The Stochastic Human Exposure and Dose Simulation Model: High-Throughput

SHEDS-HT Beta Version 0.1.10

Quick Start Guide

April 2024

Prepared by: Kristin Isaacs

US Environmental Protection Agency, Office of Research and Development, Center for Computational Toxicology and Exposure

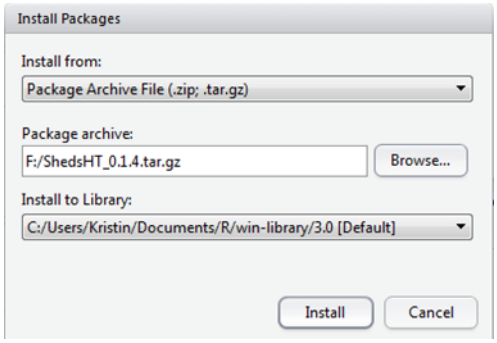
ACKNOWLEDGMENTS AND DISCLAIMER

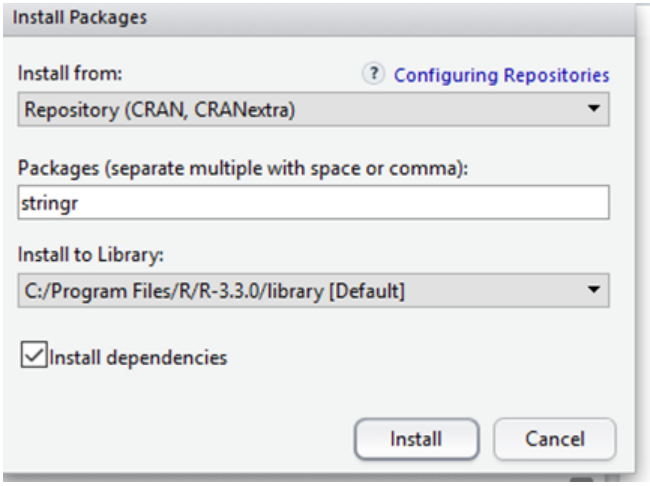
The United States Environmental Protection Agency through its Office of Research and Development funded and collaborated in the research and development of this software, in part under Contract EP-C-14-001 to ICF International. The model is publicly available in Beta version form. All input data used for a given application should be reviewed by the researcher so that the model results are based on appropriate data sources for the given application. This model, default input files, and R package are under continued development and testing. The model equations and approach are published in the peer-reviewed literature (Isaacs et al. Environ. Sci. Technol. 2014, 48, 12750-12759). The data included herein do not represent and should not be construed to represent any Agency determination or policy.

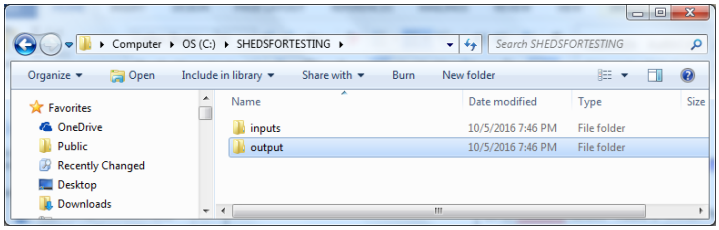
This tutorial will guide you through 1) Installing the SHEDS-HT R Package and corresponding data and 2) running an example run included in the R package (a run of chemicals identified via MSDS sheets as present in various categories of consumer products). This tutorial assumes that you have already:

- Installed R
- Installed RStudio
- Downloaded the most recent version of the SHEDS-HT R package. The most recent SHEDS-HT release can be obtained https://github.com/HumanExposure/SHEDSHTRPackage. Navigate to the folder "R Package", click on the current file name (e.g., "ShedsHT_0.1.X.tar.gz), and select the "Download" button:

Most of the below steps will only need to be done once. Also included with the SHEDS-HT distribution package is a script titled that, with editing for paths, will perform the steps in this tutorial.

1. Install SHEDS-HT within RStudio	
Start R Studio. Install SHEDS-HT package into your R environment. You will only need to do this once for each computer, or when you are installing an updated version of the R package.	Type the following command into the RStudio command line, giving the path to where ever you stored the SHEDS-HT package file. `>install.packages("F:/ShedsHT_0.1.10.tar.gz", repos = NULL, type = "source")` OR in RStudio, select "Tools"; "Install Packages"; and "Install from: Package Archive File" and browse to the file location. 

