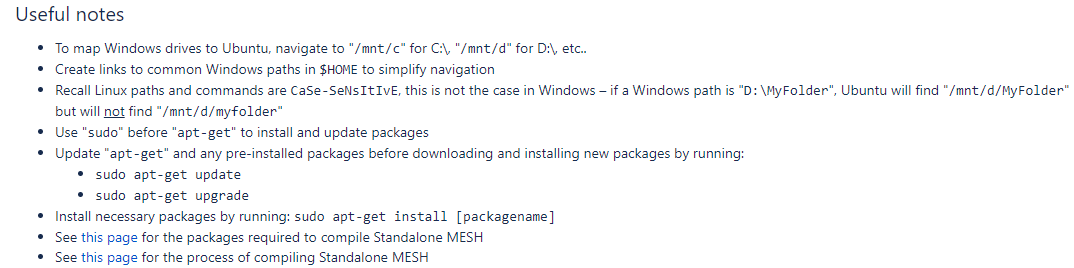
# How-to’s and Helpful Information

## Configuring Usask Wired Internet on BYOD

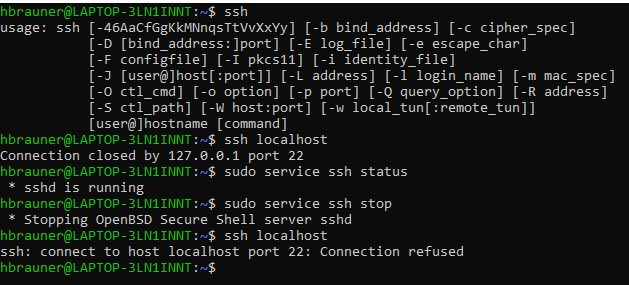
<https://wiki.usask.ca/display/public/ISDKB/How+to+Configure+Wired+802.1x+for+Windows>

## Setting up Linux on Windows

<https://wiki.usask.ca/display/MESH/Installing+and+Activating+Windows+Subsystem+for+Linux+on+Windows+10>



### SSH – Check/Enable/Disable

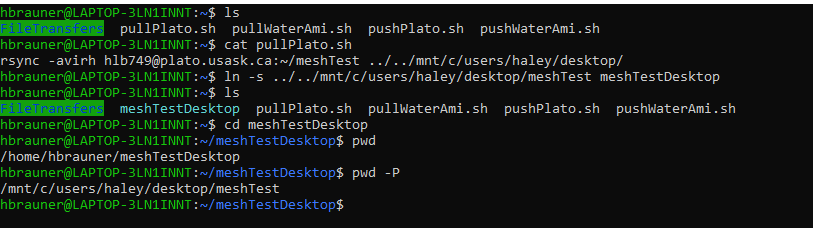


Error: Trying to connect to the server:



* Solution: must be connected to usask secure server; VPN doesn’t work
* If on GC or VPN internet, must use PuTTy / WinSCP

#### Symbolic Links on Ubuntu



## Downloading ECCC Weather Data using R

<http://ropensci.github.io/weathercan/>

## .rdb Error in RStudio

#### Error: was unable to find the file/directory with a .rdb extension

* Tried installing the package specified, but it just gave the error for another and another package

Solution: delete the “3.6” folder at C:\Users\haley\Documents\R\win-library and updated the packages and re-installed and it was fine; Found the description of the solution here: <https://github.com/rstudio/blogdown/issues/115>

## Plato

Email from IT:

Please refer to following documentation for additional information about use of HPC Cluster Plato:

1. Documentation for use and access to Plato <https://www.usask.ca/ict/services/research-technologies/advanced-computing/plato/index.php>

2. How to use software on Plato: <https://www.usask.ca/ict/services/research-technologies/advanced-computing/plato/index.php#AvailableSoftware>

3. How to set up environment for software:  <https://www.usask.ca/ict/services/research-technologies/advanced-computing/plato/index.php#SetupEnvironment>

All software you indicated already available on Plato via environmental modules (see above).

All users have 200GB disk quota by default and we have no limitation for memory use. Plato has nodes with 32GB of RAM, 192GB of RAM and one node with 2TB of RAM.

## Data Inventory

See this file: [C:\Users\haley\OneDrive\Documents\1.MWS2018-2019\T2\Project\ECCC\_Project\Data\Data Inventory.xlsx](file:///C:\Users\haley\OneDrive\Documents\1.MWS2018-2019\T2\Project\ECCC_Project\Data\Data%20Inventory.xlsx)

#### Also: Google My Maps for location and basic info

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Driving Data** |  |  |  |  |
| **Data Type** | **Source** | **Interval** | **Notes (dates, missing, location, etc.)** | **File Location** |
| Precipitation (P) |  |  |  |  |
| Temperature (T) |  |  |  |  |
| Wind speed (u) |  |  |  |  |
| Incoming shortwave (SWin) | ESSD |  |  |  |
| Incoming longwave (LWin) | ESSD |  |  |  |
| Air Pressure (pa) |  |  |  |  |
| Specific Humidity |  |  |  |  |
| **Calibration / Verification Data** | | |  |  |
| Streamflow | WSC |  | 1983-01-01 to 2016-12-31  To 2019 is available (from ECCC) |  |
| ET | ESSD |  |  |  |
| Soil Temp | ESSD |  |  |  |
| Soil Moisture | ESSD |  |  |  |
| Snow Depth | ESSD |  |  |  |

## MESH Structure

### MESH Prep Checklist

#### MESH\_input\_run\_options.ini

* Hourly flag divisible by 30 and = the timestep of the forcing data?
* File format of the forcing files specified by the flag values?
* Different forcing parameter timestep specified with hf=X?
* Any flags that you want to change the default parameter? (see the Wiki page for a list of all options) <https://wiki.usask.ca/display/MESH/MESH_input_run_options.ini>
* Do the output directories specified here exist in folder?
* Have you chosen to start the model at a time of year when setting initial parameters (CLASS.ini file lines 17-19) are easier?
  + i.e. no snow, no ice in soil
  + unrealistic initial values can result in long-term bias in the model

#### MESH\_soil\_levels.txt

* Set the soil layer thicknesses to be between soil moisture/temp measurements (if have)

#### Matching Items

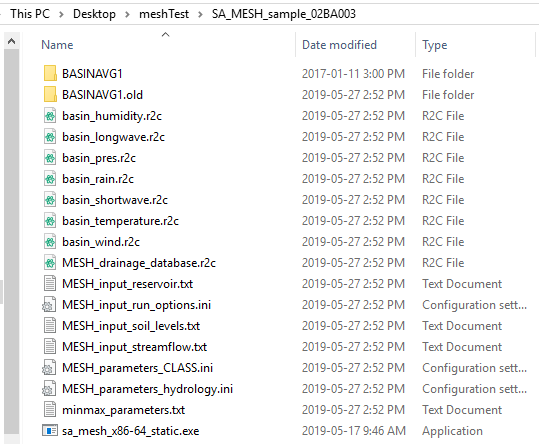
* Number of river classes (IAK) in drainage database file && River roughness values in the hydrology options file
* Number of GRUs (CLASS.ini file) && number of classes (-1 “dummy” class) in drainage database && number of GRUs in the hydrology parameters .ini file
* Number of grid cells in CLASS.ini && drainage database.r2c
* Output directory specified in run options file && directory exists in the model folder
* In run options file:
  + If FROZENSOILINFILFLAG enabled -> set parameters in the hydrology.ini file
  + If PDMROF enabled (IWF=2) -> set parameters in the hydrology.ini file
  + If PBSMFLAG enabled -> set parameters in the hydrology.ini file
* If more than 3 soil layers:
  + Define layers in the soil\_levels.txt file
  + Define parameters in the CLASS.ini file (additional columns)
  + Turn on the “NRSOILAYEREADFLAG” in run\_options.ini
* ROOTS (max. rooting depth) in CLASS.ini and soil layers in soil\_levels.txt
  + If roots just enter a layer, they theoretically have access to all the moisture in that layer which can result in over-estimating ET; tweak rooting depth as necessary.
* LOCATIONFLAG in run\_options file && format of locations specified in the input\_reservoir and input\_streamflow files

## Ostrich

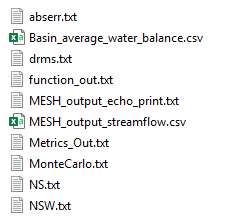
* How to use Ostrich on the MESH Wiki: <https://wiki.usask.ca/display/MESH/Calibration+using+Ostrich>
  + How-to manual and program documentation

## MESH Sample

* Pre-running:



* + BASINAVG1 is empty
  + BASINAVG1.old contains:



## Compiling / Configuring MESH (from training May 28-29)

Point mode:

* <https://wiki.usask.ca/display/MESH/Compiling+Standalone+MESH>
* Use the old code: version 1149
* Has many versions of the makefile
  + make -f makefile.ifort (for Plato)
    - The “-f” tells the program that we want to specify which makefile to use
  + make -f makefile.gfortran clean
    - Run “make -f makefile.gfortran clean” afterward to get rid of the extra .o and .mod files and leave just sa\_mesh, which is what we use to run the model

Non-point mode:

* Use the new way of compiling – 1398 or 1555
* Interim releases: <https://wiki.usask.ca/display/MESH/Interim+releases>
* Has 2 versions of the makefile
  + Use “make gfortran” (to use gfortran on ubuntu, or target ifort for plato)
  + Can also add a debug target to compile without optimizing the code in order to get information during crashes

Use the new version of the MESH\_parameters\_hydrology.ini file (2.0)

<https://wiki.usask.ca/pages/viewpage.action?pageId=1176797394>

#### Input Files

* Temp in K
* Specific humidity
* Wind speed: there’s a lower limit to the value (ex. 0.1; and zeroes will assume the minimum value)
* CLASS: reference height can change the physics (ex. using a height below the canopy)
  + Must decide whether reference heights are within or below the canopy at the start of the project, set the flags in CLASS (ex. izref), and then **keep the meas. reference heights the same for the whole project**
    - Recommended to use a reference height above the canopy
* CLASS.ini
  + Use a combination of the rooting depth and permeable depth to control the plant-available water for transpiration;
    - Don’t allow a thin layer under thick layers (i.e. setting a permeable depth just below a soil type interface)
    - May want to set the permeable depth at a soil layer interface
  + Peat soil: see the wiki and the CLASS documentation for setting soil type as non-mineral soil
    - Rule of thumb: if organic matter is >30% then use the special flag
  + Hydraulic and thermal conductivity are calculated based on texture specs
    - See CLASS B manual
  + Rule of thumb: start the model when the basin is hydrologically inactive
    - Eliminate: ponded water, snow on ground, heavy precip to make the initialization variables
    - Initial temps: canopy=same as air, top soil = ???, soil layer 2 = monthly avg, soil layer 3 = yearly avg.
  + Soil texture
    - Sand+clay+organic (the difference from 100 is the silt content)
    - Value of the 3rd layer carries through to anything deeper
      * Need to configure the model differently if you want a different layer to carry to depth
    - Don’t need more soil layers to get a better representation; it depends on your purpose
      * More layers makes the model run slower and increases the chance of instability
    - Generally want to increase the soil layer thickness with depth
  + Do a spin-up -> run in a circle until steady state is reached
  + 2 reference heights: 1 for temp and RH, and one for wind speed (

#### Initial running

* Start with a sample that works
* Then change the forcing file(s)
* Then change the parameter

