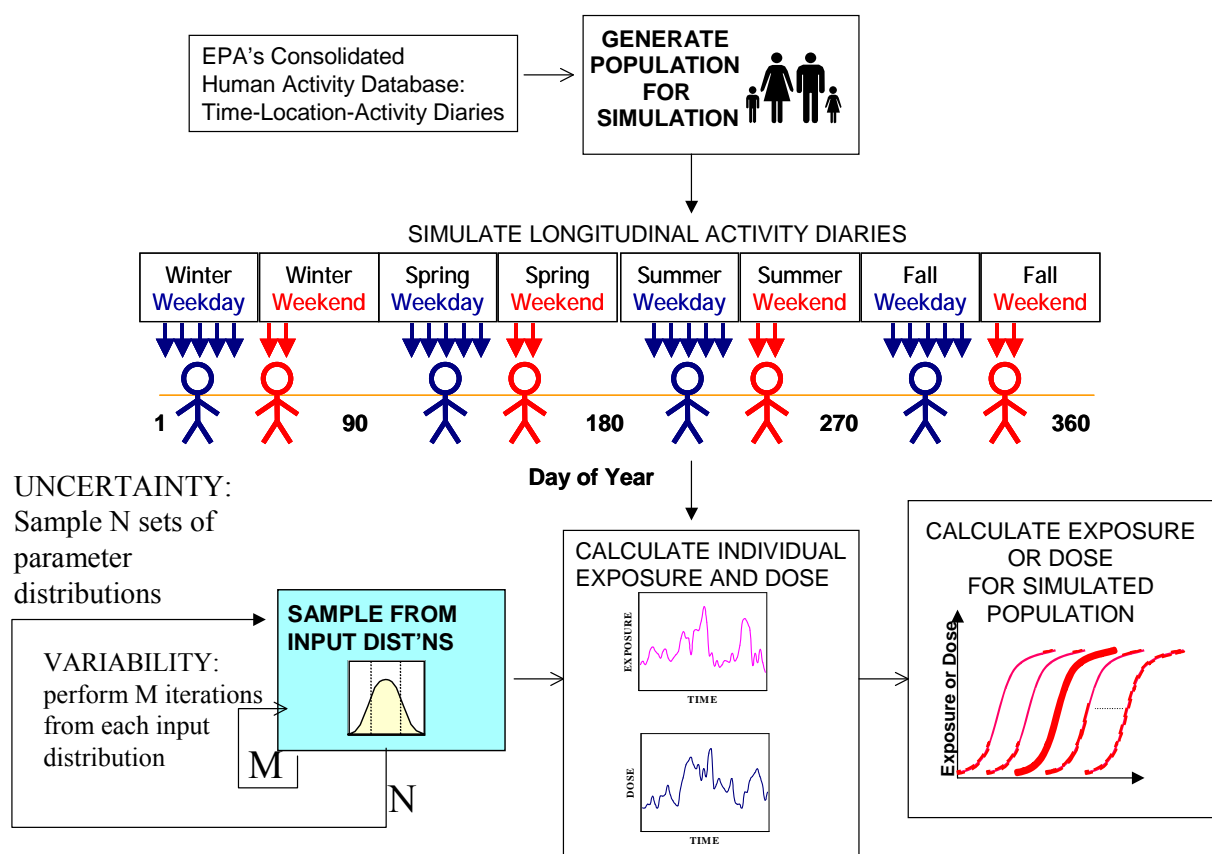


Figure 3. Overview of SHEDS methodology.

Features and Advantages of SHEDS-Multimedia Version 3

Key features of the aggregate version of SHEDS-Multimedia (version 3) include the following:

- physically-based, probabilistic exposure model coded in SAS with modularized code, flexible user input files, user-friendly GUI;
- generates variability and/or uncertainty distributions of exposure and/or dose;
- simulates route-specific (dietary, hand-to-mouth, object-to-mouth, dermal, inhalation) and aggregated (single chemical) exposures/doses;

- time series approach uses EPA's CHAD macroactivities as basis for longitudinal activity patterns and time steps; accounts for variability within a day;
- includes co-occurrence of chemical usage for single chemical, multiple application methods;
- exposure scenarios for chemical applications include post-application residential indoor crack and crevice, pets, indoor fogger, lawn, and garden (for different application methods);
- 3 options for source-to-concentration module: read time series; sample from probability distributions; use built-in decay/dispersion model;
- 2 options for exposure-to-dose: export exposure time series to PBPK model; built-in simple PK model;
- 2 options for soil/dust ingestion: direct ingestion; hand-to-soil/dust-to-mouth;
- 2 options for dermal exposure: transfer coefficient or transfer efficiency;
- code permits correlation of randomly sampled model inputs;
- built-in decay/dispersion model includes treated and untreated room; concentrations/residues for surfaces (smooth or textured), air, dust, as well as outdoor soil, lawn, and vegetable garden;
- separates dermal hand and body exposure profiles and links hand-to-mouth ingestion with dermal exposure; and
- accounts for sequential activities, dermal replenishment, and removal processes (hand-to-mouth, bathing, hand washing, absorption).

QUALITY ASSURANCE (QA)

SHEDS-Multimedia version 3 was conceptually developed (model approach and algorithms), designed (GUI), and quality assured by exposure modeling researchers in EPA/ORD/NERL. Writing the SAS computer code to implement the model, developing some of the model algorithms, and constructing the GUI was conducted by Alion Science and Technology under Contract EP-D-05-065. Integrated throughout this entire process was a heavy emphasis on QA, both by EPA and Alion Science and Technology. This emphasis resulted in a high degree of reliability in both model operation and GUI performance.

Development QA

As the model structure was being developed, the equations were checked both individually and collectively for consistency and to ensure that they were combined properly so that the model itself behaved as intended. The model has been implemented in SAS and requires only the BASE, STAT, and GRAPH modules. Development of the GUI also required the AF module, though this piece of software is not needed to run the model or GUI. SAS software is universally recognized as an extremely reliable statistical programming language and is thoroughly quality assured before its release. In addition, SAS provides technical support to its users, and the SHEDS-Multimedia developers availed themselves of this utility during model code and GUI construction.

As the model code and GUI were being built, the developers cross-checked each other's work and as problems were revealed they were corrected. This checking procedure was done on individual modules and sections of code (e. g., diary assembly, random distribution sampling, exposure module), and as portions of new code were added to existing programming to make certain that the assembled unit performed correctly. Finally, the complete package was also tested.

Next, individuals in ORD/NERL, OPP, and Alion tested SHEDS-Multimedia, including evaluating the user-friendliness and layout of the GUI. They recorded any problems or suggestions for improvement and communicated these to the developers. The model code and GUI were then appropriately revised.

Finally, SHEDS-Multimedia was released to several interested external beta testers who provided valuable feedback; their comments led to further revisions.

One-person, one-day simulation for code verification

After all these steps, the implementation of the exposure equations in the final code was numerically evaluated by comparing the output from a one day simulation with selected hand calculations. The equations, inputs used for this, and the resulting outputs are all provided in Appendix D.

A model run of SHEDS-Multimedia covering just one person for one day was made. This individual was a 10-year old girl. An artificial diary day with 30 events was constructed, with enough variety to provide examples of nearly all of the exposure equations. The Post-Application method for chemical concentrations was used, with the concentrations set to integer point values for simplicity. In the exposure code, 'ConcR' is the variable indicating the loading of residues on the contacted surface in $\mu\text{g}/\text{cm}^2$, 'ConcM' gives the concentration in soil or dust in $\mu\text{g}/\text{g}$, and 'ConcA' gives the air concentration in $\mu\text{g}/\text{m}^3$. The following lists the values used in this example for each of these variables:

ConcR = 30 [if residue = 'RY|lawn']

	= 10	[if residue = 'RThard']
	= 5	[if residue = 'RTsoft']
	= 5	[if residue = 'RUhard']
	= 2	[if residue = 'RUsoft']
	= 0	[otherwise]
ConcM	= 20	[if matter = 'MTdust']
	= 3	[if matter = 'MUdust']
	= 0	[otherwise]
ConcA	= 2	[if air = 'ATair']
	= 1	[if air = 'AUair']
	= 0	[otherwise]

The data set 'expo_ug' in the output directory was examined in detail, and two tables were constructed to summarize its contents. Table D-1 is 'Exposure_inputs', which lists in alphabetical order the variables input to the exposure module. This table has 30 rows, one for each of the 30 diary events. Each of the 50 columns represents one input variable and are spread over multiple pages. Some of these inputs retain the same value over the course of the day, while others may change hourly or change on every event. Table D-2 provides the exposure equations, and Table D-3 reports the model output.

All of the concentration values (ConcR, ConcM, and ConcA) and all the ingestion values were checked. Sleep events were visually scanned as blocks and no problem was found. Spot checks on other variables were made which included at least one example from every row and every column in Table D-3. All of these checks were successful. Without automating the process (for example, by entering all the data and equations in a spreadsheet), it is not practical to evaluate every item in these tables.

The reader may wish to perform additional spot checks at their own discretion. Instructions for using the equations in Table D-2 in conjunction with the inputs in Table D-1 are provided in Appendix D. These results may then be checked against the outputs reported in Table D-3.

Generating Concentration Time Series for the Contact Media

Chemical usage patterns (applications)

Unless the user selects the 'TimeSeries' input option for concentrations, the environmental chemical concentrations will be the result of chemical usage patterns. There are two methods of specifying chemical usage patterns. One is for the user to specify the calendar dates when chemicals are applied (the 'UserDates' option). Here, all simulated individuals will share the same usage patterns. However, the fraction of the house treated and the mass of chemical applied may vary from person to person. The UserDates option is useful, for example, for estimating the variability in exposures that may follow an application. All persons in the model run may experience an application on day 1 (or on day 0), and the comparison of individuals on similar days of their time series provides the exposure distribution for that point in time. The UserDates option also supports multiple applications of one type, as well as multiple types. For the UserDates option, if a lawn application is selected, then automatically all simulated individuals are assumed to have a lawn. Similar logic applies to pets and to the vegetable garden, when applications of suitable type are specified.

The second method for determining chemical usage patterns is called the 'ModelDates' option. Here, SHEDS randomly generates chemical usage frequencies and dates according to user-specified rules. Output data analyses from ModelDates runs often differ from those of UserDates runs, since individuals are not necessarily exposed at the same times. For ModelDates, summary statistics over each time series are computed first, then the results are compared across the population. For example, one may be interested in the mean daily exposure, maximum single daily exposure, or total number of days over some exposure threshold. Such quantities are not directly dependent on the exact timing of the chemical usage.

Apart from the determination of chemical usage frequencies and application dates, the UserDates and ModelDates options are very similar. Most sections discussed below apply to both methods. Exceptions are noted, with an explanation of the differences.

