**Brief introduction to HYSTAR**

**R and HYSTAR**

HYSTAR was developed in R (<http://www.r-project.org/>), which is a free software environment for statistical computing and graphics. Thus, a modeler needs to install R and the associated libraries and packages in a computer to run HYSTAR. The default installation of R should be enough to run the main model. However, model calibration and uncertainty analysis requires some libraries and packages including ‘LHS’. R libraries and packages would be found and installed through ‘Install package(s)’ in the ‘Packages’ menu of RGui. HYSTAR was developed with R version 2.11.x and tested in the Window operating system. Any higher version might be compatible with HYSTAR. The model source code has been tested in both 32 and 64 bit systems.

Once a modeler installed R in a computer, one will need to open the main script, HYSTAR.r, to run the model. All the required input data and files such as DEM, land cover and use map, and soil data should be in the same directory as that of HYSTAR.r. Yet, GUI for the model has not been developed so a modeler needs to manually change the names of the input files in HYSTAR.r, which includes all the input information for modeling. When you open the main script, HYSTAR.r, you will find a similar window in the screen to Fig. 1.

HYSTAR consists of several sub-modules for hydrology and sediment transport simulations including calibration and uncertainty analysis written by R script. Once a modeler (1) installed R and the required libraries and packages in the computer, (2) placed all the associated R scripts for sub-modules and the prepared input files in the predefined format under the same directory, and (3) opened RGui and the main script (HYSTAR.r), one will be ready to run the model. The sub-module scripts and format of the input files are described in the following sections.

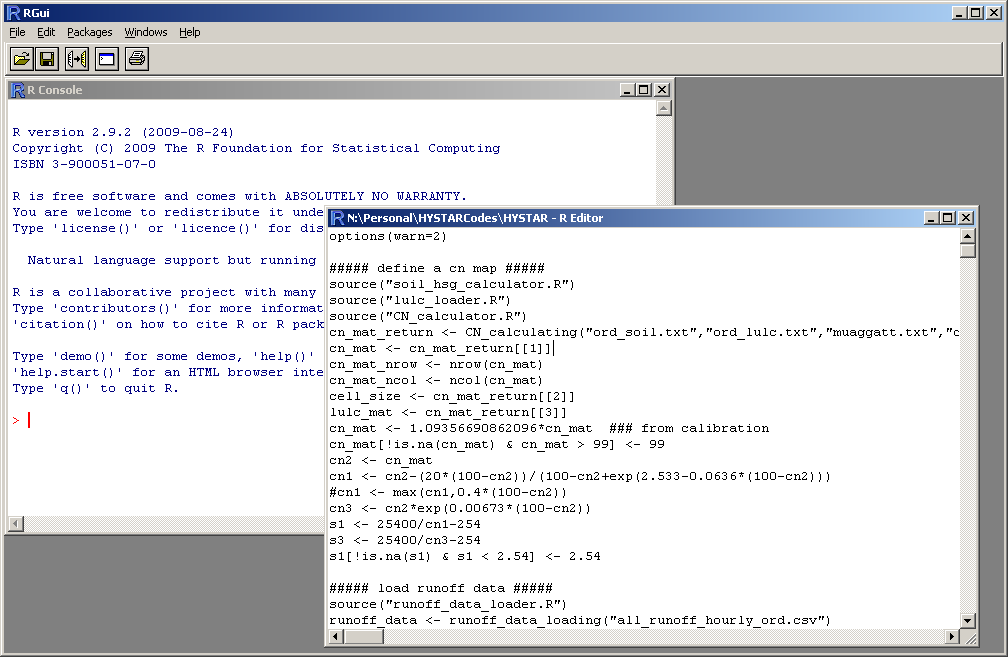


Fig. 1. The script, HYSTAR.r, in R.

