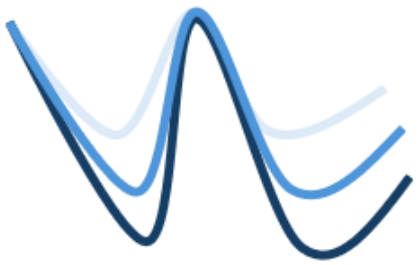
WATSON
COST ACTION

WITS- toolbox quick guide:



Water
Isotope modeling for
Transit time and **S**torage

Quick guide to the WITS (Water Isotope modelling for Transit time and Storage) toolbox. Final report of working group 3 of COST Action CA19120: WATer isotopeS in the critical zONe from groundwater recharge to plant transpiration (WATSON)

https://watson-cost.eu/

This publication is based upon work from COST Action WATSON, CA19120, supported by COST (European Cooperation in Science and Technology).

COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation.

www.cost.eu

Quick guide to the WITS software

DISCLAIMER

This report documents version 1.0 of the software WITS, a toolbox for estimating water transit times and storage volumes in the hydrological cycle. The toolbox is a collection of publicly available model codes. The different model codes implemented in this toolbox have been verified against a number of test cases. However, no warranty is given that the toolbox is completely error-free. If you do encounter problems with the code, find errors, or have suggestions for improvement, please contact info-wits@uu.nl.

The latest version of WITS can be downloaded from Github (https://github.com/srlutz/wits) and from the WATSON website (https://watson-cost.eu/storage-and-transit-time-estimation-in-hydrological-systems/).

The use of WITS can be acknowledged as follows:

Correa, A., Farlin, J. Lutz, S., Müller, S., Stockinger, M., The WITS-toolbox: Water Isotope modeling for Transit time and Storage. User manual, WITS, Version 1.0, Florence University, 27 pp., 2024.

Before you start

Please download the R programming language (https://cran.rstudio.com/). We recommend using Rstudio which can be downloaded here: https://posit.co/download/rstudio-desktop/). Also download the most recent version of the WITS toolbox from Github (https://github.com/srlutz/wits) or the WATSON website (https://watson-cost.eu/storage-and-transit-time-estimation-in-hydrological-systems/).

WITS can be used in any operating system supporting R. The developer team recommends using RStudio to avoid graphical issues with the predefined plots.

Contents

General Information

The current Quick Start guide will guide you briefly through a test case, we chose to simplify the first steps into the toolbox. You will be guided through all ´prompt windows´, which occur after executing the main script. This will give you a brief overview of all potential input parameters used for the variety of models incorporated in this toolbox. Detailed information on toolbox structure, model setup, data requirements and user recommendations can be found in the User-Guide.