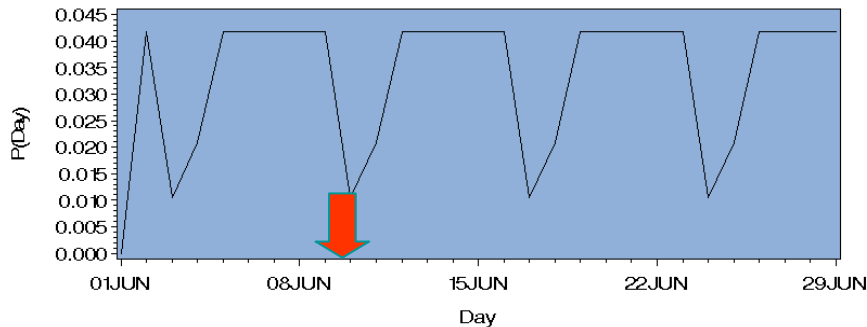
Application types

An "application type" refers to a specific combination of chemical formulation, application method, and target location. An example of target location is the lawn. Pesticides may be applied to the lawn either in granular form (e.g., using a mechanical spreader), or liquid form (e.g., using a pressure sprayer). Other alternatives may be possible as well. Indoors, there are even more choices for application types. The user selects the application types for a particular model run from a master list provided by the interface. In batch runs, it is possible to add new application types without changing the SHEDS code, but this is fairly difficult and is not recommended for most users, nor is it supported by the interface.

For a Single Treatment Type the User Specifies

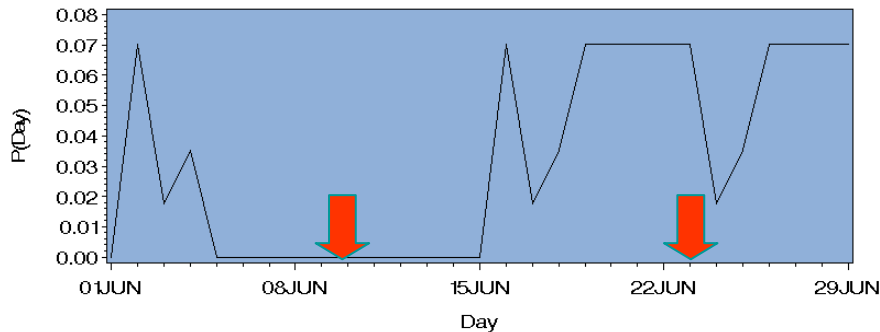
- Treatment type and priority for co-occurrence
- Probabilities of applications based on month and day of week

Initial Application Probabilities for Treatment Type 1
Initial Application Chosen on June 10th



Initial probabilities for treatment type 1. First Application date chosen on June 10th.

Adjusted Probabilities for Second Application of Treatment Type 1
First Application on June 10th: 5 Day Minimum Between Applications

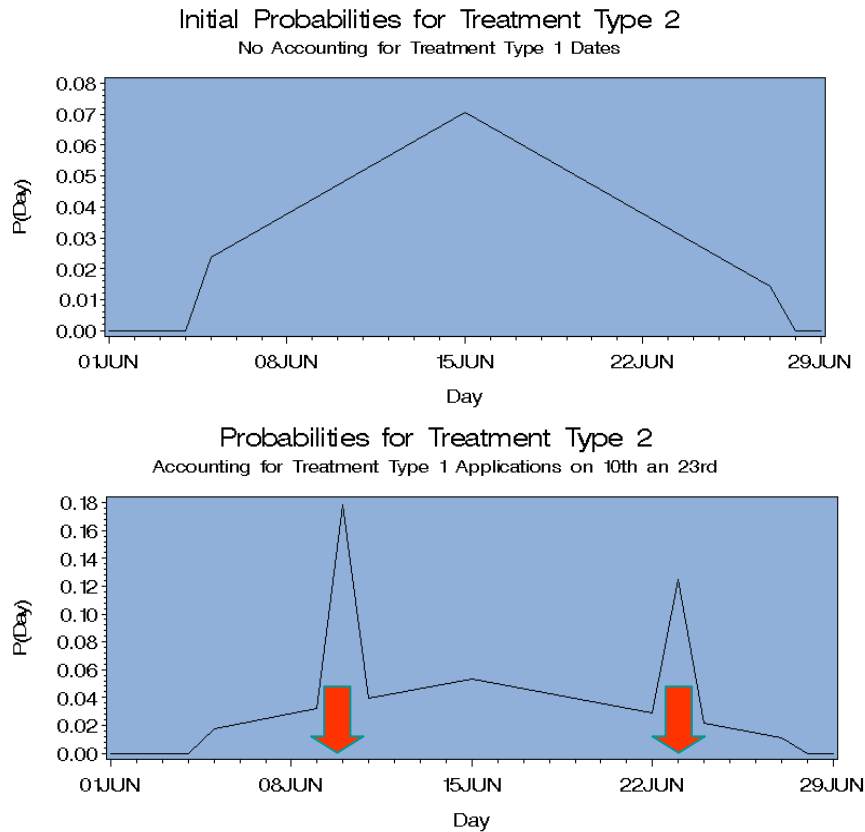


Probabilities adjusted for initial June 10th application date assuming a 5 day minimum time between applications. Second date chosen on June 23rd.

Figure 5. Choosing application dates for a single treatment type

User Specifies Treatment Types and Relationships

- All quantities needed to define individual treatment types
- Priority among treatment types (type 1 was first priority, type 2 is second, ...)
- Influence factor and influence width for each pair of related treatment types



Initial probabilities for treatment type 2 not accounting for treatment type 1 application dates.

Probabilities adjusted for treatment type 1 application dates, June 10th and 23rd. These probabilities will be used to pick first date for treatment type 2.

Figure 6. Co-occurrence of treatment types, choosing application dates for treatment type 2

User-Specified Concentration Time Series

This is the simplest option for determining the concentration time series. The user supplies a file (a SAS data set) containing daily time series for each contact medium. The user may supply multiple examples of each contact medium, corresponding to a set of houses. This data set has one record per day. The list of variables must include one called “date” which indicates the date for each record in SAS date format. It also requires variables whose names are in the form “medium_house#”. For example, if there were data from three houses, then the air concentrations in the treated room would be under the variables named “ATair_1”, “ATair_2”, and “ATair_3”. The corresponding dust concentrations would be “MTdust_1”, “MTdust_2”, and “MTdust_3”. The first letter of these names indicates the type of media: A for air, S for surfaces, or M for matter. The types indicate the necessary units: ($\mu\text{g}/\text{m}^3$) for air, ($\mu\text{g}/\text{cm}^2$) for surface residues, or ($\mu\text{g}/\text{g}$) for matter. The units do not explicitly appear on the input files; the user must ensure that they are correct. The suffix 1, 2, or 3 indicates the house number. SHEDS selects one house at random for each person, and uses all the concentrations from that house together. Other variables may be supplied apart from the ones listed above; they are not read by SHEDS. It might be useful to have a variable for comments or notes to help document the data set.

The houses do not need to use the same start and stop dates, but must have no gaps (days with missing data) between their start and stop dates. All of the time series for any given house must have the same start and stop dates. As an example, suppose there are just three houses and two media. The following input file would be valid:

Table 4. Example input file for concentration time series.

Date	Thard_1	Tsoft_1	Thard_2	Tsoft_2	Thard_3	Tsoft_3
14JUL2006	.	.	5.4	1.2	.	.
15JUL2006	.	.	3.3	0.6	.	.
16JUL2006	6.8	4.2	1.7	0.4	.	.
17JUL2006	5.8	4.7
18JUL2006	3.7	2.6
21JUL2006	7.4	5.2
22JUL2006	6.1	3.5

Note that there are two days missing between the end of house 1 and the start of house 3. These records may either be present or omitted on the input file; it does not matter. If present, then of course all the concentrations on these days would be indicated by periods. Each person simulated will be assigned to one of the three houses at random. The simulation period for that individual is then set according to the start and stop dates for that house, provided they fall within the start and stop dates of the simulation period as a whole.

The user does not need to supply data for more than one house. If the input file contains data for N houses, then roughly (1/N) of the people will be randomly assigned to each house. For each simulated person, all the concentrations are taken from the same house. After the model run is complete, the user can examine the exposure distribution within houses, between houses, or overall. This may provide information on the contribution of variance in concentration to the variability seen in exposure or dose.

When defining a SHEDS-Multimedia job, the user specifies a start date and a number of days, defining the simulation period. For the UserDates and ModelDates options, all simulated persons share the same simulation period. For the TimeSeries option, these dates are used to subset the input file. The dates on the timeseries input file may begin before and/or extend after the dates for the simulation period, but only the dates within the simulation period are utilized.

For example, suppose the start date is set to July 17, 2006 and the number of days to 7. Then only the last four records in Table 4 would qualify as belonging to the simulation period. Since there is no valid data remaining from house #2, it is removed as a possible candidate, so half of the simulated persons would be assigned to house #1, and the other half to house #3. In this case each simulation would be two days long, but in general the number of days may differ between houses.

Generating Exposure Time Series for Individuals

SHEDS equations for exposure by pathway

The estimation of exposure is the central function of the SHEDS model. Exposure is defined in SHEDS-Multimedia as the contact between a chemical agent and a human target at the skin, lung, and gastrointestinal (GI) tract exposure surfaces. The skin exposure surface is separated in SHEDS into hands and the rest of the body. The time series of exposure preserve within-day peaks and variability as a person moves through time (Figure 7 and Figure 8).

These exposure profiles can yield toxicologically relevant dose profiles (especially for chemicals with short half-life), and ultimately, improved risk estimates. Exposure is calculated for each diary event, resulting in pathway-specific time series for each individual. The skin (hands and skin on rest of body) and GI tract may be exposed to either residues (that is, chemical which is transferred from surface residues) or to matter (dust or soil). The lungs are exposed only to air. Thus, SHEDS tracks seven pathways for exposure:

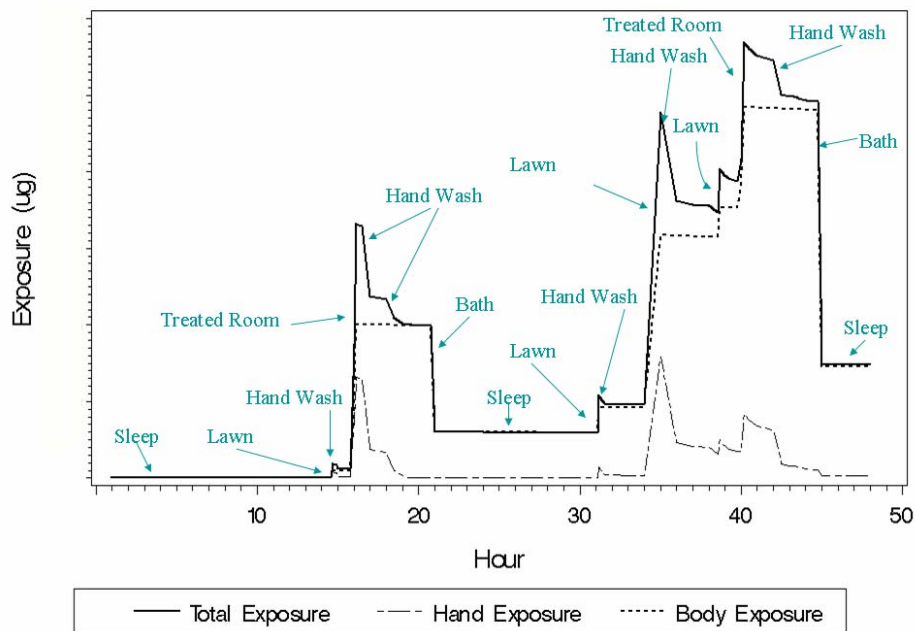


Figure 7. Example time profile of dermal exposure from surface residues

HF_e = fraction of one hand that enters the mouth [-]
 MRE = mouthing removal efficiency (fraction transferred to mouth) [-]
 FQH_e = frequency of hand-mouth activity [mouthing events/hr]
 T_e = duration of diary event [hr]
 $bioav$ = bioavailability for residues or dust/soil

The first assumption behind the hand-to-mouth transfer is that each diary event starts with the dermal hand loading spread evenly over the two hands. Each diary event is likely to have several hand-mouth contacts. Within a diary event, it is assumed that repeated hand mouthings all involve placing the same part of the hand into the mouth, and that no reloading occurs. Each repetition therefore involves less mass remaining on the hand and therefore less transfer. For example, three thumb insertions in one diary event does not result in triple the mass transfer of a single insertion, because the thumb has been partially cleaned prior to the second and third.

