# **MERRA Spatial Downscaling for Hydrology (MSDH)**

# **Software User Guide**

The purpose of this document is to provide a quick guideline to generate 3-hourly grid surfaces of temperature, precipitation, relative humidity, wind speed, and shortwave and longwave radiation over a complex terrain watershed. The spatial downscaling approach produces high-resolution climate data from Modern Era Retrospective-Analysis for Research and Applications (MERRA) climate data products and a digital elevation model of the watershed or a target area. This guide is not intended to be comprehensive in documenting the detail physical processes involved in downscaling MERRA and RFE2 data. Rather it is intended as a brief introduction to guide a reader through the steps of installing software, downloading input data and running the processor, and obtaining some results.

# **Software Installation**

## Installing CDO

Obtain latest win32 version of CDO from <https://code.zmaw.de/projects/cdo/files>. Installing CDO is also very simple and similar to the installing netCDF DLL. To install CDO.exe, you just have to leave it in some directory in your computer, and tell your compiler in which directory to look for it. If you place the CDO.exe in Windows\System32 folder, you will be able to use this without further work. But if you put it in some other place, add the folder address that contains CDO.exe (Right click Computer→Properties→Advanced System Settings→Advanced→Environment variables).

## Installing GTK2+

1. Download and install GTK2+ from <http://sourceforge.net/projects/gtk-win/files/latest/download> using standard installation procedure.
2. The bin directory of GTK2+ installation must be set as path in environment variable.

## 1.3. Installing R and NCO

Download latest windows version of R from <http://www.r-project.org/>, and NCO from [http://nco.sourceforge.net/src](http://nco.sourceforge.net/src/nco-4.3.9.windows.mvs.exe). Install these software by standard installation procedure.

Please make sure to add bin directories of R and NCO installation folders as a PATH in your commuter’s environment variable.

# **Using the Software**

### **Running MERRA Spatial Downscaling for Hydrology (MSDH)**

MERRA Spatial Downscaling for Hydrology (MSDH) is available at https://bitbucket.org/AvirupSenGupta/msdh.usu. Please download and Unzipped the folder. The folder contains RunMSDH.exe and Run\_MSDH.R. To run RunMSDH.exe, we need to follow the procedure described below:

1. Make sure RunMSDH.exe and Run\_MSDH.R are under the same folder.
2. Click on RunMSDH.exe.

Rscript window (figure 2.1) should appear on your screen.

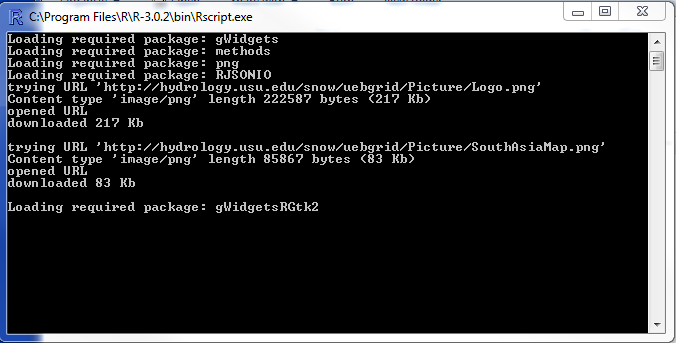


Figure 2.1. Rscript window to run R statistical and programming language.

Within few seconds, another window, the user interface for MERRA Spatial Downscaling for Hydrology (MSDH)(Figure 2.2) will appear. The Rscript.exe console may show several warnings like “Gtk-CRITICAL \*\*: gtk\_table\_attach: assertion ‘child -> parent == NULL’ failed”. You may also get few other warning Various R package related This is a widget related warning which can be ignored and we will continue working with the “MERRA Spatial Downscaling for Hydrology (MSDH)” window.