**Model components**

HYSTAR.r contains names of all the input files and sub-modules, and sub-modules may recall other input files and/or sub-modules to load data, calculate parameters, or simulate variables. Brief descriptions of sub-modules and input data are provided in Table 1. Among the scripts, there are two ‘simulator’s, which are the main sub-modules for direct runoff and sediment transport simulation: ‘surface\_flow\_nonlinearCN\_simulator’ and ‘sediment\_tranport\_simulator’. In the direct runoff simulator, amount of excess rainfall, infiltrated, and routed water is calculated. Since ET is assumed occurring only between storm events in HYSTAR, ET and the associated soil moisture depletion are directly simulated in HYSTAR. On the other hand, soil water is hypothesized always percolated and immediately contribute to base flow so that percolation and the associated soil water depletion are directly simulated in HYSTAR and direct runoff simulator sequentially.

Table 1. Sub-modules and the associated input

|  |  |  |
| --- | --- | --- |
| Sub-module Name | Description | Associated Input |
| HYSTAR | Main model script | All the sub-module and input file names |
| rainfall\_loader | Load hourly rainfall time series data | Rainfall time series file name |
| rainfall\_acc\_calculator | Calculate accumulated hourly rainfall time series for every storm event | Hourly rainfall time series of a storm event |
| Thiessen\_loader | Load Thiessen grid layer that contains identification codes of the closest rain gage | Thiessen area division map file name |
| runoff\_data\_loader | Load hourly runoff time series data | Runoff time series file name |
| temperature\_loader | Load daily max. and min. temperature (Fahrenheit) | Temperature time series file name |
| extraterrestrial\_radiation\_hourly | Calculate hourly extraterrestrial radiation for potential evapotranspiration estimate | Latitude, longitude, LLTZ, simulation time (year, month, day, and hour) |
| potential\_ET\_calculator | Calculate hourly potential evapotranspiration (PET) using the Hargreave’s equation | All the information for ‘extraterrestrial\_radiation\_hourly’ and the max. and min. daily temperature time series |
| ET\_coefficient | Estimate dual crop coefficients | Lookup tables of (1) crop coefficient, (2) ground cover fraction, (3) pavement fraction and (4) plant growing stage, (5) land cover map file name, and (6) simulation time |
| actual\_ET\_hourly\_calculator | Calculate actual evapotranspiration (AET) using the dual crop coefficient method | All the information for ‘ET\_coefficient’, the estimated PET, simulated soil moisture map, temperature time series, soil wilting point map, land cover map, 4 soil feature maps (TEW, REW, TAW, RAW), and soil evaporation and transpiration history maps. |
| growing\_stage\_loader | Load growing stage information of plants | Predefined growing stage file name |
| elevation\_loader | Load elevation data | Predefined elevation layer name |
| slope\_loader | Load slope data | Predefined slope layer name |
| flow\_direction\_loader | Load flow direction data | Predefined flow direction layer name |
| stream\_network\_loader | Load stream network | Predefined stream network layer name |
| flowpath\_loader | Calculate flowpath connection information for efficient routing and write the information in files | Flow direction map |
| lulc\_loader | Load land cover and use layer | Land cover layer file name |
| soil\_hsg\_calculator | Get hydrologic soil group (HSG) of all the soil polygons from the SSURGO database | (1) File names of soil component (MUKEY: Map Unit KEY) of SSURGO (comp.txt) and (2) File name of ‘muaggatt’ of SSURGO (muaggatt.txt) |
| CN\_calculator  (contain lulc\_loader and soil\_hsg\_calculator) | Calculate curve numbers (CNs) for all the cells within a watershed of interest | All the input file names for ‘lulc\_loader’ and ‘soil\_hsg\_calculator’ and curve number (CN) lookup table file name |
| manningN\_calculator | Calculate Manning’s roughness coefficient for every land cover class | (1) Land cover map and (2) Manning roughness coefficient lookup table file name |
| soil\_loader | Load soil MUKEY map | File name of soil component (comp.txt) |
| soil\_depth\_loader | Load soil depth data for soil root depth estimation from SSURGO | (1) Soil MUKEY map, (2) file names of soil component (comp.txt), and (3) soil horizon (chorizon.txt) of SSURGO |