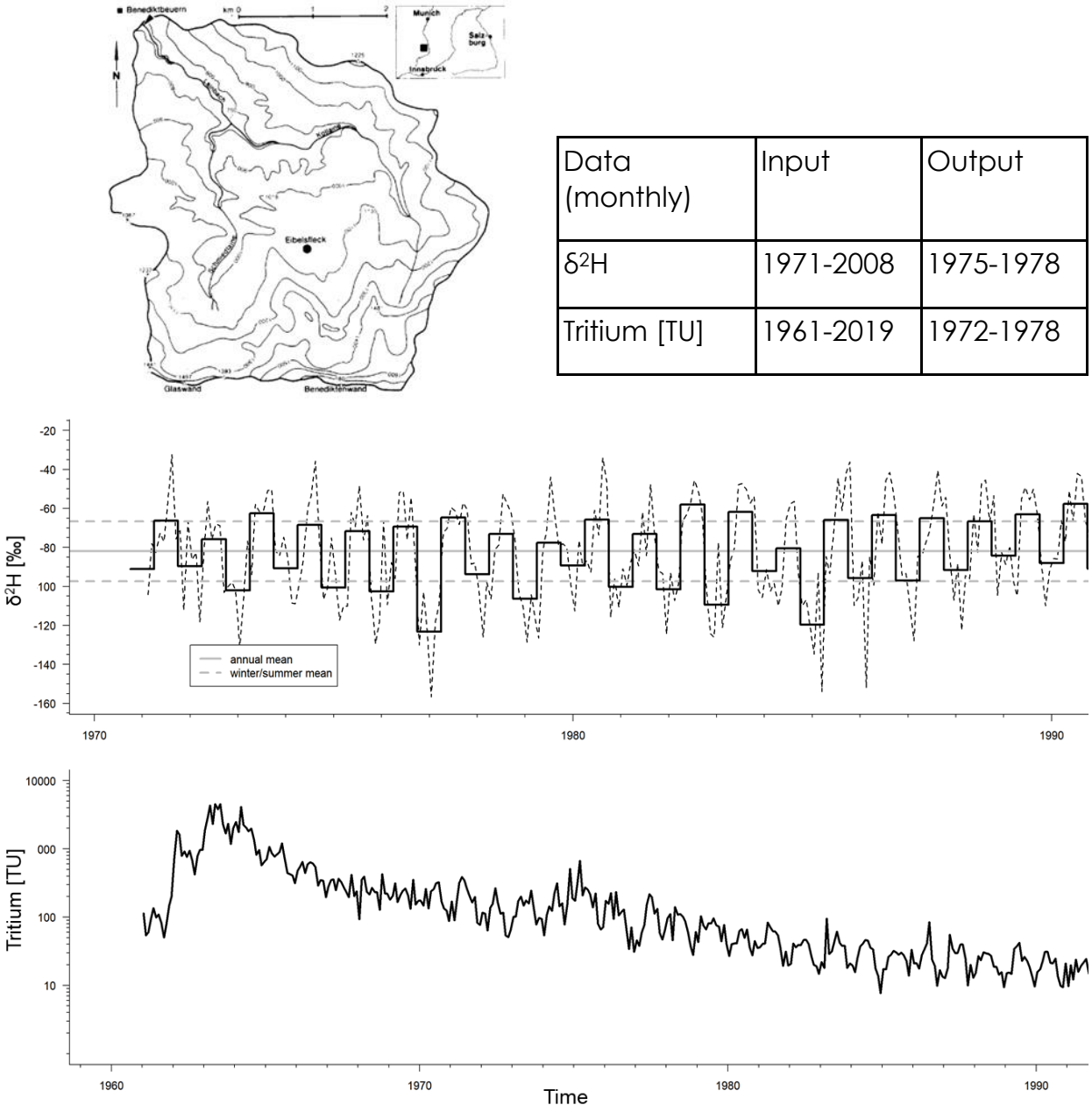
Running WITS

For the example provided in this quick guide, the data is in the folder **´\data\additional_data´**. The data folder used by WITS is **´\data\raw´,** so the sample data is also already copied into that folder. The main executable script and necessary model functions can be found in **´\src…´.** Additional literature and the User Guide can be found **´\docs…´.** The results of successful model executions are stored in **´\results…´.** The toolbox provides a basic graphical output for ad-hoc visual evaluation of the model results, but the users are encouraged to create their own graphical output once satisfied with the results. The raw data of the simulated output is stored in ´.txt-file format´.

Input data

For this example, we use data from the Lainbach catchment (

Figure 1) to demonstrate how the WITS operates with real-world datasets.



Data (monthly)	Input	Output
δ^2H	1971-2008	1975-1978
Tritium [TU]	1961-2019	1972-1978

Figure 1. Example of input data to the model. This quick guide uses data from the Lainbach catchment south of Munich, Germany (detailed map, top row left). Various isotope data is available for that station or the nearest tritium precipitation sampling station (Vienna). The center plot shows a 20-year deuterium (δ^2H) time series (dashed lines) and their respective six-month mean values for summer and winter periods. The lower plot shows the atmospheric tritium concentration reconstructed for the closest tritium measurement station in Vienna.

Running the script

First, browse to the respective folder, where you installed the toolbox. Go to ´**\src…**´ and load and execute the main script in RStudio: ´**main_script_0_7.R**´.

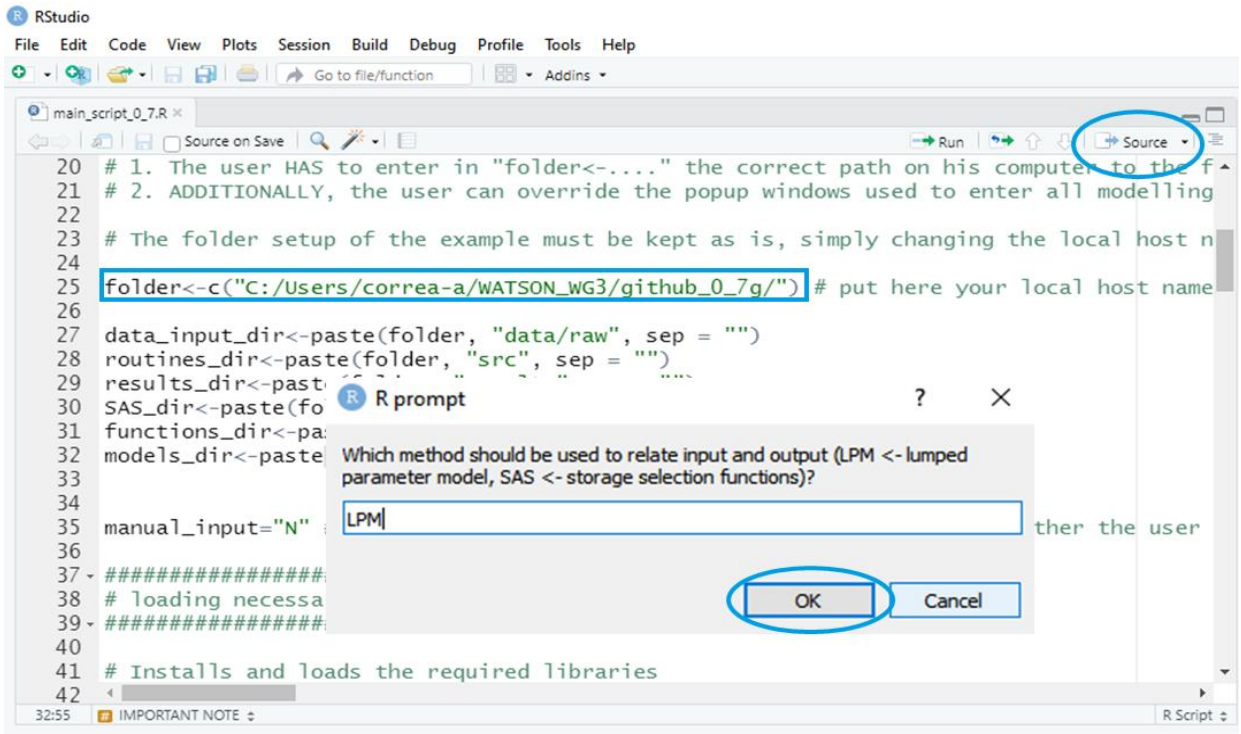
The folder setup of the example must be kept as is, simply changing the local host name to the main folder

