



**JNCC Report  
No. 664**

**Darwin Plus 081 Project**

**Using earth observation to map terrestrial habitats of  
the Turks and Caicos Islands**

**Jones, G. & Lightfoot, P.**

**July 2020**

**© JNCC, Peterborough 2020**

**ISSN 0963-8091**

**For further information please contact:**

Joint Nature Conservation Committee  
Monkstone House  
City Road  
Peterborough PE1 1JY  
[www.jncc.gov.uk](http://www.jncc.gov.uk)

**This report should be cited as:**

Jones, G. & Lightfoot, P. (2020). Darwin Plus 081 Project - Mapping Terrestrial Habitats of the Turks and Caicos Islands. JNCC Report No. 664. JNCC, Peterborough, ISSN 0963-8091.

**JNCC EQA Statement:**

This report is compliant with the JNCC Evidence Quality Assurance Policy

<https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/>

**Acknowledgements**

This research was carried out with support from the Darwin Initiative. We are very grateful to Bryan Manco at the Department of Environment and Coastal Resources (DECR) in the Turks and Caicos Islands for providing expert advice and sharing his extensive knowledge of the island's ecosystems and habitats. Our thanks to both Eric Salamanca and Bryan Manco at DECR for their comments which helped to improve this report. We are also grateful to our colleagues at JNCC and staff at Environment Systems Ltd for their support and valuable input throughout this project.

## Summary

This report outlines the process of generating a 'living' terrestrial habitat map for the Turks and Caicos Islands (TCI), which will be used to support conservation monitoring and provide firm evidence for policy development and planning. The diverse habitats of the Turks and Caicos Islands support many rare, threatened and endemic species, but are at risk from the impacts of development and climate change. The 'living maps' concept (different to the Living Maps Method), allows for the habitat map to be refined using ecological knowledge over time, so that misclassifications can be corrected, and any changes to the landscape recorded. The most recent version of the map will be available at the TCI Department of Environment and Coastal Resources (DECR).

### Outline of method

- **Selection and processing of input data**

Optical satellite imagery with minimal cloud cover was identified for the area of interest and processed to create analysis-ready data using JNCC's pre-processing workflows. Ancillary datasets on topography, roads and buildings were also collated.

- **Image mosaicking and segmentation**

Very high resolution (WorldView-2, Pleiades) imagery was used to produce a mosaic covering most of the TCI. Sentinel-2 imagery was used to fill gaps caused by cloud in the VHR imagery. The imagery was segmented at an appropriate scale to create objects (polygons) representing homogeneous areas of habitat.

- **Habitat map produced using the Living Maps Method (LMM)**

These objects were classified using the Living Maps Method (LMM). LMM uses random forest machine learning to classify objects based on spectral values and other features. LMM is open source, objective and repeatable. Training data samples were selected with reference to the 2010 habitat map and local knowledge.

- **Habitat map refined using local ecological knowledge**

The first iteration of the habitat map was refined using local ecological knowledge. This involved correcting misclassifications and adding classes that could not be detected from EO data (archaeological artefacts).

The final iteration of the habitat map, the input data and the code used to create the map have been shared with the TCI Government. The mapping methods used and outputs have been discussed at several workshops over the last 2 years (funded by the Conflict Security and Stability Fund's (CSSF) 'Natural capital in the Caribbean and South Atlantic Overseas Territories: Valuation, vulnerability and monitoring change project'; in collaboration with the Darwin Plus project 'Mapping for evidence-based policy, recovery and environmental resilience' with key staff from the TCI Government. The map is considered an output of the Darwin Plus project which has helped to build local capacity in the use of GIS and earth observation (EO) for ongoing habitat mapping and monitoring. This will enable the relevant authorities in the TCI to take full advantage of the unprecedented opportunities offered by EO data to improve environmental decision-making and meet operational evidence needs.

# Contents

<b>1</b>	<b>Introduction.....</b>	<b>1</b>
<b>2</b>	<b>Turks and Caicos Islands.....</b>	<b>1</b>
<b>3</b>	<b>Input data.....</b>	<b>2</b>
3.1	VHR Satellite Imagery.....	2
3.2	Sentinel-2 .....	4
3.3	Band indices .....	5
3.4	Digital Elevation Model (DEM) .....	6
3.5	Open Street Map data.....	6
3.6	Training Data .....	6
<b>4</b>	<b>Habitat map workflow.....</b>	<b>7</b>
4.1	Segmentation.....	8
4.2	Classification.....	8
4.3	Refinement .....	9
<b>5</b>	<b>Project outputs.....</b>	<b>11</b>
5.1	Habitat map .....	11
5.2	Data, code and knowledge transfer.....	13
<b>6</b>	<b>References .....</b>	<b>15</b>
	<b>Appendix 1 – R script .....</b>	<b>16</b>

## 1 Introduction

Habitat maps provide easily accessible information on the distribution and extent of features across the landscape. This information is used to facilitate better management practices, ecosystem service mapping, biodiversity monitoring and to develop and deliver sustainable development policies. Many global land cover and land use maps exist, but do not provide enough spatial or thematic detail to support environmental decision making on a regional level.

