

Demo Run Instructions

Population PFAS TK Model

Updated 2026-06-26

This document provides a worked example using the Population PFAS Toxicokinetic Model from the Journal of Environmental Exposure Assessment publication *Estimating Drinking Water PFAS Exposures Associated with Serum Clinical Action Levels Using a Probabilistic Toxicokinetic Model* by Nielsen et al., 2026. Model limitations and disclaimers regarding the model are also presented and should be read carefully before the model is used.

Limitations and Disclaimers. This code is made available to support transparency, reproducibility, scientific review, and appropriate implementation of the methods described in the publication.

This model and code are research materials only. They are not a validated clinical, diagnostic, or medical decision-making tool. They are not intended to provide medical advice, diagnosis, treatment recommendations, or individualized exposure or health determinations. The model should not be used as the sole basis for determining whether any individual should receive clinical testing, treatment, medical monitoring, or follow-up. Any health-related interpretation or application of model outputs should be conducted by qualified professionals using appropriate clinical, exposure, toxicological, and public health information.

Publication of this code does **not** establish, revise, or supersede any regulatory standard, guidance value, imminent hazard value, policy, enforcement position, permit condition, or other official determination of the Massachusetts Department of Environmental Protection (MassDEP), the Executive Office of Energy and Environmental Affairs (EEA), or the Commonwealth of Massachusetts. Any regulatory, enforcement, site-specific, or public health use of the model requires independent professional judgment, review of applicable law and guidance, and consideration of relevant facts, assumptions, uncertainties, and limitations.

Model outputs are estimates and are subject to uncertainty. Users should review the associated publication, documentation, assumptions, limitations, and uncertainty analyses before applying or interpreting the model or code.

The code and associated materials are provided “**as is**”, without warranty of any kind, express or implied, including but not limited to warranties of accuracy, completeness, reliability, security, compatibility, merchantability, fitness for a particular purpose, or non-infringement. Potential incompatibilities with other software, systems, packages, or computing environments have not been fully assessed. Users are responsible for evaluating the code, validating results, complying with applicable laws and requirements, and determining whether and how the code is appropriate for their intended use.

The views, conclusions, and model outputs generated by users of this code do not necessarily represent the views, policies, or positions of MassDEP, EEA, or the Commonwealth of Massachusetts. Peer review of the associated publication and code does not constitute approval, certification, or validation of the model for any particular clinical, regulatory, legal, commercial, or site-specific application.

Overview

The population PFAS Toxicokinetic model predicts the distribution of serum PFOA, PFOS, PFHxS, PFNA, PFDA, and PFHpA levels for defined populations. Each defined population for modeling along with the drinking water exposure conditions is referred to as an ‘exposure scenario’. Exposure scenarios are specified using an inputdata spreadsheet where users enter details on the populations’ drinking water exposures, maternal serum information, birth date, modeling end date, and number of Monte Carlo iterations. A detailed description of how to use the inputdata spreadsheet is available in the spreadsheet’s ReadMe tab. Modeling is performed in R using data entered in the inputdata spreadsheet.

Demo Exposure Scenario

The defined population for the model demo is males with lifetime exposure to 20 ppt of PFOA, PFOS, PFHxS, PFNA, PFDA, and PFHpA with all chemicals modeled individually. Modeling occurs from birth through age 70. The maternal serum exposure uses the model’s feature to input a distribution of maternal serum PFAS concentrations for each chemical (in this case, the maternal serum input is the modeled distribution of serum levels for six PFAS for a population of women age 40 with lifetime exposure to 20 ppt for each PFAS.) The breastfeeding duration for the male population is set to six months. The modeling set up for this population is entered in the inputdata spreadsheet.

Files and folders for demo model run in the Population-PFAS-TK-Model folder:

[V2.4/2026-06-25_Set01_Demo.R](#)

This script runs the model demo.

[V2.4/Set01_Demo/inputdata_Set01_Demo.xlsx](#)

The set up for this demo exposure scenario is entered in the “Scenario” tab of the “inputdata_Set01_Demo.xlsx” excel spreadsheet. There are six exposure scenarios listed in rows 2-7, one for each PFAS. The columns for the demo exposure scenarios are as follows:

id: A unique identifier for each exposure scenario. These are constructed in the ScenSummary tab with the Set number, the Exposure Scenario number, and abbreviations for key exposure information. In this example, the chemical is the key exposure information.

filegroup: An identifier organizing output files. Here, “F2_20ppt” for all six scenarios to save into a single raw data file.

chemical: The PFAS modeled in each exposure scenario. The demo has one scenario for each of the six modeled PFAS.

water_concs: The drinking water concentration(s) of the PFAS, in ng/L. In this example, 20 ng/L.

water_concs_dates: A start date for the drinking water PFAS concentration(s) typically set to a date before modeling begins. In this example, set to January 1, 1950.

water_concs_sds: The standard deviation(s) of the drinking water PFAS concentration(s). Set to 0 in this example for this scenario with a constant lifetime exposure.

water_concs_dists: The distribution shape(s) for the water concentration. Set to none for this modeling scenario.

mom_water_conc: Information about maternal serum or drinking water concentration. In this example, we use serum distributions of an adult female model as the maternal serum values. Each row contains an identifier matching list names inside the file

LogsplineFits_Set00_MaternalSerumDemo.RData

mom_serum_method: Information on the format of the maternal serum. In this example, ‘logspline’ indicates that the maternal serum is a logspline-fit distribution.

gender: Male for this modeling run. Other options include “Female” and “Unknown.”

fed_mo: The duration of breastfeeding. Set to 6 months in this example.

niter: The number of Monte Carlo iterations for distributional parameters. Set to 10000 for this demo. Decrease to 1000 for all chemicals for faster run time.

sd: The start date for modeling, corresponding to the population’s birth date. Set to 2025-01-01 in this demo.

ed: The end date for modeling. Set to 2095-01-01 to model 70 years of exposure.

timezoom_start: A beginning date for more frequent calculation of serum PFAS. Set to 2026-01-01 in this example.

timezoom_end: An end date for more frequent calculation of serum PFAS. Set to 2035-01-01 in this example.

timezoom_interval: Set to '1 month' in this example. These timezoom settings cause serum to be calculated monthly from ages 1 through 10 years, instead of the default annual calculation. It allows a chart focusing on childhood to have better time resolution. This more frequent calculation will not affect total serum values.

saved_dates: Set to 'all' in this example. This means serum levels will be saved at all the dates where serum is calculated.