## Software I’ve Installed for this Project

* R language
* R Studio
* Ubuntu – for executing bash commands from my computer, easily accessing the server/cluster
* Green Kenue – for setting up the watershed
* Qgis – open source gis software (similar to Arc GIS)
* Winmerge – for checking the differences between folders and files and merging them
  + Allows you to compare 2 folders or files and highlights differences
  + Useful for checking differences between outputs, setup files (parameters), etc.
* Atom – text editor (windows standard notepad is not suitable)

# MESH Modelling Journal

## Driving Data Preparation Notes

Script: Prepare\_Driving\_Data.R

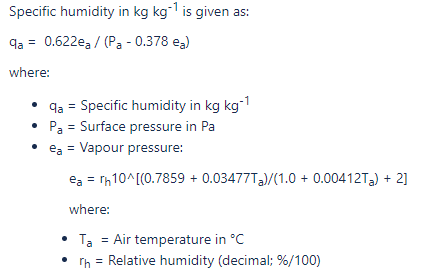
### Driving Data Required for the Model:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Precip\_rate | Kin | Lin | Tair | u | Pair | q |
| mm/s | W/m2 | W/m2 | K | m/s | Pa | kg/kg |

### Vital Data

**Source**: Spence, C., & Hedstrom, N. (2018). Hydrometeorological data from Baker Creek Research Watershed, Northwest Territories, Canada. *Earth System Science Data*, *10*(4), 1753-1767   
<http://dx.doi.org/10.5194/essd-10-1753-2018>

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| DateTime | Kin | Kout | Lin | Lout | Tair | Tair | u | RH | RH | Pair | Q\* | Qe | Qh | rain |
| ½ hourly  Apr. 12/05 8:30 to Dec. 31/16 23:30 | W/  m2 | W/  m2 | W/  m2 | W/  m2 | oC | oC | m/s | % | % | kPa | W/  m2 | W/  m2 | W/  m2 | mm |
| Not winter |  |  |  |  | 1.1m | 4.4m | 4.4m | 1.1m | 4.4m |  |  |  |  |  |

* Converted air pressure to Pa by multiplying by 1000
* Converted RH at both heights to specific humidity using the equation on the MESH Wiki
* 
* Converted rain to a precipitation rate (mm/0.5h/3600s/h)

### Landing Data

**Source**: Spence, C., & Hedstrom, N. (2018). Hydrometeorological data from Baker Creek Research Watershed, Northwest Territories, Canada. *Earth System Science Data*, *10*(4), 1753-1767   
<http://dx.doi.org/10.5194/essd-10-1753-2018>

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| DateTime | Kin | Kout | Tair | T water | u | u dir | e | Pair | Q\* | Qe | Qh | rain |
| ½ hourly  Jan. 1/08 0:00 to Dec. 31/16 23:30 | W/  m2 | W/  m2 | oC | oC | m/s | deg | kPa | kPa | W/  m2 | W/  m2 | W/  m2 | mm |
| Not winter |  |  | 1.4m | 0m | 1.1m |  | 1.4m |  |  |  |  |  |

### GEM Data

#### Source: extracted from the GEM model by Dan Princz

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| DateTime | Kin | Lin | Tair | Tair | u | u |  |  | RH | RH | Pair | Q\* | Qe | Qh | rain |
| hourly  May 19, 2004 to Dec. 31, 2018 23:00 | W/  m2 | W/  m2 | oC  From Oct. 1, 2011 | oC | m/s | m/s |  |  | % | % | kPa | W/  m2 | W/  m2 | W/  m2 | mm |
| Not winter |  |  | 2m | 40m | 10m | 40#m |  |  | 1.1m | 4.4m |  |  |  |  |  |

* Given hourly GEM data for Kin, Lin, Tair, u, Pa, and q. Will run the model at half-hourly since we have half-hourly observations at the Vital and Landing Tower stations. Therefore, will scale the GEM data down to half-hourly.
  + Methodology will be to use the value at the hour for the half-hour value. I did create a script to calculate the average between, but this would not preserve the overall average value of the data. (Note: I kept the formula commented in the script “Combine\_Driving\_Data.R” if we want to change to the interpolation method

### CAPA Data

### Yellowknife A Data

## Baker Creek Configuration Notes / Decisions

### Replication:

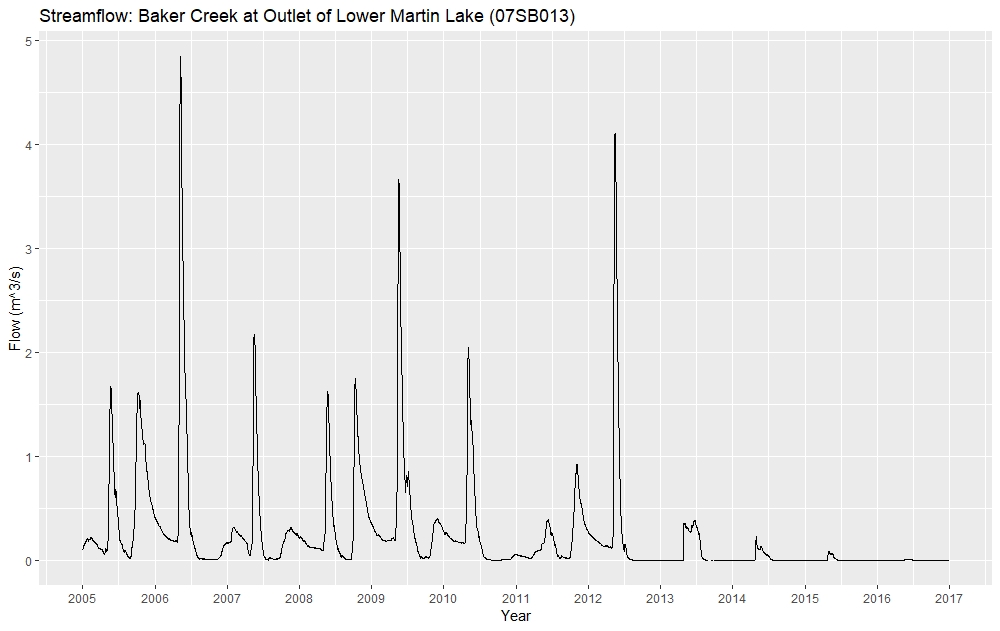
* Set up the same as Herbert but adjusting the parameters to reflect the different landscape, as necessary

### Met Data period:

* GEM/CaPA data for 2005-01-01 00:00:00 to 2018-12-31 16:00:00

### Calibration Period(s):

* Want to cover both wet and dry periods
  + Wet period: 2008-2011
  + Dry period: 2012-216
* Have streamflow to the end of 2016
* Baker Creek WSC Gauge Streamflow:



Validation

Validation

Calibration

Calibration

Spin-up

#### Calibration/Validation/Spin-up

Spin-up: fall 2006-fall 2007 (1 year)

Calibrate: fall 2007 – fall 2010 and fall 2013 – fall 2015 (5 years total)

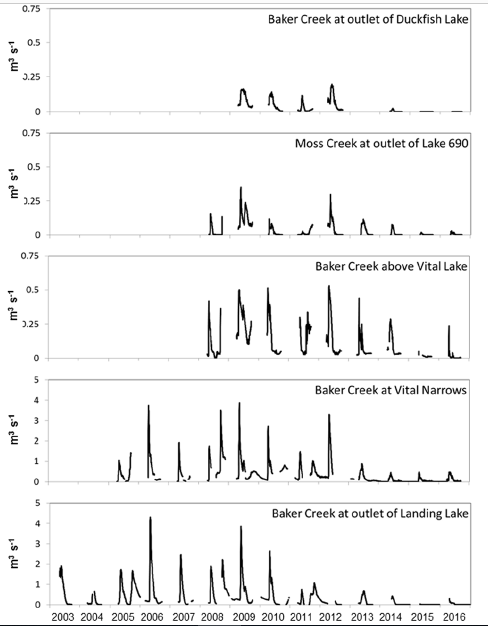
Validate: remaining: fall 2005-fall 2006, fall 2010-fall 2013, fall 2015 - fall 2016 (5 years total; don’t have streamflow past 2016)

When running:

Calibration: 2006-258 to 2016-258 but have streamflow as -ve values for spin-up and validation periods

Validation: 2005-285 to 2016-258 but have streamflow as -ve values for spin-up and calibration periods

Q: is this right? Or can I still validate during the spin-up period?

* Streamflow from the research gauges:
* 

### Select Baker Creek model start date

#### Combined Precip Data (used in the model):

|  |  |
| --- | --- |
| **Date** | **Precip (mm)** |
| 2005-09-25 | 10.73 |
| 2005-09-26 | 0.68 |
| 2005-09-27 | 0.49 |
| 2005-09-28 | 8.10 |
| 2005-09-29 | 1.58 |
| 2005-09-30 | 0.00 |
| 2005-10-01 | 0.00 |

|  |  |
| --- | --- |
| **Date** | **Precip (mm)** |
| 2005-09-10 | 0.00 |
| 2005-09-11 | 7.90 |
| 2005-09-12 | 4.90 |
| 2005-09-13 | 0.00 |
| 2005-09-14 | 0.00 |
| 2005-09-15 | 0.00 |

Since there was a fair bit of rain a few days before October 1st, would prefer to use September 15, 2005 as the model start date

#### Temperature Data (Combined; used in the model):

|  |  |
| --- | --- |
| **Date** | **Daily Avg. T (oC)** |
| 2005-09-10 | 8.2725375 |
| 2005-09-11 | 7.4647667 |
| 2005-09-12 | 3.5984542 |
| 2005-09-13 | 1.7400375 |
| 2005-09-14 | 3.5648083 |
| 2005-09-15 | 4.6777458 |
| ~ | ~ |
| 2005-09-26 | 1.1207083 |
| 2005-09-27 | -1.6074167 |
| 2005-09-28 | -1.1949167 |
| 2005-09-29 | 0.1040417 |
| 2005-09-30 | -3.1303333 |
| 2005-10-01 | -2.2324167 |

Daily average temperature on Oct. 1 is below zero, so prefer to use Sept. 15

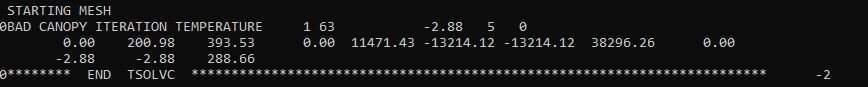
\*\*\*Choose to start the model Sept. 15, 2005 (day 258) \*\*\*

