1. Run\_cleanLidarProfiles.py – does a lot of removal/smoothing of weird data in profiles
   1. Final acceptable profiles saved as cleanLidarProfiles.pickle; obj = [lidar\_xFRF,time\_fullspan,final\_profile\_fullspan\_best]
2. Run\_blendLidarBathy.py – combines best-available lidar and available bathy surveys and fills nans; blend between lidar and bathy overlap of 0.5m
   1. Lidar comes from cleanLidarProfiles.pickle
   2. Bathy comes from tidalAveragedMetrics.pickle
   3. Combined output saved to blendedLidarBathy.pickle
3. Run\_prepDatasetsForML\_V2.py – fill gaps in bathy and hydro assuming datasets will be 60 hrs long
   1. ZbFull\_LidarBathyBlended gets written into data\_fullspan\_addBlendedLidarBathy.pickle
   2. Save data in “dataset” structure” as topobathyhydro\_ML\_final\_25Mar2025\_Nlook60.pickle
4. Run\_PCA\_ExtendedDatasets\_V2.py
   1. Shifts profiles in x-shore (slightly) so they all start at x = 0, z = 6
   2. Performs PCA
   3. Creates shifted elevations, non-shifted elevations, PCs, and hydro datasets as full-span (t = 0:10yrs) so length of ML input data can be varies
   4. Checks that dVol between time steps is small
   5. Saves datasets to topobathyhydro\_ML\_final\_25Mar2025\_Nlook60\_PCApostDVol\_shifted.pickle
5. Run\_ML\_multivariate\_variableNLook.py