The initial chemical mass on the part of the hand entering the mouth is simply the loading on one hand ($DE_{h,e}/2$), times the fraction of the hand which enters the mouth (HF_e). The second part of the transfer equation (the term $[1 - (1-MRE)^{(FQH_e * T_e)}]$) represents the fraction of this mass that is transferred. The fraction transferred is calculated as follows. The variable FQH gives the frequency of hand mouthing events per hour. For a diary event of duration M minutes, the expected number of hand mouthings during this event is $(FQH_e * M / 60)$, or $(FQH_e * T_e)$, where T is the event duration in hours. As an example, suppose each hand-mouthing event results in removal of $MRE = 20\%$ of the available loading (that which enters the mouth). If one starts with a loading of 1000 (units are irrelevant for this example), then the first mouthing removes 200, leaving 800 behind. The second mouthing now involves a loading of 800, of which 20% (or 160) is removed, leaving 640 behind. The third mouthing results in $640 \times 20\% = 128$ being removed, leaving 512 behind. It should be clear that every mouthing means that the amount left behind is multiplied by another factor of 0.80, or $(1-MRE)$. After N mouthings, the amount left behind would be 1000×0.80^N . Using variables instead of constants, the fraction removed on N mouthings is $(1-MRE)^N$, and here $N=(FQH_e * T_e)$. The fraction transferred to the mouth is one minus the fraction left behind, so the transfer fraction is $[1 - (1-MRE)^{(FQH_e * T_e)}]$.

The scenario presented here for hand-mouthing would imply that the power $(FQH_e * T_e)$ should logically be an integer. In practice, diary events have variable duration, and $(FQH_e * T_e)$ will often have a non-integer value. Suppose it turns out to be 3.7 on some diary event. An exponent of 3.7 will remove more than an exponent of 3, but will remove less than an exponent of 4. The 'fractional' number of mouthings can also be interpreted as 3 full mouthings (each of HF_e) and one partial mouthing (an area somewhat less than HF_e).

Note that at most, only an amount $DE_{h,e}/2 * HF_e$ can be removed by hand-to-mouth transfer during one diary event. For example, if $HF_e = 20\%$ then a maximum of 10% of the hand loading can be removed. However, at the end of each diary event, the remaining loading is redistributed so more could be removed on the next diary event.

' $Bioav_c$ ' is the bioavailability term. The bioavailability refers to the fraction of the ingested chemical mass that is in a form that makes it available for absorption through the GI tract lining into the blood. Whether this fraction should be applied to the exposure or to the dose calculation depends on the definition of the agent of the GI tract exposure. If the agent is defined narrowly (that is, as the amount present in a form that can be absorbed) but the total chemical mass is being tracked when it is outside the body, then the bioavailability fraction should be set to a value below 1. Otherwise, the bioavailability should be set to 1. The user should ensure that the GI tract absorption rates are consistent with the selected interpretation of the bioavailability term.

The hand-to-mouth transfer equation is always used for surface residues, but is one of two options for calculating soil/dust ingestion. This is the indirect method, as opposed to direct soil ingestion. The indirect method still uses the above equation as a hand removal rate, but does not add this to the GI tract loading, as that would result in a “double counting” of soil/dust ingestion.

Object-to-mouth exposure transfer (transfer efficiency method)

Equation 9

$$E_{om, e} = C_{surf, e} * object_ratio * om_area * [1 - (1 - om_transfer)^{(om_freq_e * T_e)}] * bioavr$$

where

$E_{om, e}$	=	new object-to-mouth ingestion exposure for event e for residues [μg]
$C_{surf, e}$	=	chemical residue concentration on contacted surface for event e [$\mu\text{g}/\text{cm}^2$],
$object_ratio$	=	concentration ratio for object to nearby surface [-]
om_area	=	area of the part of the object that enters the mouth [cm^2]
$om_transfer$	=	fraction transferred to mouth from object surface per mouthing event [-]
om_freq_e	=	frequency of hand-mouth activity for event e [mouthing events/hr]
T_e	=	duration of diary event [hr]
$bioavr$	=	bioavailability in the GI tract for residues [-]

The object-to-mouth transfer uses exactly the same logic as the hand-to-mouth transfer presented above. The only differences are in the names of the variables. These objects are generally assumed to be toys that children leave on the floor and other exposed places, later to be played with and chewed. The amount of chemical on the toy depends on the variable “object_ratio”, which gives the typical concentration ratio for toy surfaces to room surfaces. The variable $C_{surf, e}$ will be event-dependent, so for example if a child is contacting the ‘Thard’ surfaces during some diary event, then the toys the child plays with at that time are with are assumed to have concentrations proportional to the Thard concentration. When the child plays in the untreated area, the toys are likely to be cleaner. This effectively implies that the toys are not moved from place to place with great frequency.

If the dermal transfer coefficient approach is used instead of the dermal transfer efficiency approach, then object-to-mouth transfer is calculated using the same equation as for hand and body dermal contacts, but with a different transfer coefficient.

SHEDS Exposure Adjustments

In SHEDS the running exposure is the current mass of the agent chemical being carried on the exposure surface. For example, the running dermal exposure is the mass of chemical (μg) on the skin, summed over both hands. Mass units are used in preference to others such as mass per unit area or mass per unit body weight, to facilitate conservation of mass during the acquisition and removal processes.

The exposure undergoes a series of adjustments on each diary event. It starts with the carryover exposure from the prior diary event. New exposure is then added. The totals are then checked against the maximum dermal loading limits. Next SHEDS implements absorption or other removal processes, which move some of the chemical from the given body part into the blood. Then hand-to-mouth transfer is calculated. Finally, removal processes such as washing and bathing are simulated.

The adjustments for each diary event are estimated and are made “instantaneously.” For example, all new exposure occurs before any absorption, and all absorption occurs before any hand-to-mouth transfer, and so on. While these processes ought to be ongoing and simultaneous, using discrete time steps in modeling is a well-established practice that should be reasonable unless the diary events become too lengthy.

Not all of the SHEDS adjustments to running exposure affect all seven pathways. For example, the carryover loading does not apply to the lungs. SHEDS assumes that any chemical not immediately absorbed into the blood from the lungs will be exhaled, so the lungs start every diary event with zero running exposure. Also, at 6 a.m. each day the GI tract is voided and the running exposures set to 0.

New exposure

As discussed above, the “new exposure” in SHEDS is the mass of chemical transferred from the contact media to the person on the current diary event (equivalent to the “exposure mass” definition in Zartarian et al., 2005). New exposure only occurs during contact events with media that have nonzero concentration. Two methods are available in SHEDS for determining new exposure: either transfer coefficients or transfer efficiencies. The choice is made at the start of the model run and applies thereafter; the two approaches cannot be mixed. Transfer coefficients are expressed in units of (cm²/hr). There are three transfer coefficients in SHEDS, one for the hands, one for the rest of the body, and one for the mouthing of objects (other than the hands). The transfer coefficients are multiplied by the event duration (in hours) and the surface residue concentrations (in µg/cm²) to give the new exposures in (µg).

The transfer efficiency approach calculates new exposure as the product of a fractional dermal contact rate (1/hr), the event duration (hr), the available (unclothed) area of the body part (cm²), the surface concentration (µg/cm²), and a surface-skin transfer efficiency (no units). For the mouthing of objects, an equation analogous to the hand-mouthing equation is used. The inhalation exposure is the same regardless of whether transfer coefficients or transfer efficiencies are used; it is given by the product of the METS level (no units), the basal ventilation rate (m³/hr), the event duration (hr), and the air concentration (µg/m³). The METS levels are part of the CHAD system and they represent the ratio of activity-specific metabolic rates to basal metabolic rates. These ratios are more stable across the population than the activity-specific metabolic rates are alone.

Maximum dermal loading

This check ensures that the dermal loading does not exceed some specified upper limit. For this purpose, the combined loading for both residues and matter (dust/soil) is examined. As a model input, the bound is given as maximum loading per unit skin area (micrograms of chemical per square centimeter of skin). Separate distributions may be entered for the hands and for the rest of the body, although generally one might expect them to be the same. For checking the dermal loading, the maximum dermal loading is multiplied by the hand skin area to obtain a maximum permitted mass on the hands; and similarly by the unclothed body skin area to obtain a maximum mass on the body. If the sum of the dermal loadings exceeds the maximum dermal loading, both components (residue and matter) are reduced proportionally until the limit is reached. If the actual loading is below the limit, no adjustment is made.

Absorption

On each diary event, the amounts absorbed into the blood are determined and removed from the exposures (preventing absorbing them again later). For the lungs, the absorption is a simple fraction of the amount inhaled; anything not absorbed is considered to be exhaled and is not available as carryover exposure to later events. Note that the time scale for pulmonary absorption is relatively fast compared to the diary event duration. For the other types of exposure, the absorption rates are specified on the input files as fractions per day. These are multiplied by the diary event duration (converted to days) to determine the fraction absorbed. For many chemicals, dermal absorption time scales are slow compared to diary events, so these fractions are never close to one for a single diary event. Due to the smallness of these fractions per event, the amount absorbed is considered to be directly proportional to the diary event duration, without any corrections for non-linear effects.

Hand-to-mouth transfer

Hand-to-mouth transfer of loading is both a source of new exposure for the GI tract and a removal process for the hands. In SHEDS it is determined after the absorption but before other removal processes. If the hand-to-mouth transfer calculations were positioned with other forms of new exposure, then the potential would exist for supra-maximal loadings to occur and be transferred to the mouth before the maximum dermal loading check could be made. This would be incorrect, as the maximum dermal loading should apply instantaneously. That is, if the new exposure results in excessive loading, then it essentially does not stick to skin at all but just falls off immediately, before there is time to transfer it to the mouth.

Given that the hand-to-mouth transfer is a dermal removal process, the question arises whether it should be considered before, during, or after the other removal processes. The most relevant competing removal process is hand washing. When children play and get their hands dirty, it is likely that they have enough time to mouth their fingers before their hands can be washed. Hence, the hand-to-mouth transfer should be evaluated before the hand loading is reduced by washing.

The equation for hand-to-mouth transfer is discussed in detail in the section on new exposures above; it is presented there because it amounts to a form of new exposure for the GI tract, and it also serves as a template for discussing the object-to-mouth exposure equation.

Other removal processes

There are three other removal processes for dermal loadings allowed in SHEDS that have not yet been discussed. Bathing (including showers and swimming) removes a fraction of the dermal loading from both the hands and the body. Hand washing removes a fraction from the hands only and does not affect the rest of the body. The third removal process is called “dry removal” or “brush-off,” which removes a fraction of the loading on all diary events, not just those involving water.

