SWAT+ Automatic Workflow (v 1.0.2) User Instructions

Introduction

SWAT is a time-continuous, semi-distributed hydrological model that has been widely used in various hydrological studies across the world. SWAT+ is a restructured and improved version of the SWAT model. QSWAT+ and SWAT+ Editor are currently the only interfaces for setting up SWAT+ models.

Reproducibility of science has come under scrutiny recently (Marcus, 2015) as it has been discovered that a large proportion of science is not reproducible (McNutt, 2014; Vasilevsky et al., 2013). In hydrology, modelling work is often not reproducible because of missing details on model setup. This software uses a namelist, which is a settings file, to build and run SWAT+ models. Users can share the namelist along with their published works to promote transparency as all model options are well documented in the file. Sharing the namelist and data also allows others to reproduce the models from scratch.

Setting up software

Download and install QGIS 3.10 – 64-bit version from QGIS website (Download QGIS 3.10.3). Download and install SWAT+ Automatic Workflow (SWAT+ AW) using default settings. The installer can be found on the repository (SWAT+ AW Releases). This installer will install all the dependencies and set environmental variables. It is recommended that you use the installer but you an also setup all the dependencies and environmental variables manually.

Pandas, geopandas and *playhouse* python packages do not ship with the QGIS python but are used by the workflow. If you do not want to use the installer from the repository, you will need to install *pandas, geopandas* and *playhouse* python packages for the Python that ships with QGIS 3.10 (C:\Program Files\QGIS 3.10\apps\Python37). You will also need to install TauDEM5 binaries to C:\SWAT\SWATPlus\TauDEM5Bin (View this page). The workflow should be extracted to the Workflow Directory (C:\SWAT\SWATPlus\Workflow). The Workflow Directory should be appended to the "path" environmental variable. It should also be added to the environmental variable "swatplus_wf_dir".

Data

The data requirements for the SWAT+ AW are the same as data requirements for QSWAT+ and SWAT+ Editor. Thus, users can use the same data formats as the GUI counterparts. For more information on SWAT+ data formats refer to QSWAT guide (Dile, Srinivasan, & George, 2017).

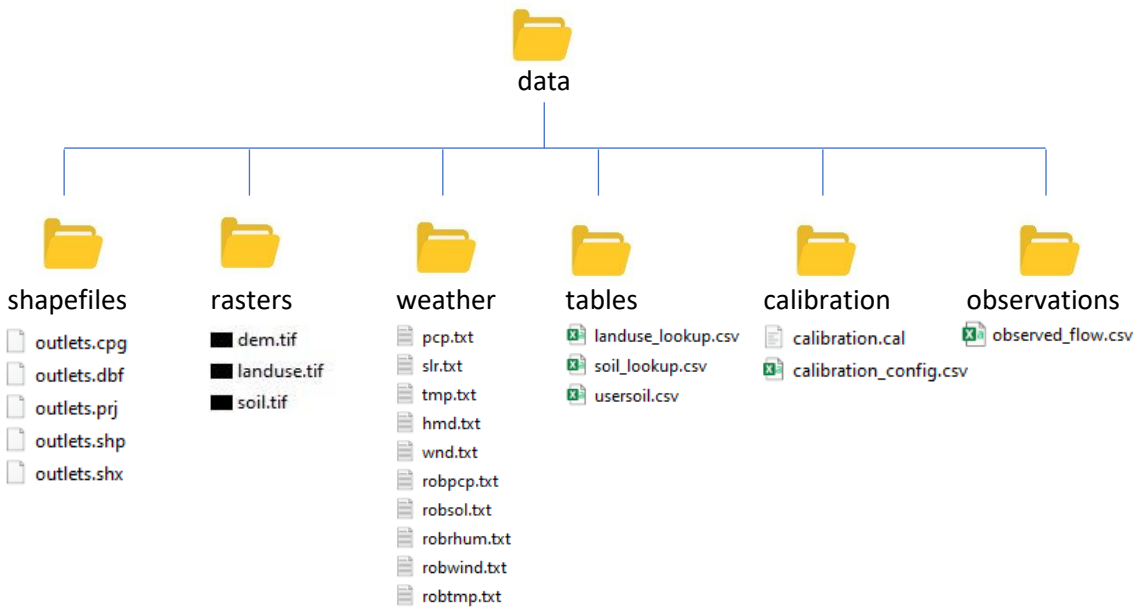

Figure 1: Directory structure of the SWAT+ WF

Rasters files. i.e. Digital elevation model (DEM), land use map and soil map should be placed in the *rasters* directory within the *Data* folder.

