User's Guide for Multi-Modular Penn State Integrated Hydrologic Model

User's Guide for MM-PIHM Version 1

Yuning Shi
yshi@psu.edu

September 2, 2020

Contents

Chapter 1

Overview

The Multi-Modular Penn State Integrated Hydrologic Model (MM-PIHM) is a physically based watershed model with multiple optional modules. The current release of MM-PIHM contains the source code for PIHM (Qu and Duffy 2007), PIHM with a surface heat flux module (Flux-PIHM; Shi et al. 2013), PIHM with surface heat flux and biogeochemistry modules (Flux-PIHM-BGC; Shi et al. 2018), and PIHM with surface heat flux and reactive transport modules (bioRT-Flux-PIHM; Bao et al. 2017; Zhi et al. 2020). There is also an optional deep groundwater module (DGW) that works with PIHM, Flux-PIHM, and bioRT-Flux-PIHM (Brantley et al. 2018).

PIHM is a spatially-distributed, physically based hydrologic model. Flux-PIHM adds a land surface model (adapted from the Noah land surface model; Chen and Dudhia 2001) to PIHM for the simulation of land surface processes. Flux-PIHM-BGC couples Flux-PIHM with a terrestrial ecosystem model (adapted from Biome-BGC; Thornton et al. 2002) that enables the simulation of carbon and nitrogen cycles. bioRT-Flux-PIHM couples Flux-PIHM with a reactive transport module that enables the simulation of transport and chemical processes.

MM-PIHM is open source software licensed under the MIT License. The development of MM-PIHM is coordinated via the MM-PIHM GitHub page (`https://github.com/PSUmodeling/MM-PIHM`). All bug reports and feature requests should be submitted using the GitHub "Issues" page.

Chapter 2

Installation

The source code provided in the MM-PIHM package can be used to compile executables for PIHM, Flux-PIHM, bioRT-Flux-PIHM, and Flux-PIHM-BGC. The instructions below show how to compile MM-PIHM executable on Linux or macOS systems. To install MM-PIHM on Windows, please refer to this guide: `https://gist.github.com/shiyuning/867d5af0a3a6345b50ec1b193a71e4be`.

2.1 Installing CVODE

MM-PIHM requires the SUNDIALS CVODE v2.9.0 implicit solver (Hindmarsh et al. 2005). The CVODE Version 2.9.0 source code is provided with the MM-PIHM package for users' convenience. SUNDIALS (©2012–2016) is copyrighted software produced at the Lawrence Livermore National Laboratory. A SUNDIALS copyright note can be found in the `cvode` directory.

If you already have CVODE v2.9.0 installed, you can edit the `Makefile` and point `CVODE_PATH` to your CVODE directory. Otherwise, you need to install CVODE before compiling MM-PIHM, by doing

```
$ make cvode
```

in your MM-PIHM directory.

Currently CMake (version 2.8.1 or higher) is the only supported method of CVODE installation. If CMake is not available on your system, CMake Version 3.7.2 binary for Linux (or macOS, depending on your OS) will be downloaded from `http://www.cmake.org` automatically when you choose to `make cvode`.

2.2 Installing MM-PIHM

Once CVODE is installed, you can compile MM-PIHM models from the MM-PIHM directory by doing

```
$ make <model>
```

The <model> can be replaced by the name of model that you want to compile, which could be pihm, flux-pihm, flux-pihm-bgc, or rt-flux-pihm. The users can also run

```
$ make test
```

to compile Flux-PIHM, and run the embedded ShaleHills simulation using Flux-PIHM, to get familiar with MM-PIHM interface.

You can compile PIHM, Flux-PIHM and bioRT-Flux-PIHM with the optional deep groundwater module (DGW), using

```
$ make DGW=on <model>
```

The command

```
$ make clean
```

will remove the executables and object files. Note that if you want to switch from one MM-PIHM model to another, you must make clean first. To get help, run make to show a help message.

2.3 Installation options

By default, MM-PIHM is parallelized using OpenMP, which significantly improves the computational efficiency of MM-PIHM models. CVODE, however, is not implemented using OpenMP by default. According to CVODE documentation, CVODE state variables (i.e., unknowns) "should be of length at least 100,000 before the overhead associated with creating and using the threads is made up by the parallelism in the vector calculations." In other words, you should only use OpenMP for CVODE if your model domain has $> 30,000$ model grids. If you do want to test using OpenMP for CVODE, you can compile MM-PIHM models using

```
$ make CVODE_OMP=on <model>
```

Note that in order to use OpenMP for CVODE, you also need to turn on the `OPENMP_ENABLE` option when using CMake to install CVODE.

You can also turn off OpenMP for MM-PIHM (NOT RECOMMENDED):

```
$ make OMP=off <model>
```

By default, MM-PIHM compilation is using the `-O0` gcc option. If you wish to further accelerate MM-PIHM models, you may want to use

```
$ make DEBUG=off <model>
```

which will compile using `-O2` gcc option.

Chapter 3

Running MM-PIHM

3.1 Setting up OpenMP environment

To optimize PIHM efficiency, you need to set the number of threads in OpenMP. For example, in command line

```
$ export OMP_NUM_THREADS=12
```

will enable MM-PIHM model simulations using twelve (12) OpenMP threads.

If you use a PBS script, you must require the right number of ppn (processor cores per node) before setting the number of threads. The ppn should be the same as the number of threads you want to use. For example, your PBS script may look like

```
#PBS -l nodes=1:ppn=12
#PBS -l walltime=1:00:00
#PBS -j oe
#PBS -l pmem=1gb

cd $PBS_O_WORKDIR

export OMP_NUM_THREADS=12
./pihm ShaleHills
```

3.2 Running MM-PIHM models

Now you can run MM-PIHM models using

```
$ ./<model> [-b] [-c] [-d] [-f] [-o dir_name] [-s] [-V] [-v] <project>
```

where `<model>` is the installed executable, `<project>` is the name of the simulation, and `[-bcdfosVv]` are optional parameters.

Parameter `-b` will turn on brief mode, with minimal screen output.

Parameter `-c` will turn on elevation correction mode. Surface elevations of all model grids will be checked to find surface sinks. If surface sinks exist, the elevations of sink grids will be elevated to be the same as its lowest neighbor.

Parameter `-d` will turn on debug mode. In debug mode, helpful information is displayed on screen and a CVODE log file will be produced.

Parameter `-f` will turn on fixed length spin-up. By default, spin-up stops when the model reaches equilibrium or the specified maximum number of spin-up years. When fixed length spin-up is turned on, spin-up only stops when reaching the specified maximum spin-up years.

Parameter `-o` will specify the name of directory to store model output. All model output variables will be stored in the `output/dir_name` directory when `-o` option is used. Otherwise, model output will be stored in a directory named after the project and the system time when the simulation is executed (Section 5).

Parameter `-s` will turn on silent mode, with no screen output.

Parameter `-V` will display model version. Note that model will quit after displaying the version information. No simulation will be performed when using the `-V` parameter.

Parameter `-v` will turn on the verbose mode, with verbose screen output for users to understand the simulation processes.

When MM-PIHM model runs, the model will look into the `input/<project>` directory for the input files (Section 4).

Chapter 4

Input files

MM-PIHM input files are organized by projects. The input files for each project should be put into the same directory, under `<MM-PIHM_DIR>/input/<project>`, where `<MM-PIHM_DIR>` is the directory in which MM-PIHM is installed, and `<project>` is the name of the project. The `vegprmt.tbl` input file is provided with the source code, and is stored in the `input` directory. The `.epc` input files (required by Flux-PIHM-BGC; see Section 4.3.2) are also shared by different projects, and are stored in the `input/epc` directory. Different MM-PIHM models can share the same input files, with a few exceptions. For example, input files for a Flux-PIHM simulation, except for the initial condition file (`.ic` file; see Section 4.1.10), should also support a PIHM simulation.

4.1 Hydrologic module input files

4.1.1 Mesh structure file

The mesh structure (`.mesh`) file is a required input file for all MM-PIHM models. It describes the irregular mesh (triangular irregular network, or TIN) structure of the model domain. The mesh structure file consists of the element block and the node block. The element block describes the total number of elements, and the three vertex nodes and three neighbor elements of each triangular element. The node block describes the location and elevation of each node. A mesh structure file looks like this:

```
 NUMELE  535
```

```
INDEX   NODE1   NODE2   NODE3   NABR1   NABR2   NABR3
1       139     6       179     239     112     238
2       50      49      46      45      4       23
3       16      249     248     442     439     440
...
533     295     298     268     204     475     526
534     218     299     217     499     383     225
535     281     299     291     225     515     518
NUMNODE  299
INDEX   X              Y              ZMIN      ZMAX
1       254488.5000    4505385.500    269.576   271.617
2       254469.5000    4505386.500    268.968   270.577
3       254381.5000    4505355.500    264.050   266.320
...
297     254583.3477    4505486.753    292.435   294.552
298     254535.7270    4505398.073    278.241   279.759
299     254598.6925    4505415.428    286.376   287.880
```

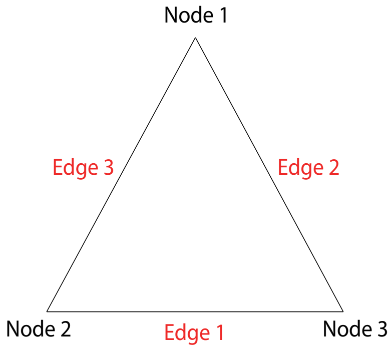

Figure 4.1: Demonstration of grid node and edge numbering.