Bathing events are sometimes indicated as such on the original CHAD activity diary. CHAD was assembled from a dozen or more separate studies, and these were not consistent in how bathing events were recorded. Some studies simply included these among “personal time” or some similar non-specific description. The consequence is that SHEDS diaries may sometimes indicate baths, but other diaries may have long stretches without any explicit mention of them. Each individual is assigned a value for the maximum number of days between baths. Once this time is reached without an explicit bathing event, a bath is forced onto the end of the current diary event. The fraction of loading removed by bathing is given by the input variable ‘remv_bath,’ applied to both the body and the hands.

Hand washing is never explicitly indicated in CHAD and must be added by the model. Each person is assigned a value for the ‘washprob’ variable drawn from a distribution, which indicates the mean number of hand washings per day for that individual. It is assumed that the hourly probability for hand washing is 1/16 of this daily value, while awake. Thus, if a person had a daily mean of 4.0 hand washes, then the likelihood of hand washing during any waking hour is taken to be 25%. On diary events shorter than one hour, the probability is reduced proportionally to the event duration. In this example, a 12 minute event would have a 5% hand washing probability, since it lasts 1/5 of an hour. The efficacy of removal for hand washing is given by the ‘remv_wash’ variable.

This methodology for hand washing results in a variable number of washes from day to day. By the same token, the hand washes occur at different times each day, even when the same CHAD diary is used for the activities. Since the hourly probabilities are never integers, there is no need or advantage in requiring the daily mean to be an integer either. A person could average 1.6 hand washings per day, giving a 10% hourly probability. Therefore, the ‘washprob’ variable can be sampled from a continuous distribution, without needing to be restricted to integers.

MODEL INPUTS

SHEDS-Multimedia is a very flexible model which may be applied to a wide range of chemical exposure scenarios. To target the specific scenario(s) desired, the user must provide a set of appropriate inputs (Table 6). These describe, for example, the chemical usage and properties, human behavioral characteristics, and various exposure and dose factors. Many of these inputs allow random sampling from probability distributions, although point values can also be used to fix certain inputs.

Table 6. Summary of SHEDS user-specified input variables (not all apply to all model runs).

- Chemical Use-Related
 - Probability of using target chemical (on lawn, garden, pet, or indoors)
 - Pr (# applications per year) for each application type
 - Pr (month of application)
 - Pr (day of week application)
 - Pr (hr of application)
 - Probability of having a lawn, garden, pet
 - Pr (re-entry time indoors)
 - Pr (re-entry time outdoors)
- Activity-Related
 - Contact probability for various potential contact media
 - Probability of being in treated room while in home and awake
 - Max # days between baths
 - Soil/dust ingestion rate
 - Mean # hand washes per day per person
 - Fraction body unclothed
- Chemical Concentration/Residue-Related
 - Background and initial concentration/residues (Smooth and textured surfaces, dust, air, pet, lawn, garden, treated and untreated areas)
 - Decay rates (fraction lost per day)
 - Ratios for untreated to treated room concentrations
 - Object-surface concentration ratio
 - removal efficiency during bath/shower
 - removal efficiency during hand mouthing
 - removal efficiency during hand washing
 - residue-skin transfer efficiency
 - object-mouth transfer efficiency
- Non-Chemical Specific Exposure Factors
 - Body weight and surface area of hands and body
 - Dermal transfer coefficient for hands and body
 - Obj-mouth TE
 - soil-skin adherence factor
 - maximum dermal loading
 - Fraction of hand surface area that enters mouth
 - Obj-mouth contact SA
 - Body-surface and hand-surface fractional contact rate
 - object-mouth contact frequency
 - hand mouthing events per hour

Stochastic Sampling

As a Monte Carlo model, SHEDS relies extensively on random or stochastic sampling. It estimates the population variability by modeling many different people who collectively both represent the target population and who meet the specified distributions for a suite of input variables. The user can control many aspects of the modeling through these input variables, but one needs to be careful to specify all these inputs sensibly. Note that values in the demonstration file for SHEDS version 3 should NOT be assumed to be appropriate for actual modeling scenarios; they are for demonstration purposes only.

There are three types of stochastic sampling in SHEDS, two of which are under control of the user. For variables with a limited set of discrete output values, probability vectors are used. For variables that can take on any real-numbered value (or any value in a bounded range), continuous probability distributions are used. The third type is not under the control of the user; these involve hard-coded logic in the program code. An example of this last type is the selection of specific activity diaries from the diary pools.

Probability vectors

Input variables in SHEDS that permit more than two discrete outcomes are sampled from probability distributions. The user supplies a list or vector of numbers that represent the probability for each outcome. For example, the vector (0.3, 0, 0.5, 0.2) would mean a 30% chance for the first outcome, no chance for the second, a 50% probability for the third, and a 20% chance for the fourth. The numbers in a probability vector must sum to one.

The interpretation of these outcomes depends on the variable in question. The probability vector as a type does not have a maximum length; any number of outcomes are permitted. However, for each input variable there are only a set number of outcomes that have meaningful interpretations, so there is a variable-dependent maximum length for the vector. Variables that allow only two outcomes (which may be interpreted as a yes/no test) are not sampled from probability vectors. The user simply furnishes point value probability for one outcome, and the other outcome is automatically assigned if the test fails. For example, the variable “Do you have a pet?” has a point value for a “yes” outcome, not a probability vector.

In the SHEDS interface, the length of probability vectors is pre-set, and a check is made that the probabilities sum to one. If SHEDS is run in batch mode, the user must apply these checks. Many of the probability vectors in SHEDS refer to chemical usage, especially when the model is asked to generate the application dates. For each application type, probability vectors are assigned for “Day of Week,” “Month of Year,” and “Number of Applications per year.” If the model is run with 4 application types, then 12 probability vectors are needed for these aspects alone. The five other probability vectors in standard use are listed below. In addition, the SHEDS model constructs probability vectors internally for the determination of contact with the various contaminated media. The user enters the individual probabilities that are used to construct these last vectors (not the vectors themselves), because the list of possible contact media is quite fluid in SHEDS and depends on other inputs or probability tests. For example, the pet is not a potential contact medium for persons without a pet, but it can be for other persons.

The full list of input variables for which the user must explicitly supply probability vectors is given in Table 7.

Table 7. List of user-supplied probability vectors.

Variable	Purpose
NumApps	Probability of exactly 1, 2, 3, ... applications of same type at same house in one calendar year, given that at least one occurs at this house. In batch mode this vector can have any length, but the interface limits it to 12 per year. There is one NumApps variable for each application type used in the model run.
Month	Fraction of applications of this type occurring in each month. This vector has length 12, with the first cell representing January, the second February, and so on. For example, 0.05 in January means that 5% of all applications of that type are made in January, regardless of the length of the simulation period. Note that applications that occur outside the simulation period may result in chemical concentration that persists until the simulation period begins. There is one Month variable for each application type used in the model run.
DayOfWeek	Fraction of applications of this type occurring on each day of the week. This vector always has length 7, with the first cell representing Sunday, the second Monday, and so on. The model assumes that the same fractions apply throughout the year. There is one DayofWeek variable for each application type used in the model run.
Reentry(outdoors)	Immediately after any outdoor application, people may be prohibited from entering the treated area. This is called the reentry prohibition. There are 9 categories: 0 hours (no prohibition), 1, 2, 3-4, 5-6, 7-12, 13-24, 25-48, and 49-96 hours. Each category has an overall probability for selection that the user may alter. Within each category, the possible outcomes are all equally likely and cannot be altered individually. For example, if the 5-6 category is selected, then the actual reentry time is equally likely to be 5 hours or 6 hours (50% chance for each). There is one vector for all outdoor application types, although it is sampled again for each application. The actual reentry time is an integer number of hours between 0 and 96 inclusive.
Reentry (indoors)	The same as above, but applied to all indoor applications.
TimeOfDay(outdoors)	The hour of the day when an application occurs. To keep the length of the vector and the randomly-generated application times reasonable, the hours are restricted to 7 a.m. to 6 p.m. The same vector is used for all outdoor application types.
TimeOfDay(indoors)	The same as above, but applied to all indoor applications.
BathDays	The maximum number of days between baths. The CHAD diaries are inconsistent with regard to recording baths/showers. If the model assembles a longitudinal diary with a long stretch between baths, then baths are added to diary events whose duration causes the accumulated time since the previous bath to equal or exceed the value of the BathDays variable. This is an integer number of days with a maximum value of 7.

Continuous Probability Distributions

Table 8 lists the continuous probability distribution types and parameters used in SHEDS. The parameters are given generic names v1, v2, and v3 in SHEDS.

The geometric mean (GM) of a lognormal distribution is also its median. $\text{Log}(\text{GM})$ is the mean of the distribution of $\log(x)$. $\text{Log}(\text{GSD})$ is the standard deviation of $\log(x)$. Since standard deviations must be positive, then $\text{Log}(\text{GSD}) > 0$, which implies $\text{GSD} > 1$. The PDF of the SHEDS lognormal is

Equation 13

$$p(x) = \text{Exp}[-(1/2) (\text{Log}[x/v1] / \text{Log}[v2])^2] / (x \text{ Sqrt}[2 \pi] \text{ Log}[v2]), \text{ for } x > 0$$

If GM and GSD are given, then the lognormal has arithmetic mean and standard deviation

Equation 14

$$\begin{aligned} \mu &= \text{GM} \text{ Exp}[(1/2) (\text{Log}(\text{GSD}))^2], \\ \sigma &= \text{GM} \text{ Sqrt}[\text{Exp}([\text{Log}(\text{GSD})]^2) (\text{Exp}([\text{Log}(\text{GSD})]^2) - 1)]. \end{aligned}$$

If the user knows the arithmetic mean μ and arithmetic standard deviation σ of the lognormal instead of the GM and GSD, then these can be converted as follows:

Equation 15

$$\begin{aligned} \text{GM} &= \mu / \text{Sqrt}(1 + \sigma^2 / \mu^2), \\ \text{GSD} &= \text{Exp}(\text{Sqrt}(\text{Log}(1 + \sigma^2 / \mu^2))). \end{aligned}$$

If instead, one has the mean μ_{\log} and standard deviation σ_{\log} of $\log(x)$, then use

Equation 16

$$\begin{aligned} \text{GM} &= \text{Exp}(\mu_{\log}), \\ \text{GSD} &= \text{Exp}(\sigma_{\log}). \end{aligned}$$

Normal

This is the normal or Gaussian distribution commonly used in statistics. The normal has two parameters: the mean ($v1$) and the standard deviation ($v2$), with $v2 > 0$. Note that the normal is unbounded, so it is a good idea to provide lower and upper truncation points to prevent physically impossible values from being returned. The PDF of the normal is

Equation 17

$$p(x) = \text{Exp}[-(x-v1)^2 / (2 v2^2)] / (\text{Sqrt}[2 \pi] v2)$$

Point

A point value means that the same value is always returned. This is sometimes called a *fixed* or *constant* form. The point has one numeric argument ($v1$) which is the value that is to be returned. The mean is $v1$ and the standard deviation is zero. The sampling frequency does not matter for points. While points are technically discrete, here they are classified with the continuous distributions since they are applied to variables that are expected to reside on a continuous scale, but happen to be assigned no variability.

Table 10. Rules for diary pre-processing in SHEDS.