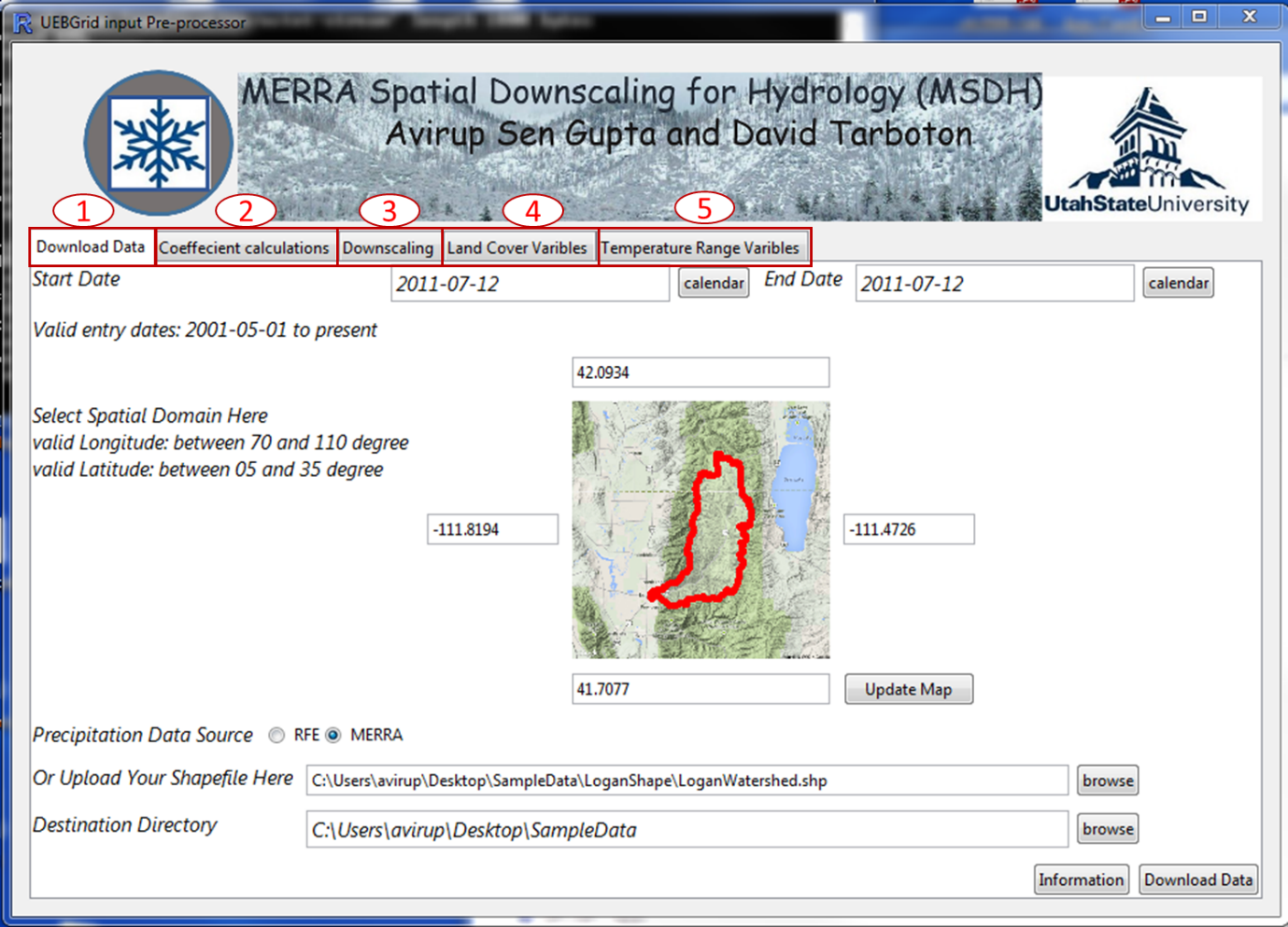


Figure 2.2. User interface for MERRA Spatial Downscaling for Hydrology (MSDH).

The window (figure 2.2) contains four different tabs (1) download data, (3) coefficient calculation, (3) downscaling, (4) vegetation variables and (5) temperature range variables.

The first tab “Download Data” is to access the MERRA and RFE22 on-line sources and automatically downloads the data based on a user-defined geographic extent and temporal range. The user will need an internet connection to successfully run this script.

As shown in Figure 2.2, to download the MERRA and RFE2 data for a range of dates and a spatial domain, the user needs to provide start and end date as well as latitude-longitude extent of the areal domain. To enter a date, a user need to click on the calendar button next to the textbox of start and end dates. The default calendar widget appears that allows users to select date visually. User also need to fill the Latitude North, Latitude South, Longitude East and Longitude West textboxes to assign a spatial domain. Instead of filling out these boxes, user may also upload a shapefile (as shown in Figure 2.2) of a target watershed or an area to allocate the spatial domain. It is worth to note that the shapefile must be in geographic coordinate system. At the end the user should indicate the directory where the downloaded data will be saved. This can be done using the file browsing widget corresponding to “Destination directory”. The precipitation can be downloaded from two sources (1) MERRA and (2) RFE2. RFE2 precipitation is only available for South Asian domain. On the other hand, MERRA is a global dataset. Based on the needs, experience, accuracy or watershed location, the user needs to select either of these two precipitation sources. Once all these information are provided, “download data” button at the right-bottom corner of the window needs to be clicked. If the destination directory do not contain three subfolder called MERRAcli, MERRArad and RFE2, they will be created. If these three subdirectories already exists, they will not be overwritten. Instead downloaded data will be stored in this folder.