A description of the input variables can be found below:

Element block

`NUMELE` (type: integer, unit: -) Total number of elements.

`INDEX` (type: integer, unit: -) Element index.

`NODE`i (type: integer, unit: -) The ith node of element.

`NABR`i (type: integer, unit: -) The neighbor element on the ith edge. The ith edge is the opposite to the ith node (Figure 4.1).

Node block

NUMNODE (type: integer, unit: -) Total number of nodes.

INDEX (type: integer, unit: -) Node index.

X (type: double, unit: m) x coordinate of node.

Y (type: double, unit: m) y coordinate of node.

ZMIN (type: double, unit: m) Bedrock (bottom of soil) elevation of node.

ZMAX (type: double, unit: m) Surface elevation of node.

4.1.2 Element attribute file

The element attribute (`.att`) file is a lookup table which stores the attribute of each mesh element, including soil type, land cover type, meteorological forcing time series, boundary condition type, etc. It allows efficient data storage. It is required by all MM-PIHM models. An element attribute file looks like this:

```
INDEX  SOIL  GEOL  LC  METEO  LAI  BC1  BC2  BC3
1      5     0     5   1      1    0    0    0
2      1     0     5   1      1    0    0    0
3      1     0     5   1      1    0    0    0
...
533    3     0     4   1      1    0    0    0
534    1     0     4   1      1    0    0    0
535    1     0     4   1      1    0    0    0
```

A description of the input variables can be found below.

INDEX (type: integer, unit: -) Element index.

SOIL (type: integer, unit: -) Soil (shallow zone) class.

GEOL (type: integer, unit: -) Geology (deep zone) class.

LC (type: integer, unit: -) Land cover class.

METEO (type: integer, unit: -) Index of meteorological forcing time series in the `.meteo` file (Section 4.1.6).

LAI (type: integer, unit: -) Index of leaf area index forcing time series in the .lai file (Section 4.1.11). If 0, NLCD climatological LAI will be used.

BC*i* (type: integer, unit: -) Boundary condition forcing time series on Edge *i*. Default values for boundary condition is 0, which indicates no flow condition and does not link to any time series.

4.1.3 Soil parameter file

The soil parameter (.soil) file contains the hydraulic and thermal properties of all soil classes, for both soil matrix and macropore, in the model domain. It is required by all MM-PIHM models. A soil parameter file looks like this:

```
NUMSOIL  5
INDEX  SILT   CLAY   OM    BD    KINF      KSATV     KSATH     MAXSMC  MINSMC  ALPHA  BETA   MACHF  MACVF  DMAC  QTZ
1      41.67  14.44  4.02  1.15  9.098E-5  1.625E-5  1.202E-5  0.3698  0.0370  8.80   1.240  0.01   0.01   1.0   0.25
2      33.90  15.85  3.44  1.53  1.516E-4  1.862E-5  9.802E-6  0.4032  0.0403  6.45   1.212  0.01   0.01   1.0   0.25
3      33.90  15.85  3.44  1.49  9.833E-5  1.086E-5  2.263E-5  0.4247  0.0425  6.50   1.258  0.01   0.01   1.0   0.25
4      39.70  19.86  3.66  1.64  1.506E-5  6.762E-6  3.047E-5  0.4179  0.0418  5.34   1.260  0.01   0.01   1.0   0.25
5      39.70  19.86  3.66  1.48  8.281E-5  3.705E-5  6.962E-5  0.4928  0.0493  5.82   1.220  0.01   0.01   1.0   0.25
DINF      0.10
KMACV_RO  100.0
KMACH_RO  1000.0
```

A description of the input variables can be found below.

`NUMSOIL` (type: integer, unit: -) Number of soil classes described.

`INDEX` (type: integer, unit: -) Soil class index.

`SILT` (type: double, unit: %) Silt percentage.

`CLAY` (type: double, unit: %) Clay percentage

`OM` (type: double, unit: %) Organic matter percentage.

`BD` (type: double, unit: g cm^{-3}) Bulk density.

`KINF` (type: double, unit: m s^{-1}) Vertical saturated infiltration hydraulic conductivity.

`KSATV` (type: double, unit: m s^{-1}) Vertical saturated hydraulic conductivity.

`KSATH` (type: double, unit: m s^{-1}) Horizontal saturated hydraulic conductivity.

`MAXSMC` (type: double, unit: m^3 m^{-3}) Maximum soil moisture content.

`MINSMC` (type: double, unit: m^3 m^{-3}) Residual soil moisture content.

`ALPHA` (type: double, unit: m^{-1}) Van Genuchten soil parameter (related to the inverse of the air entry suction).

`BETA` (type: double, unit: dimensionless) Van Genuchten soil parameter (measure of the pore-size distribution).

`MACHF` (type: double, unit: m^2 m^{-2}) Horizontal area fraction of macropore.

`MACVF` (type: double, unit: m^2 m^{-2}) Vertical area fraction of macropore.

`DMAC` (type: double, unit: m) Macropore depth.

`QTZ` (type: double, unit: 100%) Quartz content.

`DINF` (type: double, unit: m) A virtual top soil layer thickness across which infiltration is calculated.

`KMACV_RO` (type: double, unit: dimensionless) Ratio between vertical macropore hydraulic conductivity and vertical saturated infiltration hydraulic conductivity.

KMACH_RO (type: double, unit: dimensionless) Ratio between horizontal macropore hydraulic conductivity and horizontal saturated hydraulic conductivity.

Values of KINF, KSATV, KSATH, MAXSMC, MINSMC, ALPHA, BETA, and QTZ can be -999, in which case, pedotransfer functions will be used to calculate these properties. Values of SILT, CLAY, and OM can be -999 if all the parameters above are provided. Parameters DINF, KMACV_RO, and KMACH_RO apply to all soil types.

