HPAT v1.1 User Manual and Test Cases

The Hydropower Potential Assessment Tool (HPAT) is a coupled set of tools for estimating streamflow and associated hydropower potential for any region where the necessary importants are available (Mosier et al., 2016a). HPAT is written in Matlab and is initiated by running the “main script”, e.g. “HPAT\_main\_v1.m”, where v1 refers to the version of HPAT and therefore may change over time. The streamflow (i.e. hydrologic) model is a simple temperature index instantiation of that described in Mosier et al. (2016b).

Within the main script, there are several options that must be set. The following provides an overview of information needed to understand how to run HPAT, including a test case that is bundled with the code distribution.

**INPUTS**

The three types of inputs are:

* Climate time-series
* Watershed geometry inputs (i.e. a digital elevation model, corresponding flow direction raster, and glacier cover – if applicable)
* Calibration data (primarily streamgage data, although additional variables can be utilized as well)

Climate data can be from any source, provided that they are formatted properly. The minimum climate input requirements are time-series of precipitation and mean temperature. Often the easiest source of climate data is the Global Climate Data Downscaling Package, distributed at [www.GlobalClimateData.org](http://www.GlobalClimateData.org) (Mosier et al., 2014). These data can be produced for any global land area, for historic or future projection periods. One limitation of these data are that they have a monthly time step, which means that they do not capture extreme events. For many regions, there may be better sources of climate data. The climate inputs must be gridded with equal and uniform latitude and longitude grid steps in geographic coordinates. They can be in ESRI ASCII or NetCDF format (following CMIP5 climate model formatting specifications). *Each climate variables should be contained in a separate subfolder.*

A digital elevation model (DEM) and flow direction raster (FDR) must be supplied in ESRI ASCII format and must have the same grid values and properties as the climate inputs. In many cases the DEM will need to be conditioned in a program such as ArcMap. The main conditioning step is to “fill” the DEM, meaning to remove internal sinks, which can sometimes be a numeric artifact of the spatial sampling and can cause unrealistic streamflow results in HPAT. The FDR is then calculated from the sink-filled DEM using a program such as ArcMap.

The most desirable calibration data are daily streamflow observations from one or more streamgages. In addition, 8-day MODIS snow covered area can be used for calibration as well. These observations are often inaccurate in forest-covered regions, though, which limits their utility to alpine terrain and other environments with minimal ground cover.

**CALIBRATION**

The hydrologic model must be calibrated before HPAT can be used to simulate run-of-river hydropower potential. Calibration mode is turned on by setting “runType” equal to ‘calibration’ in the user inputs section of the HPAT main script (see User Inputs section below). Calibration requires running HPAT many thousands of times and is therefore computationally very expensive. It is therefore recommended that calibration is performed on a small geographic subset of the overall domain of interest. For example, calibration performed on a high-end desktop computer using a daily time step for ten years and a grid that is 10 by 20 may take close to 24 hours.

Several outputs will be generated during calibration, which will be written to a subfolder in the folder containing the input DEM. The prefix of the output folder will be “calibrate\_”. Some of the outputs are:

* “model\_performance.txt”: reports hydrologic model performance relative to input calibration data. Several metrics are computed, including the Kling-Gupta Efficiency (KGE value; ) and mean absolute error (MAE).
* “./mod\_v\_obs\_plots”: subfolder containing time-series and scatter plots comparing calibration data and model output.
* “./model\_outputs\_plots”: subfolder containing any time-series plots produced during the model run.
* “coefficients\_\*\_fittest.txt”: text file containing the best-fitting set of model parameters. The asterisk will be the region name set by the user. This file is used in any validation or simulation model runs to set the model parameters.
* “100\_parameter\_sets-stage\_1.csv”: Text file containing the 100 best-fitting sets of model parameter values. This file can be used during simulation mode in order to simulate hydrologic and hydropower conditions for each of the sets of parameters (e.g. to assess equifinality or other uncertainties related to model parameter selection).
* “coefficients\_\*\_fittest\_parameter\_plot.png”: Graphical representation of best fitting model parameters with respect to the search space for each parameter.
* “parameter\_each\_gen\_\*.txt”: Text file containing the best-fitting set of model parameter values during each generation of calibration.
* “processing\_log\_\*.txt”: Text file with all content written to the command window during the model run.