The downloaded files will appear in the following format:

*MERRAcli = merra.prod.assim.YYYYMMDD\_dwd.nc (NetCDF);*

*MERRArad = MERRA300.prod.assim.tavg3\_2d\_chm\_Fx.YYYYMMDD.SUB\_long.nc (NetCDF);*

*RFE2 = cpc\_RFE2\_v2.0\_sa\_dly.bin.20010501.gz (zipped binary).*

For this example, we will download MERRA for

### **Preprocess and Downscale the MERRA and RFE data**

Most hydrologic and ecological models run on a much finer spatial grid than MERRA and RFE2 data for better capturing the local hydro-climatic condition. The choice of the resolution of a DEM is left to the users depending of the nature of the problem, source of the DEM, computer disk space availability, resource constraint, and usage of the data. The downscaling methods to interpolate and adjust, based on topography, the coarse scale MERRA and RFE2 data to the spatial scale of the DEM. The temporal resolution of the downscaled data is always 3 hours. Our choices of temporal resolution was largely influenced by the need of the forcing variables in a typical physically based energy model to compute energy content at the earth’s surface. To capture the diurnal pattern of a variable, it is preferable to obtain data at a time resolution of 6 hours or less. In this application, we have chosen 3-hours as the time interval that, to our best judgment, would maximize the information but would not increase the data volume to an unmanageable level for a desktop application.

The second tab “Downscaling” spatially downscales the MERRA and RFE data to DEM resolution and converts the temporal resolution to every 3-hours based on a user-defined DEM and temporal range. For Lantang Khola watershed, a DEM is provided in <http://hydrology.usu.edu/snow/uebgrid/DEM.zip>. Unzipped folder contains, *srtm\_elevation\_LangtangKhola\_LambertAzimuthalEqualArea (\*.prj,\*.tfw,\*.tif,\*.tif.aux.xml,\*.tif.vat.dbf)*: ~98-meter Digital Elevation Model. The DEM that will be used throughout the project must be prepared prior to the preprocessing and downscaling procedure so that all of the input layers into UEB have consistent extents and resolutions. The numbers of rows and columns MUST match across datasets for UEBGrid to run properly. This may require using BASINS before running the R scripts in order to prepare the DEM by clipping it for example.

