**Lab 8: Preparing geospatial data for watershed modeling (R coding)**

**Background information on watershed modelling**

A variety of hydrologic and water-quality models have been used to assess the impacts of land-use changes. For example, Weller and others ([2003](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR48)) developed an empirical linear model to predict water quality using the proportion of cropland and developed land as independent variables. Bhaduri and others ([2000](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR5)) integrated GIS with an NPS pollution model to assess the long-term runoff and NPS pollution. Tang and others ([2005](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR41)) implemented the Long-Term Hydrologic Impact Assessment model (Harbor [1994](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR24)) to estimate the impacts of land-use changes on surface runoff and NPS pollution. Empirical water-quality models have advantages in data preparation because the input data are readily available, and the model can be routinely used for operational applications.

Recently, distributed watershed models have been increasingly used to assess hydrologic responses to different land-cover changes. The process-based watershed models can be useful for improving the understanding of interactions between land-use change, water balance, and water-quality issues. The SWAT model is a physically based, continuous time-step model (Arnold and others [1998](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR4); Neitsch and others [2002](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR32)). The model was developed by the USDA Agricultural Research Service (ARS) to assess the impact of agricultural management practices on water balance, sediment, and nutrient loadings for nongauged watersheds. The SWAT model has been widely used in both the United States and international sites. Borah and Bera ([2004](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR6)) provided an overview of SWAT model applications for 17 case studies. Most SWAT models were calibrated and validated at monthly intervals. Good results were achieved for both small [Warner Creek (3.46 km2)] and large watersheds [Upper Mississippi River Basin (491,700 km2)]. The daily estimations from SWAT model are generally considered less accurate compared with monthly estimations (Borah and Bera [2004](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR6)). A thorough review of the SWAT model is also provided by Gassman and others ([2007](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR20)).

One of the main drawbacks for the application of SWAT model is its significant data requirements. Primary SWAT model input data include a land-cover, digital-elevation model (DEM), soil map, daily precipitation and temperature, and detailed agricultural management information (land use). The SWAT model’s calibration and validation procedures require aditional data sets, such as stream flow, sediment, and nutrient loadings. Due to limited data availability, thorough model calibration and validation were not possible for many applications (Stonefelt and others [2000](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR40)). However, recent advances in remote sensing and GIS have resulted in improved data availability, and continuous software development (i.e., ArcSWAT) has made the SWAT model toolbox more user-friendly (Neitsch and others [2002](https://link.springer.com/article/10.1007/s00267-012-9903-9#CR32)). As a result, it is expected that the SWAT model will be widely used for future watershed assessments, particularly those linking land cover and water quality.

**Objective**

The main task of this lab is to develop an R script that can be used to extract and process geospatial data for any user-specified watershed. All resultant geospatial data should be in SWAT file format (i.e., Analysis Ready Data).

**Study Area:**

****

The study region includes six 8-digit watersheds in VA (Fig. 1).

**Data**

DEM (lab8/inputdata/dem.tif)

NLCD (lab8/inputdata/nlcd2001.tif)

Watershed shapefile (lab7/inputdata/watersheds.shp)

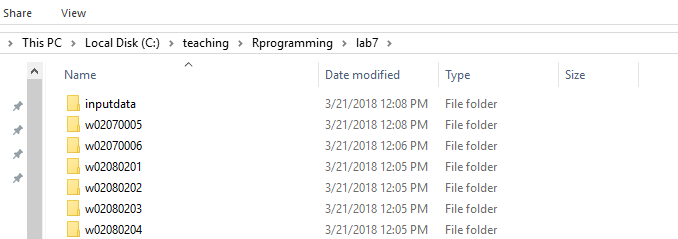
Daily Precipitation (2021 only): please download daily precipitation data from PRISM website: <http://prism.oregonstate.edu/recent/>)

Daily temperature (minimum and maximum, 2021 only): please download daily data from PRISM website: <http://prism.oregonstate.edu/recent/>)

PRSIM website has 1981-current data (probably a few hundred gigs to a terabyte). We will use 2021 data only for this lab.

**Data processing**

1, Use R script to create 6 output folders in your lab8 folder. Each watershed has a unique huc8 ID (e.g., 02070005, 02070006,…). So your output folders could be named using the huc8 codes (see the following example).



The following R function creates a new folder:

dir.create("C:/Rprog/lab8/w02070005")

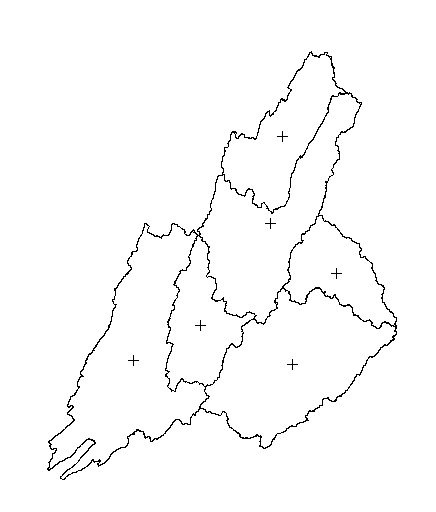
Please use a ‘for loop’ to create 6 output folders. Specific watershed id (i.e., HUC8) can be listed from the watersheds shapefile in the inputdata folder. You may use readOGR() and paste() functions to generate output folder names.



2. Read dem and nlcd raster layers from inputdata folder

3. Clip DEM and NLCD data using each of six watershed boundaries. Name the output dem as demsub. Name the output nlcd as nlcdsub. Export demsub and nlcdsub as a GTIFF files to output folders. Each output folder has it’s own demsub.tif and nlcdsub.tif files.

4. Before you extract daily precipitation values for each watershed, you need to identify the polygon centroid in R.



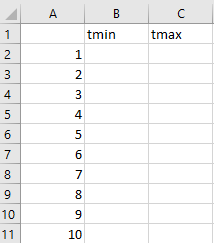
library(rgeos)

centroids = gCentroid(watershed\_p,byid=TRUE)

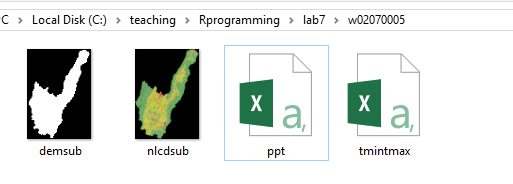
5. Extract daily precipitation value (2021) for each polygon centroid and export the result as a csv file. See the following R command as an example:

write.csv(ppt,'C:/Rprog/lab8/w02070005/ppt.csv')

6. Extract daily temperature (2021). Please include both tmax and tmin values in your csv file. The format of the csv file should look like:



7. Check your final output. For each watershed folder, you should have the following files:



**Deliverables**

**Upload your R script to Canvas site**