All shapefiles. i.e. outlet shapefile and river shapefile for 'burn in' should be placed in *shapes* directory

Files that have information that needs to be imported into project databases later should be placed in the *tables* directory. These files include land use and soil lookup tables, and soil properties table.

All weather files should be placed in the *weather* directory.

If you want to include calibrated parameters that should be applied to the model after set up, you should include the parameters file in the *calibration* directory. Remember to include the name of the calibration file in the namelist.

Setting up the namelist

The configuration information for setting up the model is in a settings file called 'namelist.py'. Users can open and make changes to this file in any text file editor such as Notepad, Notepad++ and Visual Studio Code (recommended).

Go download the example_dataset.zip file to view an example of a filed namelist file. The following is a description of each section and how to fill it.

Project identification

This section identifies the name and type of the project (Figure 2). The project name should be filled on Project_Name. Model_2_namelist option identifies what type of project it is. If the SWAT model was set up in the QSWAT+ and SWAT+ Editor GUIs and the user wants to retrieve a namelist and dataset that will reproduce that model, the Model_2_namelist option should be set to *True*. To go from namelist and data to a SWAT+ model, set this option to *False*.

File names

In this section, all files that are to be used in the model setup are listed. This includes the names of the soil, land use and DEM raster files present in rasters directory (Figure 1). Note that file names should include file extensions if rasters are not in *grid* format, as shown in Figure 2.

```
"""---------------- QSWAT+ Workflow v1.0 Settings File --------------"""

# Project Identification
Project_Name        = "robit"
Model_2_namelist    = False        # True = get settings from existing model
                                   # False = get model from current settings

"""--------------------------   File Names  -------------------------"""
# Raster files (Should be projected with the same projection)
Topography          = "srtm_30m"
Soils               = "mowr_soil90"
Land_Use            = "roblandusenew"


# Database Files
Soil_Lookup         = "soil_lookup.csv"
Landuse_Lookup      = "landuse_lookup.csv"
Usersoil            = "usersoil.csv"

# Shape Files
Outlets             = "outlets.shp" # it should have same format as in the exa

"""--------------------------  Project Options  ---------------------"""
```

Figure 2: Project identification and file names sections in namelist.py

Lookup files contain the soil type or land use classes represented by the values in the soil and land use raster files. The lookup files are contained in the tables' directory (Figure

1) and should be listed in the *Soil_Lookup* and *Landuse_Lookup* lines for soil lookup file and land use lookup file respectively.

Database table file for soil properties (usersoil file) should be specified in the on the *Usersoil* variable as demonstrated in Figure 2. You should also list the name of the file that contains locations where outlets should be created within the model in the *Outlet* line under Shape Files section.