| soil\_awc\_loader | Calculate field capacity, wilting point, and saturated hydraulic conductivity of soil | (1) Soil MUKEY map, (2) file names of soil component (comp.txt), and (3) soil horizon (chorizon.txt) of SSURGO |
| soil\_porosity\_calculator | Calculate soil porosity | (1) Soil MUKEY map, (2) file names of soil component (comp.txt), (3) soil horizon (chorizon.txt), (4) soil texture group (chtextgrp.txt), and (5) soil texture (chtexture.txt) of SSURGO |
| effective\_porosity\_calculator | Calculate effective soil porosity | (1) Soil porosity map, (2) soil wilting point map, and (3) stream network map |
| initial\_soil\_moisture | Distribute initial soil moisture content based on the flow accumulation map over a watershed | (1) Flow accumulation map file name and (2) max. and min. initial soil moisture contents |
| effective\_flow\_vol\_acc\_calculator | Calculate accumulated effective flow volume based on the modified CN method | (1) Flow volume map and (2) the retention coefficient map of the CN method |
| traveltime\_calculator | Calculate travel time of flow on every overland cell using the Manning’s equation under the very shallow flow assumption | (1) Calculated effective flow volume map for overland, (2) stream network map, (3) slope map, (4), Manning’s roughness coefficient maps for overland, and (5) flowpath connection for efficient overland routing |
| st\_traveltime\_calculator | Calculate travel time of flow on every channel cell using the Manning’s equation | (1) Calculated effective flow volume map for channel, (2) stream network map, (3) slope map, (4), Manning’s roughness coefficient maps for channel, and (5) flowpath connection for efficient channel routing, and (6) the predefined stream depth and width maps |
| routing\_calculator | Route the calculate effective flow volume along the predefined flow paths | (1) Calculated travel time map, (2) calculate effective flow volume map, (3) flow direction map, (4) stream network map, (5) effective reservoir storage information, (6) transported sediment mass map, and (7) effective reservoir sediment trapping map |
| surface\_flow\_nonlinearCN\_simulator | Simulate surface flow, soil moisture, and percolation |  |
| soil\_particlesize\_loader | Load characteristic particle diameter for the Yalin equation | (1) Soil MUKEY map, (2) file names of soil component (comp.txt), and (3) soil horizon (chorizon.txt) |
| soil\_characteristic\_particlesize\_loader | Load characteristic particle sizes for the equations | Fraction of (1) fine sand, (2) very fine sand, (3) silt, and (4) clay |
| sediment\_transport\_capacity\_calculator | Calculate transport capacity of overland and channel flow | (1) Slope map, (2) Manning’s roughness coefficient map, (3) overland and channel flow velocity, width, and depth maps, (4) stream network map, (5) soil characteristic diameter, (6) soil particle settling velocity, (7) CSO and CSC |
| soil\_EUROSEM\_loader | Load the EUROSEM parameters from SSURGO | File names of (1) soil component (comp.txt), (2) soil horizon (chorizon.txt), (3) soil texture group (chtextgrp.txt), and (4) soil texture (chtexture.txt) of SSURGO |
| sediment\_tranport\_simulator | Simulate sediment transport |  |

HYSTAR requires 6 lookup tables: Curve number (CN), Manning’s roughness coefficient (Manning N), plant growing stage, crop coefficient, ground cover fraction, and impervious area fraction. There lookup tables should contain values of the corresponding parameters for all the land cover classes. An example is provided in Table 2. Plant growing stage, ground cover fraction, and impervious area fraction of each land cover should be approximated based on modeler’s logic, experience, and literature while lookup tables of CN, Manning N, and crop coefficient could be easily obtained from textbooks, reports, and/or manuals. Because the three lookup tables (plant growing stage, ground cover fraction, and impervious area fraction) control ET and subsequently soil water content, the approximation would be one of the biggest uncertainty and error sources of the model, in particular long-term simulation. All the lookup tables used for the test watershed, Owl Run Watershed, located in Northeastern Virginia, are provided in Center for Watershed Studies of Virginia Tech.