Load other R packages required by SHEDS-HT (**data.table**, **stringr**, **plyr**, and **ggplot**). You will only need to do this once for each computer.	Type the following commands into the RStudio command line
	
	`>install.packages("data.table")`
	`>install.packages("stringr")`
	`>install.packages("plyr")`
	`>install.packages("ggplot2")`
	
	OR
	
	in RStudio, select "Tools"; "Install Packages"; and "Install from: Repository" and type in a package name. Repeat for all 4 packages.
	
	 

2. Set up a SHEDS-HT Run	
Load the ShedsHT Package functions into the current R Studio session. (This must be done every time you start a new R session.)	Type the following command into the RStudio command line: `>library(ShedsHT)`
Decide on a SHEDS home folder location for storing your input and output files for any runs. This is where SHEDS will create "**inputs**" and "**output**" subfolders. You can create as many SHEDS-HT home folders as you want (with any name), accessing them at different times via the setup() step below. Here, for example, is a SHEDS home folder called "C:/SHEDSFORTESTING". (Note that the **inputs** and **output**	

folders will be created by SHEDS).	

Run the "**setup**" function on the SHEDS home location folder that you created above, so SHEDS knows where to store materials. (This must be done every time you start a new R session or want to switch between Home locations.)
 	
If when running "setup()" you receive the message
 	
`Error: cannot open file 'R/Sheds_HT.R': No such file or directory`
 	
Or on some systems:
 	
`Error in file(filename, "r", encoding = encoding) : cannot open the connection`
 	
then you have an old version of SHEDS-HT functions loaded in your R workspace. From the RStudio menu select" Session->Clear Workspace" and try again.
	Type the following command into the RStudio command line, using the name of your home directory that you created in Step 1:
	
	`> setup("C:/SHEDSFORTESTING")`
	
	The version information and EPA Disclaimer will display, for example:
	
	ShedsHT Version 0.1.10 (04/23/2024)
	Disclaimer
	The United States Environmental Protection Agency through its Office of Research and Development
	funded and collaborated in the research and development of this software, in part under Contract EP-C-14-001
	to ICF International. The model is publicly available in Beta version form. All input data used for a given
	application should be reviewed by the researcher so that the model results are based on appropriate data
	sources for the given application. This model, default input files, and R package are under continued develop
	and testing. The model equations and approach are published in the peer-reviewed literature
	(Isaacs et al. Environ. Sci. Technol. 2014, 48, 12750-12759). The data included herein do not represent
	and should not be construed to represent any Agency determination or policy.
	
	Inputs folder was created in the working directory.

If this is the first time you are using SHEDS, or if you have created a new home location, call **unpack**(). The default SHEDS input files included in the R package are written into the "inputs" directory. **Note: If you run unpack() again in the same home location it will overwrite any of the ShedsHT package default input files you have altered. (It will not overwrite new files you may have created - only those with the same names as files included with SHEDS).**	`>unpack()` This will produce the output: Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/activity_diaries.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/chem_props.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/diet_diaries.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/exp_factors.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/fugacity.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/media.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/physiology.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/population.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/run_artsandcrafts.txt Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/run_food_residue.txt Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/run_other_sources.txt Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/run_cpdatv3max.txt Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/source_chem_ac.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/source_chem_food.csv Unpacking C:/Users/kisaacs1/AppData/Local/Programs/R/R-4.2.2/library/ShedsHT/extdata/source_chem_cpdatv3_max.csv Etc.
3. Perform a SHEDS-HT Run	

Call the SHEDS run function with a SHEDS **Run file** as argument. SHEDS comes with several example run files described in the Technical Manual. They are located in the **input** directory of your SHEDS home location. Here, we run the "**run_artsandcrafts**" example, which runs all the current SHEDS-HT default product composition data for arts and crafts products, which was developed from EPA's CPDat data. Alternatively, one could run all of the default SHEDS-HT data from EPA's CPDat V3 (all product types and chemicals) by calling the run "**run_cpdatv3**" run file. Note: the **run_cpdatv3** run may take 2 or more hours to complete, depending on your computational resources.