Prior to this project, the most recent terrestrial habitat map of the Turks and Caicos Islands was produced in 2010 through field survey and manual interpretation of aerial imagery (Wood 2010). These methods are resource intensive and cannot be used to update the habitat map on a regular basis.

Earth observation (EO) has considerably progressed habitat mapping for conservation management, and can give insight into land use, cover and change at a variety of scales. EO can generate data for large areas in a non-destructive, rapid, repeatable way. It provides information beyond human perception, access to inaccessible locations and enables automated and objective data interpretation. The quality and quantity of EO data are both rapidly increasing. Some missions, such as the European Union's Copernicus Programme, make data and products freely available to users.

The aims of this project were to:

- (1) create an updated terrestrial habitat map of the Turks and Caicos Islands by combining a machine learning approach with local ecological knowledge to interpret multi-source EO data and contextual data.
- (2) build local capacity to use freely available EO data for ongoing habitat mapping and monitoring, by using open source tools where possible and sharing the code and input data used in map creation.

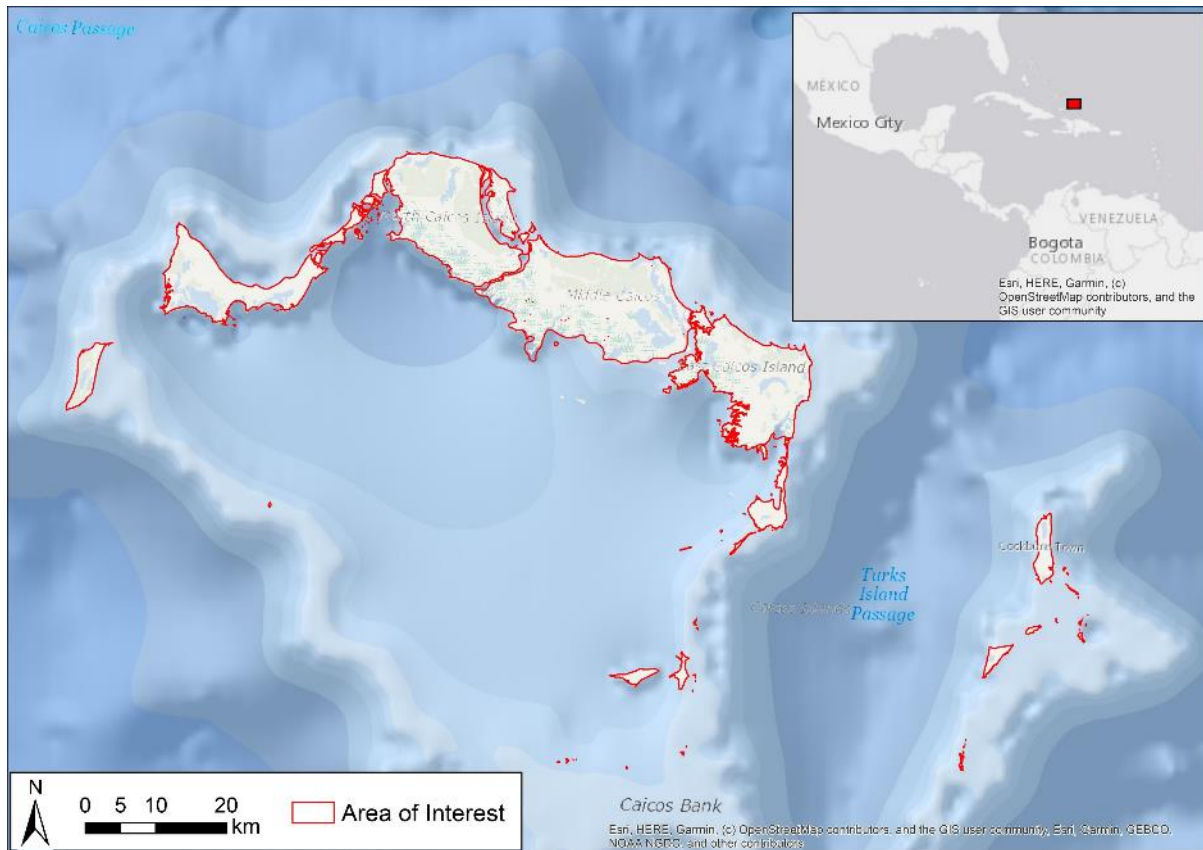
## 2 Turks and Caicos Islands

The Turks and Caicos Islands (TCI) are a UK Overseas Territory located in the tropical Atlantic Ocean around 900km southeast of Florida (Figure 1). The archipelago of over 40 islands has a total land area of 950km<sup>2</sup> and supports numerous endangered, endemic or range-restricted species and rare habitats. The dry tropical forest ecosystem is one of the most threatened habitats in the world and supports many endemic species. Extensive mangroves provide critical ecosystem services, such as nursery grounds for fish and protection from storm surges. Freshwater seasonal habitats and salinas (salt-pans) provide important nesting and foraging sites for diverse bird populations.