folder<-c("C:/exampleuser/exampleuser/exampleWITS/exampleWITSversion/")

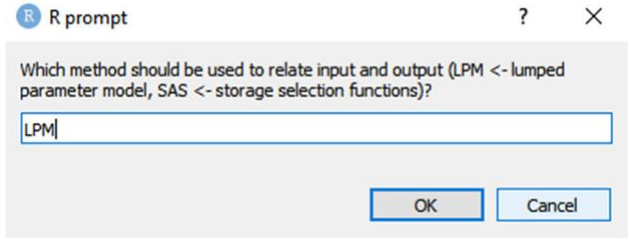
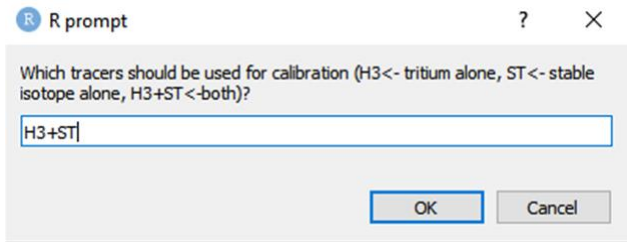
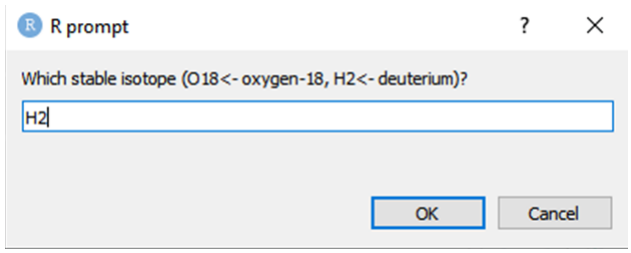
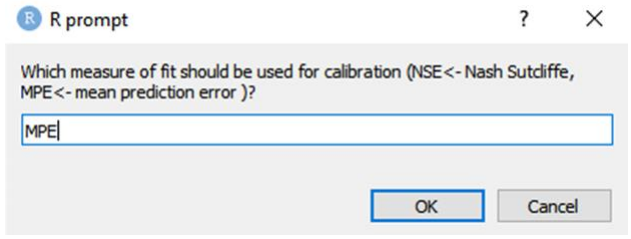
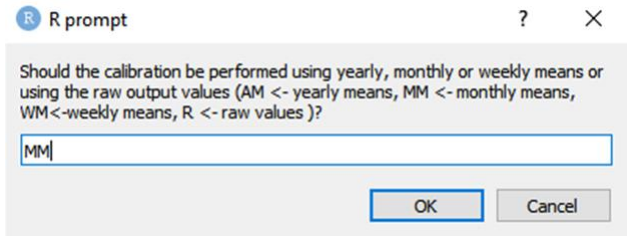
put here your local host name.