*As noted above, the “coefficients\_\*\_fittest.txt” file is important because the user will be prompted to select it during any validation or simulation model runs.*

**VALIDATION AND SIMULATIONS**

After the model is run in calibration mode, it can be run in either validation or simulation mode. The primary difference between these two modes is that validation mode utilizes available observation data to assess model performance, while simulation mode does not. Therefore, simulation mode must be used for all future simulations or when the user does not possess observation data to assess model performance. Expect when “100\_parameter\_sets-stage\_1.csv” is used as the input parameter file, the model will only be run once for each of these modes. Therefore, much larger spatial and temporal domains can be used during validation or simulation runs relative to calibration runs. For instance, a spatial grid with 240 by 240 grid cells and 600 time steps takes roughly forty minutes to run on a high-performance desktop computer.

The main outputs during simulation runs are files assessing run-of-run hydropower potential. These are written in either NetCDF or ASCII formats and are output to the subfolder “./hydro\_stats”. The files include:

* “power\_quality.\*”: Spatially gridded file in which the outputs correspond to the run-of-river hydropower resource potential quality metric defined in the function “power\_quality.m”. Power quality is assigned based on multiple run-of-river hydropower criteria calculated in “power\_potential.m”.
* “power\_mean.\*”: Volumetric flowrate of water times the change in height between grid cells along the flow path times density of water times the gravitational constant. Calculated in “power\_potential.m”.
* “power\_density.\*”: Average power divided by the horizontal distance of the grid cell. Calculated in “power\_potential.m”.
* “power\_stability\_metric.\*”: Gridded power stability metric, as defined and described in Mosier et al. (2016). Calculated in “power\_potential.m”.
* “power\_standard\_deviation.\*”: Standard deviation of power over the simulation period. Calculated in “power\_potential.m”.
* “./grids/flow\_\*.\*” Gridded time-series of modeled streamflow. These are used for calculating run-of-river hydropower resource potential estimates.

**TEST CASES**

HPAT is distributed with a test case using a portion of the state of Oregon, USA. It is recommended that the user run both parts of the test case in order to familiarize themselves with the code; however, it is possible to skip directly to the simulation portion of the test case.

*CALIBRATION*

To calibrate HPAT for this test case, set:

* “region = ‘USGS \_14158790’”,
* “runType = ‘calibrate’”,
* “nGage = 1”
* “startDate = [2000, 10]” and
* “endDate = [2010, 9]”.

Run the model with the other default settings, which should be set to

* “monthsSpin = ‘12’”
* “timeStep = ‘monthly’”
* “dataRes = ‘monthly’”
* “printHydro = ‘asc’” and
* “output = {'flow', {'all'}}”

HPAT will create a unique output folder and prompt the user to select several inputs, including to select

* Digital Elevation Model: Select “./HPAT\_Oregon\_testcase/USGS\_14158790/terrain/ SRTM\_CGIAR41\_USGS14158790\_dem.asc”.
* Flow direction grid: Select “./HPAT\_Oregon\_testcase/USGS\_14158790/terrain/ SRTM\_CGIAR41\_USGS14158790\_fdr.asc”.
* Glacier presence grid or shapefile: Click “cancel” to indicate that glacier modeling will not be included in the run.
* Precipitation time-series: Select “./HPAT\_Oregon\_testcase/USGS\_14158790/climate/NetCDF/ delta\_his\_pchip\_pre\_USGS14158790”.
* Near-surface air temperature time-series: Select “./HPAT\_Oregon\_testcase/USGS\_14158790/climate/NetCDF/ delta\_his\_pchip\_tmp\_USGS14158790”.
* Observation data: Select “./HPAT\_Oregon\_testcase/USGS\_14158790/streamgage/USGS\_14158790.txt”
  + These data have been downloaded from the USGS website in an unaltered text file format. HPAT recognizes that they are USGS data by the “USGS\_” in the file name.
  + The user will be promted to enter the longitude (-122.046) and latitude (44.3314) of the streagage. If there were multiple observation files, the user would be prompted to select multiple in a row and enter any necessary information for each. HPAT will format all of the observation data into a single text file with the name “CCHF\_gage-\*.txt”.