Habitat loss, degradation and fragmentation has occurred in recent years due to residential, commercial, industrial and tourism-related development, particularly on the most highly populated island of Providenciales. Many of the TCI's species and habitats are highly sensitive to disturbance from human activities, particularly nesting birds and sea turtles. Furthermore, much of the TCI's land area is less than 50m above sea-level, making coastal communities and wildlife vulnerable to hurricane-generated storm surges and to longer term sea-level rise caused by climate change.

The TCI have an extensive network of terrestrial and marine protected areas designated under local legislation and managed by the Department of Environment and Coastal Resources (DECR). The North, Middle and East Caicos wetland complex is internationally

designated as a Ramsar site covering over 58,000ha, and as a Nature Reserve under the Protected Areas System, it is the largest protected site on the islands. Several active non-governmental organisations support the DECR in the monitoring and management of conservation areas. The habitat map produced by this project will be used to support these activities and provide firm evidence for policy development and planning.



**Figure 1.** Location of the Turks and Caicos Islands and outline of the area of interest used for mapping.

### 3 Input data

The datasets used for habitat classification are listed below. All datasets were georeferenced to WGS 84 / UTM zone 19N (EPSG:32619).

#### 3.1 VHR Satellite Imagery

Very High Resolution (VHR) optical satellite imagery was obtained for the area of interest. WorldView2 and WorldView3 was purchased from Maxar Technologies (formerly DigitalGlobe)<sup>1</sup>; and Pleiades imagery was purchased from Airbus<sup>2</sup> under a group license to enable its re-use by the TCI Government. All data are transformed to produce a surface reflectance product using the Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) software<sup>3</sup>, and orthorectified using ORFEO Toolbox<sup>4</sup>. Only the red, green, blue and Near Infra-Red (NIR) bands were used to create the mosaic for mapping (band names and wavelengths shown in Table 1). The Worldview bands were also re-scaled to match the

<sup>1</sup> <https://www.maxar.com/>

<sup>2</sup> <https://www.airbus.com/space.html>

<sup>3</sup> <http://www.rsgislib.org/arcsi>

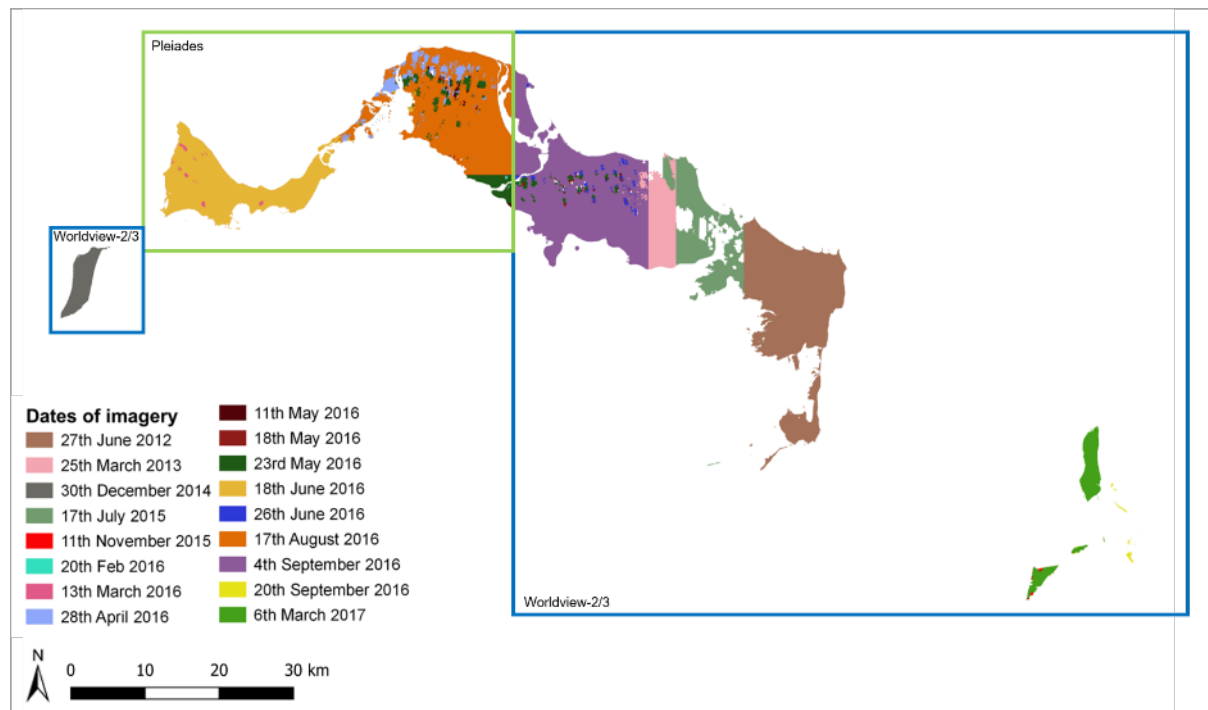
<sup>4</sup> <https://www.orfeo-toolbox.org/>

