**MBL Ensemble Approach Documentation**

*Example Folder*

/mnt/data2/disk1/soilcarbon/crivard/predEnsemble/ensemble\_example/

^All scripts have setwd() to this folder so you can change to wherever you copy it

*Scripts*

* **setname\_prep.R**
  + Performs the calibration transfer on the spectra and saves as RData file in ‘spc’ folder
  + Change the input csv file, the columns being selected as spectra (lines 11-12), and output name/location
* **setname\_oc.R**
  + Submit as a job through cloudops, creates all the mbl models with different parameter combinations, to output/oc folder
  + Change input validation and calibration sets (line 32-38), property (oc) throughout the file, output location (line 107) and create output folder for soil property
* **setname-fratio.R**
  + Calculates the fratio for all samples in the calibration and validation sets and outputs a list of outlier indices from the combined dataset. Ex: calibration set indices are 1-15000, validation set indices are from 15001-15240
  + Change input calibration and validation spectra (5-6 and throughout), number of directories in line 8 as needed, output location- currently ‘fratio’ subfolder.
* **setname-extract.R**
  + Creates comprehensive files containing all predictions for each mbl model by property. (ie. pred.oc.csv, pred.bd.csv)
  + Creates a file containing the lower, mean and upper prediction estimates for each property across all models (all-predictions.csv)

*Getting Started*

I would start by copying this folder into one of your directories and changing the setwd() path at the top of each script. I kept output examples for you to check out but you should delete them before running the scripts yourself:

* ./spc/rod.pds.spc.RData ← created by rodale-prep.R
* ./output/oc/mbl.waspls1.euclid.none.RData ← created by rodale\_oc.R
* ./output/oc/mbl.waspls1.cosine.none.RData ← created by rodale\_oc.R
* ./fratio/pca.oc.RData ← created by rodale-fratio.R
* ./fratio/oc.csv ← created by rodale-fratio.R
* ./output/pred.oc.csv ← created by rodale-extract.R
* ./output/all-predictions.csv ← created by rodale-extract.R

*MBL Documentation* (New Zealand guide)

<https://whrc.github.io/Soil-Predictions-MIR/mbl-models.html>