### Transition to replication

* General Procedure, Scenario 1 (simplest):
  + Run Herbert’s non-calibrated Scenario 1 setup, gradually changing the input files to Baker Creek; may need to change information in runoptions file each time
  + Do for the time period 2005-2006 so that streamflow and driving data are available for both Herbert and Baker Creek
  + Ran using Linux (Ubuntu) on Haley’s computer
  + MESH Code used: r1024 (from Herbert’s files) compiled using gfortran
* CLASS Parameters
  + Vegetation parameters: based on discussions with Chris Spence and the hillslope setup provided (see the “Baker Creek Configuration Details” xlsx document)
  + Soil Parameters: based on discussions with Chris and on the hillslope setup provided. The provided setup included 5 soil layers, but didn’t include a soil layers document detailing the thickness of the layers. Assumed 4 layers with division at: 0.1m, 0.35m, 1.1m, and 4.1m; it is assumed that maximum rooting depth is 1m and permafrost may be present below 1m (or shallower) in wetland or peatland areas. Note that for scenario 1, the soil layers are all the same for the whole watershed (i.e. the difference between bedrock and peat is not explicitly represented)
  + Initial values
    - The values were calculated using the “SoilTempMoisture.R” script
    - Soil temperatures: obtained from the weighted average by relative landcover percent (Bedrock, Hillslope, Wetland, and Peatland) of the mean soil temperature on October 1 for the years of observation available. Since only surface and 25cm depth values were available for the hillslope, Wetland, and Peatland landcovers, deeper values were estimated assuming the soil gets cooler with depth (though this is not the case at the Bedrock site)
    - TCAN: assumed to be the average air temperature on October 1, 2005 at the Yellowknife station (Vital data not available on that day)
    - TSNO=0 (assuming no snow on the ground)
    - TPND: obtained as the average Twater at the Landing tower observed on October 1
    - THLQ: obtained from the weighted average by relative landcover percent (Hillslope, Wetland, and Peatland) of the mean soil moisture on Oct. 1 for the years of observation available. Since only surface and 25cm depth values were available, made the assumption that soil moisture is consistent with depth (see Morse et al, 2016, Figure 4). It is noted that the average soil moisture contents seem high; may want to consider looking at the data for the few weeks prior to see if the high values are due to partial soil freezing.
    - THIC: assume no frozen portion of soil, so all =0.0
    - ZPND=RCAN=SCAN=SNO=0 -> assume no ponding of water in any form on the ground or in the canopy
    - ALBS and RHOS: though they don’t matter, assume values of 0.2 and 100.0
    - GRO: assume 1.0 (full leaf-out), though may want to consider a smaller fraction if trees are already dormant.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Source** | | | | | | |
| **Setup File(s)** | **Change Order** | **Step 1** | **Step 2** | **Step 3** | **Step 4** | **Step 5** | **Step 6** | **Step 7** |
| **Purpose** | Change file in Step X | Test Herbert’s setup before changing things | Change input\_reservoir.txt | Change the hydrology.ini file | Change the soil levels and CLASS files | Use the Baker Creek Streamflow (07SB013) | Use the Baker Creek driving data (all 7 files) | Use the Baker Creek drainage database info |
| \_reservoir.txt | 2 | Herbert | Baker Creek | Baker Creek | Baker Creek | Baker Creek | Baker Creek | Baker Creek |
| \_run\_options.ini | 1 | Baker Creek | Baker Creek | Baker Creek | Baker Creek | Baker Creek | Baker Creek | Baker Creek |
| \_soil\_levels.txt | 4 | Herbert | Herbert | Herbert | Baker Creek | Baker Creek | Baker Creek | Baker Creek |
| \_streamflow.txt | 5 | Herbert | Herbert | Herbert | Herbert | Baker Creek | Baker Creek | Baker Creek |
| \_CLASS.ini | 4 | Herbert | Herbert | Herbert | Baker Creek | Baker Creek | Baker Creek | Baker Creek |
| \_hydrology.ini | 3 | Herbert | Herbert | Baker Creek | Baker Creek | Baker Creek | Baker Creek | Baker Creek |
| Drainage\_database.r2c | 7 | Herbert | Herbert | Herbert | Herbert | Herbert | Herbert | Baker Creek |
| driving data (.csv) | 6 | Herbert | Herbert | Herbert | Herbert | Herbert | Baker Creek | Baker Creek |
| **Notes** |  | - Only thing changed was run period (2005-2006 day 274=Oct. 1)  - Ran fine, no errors or weird results  - No NS or NSW values output | - Ran fine, same results (since both watersheds had no reservoirs)  - No NS or NSW values output | - Changing WF\_R2, ZSNL, ZPLS, ZPLG had only a small change on the water balance values; slight increase in interflow and decrease in baseflow  - No NS or NSW values output | - Ran fine; Total precip is the same as the previous run; had way more evap and a lot less change in storage, interflow, and baseflow  - No NS or NSW values output | - Ran fine; Note that the lat-long in the input\_streamflow file kept the same as the Whitegull setup since haven’t changed the drainage database file yet  - No NS or NSW values output | - Updated the start date and time of the met data in CLASS.ini  - Error: see note below; forgot to convert T driving data to K  - Ran fine  - Very little overland and interflow, and no baseflow | - Need to change the reference height  - Consider the start time of the model wrt the reality of the initial parameters  - Change the location of the streamflow gauge so within basin |

Step 6 Errors:



* TCAN: assumed to be the average air temperature on October 1, 2005 at the Yellowknife station (Vital data not available on that day)
  + Try using the maximum air temperature on Oct. 1, 2005 = 0.4oC (Yellowknife Station @ 2m)
    - Still didn’t work
  + Forgot to convert temperature of the driving data file to kelvin; convert and retry

# MESH Model Folder Structure

* On personal computer, water-ami, and Plato
  + Folder
  + Files

### File Flow:

# Model Runs and Results

## Scenario 1

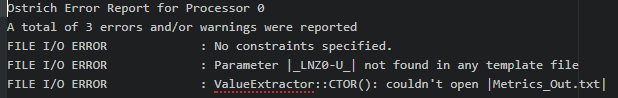
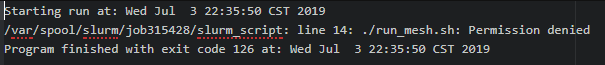
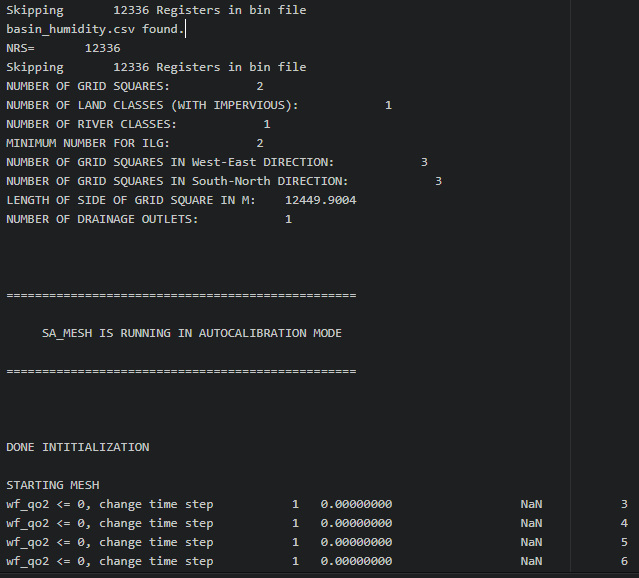
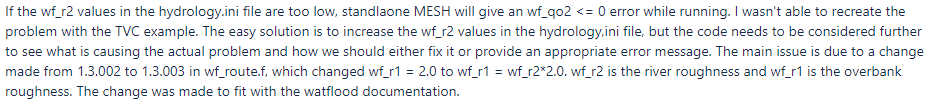
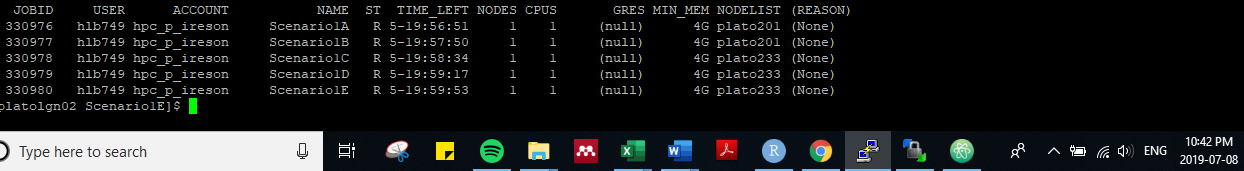
Date: July 3, 2019

Run location: Plato -> ran from the RunMesh/Scenario1 folder

Job Number: 315422

Ostrich code: OstrichGCC

Notes:

* I submitted the job using “sbatch submitjob.sh” but then nothing appeared in the queue
* Ostrich output/error files were created. From the OstErrors0 file:
* 
  + No constraints specified: missed the “minmax\_parameters.txt” file in the folder; added it from the Step 7 folder; but that didn’t help
  + LNZ0 error: had used the letter “O” instead of a zero; fixed
  + Checked the run\_mesh.sh file – added tabs before the ending quotes directing a dummy Metrics\_Out.txt file to be written
  + Also found an error that it couldn’t open run\_mesh.sh -> “permission denied”
    - Later found out this is because the file was not marked as “executable”; from WinSCP, right click, select “properties”, and check all the “x” boxes, or use chmod +x from the command line; if you copy from windows to the linux system, it will likely shut off the executable permission again
* Copy the files back to Plato and try again
  + Still didn’t work; getting the 1st and 3rd errors still, as well as an error that ./run\_mesh.sh can’t be opened
  + 
    - Fixed: syntax is simply “run\_mesh.sh” or the hard path “/home/hlb749/RunMesh/Scenario1/run\_mesh.sh”
* Ran mesh-only Step 7 on the login node (on Plato) and it worked fine
  + So the problem is not the mesh executable file
* Try running Scenario 1 in static mode to see if everything works before calibrating
  + Yes, it does
* Try running Ostrich and calling sa\_mesh directly, rather than the run\_mesh script (to determine if there’s a problem with the script)
  + Works; also fixed the problem with the run\_mesh.sh script so back to using that instead
* BASEFLOWFLAG: had this option as “1” (in the run\_options.ini file) for the transition and initial runs of Scenario 1; wasn’t seeing baseflow in the results; took this out and now there’s baseflow showing up in the model results
* Tried running the model for the full duration but with only 10 iterations and 10 trials for Ostrich; “permission denied” error when trying to open run\_mesh.sh again
* Tried running for the full duration using only mesh
  + The slurm output (what is normally printed on the screen when running mesh) showed the following:
  + 
  + I checked the driving data in Excel (using the .xlsx.csv file) and it seems that the driving data is off by 1 hour (2 timesteps)… time change error when creating the data in R? But how would the model know this, as we only specify a start date and time for the driving data (which was set to Jan 1 2005)
  + Found this on the wiki:
  + 
  + When the error showed up, WF\_R2 was set to 0.149 ish (I think?), so changed it to 0.349, but still got the error so changed it to 3.49 and tried running again
    - For code version 1024 that Herbert used, was the overbank and river roughness read from the r2c file? Dan said that those values aren’t actually read in the model so I put nonsense values in; if still getting the wf\_qo2 error, try putting more meaningful values in the r2c file
  + Error still showed up, so in the r2c file I changed Bankfull value to 100 and IntSlope value to 0.00999 (previously both were set to 999.0 for null purposes).
    - Still got the error!
  + Try turning off the AUTOCALIBRATION flag
    - Nope, didn’t fix it. Turn the flag back on.
  + Maybe MESH doesn’t like starting with no streamflow? Could try running from 2007-258 to the end (2016-257)
    - SUCCESS!!! I think the model doesn’t like starting running during a validation period
  + Now try running Ostrich again
    - Getting the “sh” run\_mesh.sh permission denied” error again
    - Maybe this script is being read, but one of the executables in the file isn’t being allowed?
      * Tried commenting out the section that writes a dummy Metrics\_Out.txt file
        + Still get the error
      * Try commenting out the section where the existing MESH output files are removed
        + Nope, still get the error
      * Try commenting out the section that calls mesh to run
        + Nope, still get the same error
    - Ok, so if in the “ostIn.txt” file I call run\_mesh.sh without the ./ in front, I get an error saying it can’t read the file, but if I include the relative path then I get the permission denied error. What if I write the hard path to the file?
      * Still get the permission denied error
    - Noticed that the ostOutput00X folders are still there; try deleting and running again
      * Nope, same error
    - Try deleting files/folders to create a fresh state and running again
      * Also added some script to the submitjobtrials.sh file to remove ostOutput folders if they exist
    - Try running without the trials again
* I’m at a loss as to why this isn’t working. Went back to the folder Scenario1TestTrials10iter and ran Ostrich again and it worked
  + Try copying the folder and incrementally changing things, running Ostrich again each time
    - Only run for 2 iterations = good!
    - Change the simulation run time to 2007-2016 = good!
    - Went back to the 1 year sim. run time (to test the model faster), kept number of trials at 2 and number of iterations at 10, and removed the Baseflowflag (deleted the line and changed the # of options to 17 from 18) = good (ostOutput folders have all MESH and Ostrich outputs specified, MESH\_output\_echo\_print has values and message looks ok). HOWEVER, the NSD was on the order of -300 and -700, so something wrong there. Maybe need to look into specifying constraints for GCOP?
      * The OstExeOut.txt file said that the parameter XDrainH (aka GRKF) was out of range.
        + This value is the ratio of the saturated horizontal conductivity at a depth of 1m to that at the surface; changed to a solid parameter of 0.01 (after Davison et al 2016). Ran ok
        + Ran the model again; no parameter range errors this time (for the 2nd trial anyway (as per the OstExeOut.txt file in the working directory); for trial 1, NSD is -255, for trial 2 NSD is -149; possibly because only 10 iterations; ok for now – don’t change any parameters
    - Add the “remove ostOutput folders” to the submitjobtrials script = ran fine, timestamp on the ostOutput folders reflects the model run time; no errors in OstExeOut or the slurm file and all the expected files exist in the ostOutput folders
    - Changed the GRKF to calibration parameter again, range to 0.01 to 0.5 (Davison et al 2016 and ini files from Chris)
      * Trial 1 still had very large -ve NSD, but Trial 2 had better results
    - Change ostIn.txt to select random start rather than initial
      * Ran well, large NSD errors again but that’s ok for now (short modelled period, small number of iterations and trials)
    - Check the dates of the driving data (in R)
      * Initial inspection: there is a 1 hour time shift (ahead) in the humidity driving data on 2015-04-03 after 01:30:00. Need to look into this!!! Likely a timezone thing in R;
* Start working up to the full calibration
  + Save the contents of folder Scenario1TestTrials10iter2 to a new folder “Scenario1”
  + See the time trials below
  + Both the 8hr, 2G memory requests and the 16hr 4G memory requests started right away
* It looks like the model is running ok, and we now have a good representative time trial. However, I don’t think I’ve seen an NS value above zero for any of the tests so far; THEREFORE, I AM GOING BACK TO REVISIT THE PARAMETER RANGES. Will look at the parameter ranges chosen for the distributed setup and set the ranges for Scenario 1 as the full range
* Adding a spin-up period
  + Scenario1\_3: add “METRICSSPINUP 365” to run options, start running from 2006 day 258, and change streamflow values from 2006 Sept. 15 to 2007 Sept. 14 to positive values; slurm 328184 -> seems to be running fine
  + Scenario1\_4: add “METRICSSPINUP 365” to run options, start running from 2006 day 258, and only change streamflow value on Sept. 15 2006 to +ve to get the model running (keep the remainder -ve between 2006-2007-258); slurm 328186; used the same seed number as the spin-up trial for Scenario 1\_3 for comparison purposes.
  + Results:
    - Scenario1\_3: NS=0.202
    - Scenario1\_4: NS=0.237
    - Already seeing better results just by initializing the model
* Changing the Scenario 1 parameter ranges (in folder Scenario1\_5)
  + Most parameters for Scenario 1 were pretty close to the full range for Scenario 2; noted the ones adjusted in the “Baker Creek Calibration Details” spreadsheet
  + Soils might be the issue; going to try limiting the sum of all the %’s to 1 by setting the SOILINIFLAG=4 (if sum>1, then reduce all proportionally to equal 1)
  + Calculated new proportions for soil composition by multiplying the lower and upper bounds for each layer and each landcover type (based on Scenario 2 ranges) by the % landcover to get a weighted value
    - Assumed bedrock behaves as 100% clay (this may be way off, but we’ll see)
    - Where the SAND=-2 flag was used for organic matter, considered that layer 100% organic matter
  + Test for 2006-2008 (with 1 year spin-up) to see what kind of NS we get
    - Folder: Scenario1\_5
    - This involved:
      * Run\_options.ini -> change the SOILINIFLAG to 1
      * ostIn.txt -> change bounds for ZPLG, LAMX-G, ALIC-U, SDEP, GRKF, MANN, and changed which soil layers/textures are calibrated and their ranges
      * …CLASS.tpl -> changed the soil values, both static values and calibration symbol
      * Minmax\_parameters.txt -> Increased the maximum allowable range of ZPLGROW from 1.0000 to
    - Slurm: 330038
      * Didn’t start right away – waited in the queue for about 1 minute
  + I just realized I’ve been using decimal instead of whole percent for soil proportion values!!!!
    - Essentially this was making the soil very high in silt
    - Will incorporate the updated values to see how the results will change
      * Originally: NS=0.202
      * Updated: NS=0.520 -> much better!!
  + Note: The streamflow value on the first day of the simulation but be >0, not the first entry in the streamflow file (in case it starts before the simulation)
* Preparing the final calibration run
  + **Pre-run checks:**
  + ~~Streamflow<0 for all non-calibration years and >0 for simulation start date~~
  + ostIn:
    - RandomSeed NOT specified
    - MaxIterations=1000
    - \*\*\*Consider calibrating to a tied NSD and lnNSD parameter(since lots of low-flow)
    - Number of trials
  + ~~run\_options~~
    - ~~Simulation run time~~
  + submitjob.sh
    - number of trials = (20)
      * Will divide into 5
      * Numbered all trials sequentially from 001 to 100 -> 20 in each folder
    - mem=4G (Can I ask for more? Dunno – they took the support pages down)
    - time=140
    - Remove copying CLASSOUT/\*
  + Separate into 1A, 1B, 1C, 1D, and 1E folders and run from each
* 

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Folder Name** | **Purpose** | **Modelled Years** | **Num. Iterations** | **Num Trials** | **Run time** | **Notes** |
| Scenario1 | Time trial for the full period | 2007-2016 | 1000 | 1 | ~6.5h |  |
| Scenario1\_2 | Time trial for 1 year | 2007-2008 | 1000 | 1 | ~34 min | Main folder is without printing CLASS outputs; subfolder is with running CLASS (took longer) |
| Scenario1\_3 | Test out including a spin-up period | 2006-2008 | 1000 | 1 | ~1.25hr | Spin-up incorporated by starting at 2006-258, adding a METRICSSPINUP flag of 365, and chaning the streamflow values to +ve for the 2006-2007 year |
| Scenario1\_4 | Test out including a spin-up period | 2006-2008 | 1000 | 1 | ~1.25h | Spin-up incorporated by starting at 2006-258, adding a METRICSSPINUP flag of 365, and changing only the very first value (2005-01-01) to a +ve value |
| Scenario1\_5 | Test the updated parameter ranges and soil % (now whole instead of fraction) | 2006-2008 | 1000 | 1 | ~1.25h | Based on the Scenario1\_4 setup |
| Scenario1A through Scenario 1E | Run 1/5 of the calibration trials | 2006-2008 | 1000 | 20 ea | <140? |  |

#### Run Time Calculation

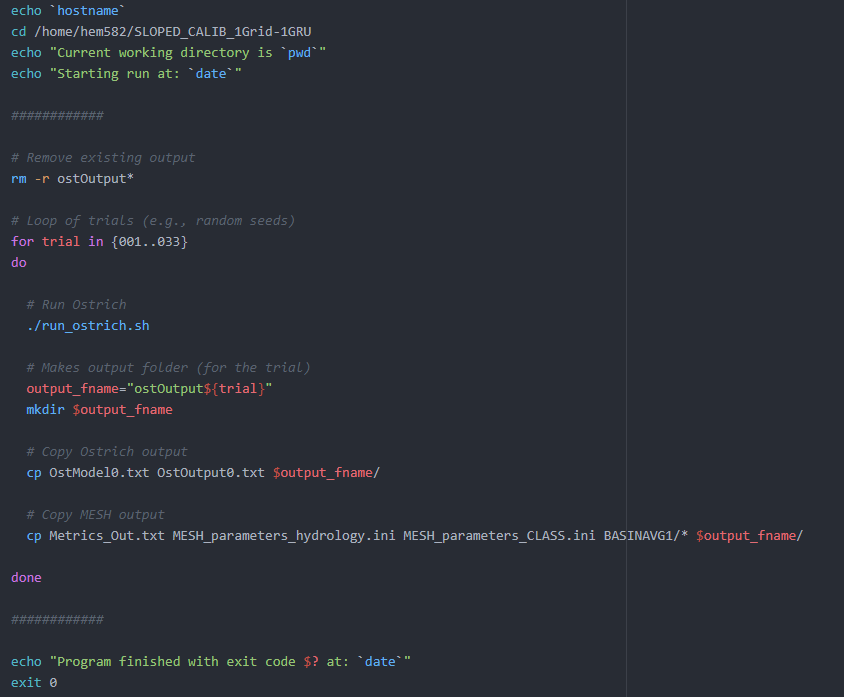
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Scenario** | **Num. Pars** | **Modelled Years** | **Modelled Duration** | **Cal Duration** | **Num. Iterations** | **Num Trials** | **Ostrich run time** | **Notes** |
| 1 | 41 | 2007-2008 | 1 year | 1 year | 10 | 10 | 4 mins | Folder: Scenario1TestTrials10iter |
| 1 | 41 | 2007-2008 | 1 year | 1 year | 10 | 2 | ~2 mins (~ 1 min /trial /year) | Folder: Scenario1TestTrials10iter2  - Noticed that MESH takes between 40-50s for each run (shown on the output\_echo file) |
| 1 | 41 | 2007-2016 (day 258 to 257) | 9 years | 5 years | 10 | 1 | 8 min 21 sec  (~ 1 min /trial /year) | No errors, all files look good; NS=-21.290; slurm-328138.out |
| 1 | 41 | 2007-2008 | 1 year | 1 year | 1000 | 1 | 1 hr 21 minutes | slurm-328140  NS: -0.61  Saved in Scenario1\_2/WithCLASSOut |
| 1 | 41 | 2007-2008 | 1 year | 1 year | 1000 | 1 | 19.6% after 425 s (est: 36min)  34min 19 sec | No CLASS output printed (on purpose)  slurm-328167  NS: -0.586  Folder: Scenario1\_2 |
| 1 | 41 | 2007-2016 | 9 years | 5 years | 1000 | 1 | NC; 54.2% complete after 7 hrs | slurm-328139  Aborted at approx. 7 hours running  Folder: Scenario 1 |
| 1 | 41 | 2007-2016 | 9 years | 5 years | 1000 | 1 | 0.8% after 160s  (est: 5.5h)  6h 32min 33s | Change run time to 16 hours, request 4G mem (up from 2G), and try running without printing a CLASS output  Folder: Scenario1\_longtrial  Slurm: 328168  NS: -0.175 |