radiometric resolution of Pleiades data (reduce the bit depth from 16 bit to 12 bit) using The Remote Sensing and GIS Software Library (RSGISLib)<sup>5</sup>.

**Table 1.** Band names, and wavelengths of Worldview-2, Worldview-3 and Pleiades A & B. All sensors come with a panchromatic band which was not used in this project and therefore omitted from the table.

Band name	Worldview-2 Central wavelength (nm)	Worldview-3 Central wavelength (nm)	Pleiades-A & -B wavelength range (nm)
Coastal	427	426	-
Blue	478	481	430-550
Green	546	547	490-610
Yellow	608	605	-
Red	659	661	600-720
Vegetation red edge	724	724	-
NIR 1	833	832	750-950
NIR 2	949	948	-

Imagery from 2012 to 2017 (see Figure 2 for spatial distribution of image dates) was mosaicked, resampled to 2m resolution and clipped to the boundary of the 2010 habitat map plus a 20m buffer. The majority of the imagery used was from 2015-2017, but older imagery from 2012-2014 was needed to fill some gaps, predominantly in East and Middle Caicos.



**Figure 2.** Dates of the source imagery for the VHR mosaic.

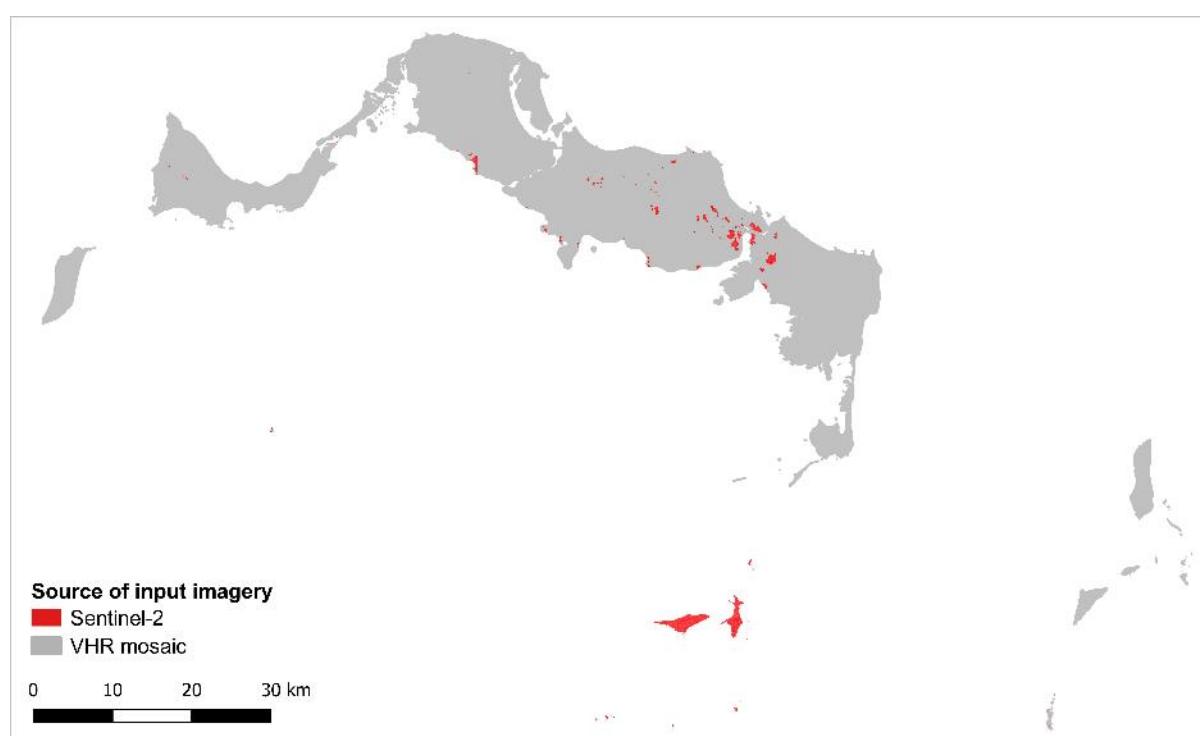
<sup>5</sup> <http://www.rsgislib.org/>

## 3.2 Sentinel-2

The VHR mosaic covered most of the area of interest, but a few areas were obscured by cloud, notably Little Ambergris Cay, Big Ambergris Cay and parts of Middle and East Caicos. Sentinel-2 imagery acquired on three dates was used to fill these gaps (Table 1 and Figure 2).

**Table 1.** Sentinel-2 granules used to fill gaps caused by cloud in VHR mosaic.

Sensor	Orbit	Acquisition date
Sentinel-2A	68	24/12/2015
Sentinel-2A	25	18/06/2016
Sentinel-2A	68	06/06/2017



**Figure 2.** Source of input imagery used to create habitat map.

The Sentinel-2 imagery was processed to analysis-ready data (ARD) using processing chains established through previous project work at JNCC (Jones *et al.* 2017). Processing was carried out on Amazon Web Services cloud computing facilities. Sentinel-2 data are transformed to produce a topographically corrected surface reflectance product using the Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) software<sup>6</sup>. During processing, the 20m image bands are sharpened to 10m through application of linear regression models. The 60m bands are primarily used for atmospheric aerosol correction and processes and are therefore removed from the final surface reflectance product. Table 2 shows the central wavelengths, capture resolution and band numbers of Sentinel-2 ARD.