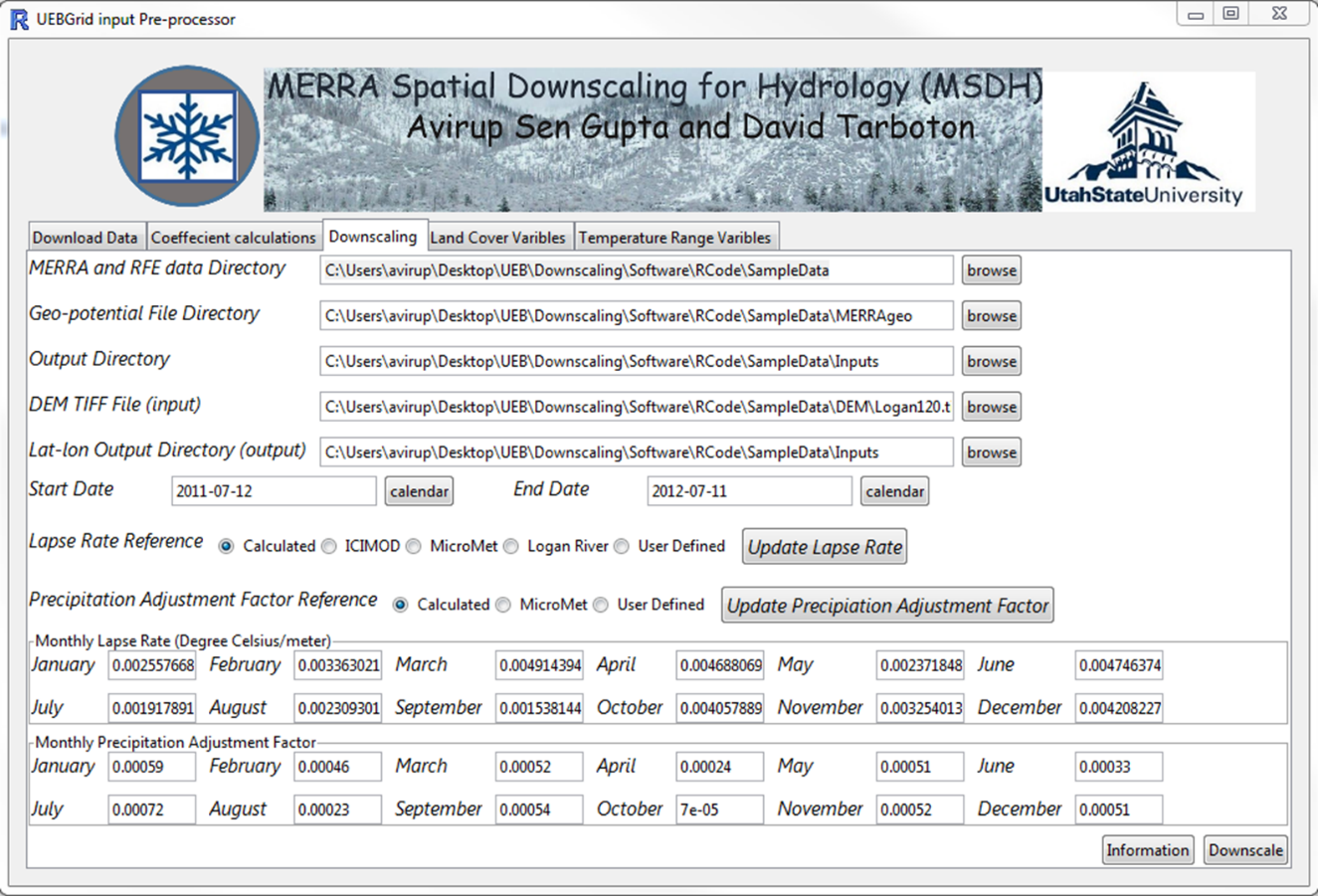


Figure 2.3. Downscaling interface of MERRA Spatial Downscaling for Hydrology (MSDH).

Figure 2.3 is the interface for downscaling MERRA and RFE2 data. The user will need to set directories, including the top directory where the downloaded data (i.e. where the *MERRAcli*, *MERRArad*, and *RFE* directories) resides, the *MERRAgeo* (geopotential heights) data are saved, the destination directory for the downscaled data, the DEM, and a output text file which lists list the center lat/lon values for the DEM. By Default, the program assumes that the “Destination Directory” in the “data Download” tab is the top most directory for this project and *MERRAcli*, *MERRArad*, *RFE* and *MERRAgeo* are the subdirectories. Based on this assumption, the program sets all the input directories. It also assumes a folder “UEBinput” exists in the UEB project directory and sets it as output directory to store downscaled netCDF files and lat-lon text file. Program also assumes downscaling temporal range would be same as download temporal range and sets the start and end date as previous tab. But the program also allow a user to change the start and end dates if s/he wishes to. Monthly lapse rate is the linear temperature lapse rate for temperature adjustment due to elevation difference and visible to the user. There are three default sets of lapse rates provided by the software:

1. Calculated: Use the coefficients calculated in the previous step “Coefficient calculations”
2. ICIMOD: [International Centre for Integrated Mountain Development](http://en.wikipedia.org/wiki/International_Centre_for_Integrated_Mountain_Development)(ICIMOD) installed experimental stations in Langtang Khola watershed and obtained a set of monthly lapse rate for that watershed.
3. MicroMet: Global Monthly lapse rate provided by ([Liston and Elder, 2006](#_ENREF_1)) in MicroMet downscaling approach. If a user is unaware of lapse rate information in his/her study area, this lapse is recommend.
4. Logan River: By studying the temperature data from SNOTEL station within Logan River area, Authors have derived this lapse rate.

If a user wish to use his/her own lapse rate, “User Defined” should be selected. After selecting any of these four options, “Update Lapse Rate” option should be selected to update the values. If user selects “User Defined”, monthly lapse rate values disappear and user needs to input those values.

Similarly, user also have an option to adjust the monthly using a relationship precipitation-elevation (equation 1)

(1)

Here, is precipitation at target DEM after downscaling (elevation adjustment)

is MERRA precipitation at DEM resolution before downscaling (elevation adjustment)

is elevation provided in DEM,

is the bilinearly interpolated MERRA elevation at DEM resolution

is the precipitation adjustment factor

Similarly, Users also have an option to choose to provide the precipitation adjustment factor to the software. If the user does not have any precipitation data within or near the target area, they are suggested to use “MicroMet” option.

Once all these information are entered and “Downscale” button is clicked, downscaling of MERRA and RFE2 will start. The resulting files will combine both MERRA and RFE2 into one NetCDF file and appear in the following format: merra.rfe.90m.YYYYMM.nc (each month will result a single netCDF file). These files will be saved in the directory as defined the user above.

### **Derive the Land Cover Characteristics from Land Cover Type**

The third tab “Vegetation Variables” creates leaf area index, canopy height, % canopy cover, and canopy structure based on the MODIS Land Cover Type data.