4.1.4 Land cover parameter file

The vegetation parameter file (vegprmt.tbl) contains the vegetation properties corresponding to different land cover classes present in the model domain. The vegetation parameters are the same as the NLCD40 category in the VEGPARM.TBL used by the Weather Research and Forecasting Model (WRF). This file is shared by all projects and the file should not be renamed. It is required by all MM-PIHM models. The default land cover parameter file looks like this:

```
NUMLC    40
INDEX    SHDFAC    DROOT     RS        RGL       HS        SNUP      LAIMIN    LAIMAX    EMISMIN   EMISMAX   ALBMIN    ALBMAX    ZOMIN     ZOMAX     ROUGH
1        0.70      1.00      125.0     30.0      47.35     0.080     5.00      6.40      0.95      0.95      0.12      0.12      0.50      0.50      0.07      # Evergreen Needleleaf Forest
2        0.95      1.00      150.0     30.0      41.69     0.080     3.08      6.48      0.95      0.95      0.12      0.12      0.50      0.50      0.07      # Evergreen Broadleaf Forest
3        0.70      1.00      150.0     30.0      47.35     0.080     1.00      5.16      0.93      0.94      0.14      0.15      0.50      0.50      0.07      # Deciduous Needleleaf Forest
4        0.80      1.00      100.0     30.0      54.53     0.080     1.85      3.31      0.93      0.93      0.16      0.17      0.50      0.50      0.07      # Deciduous Broadleaf Forest
5        0.80      1.00      125.0     30.0      51.93     0.080     2.80      5.50      0.93      0.97      0.17      0.25      0.20      0.50      0.06      # Mixed Forest
6        0.70      0.70      300.0     100.0     42.00     0.030     0.50      3.66      0.93      0.93      0.25      0.30      0.01      0.05      0.045     # Closed Shrubland
7        0.70      0.70      170.0     100.0     39.18     0.035     0.60      2.60      0.93      0.95      0.22      0.30      0.01      0.06      0.045     # Open Shrubland
8        0.50      0.70      70.0      65.0      54.53     0.040     0.50      3.66      0.93      0.93      0.25      0.30      0.01      0.05      0.04      # Woody Savanna
9        0.50      0.70      70.0      65.0      54.53     0.040     0.50      3.66      0.92      0.92      0.20      0.20      0.15      0.15      0.04      # Savanna
10       0.80      0.70      40.0      100.0     36.35     0.040     0.52      2.90      0.92      0.96      0.19      0.23      0.10      0.12      0.04      # Grassland
11       0.60      0.25      100.0     30.0      51.93     0.020     1.75      5.72      0.95      0.95      0.14      0.14      0.30      0.30      0.035     # Permanent Wetland
12       0.80      0.70      40.0      100.0     36.25     0.040     1.50      5.68      0.92      0.99      0.15      0.23      0.05      0.15      0.04      # Cropland
13       0.10      0.70      200.0     100.0     36.25     0.040     1.00      1.00      0.88      0.88      0.15      0.15      0.50      0.50      0.01      # Urban and Built-Up
14       0.80      0.70      40.0      100.0     36.25     0.040     2.29      4.29      0.92      0.98      0.18      0.23      0.05      0.14      0.04      # Cropland/Natural Veg.  Mosaic
15       0.00      0.05      999       999       999       0.020     0.01      0.01      0.95      0.95      0.55      0.70      0.001     0.001     0.035     # Permanent Snow
16       0.01      0.05      999       999       999       0.020     0.10      0.75      0.90      0.90      0.38      0.38      0.01      0.01      0.035     # Barren/Sparsely Vegetated
17       0.00      0.00      100.0     30.0      51.75     0.010     0.01      0.01      0.98      0.98      0.08      0.08      0.0001    0.0001    0.035     # IGBP Water
18       0.00      0.00      999       999       999       999       999       999       999       999       999       999       999       999       999       # Unclassified
19       0.00      0.00      999       999       999       999       999       999       999       999       999       999       999       999       999       # Fill Value
20       0.00      0.00      999       999       999       999       999       999       999       999       999       999       999       999       999       # Unclassified
21       0.00      0.00      100.0     30.0      51.75     0.010     0.01      0.01      0.98      0.98      0.08      0.08      0.0001    0.0001    0.035     # Open Water
22       0.00      0.05      999       999       999       0.020     0.01      0.01      0.95      0.95      0.55      0.70      0.001     0.001     0.035     # Perennial Ice/Snow
23       0.30      0.70      200.0     100.0     36.25     0.040     1.00      1.00      0.88      0.88      0.20      0.20      0.50      0.50      0.04      # Developed Open Space
24       0.27      0.70      200.0     100.0     36.25     0.040     1.00      1.00      0.88      0.88      0.20      0.20      0.70      0.70      0.03      # Developed Low Intensity
25       0.02      0.70      200.0     100.0     36.25     0.040     1.00      1.00      0.88      0.88      0.20      0.20      1.50      1.50      0.02      # Developed Medium Intensity
26       0.11      0.70      200.0     100.0     36.25     0.040     1.00      1.00      0.88      0.88      0.20      0.20      2.00      2.00      0.01      # Developed High Intensity
27       0.01      0.05      999       999       999       0.020     0.10      0.75      0.90      0.90      0.38      0.38      0.01      0.01      0.035     # Barren Land
28       0.80      1.00      125.0     30.0      54.70     0.080     1.00      5.16      0.93      0.94      0.14      0.17      0.50      0.50      0.07      # Deciduous Forest
29       0.95      1.00      140.0     30.0      44.00     0.080     3.08      6.48      0.95      0.95      0.12      0.12      0.50      0.50      0.07      # Evergreen Forest
30       0.80      1.00      125.0     30.0      51.93     0.080     2.80      5.50      0.93      0.97      0.17      0.25      0.20      0.50      0.06      # Mixed Forest
31       0.70      0.70      170.0     100.0     39.18     0.035     1.00      4.00      0.93      0.95      0.16      0.30      0.01      0.04      0.04      # Dwarf Scrub
32       0.70      0.70      300.0     100.0     42.00     0.030     0.50      3.66      0.93      0.93      0.22      0.30      0.01      0.05      0.04      # Shrub/Scrub
33       0.80      0.70      40.0      100.0     36.35     0.040     0.52      2.90      0.92      0.96      0.19      0.23      0.10      0.12      0.04      # Grassland/Herbaceous
34       0.60      0.25      40.0      100.0     60.00     0.010     1.50      5.65      0.95      0.95      0.14      0.14      0.20      0.20      0.04      # Sedge/Herbaceous
35       0.60      0.25      40.0      100.0     60.00     0.010     1.00      2.00      0.95      0.95      0.31      0.31      0.01      0.01      0.04      # Lichens
36       0.60      0.25      40.0      100.0     60.00     0.010     1.00      2.00      0.95      0.95      0.24      0.24      0.01      0.01      0.04      # Moss
37       0.80      0.70      40.0      100.0     36.25     0.040     1.56      5.68      0.92      0.99      0.17      0.23      0.05      0.15      0.02      # Pasture/Hay
38       0.80      0.70      40.0      100.0     36.25     0.040     1.56      5.68      0.93      0.99      0.20      0.25      0.02      0.10      0.04      # Cultivated Crops
39       0.60      0.25      100.0     30.0      51.93     0.020     0.70      3.50      0.95      0.95      0.14      0.14      0.40      0.40      0.04      # Woody Wetland
40       0.60      0.25      40.0      100.0     60.00     0.010     0.70      3.50      0.95      0.95      0.12      0.12      0.20      0.20      0.04      # Emergent Herbaceous Wetland
TOPT      298.0
CFACTR    0.5
RSMAX     5000.0
```

A description of the input variables can be found below.

NUMLC (type: integer, unit: -) Number of land cover classes described.

INDEX (type: integer, unit: -) Land cover class.

SHDFAC (type: double, unit: $\mathrm{m}^2\ \mathrm{m}^{-2}$) Green vegetation fraction.

DROOT (type: double, unit: m) Rooting depth.

RS (type: double, unit: $\mathrm{s\ m}^{-1}$) Minimum stomatal resistance.

RGL (type: double, unit: $\mathrm{W\ m}^{-2}$) Parameter used in radiation stress function.

HS (type: double, unit: dimensionless) Parameter used in vapor pressure deficit function.

SNUP (type: double, unit: m) Threshold water-equivalent snow depth that implies 100% snow cover.

LAIMIN (type: double, unit: $\mathrm{m}^2\ \mathrm{m}^{-2}$) Minimum leaf area index through the year.

LAIMAX (type: double, unit: $\mathrm{m}^2\ \mathrm{m}^{-2}$) Maximum leaf area index through the year.

EMISMIN (type: double, unit: dimensionless) Minimum background emissivity through the year.

EMISMAX (type: double, unit: dimensionless) Maximum background emissivity through the year.

ALBMIN (type: double, unit: dimensionless) Minimum background albedo through the year.

ALBMAX (type: double, unit: dimensionless) Maximum background albedo through the year.

ZOMIN (type: double, unit: m) Minimum background roughness length through the year.

ZOMAX (type: double, unit: m) Maximum background roughness length through the year.

ROUGH (type: double, unit: $\mathrm{s\ m}^{-1/3}$) Manning's roughness coefficient.

TOPT (type: double, unit: K) Optimum transpiration air temperature.

CFACTR (type: double, unit: dimensionless) Parameter used in the canopy interception calculation.

RSMAX (type: double, unit: $\mathrm{s\ m}^{-1}$) Maximum stomatal resistance.

4.1.5 River file

The river (`.riv`) file provides description of the river channel structure and properties. It consists of five blocks. The river attribute block describes the topographical information related to river segments (the upstream and downstream nodes and the left and right bank elements), and river shapes, materials, boundary conditions, and reservoir types. The shape and material blocks contain the parameters for different shape and material types. The boundary conditions and reservoir time series are provided in the boundary condition and reservoir blocks. A river file looks like this:

```
NUMRIV  20
INDEX   FROM   TO  DOWN   LEFT   RIGHT   SHAPE   MATL   BC  RES
1       1      2   2      118    162     1       1      0   0
2       2      97  3      153    160     1       1      0   0
3       97     93  4      120    155     1       1      0   0
...
18      41     43  19     15     63      4       2      0   0
19      43     9   20     58     59      4       2      0   0
20      9      10  -3     17     66      4       2      0   0
SHAPE   4
INDEX   DPTH   OINT  CWID
1       0.1    1     1.5
2       0.3    1     1.5
3       0.4    1     1.5
4       0.4    1     1.5
MATERIAL  2
INDEX   ROUGH  CWR  KH
1       0.04   0.6  6.962E-5
2       0.04   0.6  6.962E-5
BC  1
RIV_TS  1  TYPE  1  # 1:  Dirichlet, 2:  Neumann
TIME              HEAD
2008-01-01 00:00  270.0
2012-07-31 00:00  270.0
RES  0
```

A description of the input variables can be found below.

Attribute block

`NUMRIV` (type: integer, unit: -) Number of river segments.

`INDEX` (type: integer, unit: -) River segment index.

`FROM` (type: integer, unit: -) From node index.

`TO` (type: integer, unit: -) To node index.

`DOWN` (type: integer, unit: -) Downstream segment index.

`LEFT` (type: integer, unit: -) Left element index.

`RIGHT` (type: integer, unit: -) Right element index.

`SHAPE` (type: integer, unit: -) Shape type.

`MATL` (type: integer, unit: -) River bank and bed material type.

`BC` (type: integer, unit: -) Boundary condition.

`RES` (type: integer, unit: -) Reservoir type.

Shape block

`SHAPE` (type: integer, unit: -) Number of shape types.

`INDEX` (type: integer, unit: -) River shape type index.

`DPTH` (type: double, unit: m) Depth of river.

`OINT` (type: double, unit: dimensionless) Interpolation order.

`CWID` (type: double, unit: dimensionless) Width coefficient.

Interpolation order (b) and width coefficient (a) are parameters defining the relationship between river width and depth as $D = ax\left(\frac{W}{2}\right)^b$.

Material block

`MATERIAL` (type: integer, unit: -) Number of material types.

`INDEX` (type: integer, unit: -) River material type index.