Test	Interpretation
(gender) in ('M','F')	Gender of diary respondent must be known.
(age>=0) and (age<=99)	Age must be known and not over 99.
sleep>0	There must be some sleep time.
LocUX<60	Less than 60 minutes in unknown locations.
ActUX<120	Less than 120 minutes in unknown activities.

A diary that fails any of these tests is removed from consideration for use in SHEDS. All other CHAD diaries are acceptable. This list of criteria may be modified or expanded.

If a rule contains any logical connectives like 'OR' or 'NOT', then the terms on either side of the logical operator should be put in parentheses, for example: (gender='F')or(gender='M'). The reason for this is that blank spaces are not permitted inside a rule, and without the parentheses the operator would not be parsed correctly. The diary QA rules may be used to limit the pool of available diaries. For example, if the model is being run only on children ages 1-6 years, then one could add the QA rule (1<=age<=6), to limit the diaries to the correct age range.

Similarly, one can require that all diaries have some outdoor time by using the rule (outside>0). Restricting the pool of diaries results in both space and execution time efficiencies. Any variable on the dataset RawQuest2 can be used in a QA rule. These include both variables from the CHAD questionnaire file and some additional variables like sums of time in various locations that are added in the 'diarysums' step. Note that every micro has its own variable of the same name, which contains the number of minutes spent in that micro. Assume that the user has specified five micros: in_home, in_veh, in_oth, out_home, and out_oth (see Table 2). Table 11 shows a complete list of variables that may be used in QA rules:

Table 11. Variables that may be used in diary QA rules.

Variable	Meaning
actUX	total time (minutes) in CHADact 'U' or 'X'
age	age in full years
age1	age as in CHAD (decimal years for ages below 2)
airCond	Y=has air conditioning at home, N=no, X=unknown
asthma	Y=has asthma, N=does not have asthma, X=unknown
avgtemp	daily average temperature (Fahrenheit)
bath	total time (minutes) spent in bath/shower
chadid	unique ID for each CHAD diary
county	county of residence
date	SAS date of diary day

Variable	Meaning
day1	day of month (range 1-31, unless missing)
dayNum	day sequence number for this individual
educat	2-character code for educational level
employed	Y=employed, N=not employed, X=unknown
fuel	1-character code for type of heating fuel
fulltime	Y=employed full time, N=not fulltime, X=unknown
garage	Y=has a garage, N=no garage, X=unknown
gasstove	Y=has a gas stove, N=no gas stove, X=unknown
gender	M=male, F=female, X=unknown
heartlung	Y=has a heart or lung condition, N=no, X=unknown
heating	2-character code for system for home heating
hourrain	hours of rainfall (0-24) on diary day
housing	2-character code for type of housing
in_home	total time (minutes) in in_home micro
in_oth	total time (minutes) in in_oth micro
in_veh	total time (minutes) in in_veh micro
inchrain	inches of rainfall on diary day (to 0.01 inch)
income	household annual income category (see CHAD)
jobhours	hours worked at job in prior week (see CHAD)
locUX	total time (minutes) in CHADloc 'U' or 'X'
maxtemp	daily maximum temperature (Fahrenheit)
month1	month of year (range 1-12, unless missing)
nDays	number of diary days in CHAD from this individual
nomets	total time with missing 'metslink' variable
nomicro	total time with missing (blank) micro assigned
occup	occupational category, X=(unknown or not working)
out_home	total time (minutes) in out_home micro
out_oth	total time (minutes) in out_oth micro
outside	total time (minutes) in any outdoor micro
pesticide	Y=used pesticides recently, N=no, X=unknown
pid	personal ID from original study
qaflag	# of failed QA test (qaflag=0 means it passed)
qcActLoc	CHAD flag for activity-location mismatch

Variable	Meaning
qcEatime	CHAD flag for minutes spent eating
qcHeavy	CHAD flag for minutes with heavy breathing
qcInfer	CHAD flag for minutes in inferred diary events
qcLong	CHAD flag for most minutes in same act and loc
qcMeals	CHAD flag for number of meals
qcMetab	CHAD flag for minutes at high METS (>3.5)
qcMiss	CHAD flag for minutes with missing act and loc
qcSleep	CHAD flag for hours spent sleeping
qctravel	1=mismatch between morning and evening travel, 0=OK
race	A=Asian,B=black,H=Hispanic,O=other,W=white,X=unknown
reccount	# of event records on diary day
school	total time (minutes) spent at school
season	P=spring, S=summer, F=fall, W=winter, X=unknown
sleep	total time (minutes) spent sleeping
smoker	Y=person is smoker, N=not smoker, X=unknown
smoker2	Y=person lives with smoker, N=no, X=unknown
state	20-character political state of residence
student	Y=is a student, N=not a student, X=unknown
travel	total time (minutes) spent in travel
wdwe	WD=Monday-Friday, WE=Saturday or Sunday
weekday1	3-character day of week (SUN,MON,TUE,etc.)
weekend	1=Saturday or Sunday, 0=other day of week
weight	person's weight in kilograms
work	total time (minutes) spent working at job
wraptime	actual start time for diary day
year1	4-character year for diary day
zipcode	5-character zip code, X=unknown

Some of the above variables such as 'fuel,' 'heating,' and 'occup' have codes or settings that can be found in the CHAD documentation.

The use of these rules can provide the user with control over the pools of acceptable diaries. However, at present this feature is not available through the user interface. If the rules are changed, then the %DiaryPreProcessing macro must be invoked, either before or during the SHEDS model run.

Contact Probability-related Inputs

The contact probability-related inputs control the likelihood of the simulated person coming into contact with particular media. For CHAD diary events that take place inside the residence, three input variables are needed to determine the “sub-micro” or room type (treated, untreated, or neither). These are:

- Probability of being in treated room while in home and awake
- Probability of being in untreated room while in home and awake
- Probability of being far from a source while in home and awake

As discussed in the Methods section, the “treated room” is any room where the chemical has been applied (or will be applied). The “untreated room” is any room into which the chemical may disperse. The third category refers to locations with no possibility of new exposure.

While individuals are sleeping, SHEDS assigns them to an untreated room, and it is assumed that no dermal contact or soil/dust ingestion occurs while sleeping.

The above three variables may be assigned distributions. These may be point values, but this is not necessary. In a general population, some individuals will probably treat a larger fraction of their house than others. Correspondingly, the probability of being in a treated area should be larger for such persons. For each person, a sample is drawn from each distribution. The sum of the three values is determined, and each value is divided by the sum. The three adjusted values then sum to unity.

In addition, the following probabilities are needed to determine the specific media being contacted. In Table 12 these are grouped according to the probability vectors formed from them.

Table 12. Contact probability variables.

Group	Variable
1	Contact probability for smooth surface in treated room Contact probability for textured surface in treated room Contact probability for pet while in treated room Contact probability for no surface while in treated room
2	Contact probability for dust while in treated room
3	Contact probability for smooth surface in untreated room Contact probability for textured surface in untreated room Contact probability for pet while in untreated room Contact probability for no surface while in untreated room
4	Contact probability for dust while in untreated room
5	Contact probability for lawn while outdoors at home Contact probability for garden while outdoors at home Contact probability for pet while outdoors at home Contact probability for no surface while outdoors at home
6	Contact probability for soil while outdoors at home

The odd-numbered groups contain four probabilities each. As for the sub-micro probabilities, one value is drawn from each distribution, per person. Each value is divided by the sum of the four, to create a probability vector whose probabilities sum to unity. For dust and soil contacts, only one probability is given in each location. This is an example of a yes/no decision (either contact occurs on this diary event, or it does not). The probability of a “no contact” outcome for dust or soil is just one minus the probability of a “yes” outcome, so it does not need to be specified separately.

Concentration-Related Inputs

As discussed above, there are three options for specifying concentrations used in SHEDS. The following sections describe the relevant inputs for these different methods.

Decay-dispersion method

Indoor application types

For each combination of selected application type and directly affected contact medium there are three inputs for the decay-dispersion method. The indoor application types (InCCAer, InCCliq, InCarRm, InFIK, and InFog – see Table 3 for a description of these) each affect four contact media (Thard, Tsoft, Tdust, Tair). Suppose the user selects the indoor crack-crevice aerosol (InCCAer) application type. The distribution for the initial concentration resulting from an InCCAer application is needed for each of the four contact media in the treated room. The concentration units are ($\mu\text{g}/\text{cm}^2$) for the surfaces Thard and Tsoft, the units are ($\mu\text{g}/\text{g}$) for Tdust, and are ($\mu\text{g}/\text{m}^3$) for Tair. The decay rate for the concentration is needed in each of the four media, expressed as the fraction lost per day. The third variable is used to determine the rate of dispersion from the treated to the untreated room. This variable requires some explanation.

Suppose both the treated and untreated rooms start at zero concentration and an application occurs in the treated room. The concentration in the treated room jumps to some initial value and then follows a first-order exponential decay over time, using the distributions discussed above. Once the application has occurred, the chemical starts to disperse into the untreated room. The concentration in the untreated room continues to increase for several days, but it is also subject to decay (using the same rate constant as for the treated room). As the concentration in the treated room drops, so does the amount of chemical entering the untreated room. Eventually, the rate at which chemical enters the untreated room drops below the rate at which it decays, and the concentration in the untreated room starts to drop. From this point on, the concentrations in both rooms continue to decline towards zero. Thus, there is a unique point in time when the untreated room is at maximum concentration. There is also a particular ratio of concentrations between the untreated and treated rooms at this time.

The exact relationship between the treated and untreated rooms is quite complicated, depending on room geometry, size, flow rates, and other factors. A proper analysis would require a detailed indoor mass-balance model for the chemical. The simpler decay-dispersion logic assumes that the treated and untreated rooms are essentially equivalent in all respects except for the chemical concentration. In this case, there is a clear relationship between the flow rate, the time until the untreated room reaches maximum concentration, and the concentration ratio at this time (when the untreated room reaches its maximum). Specifically, the ratio of the fraction in the treated room lost to dispersion to the fraction lost to decay equals the ratio of concentrations in the untreated room to the treated room at the time of the untreated room maximum (assuming both these ratios are much smaller than one). The latter ratio is easily measured in practical situations, whereas the fraction lost to dispersion would be difficult to measure directly. Therefore, SHEDS requests as input the variable “UTratio,” which is the ratio of concentrations in the untreated room to the treated room at the time of the untreated room maximum. SHEDS uses four such ratios, one for each contact medium.

Transfer-related exposure factors

The transfer-related exposure factors are variables that describe the tendency for the chemical to move from one compartment to another (generally on or off the human body). These variables include:

- soil-skin adherence factor
- maximum dermal loading
- object-mouth contact area
- object-surface concentration ratio
- fraction of body unclothed
- removal efficiency during hand washing
- removal efficiency during bath/shower
- surface residue to skin transfer efficiency
- removal efficiency during hand mouthing (saliva transfer efficiency)
- object-mouth residue transfer efficiency
- brush-off removal efficiency
- surface residue-to-skin dermal transfer coefficient for hand
- surface residue-to-skin dermal transfer coefficient for body
- transfer coefficient for object mouthing

The soil-skin adherence factor measures the tendency for the skin to accumulate soil loading while outdoors. In SHEDS, this variable is not dependent on event duration, but it is applied only when contact occurs. Therefore, when used it is multiplied by the hand-surface or body-surface contact rate and by the event duration. The units should therefore be interpreted as milligrams of soil per square centimeter of skin per contact. Note that the literature usually reports soil loading effectively as ($\text{mg}/\text{cm}^2/\text{day}$), as the loadings are often measured after accumulating over several hours of outdoor activity. Soil-skin adherence is sampled once per person.