### Setting up the final run

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Scenario** | **Num. Pars** | **Modelled Years** | **Modelled Duration** | **Cal Duration** | **Num. Iterations** | **Num Trials** | **Ostrich run time** | **Notes** |
| 1 | 41 | 2006-2016 | 10 yrs | 5 yrs | 1000 | 100 | Start: Jul 8 22:39 (1A), End: Jul 14 13:19 (1B), Total: 5 days 14 hrs 40 mins |  |

### Running 100 trials

#### Create a looping script in the submitjob.sh file

* Herbert’s script:
* 
* Was able to run the Scenario 1 model using a similar script in the “submitjobtrials.sh” script which incorporates the loop (above) for 10 trials of 10 iterations each
  + The “OstExe.txt” file shows the MESH outputs that are usually printed to the screen
* Need to clean up my scenario 1 folders once everything is working properly
* Need to write an R code to gather all the results from the Ostrich files and plot them

Checking Scenario 1 for Crashes (“dummy” Metrics\_Out file had incorrect values that wouldn’t have lead the model away from the crash)

* OstOut001 through OstOut045: checked the OstModel0 file for 0.999 -> none (i.e. no crashes)

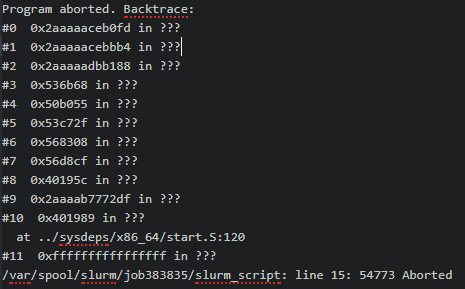
### Running Scenario 1 Again (Scenario1\_Calibrated\_2)

* Because:
  + Run\_mesh was not used, so the previous version of Metrics\_Out was used in the next run
  + Out of 100 trials, 82 of them had at least 1 crash
  + The number of crashes out of 1000 iterations ranged from 0 to 167 (16.7%)
* Fix the values in the dummy Metrics\_out in run\_mesh.sh
* Enable executable for run\_mesh.sh
* Change ostIn so run\_mesh.sh is called
* Change the range of ZSNL to 0.001-0.2
* 1st run saved in the folder “Scenario1\_Calibrated\_1”
* Scenario1D stopped running because there was an error in the ostIn.txt file; corrected and started running again.
* Not all the trials ran because didn’t change the from:to numbers in the “submitjob.sh” file; during the first run, had stopped part-way through and started again. Once finished, run the remaining trials
* This run is saved in the folder “Scenario1\_Calibrated\_2”

### Running Scenario 1 a 3rd time (Scenario1\_Calibrated\_3)

* Because:
  + Correct some inconsistencies in the parameter ranges compared to Scenario 2
  + Make the soiliniflag = 1 (instead of 4) just like the other scenarios

### Validation

* Changed the streamflow data to reflect the validation period as Sept.15 2005-Sept. 14 2006, Sept. 15 2010 - Sept. 14 2013, and Sept. 15 2015 – Sept. 14 2016
* Obtained the top 10 best parameter sets based on the calibration using R code from the “MESHOutput.Rmd” (RCode folder) file to write the “10BestCalParams.csv” file, and then calling it with “Find\_Replace\_in\_Text\_Val.R” file (RCode/ValTest folder) to write the CLASS.ini and hydrology.ini files (also saved here)
  + \*\*\* Should figure out a more streamlined way of doing this for each scenario; maybe calling the Find and Replace code from inside the R notebook and saving to their respective folders?
* Use an updated version of the submitjob.sh file (called “submitjobval.sh”) to rename the CLASS and hydrology.ini files without the trial numbers, run the model, and save the results in a folder for each validation run
* Changed in runoptions:
  + Removed the spin-up period (and number of control flags)
  + Start in 2005 day 258
* Test running with the top 2 only to see if it works; should take about 2-3 minutes or so
  + The run with the best calibration parameters crashed at 2014-132
    - 
    - ZSNL was 0.6 -> this was likely the reason
    - Check if this calibration run actually completed: No, it did not complete
    - Went back and changed the code (in MESHOutput.Rmd) so that only the top 10 runs that finished are represented
    - Had much much better NSE results! In the +ve, and larger than the best run!
  + The run with the 2nd best calibration parameters ran to completion, but the NSE was terrible
    - Likely should put in a spin-up period; consider starting earlier in 2005 (maybe 6 months?) and having a 6 month spin-up period. But will this screw with the initial conditions? Maybe still start Sept. 2005 -> can have a spin-up of 6 months and still catch the spring runoff in 2006

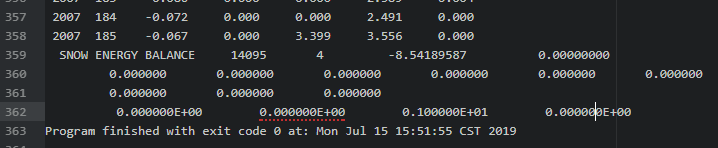


* + Try running again, but with a 6 month (182 day) spin-up period (specified in the runoptions file)

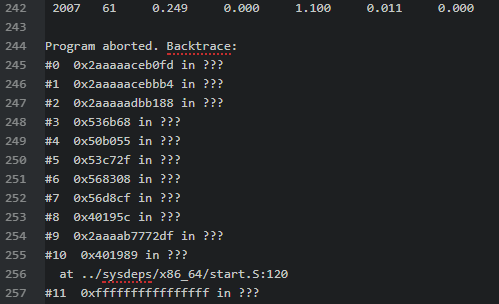
## Scenario 2

1. Run static to make sure it works (Scenario2\_MeshOnly folder)

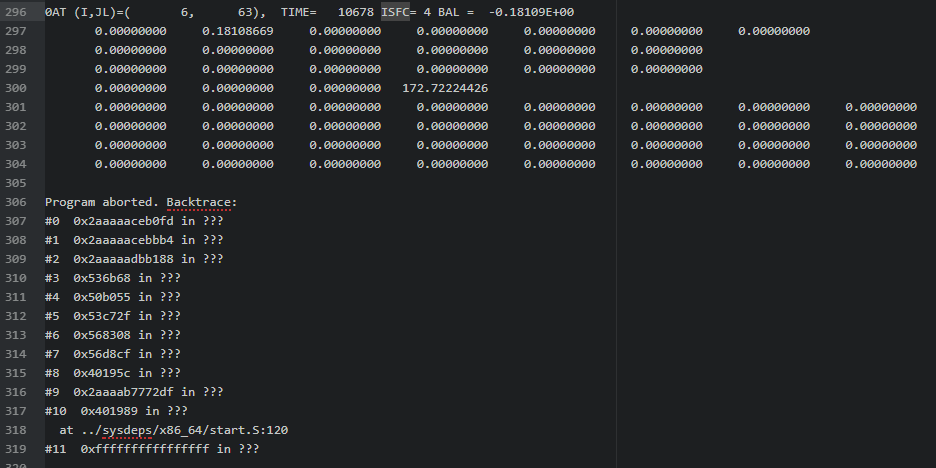
* Ran from 2006-258 to 2008-275 and then stopped; printed “Reached end of forcing data” at the end
* Getting the error: Note: The following floating-point exceptions are signalling: IEEE\_INVALID\_FLAG IEEE\_DIVIDE\_BY\_ZERO IEEE\_UNDERFLOW\_FLAG IEEE\_DENORMAL following 2008-243 result
  + Fixed an error value in the drainage database file
  + Try changing the KSAT of bedrock to >0 (used 1e-9) -> This fixed the INVALID and DIVIDE\_BY\_ZERO flags! Still have the UNDERFLOW and DENORMAL flags
  + For the static run, initially ran with the smallest values of all the cal parameters; tried using the average of the range but didn’t fix the error. Changed the ponding depths to something more realistic, and still didn’t fix the problem
  + Try changing the following (using folder Scenario2\_MeshOnly for mesh runs and Scenario2\_OstTest1 for ostrich tests):
    - MESH only: WFR2 to 2.5: slurm 331455; still ended in 2008
    - MESH only: SDEP reduced to 1m (if was larger): slurm 331456; same error
    - MESH only: DRN: Change the BR value to 0.25 – slurm 331457; same error, very similar results
    - MESH only: DDEN: typically range from <2km/km2 to 100km/km2; try changing DDEN for all GRUs to 1.0 (except Water, which was already at 1.5) – slurm 331459; same error; this increased the amount of overland flow and slightly more interflow
    - Try running Ostrich with 10 iterations to see if it can find a solution (make the KSAT change but not the rest of the changes listed above) -> still had the error where the model only ran unto 2008-Oct-1
    - Try running with MESH version 1552 -> same problem. This tells me it’s not a bug but a problem with the parameterization
    - I noticed that through these trials I have been getting very very low NS values (near zero, + or – a bit)
    - MESH only: Try changing DRN to 0 for peatlands, wetlands, and water since high water table (see the see the CLASS 3.6 manual); slurm 331466; same issue with the model run time; storage changed from a net negative value to net positive and baseflow reduced drastically; NS=-0.062
    - MESH only: Set MID to 0 for the water GRU (inland lake; see the CLASS 3.6 manual); slurm 331467; same issue; really bad NSE
    - MESH only: Try using DDEN=20; slurm 331468; same error
    - MESH only: Try using the full streamflow dataset (no missing values specified); slurm 331475; still error
* Tried parameterizing each GRU the same as a “best” version of Scenario 1 (Trial 27, which was the best of those that had completed), and then gradually start changing parameters until the error shows up again.
  + Giving it a try; folder: Scenario2\_Transition
  + Starting from the “best so far” values from Scenario 1 (trial 27) -> copied the CLASS.ini and hydrology.ini files
  + Copy and paste the CLASS.ini section 5x (1 per GRU); leave all values except the FCAN values; this will essentially parameterize the basin the same as in Scenario 1
  + Used the drainage\_database file I made for Scenario2
  + Hydrology.ini – copied the GRU-dependent parameters for each GRU
  + Same error (ended normally but in 2008)
* Talked to Dan – he informed me that each GRU must have a column in the forcing data files
  + Model runs for the full time!!!!
    - Run time: 41 sec
    - NSD = -7.479
    - Ran with originally-parameterized CLASS.ini file; Snow Energy Balance error: occurs on day 185 (Jul 10?) 2007; why? Run again but write CLASS output files to look into the energy balance data



* + - * Searched on the WIKI and suggested to check the ZSNL values;
        + In the hydrology.ini file, ZSNL was 0.525 and ZPLS was 0.25; changed ZSNL to 0.1 and ZPLS to 0.01 for all GRUs = SUCCESS!
        + However, NSD=-21.03
        + May need to consider lowering the range(s) since the error occurred at the average value
  + Try a calibration run using OstTest1 folder
    - Try reducing the max. range of ZPLS to 0.25
      * Caused this error:

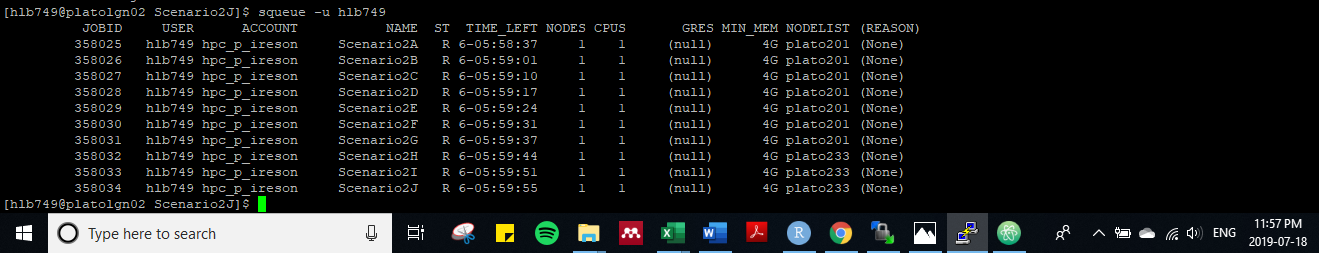


* + - Change ZPLS back to 0.5 and reduce the max ZPLG-WL and ZPLG-W to 1
    - Changed DRNROW for peatland back to 1.0 (from 0.0)
    - Is there an easy way to use Green Kenue to calculate the drainage density for each GRU?
    - Found another error when running OstTest1 (10 iterations, 1 trial):

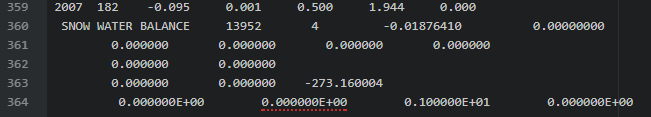


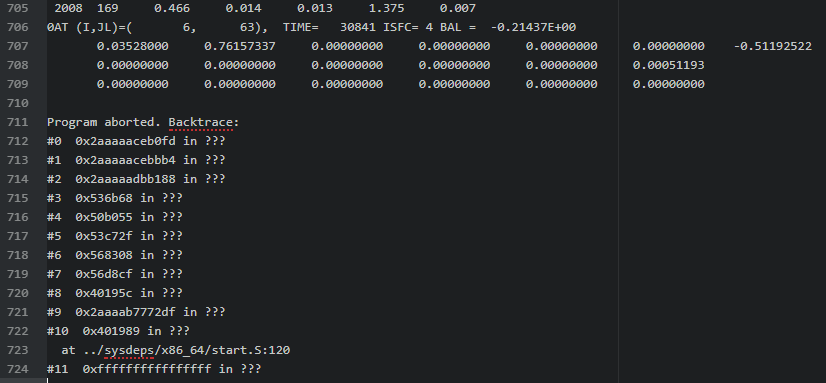
* Try running MESH alone (not Ostrich) for the last-run param set.
  + Same error (of course).
  + Plotting the water balance in R:
    - P-E-R-deltaS is approx. steady at 25mm… but should be about zero
    - Cumulative storage follows the cumulative precip, which makes sense for the winter since snow is accumulating and virtually no runoff
    - Looks like there’s negative evap in December 2006??
    - From the hydrograph, there is no simulated streamflow in the winter, even though there is recorded streamflow during that time (~0.25m3/s)

1. Do a time trial for 1000 iterations, whole time period, 1 trial to calculate full run time (OstTest1 folder)
   * Took about 2 hours, but had lots of errors/aborting, and the NSE didn’t change from 1 iteration to the next even though the params did (likely because didn’t run the model)
     + Error running; last MESH run showed a SNOW ENERGY BALANCE error; reduced max ZSNL to 0.7 and ran again but for 10 iterations
   * Run again using a tiedNSD (0.5 negNSD, 0.5 neglnNSD)
     + Took 4286s (1.2 hours)
     + Snow energy balance is causing issues again; need to look into this; could be the ZSNL or ZPLS values, or maybe the initial soil temp. values?? Maybe print CLASS output for more info? Slurm 351449
       - Explore the snow energy balance; need to re-run to print CLASS outputs (re-run just the last iteration)
   * Talked to Dan about the snow energy balance errors and he suggested to check the OstModel0 output -> if a large number of iterations are failing (i.e. the “dummy” NSE values are written to the file” then there could be a problem with the model; otherwise it’s probably just the model finding the edges of the valid model runs
   * Realized I had put very good instead of very bad dummy values in the metrics out part of the runmesh.sh script; corrected in the F: drive, RunMesh\_MeshOnly, and RunMesh\_OstTest1 and 2 folders
   * Ran the 1000 iteration 1 trial test again (OstTest2, slurm 353942)
     + Time: 6hrs 11 minutes
     + Model aborted 57 times
   * For the next run (Scenario2\_2):
     + Adjusted some of the parameters of the model (see the Scenario2\_2 tab in the Calbiration Details excel file)
     + Removed the metricspinup option from the runoptions file
     + Increased the DDEN of Conifers, Wetlands, and Water to 2 ( since most of the GRU areas have a blue line channel running through them)
     + Set DRNROW of wetlands, peatlands, and water to 0.25
     + MID value for the Water GRU is zero since it’s an inland lake
     + Ran the MESH\_only version -> good, ran
2. Set up and run the full calibration

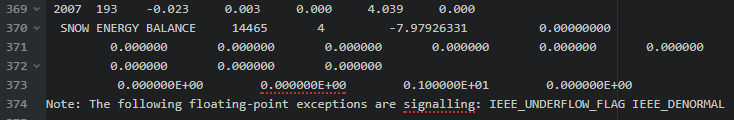


* + Had to restart Scenario2G since missed copying the BASINAVG1 folder into the directory
  + Results:
    - Of the 100 trials, only 8 ran at least to the end of the calibration period, 7 ran to the end of the simulation range
    - Max NSE (tied; 0.25 NSE, 0.75lnNSE) of the trials that ran the full cal length was ‑0.247
    - The errors present in the OstExeOut.txt files that remain (every 10th trial) are:

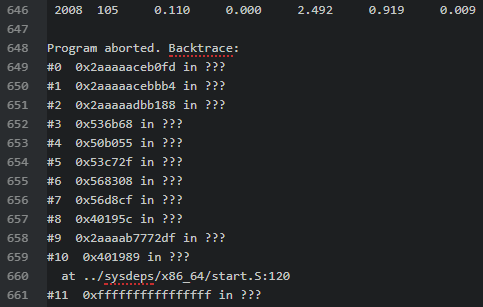
10. 

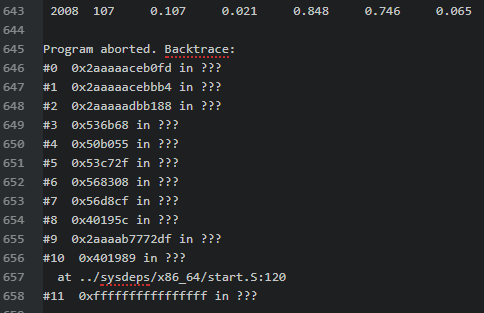
20. 

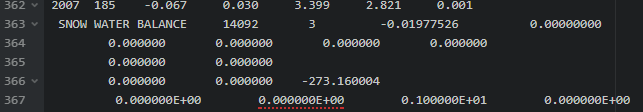
* Re-Run this setup in the folder “Scenario2\_Debug” using the debug version of MESH
  + Slurm 366706
  + Produced the same error
  + Only additional information was that the program aborted during running the program (line 7 of ./run\_mesh.sh)
  + Still ???

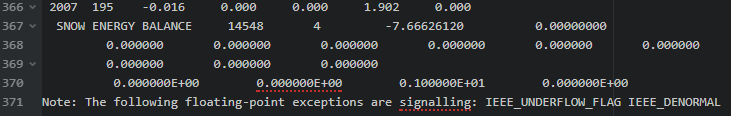
30. 

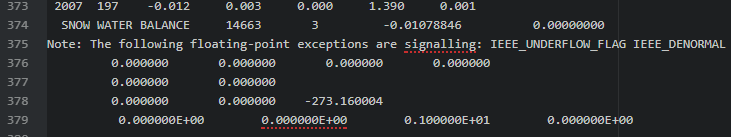
* Re-Run this setup in the folder “Scenario2\_Debug” using the debug version of MESH
  + Slurm 366708
  + Same error, no additional information provided from the debug version
  + Still ???
* Run again, this time print CLASS outputs for GRUs 1,3,4,5,6
  + Slurm 366709
  + Only 1 set of CLASSOUT files printed (need a folder for each GRU???)
  + CLASSof6

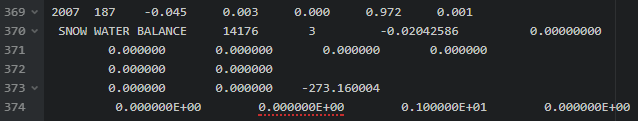
40. 

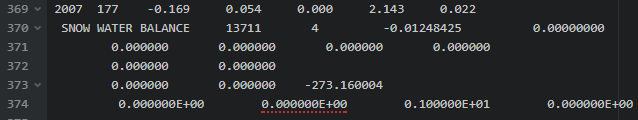
50. 

60. 

70. 

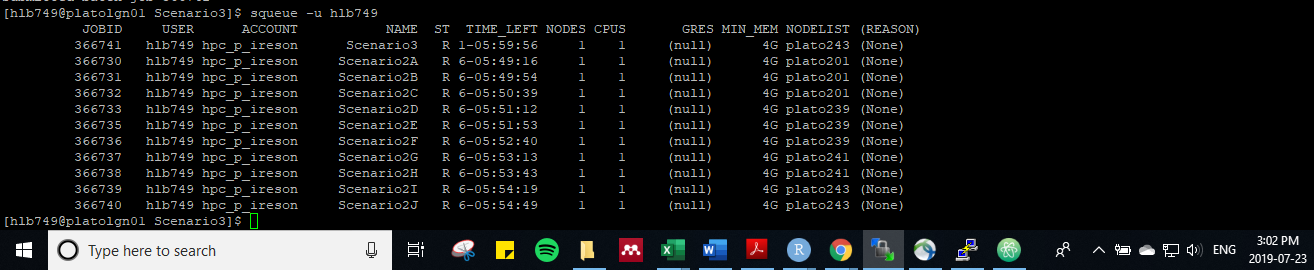
80. 

90. 

100. 