`ROUGH` (type: double, unit: s m$^{-1/3}$) Manning's roughness coefficient.

`CWR` (type: double, unit: dimensionless) Discharge coefficient.

`KH` (type: double, unit: m s^{-1}) River bank hydraulic conductivity.

Boundary condition block

BC (type: integer, unit: -) Number of boundary conditions.

RIV_TS (type: integer, unit: -) River boundary condition time series index.

TYPE (type: integer, unit: -) Type of boundary condition. Use 1 for Dirichlet boundary conditions, and 2 for Neumann boundary conditions.

Note that the numbers for water heads in Dirichlet boundary conditions should be the height above the datum. The numbers for water fluxes in Neumann boundary conditions should be positive if representing sources (fluxes into the river channel), and negative if representing sinks (fluxes out of the river channel).

Reservoir block

RES (type: integer, unit: -) Number of reservoir time series.

4.1.6 Meteorological forcing file

The meteorological forcing (.meteo) file contains the meteorological forcing for the model, including precipitation rate, air temperature, relative humidity, surface wind speed, downward solar radiation, downward longwave radiation, and surface pressure. Note that the meteorological forcing file can contain multiple meteorological forcing time series. It is required by all MM-PIHM models. The forcing file can use almost any desirable time intervals. The model will interpolate between the two forcing time steps to drive the simulation. A meteorological forcing file looks like this:

```
METEO_TS  1  WIND_LVL  10.0
TIME              PRCP        SFCTMP  RH     SFCSPD  SOLAR  LONGWV  PRES
2008-01-01 00:00  0.00047200  272.35  83.88  2.62    0.00   227.36  97614.16
2008-01-01 01:00  0.00047200  271.94  84.75  1.75    0.00   224.42  97625.35
2008-01-01 02:00  0.00047200  270.75  90.53  2.77    0.00   231.55  97628.40
2008-01-01 03:00  0.00047200  270.82  90.02  3.79    0.00   231.40  97576.52
```

A description of the input variables can be found below.

METEO_TS (type: integer, unit: -) Index of meteorological forcing time series.

WIND_LVL (type: double, unit: m) Height of wind speed observation.

TIME (type: string, unit: -) Time of forcing observation.

PRCP (type: double, unit: kg m^2 s^{-1}) Precipitation rate.

SFCTMP (type: double, unit: K) Surface air temperature.

RH (type: double, unit: %) Relative humidity.

SFCSPD (type: double, unit: m s^{-1}) Wind speed measured at WIND_LVL.

SOLAR (type: double, unit: W m^{-2}) Downward solar radiation.

LONGWV (type: double, unit: W m^{-2}) Downward longwave radiation.

PRES (type: double, unit: Pa) Surface pressure.

4.1.7 Control parameter file

The control parameter (.para) file consists of the model control block, CVODE control block and output control block. The model control block provides all the control parameters to the model, including model start and end time, time steps, etc. The CVODE parameter block provides control parameters to the CVODE solver. The output control block controls which variables should be printed for output with what output interval. A description of the input variables can be found below.

Model control parameter block

SIMULATION_MODE (type: integer, unit: -) Simulation mode. 0: Normal model simulation; 1: Spin-up simulation. In spin-up simulation mode, model will run repeatedly and forcing will be recycled until model reaches equilibrium, or reaches maximum of years of simulation (MAX_SPINUP_YEAR).

INIT_MODE (type: integer, unit: -) Initialization mode. 0: Relaxation mode (model starts from saturation); 1: Hot start mode (model starts from the initial condition described in .ic file).

ASCII_OUTPUT (type: integer, unit: -) ASCII output switch. 0: Do not print ASCII output; 1: Print ASCII output.

WATBAL_OUTPUT (type: integer, unit: -) Diagnostic water balance output switch. 0: Do not print water balance output; 1: Print water balance output file (.watbal.plt).

WRITE_IC (type: integer, unit: -) Write initial conditions to file switch. 0: Do not write .ic files; 1: Model state variables will be written into .ic files in the OUTPUT_DIR/restart directory. The interval of writing initial condition files can be set using the IC parameter in the print control block. If the IC parameter is 0, only the last model step output will be written into a restart file.

START (type: string, unit: -) Simulation start time (yyyy-mm-dd HH:MM).

END (type: string, unit: -) Simulation end time (yyyy-mm-dd HH:MM).

MAX_SPINUP_YEAR (type: integer, unit: year) Maximum number of years of simulation in spin-up mode.

MODEL_STEPSIZE (type: integer, unit: s) Model step size. This is the step size that model state variables are reported, but not the step for CVODE integration.

LSM_STEP (type: integer, unit: s) Interception, evapotranspiration, and snow physics calculation time step. When the Noah land surface model (LSM) is used, this is the time step for the Noah LSM. Must not be smaller than MODEL_STEPSIZE.

CVODE parameter block

ABSTOL (type: double, unit: m) Absolute integration tolerance. See CVODE manual for details.

RELTOL (type: double, unit: dimensionless) Relative integration tolerance. See CVODE manual for details.

INIT_SOLVER_STEP (type: double, unit: s) Initial solver step size. See CVODE manual for details.

NUM_NONCOV_FAIL (type: double, unit: -) Tolerance of CVODE non-convergence failures per CVODE integration step.

MAX_NONLIN_ITER (type: double, unit: -) Threshold number of non-linear iteration per CVODE integration step to decrease the upper bound of the CVODE step size.

MIN_NONLIN_ITER (type: double, unit: -) Threshold number of non-linear iteration per CVODE integration step to increase the upper bound of the CVODE step size.

DECR_FACTOR (type: double, unit: dimensionless) Decrease factor of the upper bound of the CVODE step size.

INCR_FACTOR (type: double, unit: dimensionless) Increase factor of the upper bound of the CVODE step size.

MIN_MAXSTEP (type: double, unit: s) Minimum value of the upper bound of the CVODE step size.

In MM-PIHM, the upper bound of the CVODE step size is adaptable to minimize the oscillation of CVODE solutions. Within each model time step, when the average number of non-convergence failures per CVODE integration step is larger than NUM_NONCOV_FAIL, or when the average number of non-linear iteration per CVODE integration step is larger than MAX_NONLIN_ITER, the upper bound of the CVODE step size will be decreased by DECR_FACTOR to assure stability; when the average number of non-convergence failures per CVODE integration step is smaller than NUM_NONCOV_FAIL, and at the same time the number of non-linear iteration per CVODE integration step is smaller than MIN_NONLIN_ITER, the upper bound of the CVODE step size will be increased by INCR_FACTOR to accelerate the model. The upper bound of the CVODE step size, however, is always larger than MIN_MAXSTEP, and smaller than MODEL_STEPSIZE.

Output control block The output intervals can be YEARLY, MONTHLY, DAILY, HOURLY, or a positive integer with the unit of second. If the output interval for a variable is 0, the variable will not be printed for output. See Table 5.1 for a list of available output variables and their output units.

4.1.8 Calibration file

The calibration (.calib) file contains the global calibration coefficient of key model parameters and climate scenario parameters. All calibration coefficients in the .calib file are the calibration multiplier to the original corresponding variables. The climate scenario begins with the keyword SCENARIO, and the parameters can be a multiplier or an offset. A description of the variables can be found below.

Calibration block

`KSATH` (type: double, unit: dimensionless) Calibration coefficient for horizontal saturation conductivity (Section 4.1.3).

`KSATV` (type: double, unit: dimensionless) Calibration coefficient for vertical saturation conductivity (Section 4.1.3).

`KINF` (type: double, unit: dimensionless) Calibration coefficient for vertical infiltration saturation conductivity (Section 4.1.3).

`KMACSATH` (type: double, unit: dimensionless) Calibration coefficient for horizontal macropore saturation conductivity (Section 4.1.3).

`KMACSATV` (type: double, unit: dimensionless) Calibration coefficient for vertical macropore saturation conductivity (Section 4.1.3).

`DROOT` (type: double, unit: dimensionless) Calibration coefficient for rooting depth (Section 4.1.4).

`DMAC` (type: double, unit: dimensionless) Calibration coefficient for macropore depth (Section 4.1.3).

`POROSITY` (type: double, unit: dimensionless) Calibration coefficient for porosity (Section 4.1.3).

`ALPHA` (type: double, unit: dimensionless) Calibration coefficient for van Genuchten α (Section 4.1.3).

`BETA` (type: double, unit: dimensionless) Calibration coefficient for van Genuchten β (Section 4.1.3).

`MACVF` (type: double, unit: dimensionless) Calibration coefficient for vertical areal fraction of macopore (Section 4.1.3).

`MACHF` (type: double, unit: dimensionless) Calibration coefficient for horizontal areal fraction of macopore (Section 4.1.3).

`VEGFRAC` (type: double, unit: dimensionless) Calibration coefficient for vegetation fraction (Section 4.1.4).

`ALBEDO` (type: double, unit: dimensionless) Calibration coefficient for albedo (Section 4.1.4).

ROUGH (type: double, unit: dimensionless) Calibration coefficient for Manning's n for landcover (Section 4.1.4).

ROUGH_RIV (type: double, unit: dimensionless) Calibration coefficient for river Manning's n (Section 4.1.5).