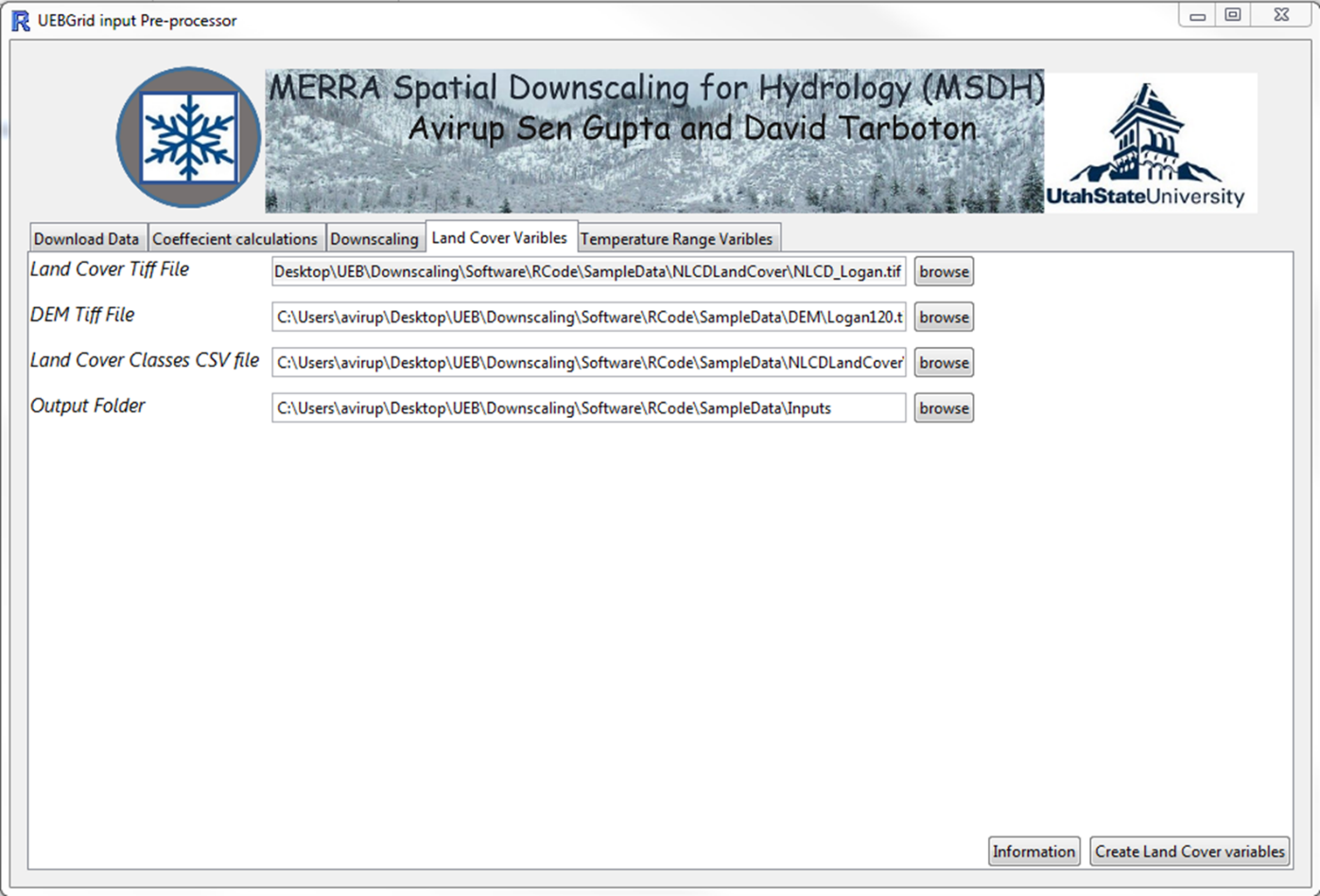


Figure 2.4. Interface to create land cover related variables.

If a user follows default folder structure shown in section 2.2, all the files including the land cover file and lookup table, the DEM, and the output directory will already be set by the program as shown in Figure 2.5. If a user alternative land cover data and/or own lookup table (provided to reflect land cover classes), s/he needs to provide location of those files in “Land Cover Tiff File” and “Land Cover Classes CSV file” textboxes, respectively. Conversion of land cover type to the land cover derivatives is started by Clicking “Create vegetation variables” button. The resulting files will appear as ccgrid.nc, hcangrid.nc, LAIgrid.nc, and ycagegrid.nc.

# **Worked out example with sample data**

### **Sample Data**

Folders:

DEM (inputs to the software): Digital elevation Model for Logan River Watershed

Inputs (holds the outputs. Should be created by the user): Contains downscaled outputs. The netCDF file inside is folder is the final output of this

LoganShape (inputs to the software): Contains shapefile for Logan River watershed in geographic coordinate system.