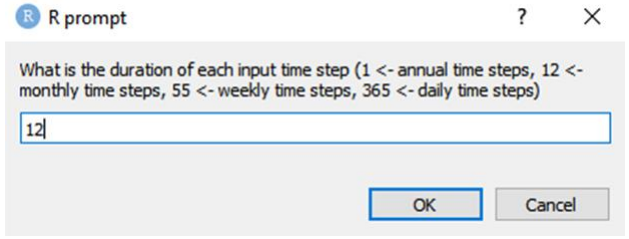
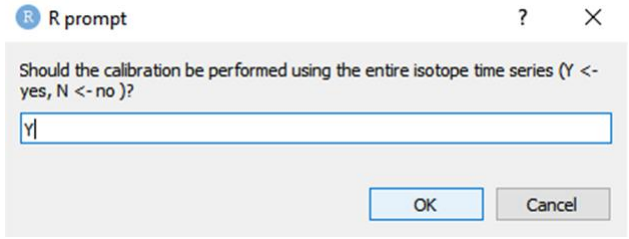
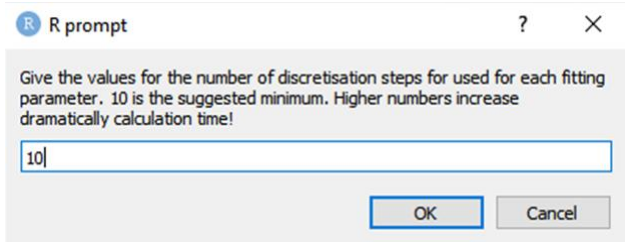
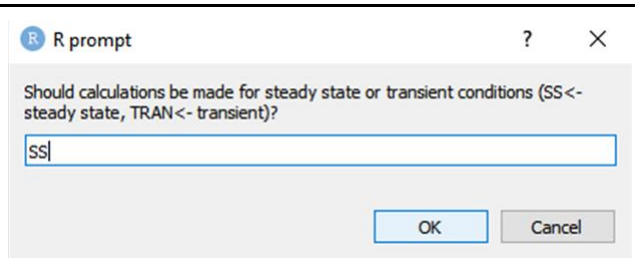
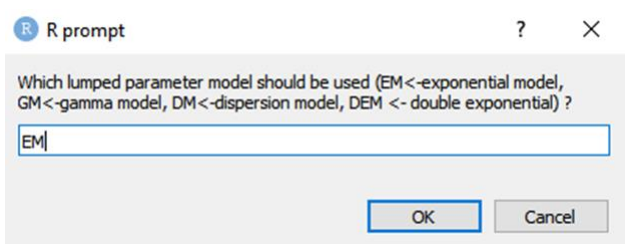
The options selected in this guide are the ones **used by the authors**, we recommend using the same settings for your initial run to familiarize yourself with WITS. This will help you to follow along with the guide and understand the process step by step.

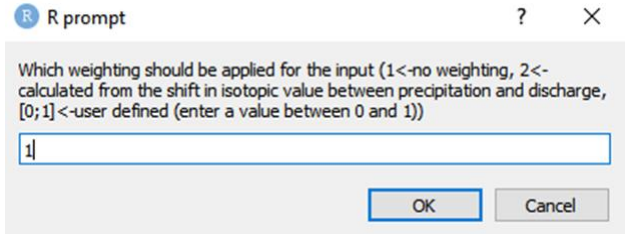
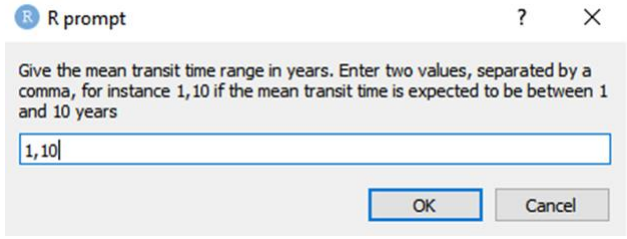
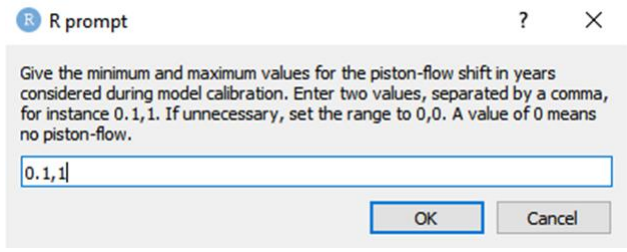
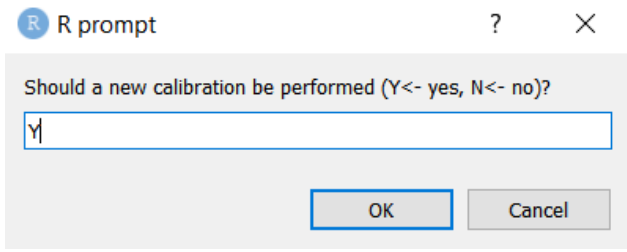
To run the script, click on "Source." Pop-up windows will appear, allowing you to customize your preferences for running the script. Once you've selected your options, click "OK" to proceed with the next steps (as shown in the graphic below).

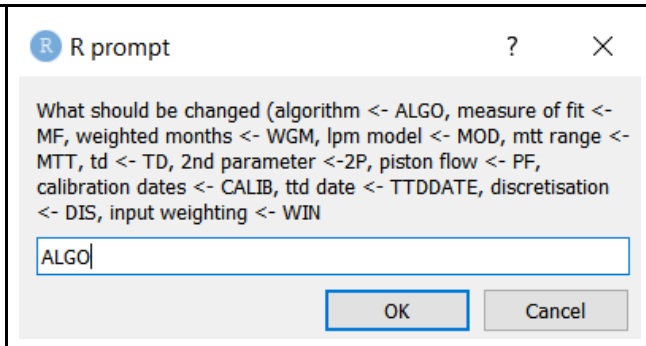


Please continue selecting your options in the following windows, adjust them according to your needs. Follow the prompts until the process is complete.