For watershed delineation, you need to specify the channel and stream delineation thresholds using the Channel_Threshold and Stream_Threshold variables respectively. The maximum distance for snapping outlets to streams should also be entered in metres in the *OUT_Snap_Threshold* line under the *Watershed Delineation* section (Figure 3).

```
24  # Watershed Deliniation (1 = Cells)
25  Ws_Thresholds_Type  = 1
26  Channel_Threshold   = 200
27  Stream_Threshold    = 200
28  Out_Snap_Threshold  = 300              # metres
29  Burn_In_Shape       = ""               # leave as "" if none
30
31  #  -------------------  HRU Definition  -----------------
32  Slope_Classes       = "0, 9999"
33
34  # HRU creation method   (1 = Dominant landuse, soil, slope , 2 = Dominant HRU,
35  #                        3 = Filter by Area,                4 = Target Number
36  #                        5 = Filter by landuse, soil, slope)
37
38  HRU_Filter_Method     = 3
39
40  # Thresholds_Type           (1 = Total Area (ha) , 2 = Percent)
41  HRU_Thresholds_Type   = 2
42  Land_Soil_Slope_Thres = "12, 10, 11"    # Thresholds for Landuse, Soil and Slo
43  Target_Area           = 0               # used if HRU_Filter_Method 3 is selec
44  Target_Value          = 110             # used if HRU_Filter_Method 4 is selec
45
46  # Routing and ET and infiltration
47  ET_Method             = 2          # 1 = Priestley-Taylor, 2 = Penman-Monteit
48  |  |  |  |  |  |  |  |  |  |     # 3 = Hargreaves,      (4 = Observed - Not
49  Routing_Method        = 1          # 1 = Muskingum,        2 = Variable Stora
50  Routing_Timestep      = 1          # 1 = Daily Rainfall/routing, curve number
51  |  |  |  |  |  |  |  |  |  |     # 2 = Sub-daily Rainfall/routing, Green &
52
53  # model run settings
```

Figure 3: Project options section in namelist.py

The HRU Definition section offers options to be used for creating HRUs. Slope classes for model setup can be specified in the *Slope_Classes* variable. Separate the slope demarcation values by commas as shown in Figure 3.

You can specify one of different *HRU_Filter_Method* values (1 through 5) that are also available in the QSWAT+ GUI. For more information on how they work, refer to QSWAT+ guide #Ref!. The HRU creation methods are listed in Table 1:

Table 1: Available HRU creation options

Number	Method
1	Dominant land use, soil, slope
2	Dominant HRU
3	Filter by Area
4	Target Number of HRUs
5	Filter by land use, soil, slope

If HRU creation method 4 is selected, *Target_value* line should be set to a value. The QSWAT Workflow will create the number of HRUs specified in the *Target_value* line.

If method 5 is selected, the user should also specify the threshold for land use, soil and slope in *Land_Soil_Slope_Thres* as demonstrated in Figure 3.

If method 3 or 5 is selected, set *HRU_Thresholds_Type* to either 1 or 2 to interpret threshold values as areas (hectares) or per cent of landscape units respectively. Note that if method 3 is selected, *Target_Area* will be used as the area threshold.

Routing, ET calculation method and Infiltration options can be filled as shown in Figure 4 with reference to options presented in Table 2.

Table 2: Routing Options, ET calculation method and Infiltration options

Setting	Options
ET_method	1 = Priestley-Taylor 2 = Penman-Monteith 3 = Hargreaves
Routing_Method	1 = Muskingum 2 = Variable Storage
Routing_Timestep	1 = Daily Rainfall/routing, curve number 2 = Sub-daily Rainfall/routing, Green & Ampt

For model run options, specify the duration of the run period in the *Start_Year* and *End_Year* variables. Set the model warm up period in the *Warm_Up_Period* variable, as shown in Figure 4. You can get results in Comma Separated Values (CSV) format by setting the *Print_CSV* option to 1, else, set it to 0.

You have the option to run the model after setup using the *Release* or *Debug* executables by setting *Executable_Type* to 1 or 2 respectively. Set it to 0 to setup the model without running it.

You can set the model to give results at specific timesteps using the *Print_Objects* variable. Set the name of the output variables you want to specify and the list of the results timesteps needed as demonstrated in Figure 4. Include 1, 2, 3 and 4 to print daily, monthly, yearly and annual average results respectively. The variables available for selection are listed as follows:

```
# model run settings
Start_Year            = 1992
End_Year              = 1997
Warm_Up_Period        = 1           # the number of years for running the mode
                                    # 0 = no, 1 = yes, selection to output csv
Print_CSV             = 1
                                    # 1 = daily, 2 = month, 3 = year, 4 = annu
Print_Objects         = {           # default prints yearly results if not spe
                        "channel_sd"    : [1, 2, 3, 4],
                        "basin_nb"      : [2, 3, 4],
                        "basin_ls"      : [2, 3, 4],
                      }
Executable_Type       = 1           # 1 = Release, 2 = Debug, 0 = Do not run
Cal_File              = ""          # a calibration.cal file with parameters f
                                    # leave as "" if there is no file to be us
Calibrate             = False       # set to "True" to perform calibration, "F
Calibration_Config_File = "calibration_config.csv"            #
Number_of_Runs        = 4       # Set the number of runs for calibration
Number_of_Processes   = 7       # Set the number of parallel processes to ma
Make_Figures          = False       # set to "True" to create maps, "False" to
# Log progress or not? If yes, you will not see updates
log                   = False       # True or False
"""--------------------------  Settings End  ------------------------"""
```

Figure 4: Model run and calibration settings section

basin_wb, basin_nb, basin_ls, basin_pw, basin_aqu, basin_res, basin_cha, basin_sd_cha, basin_psc, region_wb, region_nb, region_ls, region_pw, region_aqu, region_res, region_cha, region_sd_cha, region_psc, lsunit_wb, lsunit_nb, lsunit_ls, lsunit_pw, hru_wb, hru_nb, hru_ls, hru_pw, hru-lte_wb, hru-lte_nb, hru-lte_ls, hru-lte_pw, channel, channel_sd, aquifer, reservoir, recall, hyd, ru, pest. For more information on this variables, see the SWAT+ Input Output Documentation (Click here to view).

If you want to apply parameters to the model after setup, specify the name of the parameter file in the *Cal_File* variable; leave “” if you do not want to apply parameters. Remember to have the parameter file in the *calibration* folder under *data*.

You can perform calibration on the model if you already know the outlet number that will need to be calibrated. Information on the calibration such as observation file name, parameter names and ranges are all set in a separate CSV file placed in the *calibration* directory. The name of this configuration file should be placed on the Calibration_Config_File. Figure 5 shows an example of how the Calibration_Config_File looks. It is important to know that the calibration is solely done by trying out parameter sets generated by lating hypercube sampling and chosing the parameter set that gives the highest Nash-Sutcliffe Efficiency Value. The Number_of_Processes variables allows you to perform the calibration using parallel processing which speeds up the calibration by trying multiple parameter sets at a time.

	A	B	C	D	E	F	G	H
1	Parameter	Minimum	Maximum	Change Type				
2	1 = pctchg, 2 = abschg, 3 = absval							
3	File Name:		observed_flows.csv					
4	Channel Number:		29					
5	Timestep:		2		1 = day, 2 = month, 3 = year			
6	Calibration Variable:		1		1 = flow, 2 = evapotranspiration			
7								
8	Parameter	Min	Max	Change Type				
9	cn2	-25	25	pctchg				
10	esco	-0.6	0.6	abschg				
11	perco	-20	20	pctchg				
12	awc	-20	20	pctchg				
13	surlag	-20	20	pctchg				
14	alpha	-20	20	pctchg				
15								

Figure 5: Example of Calibration_Config_File

Running the workflow

Setting up the model

Once the namelist has been filled, it should be placed in the same directory as the *data* directory (Figure 6)

User can run the workflow using Command Prompt or PowerShell by following these steps

1. Open the directory where the data directory and namelist are located.

2. Click the address bar and type ‘cmd’ or ‘powershell’

3. Press the ENTER key to open the Command Prompt window in the current directory (Figure 6). Note that you can open Command Prompt outside the current directory from the start menu, but the directory needs to be changed to the location of the namelist using the CD command.