MERRAcli (output at data download process): Contains MERRA temperature, pressure, dew point temperature, and wind speed data. Each netCDF file contains hourly values for each day.

MERRAgeo: Downloaded MERRA global surface geo-potential data.

*merra\_global\_chm.nc:* chemistry forcing surface geo-potential

*merra\_global\_slv.nc:* single level surface geo-potential

MERRAPrecip (output at data download process): Contains series of netCDF files where each file stores MERRA daily total precipitation. UEBinput: An empty folder to stored downscaled data.

MERRArad (output at data download process): contains MERRA shortwave and longwave radiation data. Each netCDF file contains 3- hourly values for each day.

NLCDLandCover (inputs to the software): Contains land cover information to produce netCDF files for Leaf Rea Index (LAI), canopy height (hcan), canopy coverage (cc) and canopy structure (ycage) at DEM resolution.

* *landcover.csv*: lookup table for mapping NLCD land cover classes onto vegetation parameters (cc, hcan, lai, and ycage).
* *NLCD\_Logan.tif*: 30-meter land cover type dataset in Albers Conical Equal Area representing with Values from 0 to 255 each value representing a specific land cover class.

SNOTELData (inputs to the software): Contains SNOTEL station information and temperature and precipitation data at each of the given SNOTEL stations.

All the SNOTEL station and data should be stored in csv files. The columns of the station table must follow this structure (1) Station\_ID, (2) Elevation, (3) Latitude, (4) longitude and (5) Station Name. These column headings must be a single word (for example, station and ID can be written as “Station\_ID”).

csv file with temperatures data should just be named with their stationID. For example, if the station ID is 823, then the csv file containing the temperature data must be named as “823.csv”. Similarly, csv file containing precipitation data should start with “Prec” followed by the station ID. For example, precipitation csv file for station 823 should be named as “Prec823.csv”. In both cases, files should contain only 2 columns, first column representing data in “MM/DD/YYYY” format and second column representing temperature or precipitation data. Temperature data should be in ˚C and precipitation should be in inches unit.

# **Supplied Data and Folder structure for South Asian Region**

The following folders (bold) and files (Italic) have been provided with the HIMALA-BASINS system from <http://hydrology.usu.edu/snow/uebgrid/SAsiaBaseData.zip>. If we unzip the downloaded folder as SAsiaBaseData, we find seven subfolders, named as: MERRAcli, MERRAgeo, MERRArad, MERRAtrange, MODIS, RFE and UEBinput. In the manual, we assume this folder was unzipped in a folder called “UEBERRARFE” in C drive. Therefore, C:\UEBERRARFE\SAsiaBaseData is the directory location of the unzipped folder.

SAsiaBaseData: Contains Data for South Asia domain (North: 5-35˚, east: 70-110˚). Unless otherwise noted, the data are in the Lambert Azimuthal Equal Area projection with datum = WGS84, latitude of center = 45, longitude of center = 100). Subfolders of SAsiaBaseData are described below.

MERRAcli: Contains MERRA temperature, pressure, dew point temperature, and wind speed data. Each netCDF file contains hourly values for each day.

MERRAgeo: Downloaded MERRA global surface geo-potential data.

*merra\_global\_chm.nc*: chemistry forcing surface geo-potential

*merra\_global\_slv.nc*: single level surface geo-potential

MERRArad: contains MERRA shortwave and longwave radiation data. Each netCDF file contains 3- hourly values for each day.

MERRATrange: contains a single netCDF file (i.e. *MERRAAvgMonthlyTrange.nc*) that stores MERRA monthly diurnal temperature ranges for South Asia domain.

MODIS (MODIS MOD12Q1 IGBP ~ 500-meter Land Cover Type v051 for South Asia)

*landover.csv*: lookup table for mapping MODIS classes onto UEB parameters (cc, hcan, lai, and ycage) used in land cover.

*MOD12Q1\_LandCoverType\_IGBP\_SouthAsia\_Geographic.tif*: ~500-meter land cover type dataset in Geographic WGS84 representing 17 classes for South Asia- see p. 54.)

*MOD12Q1\_LandCoverType\_IGBP\_SouthAsia\_LambertAzimuthalEqualArea (\*.prj, \*.tfw,\*.tif,\*.tif.aux.xml,\*.tif.vat.dbf)*: ~500-meter land cover type dataset representing 17 classes for South Asia- see p. 54.)

*MOD12Q1\_SouthAsia\_Geo\_Clip (\*.tfw,\*.tif,\*.tif.aux.xml,\*.tif.vat.dbf)*: a subset of the above layer in Geographic Coordinate System

RFE2: Contains series of netCDF files where each file stores RFE2 daily total precipitation.