HPAT will try to automatically detect if it is possible to perform the calibration in parallel. If so, it will utilize all available local processor cores. After each generation, the command console will display performance metrics for that generation and updates regarding run time. HPAT will then generate overall performance statistics once the model calibration run is complete. Note that HPAT sets an upper bound on the number of calibration generations, however, in all practical cases calibration terminates well before the upper bound (often calibration takes approximately 40-50 generations).

The output from a completed calibration run are distributed with the testcases in the folder “./HPAT\_Oregon\_testcase/USGS\_14158790/terrain/calibrate\_2000thru2010”. This folder contains all the files necessary to skip directly to a validation or simulation run.

*VALIDATION OR SIMULATION*

To validate HPAT for the single streamgage test case used in calibration, set:

* “region = ‘USGS \_14158790’”,
* “runType = ‘validate’”,
* “nGage = 1”
* “startDate = [1990, 10]” and
* “endDate = [2000, 9]”.

Then select the same the same input files outlined above in the calibration test case. The one additional file that you will be prompted to select is a set of model parameters produced during a calibration run. The path for this file is “./HPAT\_Oregon\_testcase/USGS\_14158790/terrain/calibrate\_2000thru2010/coefficients\_OR\_14158790\_test\_fittest.txt”.

The user can also run HPAT in simulation mode. The primary difference between validation and simulation is that no observation data are used in simulation mode. The test case folders include monthly climate time-series from 1980 – 2014. Thus, the dates can be set to any period within this range.

Note that inputs are provided that allow validation and simulation runs to be carried out for a broader Oregon, USA test case. The inputs for this test case are in the same formats as described above, but under the folder “./HPAT\_Oregon\_testcase/Oregon\_large\_test”. There are eight streamgages that can be selected for performing a validation run on this domain. In order to utilize all eight of these streamgages, set “nGage = 8”.

**USER INPUTS (under heading: %%USER INPUTS in HPAT script)**

region = 'USGS\_14158790';

%Select the mode to run the model in:

runType = 'simulate'; %Either 'default' (guess a parameter set),

%'calibrate' (for optimizing parameters),

%'calibrate\_resume' (use this if a calibration

%routine was interuptted)

%'validate' (apply optimized parameter set and

%compare to observation data)

%'simulate' (apply optimized parameter set but

%do not compare to observations)

nGage = 1; %Number of gage files to load (if using 'CCHF\_gage-\*.txt' compiled

%gagedate written in run, set nGage to 1);

%Select the time-series elements to run the mdoel for:

startDate = [1981, 10]; %Vector with format '[year, month]' specifying date

%to begin model run (run will include this date).

endDate = [2010, 9]; %Vector with format '[year, month]' specifying date

%to end model run (run will include this date).

monthsSpin = '12 months'; %String defining number of months to

%run model for prior to start date

%'startDate'.

timeStep = 'monthly'; %String specifying time resolution. Can be

%'daily', 'monthly', 'hourly'.

dataRes = 'monthly'; %String specifying resolution of input time-series

%data

printHydro = 'asc'; %Print run-of-river hydropower potential information in ESRI

%ASCII ('asc')

output = {'flow', {'all'}}; %surface flowrate through each grid cell

**REFERENCES**

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>

Mosier, T. M., Hill, D. F., & Sharp, K. V. (2014). 30-Arcsecond monthly climate surfaces with global land coverage. International Journal of Climatology, 34(7), 2175–2188. https://doi.org/10.1002/joc.3829

Mosier, T. M., Hill, D. F., & Sharp, K. V. (2016a). How much cryosphere model complexity is just right? Exploration using the conceptual cryosphere hydrology framework. The Cryosphere, 10(5), 2147.

Mosier, T. M., Sharp, K. V., & Hill, D. F. (2016b). The Hydropower Potential Assessment Tool (HPAT): Evaluation of run-of-river resource potential for any global land area and application to Falls Creek, Oregon, USA. Renewable Energy, 97, 492–503. https://doi.org/10.1016/j.renene.2016.06.002