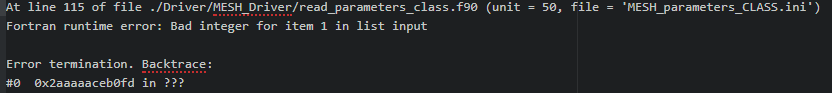
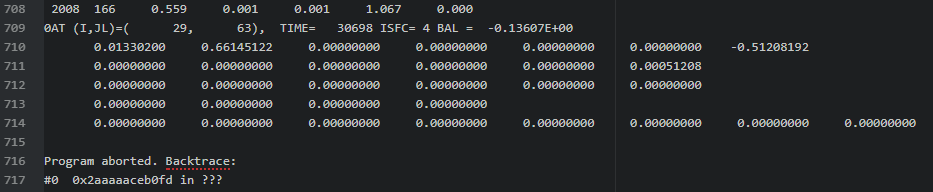
* The snow energy balance errors are occurring between late June and mid-July

1. Re-running Scenario 2
   * Original run files and results saved to folder “Scenario2\_Calibration\_Attempt1”
   * Calibrating to NSE only
   * Updated the following items:
     + ostIn:
       1. Change the model executable to run\_mesh.sh (rather than the direct path)
       2. Change ZSNL limit to 0.001-0.2
       3. Change the response variable to negNSD and remove tied parameters section
     + run\_mesh.sh
       1. ensured the regular (not debug) version is being used
       2. change to executable
     + run options:
       1. Changed the SOILINIFLAG to 1 (from 4)
   * Use the version that has the Snow Energy Balance error resolved (ZSNL 0.001-0.2)
   * Copy the OstExeOut.txt file to each trial of Ostrich as well

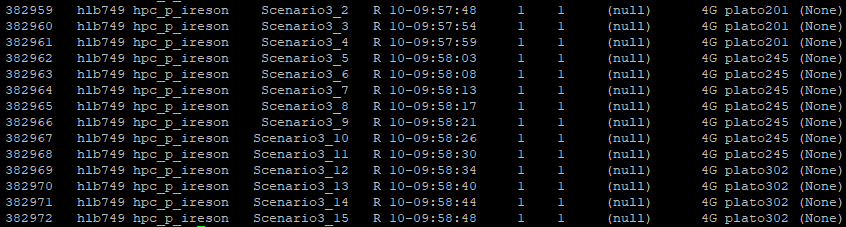


* + The model stopped running at about 12:30pm on July 24 because I accidently copied the Scenario 3 CLASS.ini and CLASS.tpl files into here this morning, and when it tried to start running again the number of grids was wrong; Corrected the files and started running again (changed to start at trial 003 in the submitjob.sh file)

## Scenario 3

* See the MESH Training document for details on how to create the r2c drainage database for the gridded setup
* Create Scenario 3 (Scenario 3 folder) by:
  + Copying the last run of Scenario 2 over to a new folder and run with mesh only to make sure it works -> good; slurm 358035, took about 34 seconds to run
  + Swap out the r2c file for the gridded version and change the number of grids in the CLASS.ini and .tpl files from 2 to 8 -> slurm 358036, took 2min 13 seconds (about 4.5 times as long)
  + Run with Ostrich, 5 iteration (expect to take about 15 min) -> slurm 358037, 3 “aborted” lines, and the last run of MESH ended with a backtrace error. All the MESH runs had the same response variable values, but most of the calibration parameters were different. Maybe it’s ok? Maybe not? Try running the 1000 iteration time trial overnight anyway since going to bed soon
  + Finished the 1000 iterations, but had the aborted and backtrace errors again
    - Ran with debug version that Dan made on Wednesday. Said there was a “wf\_qo2” error; added the metricspinup back into the run options file
    - Ran again; the model ran to completion this time, but the response variable didn’t change during calibration
    - Noticed that the drainage\_database file may have some errors
      * Rank and Next have decimal values; changed to integers
      * There was no drainage area listed for the cell containing the outlet; changed the number to 155.0000. Bankfull also showing the “outside the basin” value of 0.100 in the last cell (8); maybe ok? Try running
      * Try running again; still not changing the response variable; noticed that there is an “INVALID\_FLAG” in the mesh output; the wiki says this is either due to bad parameterization or bad drainage values
      * Changed the Bankfull value for cell 8 to the same value as cell 7 and tried running again (Mesh only)
* Had errors when calibrating Scenario 2 due to a snow balance error because ZSNL was too high (shouldn’t be larger than about 0.3 or 0.2).
* Also learned that I was running MESH by directly pointing to the executable from ostIn instead of pointing to run\_mesh.sh
  + Need to make run\_mesh.sh executable (by right-clicking and selecting the x’s in “properties” or using chmod +x from the command line
  + Changed the line in ostIn so it calls run\_mesh.sh
  + Essentially when the model crashed, it was keeping the same NSE value rather than reading an arbitrarily small value to penalize the model for selecting error values
  + Start running again around 5pm on July 23
  + Based on the status (OstStatus0.txt) the next morning, the calibration (1000 iterations, 100 trials) will take about 55 hours to complete; only gave the model 30 hours to run; likely won’t run to completion but at least I’ll get some results
  + Still may have the error with the r2c file, so this one likely isn’t useful anyway. Also, so far the NSE values aren’t changing for a number of the iterations. Also realized I was running the debug version, but should be ok, but changed the run\_mesh.sh file to run the normal version of MESH after about 332 iterations (morning of July 24)
* Talked to Dan about the r2c file. He showed me how to modify the grid so that the outlet is actually outside the basin and isn’t including a large portion of the watershed in that cell (and isn’t excluding a lot of drainage area
* Created a new r2c file to make sure the outlet cell was not being considered as a large part of the basin
  + Manually adjusted the xdelta and ydelta, as well as the origin to place the outlet in a good spot. Also had to manually include the middle-easter cell in the basin; manually adjusted the FRAC for a few of the cells so that drainage area was approximately correct. Checked the drainage directions
* Tried running (Scenario3\_newr2c folder) and got this error:
* 
  + Checked the class.ini and class.tpl files, re-copied over from the version of Scenario 3 that’s running, and still the error
  + Check the r2c file; realized there were no GRU-specific parts of the file because I forgot to map the landcover types to the map file. Go back and fix.
* Running MESH once, and got a water balance error:
* 
* Try calibrating with 1 trial, 5 iterations to see if we get any successful results
  + Best NSE: -1.09
  + Time to run: 9 min, 50 sec. If the same time, 1000 iterations will take 33.33 hours
  + Slurm: 381632
* Set up and start the 1000 iteration time trial
  + Gave it 60 hours to run
  + Slurm: 381637
  + Completed with only 2% crashes, optimal NSE=-0.145
  + Took 31 hrs 44 min;
  + For the full run, assume each trial will take 35 hours

### Full Run

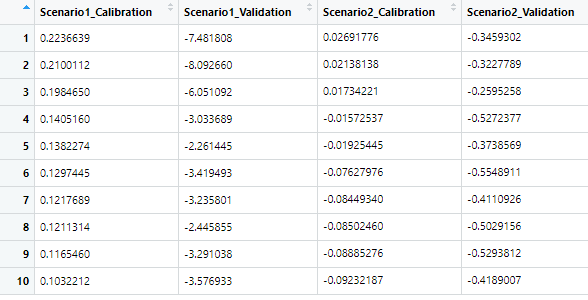
* Started at about 9:40 am on July 26th, divided into 15 runs (1 run of 2 trials, and 14 runs of 7 trials). Should take about 10 days to run.
* 
* 

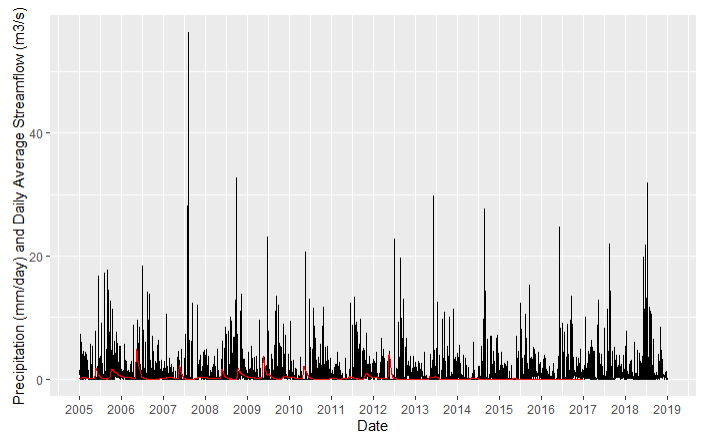
## PDMROF

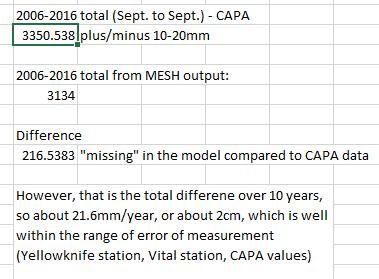
* Setting up trials using the parameterization of Scenario 1 and 2 but using PDMROF instead of WATROF (IWF set to 2 instead of 1)
* Discussed with Bruce end-of-day Tues. Aug 6
* Took a bit to get the parameterization figured out and the model files set up. Tried to start running Scenario 1-P on Wednesday Aug. 7, but because Plato was schedule to be down for maintenance on Friday Aug 9, they wouldn’t start until after
* \*\*\*Plato was down for maintenance from Friday Aug. 9 through about 3:30pm on Tuesday Aug. 13\*\*\*
* Started running the 10-subfolders of Scenario1-P and Scenario2-P by 5:15pm on Tues Aug. 13
  + Only calibrating CMAX and B; CMIN, K1, and K2 set to zero (as per Dan’s email)
  + Using r1552 (new point-mode MESH code)
  + Using a modified submitjob.sh and runmesh.sh script to reflect the differences compared to r1024