The maximum dermal loading allows separate distributions for the hand and body. It is measured in units of micrograms of target chemical per square centimeter of skin. It is applied to the combined loading of residue, dust, and soil. When new exposure pushes the dermal loading over this limit, the loadings are reduced proportionally until they no longer exceed the maximum. The maximum dermal loadings are sampled once per person.

The object-mouth contact area is only used when the transfer efficiency option is selected by the user (see User Manual). This represents the effective surface area of toys and other objects that is in contact with the mouth during object mouthing events. If a pliable object is crumpled or folded to get it into the mouth, then only the external surface area should be counted. This variable is sampled once per person.

The object-surface concentration ratio is used to calculate the chemical concentration on the objects being mouthed. This is expressed as a ratio to the concentration on floor surfaces in the same location. It is unitless and is sampled once per person.

There are two important features of stochastic models that should be noted. First, no two model runs will produce the same results, even if all the input settings are identical. The “stochastic variation” is the variation in the same output statistic when the same model run is repeated. More individuals in each analysis leads to less stochastic variation. The second important feature is that there is less stochastic variation in the outputs near the center of the distribution than for points in the tails of the distribution. The mean or median exposure has less stochastic variation than the 90th percentile, which in turn has less stochastic variation than the 99th percentile.

For results to be interpreted correctly, the stochastic variation should be smaller than the real effects that are being estimated. For example, suppose a given statistic such as the 95th percentile of ADD is being compared across genders or across age groups. A statement that one group was higher than another is really only meaningful if the difference is not due to stochastic variation. Otherwise, another pair of model runs might show no effect, or even one in the opposite direction. As a rule of thumb, the stochastic variation in the mean or median of 1000 persons might be as large as 5%, while the stochastic variation in the 90th or 95th percentiles might be closer to 10%. If the user is interested in differences of this magnitude, then larger sample sizes should be used.

A standard SHEDS-Multimedia 3 variability run produces the following files in the output directory as shown in Table 13. An example CDF from a SHEDS variability run is given in Figure 9.

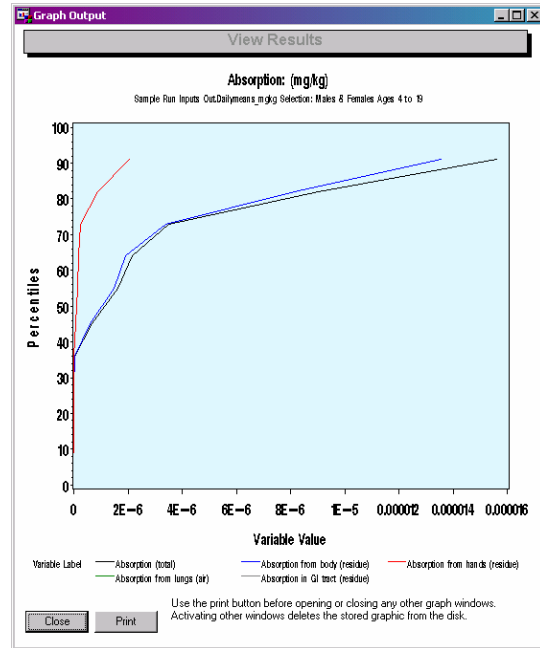


Figure 9. Example cumulative density function (CDF) plot from a SHEDS variability run.

Table 13. Output files for SHEDS-multimedia version 3 variability run

File	Contents
allpersondays_mgkg	one record for each simulation day, for each person, with exposures and doses in units of (mg/kg/day).
allpersondays_ug	one record for each simulation day, for each person, with exposures and doses in units of micrograms per day.
allvardaily_ug	one record in total, giving the averages across all persons and days for the variables on allpersondays_ug.
daily_mgkg	daily summaries for all output variables, for just the last person simulated, in units of (mg/kg/day).
daily_ug	daily summaries for all output variables, for just the last person simulated, in units of micrograms per day.
dailymeans_mgkg	means of daily summaries for all output variables, averaged across simulation days but not across persons, in units of (mg/kg/day).
dailymeans_ug	means of daily summaries for all output variables, averaged across simulation days but not across persons, in micrograms per day.
dose_mgkg	event-level values for all absorption, blood dose, and elimination variables, for just the last person simulated, in (mg/kg/day).
dose_ug	event-level values for all absorption, blood dose, and elimination variables, for just the last person simulated, in micrograms per day.
expo_mgkg	event-level values for all new exposure, exposure, and absorption variables, for just the last person simulated, in (mg/kg/day).
expo_ug	event-level values for all new exposure, exposure, and absorption variables, for just the last person simulated, in micrograms per day.
persons	one record for each simulated person, giving time-independent personal data, including age, gender, and input variables that are evaluated once per person.

Table 14. Hypothetical stepwise regression and percentile scaling sensitivity analysis results.

Input Variable	units	Step	Ratio between 95th and 5th percentiles			
			Mean	Median	95th	99th
hand mouthing events per hour	events/hr	1	2.22	3.42	1.92	1.84
GI tract absorption rate per day for surface residues	1/day	2	2.20	3.75	1.87	1.70
fraction of body unclothed	[-]	4	2.13	1.43	2.43	2.50
fraction of surface of one hand that enters mouth	[-]	8	1.90	2.83	1.68	1.58
maximum dermal loading for hands	ug/cm ²	9	1.53	1.40	1.54	1.39
removal efficiency during mouthing	[-]	14	1.50	1.80	1.39	1.35
maximum dermal loading for body surface-skin transfer coefficient for body (unclothed)	cm ² /hr	5	1.37	1.34	1.37	1.27
surface-skin transfer coefficient for hand	cm ² /hr	6	1.31	2.00	1.14	1.14
object-mouth contact area	cm ²	20	1.08	1.17	1.06	1.09
soil ingestion rate (outdoor, direct only)	mg/hour	16	1.06	1.25	1.03	0.99
GI tract absorption rate per day for dust or soil	1/day		1.02	1.26	0.98	1.02
dermal absorption rate per day for dust or soil	1/day	13	1.02	1.06	1.01	1.01
dust ingestion rate (indoor, direct only)	mg/hour	21	1.02	1.15	1.00	0.99
object-mouth transfer efficiency	[-]	15	1.01	0.95	1.04	1.00
body-surface fractional contact rate	1/hr		1.01	1.07	1.01	1.00
hand-surface fractional contact rate	1/hr	19	1.01	1.00	1.00	1.00
soil-skin adherence factor	mg/cm ²	18	0.98	1.00	0.99	0.97
residue-skin transfer efficiency	[-]	17	0.97	0.91	0.98	0.99
removal efficiency during hand washing	[-]	10	0.95	0.87	0.96	0.99
removal efficiency during bath/shower	[-]	7	0.77	0.71	0.80	0.86
mean # hand washes/day per person	1/day	12	0.68	0.48	0.77	0.77
elimination rate from the blood	1/day	3	0.52	0.47	0.54	0.53

The Percentile Scaling Approach

The second major form of sensitivity analysis is the percentile scaling approach. This differs greatly from the above methods in that the continuous input variables are not allowed to vary randomly, but always use preset values. Each input variable is still assigned a distribution as before, but instead of randomly sampling, the model is told to use one of three values: the “low” being the 5th percentile, the “medium” being the median or 50th percentile, and the “high” being the 95th percentile. The SHEDS model locates the values of these percentiles and assigns them as needed.

The user selects which variables are to be analyzed in a given model run, by setting the variable called “sensitivity” to unity, on the “variables” input file. Variables not to be analyzed are set to sensitivity=0. If N variables are selected, then SHEDS performs a set of (2N+1) model runs, each of the usual number of M persons. The first of these runs, called the baseline run, has all the input variables set to their “medium” values. The M persons do not share the same output because the non-continuous inputs may still differ. For example, the M persons will not necessarily realize the same demographic variables, even though they are drawn from the same population. The application dates may be different, unless the model is run in userdates mode (which is recommended). The activity diaries will likely be different. Finally, the contacts with media are sampled from probability vectors (which are not subject to percentile scaling), and so will be different. Hence, a sample of M persons is needed to represent the exposure distribution resulting from these other effects, even for a fixed set of continuous input variable settings.

For each of the N selected input variables, one model run is performed with this input at its low setting (5th percentile) for all persons, while all other continuous inputs are at their medium settings. Another run is then done with this input at its high setting (95th percentile) for all persons, while all other continuous inputs are at their medium settings. Thus there are two additional runs for each input, apart from the baseline run.

The mean output statistic (for example, ADD) is calculated across the M persons, for each run. Ratios of these means are then tabulated as follows. Considering one input at a time, the ratio of mean output from the high setting to the mean output at the low setting is calculated. Similarly, the ratios of high/medium and medium/low may be calculated, using the baseline run as the medium value (see). In all these ratios, all but one of the inputs are held fixed at its medium value in both runs. Thus, the ratio itself should reflect differences in exposure resulting from the difference in the setting for the one specified input. In practice, the ratios also contain some stochastic variation from the non-continuous SHEDS inputs. The sample size must be large enough that this stochastic variation does not overwhelm the systematic variation from the selected input, or else the latter cannot be quantified. A strong signal that the stochastic variation is too large occurs when the “high/medium” and “low/medium” ratios both exceed unity (or both are less than unity)

The percentile scaling approach is amenable to straightforward interpretation, due to the “one-at-a-time” nature of the variation of the inputs. By contrast, all the other sensitivity analysis methods allow all inputs to vary simultaneously. The principal danger with the percentile scaling approach is the possibility that the remaining stochastic variation from the other inputs might mask the true effect of the input in question. This issue may be resolved in Version 4 by the use of fixed random number seeds.

At the start of each uncertainty simulation, one of the K parameter pairs is randomly selected for each input variable. The selected (v1,v2) pairs define the variability distributions to be used for this simulation. All of the individuals within one uncertainty simulation randomly draw values from these variability distributions. This approach avoids potential problems arising from determining v1 and v2 independently.

Steps 5 and 6 above require one to compare the input variable uncertainty with that seen in the original studies, leading to a choice for B. Typically, it was found that B= 5 was suitable for very small or highly uncertain datasets; B=10 for somewhat larger datasets; and B=15 or B=20 for variables with fairly little uncertainty. The size of B depends on the sample size of the original data and the how well the bootstrap scatter plots fit over the data points among the various data sources (step 5).

The above applies to those inputs having two parameter distributions. All of the forms currently supported in SHEDS are two-parameter, except for the point (1-parameter) and triangular (3-parameter). Point values may have uncertainty (although they have no variability), as different surveys or studies may provide different estimates. In SHEDS, point distributions are used when all persons are intended to share the same value. For example, the probability of having a lawn may be set to 0.79, which means that each person has a 79% chance of having a lawn (not that the outcome of this test will be the same for all persons). However, this 79% chance may be uncertain, especially if SHEDS is applied to a subpopulation for whom the chance may be different.