UEBinput: An empty folder to stored downscaled data.

### **Derive the Mean Monthly Temperature Range**

If a user follows default folder structure shown in section 2.2, South Asian temperature range netCDF file, the DEM, and the output directory will already be set by the program. When the user runs the program by clicking “Create climatology Files”, 12 new netCDF files will be created in the output directory. Thee netCDF files are named as merra.trange.climatology.MM.nc, where MM indicates the month of the year.

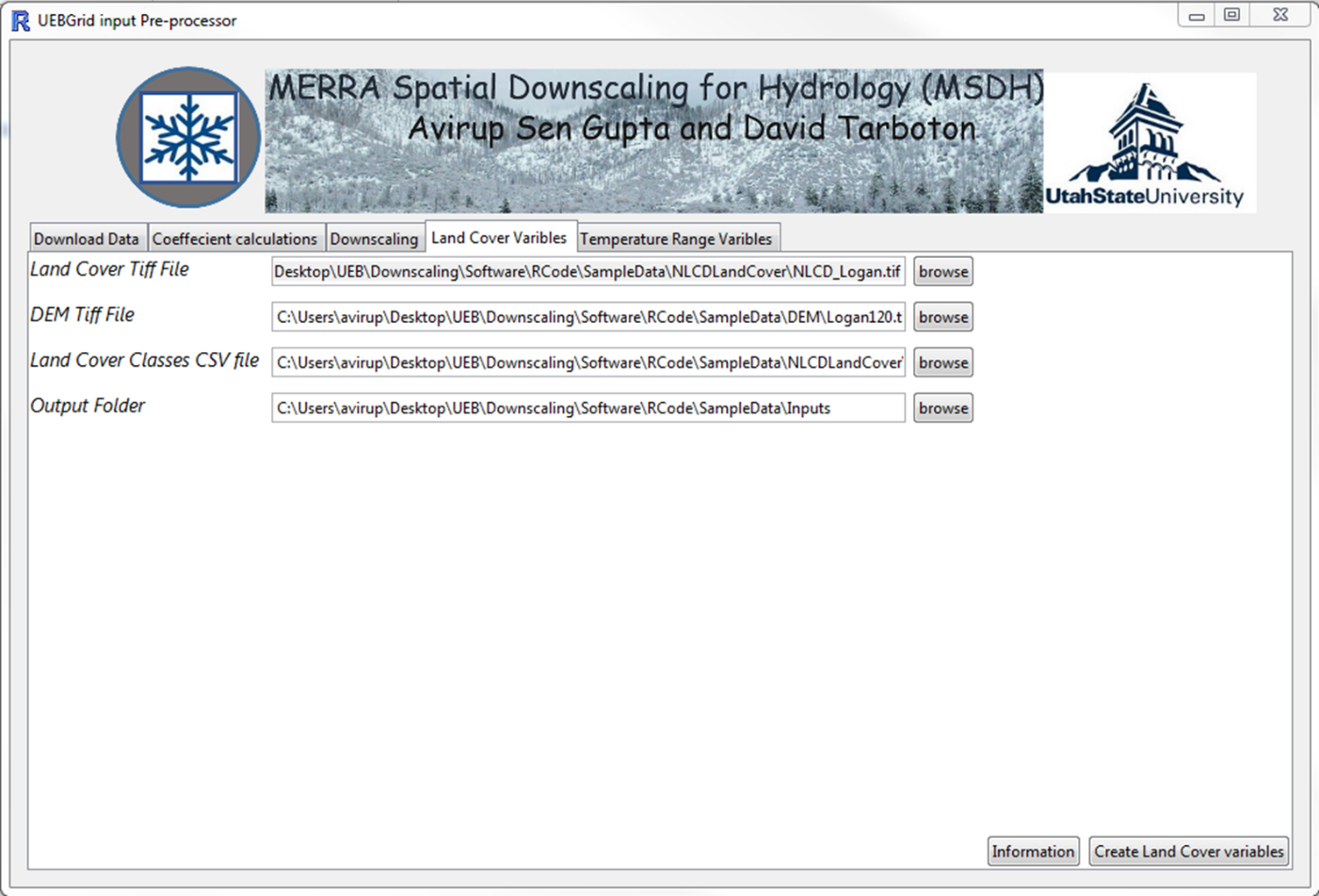


Figure 2.5. Interface to create monthly mean temperature range netCDF files.

The fourth tab “Temperature Range variables” calculates monthly mean temperature range for the given watershed. The output is necessary when using the shortwave and longwave radiation option in UEBGrid.