## Results Reporting

### Follow-up to Results Update Meeting – Aug 1

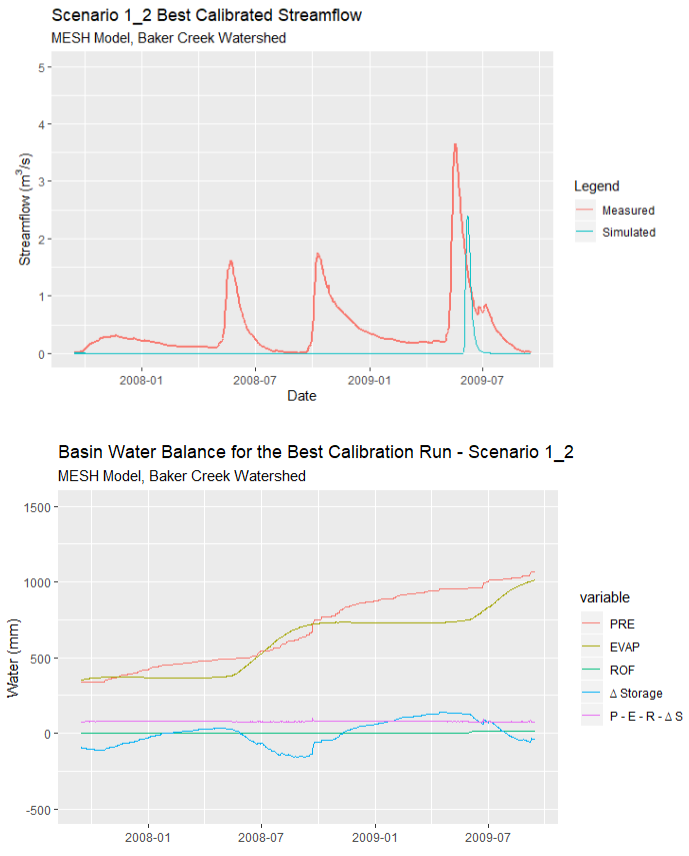
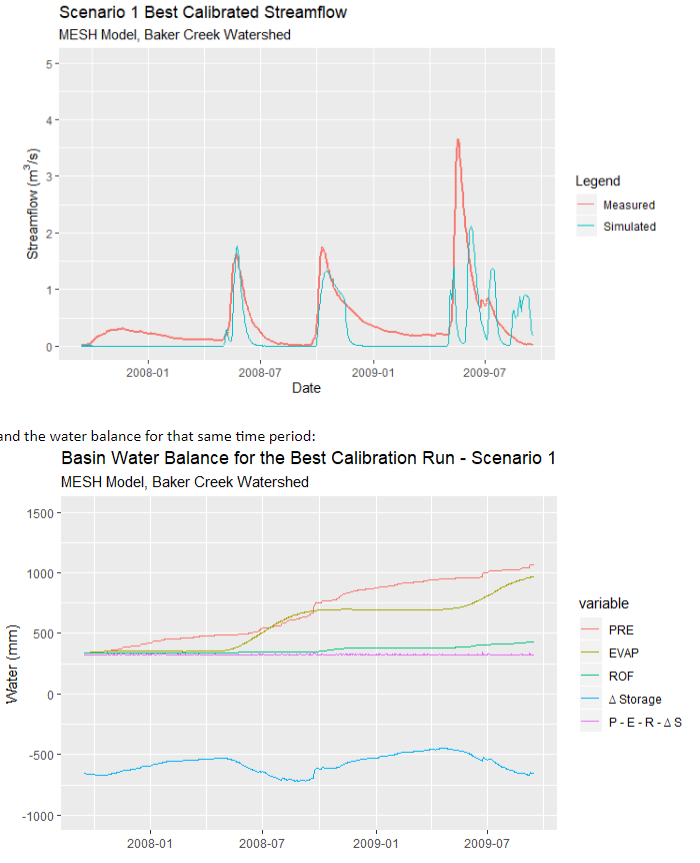
* Met with Bruce to present the results from Scenarios 1 and 2
* What was presented is saved here: [MESH Model\Baker Creek Model Files\Baker Creek - MESH Output Visualization and Analysis - Aug 1 2019.html](MESH%20Model/Baker%20Creek%20Model%20Files/Baker%20Creek%20-%20MESH%20Output%20Visualization%20and%20Analysis%20-%20Aug%201%202019.html)
* Action items:
  + Q. How often did the 1st calibration of Scenario 1 crash? Would that affect the results?
  + Q. For the re-run of the Scenario 1 calibration (purpose being to incorporate the “dummy metrics on crash” script): is the full range of parameters being represented consistent with Scenario 2?
    - A. For the most part, yes. Changed the upper bound of ZPLG and lower bound of MANN to represent the full-range of Scenario 2 ranges. Also removed the “Metricspinup” flag (still have -ve streamflow for 2006 to create a spin-up period). Updated these values in the Scenario 1 re-run and started again around 4pm Aug 1; separate into 10 folders (instead of 5 for the initial run) so it will be done in 2.5 days instead of 5 (early Sunday finish instead of end-of-day Tuesday) One significant item of note is the vast difference in the parameterization of the soils between scenarios 1 and 2; May want to explicitly talk about the rationale for this and possible effects in the report.
    - Started running the re-run of Scenario 1 (with the changes above plus using the run\_mesh script) at 9:38 pm on August 1st.
    - Q: What were the sand, silt, clay, and OM values from the Scenario 1 results?
    - A: Took a spatial-weighted average of the soil texture characteristics used in Scenario 2; This may not be the right approach; need to look into what the model does with the soil texture
  + Q. Change the calculation of the validation NSE to reflect only the period from 2006-258 onward; use the output of calibration to calculate
    - A. The new NSE results are as follows:
    - 
    - Validation NSE are all less than Calibration NSE; For Scenario 1, the validation NSE is an order of magnitude larger and negative as compared to the calibration; makes the box plot and run time plot look wonky
  + Q. From the hydrographs, it appears that the simulated streamflow does not capture the fall “peak” events. Is there precipitation in the driving data corresponding to these events? If not, then it appears we have a case of “garbage in – garbage out”. If yes, then what hydrological processes are not being well-represented in the model? \*\*\*Talk to Chris about this \*\*\*
    - A. Yes, it seems that there is precip around when fall streamflow occurs.



* + - Q. Why so much fall/winter streamflow in 2008 and 2011, and not in other years? See Spence et al 2010 for the 2008 rationale (in short, its because storage capacity was full late in 2007, hence the winter streamflow through 2007/2008), and then heavy rains in September 2008
    - Check the driving precip data -> make sure you can work back to the original data (i.e. CAPA); Yes; the yearly precip calculated from the driving data (see MESHcheck.Rmd doc) is approx.. equal to the CAPA annual precip, and the total for the modelling period of 3134 mm (MESH echo output) is approx. equal to the CAPA data
    - 
  + Q. Calculate and summarise the % crashes for each trial for each scenario (may not report, but check to make sure the model is running properly)
  + Q. Are there any other process-related (not core model-related) reasons for the poor performance?
    - Driving data?
    - Model setup?
      * Soils?
      * Other?
  + Q. Why is there no initial peak streamflow at the start of Scenario 2 when there is for Scenario 1? (look at the initial conditions, esp. soil moisture)
  + Q. Plot the ranked NSE values vs trial (same as in Herbert’s report; one plot for all scenarios).
  + Q. Change the limit of the y-scale on the hydrographs
* Once have the answers to these, email Bruce and the meet up again early next week to discuss.

### Follow-up to the Follow-up

Comparing the results of Scenario 1 and Scenario 1\_2 (re-run which reduced the range of ZSNL, slightly modified the ranges of KSAT and MANN, and used the run\_mesh.sh script to steer the model away from crashes)

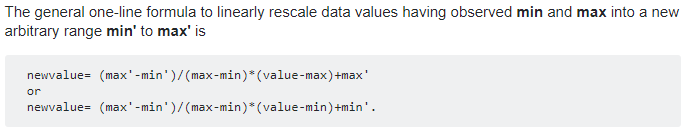


* What was different about Scenario 1\_2 that:
  + Made the model miss the peaks in spring and fall 2008
  + Result in the net water balance being around 300mm initially (and initial deltaS around -700mm)

### Identifiability:

* “a measure of how uniquely a parameter is found by the calibration process” (Herbert)
* Need an original reference (ask Andrew)
* Alternatively, an unidentifiable is when “two or more parameterizations are observationally equivalent”

The way it was calculated in Herbert’s report:

1. Normalize the data so that each parameter’s values range from 0 to 1
   1. Normalization equation: 

<https://stats.stackexchange.com/questions/70801/how-to-normalize-data-to-0-1-range>

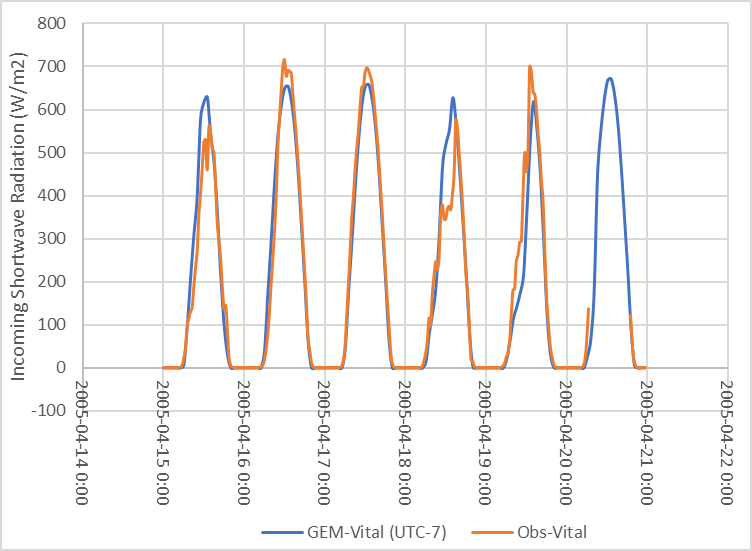
1. Identify the 10th and 90th percentile
2. The span of those percentiles (difference between the 90th and 10th percentiles) is the Normalized Parameter Range
   1. This is essentially the normalized span of 80% of the optimal values found during calibration
3. Set a “threshold” value, below which a parameter is said to be identifiable, and above which it is not.

## Dataset Times

#### **GEM/CaPA:** is in UTC/GMT, Yellowknife standard time is UTC-6

**Vital tower:** The dataset starts April 12, 2005 at 08:30. Checked the Kin values for 24 hours on April 12-13 2005 and K peaked at 12:30 pm. Therefore, the dataset is in local time, which is Montain Daylight Time (MDT), UTC-6

Comparison between GEM UTC-7 and Vital SWin:



Landing tower:

## June 3 – Setting up input files

#### Soil levels

* Kept the soil thicknesses the same as Herbert’s report: 0.10, 0.25, and 3.75 m

#### Hydrology.ini

* Not sure what to select for WF\_R2… this will be a calibrated parameter again
* Kept the ZSHN, ZPLS, and ZPLG the same as Herbert’s
  + Note: it looks like he selected these values based on the “sloped-calibrated” values from the Davison 2016 report; may need to vary for Baker Creek, keeping in mind that bedrock/water makes up over 60% of the landscape area

#### Streamflow

* Going to include the full dataset from 2005 through 2018, knowing that the model will only use the data for the modeled time period (check on this)

#### Run Options

* Still need to decide on:
  + Driving data timestep (HOURLYFLAG)
  + Simulation run times
    - Should be the start of the hydrologic year; need to select based on the basin
* Will use .csv files for each type of forcing data
  + Or use a met file? See the “interim release” notes update for the BASIN\*FLAG, as well as the ojp point example (provided by Dan)

## May 30 – Run Herbert Scenario 1 Validation

Use the files from the folder “SLOPED\_VALID\_1Grid-1GRU\_AverageForcing\_1” including:

* + The 7 .csv driving data files
  + The drainage database (r2c)
  + Input\_reservoir
  + Run options
  + Soil levels
* For the CLASS and hydrology parameters, modified the .tpl file (template file for Ostrich) using an R script that reads a csv file containing the best parameter values (created by copying the optimal parameter set from the “OstOutput0.txt” file from the validation folder) and replaces the value tags with the best parameter values
  + \*\*\* Note that for scenario 1 (not sure about the rest of them), the ostrich output file with the best parameters used in the validation run is located in the validation folder, NOT the calibration folder
  + Haven’t figured out how to make R preserve more than 3 significant figures when it does the substitution
  + Checked the .ini files and the parameters were all replaced successfully!

#### Checking the Files

* CLASS and hydrology .ini files:
  + Compared the “best parameter” .ini files with the .ini files in Herbert’s validation folder
  + There are fewer decimal points in the file I created using R, but the values are right
* Driving Data:
  + Confirmed that the driving data is for the full cal and val period
* Drainage Database: ok
* Run options:
  + Half hourly time step, IWF=1, runs from 1999-274 to 2014-364
* Soil levels:
  + 3 levels for the whole basin (1 GRU)
* MESH\_parameters\_CLASS.ini
  + Best value parameters checked above, start date of met data assumed correct
* MESH\_parameters\_hydrology.ini
  + Best value parameters checked above, the rest is assumed correct

#### Running the Model

* Run the model on Andrew’s server
  + Used MESH 1024 (the version Herbert used), compiled using gfortran

Run time: 7.44 seconds Results:

* + NSE values for cal and val periods don’t match what’s in the report
  + Water balance: Haley’s results are nearly double Herbert’s file (based on the MESH\_output\_echo\_print.txt file
  + Will try running the model with all of Herbert’s original files
* Ran Herbert’s model with ALL his original files in this folder
  + The “MESH\_output\_echo\_print.txt file” in the main file was the same as before, but the one in the BASINAVG1 folder matched my 1st trial results; therefore, the file in the main folder must be an old version from another trial
  + Slight difference in flow between the two versions (Herbert’s original vs. version where I used a script to insert the best values) but not significant.