Table 2. Curve number lookup table

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Code | Description | A | B | C | D | Cover |
| 11 | Open Water | 99 | 99 | 99 | 99 | Water |
| 12 | Perennial Ice/Snow | 98 | 98 | 98 | 98 | Ice |
| 21 | Low-Intensity Residential | 54 | 70 | 80 | 85 | 1/2 acre |
| 22 | High-Intensity Residential | 77 | 85 | 90 | 92 | 1/8 acre |
| 23 | ommercial/Industrial/Transportation | 89 | 92 | 94 | 95 | Commercial |
| 31 | Bare Rock/Sand/Clay | 77 | 86 | 91 | 94 | Bare soil |
| 32 | Quarries/Strip Mines/Gravel Pits | 98 | 98 | 98 | 98 | Impervious |
| 33 | Transitional | 68 | 79 | 86 | 89 | Poor open scpae |
| 41 | Deciduous Forest | 30 | 55 | 70 | 77 | Good woods |
| 42 | Evergreen Forest | 45 | 66 | 77 | 83 | Poor woods |
| 43 | Mixed Forest | 36 | 60 | 73 | 79 | Fair woods |
| 51 | Shrubland | 35 | 56 | 70 | 77 | Fair brush |
| 61 | Orchards/Vineyards/Other | 43 | 65 | 76 | 82 | Fair orchard |
| 71 | Grasslands/Herbaceous | 49 | 69 | 79 | 84 | Fair grassland |
| 81 | Pasture/Hay | 30 | 58 | 71 | 78 | Meadow |
| 82 | Row Crops | 67 | 78 | 85 | 89 | Good straight row crops |
| 83 | Small Grains | 63 | 75 | 83 | 87 | Good small grain |
| 84 | Fallow | 76 | 85 | 90 | 93 | Poor fallow |
| 85 | Urban/Recreational Grasses | 39 | 61 | 74 | 80 | Good open space |
| 91 | Woody Wetlands | 98 | 98 | 98 | 98 | Wet soil |
| 92 | Emergent Herbaceous Wetlands | 96 | 96 | 96 | 96 | Wet soil |

In the sediment transport simulator, ‘sediment\_transport\_capacity\_calculator’ receives hydraulic characteristics of the simulated flow such as depth and velocity from the direct runoff simulator and then calculates sediment transport capacity of the flow. Because the sediment transport capacity concept HYSTAR employs hypothesizes that flow is always carrying sediment as much as its capacity, the calculated sediment transport capacity should be equal to sediment concentration of the corresponding flow in the model. ‘soil\_EUROSEM\_loader’ estimates values of the parameters adapted from EUROSEM such as soil detachability and cohesion ratios based on SSURGO. Then, the calculated amount of sediment loaded in the flow is routed along the stream network in ‘routing\_calculator’.

The model directly reads the prepared soil, land cover and use, elevation, slope, stream network, flow direction, and Thiessen polygon maps in the ASCII format. The ASCII files must contain headlines that show the numbers of columns and rows, cell size, nodata value, and geographic location of the map origin, and the headlines will be automatically created when the ASCII files are exported from a gird file format in ArcMap. The current version of HYSTAR assumes all the input maps have the same cellsize, map origin, and the numbers of columns and rows, thus all the maps should have the same headlines in the ASCII format. An example of the ASCII file is shown in Fig. 2.