<sup>6</sup> <http://www.rsgislib.org/arcsi>

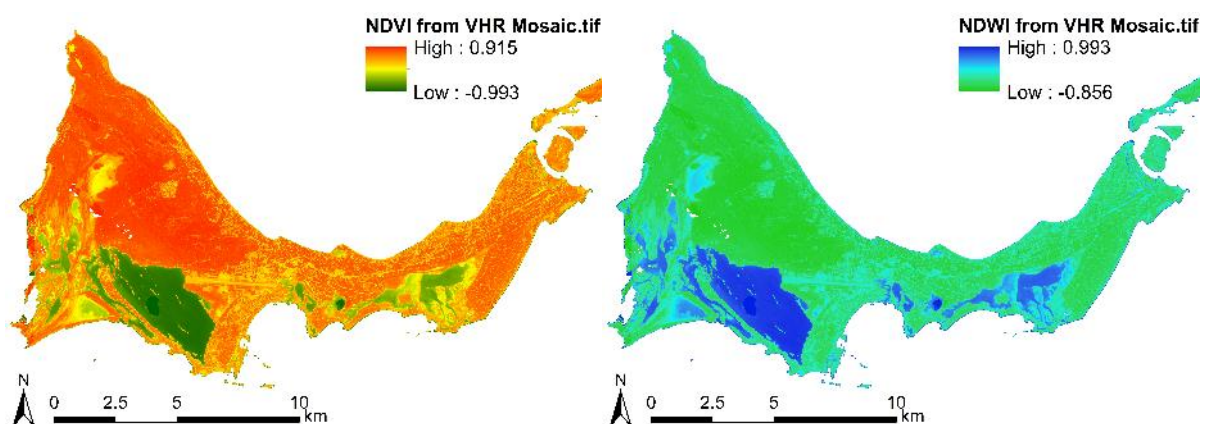


**Table 2.** Band numbers, spatial resolution and wavelengths of Sentinel-2 original data and ARD.

Band number and name in original data	Sentinel 2A Central wavelength (nm)	Sentinel 2B Central wavelength (nm)	Capture resolution (m)	Band number in ARD
Band 1: Coastal aerosol	442.7	442.2	60	
Band 2: Blue	492.4	492.1	10	<b>1</b>
Band 3: Green	559.8	559.0	10	<b>2</b>
Band 4: Red	664.6	664.9	10	<b>3</b>
Band 5: Vegetation red edge	704.1	703.8	20	<b>4</b>
Band 6: Vegetation red edge	740.5	739.1	20	<b>5</b>
Band 7: Vegetation red edge	782.8	779.7	20	<b>6</b>
Band 8: NIR	832.8	832.9	10	<b>7</b>
Band 8A: Narrow NIR	864.7	864.0	20	<b>8</b>
Band 9: Water vapour	945.1	943.2	60	
Band 10: SWIR – Cirrus	1373.5	1376.9	60	
Band 11: SWIR	1613.7	1610.4	20	<b>9</b>
Band 12: SWIR	2202.4	2185.7	20	<b>10</b>

### 3.3 Band indices

Normalised difference vegetation index (NDVI) rasters were created from both the VHR mosaic and the Sentinel-2 mosaic using the formula  $(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$ . Normalised difference water index rasters were also created for both mosaics using the formula  $(\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR})$ . Band 8A was used as the NIR input for both indices derived from the Sentinel-2 imagery. Extracts of the NDVI and NDWI rasters produced from the VHR mosaic are shown in Figure 3.



**Figure 3.** Extracts from the NDVI (left) and NDWI (right) rasters produced from the VHR mosaic. The area shown is the island of Providenciales.

### 3.4 Digital Elevation Model (DEM)

A digital elevation model (DEM) was sourced from Airbus WorldDEM™ database<sup>7</sup>. WorldDEM™ elevation models are provided as 12m rasters with 2m relative and 4m absolute vertical accuracy. The DEM was clipped to the area of interest and resampled to 10m spatial resolution. Slope, aspect and hillshade rasters were derived from the clipped, resampled DEM.

### 3.5 Open Street Map data

OpenStreetMap® are open data, licensed under the Open Data Commons Open Database License (ODbL) by the OpenStreetMap Foundation (OSMF). Data layers for buildings and roads were downloaded and used to create proximity rasters at 2m resolution. Proximity raster cell values range from 0 – 20 (buildings) and 0 – 10 (roads) based on the distance in metres from the nearest feature.

### 3.6 Training Data

A proposed national standardised vegetation classification was developed for the TCI in 2010 (SWA Ltd *et al.* 2010), which was used as a basis for the current mapping project. A simplified classification comprising 21 classes was used for the first iteration of the map. Table 3 shows the classes used in the first iteration and how these relate to the national standardised classification. Samples of each type of habitat were selected through visual interpretation of the VHR satellite imagery, reference to the 2010 habitat map and consultation with local ecological experts.

Fifty samples were created per class for the area covered by the VHR mosaic. Due to the smaller area covered by the Sentinel-2 mosaic (Figure 2), it was only possible to create 50 samples per class for 14 of the 21 habitat classes, the remaining seven habitat classes had a lower number of samples.