Uncertainty can and should be assigned to the one and three parameter distributions in SHEDS, as well. The bootstrap method could be used, except that the comparison of scatter plots is simplest and most natural in two dimensions. Obtaining uncertainty clouds for points is straightforward (and simpler than the bootstrap method) as long as there are multiple sources of estimates. For the triangle, the bootstrap method could still be used if the user required that one point of the triangle be fixed (for example, the minimum could be always kept at zero). A three-parameter variant of the bootstrap could be used, but this would require plotting and matching clouds of data points in a three-dimensional space.

Each of the K points in the uncertainty cloud represents a fully-specified variability distribution. For a two-parameter distribution, each of these K points consists of a (v1,v2) pair that are to be used together.

The SHEDS-Multimedia 3 interface does not have the built-in capability for bootstrap analysis. The user must therefore conduct this analysis outside of SHEDS, or else use pre-generated uncertainty clouds from other sources.

Uncertainty inputs and outputs

For each of the N sets of inputs in a two-stage SHEDS run, one of the K points in the uncertainty cloud is selected at random, for each input. There is no need for the value of K to be the same for all inputs. In fact, if a certain input had no uncertainty, or were to be excluded from the uncertainty analysis for some reason, one could set K=1 for that input, so that the same variability distribution was always used.

Since one of the K points is selected at random, there is a chance that the same point in the uncertainty cloud will be selected for more than one of the N sets of inputs. If $K < N$, then this becomes certain. However, this is not a concern since the other input variables are likely to be different. Suppose that the model has 20 input variables, each with K=10 for their uncertainty clouds. Then there are 10^{20} possible combinations of input distributions, while only N of these (typically N is a few hundred) are actually sampled. Even with the unreasonably small value of K=2 for all inputs, there is no danger of running out of degrees of freedom, as $N \ll 2^{20}$. In practice, there is virtually no chance that any two of the N sets of inputs will be identical.

Sensitivity Analysis on Results from a Two-Stage Monte Carlo Model Run

The Pearson correlation, Spearman correlation, and stepwise regression methods of sensitivity analysis discussed in the Sensitivity Analyses section above may also be applied to the results from a two-stage SHEDS run (see Table 15 and Table 16). For this purpose, the mean value for each input variable and the mean output statistic are recorded for each of the (MxN) persons in the uncertainty run. The above methods are then applied, just as if the entire set of (MxN) persons belonged to a single-stage run.

This type of sensitivity analysis reflects the overall importance of the combined uncertainty/variability variation in each input on the output statistic. If a given input variable exhibits little importance in this regard, then the model will not benefit much from improvements in the characterization of the distribution for that input. For variables identified as important, the next step is to ascertain whether the variability or the uncertainty is predominant. This can be achieved by performing a sensitivity analysis on a single stage run. If the uncertainty is large, then the SHEDS exposure estimates can be improved by refining the knowledge of this input variable. If variability (rather than uncertainty) predominates, then this variable has an important influence on exposure, but this variation mostly reflects real differences among the target population and will not be much reduced by conducting further research into the input distributions.

Table 15. Hypothetical Spearman correlation uncertainty analysis results.

Input Variables	R ²
maximum dermal loading for hands	0.0989
maximum dermal loading for body	0.0959
GI tract absorption rate per day for surface residues	0.0921
fraction of surface of one hand that enters mouth	0.0859
fraction of body unclothed	0.0835
Body-surface fractional contact rate	0.0826
removal efficiency during mouthing	0.0823
object-surface concentration ratio	0.0822
probability of having a vegetable garden	0.0822
probability of having a lawn	0.0822
probability of having a dog or cat	0.0822
mass ratio (metabolite/pollutant)	0.0822
object-mouth contact rate	0.0822
removal efficiency during events without water	0.0822
transfer coefficient for object mouthing	0.0822
Hand-surface fractional contact rate	0.0821
removal efficiency during hand washing	0.0821
soil-skin adherence factor	0.0821
dermal absorption rate per day for dust or soil	0.0821
object-mouth transfer efficiency	0.0820
removal efficiency during bath/shower	0.0820
residue-skin transfer efficiency	0.0818
GI tract absorption rate per day for dust or soil	0.0813
elimination rate from the blood	0.0730
dust ingestion rate (indoor, direct only)	0.0487
object-mouth contact area	0.0284
hand mouthing events per hour	0.0246
soil ingestion rate (outdoor, direct only)	0.0076
surface-skin transfer coefficient for hand	-0.0034
surface-skin transfer coefficient for body (unclothed)	-0.0082
mean # hand washes/day per person	-0.0351

Table 16. Hypothetical stepwise regression uncertainty analysis results.

Input Variables	Stepwise Rank	Partial R ²	Model R ²	P value
maximum dermal loading for hands	1	9.78E-03	9.78E-03	0.00
elimination rate from the blood	2	1.15E-02	2.13E-02	0.00
GI tract absorption rate per day for surface residues	3	2.56E-02	4.69E-02	0.00
mean # hand washes/day per person	4	6.65E-03	5.35E-02	0.00
fraction of surface of one hand that enters mouth	5	5.72E-03	5.93E-02	0.00
object-mouth contact rate	6	7.58E-03	6.68E-02	0.00
maximum dermal loading for body	7	5.30E-03	7.21E-02	0.00
fraction of body unclothed	8	2.70E-03	7.48E-02	0.00
hand mouthing events per hour	9	5.09E-04	7.53E-02	0.00
surface-skin transfer coefficient for body (unclothed)	10	1.65E-04	7.55E-02	0.00
removal efficiency during mouthing	11	1.09E-04	7.56E-02	0.00
surface-skin transfer coefficient for hand	12	1.01E-04	7.57E-02	0.00
removal efficiency during bath/shower	13	9.25E-05	7.58E-02	0.00
residue-skin transfer efficiency	14	1.02E-05	7.58E-02	0.01
removal efficiency during hand washing	15	7.59E-06	7.58E-02	0.02
Max mets	16	7.07E-06	7.58E-02	0.02
object-mouth transfer efficiency	17	5.97E-06	7.58E-02	0.04
GI tract absorption rate per day for dust or soil	18	5.92E-06	7.59E-02	0.04

- U.S. Census Bureau. (2002a). Census 2000 Summary File 3 – United States.
- U.S. Census Bureau. (2002b). Census 2000 Summary File 3 Technical Documentation.
- Xue J. (2007). SHEDS-dietary SAS code.
- Xue J, McCurdy T, Spengler J, and Özkaynak H. (2004). Intra- and inter-individual activity considerations in human exposure modeling. *J. Expos. Anal. Environ. Epidemiol.* 14: 221-233.
- Xue J, Zartarian V, Özkaynak H, Dang W, Glen G, Smith L, and Stallings C. (2006) A probabilistic arsenic exposure assessment for children who contact chromated copper arsenate (CCA)-treated playsets and decks: Part 2: Sensitivity and uncertainty analyses. *Risk Analysis* 26: 533-541.
- Zartarian VG, Özkaynak H, Burke JM, Zufall MJ, Rigas ML, and Furtaw Jr. EJ. (2000). A Modeling Framework For Estimating Children's Residential Exposure and Dose to Chlorpyrifos Via Dermal Residue Contact and Non-Dietary Ingestion. *Environmental Health Perspectives* 108(6): 505-514.
- Zartarian, V.G., J. Xue, H. A. Özkaynak, W. Dang, G. Glen, L. Smith, and C. Stallings. (2005). A Probabilistic Exposure Assessment for Children Who Contact CCA-treated Playsets and Decks Using the Stochastic Human Exposure and Dose Simulation Model for the Wood Preservative Scenario (SHEDS-WOOD), Final Report. U.S. Environmental Protection Agency, Washington, DC, EPA/600/X-05/009.
- Zartarian V, Xue J, Özkaynak H, Dang W, Glen G, Smith L, and Stallings C. (2006) A Probabilistic Arsenic Exposure Assessment for Children Who Contact CCA-Treated Playsets and Decks, Part 1: Model Methodology, Variability Results, and Model Evaluation. *Risk Analysis* 26: 515-531.
- Zartarian V, Bahadori T, and McKone T. (2005). Feature Article: Adoption of an Official ISEA glossary. *Journal of Exposure Analysis and Environmental Epidemiology* 15: 1-5.

Concentration Time Series	A time series with time on the x-axis and contact media concentration on the y-axis.
Correlation	The SHEDS user may request correlation between pairs of continuous input variables. This is done by selecting the pair and specifying the Spearman (or rank) correlation between them. The marginal distributions of each variable will remain exactly as specified elsewhere. For example, one might be triangular and the other a gamma. The method allows the two correlated variables in the pair to be sampled at different frequencies. The user may specify as many variable pairs as desired. Some combinations of pairwise correlations are contradictory and cannot be achieved. SHEDS determines when this occurs and informs the user.
Decay/dispersion	One of the built-in source-to-concentration modules in SHEDS. The treated area undergoes first-order exponential decay in concentration, while nearby untreated areas may acquire chemical via dispersion from the treated area.
Dermal	SHEDS separates the skin into two regions: hands and body.
Diary event	The time basis on which all exposure and dose variables are calculated in SHEDS. Diary events have variable duration, from 1 to 60 minutes. Diary events never cross the top of a clock hour. The user may lower this upper bound for duration on the “Specify Simulation Information” screen of the model interface. The number of diary events in the simulation period generally differs between individuals.
Distribution	One of two ways for specifying the form of stochastic or randomly-determined inputs (the other being <i>probability vector</i>). Distributions are used for continuous variables in SHEDS, which may be sampled from the beta, gamma, exponential, lognormal, normal, point, triangle, uniform, and Weibull statistical distributions.
Dose	The mass of chemical absorbed (absorbed dose) or metabolite in the blood (blood dose) after the chemical agent has crossed the exposure surfaces. The pharmacokinetic model built into SHEDS assumes that upon absorption, all of the mass is immediately converted to a single metabolite and stored in the blood. The input variable “metab_ratio” gives the mass ratio of the metabolite to the parent compound (the chemical agent). Setting metab_ratio=1 means that the dose measures the parent chemical. Each event, the seven blood doses are each increased by corresponding “abs” term (multiplied by metab_ratio), and reduced by the corresponding “elm” term. Blood dose is measured in micrograms of metabolite (μg). When aggregating to longer time periods, doses must be averaged, not summed.

Post-application	Following (but not during) a residential chemical application. Also refers to one of the built-in methods for source-to-concentration modeling in SHEDS.
Probability vector	One of two alternatives for specifying the form of a stochastic input (the other being a <i>distribution</i>). It is used for variables that have a discrete set of outcomes. Each outcome is assigned a probability. These should total to one, as each time the distribution is sampled, one of the outcomes must occur.
Residue	One of the phases in which the target chemical may be present, abbreviated “R”. Generally, it refers to chemical residue that may be encountered when a surface is contacted.
Running exposure	The current mass of chemical on/in the given body part. This includes both new exposure and the amount carried over from prior events. The ones currently tracked in SHEDS are: ldgHR, ldgHM, ldgBR, ldgBM, ldgGR, ldgGM, ldgLA. All loadings are in micrograms (μg) of chemical. Care must be taken not to confuse “exp” and “ldg” variables.
Sampling frequency	The rate at which new samples are drawn for a given input variable. The most common sampling frequency is once per person. This means that each individual is given one value drawn at random, which remains fixed for that person. Other sampling frequencies include monthly, daily, hourly, or every event. Such variables change values over time for a single person, as well as differing between persons.
Simulation period	The time period for which exposures are estimated. This may range from one day to a year or more, measured in whole days. Each individual is modeled for the entire simulation period, which is the same for everyone. The exception to this is when the user specifies the “timeseries” method for concentrations and the input file contains more than one set of concentration measurements, with unequal measurement periods. In that case, each person is modeled only for the period with non-missing values in the set of measurements assigned to them.
Stochastic	An adjective meaning “randomly determined” or probabilistic. In SHEDS the sequence of persons and the values for many of the input variables are stochastic, so the individuals are independent and there is no significance to their order.
Sub-micro	A sub-microenvironment is a division of a microenvironment (see “micro”) into regions of differing concentration, usually because one part is “treated” while the rest is “untreated”. The “in-home” micro can optionally be divided into a third sub-micro called “neither”, which can indicate other houses than the one being treated. While the micro is determined deterministically by the codes on the

CHAD diary, the sub-micro is determined probabilistically using distributions input into SHEDS.