KRIVH (type: double, unit: dimensionless) Calibration coefficient for river bank conductivity (Section 4.1.5).

RIV_DPTH (type: double, unit: dimensionless) Calibration coefficient for river depth (Section 4.1.5).

RIV_WDTH (type: double, unit: dimensionless) Calibration coefficient for river width coefficient (Section 4.1.5).

Climate scenario block

PRCP (type: double, unit: dimensionless) Multiplier for precipitation forcing.

SFCTMP (type: double, unit: K) Offset for surface air temperature.

Land surface module calibration block When the surface heat flux module is turned on, a land surface calibration block should be added beginning with the keyword LSM_CALIBRATION after the hydrologic calibration, but before the climate scenario block, with the following variables:

DRIP (type: double, unit: dimensionless) Calibration coefficient for through-fall of precipitation and/or dew.

CMCMAX (type: double, unit: dimensionless) Calibration coefficient for maximum canopy water capacity.

RS (type: double, unit: dimensionless) Calibration coefficient for canopy resistance (Section 4.1.4).

CZIL (type: double, unit: dimensionless) Calibration coefficient for Zilitinkevich constant (Section 4.2.1).

FXEXP (type: double, unit: dimensionless) Calibration coefficient for soil evaporation exponent used in direct evaporation (Section 4.2.1).

`CFACTR` (type: double, unit: dimensionless) Calibration coefficient for parameter used in the canopy interception calculation (Section 4.1.4).

`RGL` (type: double, unit: dimensionless) Calibration coefficient for reference incoming solar flux (Section 4.1.4).

`HS` (type: double, unit: dimensionless) Calibration coefficient for parameter used in vapor pressure deficit function (Section 4.1.4).

`REFSMC` (type: double, unit: dimensionless) Calibration coefficient for soil moisture threshold where transpiration begins to stress.

`WLTSMC` (type: double, unit: dimensionless) Calibration coefficient for wilting point.

Deep groundwater module calibration block When the deep groundwater module is turned on, a deep groundwater calibration block should be added beginning with the keyword `DGW_CALIBRATION` after the land surface module calibration block, but before the climate scenario block, with the following variables:

`KSATH` (type: double, unit: dimensionless) Calibration coefficient for deep zone horizontal saturated hydraulic conductivity (Section 4.5.2).

`KSATV` (type: double, unit: dimensionless) Calibration coefficient for deep zone vertical saturated hydraulic conductivity (Section 4.5.2).

`POROSITY` (type: double, unit: dimensionless) Calibration coefficient for deep zone porosity (Section 4.5.2).

`ALPHA` (type: double, unit: dimensionless) Calibration coefficient for deep zone van Genuchten α (Section 4.5.2).

`BETA` (type: double, unit: dimensionless) Calibration coefficient for deep zone van Genuchten β (n) (Section 4.5.2).

`KMACSATH` (type: double, unit: dimensionless) Calibration coefficient for deep zone macropore horizontal saturated hydraulic conductivity (Section 4.5.2).

KMACSATV (type: double, unit: dimensionless) Calibration coefficient for deep zone macropore vertical saturated hydraulic conductivity (Section 4.5.2).

DMAC (type: double, unit: dimensionless) Calibration coefficient for deep zone macropore depth (Section 4.5.2).

MACVF (type: double, unit: dimensionless) Calibration coefficient for deep zone macropore area fraction on a vertical cross-section (Section 4.5.2).

MACHF (type: double, unit: dimensionless) Calibration coefficient for deep zone macropore area fraction on a horizontal cross-section (Section 4.5.2).

Biogeochemical module calibration block When the biogeochemical module is turned on, a biogeochemical calibration block should be added beginning with the keyword BGC_CALIBRATION after the land surface module calibration block, but before the climate scenario block, with the following variables:

MORTALITY (type: double, unit:dimensionless) Calibration coefficient for mortality turnover rate.

SLA (type: double, unit:dimensionless) Calibration coefficient for canopy average projected SLA.

Reactive transport module calibration block When the reactive transport module is turned on, a reactive transport calibration block should be added beginning with the keyword RT_CALIBRATION after the deep groundwater module calibration block, but before the climate scenario block, with the following variables:

RATE (type: double, unit: ?) Offset for rate of kinetic reaction.

SSA (type: double, unit: dimensionless) Calibration coefficient for specific surface area.

XSORPTION (type: double, unit: ?) Offset for cation exchange equilibrium constant.

4.1.9 Boundary condition file

The boundary condition (.bc) file contains boundary condition time series. It is only required when one or more elements have non-zero BCi values. When the bioRT module is not activated, a boundary condition file looks like this:

```
BC_TS  1  TYPE  1      # 1:Dirichlet, 2:Neumann
TIME              HEAD
2008-01-01 00:00  270.0
2012-07-31 00:00  270.0
BC_TS  2  TYPE  2      # 1:Dirichlet, 2:Neumann
TIME              FLUX
2008-01-01 00:00  1.0E-5
2012-07-31 00:00  1.0E-5
```

When bioRT module is activated, the specified boundary conditions should also specify the associated mineral concentration of all primary species. The species can be described in any order in the file. A boundary condition file looks like this when bioRT module is activated:

```
BC_TS  1  TYPE  1      # 1:Dirichlet, 2:Neumann
TIME              HEAD   pH    Cl-
2008-01-01 00:00  270.0  4.30  1.93E-5
2012-07-31 00:00  270.0  4.30  1.93E-5
```

A description of the input variables can be found below.

`BC_TS` (type: integer, unit: -) Index of boundary condition time series.

`TIME` (type: string, unit: -) Time of specified boundary condition.

`TYPE` (type: integer, unit: -) Type of boundary condition. Use `1` for Dirichlet boundary conditions, and `2` for Neumann boundary conditions.

`HEAD` (type: double, unit: m) Hydraulic head for Dirichlet boundary conditions.

`FLUX` (type: double, unit: m^3 s^{-1}) Flux for Neumann boundary conditions.

Note that the numbers for water heads in Dirichlet boundary conditions should be the height above the datum. The numbers for water fluxes in Neumann boundary conditions should be positive if representing sources (fluxes into the river channel), and negative if representing sinks (fluxes out of the river channel).

4.1.10 Initial condition file

The initial condition (`.ic`) file is a binary file which contain the initial state variables for hydrology and land surface. It is only required when starting the model using the "hot start" mode (i.e., INIT_MODE set to 1; see Section 4.1.7). The best way to generate initial condition files is to run the MM-PIHM models from the relaxation mode, with the WRITE_IC option set to 1. When the WRITE_IC option is turned on, .ic files will be written into the restart directory within the simulation output directory. They can then be renamed, moved to the input directory, and used to "hot start" the model.

The .ic files generated by PIHM can only be used to initialize PIHM runs. The .ic files generated by Flux-PIHM can only be used to initialize Flux-PIHM runs. Errors will be reported if PIHM initial conditions are used for Flux-PIHM, or vice versa. PIHM initial conditions include the equivalent surface water storage, unsaturated zone storage, groundwater storage, canopy water storage, and snow water equivalent storage of all triangular elements, and the stream water level of all river segments. Flux-PIHM initial conditions also include actual snow depth, soil water and moisture contents of all Flux-PIHM soil layers, soil temperature of all Flux-PIHM soil layers, and land surface temperature of all triangular elements, in addition to PIHM initial conditions.

4.1.11 Leaf area index forcing file

The leaf area index forcing (.lai) file contains the leaf area index forcing for the model. Like meteorological forcing files, LAI files can also contain multiple LAI forcing time series. If the LAI attributes of all elements in the element attribute (.att) file (Section 4.1.2) are 0, LAI forcing file is not needed. A description of the input variables can be found below.

LAI_TS (type: integer, unit: -) Index of LAI forcing time series.

TIME (type: string, unit: -) Time of forcing observation.

LAI (type: double, unit: m^2 m^{-2}) Leaf area index.

4.2 Land surface module input files

4.2.1 Land surface model control file

The land surface model control (.lsm) file contains the parameters used for the modified Noah LSM. This is only needed by the surface heat flux module (i.e., Flux-PIHM, bioRT-Flux-PIHM and Flux-PIHM-BGC). A description of the input variables can be found below.

LATITUDE (type: double, unit: degree) Latitude of model domain.

LONGITUDE (type: double, unit: degree) Longitude of model domain.

NSOIL (type: integer, unit: -) Default number of soil layers in the land surface model. The number of soil layers is subject to change to adapt to the bedrock depths of different model grids.

SLDPTH_DATA (type: double, unit: m) Default depths of soil layers in the land surface model. Note that the depths are subject to change to adapt to the bedrock depths of different model grids.

RAD_MODE_DATA (type: integer, unit: -) Solar radiation mode. 0: Uniform downward solar radiation; 1: Topographic solar radiation (a radiation forcing file is needed).

SBETA_DATA (type: double, unit: dimensionless) Used to compute vegetation canopy effect on ground heat flux as a function of greenness.

FXEXP_DATA (type: double, unit: dimensionless) Bare soil evaporation exponent.

CSOIL_DATA (type: double, unit: J m^3 K^{-1}) Soil heat capacity.