---

<sup>7</sup> <https://www.intelligence-airbusds.com/elevation-models/#worldDEM>

**Table 3.** Classes used to create first iteration habitat map of the Turks and Caicos Islands

Classes used in first iteration habitat map	National standardised vegetation classification: class name and code
Algal Flat	Estuarine nonvascular (613), Estuarine nonvascular - black algae (613.1)
Anthropogenic	Urban (710), Clear-cut land (720), Human altered aquatic habitats (740)
Beach	Coastal nonvascular (612), Coastal nonvascular – sand (612.1) Coastal nonvascular – rock (612.2)
Blue Hole	Palustrine permanent pond karst feature (615)
Coastal coppice	Coastal broadleaf evergreen woodland (212), Coastal broadleaf mixed evergreen/drought deciduous woodland (232)
Coastal scrub	Coastal broadleaf evergreen shrubland (312), Coastal broadleaf mixed evergreen/drought deciduous shrubland (332)
Dwarf Coastal Scrub	Coastal evergreen dwarf shrubland (412), Coastal microphyllous evergreen dwarf shrubland (412.1), Coastal broadleaf mixed evergreen/drought deciduous dwarf shrubland (432)
Dwarf Mangrove	Estuarine broadleaf evergreen dwarf shrubland (413), Estuarine broadleaf mixed evergreen/drought deciduous dwarf shrubland (433)
Dwarf scrub	Upland broadleaf mixed evergreen/drought deciduous dwarf shrubland (431)
Dwarf Swamp	Palustrine broadleaf evergreen dwarf shrubland (414), Palustrine broadleaf mixed evergreen/drought deciduous dwarf shrubland (434)
Grassland	Upland perennial graminoid (511), Upland perennial mixed graminoid/forb (531)
Herbaceous Dune	Coastal perennial graminoid (512), Coastal perennial forb (522) Coastal perennial mixed graminoid/forb (532)
Invasive Species	Impacted by nuisance exotic species (730)
Limestone thicket	Upland broadleaf drought deciduous woodland (221), Upland broadleaf mixed evergreen/drought deciduous woodland (231)
Mangrove	Estuarine broadleaf evergreen forest (113), Estuarine broadleaf evergreen woodland (213), Estuarine broadleaf evergreen shrubland (313), Estuarine broadleaf mixed evergreen/drought deciduous shrubland (333)
Marsh	Estuarine perennial graminoid (513), Estuarine perennial forb (523), Estuarine perennial mixed graminoid/forb (533), Estuarine annual forb (553), Palustrine perennial mixed graminoid/forb (534), Palustrine perennial mixed graminoid/forb fresh/brackish (534.1), Palustrine perennial mixed graminoid/forb saline (534.2)