Target

In SHEDS, a simulated person that receives an exposure.

Treated room

An area in a simulated individual's residence which has been (or will be) treated at some point during the simulation. The definition does not change over time for one person, but may vary from person to person.

Untreated room

An area of a simulated individual's residence which is not directly treated, but which may become contaminated by diffusion/dispersion from a treated area.

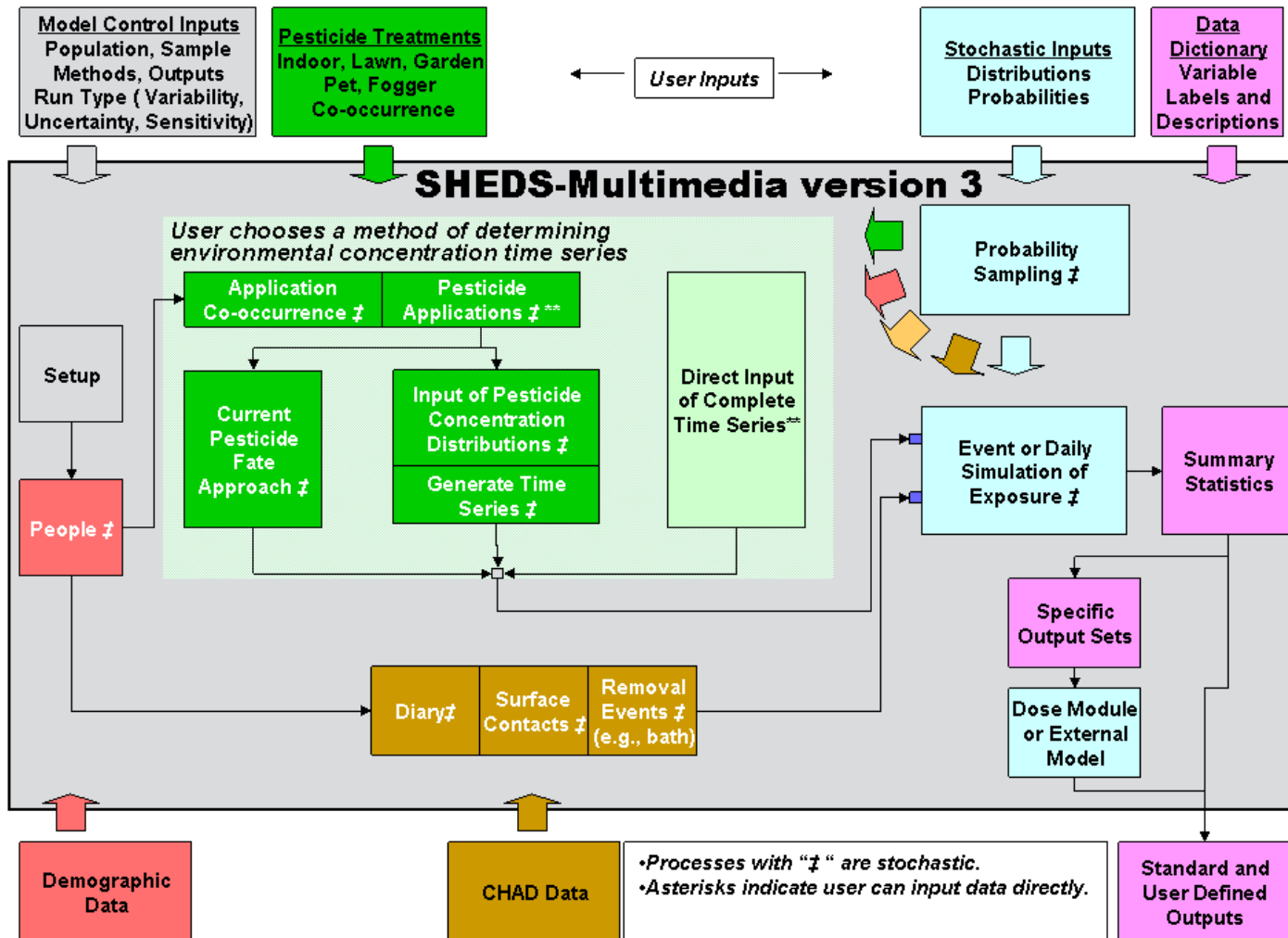


Figure B-1. Conceptual overview of SHEDS-Multimedia version 3 module linkages.

APPENDIX C: Description of Residential Code Input Files

List of Input Files

The following are the eight key input data files that the user works with to run the SHEDS-Multimedia version 3 model:

variables:	contains names, units, and descriptions of personal input variables, exposure factors, dose factors, and the re-sampling frequencies for these variables
distributions:	contains one or more probability distributions for each variable listed in the variables file, and the conditions indicating when each of these distributions should be used
correlations:	contains pairs of stochastic input variables and the desired Spearman correlation between them
contactmedia:	contains information on all the contact media (surfaces, air, or matter) which may expose a person to the target chemical
modeldates:	the list of application types and the rules for having the model randomly determine the dates when these applications occur
userdates:	the list of application types and the specific dates on which these occur (same dates for all persons)
appmedia:	a complete list of all combinations of application types and affected contact media that can be included in a model run
diarydetails:	provides distributions for the contact probabilities for the various contact media.

In addition, the following files are used internally by the model. These are applicable for a wide range of scenarios and may be used without alteration in most cases:

bwsa:	This file contains distributions for body weight, height, and basal metabolic rate, as functions of age and gender.
diarycohorts:	This file is used to group the activity diaries into pools.
diaryevents:	This is the large file containing all the diary events from all persons in the CHAD database.
diarylevels:	This file contains the probability matrix for selecting CHAD diaries with similar amounts of outdoor time, based on the overall tendency assigned to each individual.
diarylocs:	This file contains the mapping from CHAD location codes to the microenvironments used in SHEDS.
diarymets:	This file contains the activity-specific distributions for METS, which is the multiple of the basal metabolic rate needed to sustain the current level of activity.
diaryQARules:	This file contains a list of logical rules for accepting a CHAD diary for use in SHEDS-Multimedia 3.
diaryquest:	This file contains the personal data from CHAD that are not changing over the course of the diary day, like age, gender, home state, diary date, day of week, season, temperature, and many other variables.
pop2000:	This file contains a age-gender summary of the 2000 U.S. Census, without any geographical detail.

Files containing concentration time series may also be input, if the user requests it. For formatting details, see the chapter titled “MODEL INPUTS”.

Detailed Description of User-Specified Input Files

Each input file is a SAS data set, which is organized like a spreadsheet. The columns are called “variables” and the rows are “observations” or “records”. Each variable has a format, either a character string or numeric. Character strings are indicated by “\$n”, where “n” is the maximum number of characters the variable may hold. Numeric variables are indicated below by “num”.

The number of records on each file is flexible in SAS. Records may therefore be added or deleted without changing the code. However, additional records will not normally be of significance unless the code is altered to make use of them. Deleted records may cause missing values to be generated, if references to these are encountered in the code.

Variables file

The “variables” file contains background information on most of the stochastic input variables. It does not normally require changes from the user, and cannot be altered using the SHEDS interface. The “variables” file contains the following SAS variables:

Table C-1. Column definitions for the variable file.

column name	format	purpose
variable	\$24	name of input variable, as appearing in the code
label	\$90	description of the input variable
module	\$8	name of SHEDS module where the input is used
priority	num	order in which the records are to be processed
freq	\$8	the resampling frequency
units	\$16	the units for this input variable

The units are for informative purposes only; the code makes no use of them. The priority is normally not important, but the order of evaluation could matter if the descriptions file makes use of conditions that refer to other input variables. The default version of the variables file contains 40 records; that is, information on the following 40 input variables:

Table C-2. Definitions for variables from the distributions and variables files.

Variable	Description	Frequency	Units
absf_lung	absorption fraction for lungs	person	[-]
absr_dm	dermal absorption rate per day for dust or soil	person	1/day
absr_dr	dermal absorption rate per day for surface residues	person	1/day
absr_gm	GI tract absorption rate per day for dust or soil	person	1/day
absr_gr	GI tract absorption rate per day for surface residues	person	1/day
adherence	soil-skin adherence factor	person	mg/cm2
bathdays	maximum number of days between baths	person	days
bioavm	bioavailability fraction for dust/soil	person	[-]
bioavr	bioavailability fraction for surface residues	person	[-]
contactb	Body-surface fractional contact rate	hour	1/hr
contacth	Hand-surface fractional contact rate	hour	1/hr
dermaxb	maximum dermal loading for body	person	ug/cm2
dermaxh	maximum dermal loading for hands	person	ug/cm2
elimr_blood	elimination rate from the blood	person	1/day
f_uncloth	fraction of body unclothed	day	[-]
has_garden_p	probability of having a vegetable garden	person	[-]
has_lawn_p	probability of having a lawn	person	[-]
has_pet_p	probability of having a dog or cat	person	[-]
hm_fraction	fraction of surface of one hand that enters mouth	person	[-]
hm_freq	hand mouthing events per hour	hour	events/hr
ingestion_indoor	dust ingestion rate (indoor, direct only)	person	mg/hour
ingestion_outdoor	soil ingestion rate (outdoor, direct only)	person	mg/hour
metab_ratio	mass ratio (metabolite/pollutant)	person	[-]
object_ratio	object-surface concentration ratio	person	[-]
om_area	object-mouth contact area	person	cm2
om_freq	object-mouth contact rate	hour	events/hr

APPENDIX D: Tables for the One-Person, One-Day Simulation for Code Verification

The first table below (TableD-1) is 'Exposure_inputs', which lists in alphabetical order the variables input to the exposure module. This table has 30 rows, one for each of the 30 diary events. Each of the 50 columns represents one input variable and are spread over multiple pages. Some of these inputs retain the same value over the course of the day, while others may change hourly or change on every event.

Table D-1. Inputs to exposure for one-day calculation.

	Absf_lung	Absr_dm	Absr_dr	Absr_gm	Absr_gr	Adherence	Air
1	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
2	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
3	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
4	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
5	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
6	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
7	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
8	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
9	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	ATair
10	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
11	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	ATair
12	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	ATair
13	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	Anone
14	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AYair
15	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	Anone
16	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	ATair
17	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	ATair
18	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AYair
19	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AYair
20	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
21	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	Anone
22	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	Anone
23	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AYair
24	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AYair
25	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
26	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	ATair
27	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	Anone
28	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
29	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair
30	1	0.2360227506	0.03	0.1004152252	0.493773219	0.1691767658	AUair