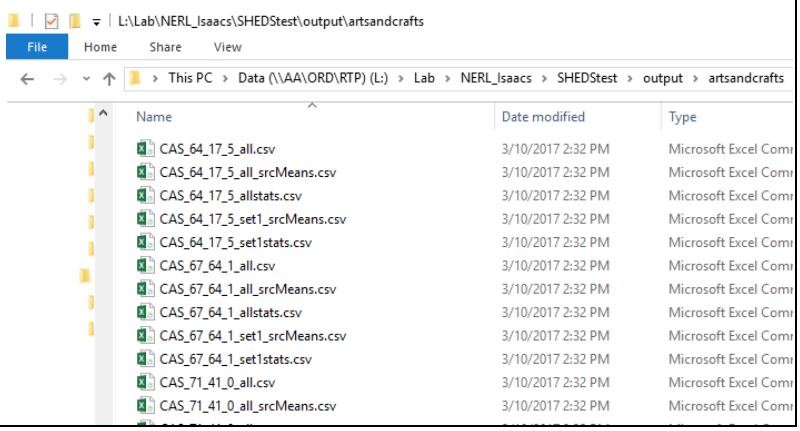
```
>run("run_artsandcrafts.txt")
```

This will produce the output:

```
run.name          = artsandcrafts
n.persons         = 10000
person.output     = 1
source.output     = 1
min.age           = 0
max.age           = 99
genders           = M F
season            = P S F W
details           = 1
age.match.pct     = 20
run.seed          = 876144637
set.size          = 10000
act.diary.file    = Activity_diaries.csv
chem.props.file   = Chem_props.csv
diet.diary.file   = Diet_diaries.csv
exp.factor.file   = Exp_factors.csv
fugacity.file     = Fugacity.csv
media.file        = Media.csv
physiology.file   = Physiology.csv
population.file   = Population.csv
source.vars.file  = Source_vars_products.csv
source.scen.file  = Source_scen_products.csv
source.chem.file  = source_chem_ac.csv
# chemicals       = 0

 Reading Activity Diaries completed
 Reading Chemical Properties completed
 Reading Dietary Diaries completed
 Reading Exposure Factors completed
 Reading Media File completed
 Reading Physiology File completed
 Reading Population File completed
 Reading Source.chemicals file completed
 Reading Source.variables file completed
 Activity Diary Pooling completed
 Dietary Diary Pooling completed
 General Factor Tables completed
 Media-specific Factor Tables completed

 Starting source  1  of chem  1  ( 100_41_4 ) of  21 ( Ethylbenzene  -  arts and crafts finish spray )
 set= 1 / 1  chem= 1 / 21     100_41_4    Ethylbenzene
      stat exp.der     exp.ing  exp.inh dos.inh   dos.intak     abs.der   abs.ing  abs.inh   tot.ug    tot.mgkg
1 pop.mean 1.06119 0.000417409 0.168857 1.49374 3.81011e-05 0.00299791 0.0160175 0.238999 0.258014 6.38525e-06
     ddd.mass
1 8.90311e-05

 Starting source  1  of chem  2  ( 106_97_8 ) of  21 ( Butane  -  arts and crafts adhesive spray )
 Starting source  2  of chem  2  ( 106_97_8 ) of  21 ( Butane  -  arts and crafts finish spray )
 Starting source  3  of chem  2  ( 106_97_8 ) of  21 ( Butane  -  body paint spray )
 set= 1 / 1  chem= 2 / 21     106_97_8    Butane
      stat exp.der     exp.ing exp.inh dos.inh  dos.intak  abs.der  abs.ing abs.inh  tot.ug   tot.mgkg  ddd.mass
1 pop.mean 69.8995 0.00652515 38.9344 416.709 0.00876872 0.366655 0.665219 66.6735 67.7054 0.00143334 0.0203228

 Starting source  1  of chem  3  ( 107_83_5 ) of  21 ( 2-Methylpentane  -  arts and crafts adhesive spray )
 set= 1 / 1  chem= 3 / 21     107_83_5    2-Methylpentane
      stat exp.der   exp.ing exp.inh dos.inh  dos.intak abs.der  abs.ing abs.inh  tot.ug    tot.mgkg  ddd.mass
1 pop.mean  35.414 0.0118858 16.9469 168.905 0.00420765 0.17051 0.489024 27.0247 27.6843 0.000686956 0.0134706
```