[V2.4/Set00_MaternalSerumDemo/LogsplineFits_Set00_MaternalSerumDemo.RData](#)

The LogsplineFits_Set00_MaternalSerumDemo.RData file contains the maternal serum distribution data for each PFAS. It is added using the 'maternalID' in the execution script.

Demo Model Run

The exposure scenario setup has already been completed for this demo and the execution script should run as-is. Each exposure scenario is assigned a unique scenario ID and usually abbreviated "ScenID" (for this model run, each chemical has a unique ScenID). Each model run (which may contain several exposure scenarios) is organized using a unique setID. Here, the setID is "Set01_Demo". The model is executed from the working script "2026-06-25_Set01_Demo.R". This script references inputs and creates outputs inside the directory "Set01_Demo".

The demo is set to run 10,000 Monte Carlo iterations for each chemical which takes a total of ~45 minutes to run on an Intel Core Ultra 7 155U processor with 32 GB of RAM. Decreasing the number of iterations allows the model to run faster but increases run-to-run variation in output due to the nature of the Monte Carlo modeling. The number of iterations can be changed in the inputdata spreadsheet "Scenario" tab column label "niter". All exposure scenarios in a model run should have the same "niter" to avoid problems in the automatic post-processing.

Demo Modeling Instructions

Download the “Population-PFAS-TK-Model” folder to your computer.

Open the project “PFAS_TK_Model_MassDEP_V2.4.Rproj” in RStudio (file path: V2.4/PFAS_TK_Model_MassDEP_V2.4.Rproj).

Open the working script “2026-06-25_Set01_Demo.R” (file path: V2.4/2026-06-25_Set01_Demo.R).

Run the functions in 2026-06-25_Set01_Demo.R starting from the top, line-by-line. The code is well-commented to follow the purpose of specific code lines and functions.

Additional notes to facilitate the demo model run are below:

- The package.check process loads packages and installs any required packages that are not present.
- The setID object defines the ID for the set, in our case “Set01_Demo”. There must be a directory exactly matching this setID with a file named “inputdata_SetID.xlsx” inside. Model runs to consider new scenarios should use a new setID with a new associated inputdata Excel file.
- The maternalID object defines the ID for maternal serum used in the exposure scenarios, if they are determined using a previous model run. The maternalID should be the setID of that original model run.
- Model.All.Scenarios() models all the exposure scenarios defined in the rows of the inputdata spreadsheet ‘Scenarios’ tab. Run time depends on the number of iterations ‘niter’ and number of scenarios. If interrupted, processing can be resumed by rerunning this line, but this can lead to unusual overwrite behavior (see the ReadMe_Population-PFAS-TK-Model_V2.4.pdf). When this function finishes, the results are stored in files with names formatted as RawData_SetID_filegroup.RData. ‘filegroup’ is an entry in the inputdata spreadsheet under tab “Scenario”. All downstream functions will load data from these RawData files, so the time-consuming Model.All.Scenarios() will not need be rerun later.
- Extract.Serum.at.Dates() extracts the modeled serum levels of each Monte Carlo iteration for a small number of dates of interest in each exposure scenario. These scenarios and dates are determined by entries in the inputdata spreadsheet tabs “FitDistInfo”, “ProjectionInfo”, and “EvalInfo”. In this demo, only “FitDistInfo” is used, which highlights two dates for each exposure scenario. These correspond to the population at age 6 months and 40 years. For large groups of scenarios or high niter values, this function may take several minutes to run. When this function finishes, the results are stored in a file with name formatted as

SerumatDates_SetID.RData. Downstream analysis functions such as Check.Stability() will load from this file, so the time-consuming Extract.Serum.at.Dates() will not need to be rerun later.

- Prep.Lifetime.Plot.Data() calculates specified quantiles of serum PFAS for each exposure scenario at all dates modeled. Default quantiles are 0.05, 0.25, 0.5, 0.75 and 0.95. Additional quantiles can be specified in the argument 'addlquants'. This function may take several minutes to run for large groups of scenarios or high niter values. When this function finishes, the results are stored in AllDatesQuants_SetID.RData, so it does not need to be re-run later. An identical .csv format file is created for manual review or export to other analysis software, like Excel.

Demo Model Output

After running all functions in “2026-06-25_Set01_Demo.R”, the Set01_Demo directory should contain the following files:

- AllDatesQuants_Set01_Demo.RData
- AllDatesQuants_Set01_Demo.csv
- Inputdata_Set01_Demo.xlsx
- RawData_Set01_Demo_F2_20ppt.RData
- SerumatDates_Set01_Demo.RData
- StabilityCheck_Set01_Demo.csv
- Subdirectory “SerumDistOverTimePlots” containing six .png images

Model output for this demo is similar to the lifetime serum trajectories shown in Figure 2 of the publication *Estimating Drinking Water PFAS Exposures Associated with Serum Clinical Action Levels Using a Probabilistic Toxicokinetic Model* but run with fewer Monte Carlo iterations for shorter run time.