SALP_DATA (type: double, unit: dimensionless) Shape parameter used in function to infer percent area snow cover from snow depth

FRZK_DATA (type: double, unit: dimensionless) A base reference value (for light clay soil type) of parameter for the frozen-soil freeze factor representing the ice content threshold above which frozen soil is impermeable.

ZBOT_DATA (type: double, unit: m) Nominal depth of TBOT_DATA: lower boundary condition on soil temperature.

TBOT_DATA (type: double, unit: K) The lower boundary condition of soil temperature applied at the depth specified by parameter ZBOT_DATA.

CZIL_DATA (type: double, unit: dimensionless) Zilitinkevich parameter (range 0.0–1.0).

LVCOEF_DATA (type: double, unit: dimensionless) Parameter that controls surface snow albedo in the presence of snow cover.

The land surface model control file also controls the output from the Noah LSM. See Table 5.2 for a list of available output variables and their output units.

4.2.2 Solar radiation forcing file

The solar radiation forcing (`.rad`) file contains the direct and diffused solar radiation time series to calculate topographic solar radiation. It is only required when the `RAD_MODE_DATA` in the land surface model control (`.lsm`) file is set to `1` (i.e., topographic solar radiation mode). A description of the input variables can be found below.

`RAD_TS` (type: integer, unit: -) Index of solar radiation time series.

`TIME` (type: string, unit: -) Time of solar radiation forcing.

`SDIR` (type: double, unit: W m−2) Direct solar radiation.

`SDIF` (type: double, unit: W m−2) Diffused solar radiation.

4.2.3 Ice thickness file

The ice thickness (`.ice`) file contains the initial glacial ice thickness of each model grid. It is not needed if the domain has no "permanent snow" or "perennial ice/snow" land use types, and is only optional when any model grid's land use type is "permanent snow" or "perennial ice/snow". When an ice thickness file is not present, all glacier model grids will be initialized with a glacier ice layer of 0.5 m thick. Note that the ice thicknesses specified in the ice thickness file for non-glacier model grids have no effects. Non-glacier model girds will always be initialized with zero ice thickness in Flux-PIHM simulations.

A description of the input variables can be found below.

`INDEX` (type: integer, unit: -) Element index.

`ICEH` (type: double, unit: m) Initial glacier ice thickness.

4.3 Biogeochemical module input files

4.3.1 Biogeochemical model control file

The biogeochemical model control (`.bgc`) file contains the control parameters used for the biogeochemical module, adapted from Biome-BGC. This is only needed by the BGC module (i.e., Flux-PIHM-BGC). The biogeochemical model control file consists of the restart block, time define block, CO_2 control block, nitrogen deposition control block, carbon state block, nitrogen state block, and output control block.

Restart block The restart block is used to determine whether a restart file should be used to initialize the model, and whether a restart file (`.bgcic`) should be written at the end of simulation. If the user chooses to write a restart file, the BGC restart file containing initial conditions will be written at the end of simulation in the `OUTPUT_DIR/restart` directory.

Time define block This block defines BGC spin-up method. In modified accelerated decomposition spin-up simulations, the base decomposition rate for the slow microbial recycling pool and the recalcitrant soil organic matter pool are accelerated by 5 and 70 times, respectively. When exiting the accelerated decomposition spin-up mode, the carbon and nitrogen stocks in the slow microbial recycling pool and the recalcitrant soil organic matter pool are multiplied by 5 and 70, respectively, to compensate for the accelerated decomposition rates. Thus the restart file produced in accelerated spin-up simulations can be used directly for native dynamics spin-up or normal simulations. The native dynamics spin-up model uses the native BGC dynamics. For Flux-PIHM-BGC spin-up, it is recommended to perform a modified accelerated decomposition spin-up first, followed by a native dynamics spin-up.

CO_2 control and nitrogen deposition control blocks The CO_2 control block and nitrogen deposition control block specify the CO_2 concentration and nitrogen deposition rate to be used in the model. If a constant CO_2 concentration (or N deposition rate) is desired, the CO_2 concentration (or N deposition rate) defined in the block will be used. If varying CO_2 concentration (or N deposition rate) is desired, the name of the file that contains the annual CO_2 concentration (or N deposition rate) time series should be specified. The deposition control block also defines the nitrogen fixation rate. Note that in spin-up simulations, constant CO_2 concentration and N fixation rates are always used. The nitrogen fixation rate is always a constant in all simulation modes.

Carbon state block The carbon state block specifies the initial carbon states when BGC restart files are not being used. The carbon states include: peak leaf carbon to be attained during the first simulation year; peak stem carbon to be attained during the first year; initial coarse woody debris carbon (dead trees, standing or fallen), initial litter carbon, labile pool; initial litter carbon, unshielded cellulose pool; initial litter carbon, shielded cellulose pool; initial litter carbon, lignin pool; soil carbon, fast pool; soil carbon, medium pool; soil carbon, slow pool; and soil carbon, slowest pool.

Nitrogen state block The nitrogen state block specifies the initial nitrogen states when BGC restart files are not being used, including litter nitrogen associated with labile litter carbon pool and soil mineral nitrogen pool.

Output control block This block controls the output from the biogeochemical module. See Table 5.3 for a list of available output variables and their output units.

4.3.2 Ecophysiological files

The ecophysiological (`.epc`) files describe the ecophysiological properties of the vegetation being simulated. These files are provided in the `input/epc` directory. Currently, ecophysiological descriptions of seven vegetation types are provided: C3 grass (`c3grass.epc`, for land cover type 10, grassland), C4 grass (`c4grass.epc`, not being used anywhere), deciduous broadleaf forests (`dbf.epc` for land cover type 4), deciduous needleleaf forests (`dnf.epc`, for land cover type 3), evergreen broadleaf forests (`ebf.epc`, for land cover type 2), evergreen needleleaf forests (`enf.epc`, for land cover type 1), and shrubs (`shrub.epc`, for land cover types 6 and 7, open and closed shrubland). Users are referred to *Users Guide for Biome-BGC, Version 4.1.2* for details of the ecophysiological files.

4.3.3 Annual CO_2 concentration and nitrogen deposition files

The annual CO_2 concentration and nitrogen deposition files are required when varying CO_2 concentration and nitrogen deposition options are turned on. These files are the same as in Biome-BGC. They must have one line for each simulation year, and the format should be:

```
yyyy  value
```

where `yyyy` is the four-digit year, the value is the CO_2 mole fraction (ppm, for CO_2 concentration file), or nitrogen deposition rate (kg N m^2 $year^{-1}$, for nitrogen deposition file).

4.3.4 Biogeochemical restart file

The biogeochemical restart (`.bgcic`) file is a binary file which contains the initial conditions for the biogeochemical module. This is an optional input file. The biogeochemical restart files should be generated by the BGC module by turning on the "write restart file" option in the biogeochemical model control file (Section 4.3.1). The generated `.bgcic` files will be written into the `restart` directory within the simulation output directory. They can then be renamed, moved to the input directory, and be used to "hot start" the BGC module. Note that in Flux-PIHM-BGC, initial conditions for hydrology and land surface and initial conditions for biogeochemistry are controlled separately. Users can "hot start" the hydrology and land surface but "cold start" biogeochemistry, or *vice versa*.

4.4 Reactive transport module input files

4.4.1 Chemical model control file

The chemical model control (`.chem`) file follows the reactive transport code CrunchFlow input file structure. This input file is required by the bioRT module. This file consists of the runtime block, global block, primary species block, secondary species block, minerals block, and precipitation concentration block.

Runtime block

`INIT_TYPE` (type: int, unit: -) bioRT module initialization type. 0: Use `.cini` file; 1: Use `.rtic` file.

`WRITE_IC` (type: int, unit: -) Write bioRT initial condition to file switch. 0: Do not write initial condition to file; 1: Write bioRT initial condition to `.rtic` file.

`ACTIVITY` (type: int, unit: -) Activity model switch. 0: Disabled; 1: Use Debye-Huckel equation.

`THERMO` (type: int, unit: -) Thermal modeling switch. If enabled (1), bioRT simulator will use temperature data to compute chemical equilibrium and kinetics.

`RELMIN` (type: int, unit: -) Mineral volume fraction mode: 0: Absolute volume fraction; 1: Relative fraction (to total mineral volume fraction)

`TRANSPORT_ONLY` (type: int, unit: -) Transport only mode switch. 0: Turn on reactions. 1: Transport only mode with reactions skipped.

`PRECIPITATION` (type: int, unit: -) Precipitation concentration flag. 0: No precipitation concentration will be specified (zero solute concentration in precipitation assumed); 1: Use precipitation concentrations specified in `.chem` file; 2: Use precipitation concentration time series specified in `.prep` file.

`RT_DELAY` (type: int, unit: s) Run bioRT simulator after certain seconds from the start of simulation.

`CONDENSATION` (type: double, unit: dimensionless) The concentration ratio between throughfall and rainwater, generally 1.0–5.0

`AVGSCL` (type: int, unit: s) Reaction time step. If reaction error is encountered, reduce reaction time step until no error is reported.

`OUTINTVL` (type: int, unit: s) Chemical output interval. Can also be “YEARLY”, “MONTHLY”, “DAILY”, or “HOURLY”.

See Table 5.4 for a list of chemical output variables and their output units.