# **Appendix A: Data Sources**

MERRA (<http://gmao.gsfc.nasa.gov/merra>): Time varying forcing data such as: temperature, relative humidity, wind speed, long and shortwave radiation are derived from MERRA. MERRA is a recent near-real time global climate reanalysis product developed at NASA derived from the Goddard Earth Observing System version 5 (GEOS-5), NASA general circulation model ([Rienecker et al., 2011](#_ENREF_3); [Suarez et al., 2008](#_ENREF_4)) and National Centers for Environmental Prediction (NCEP) Gridpoint Statistical Interpolation (GSI) analysis ([Wu et al., 2002](#_ENREF_5)). It is available from 1979 to present ([Lucchesi, 2012](#_ENREF_2)). Temperature, wind speed, relative humidity are available at spatial resolution of 2/3˚ longitude by 1/2˚ latitude and incoming shortwave and longwave radiation are available at coarser resolution of 1.0˚ by 1.25˚ ([Lucchesi, 2012](#_ENREF_2)). Temperature, wind speed, relative humidity are reported at hourly time resolution at height of 2 m from the ground and incoming shortwave and longwave radiation at the surface are reported at 3-hourly time step.

RFE2 (<http://www.cpc.ncep.noaa.gov/products/fews/index.html>): Precipoiation is derived from RFE2 data products. Rainfall Estimation (RFE2), daily total precipitation estimates are constructed using four observational input data sources, namely: approximately 280 GTS stations, geostationary infrared cloud top temperature fields, polar orbiting satellite precipitation estimate data from SSM/I, and AMSU-B microwave sensors ([Xie et al., 2002](file:///C:\Users\avirup\Desktop\UEB\Paper2\SciencePaperDraft.docx#_ENREF_12)). Near-real time daily rainfall estimations are available for the Southern Asian domain (70˚-110˚ East, 5˚-35˚ North) at a spatial resolution of 0.1˚ by 0.1˚ beginning on May 01, 2001.

SRTM (<http://www2.jpl.nasa.gov/srtm/>): Source of Digital Elevation Model (DEM). SRTM (Shuttle Radar Topography Mission) digital elevation data, originally produced by NASA, represents the most complete high-resolution digital topographic database of Earth on a near-global scale. SRTM consisted of a specially modified radar system that flew onboard the Space Shuttle Endeavour during an 11-day mission in February of 2000. SRTM data are available at no cost and can be automatically downloaded from several different sources. The 90-meter data (srtm\_54\_07) used in this manual were downloaded from CGIAR-CSI GeoPortal (<http://srtm.csi.cgiar.org/>) and are void-filled. Only the area representing Langtang Khola watershed is included and can be downloaded from <http://hydrology.usu.edu/snow/uebgrid/DEM.zip>. Other study areas will need to be downloaded by the user.

National Land Cover Database (<http://www.mrlc.gov/finddata.php>)

MODIS (<https://lpdaac.usgs.gov/products/modis_products_table/mcd12q1>): source of land cover, including leaf area index, canopy cover, canopy height, and canopy structure derivatives. The MODIS Terra + Aqua Land Cover Type Yearly L3 Global 500 m SIN Grid product (MCD12Q1) provides data characterizing five global land cover Classification systems, which describe land cover properties derived from observations spanning a year’s input of MODIS (Terra and Aqua) data.  The primary land cover scheme identifies 17 land cover classes defined by the International Geosphere Biosphere Programme (IGBP), which includes 11 natural vegetation classes, 3 developed and mosaicked land classes, and three non-vegetated land classes. V051 products are produced with revised training data and certain algorithm refinements. Downscaling will create 90-meter derivatives.

# **Reference**

Liston, G. E. and K. Elder, (2006), "A meteorological distribution system for high-resolution terrestrial modeling (MicroMet)," Journal of Hydrometeorology, 7(2): 217-234.

Lucchesi, R., (2012), "File Specification for MERRA Products. GMAO Office Note No. 1 (Version 2.3)," 82 pp, <http://gmao.gsfc.nasa.gov/pubs/office_notes>.

Rienecker, M. M., M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs and G.-K. Kim, (2011), "MERRA: NASA's modern-era retrospective analysis for research and applications," Journal of Climate, 24(14): 3624-3648.

Suarez, M. J., M. Rienecker, R. Todling, J. Bacmeister, L. Takacs, H. Liu, W. Gu, M. Sienkiewicz, R. Koster and R. Gelaro, (2008), "The GEOS-5 Data Assimilation System-Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0."

Wu, W.-S., R. J. Purser and D. F. Parrish, (2002), "Three-dimensional variational analysis with spatially inhomogeneous covariances," Monthly Weather Review, 130(12): 2905-2916.