Pineyard	Upland coniferous evergreen woodland (241), Palustrine coniferous evergreen woodland (244)
Pond	Palustrine nonvascular (614), Palustrine pond (fresh/brackish) (614.1), Palustrine pond (saline) (614.2), Palustrine permanent pond karst feature (615)
Scrub	Upland broadleaf drought deciduous shrubland (321), Upland broadleaf mixed evergreen/drought deciduous shrubland (331)
Swamp	Palustrine broadleaf evergreen forest (114), Palustrine broadleaf mixed evergreen/drought deciduous forest (134), Palustrine broadleaf evergreen woodland (214), Palustrine broadleaf mixed evergreen/ drought deciduous woodland (234), Palustrine broadleaf evergreen shrubland (314), Palustrine broadleaf mixed evergreen/drought deciduous shrubland (334)
Tropical dry forest	Upland broadleaf mixed evergreen/drought deciduous forest (131)

## 4 Habitat map workflow

The habitat map was produced using object-based image analysis (OBIA). OBIA involves segmenting imagery by grouping pixels based on homogeneity criteria to form objects, then classifying those objects based on their spectral and geometric properties, their relationships to other objects and to contextual data (Benz 2004). This has several advantages over pixel-based approaches, especially when high resolution imagery is used (Blaschke *et al.* 2014). Image objects possess far more values than single pixels, any combination of which can be used in the classification process, aiding discrimination of spectrally similar habitats.

OBIA also enables interpretation of multiple layers of data simultaneously, including contextual data, and avoids the ‘salt-and-pepper’ effect which can be generated by pixel-based analysis. The following sections outline the segmentation and classification stages of the OBIA workflow used to create the TCI habitat map.

## 4.1 Segmentation

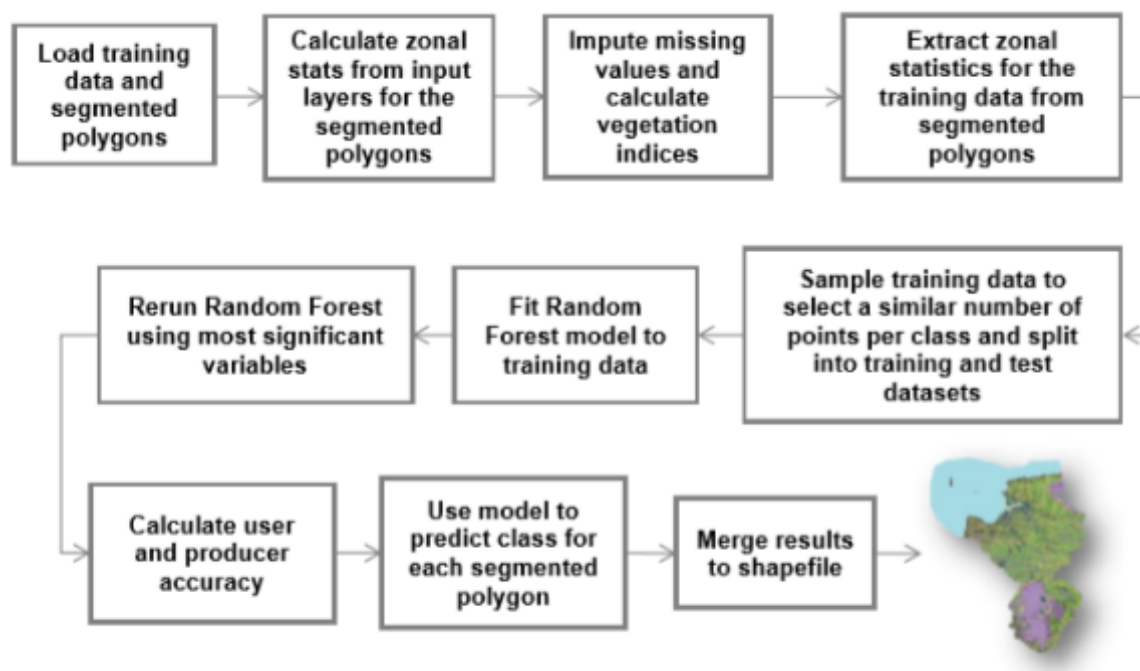
Image segmentation was carried out using Trimble eCognition version 9.2.1. The multispectral segmentation algorithm was used with the parameters shown in Table 2. Segmentation of the VHR mosaic took 77 minutes to complete and produced 44,108 objects. Segmentation of the Sentinel-2 mosaic took less than two minutes to complete and produced 33,632 objects. Although the area covered by Sentinel-2 imagery was very small (Figure 2), the number of objects in that area had to be high to maintain the required number of training points per class while preventing multiple training points in a single object.

