ED simulation and analysis instructions for work done by Logan Arnold with Dr. Medvigy in Spring 2019

GitHub Repository: <https://github.com/Larnold1997/ED2-1/tree/logan2>

Email: larnold1997@hotmail.com

Important pieces of information are bold; names of programs/directories and terminal commands are italicized; my personal notes are underlined

**Step 1: Simulations**

* **To compile ED2**: *./install.sh -p crc -k E -g* in build directory
  + Files may have the incorrect path – fix to represent what it should be for your local version
  + To “make clean” (i.e. delete the executable): *./install.sh -p crc -c –g* in build directory
* **To run ED2**: go to the run directory and type *qsub submit.sh*
  + Will receive a submit.sh.o(numbers) file with information about the execution
    - May have to wait until simulation is complete to view file
  + **Three types of submit.sh – based on the ED2IN file** 
    - All three end in a line that indicates which ED2IN to use: ./../build/ed\_2.1-opt -f *the ED2IN file*
    - All of the ED2IN files are set to run simulations from July 7, 2007 to June 30, 2037.
      * NOTE: My analysis and pre-processing scripts are often VERY dependent on these start/end dates, though it should be easy to change them if needed. My analysis scripts also treated 2008 as “year 1” since it was the first full year simulated
    - *submit.sh*: uses ED2IN\_PV\_site1\_FALL2018, which simply puts monthly files in a local folder entitled “output”
      * NOTE: be sure to change where the files are going and/or move the files to a new folder once the simulation is complete to avoid overwriting data!
    - *submit\_fast.sh*: uses ED2IN\_PV\_site1\_FALL2018\_fast, which puts monthly AND “fast” (hourly) output into a folder in your scratch365 directory. This change is necessary because a simulation that produces fast output produces A LOT of data (more than 50 GBs, for the length of simulations that I used)
      * NOTE: Will probably have to contact the CRC for them to create your scratch365 directory (if you do not already have one)
    - *submit\_regional.sh*: Uses ED2IN\_PV\_regional
      * NOTE: I never ended up exploring regional simulations, so this may or may not work, and I have no analysis scripts for the results
  + **Information in ED2IN dictates run.** Some of the variables I most often changed (see the document itself for variable descriptions):
    - ED2 File output
    - FFILOUT and SFILOUT
    - METCYC1 and METCYCF
    - NSLCON
    - **Changing PFTS**: I typically considered 3 PFTs, one at a time (thus all my simulations were really a set of three simulations for complete results). Make the following changes in the ED2IN file for the different PFTs:
      * C4 Grass:
        + INCLUDE\_THESE\_PFT = 1
        + NL%IEDCNFGF   = 'PV\_init/PV\_PFT01\_revised.xml'
        + NL%IPHYSIOL = 2
        + NL%H2O\_PLANT\_LIM = 3
        + NL%ISTOMATA\_SCHEME = 0
      * Acquisitive
        + INCLUDE\_THESE\_PFT = 2
        + NL%IEDCNFGF   = 'PV\_init/PV\_PFT02\_revised.xml'
        + NL%IPHYSIOL = 4
        + NL%H2O\_PLANT\_LIM = 4
        + NL%ISTOMATA\_SCHEME = 1
      * Conservative
        + INCLUDE\_THESE\_PFT = 4
        + NL%IEDCNFGF   = 'PV\_init/PV\_PFT04\_revised.xml'
        + NL%IPHYSIOL = 4
        + NL%H2O\_PLANT\_LIM = 4
        + NL%ISTOMATA\_SCHEME = 1
      * NOTE: Many of my graphing scripts are written for five PFTs: C4 grass (years 2-4), acquisitive and conservative shrubs (years 2-4), and acquisitive and conservative trees (years 18-20)
      * Bare Run: for a simulation with no PFTs
        + IED\_INIT\_MODE = -1 (0 for other runs)
        + REPRO\_SCHEME = 0 (1 for other runs)
        + NL%IEDCNFGF   = ‘’ (i.e. don’t provide a .xml file)