C.1		**Define the Model Type:** Choose between the Lumped Parameter Model (LPM) for Transit Time Distribution (TTD) curves and Mean Transit Time (MTT) estimates, or StorAgeSelection functions for storage volume and transit time age estimates.
C.2		**Select Tracers:** Two tracer types are available: Tritium (H3) and Stable Isotopes (ST). For this guide, Deuterium (H2) was selected. If both tracers are selected for calibration, the calibration is implemented in a stepwise manner. First, the parameters from the best 10% solutions based on tritium as calibration target are found and retained. Within this parameter subset, the best solution for stable isotopes is found.
C.3		**Please Select the stable isotope:** O^{18} or deuterium (H^2).
C.4		**Calibration Metrics:** Choose a measure of fit for calibration. We recommend Mean Predictor Error (MPE) for quicker calculations. Nash-Sutcliffe Efficiency (NSE) requires longer processing time.
C.5		**Input Data Frequency**: The preprocessing script can calculate mean values based on different time intervals: Yearly (AM) Monthly (MM) Weekly (WM). If "raw values" is selected, data gaps in the input will not be filled in by interpolation and will cause an error message.

C.6		**Time Step Duration**: Specify the duration of the input concentration time steps to match the model requirements. Downscaling (i.e., from daily to hourly) of input data is not implemented and will not work. If selected time step is larger than input time step (i.e., minute data to daily), the data is aggregated to the desired time step (the mean is calculated). If the user requires another aggregation time step, it has to be done outside the WITS toolbox.
C.7		**Isotope Time Series for Calibration:** Decide whether to use the entire isotope time series of the catchment output concentration for calibration (Y/N). If "No," is selected the user is then asked to define the specific period for calibration themselves.
C.8		**Discretization Steps:** Select the number of discretization steps for each fitting parameter. A minimum of 10 steps is recommended for adequate discretization, however, it is important to note that 10 steps is rather coarse. Higher discretization of the parameter space will increase computational speed.
C.9		**Model Conditions**: Choose between Steady-State (SS) or Transient (TRAN) conditions based on your catchment's behavior. Here, steady state (no storage and flow rate change) was selected for the Lainbach catchment. If transient is selected, a variable flow and storage is calculated.
C.10		**Lumped Parameter Model:** Choose the model you wish to apply: Exponential Model (EM) Gamma Model (GM) Dispersion Model (DM) Double Exponential Model (DEM). For variable flow, only EM and DM will be offered.

C.11		**Data Weighting:** Select whether to apply weighting to the input data: 1 = No weighting 2 = Weighting based on the shift between input and output isotope time series Or, define a custom weighting value between 0 and 1. Note, if summer rainfall does not contribute to the overall isotope signal in the output signal, weighting may be necessary. Hence, if weighting is set to a number below "1" for specific month (user defined in a follow up pop-up window), the isotopic signal is reduced by that weighting factor, compared to the non-weighted month.
C.12		**Transit Time Range:** Define the estimated range expected for transit time, using a minimum and maximum value separated by a comma. Example: (1, 10). Within this transit time range the best solution is searched for.
C.13		**Piston-Flow Consideration:** If your system should include piston-flow, specify the shift (in years) for calibration by entering a minimum and maximum value separated by a comma. Use (0, 0) if piston-flow is not present.)
C.14		**After the first execution:** Confirm if you wish to perform another run by selecting Y or N. Choose Y to rerun the process, or N to exit.

C.15		**After selecting to rerun the process: You can** configure the following alternatives: • Algorithm • Measure of Fit • Weighted Months • Lumped Parameter Model type • Mean Transit Time Range Among others.

Results:

The results will be presented in folders with the following structure (two tracers are used):