4. Locate the SHEDS-HT Output Files

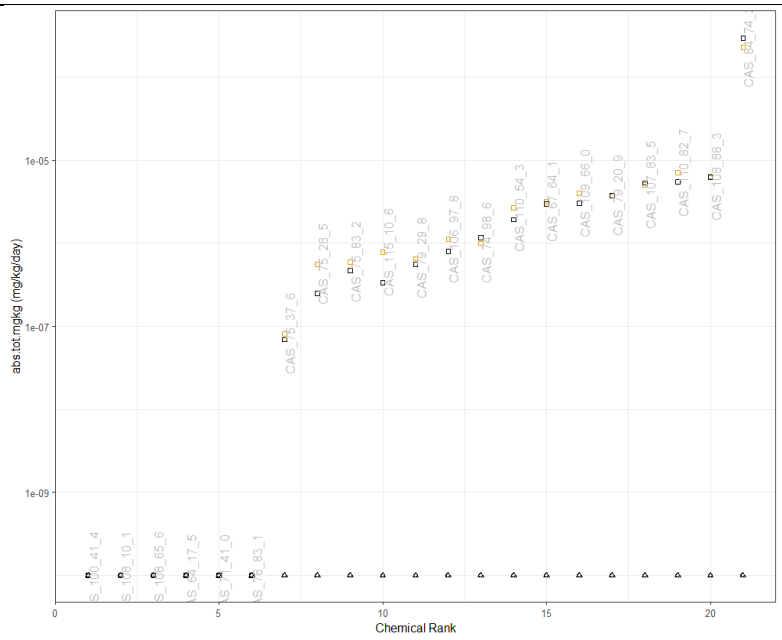
Examine the SHEDS Output files in the Output folder of the SHEDS home folder you designated with **setup**(). The files will be in a subfolder under "Output" with the run name provided in the run file (E.g., either "artsandcrafts" or "CPDatV3max").



5. Combine SHEDS Output for All Chemicals into a Summary File	

Combine the Percentile Data for all chemicals (i.e. the "AllStats" files created in the output folder) into a single file for other analyses, using the combine_output function, providing a run name and a name for the output file. The file is placed in the output folder for the run.	`combine_output(run.name="artsandcrafts", out.file="artsandcrafts.csv")`
	
	This will produce the output:
	
	`Processing chemical 1 of 21`
	`Processing chemical 2 of 21`
	`Processing chemical 3 of 21`
	`Processing chemical 4 of 21`
	`Processing chemical 5 of 21`
	`Processing chemical 6 of 21`
	`Processing chemical 7 of 21`
	`Processing chemical 8 of 21`
	`Processing chemical 9 of 21`
	`Processing chemical 10 of 21`
	`Processing chemical 11 of 21`
	`Processing chemical 12 of 21`
	`Processing chemical 13 of 21`
	`Processing chemical 14 of 21`
	`Processing chemical 15 of 21`
	`Processing chemical 16 of 21`
	`Processing chemical 17 of 21`
	`Processing chemical 18 of 21`
	`Processing chemical 19 of 21`
	`Processing chemical 20 of 21`
	`Processing chemical 21 of 21`
	`Combining data...`

5. Visualize Run Outputs	
Due to the fact that the outputs are so rich, custom visualization of SHEDS results are usually desirable, but SHEDS does provide a two simple functions (see User Manual Section 5) for visualizing results. These include **allstats.variable.rank.plot** which provides a plot of select output variables and statistical metrics across chemicals.	`allstats.variable.rank.plot("artsandcrafts", cohort.col = c("Total", "Females"),output.variable = "abs.tot.mgkg", metrics = c("5%","50%","95%"),label_chem=T)` This will produce the output:



The **puc.rank.plot** function plots the source means (across people) different source IDs (here product categories). These data reside in the Source Means output files for each chemical.

```
puc.rank.plot("artsandcrafts",label_chem = T)
```

This will produce the output: