**Source to Dose (S2D) Code User’s Guide**

|  |  |
| --- | --- |
| **To:** | Dan Vallero, Kathie Dionisio, Peter Egeghy, Kristin Isaacs, Paul Price, U.S. EPA NERL |
| **From:** | Graham Glen, Raga Avanasi, Jessica Levasseur, ICF |
| **Date:** | May 18, 2017 |
|  | |

# Introduction

The Source to Dose (S2D) module will be used in the larger Human Exposure Model (HEM) to provide higher-throughput estimates of exposure to chemicals in consumer products based on existing product chemical function and composition databases. S2D is a longitudinal exposure model that accounts for different exposure pathways and routes. S2D will expand upon the basic exposure scenario algorithm functions built into SHEDS-HT, but will operate on an hourly event time step and retain daily results.

This document describes how to run S2D with a basic example, submitted to EPA on May 18, 2017.

# Installation

After unzipping the S2D folder, users may wish to note the following significant files and folders that are referenced in the installation instructions below.

## Code

Users should have access to HEM\_s2d\_05182017.R, as this contains the main code for running the S2D module in HEM.

## Input files

Input files necessary to run S2D are listed below alphabetically:

* **basal\_vent.csv:** This file contains information for estimating basal ventilation rate for household members based on age and sex.
* **chemical\_properties\_HEM\_withfabs.csv:** This file contains a list of chemical properties relevant to calculating fugacity and human exposure.
* **composition\_for\_S2D.csv:** This file contains different formulations of PUCs as well as the weight fractions of chemicals present in those formulations.
* **Fugacity.csv**: This file contains chemical-independent variables used to evaluate the fugacity equations.
* **pophouse.csv**: This is an example output from the Residential and Population Generator (RPGen) module. This example file describes three households in a defined population, and identifies the primary person in each household.
* **PUC\_MET.csv**: This file contains the associated CHAD activity code and the mean value of METS associated with product use, as calculated by APEX.
* **PUC\_producttype\_releases\_LookUp.csv**: This file contains a list of fractional partitioning values per the 11 different compartments considered in S2D. PUCs are sorted into one of 30 different groups to accurately describe how the product will partition after use.
* **PUC\_producttype\_releases\_Values.csv**: This file assigns each PUC a different code for accurately describing how the chemicals within this PUC will partition after use.
* **PUC\_skin\_areas.csv**: This file identifies which body areas are exposed to a PUC during use, per PUC. Fraction of the PUC that is associated with dermal contact on the body area specified are provided for both adults and children, as well as a simple explanation of the variability expected in the amount of dermal exposure the body area will experience.
* **PUC\_skin\_wipe\_rinse\_data2.csv**: This file describes the fractional partitioning of PUCs on the hand and/or body and how they are affected by rinsing or wiping actions.
* **S2D\_control\_file.txt**: A control file for running S2D. Users will interact primarily with this file. *Users should note that the version of this file provided in this zipped folder of inputs and code is an example which will produce the example output also provided.*

In addition, users may have one or more ABM diaries output from the ABM module. ABM outputs one .csv file per household run in ABM, so users may have any number of ABM diaries available to them. These are all named “Household\_n.csv”, where n is the household number.

Note that these input files will be saved in folders that users will create. This is described in more detail in Section 4.

# Packages Required

S2D was developed in R (x64 version 3.2.2), and tested in R (x64 version 3.3.1) by ICF. To run S2D, the following packages are required, though these packages are automatically loaded when sourcing the R code:

* data.table
* stringr
* plyr
* dplyr
* dtplyr
* ggplot2
* rmpfr
* xlsx

Other packages loaded into a user’s R or R studio console should not impact the S2D module.

# Running the Code

After examining the unzipped file to ensure that all files listed in Section 2 are present, users should follow the directions below to run S2D.

To run S2D:

1. Users should choose a working directory, and save “HEM\_s2d\_05182017.R” to that location.
   1. Open “HEM\_s2d\_05182017.R” and edit line 8 to reflect your chosen working directory.
2. Within this working directory folder, users should create an “input” and “output” folder (note that these files names are case sensitive and should be created exactly as listed here).
3. Within the output folder, create folders for each specific S2D run. The example included with the S2D code uses the run name “new”, so create a folder within the output folder named “new”. Move “pophouse.csv” and any ABM diaries (“Household\_n.csv”) to this folder. Note that S2D requires sequential household numbers on the ABM diaries at this time, though more flexibility will be added at a later time.
4. Move “S2D\_control\_file.txt” and all remaining .csv files to the “input” folder.
5. Open “S2D\_control\_file.txt” to customize the run:
   1. Edit the run.name line to name the folder in which your output will be saved (within the “output” folder); *In the provided example, as noted in step 3 above, the run name should be “new”.* Note that if the same run name is used for multiple runs, each will overwrite the output from previous runs with the same name.
   2. Choose which chemicals you would like to run in S2D by entering multiple DTXSID values, if applicable (one per line; be sure the repeated lines begin with ‘chem’ if you choose to list more than one); if you would like to run all chemicals, delete all ‘chem’ lines
   3. Choose which PUCs you would like to run in S2D by entering PUC ID values, if applicable (one per line; be sure the repeated lines begin with ‘puc’ if you choose to list more than one); if you would like to run all PUCs, delete all ‘puc’ lines
   4. Choose the number of household diaries you would like to run by editing the ‘n.houses’ line with a numeric value (cannot exceed the amount of ABM diaries you have available)
   5. Choose a seed number for random variables by altering the value next to “init.seed”, if desired
   6. Save and close the “S2D\_control\_file.txt” when complete. Do not change the file name.
   7. **When editing the “S2D\_control\_file.txt”, be sure to use spaces and not the tab key to indent the values associated with each variable. At least one space should be used, though more may also be used.**
6. In “HEM\_s2d\_05182017.R”, source the code.
7. Call the function by typing “s2d()” in the R console.
8. Output files can be found in the folder within the “output” folder named as defined by the user (the “run.name” value described in step 5.a above). More information on these output files is found below in Section 6.

# Running the Code: Options

Users should check the “chemical\_properties\_HEM\_withfabs.csv” file to be sure that their DTXSID number of interest is available in this file. If the DTXSID number is not listed in that file, the S2D code will not work properly.

Users may also wish to investigate a particular PUC found in the diaries. If a particular PUC/chemical combination is desired for investigation with S2D, users should ensure that their PUC of interest contains the chemicals listed. This can be verified in the “composition\_for\_S2D.csv” file. Users should also ensure that the PUC is used in the ABM diaries selected by checking the “Household\_n.csv” ABM diary file(s). A message will return if an invalid PUC/chemical combination is attempted.

Users may change input file names, if the new names are also appropriately updated in “S2D\_control\_file.txt”.

# Output

Two to four types of output files will be created within the run name folder (within the “output” folder) initially created by the user, depending on the type of run performed. The two output files that will always be created are “S2D\_annual\_n.csv” and “S2D\_daily\_ n.csv”, where n is the household number. One daily and one annual file will are written for each household, except for when all values returned are zero (then neither file is written). Both files contain the variables listed below in Table 1, except for that the annual output file will not include “daynum".

Table 1. S2D Summary Output Variable Descriptions

| Variable name | Description of Output Variable |
| --- | --- |
| household | Household number from ABM output |
| daynum | The number of the day in a year |
| DTXSID | DSSTox Chemical ID |
| dir.derm.exp | Mass (mg) of chemical on skin from direct product use |
| dir.derm.max | Maximum loading (mg/cm2) on either hands or body (direct use) |
| dir.derm.abs | Mass (mg) absorbed through skin (direct use) |
| dir.inhal.exp | Average air concentration (mg/m3) for primary person (direct) |
| dir.inhal.max | Maximum air concentration (mg/m3) (direct) |
| dir.inhal.mass | Direct mass inhaled in the day (mg) |
| dir.inhal.abs | Mass (mg) absorbed through the lungs (direct use) |
| dir.ingest.exp | Mass (mg) ingested from hand-to-mouth transfer as well as transferred to the gut compartment during initial use phase (direct) |
| dir.ingest.abs | Mass (mg) absorbed through the gut (direct use) |
| release | Mass (mg) of chemical released by product use (primary person) |
| ind.derm.exp | Mass (mg) of chemical on skin from indirect exposure |
| ind.derm.max | Maximum loading (mg/cm2) on either hands or body (indirect) |
| ind.derm.abs | Mass (mg) absorbed through skin (indirect exposure) |
| ind.inhal.exp | Average air concentration (mg/m3) for primary person (indirect) |
| ind.inhal.max | Maximum air concentration (mg/m3) (indirect) |
| ind.inhal.abs | Mass (mg) absorbed through the lungs for (indirect) |
| ind.inhal.mass | Indirect mass inhaled in the day (mg) |
| ind.ingest.exp | Mass (mg) ingested from hand-to-mouth transfer (indirect) |
| ind.ingest.abs | Mass (mg) absorbed through the gut (indirect) |
| out.sur | Mass (mg) released to outdoor surfaces (outdoor product use) |
| out.air | Mass (mg) released to outdoor air (outdoor use and house losses) |
| drain | Mass (mg) going into the drain system (includes chemical in the hair and nail compartments) |
| waste | Mass (mg) disposed in trash |

In addition, if a single PUC is run in S2D then a Life Cycle Impact Assessment (LCIA) output will be created as “LCIA\_n.csv”. If more than one household is run, then an overall average LCIA output will be output as well, titled “LCIA\_avg.csv”. Neither file will be written when all variable values returned are zero. The variable names on these two output files will be the same, as seen in Table 2, though will be report averages within a household or averages across all households run, respectively.

Table 2. S2D LCIA Output Variables

| Variable name | Description of Output variable |
| --- | --- |
| DTXSID | DSSTox Chemical ID |
| source.id | Unique identifier of product use category (PUC ID) |
| mass.frac.chem.puc | Fraction of the chemical in the product (an average across all people and all uses) |
| mass.puc.use.adult | Amount of product used by an adult annually (g) |
| PIF.ingest.adult | Adult Product Intake Fraction for ingestion exposure route |
| PIF.derm.adult | Adult Product Intake Fraction for dermal exposure route |
| PIF.inhal.adult | Adult Product Intake Fraction for inhalational exposure route |
| mass.puc.use.child | Amount of product used by a child annually (g) |
| PIF.ingest.child | Child Product Intake Fraction for ingestion exposure route |
| PIF.derm.child | Child Product Intake Fraction for dermal exposure route |
| PIF.inhal.child | Child Product Intake Fraction for inhalational exposure route |
| chem.mass | Amount of chemical used annually (g) |
| mass.tot.air | Amount of chemical released into air annually (g) |
| mass.tot.water | Amount of chemical released into water annually (g) |
| mass.tot.land | Amount of chemical released into land annually (g) |