**Methods**

Three metrics were used to compare the performance of the three candidate bathymetry raster layers against the observed station depths. Percent bias (PB) of each bathymetry layer relative to the observed station depths were calculated and the mean and/or median of the distribution of PB indicates any systematic errors in the bathymetry raster values:

Mean absolute error (MAE) and root mean square error (RMSE) characterize the magnitude of the error between the bathymetry raster values and the observed station depths:

: predicted bathymetry (units: m) at station from proposed bathymetry layer

: observed bathymetry (units: m) at station

: total number of stations

**Results**

The min, median, mean, and maximum observed depths were 11 m, 130 m, 157 m, and 984 m, respectively, with the majority of observed station depth values shallower than 400 m (Figure XA). Station depths predicted from the mixed raster layer (“Mix”, Figure XB) had the best performance across all metrics, with a mean PB of -0.2%, median PB of -1%, MAE of 4.3, and RMSE of 9.2. The GEBCO layer performed similarly to the “Mix” layer (Figure XC) with mean and median PB of -3%, MAE of 5.1, and RMSE of 10.6. The ARDEM layer performed the worst of the three layers (Figure XD) in terms of relative error with an MAE of 12 and RMSE of 30 but a moderate level of bias with a mean PB of 1.6% and median PB of -1.4%. The RMSE metric is more sensitive to outliers, explaining the much larger RMSE of the ARDEM bathymetry layer compared to the GEBCO and “Mix” bathymetry layers.

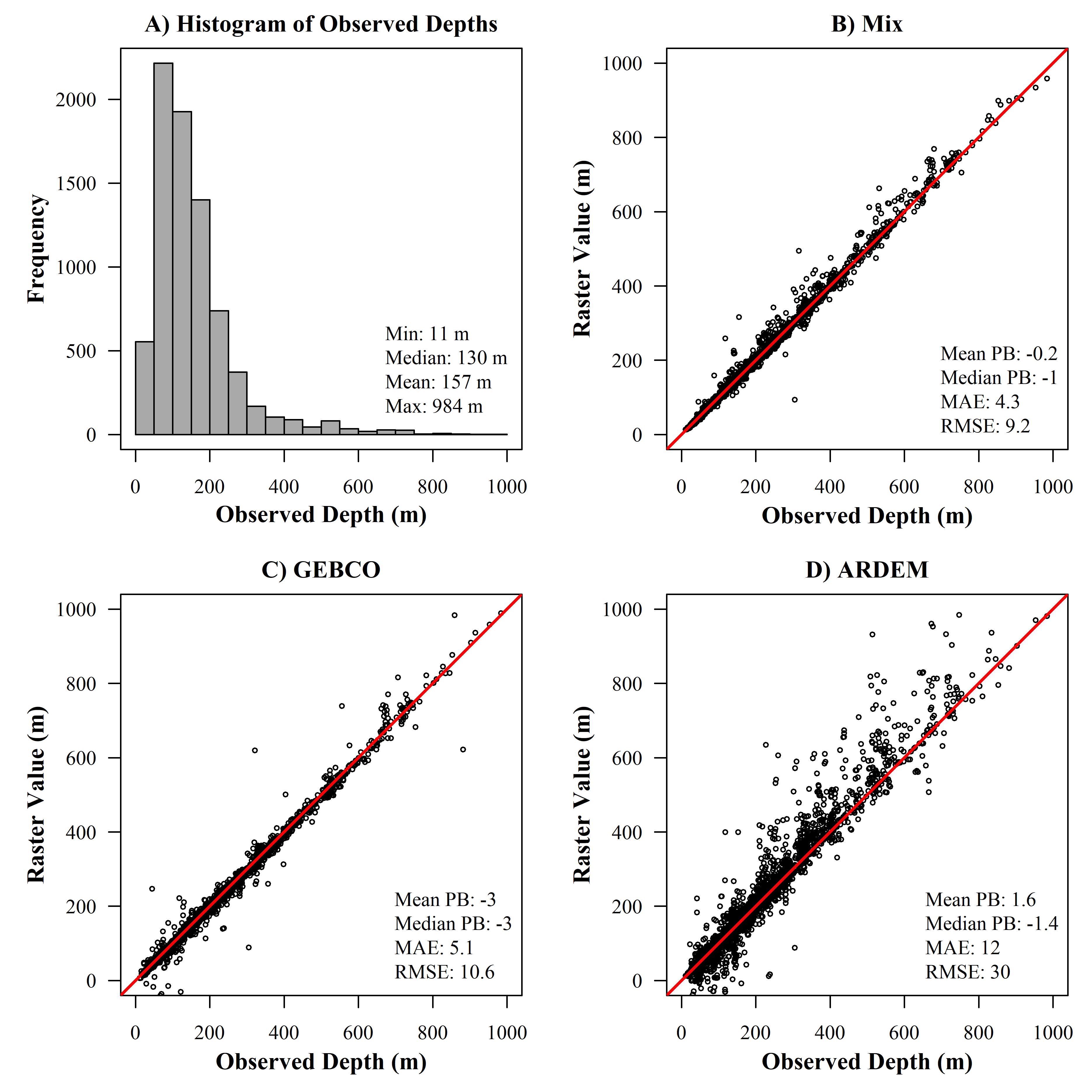


Figure X: A) Distribution of included observed station depths from the Gulf of Alaska bottom trawl survey from 19XX – 20XX. B-D is the comparison of the raster depth values from the “Mix”, GEBCO, and ARDEM bathymetry raster layers, respectively. Within each plot in B-D, the mean and median percent bias (PB), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) are provided.