Global block

`T_SPECIES` (type: int, unit: -) Number of total primary species including minerals.

`S_SPECIES` (type: int, unit: -) Number of secondary species.

MIN_SPECIES (type: int, unit: -) Number of minerals.

ADSORPTION (type: int, unit: -) Number of primary adsorption species.

CATION_EXCHANGE (type: int, unit: -) Number of primary cation exchange species.

MINERAL_KINETIC (type: int, unit: -) Number of kinetic mineral reactions.

AQUEOUS_KINETIC (type: int, unit: -) Number of kinetic aqueous reactions.

DIFFUSION (type: double, unit: $\mathrm{cm}^2\ \mathrm{s}^{-1}$) Fixed diffusion coefficient.

DISPERSION (type: double, unit: m) Fixed dispersivity.

CEMENTATION (type: double, unit: dimensionless) Cementation factor.

TEMPERATURE (type: double, unit: °C) Field temperature. Voided if thermal option is enabled.

Primary species block This block should list all primary species.

Secondary species block This block should list all secondary species the users would like to track. A complete list of the available secondary species resulted from the specified primary species could be generated via running a CrunchFlow instance with the same primary species set using the database_sweep option.

Minerals block This block should list all kinetic mineral reactions. This block follows a similar format as the corresponding block in CrunchFlow input file.

Precipitation concentration block This block should specify precipitation concentrations of all primary species in mol L^{-1}.

4.4.2 Chemical initial condition files

Two types of chemical initial condition files can be used to initialize the bioRT module. For “cold start”, .cini file is used. For “hot start”, .rtic file is required.

Cold start initial condition file

The cold start chemical initial condition file (`.cini`) is text-based. It describes a series of initial conditions and assigns them to different model grids. A `.cini` initial condition file looks like this:

```
ELEM  PRCPC  SOIL  GEOL
1     1      3     4
2     1      3     4
3     1      2     4
…
533   1      2     4
534   1      2     4
535   1      1     4
Condition RT
pH      4.56
Cl-     3.69E-5
Quartz  0.6199   -ssa  1.48
Condition MS
pH      4.56
Cl-     3.69E-5
Quartz  0.6199   -ssa  1.48
Condition VF
pH      4.56
Cl-     3.69E-5
Quartz  0.6199   -ssa  1.48
Condition BR
pH      4.56
Cl-     3.69E-5
Quartz  0.0620   -ssa  1.48
```

The `ELEM` column is the index of model grids. The `PRCPC` column specifies which precipitation time series in the `.prep` file should be used for each grid. This is only used when the precipitation concentration flag is set to 2. The `SOIL` (`GEOL`) column specifies which initial condition should be used for the shallow (deep) zone of each grid. Note that the column `GEOL` is not needed if the deep groundwater module is not turned on.

The initial condition block contains multiple total concentrations of primary species similar to CrunchFlow. The default unit for aqueous concentrations is mol L^{-1}, for volume fractions is m^3 m^{-3}, for specific surface area is m^2 g^{-1}, for surface site density is eq g^{-1}.

Hot start initial condition file

The hot start chemical initial condition file (`.rtic`) is a binary file that contains the initial conditions for the bioRT module. This is only required when bioRT initialization type in the chemical model control file (Section 4.4) is set to 1, i.e., using `.rtic` file. It can be generated by the bioRT module when the "write IC" option in the chemical control file is turned on. Model generated `.rtic` files can be found in the `restart` directory and can be moved to the corresponding input directory to "hot start" the bioRT module. Same as the BGC module, users can hot start the hydrology and land surface but cold start bioRT, or *vice versa*.

4.4.3 Chemical database file

The chemical database (`.cdbs`) file has the exactly same format as CrunchFlow's database files. In fact, the users can use bioRT's database file and CrunchFlow's database file interchangeably. For a complete description on the CrunchFlow's database file, please refer to CrunchFlow user's manual.

4.4.4 Precipitation solute concentration file

The precipitation solute concentration (`.prep`) file is an optional input file for the bioRT module. It is only required when the precipitation concentration flag in the chemical model control file (Section 4.4) is set to 2. This file can contain multiple precipitation solute concentration time series to be used for different model grids. A precipitation solute concentration file looks like this:

```
PRCP_CONC_TS  1  Pconc  1
TIME              Cl-
2008-01-01 00:00  2.500e-06
2008-01-15 00:00  2.98e-06
2008-01-22 00:00  7.50e-06
```

`PRCP_CONC_TS` is the index of precipitation solute concentration time series. `Pconc` is the number of solute species described in the time series. For each column of solute concentration, the name of the solute species should be provided in the header line (e.g., `Cl-`). The unit for solute concentration is mol L^{-1}. For those primary species that are not described in the precipitation solute concentration file, the concentrations in the precipitation concentration block in the chemical model control file (Section 4.4) will be used.

4.5 Deep groundwater module input files

4.5.1 Bedrock structure file

The bedrock structure (`.bedrock`) file is required when the deep groundwater module (DGW) is turned on. It describes the boundary conditions of the deep groundwater layer, and the elevation of the deep zone bottom. It consists of the element block, the node block, and the output control block. A description of the input variables can be found below.

Element block

`INDEX` (type: integer, unit: -) Element index.

`BC`*i* (type: integer, unit: -) Boundary condition forcing time series on edge *i* for the deep zone. The time series correspond to the time series in the `.bc` file. Default values for boundary condition is `0`, which indicates no flow condition and does not link to any time series.

Node block

`INDEX` (type: integer, unit: -) Node index.

`ZBED` (type: double, unit: m) Elevation of impermeable deep zone bottom. The deep zone is between `ZMIN` in the mesh structure file (Section 4.1.1) and `ZBED`.

Output control block This block controls the output from the deep zone. See Table 5.5 for a list of available output variables and their output units.

4.5.2 Geology parameter file

The geology parameter (`.geol`) file contains the hydraulic and properties of all geology classes for the deep zone. The file format is similar to the soil parameter file (Section **??**). A description of the input variables can be found below.

`NUMGEOL` (type: integer, unit: -) Number of geology classes described.

`INDEX` (type: integer, unit: -) Geology class index.

KSATV (type: double, unit: m s^{-1}) Vertical saturated hydraulic conductivity of the deep zone.

KSATH (type: double, unit: m s^{-1}) Horizontal saturated hydraulic conductivity of the deep zone.

MAXSMC (type: double, unit: m^3 m^{-3}) Maximum moisture content of the deep zone.

MINSMC (type: double, unit: m^3 m^{-3}) Residual moisture content of the deep zone.

ALPHA (type: double, unit: m^{-1}) Van Genuchten parameter (related to the inverse of the air entry suction) of the deep zone.

BETA (type: double, unit: dimensionless) Van Genuchten parameter (measure of the pore-size distribution) of the deep zone.

MACHF (type: double, unit: m^2 m^{-2}) Horizontal area fraction of macropore of the deep zone.

MACVF (type: double, unit: m^2 m^{-2}) Vertical area fraction of macropore of the deep zone.

DMAC (type: double, unit: m) Macropore depth of the deep zone.

KMACV_RO (type: double, unit: dimensionless) Ratio between vertical macropore hydraulic conductivity and vertical saturated infiltration hydraulic conductivity of the deep zone.

KMACH_RO (type: double, unit: dimensionless) Ratio between horizontal macropore hydraulic conductivity and horizontal saturated hydraulic conductivity of the deep zone.

Chapter 5

Output files

MM-PIHM can store model output variables in both text and binary formats. The control for PIHM output should be specified in the model control parameter file (Section 4.1.7), where users can turn on/off output for desired variables as well as change the output intervals. Table 5.1 show all available variables for PIHM output. Tables 5.2, 5.3, 5.4, and 5.5 show available variables for surface heat flux module, biogeochemical module, reactive transport module, and deep groundwater module output, respectively.

The default output format is binary, but ASCII output can be turned on in the model control parameter file (Section 4.1.7) if desired. The output files are stored in the `output/<project>.<time>/` directory by default, and are named as `<project>.<extension>.<fmt>`, where `<project>` is the name of the project, `<time>` is the system time when the simulation is executed, in the format of "yyyymmddHHMM", `<extension>` indicates the variable stored in the output file, and `fmt` can be `dat` (for binary files) or `txt` (for ASCII files). Each output variable is stored in one output file, which contains the time series of all model grid elements, or river segments. All river segment output files have `river.` in their extensions. When deep groundwater module is turned on, all deep layer output files have `deep.` in their extensions.

For the ASCII files, the first column of the output file is output time, in the format of "yyyy-mm-dd HH:MM". The time string is followed by output variables from the first element (or river channel segment) to the last element (or river channel segment). Structure of the binary output files is similar to the ASCII files, except that the output time is stored as the seconds since the Unix epoch, i.e., 00:00:00 UTC 1 January 1970, and the binary files stores the output variables (including the output time) as double-precision floating-point type.

Table 5.1: Description of PIHM output variables.