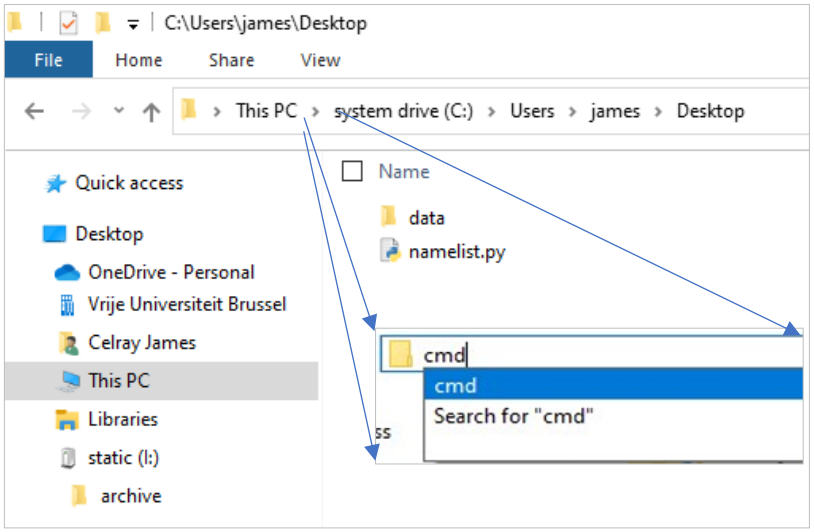

Figure 6: Directory where the model will be setup and how to open cmd or powershell

4. Type ‘swatplus_aw’ and press Enter to start running the SWAT+ AW. The workflow will set up the model based on the settings specified in the namelist (Figure 7). The model being set up will appear in the same directory as the namelist. If you change a setting in the namelist, you must change the name of the project too; otherwise, the current project will be overwritten.

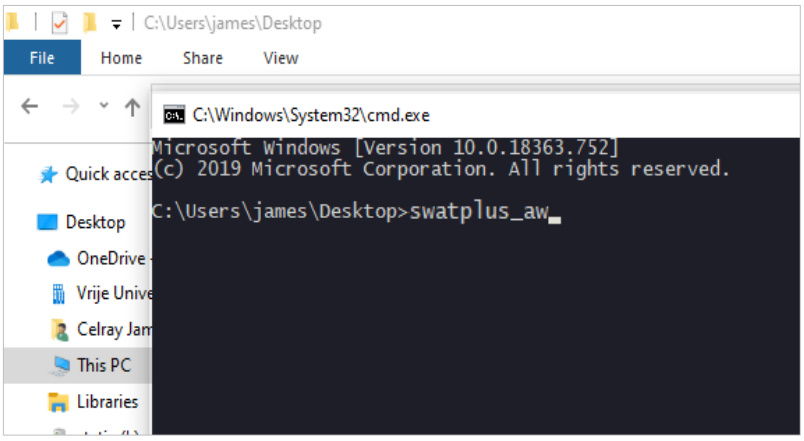

Figure 7: Running the SWAT+ AW

Once the model setup finishes, the user can open the model in QSWAT+ and SWAT+ Editor GUI for visualisation. The model project files that are created by the workflow are fully compatible with the GUI, and further changes can be made in the GUI.

Retrieving namelist and data from existing model

The procedure for retrieving the namelist and data is the same as for setting up the model. The only difference is that you should not have the namelist in the directory containing the model. However, if you have the namelist, make sure you set the *Model_2_namelist* setting to True. In that case, the namelist used to retrieve the new namelist from the existing model is saved into a directory called old_namelists. Thus the user can use previously used namelists later.

Be sure to move or rename the existing *data* directory as it will be overwritten in the process.

References

Dile, Y., Srinivasan, R., & George, C. (2017). QGIS Interface for SWAT (QSWAT), *February*(v 1.4), 77. https://doi.org/10.13140/RG.2.1.1060.7201

Marcus, E. (2015). Credibility and reproducibility. *Chemistry and Biology*, *22*(1), 3–4. https://doi.org/10.1016/j.chembiol.2014.12.008

McNutt, M. (2014). Reproducibility. *Science*, *343*(6168), 229–229. https://doi.org/10.1126/science.1250475

Vasilevsky, N. A., Brush, M. H., Paddock, H., Ponting, L., Tripathy, S. J., LaRocca, G. M., & Haendel, M. A. (2013). On the reproducibility of science: unique identification of research resources in the biomedical literature. *PeerJ*, *1*, e148. https://doi.org/10.7717/peerj.148