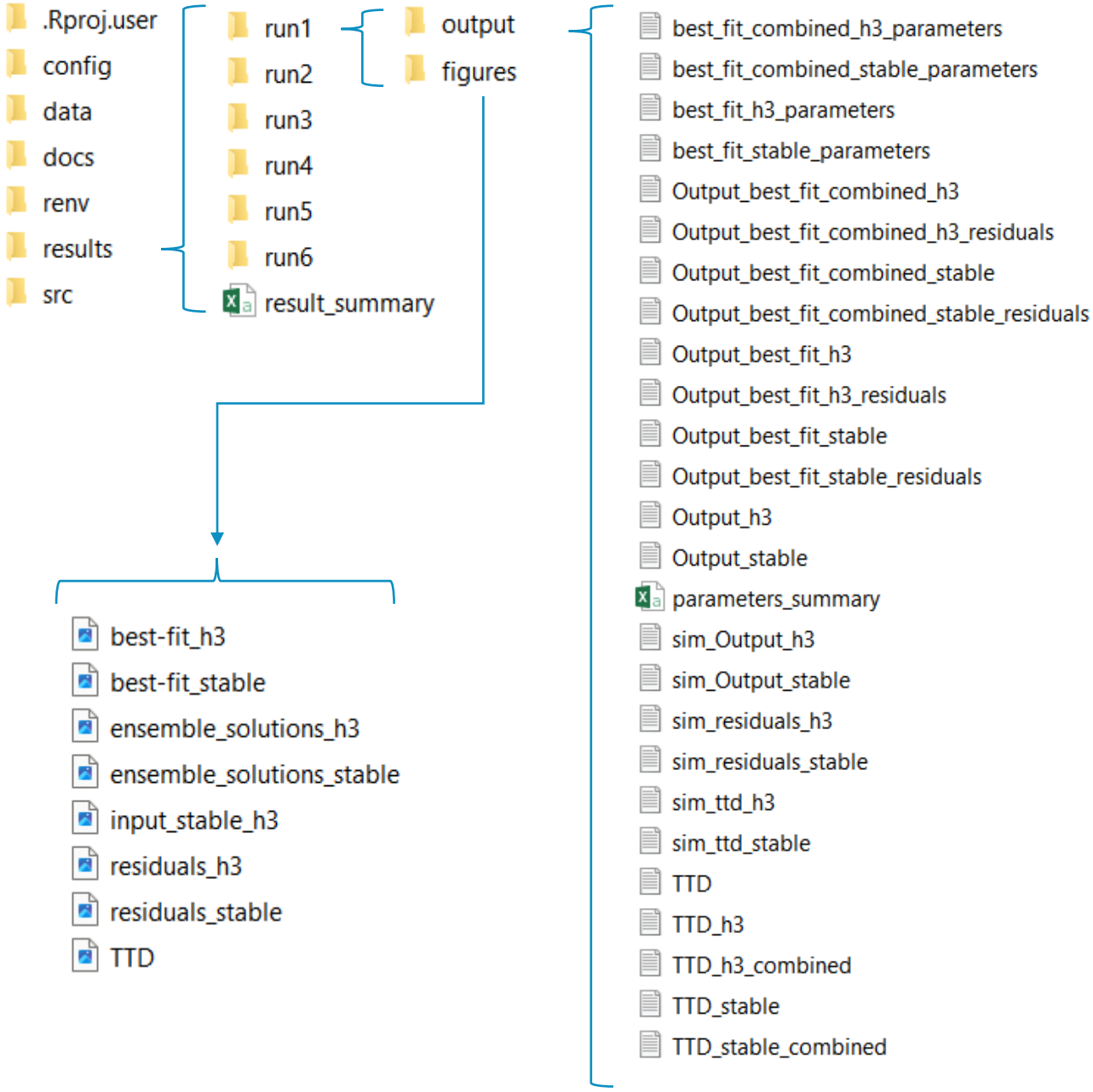


Figure 2. Structure of result folders and output files. Detailed description on each file type can be found in Table 2.

Results: Output Tables

If two tracers are used, a combined best fit table will be generated (Figure 2). This means that a combined table will summarize the best fit parameters for both tracers.

Each folder will contain multiple runs (e.g., Run 1 through Run 6 in the example). A total of up to 10 runs can be performed. In addition to the individual run folders, a summary text file will be generated, providing a summary of the results.

Inside each run folder, a series of text files are presented. These files include detailed outputs from the simulation, as well as a file with the summary of parameters (Table 1). The content of each specific files is described in Table 2.

Please note, that the "output" folder shows essentially **five categories** of files starting with **"best_fit_[…]", "Output best fit_[…]","output[…], "sim_[…]", "TTD_[…]"**.

For each category there are **three content groups**. The content in each groups is either the parameter table which is extended by **[…]_parameters**, or a timeseries of residuals extended by **[…]_residual**. The simulated best- fit isotope timeseries does **not** contain any **additional name** extension.

To those overall categories and content groups the name is finally extended depending on the chosen **isotope configuration** of the model. If only one isotope is chosen for calibration then names are either **"[…]_H3" (tritium)** or **"[…]_stable"** (deuterium or oxygen-18). If two isotopes are chosen then the name of the relative best fits output files is extended by that information such as **"[…]_combined".**

Table 1. Example of Parameter Summary for an Individual Run.

algorithm	fit	weight	Months	model	mtt	pf
BF	MPE	1	1	EM	2	0.1
BF	MPE	1	2	EM	5	1
BF	MPE	1	3	EM	2	0.1
BF	MPE	1	4	EM	5	1
BF	MPE	1	5	EM	2	0.1
BF	MPE	1	6	EM	5	1
BF	MPE	1	7	EM	2	0.1
BF	MPE	1	8	EM	5	1
BF	MPE	1	9	EM	2	0.1
BF	MPE	1	10	EM	5	1
BF	MPE	1	11	EM	2	0.1
BF	MPE	1	12	EM	5	1

Table 2. Description of the content of the text files contained in the folder "output"

