Instruction for using codes:

* PCA\_Comp\_Analysis\_Dredging
* Multicollinearity\_and\_Variable\_Ranking
* Model\_Combos\_and\_Reconstruction\_Wet
* Model\_Combos\_and\_Reconstruction\_Dry

NOTE: these codes automatically output several hundred files to local folders. Before getting started, it would be helpful to make and organize a local directory as follows:

* SD
  + rainfall\_data1
  + Predictor\_Variables2
  + State\_map\_info3
  + Composite\_Plots
  + DFs
  + Dredge\_Tables
  + Residual\_Plots
  + RSQ\_Tables
  + Correlation\_Matrices
  + VarLists\_wet
  + VarLists\_dry
  + LOOCV\_WET
    - Predict\_PC1
    - Predict\_PC2
    - Predict\_PC3
    - Predict\_PC4
    - Reconstruction
    - LOOCV\_R2
    - RMSE\_MAE
  + LOOCV\_DRY
    - Predict\_PC1
    - Predict\_PC2
    - Predict\_PC3
    - Predict\_PC4
    - Reconstruction
    - LOOCV\_R2
    - RMSE\_MAE
  + Actual\_Rain
  + Fits\_for\_Reconstruction
  + Fits\_All\_Combos\_Wet
  + Fits\_All\_Combos\_Dry
  + Fitted\_Precip\_All\_Combos\_Wet
  + Fitted\_Precip\_All\_Combos\_Dry

1contains csv file of wet season rainfall data and csv file of dry season rainfall data

2contains 16 netcdf files of predictor variables

3contains Hawaiʻi state outline (coast\_geo\_shp) and all supporting files

1. Start on code *PCA\_Comp\_Analysis\_Dredging*
2. Run lines 7 – 270
3. On line 279, set **season<-'wet'**
4. Change folder name on line 304
5. Change folder name on line 362
6. On line 380, set **pcamode=1**
7. Change folder name on line 425
8. Change folder name on line 470
9. Run all code lines 274 – 495
10. On line 498, change season and PC in file name and run

*This code will be repeated 7 times, changing the season and pcamode each time. Before moving on to the next PC (step 19), switch to “Multicollinearity\_and\_Variable\_Ranking” code to eliminate variables, or continue on to dredge code starting on line 503 (step 16) (where variable lists have already been edited based on variable elimination).*

*To assess multicollinearity:*

1. Open Multicollinearity\_and\_Variable\_Ranking code
2. Change folder name on line 26
3. Change working directory on line 38
4. Run code lines 14 – 58
5. To save correlation matrix, manually save to folder titled Correlation\_Matrix and title with season and PC number

*Steps 11 through 15 produce a ranked r-squared list of the predictor variables and a correlation matrix. The correlation matrix must be manually saved to the “correlation\_matrix” folder. The correlation matrix is visually assessed and lower ranked variables in pairwise correlation having a correlation value of 0.9 or higher are removed from the variable list. To make dredge table, return to PCA\_Comp\_Analysis\_Dredging code. In the existing code, the variable elimination has been already done, and it is possible to skip directly from step 10 to step 16.*

1. Run line 514
2. Change file name on line 537 to correct season and PC
3. Run lines 535 through 537

*Repeat code for PC2*

1. On line 380, set **pcamode=2**
2. Run code lines 380 – 495
3. On line 498, change PC in file name and run
4. Repeat steps 11 through 18, changing step 16 to line 516.

*Repeat code for PC3*

1. On line 380, set **pcamode=3**
2. Run code lines 380 – 495
3. On line 498, change PC in file name and run
4. Repeat steps 11 through 18, changing step 16 to line 518.

*Repeat code for PC4*

1. On line 380, set **pcamode=4**
2. Run code lines 380 – 495
3. On line 498, change PC in file name and run
4. Repeat steps 11 through 18, changing step 16 to line 520.

*Repeat code for dry season.*

1. On line 279, set **season<-'dry'**
2. On line 380, set **pcamode=1**
3. On line 498, change season and PC in file name and run.
4. Repeat steps 11 through 18, changing step 16 to line 524.

*Repeat code for dry season PC2.*

1. On line 380, set **pcamode=2**
2. Run code lines 380 – 495
3. On line 498, change PC in file name and run
4. Repeat steps 11 through 18, changing step 16 to line 526.

*Repeat code for dry season PC3.*

1. On line 380, set **pcamode=3**
2. Run code lines 380 – 495
3. On line 498, change PC in file name and run
4. Repeat steps 11 through 18, changing step 16 to line 528.

*Repeat code for dry season PC4.*

1. On line 380, set **pcamode=4**
2. Run code lines 380 – 495
3. On line 498, change PC in file name and run
4. Repeat steps 11 through 18, changing step 16 to line 530.

End of code.

Open code titled *Model\_Combos\_and\_Reconstruction\_Wet.*

1. Run line 5
2. Change directory line 13.
3. Change directory line 26.
4. Run code line 8 through line 63
5. Change directories lines 67 through 70
6. Run lines 67 through 70
7. Change directory line 88
8. Run lines 77 through 89

*Return to code PCA\_Comp\_Analysis\_Dredging. Change line 279 to “wet” and run lines 14 through 350 to remake the Xmatrix object.*

1. Return to code *Model\_Combos\_and\_Reconstruction\_Wet.*
2. change directory line 108
3. run code lines 102 to 125
4. change directories lines 138,148,149
5. run lines 132 to 154

*The next steps are to perform leave one out cross validation*

1. Change directories lines 177, 186, 208, 222, 236, 252, 273, 279, 285, 291, 331, 350, 359, 360
2. Run lines 173-293

*Return to code PCA\_Comp\_Analysis\_Dredging. Change line 279 to “wet” and run lines 14 through 350 to remake the Xmatrix object.*

1. Run lines 300-365

*The next steps are for calculating the median r-squared value per station*

1. Change directories lines 374, 383, 384, 385, 390
2. Run lines 373-394

Open code titled *Model\_Combos\_and\_Reconstruction\_Dry.*

1. Run line 5
2. Change directory line 13.
3. Change directory line 26.
4. Run code line 8 through line 63
5. Change directories lines 67 through 70
6. Run lines 67 through 70
7. Change directory line 88
8. Run lines 77 through 89

*Return to code PCA\_Comp\_Analysis\_Dredging. Change line 279 to “dry” and run lines 14 through 350 to remake the Xmatrix object.*

1. Return to code *Model\_Combos\_and\_Reconstruction\_Dry.*
2. change directory line 108
3. run code lines 102 to 125
4. change directories lines 138,148,149
5. run lines 132 to 154

*The next steps are to perform leave one out cross validation*

1. Change directories lines 177, 186, 208, 222, 236, 252, 273, 279, 285, 291, 331, 350, 359, 360
2. Run lines 173-293

*Return to code PCA\_Comp\_Analysis\_Dredging. Change line 279 to “dry” and run lines 14 through 350 to remake the Xmatrix object.*

1. Run lines 300-365

*The next steps are for calculating the median r-squared value per station*

1. Change directories lines 374, 383, 384, 385, 390
2. Run lines 373-394