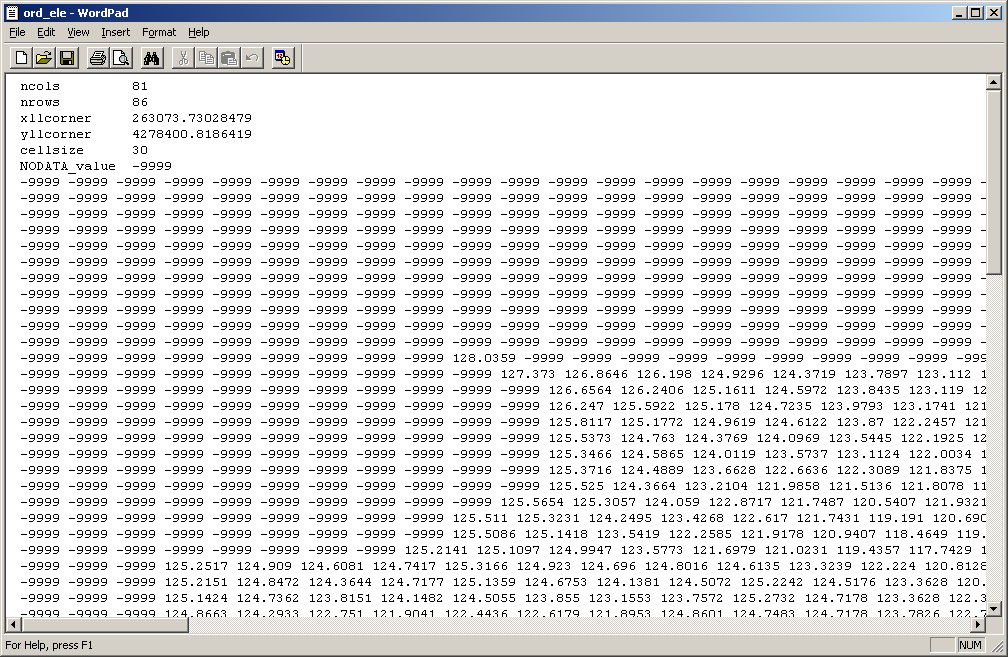


Fig. 2. Elevation map in ASCII file format, exported from ArcMap

HYSTAR requires hydrologic soil group (HSG), field capacity, wilting point, saturated hydraulic conductivity, porosity, residual soil water content, and soil (root) depth. In the current version of HYSTAR, the required soil information and data are derived from SSURGO automatically considering its database structure. However, when a modeler wants to use other soil database such as STATSGO and NATSGO than SSURGO, he/she will need to modify the associated codes with loading soil database. The required data need to be prepared in grid format (maybe in ASCII format for convenience) so that every cell within a watershed of interest should have its own soil information and data.

The model was designed to accommodate spatially distributed rainfall data in grid format such as NEXRAD. However, a modeler can use point data or converted areal data from point data using an area allocation map such as a Thiessen polygon when the distributed rainfall data are not available. The area allocation map can be created using any GIS software but should be prepared in the same format as the other GIS maps like elevation and soil. Then, ‘Thiessen\_loader‘ reads raingage station identification number for all the watershed cells. When multiple raingage data are not available, a modeler can use uniform rainfall option or Thiessen option with one raingage identification number.

The model does not create slope, flow direction, watershed boundary, and stream network maps from an elevation layer. Thus, they should be prepared in advance before running the model. In addition, a land cover map should be reclassified with unique identification numbers and then the associated look up tables should be built. In the model, stream width is an important variable for stream routing of flow and sediment, thus it is suggested to use field measurements than to estimate based on regression equations or assumptions. However, it is usually hard to get the width measurements along the stream network and to find an appropriate regression equation for a study watershed. Thus, the model provides an option to estimate the width along the stream network. When a modeler knows the maximum (maybe at the outlet of the watershed of interest) and the minimum (may be at the origin of the stream network) stream widths, the model assumes that the stream width is proportional to the drainage area or flow accumulation. Then, the stream width is continuously distributed between the maximum and minimum values along the predefined stream channel.