Variable	Extension	Description	Unit
SURF	surf	Surface water level	m
UNSAT	unsat	Unsaturated water storage	m
GW	gw	Element groundwater level	m
RIVSTG	river.stage	River stage	m
SNOW	snow	Water-equivalent snow depth	m
CMC	is	Canopy interception	m
INFIL	infil	Infiltration rate	m s^{-1}
RECHARGE	recharge	Recharge rate	m s^{-1}
EC	ec	Canopy evaporation	m s^{-1}
ETT	ett	Total transpiration	m s^{-1}
EDIR	edir	Soil evaporation	m s^{-1}
RIVFLX0	river.flx0	Longitudinal flow to river	m^3 s^{-1}
RIVFLX1	river.flx1	Longitudinal flow from river	m^3 s^{-1}
RIVFLX2	river.flx2	Lateral overland flow to river from left	m^3 s^{-1}
RIVFLX3	river.flx3	Lateral overland flow to river from right	m^3 s^{-1}
RIVFLX4	river.flx4	Lateral groundwater flow to river from left	m^3 s^{-1}
RIVFLX5	river.flx5	Lateral groundwater flow to river from right	m^3 s^{-1}
SUBFLX	subflx[0-2]	Subsurface water flux	m^3 s^{-1}
SURFFLX	surfflx[0-2]	Surface water flux	m^3 s^{-1}

Table 5.2: Description of surface heat flux module output variables.

Variable	Extension	Description	Unit
T1	t1	Ground/canopy/snowpack effective skin temperature	K
STC	stc[0-10]	Soil temperature	K
SMC	smc[0-10]	Total soil moisture content	m^3 m^{-3}
SH2O	swc[0-10]	Unfrozen soil moisture content	m^3 m^{-3}
SNOWH	snowh, iceh	Actual snow depth (.snowh) and glacier ice depth (.iceh)	m
ALBEDO	albedo	Surface albedo (including snow effect)	dimensionless
LE	le	Latent heat flux	W m^{-2}
SH	sh	Sensible heat flux	W m^{-2}
G	g	Soil heat flux	W m^{-2}
ETP	etp	Potential evapotranspiration	W m^{-2}
ESNOW	esnow	Sublimation from (or deposition) snow pack	W m^{-2}
ROOTW	rootw	Available soil moisture in root zone (fraction between wilting point and maximum soil moisture content)	dimensionless
SOILM	soilm	Total soil column moisture content	m
SOLAR	solar	Incoming solar radiation	W m^{-2}
CH	ch	Surface exchange coefficient for heat and moisture	m s^{-1}

Table 5.3: Description of biogeochemical module output variables.

Variable	Extension	Description	Unit
LAI	lai	Live projected leaf area index	m^2 m^{-2}
NPP	npp	Daily net primary production (GPP − autotrophic respiration)	kgC m^{-2} day^{-1}
NEP	nep	Net ecosystem production (NPP − heterotrophic respiration)	kgC m^{-2} day^{-1}
NEE	nee	Net ecosystem exchange (NEP − fire losses)	kgC m^{-2} day^{-1}
GPP	gpp	Gross primary production	kgC m^{-2} day^{-1}
MR	mr	Maintenance respiration	kgC m^{-2} day^{-1}
GR	gr	Growth respiration	kgC m^{-2} day^{-1}
HR	hr	Heterotrophic respiration	kgC m^{-2} day^{-1}
FIRE	fire	Fire losses	kgC m^{-2} day^{-1}
LITFALLC	litfallc	Total litterfall	kgC m^{-2} day^{-1}
VEGC	vegc	Total vegetation carbon	kgC m^{-2}
AGC	agc	Aboveground vegetation carbon	kgC m^{-2}
LITRC	litrc	Total litter carbon	kgC m^{-2}
SOILC	soilc	Total soil carbon	kgC m^{-2}
TOTALC	totalc	Total of vegetation, litter and soil carbon	kgC m^{-2}
SMINN	sminn	Soil mineral nitrogen	kgN m^{-2}

Table 5.4: Description of reactive transport module output variables.

Variable	Extension	Description	Unit
`N/A`	`conc.X`	Concentration of chemical species `X`	mol L^{-1}, or m^3 m^{-3}, or m^2 g^{-1}, or eq g^{-1}
`N/A`	`river.conc`	River concentration of chemical species `X`	mol L^{-1}
`N/A`	`river.chflx.X`	Longitudinal chemical flow of chemical species `X` from river	kmol s^{-1}

Table 5.5: Description of deep groundwater module output variables.

Variable	Extension	Description	Unit
`DEEPUNSAT`	`deep.unsat`	Unsaturated storage in deep layer	m
`DEEPGW`	`deep.gw`	Deep groundwater in deep layer	m
`DEEPINFIL`	`deep.infil`	Deep layer infiltration rate	m s^{-1}
`DEEPRECHG`	`deep.rechg`	Deep layer recharge rate	m s^{-1}
`DEEPFLOW`	`deep.flow[0-2]`	Lateral deep layer groundwater flow	m^3 s^{-1}

Chapter 6

Output visualization

Python scripts are provided in the `util` directory to visualize model output data. To run the Python scripts, Python Version 3 and the NumPy library are required.

The `pihm_func.py` script contains functions that read mesh structure input and river input files (`read_mesh` and `read_river`). It also contains a function that parses output files (`read_output`). It reads simulation time steps into a one-dimensional array, and the simulated values into a multi-dimensional array, with rows representing time steps and columns representing model grids (or river segments). Return values of the `read_output` function also include a description of the output variable and the unit, which can be used for annotation.

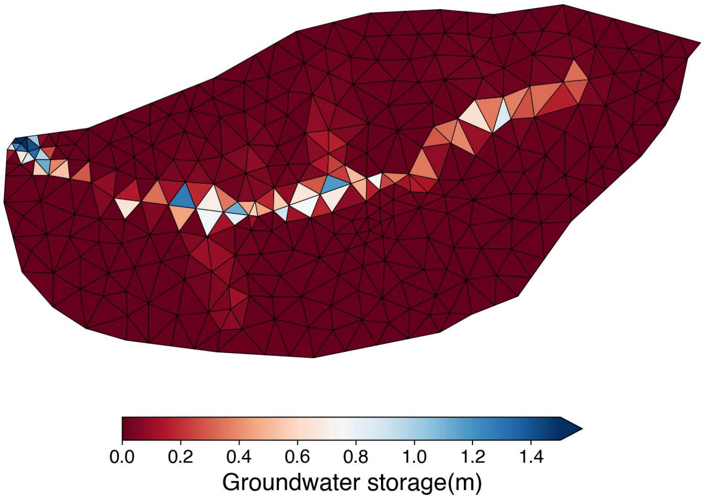

Figure 6.1: Visualization example generated by running `plot.py` script.

The plot.py uses the ShaleHills example output to demonstrate output visualization. It plots the spatial distribution of temporal average groundwater storage, and a time series of spatial average groundwater storage. To use the script, do make test first to run the ShaleHills simulation using Flux-PIHM, and than run

```
$ python3 ./util/plot.py
```

for visualization. Figure 6.1 is an example of the figures produced by the plot.py script. The code can be easily adapted to visualize other projects and variables.

Bibliography

Bao, C., L. Li, Y. Shi and C. Duffy, 2017: Understanding watershed hydrogeochemistry: 1. Development of RT-Flux-PIHM. *Water Resources Research*, **53 (3)**, 2328–2345, doi:10.1002/2016WR018934.

Brantley, S. L. and Coauthors, 2018: Susquehanna Shale Hills Critical Zone Observatory: Shale Hills in the Context of Shaver's Creek Watershed. *Vadose Zone Journal*, **17 (1)**, 180 092, doi:10.2136/vzj2018.04.0092.

Chen, F. and J. Dudhia, 2001: Coupling an advanced land surfacehydrology model with the Penn StateNCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Weather Rev.*, **129 (4)**, 569–585, doi:10.1175/1520-0493(2001)129⟨0569:CAALSH⟩2.0.CO;2.

Hindmarsh, A. C., P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker and C. S. Woodward, 2005: SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, **31 (3)**, 363–396.

Qu, Y. and C. J. Duffy, 2007: A semidiscrete finite volume formulation for multiprocess watershed simulation. *Water Resour. Res.*, **43 (8)**, W08 419, doi:10.1029/2006WR005752.

Shi, Y., K. J. Davis, C. J. Duffy and X. Yu, 2013: Development of a Coupled Land Surface Hydrologic Model and Evaluation at a Critical Zone Observatory. *J. Hydrometeorol.*, **14 (5)**, 1401–1420, doi:10.1175/JHM-D-12-0145.1.

Shi, Y., D. M. Eissenstat, Y. He and K. J. Davis, 2018: Using a spatially-distributed hydrologic biogeochemistry model with a nitrogen transport module to study the spatial variation of carbon processes in a Critical Zone Observatory. *Ecological Modelling*, **380**, 8–21, doi:10.1016/j.ecolmodel.2018.04.007.

Thornton, P. E. and Coauthors, 2002: Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. *Agric. For. Meteor.*, **113 (1–4)**, 185–222, doi:10.1016/S0168-1923(02)00108-9.

Zhi, W., Y. Shi, H. Wen, L. Saberi, G.-H. C. Ng and L. Li, 2020: Biort-flux-pihm v1.0: a watershed biogeochemical reactive transport model. *Geoscientific Model Development Discussions*, **in review**, doi:10.5194/gmd-2020-157.