Category	Content	Description
´best_fit´[…]	´best_fit_[...]_**parameters**´ ´best_fit_combined_[...]_**parameters**´	This section contains the **parameters** that yielded the best model fit. This can be either the best fit model output for each individual tracer (stable or tritium) or the combined output if two tracers are used.
´output_best_fit_[...]	´output_best_fit_stable ´ ´output_best_fit_H3 ´ ´output_best_fit_combined_stable_residuals´ ´output_best_fit_combined_H3_residuals´	This file contains the absolute best fit of the prediced **isotope concentration** per timestep for each individual tracer (stable or tritium), or if two tracers are used the relative best fit that minimizes the residuals of both tracers. If **residuals** are attached in the name, the file contains the residuals of the fit (i.e. observed versus predicted output for each observation date).
´output_[...]´	´output_H3 ´ ´output_stable ´	This file contains a list all parameter combinations and their respective error for all simulation runs.
´sim_[…]´	´sim_output_[...]´, ´sim_residuals_[...]´ ´sim_ttd_[...]´	Those files contains a **record of all simulation runs** and their respective predicted isotope value per timestep ([…]_output), **their residuals** per timestep ([…]_residuals) and the **g(t)** per timestep ([…]_ttd. Be aware this table is transposed so that the time is along rows (including header D[…] or X[…] in the first row) and columns show individual runs.
´TTD[...]´	´TTD´ ´ TTD _H3´ ´ TTD _stable´ ´ TTD _H3_combined´ ´ TTD _stable_combined´	This table gives the probability distribution **g(t)** for the **best fit** simulation. ´ TTD ´ is generated when only a single tracer is used. When two tracers are used, the 'TTD' is generated for each single isotope tracer. Note: Since in the case of the simultaneous fitting of two tracer the best 10 % solutions the all tritium simulations are retained before looking for the best stable isotope simulation amongst those, ´TTD_H3_combined´

		´TTD_stable_combined´ and will be the same as to ´TTD_H3´.
´parameter_summary´	-	Summarizes all input parameters used for this specific run.

Results: Output Figures

The output figures are stored in the *results folder (Figure 2). Each run is in a separate directory labelled "runxx". Each "run" directory contains two folders. The "figures" folder contains all the graphs showing the input and fitted output. The "output" folder contains the data tables generated.* The user finds scatter plots (XY plots) showing the best fit for stable isotopes and tritium individually, along with their residual errors presented as a time series and box plots displaying the distribution of the residuals. An example is shown in Figure 3.

