DPLUS065 Coastal Habitat Mapping: Developing fine-scale habitat maps for the Falkland Islands using a QGIS and SAGA-based Random Forest classifier modelling approach with WorldView satellite and drone imagery

Background

Upon the completion of island-wide broad-scale habitat maps for the Falkland Islands and South Georgia, fine-scale habitat models were developed for areas prioritised by stakeholders. These models used higher resolution imagery capable of capturing finer landscape features and a higher degree of nuance in classifiable habitat classes. The resultant fine-scale habitat maps from these models were created through a QGIS 3.4 and SAGA 7.2.0 based workflow (both of which are open source software). These maps incorporated both drone and commercial (WorldView, kindly provided through a Digital Globe Foundation grant) imagery into a pixel and object-based Random Forest classification.

Model inputs

- RGB drone imagery where available
- Multispectral drone imagery (RGB, Near Infra Red & Red Edge) for Port Sussex
- WorldView 2/3 satellite imagery where no drone mapping surveys were conducted
- Drone-derived digital elevation model (DEM) where available
- Shuttle Radar Topographic Mission (SRTM) based digital elevation model (DEM) from USGS.
- Slope & aspect layers calculated from DEMs
- Rasterised landward buffer initiated from the coastline
- Topographic Position Index layer
- Flow accumulation layer
- Segmentation layer from imagery created in SAGA
- Various ground validation datasets

Workflow summary

Both drone and commercial imagery were prepared for processing (including the creation of the data products listed in the "**Model inputs**" section of this document). After data preparation was completed, the imagery were segmented using SAGA's "Object Based Image Segmentation" tool. The segmented polygon shapefile was then updated with the project ground validation data, after which the groundtruthed polygons were selected and exported from the original segmented dataset as their own independent polygon shapefile. From that newly exported shapefile, 20% of the groundtruthed polygons were randomly selected and then exported to a new validation dataset (polygons to remain outside of the classification process to be used as independent verification of model results in post-processing), while the remaining 80% of groundtruthed polygons were exported into a separate shapefile to be used as the training

input of the classifier. Once the training and validation datasets were parted, a Random Forest classification was run using the training shapefile and the compiled imagery. The classification itself was managed through the dzetsaka plugin available in QGIS; an easy-to use graphical user interface capable of running various types of common imagery classifiers. Zonal statistics were then applied to results of the classifier where the raster produced by the classifier was the primary data input and the segmented polygons (of the entire map area, not the groundtruth-specific polygons) were used to define the zones over which a majority (mode) statistic were calculated. The majority (mode) of the classified raster within the boundary of each segmented polygon was then written to the segmented polygons shapefile as a new attribute. Shared boundaries between polygons classified as the same landcover type were then dissolved, creating the final map output of the modelling workflow. Classification accuracy was then assessed through the production of confusion matrices, where the groundtruthed classification values within the polygons of the validation shapefile were compared to the classifications assigned to the same polygons found in the non-dissolved classified polygons. The results, printed to confusion matrices tables, displayed the number of validation points that were correctly classified, also numbering and listing the incorrectly classified polygons by the classes into which they were misclassified.

Planned improvements
Improvements to the user interfaces and the simplification of the workflow processes through scripting and further model building would greatly simplify and expedite the overall modelling process. The expansion of groundtruthing datasets, particularly for those classes that are currently lacking in overall number of groundtruthing points, would also likely improve the classification accuracy.