Assignment #2 – Bash exercise

Due – 1pm, Sep 18, 2024

In this assignment, our goal is to 1) organize the input data, 2) prepare the configuration files, so that we can run a hydrologic model smoothly. Lastly, we will check the water balance of a Pseudo model output! (Disclaimer: I made up the data, so it is not scientifically valid!) We will use 10 Ameriflux sites in this assignment.

All data used in this assignment are in Hydroshare (https://www.hydroshare.org/resource/7034272994ee4a6bbf6d6df534c30f67/). The file is under folder *assignment/assignment_2*, with the name ***gly606_hw2.tar.gz***, you may use "**tar -xvzf gly606_hw2.tar.gz**" to extract all files.

Please feel free to download the assignment to whatever platform that you feel comfortable with. In this assignment, you will be asked to write **3** bash scripts.

1. File organization – We have all the input cramped in the same folder (*gly606_hw2*), which is not organized nor ideal for set up model run. Therefore, I would like you to write a script and organize the files by creating ONE bash script to create folders and move files around. (Ameriflux site names are stored in file ***site_list.txt***)
    a. Create a folder named *model_run*, which will be the root folder that everything goes into it (the folder *model_run* should on the same directory as *gly606_hw2*)
    b. Under *model_run*, create one folder for each site, using Ameriflux site name as the folder name.
    c. Under each site folder, create three folders, i.e., *forcing*, *init*, *static*, to host meteorological forcing data, initial condition files, and static files.
    d. Move files in the homework directory to corresponding folders we just created (Following are some hints)
        i. Please remember to revisit the naming convention for the input data in *gly606_hw2*.
        ii. Each Ameriflux site has 5-days of meteorological forcing. The forcing data are stored in daily files, starting from Jan 1st and ending on Jan 5th.
        iii. Each Ameriflux site are running for different years. The information about the corresponding year for each site is stored in **site_list.txt**.
        iv. The initial condition files are one hour before the beginning of the model run. For example, if a model run starts from 2016-01-01, the timing for the initial condition is 2015-12-31_23-00-00. You should be able to figure out the pattern of file names for initial condition files.

2. Prepare configuration files – We have successfully run a test case for site "US-SRG" (not part of this homework)! We plan to use its configuration file as an example and run it for the 10 Ameriflux cases. I would like you to write ONE bash script to finish the following tasks. (Similar to what we did in class)
    a. Prepare a template file for the configuration
        i. Find ALL site-specific information about "US-SRG" configuration file with filename "**ufs-land.namelist.US-SRG"**, including site names, input file directories, date/time for initial condition files, start date, number of run days, etc. Replace the site-specific information using unique place holders. (hint: using "sed")
    b. Create one configuration file for each Ameriflux site by replacing the unique place holders using site-specific information. The site-specific information should be consistent with Part 1. Please note that the configuration file for each site should be located in *model_run/<site_name>*. Remember to make sure that configurations are pointing to the correct folders.
3. Check water balance for the output – This part of the home is slightly disconnected from Parts 1 and 2. For each Ameriflux site, we have summarized one-year observation data about annual total precipitation, annual total runoff, annual total evaporation, and the soil column water storage change into one file, unit is millimeters (mm). The filename "**ameriflux_water_flux.txt**". It contains five columns indicating site names, precipitation, runoff, evaporation, and water storage change respectively.
    a. Write a bash script to check whether the water balance is met for each site. Print 1) site names with water balance errors and 2) the errors.
        i. Syntax for calculations of floating numbers can be referred to the "if_statement.sh" file in "inclass_practice/bash_practice/" folder on our Hydroshare folder.

Please submit the three bash scripts to Brightspace.