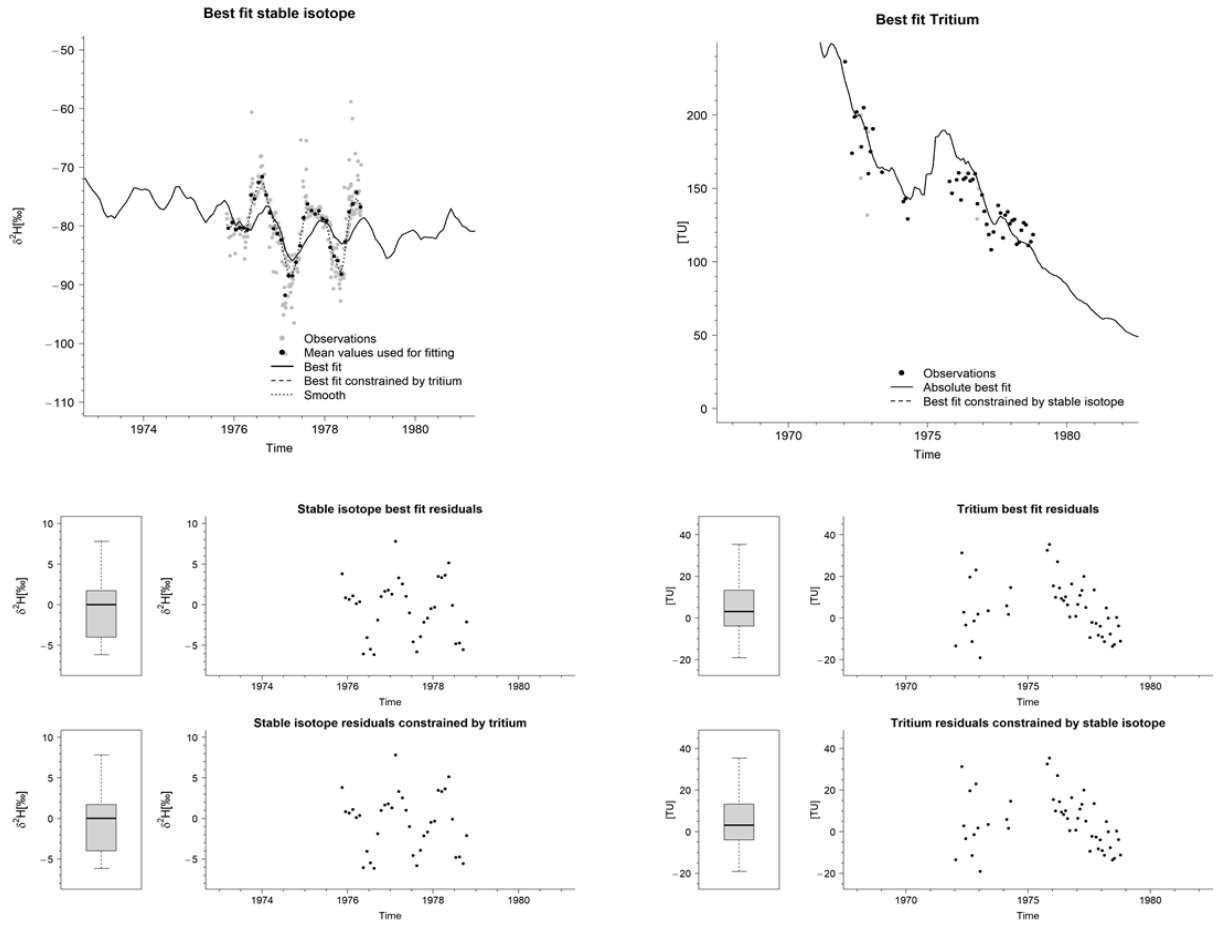


Figure 3. Various results figures produced automatically by the WITS toolbox. The header of each shows the content presented. Here the best-fit time series of the stable isotopes (upper left) and best fit tritium (upper right) is shown. If an analytical error is provided with the tritium data, it will be shown as error bars on the graph. The center row and bottom row presents residuals of the stable isotope tracer and tritium tracer, respectively. The bottom figure will be produced if both a stable and a radioactive tracer is used for calibration. The raw data of those figures is stored in the corresponding subfolder "output". Users are encouraged to develop their own plots for potential publication.

Figure 4 shows the transit time distribution, as well as the ensemble solutions for tritium and the stable isotope. For the latter, it includes observed samples and the

ensemble solutions, which represent the best 1% of all simulations, along with their respective best fit. This selection of ensemble solutions is defined based on performance criteria that identify the most accurate simulations in relation to the observed data.

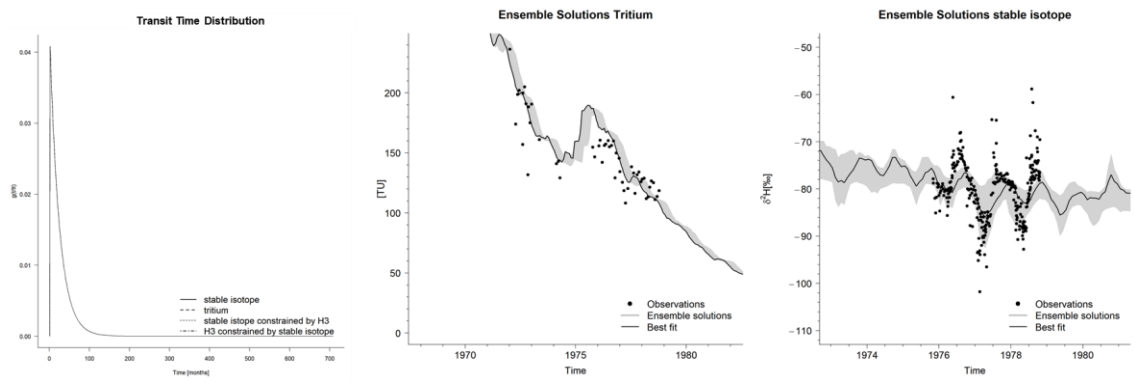


Figure 4. Other graphs stored in the "figure" folder. The left figure shows the transit time distribution of the best fit solution. The center and right plot shows an ensemble solution of tritium and stable isotope estimates. Ensemble solutions represent the best "1%" of all solutions.

Data requirements

For details on data requirements for each method please visit the **´Final report of working group 3 of the COST ACTION WATSON´.** Before using your own data for modelling please familiarize yourself with the model implementations in order to ensure a correct interpretation of the model results.

The user has to ensure a continuous data series of similar time steps. Data gaps in the time series are automatically filled in by WITS with simple linear interpolation between two data points. This may introduce unwanted errors. We encourage the User to ensure that potential gap filling errors are corrected before modelling is conducted.

The model does produce false results with mixed input data time steps (e.g.: one year of daily data, and 5 years of monthly data). Aggregation outside the WITS software to the desired time step is strongly advised.

- The user has to provide at least one isotope time series of catchment input and output to meet minimum requirement for the LPM models.
- SAS models require a minimum of four input variables for variable flow calculations: precipitation, evapotranspiration, discharge and a single isotope time series of the input and output.

When using your own data, make sure to set it up using the criteria described in the table 3. Please use the exact file name for each file as written in the table 3.

Decimal places are separated by "." (dots) while columns are separated by "," (comma) and the input file type is ".csv".

Table 3. Input files and format criteria. Decimal seperator is "." (dot) and column separator is "," (comma.

File type	File name	File format and columns names
LPM		
Stable isotope input timerseries	stable_input.csv	Date,isotope [YYYY-MM-DD],numeric value
Stable isotope output timerseries	stable_outpu.csv	Date,isotope [YYYY-MM-DD],numeric value
Tritium input timeseries	h3_input.csv	Date,H3 [YYYY-MM-DD],numeric value
Tritium output timeseries	h3_outputt.csv	Date,H3 [YYYY-MM-DD],numeric value
LPM (transient)		
Discharge	discharge.csv	Date,J [YYYY-MM-DD],numeric value
SAS		
Precipitation timeseries	rainfall.csv	Date,rain [YYYY-MM-DD],numeric value
Evapotranspiration timeseries	evapotranspiration.csv	Date,etp [YYYY-MM-DD],numeric value
Timeseries of SAS weights	weights.csv	Date,wi [YYYY-MM-DD],numeric value