Sentinel-2 waters edge dataset (SWED) is a collection of diverse coastlines for semantic segmentation of ocean and land pixels [1]. The introduction of SWED provides an opportunity to analyze the association between segmentation maps and indices across a variety of coastline types.

Both correlation and mutual information do not consider the relationships between the channels. So, we trained an XGBoost model to predict the target using the 7 indices and/or 12 bands as input. The model is limited to 100 trees with a maximum depth of 2.

A limitation of correlation is that it provides a linear measure of the association between the pixel values in the masks and channels.

Figure 1: Example of a ground truth mask and indices created using SWED. The RGB image has been created by combining the Red, Green and Blue spectral bands.

Figure 2: Box plots of the correlation between the ground truth masks and channels ordered by the absolute mean correlation. The correlation for each of the 95 test images in SWED is given.

Figure 3: Box plots of the mutual information scores between the ground truth masks and channels. The box plots have been ordered by the mean mutual information.

Figure 4: Feature importance scores for the model trained on all spectral bands and indices.

Figure 5: Feature importance scores for the model trained on only the spectral bands.

So what?

- Can relate results to interpretations of semantic segmentation models
- Provides insights for automated annotation

Figure 6: Permutation importance scores for spectral bands used as input into the coastal segmentation model (U-Net). The scores give a decrease in average accuracy when the respective band is permuted.