**Step 2a: Pre-Processing on the CRC**

* Files are in *ED/AnalysisScripts/CRCscripts*
  + *move\_files.sh*: A script that moves “fast” (i.e. hourly) ED output into a new folder.
    - This should be run FIRST when dealing with fast output
    - Files are organized by year into new folders
    - add your email as argument to –M option to be notified of start/finish
  + *Rsubmit*.*sh*: A script that will submit R code; change the file to use your desired R script
    - add your email as argument to –M option to be notified of start/finish
  + *DailySoilWater.R*: A script to be run in the same directory that holds the folders of yearly fast output. For a given simulation, it averages fast output for each day and creates a .csv file entitled “DailySoilWater.csv” with the mean value of soil water (m^3 water/m^3 soil) for the top layer of soil for each day of the simulation
    - NOTE: In the long-run, it could be beneficial to re-write this file so that it simply extracts the daily soil water value from daily output files rather than averaging hourly output
  + *DailyRain.R*: Similar to *DailySoilWater.R* but instead results in file “DailyRain.csv” with daily rain (mm)
    - NOTE: Likewise, could also re-write to use daily files
  + *MinMaxTemp.R*: A script to be run in the same directory that holds the folders of yearly fast output. Creates two output files: “Temp.csv” and “Temp\_month.csv”. Both contain three variables: average of daily maximum temperatures, average of daily minimum temperatures, and average of daily temperatures. “Temp.csv” contains annual averages, while “Temp\_month.csv” contains monthly averages
    - NOTE: Unlike *DailySoilWater.R* and *DailyRain.R*, ED2 (at time of writing) does NOT output daily values for min/max temperature, so fast output MUST be used.

**Step 2b: Pre-Processing on Your Local Machine**

* Files are in *ED/AnalysisScripts/Pre-Processing/*
* **Getting Daily Soil Water into a more useful format**
  + For the purpose of graphing, I often found it useful to combine the separate daily soil water excel files into one file corresponding to a single set of simulations (e.g. C4 grass, conservative, and acquisitive for a single soil type) into one excel file.
  + *ExtractingDailyData.R*: After reading into R each individual .csv file, this script extracts the desired data and saves it into a .csv file that should be titled appropriately by the user
* **Monthly ED2 Output**
  + Although by the end of the semester I was dealing mostly with daily soil water values, I also wrote a number of scripts that dealt with monthly and yearly data.
    - NOTE: If dealing with yearly data, I averaged simulated monthly results
  + I often found it easiest to work with this data directly on my local machine, since the number of files was relatively small. I used the application *Cyberduck* to “drag and drop” from CRC to my local machine, but the command *scp* would also work.
  + *ExtractingAnnualNaomiData.R*: Reads-in raw monthly data and produces .Rda files of annual averages of various variables for each PFT.

**Step 3: Graphing Results**

* Files are in *ED/AnalysisScripts/Graphing/DailyGraphs*
  + *DailyRain\_Graphs.R*: Takes a .csv file that contains three years of daily rain data for five total PFTs and creates histograms
  + *DailySoilWater\_Graphs*.R: Same goal as *DailyRain\_Graphs*.R, but for soil water; however, there has been some code added and some commented out.
* Files are in *ED/AnalysisScrips/Graphing/AnnualGraphs*
  + *PFT\_BarGraphs.R*: Reads-in .Rda files for each PFT and produces a bar graph of 3-year averages (years 2-4 for grass and shrubs, years 18-20 for trees) of various variables
* For some sample data files, see *ED/AnalysisScripts/ExampleFiles*
  + *PFT01.Rda*, *PFT02.Rda*, and *PFT04.Rda*: R data frames containing annual averages of various values from the months June-December for C4 Grass, Acquisitive PFT, and Conservative PFT, respectively.
  + *DailSoilWater01.csv, DailySoilWater02.csv, DailySoilWater04.csv*, and *DailySoilWater\_SiltyClay\_Dry*.*csv*: The first three files are daily soil water values for PFTS 1, 2, and 4, respectively, produced on the CRC using the *DailySoilWater.R* script from a simulation that used the silty clay soil type and a characteristic dry year for meteorological data. The last file was produced using these three data files and the script *ExtractingDailyData.R*. It can be used with the script *DailySoilWater\_Graphs*.