**Table 4.** Parameters used for segmentation of VHR and Sentinel-2 imagery in eCognition.

Segmentation Parameters		
VHR imagery	Image band weights	Blue = 1 Green = 1 Red = 1 NIR = 1 NDVI = 3 NDWI = 2
	Scale	300
	Shape	0.1
	Compactness	0.9
Sentinel-2	Image band weights	Blue = 2 Green = 2 Red = 2 Red Edge 5 = 1 Red Edge 6 = 1 Red Edge 7 = 1 NIR = 2 Narrow NIR = 1 SWIR1 = 1 SWIR2 = 1 NDVI = 10 NDWI = 5
	Scale parameter	15
	Shape	0.1
	Compactness	0.9

## 4.2 Classification

The classification method chosen for the generation of the habitat map was the Living Map Method (LMM) developed by Natural England (Kilcoyne *et al.* 2017). The classification workflow is summarised in Figure 4. The LMM uses Random Forests, an ensemble classifier that creates a number of classification and regression trees, each trained by a subset of training data with randomly sampled variables. It classifies objects as the class assigned by the majority of trees in the forest (Breiman 2001). Random Forests is a popular method of classifying Earth observation data due to its processing speed, classification accuracy, ability to handle a large number of variables and to avoid over-fitting to the data (Belgiu & Dragut 2016).



**Figure 4.** Flowchart of classification process (Kilcoyne *et al.* 2017).

The LMM code was developed in the open source R programming environment<sup>8</sup>. The code is freely available under an Open Government License. Users can apply the code to new areas and input data with little adaptation of the code required. The original code is available on GitHub<sup>9</sup> and the revised version used to create the habitat map of the TCI is provided in the [Appendix](#) to this report. This method is highly repeatable and objective, although user input into segmentation parameters, training data and feature selection can strongly influence outputs.

### 4.3 Refinement

The first iteration of the habitat map was refined by JNCC in consultation with ecological experts in the TCI. This involved carrying out manual knowledge-based corrections of misclassifications and adding classes that could not be detected from EO data. The total area changed through manual editing was 204.6km<sup>2</sup> which was 20% of the total map area. The classification change types which accounted for most of the area changed by manual edits are shown in Table 5. The classification change which accounted for the largest area changed through manual editing was from Pineyard to Swamp, which accounted for 14% of the total area edited. Archaeological artefacts from pre-Colombian or colonial eras are an important conservation feature of the TCI but cannot be mapped from EO data. A separate shapefile showing the locations of archaeological artefacts was produced with reference to the 2010 survey and supplied for use with the habitat map.

<sup>8</sup> <https://www.r-project.org/>

<sup>9</sup> <https://github.com/NE-EEOS/Living-England>

**Table 5.** Classification change types accounting for 65% of the area of the first iteration habitat map changed through manual edits.

Classification change	Area (km <sup>2</sup> )	Number of objects	Change as % of total changed area	Cumulative %
Pineyard to Swamp	28.69	521	14%	14%
Pineyard to Dwarf Swamp	15.02	324	7%	21%
Beach to Algal Flat	13.10	1182	6%	28%
Pond to Algal Flat	12.23	302	6%	34%
Pineyard to Mangrove	11.11	380	5%	39%
Blue Hole to Algal Flat	8.03	178	4%	43%
Algal Flat to Pond	7.77	608	4%	47%
Blue Hole to Pond	7.20	393	4%	50%
Pond to Tropical dry forest	4.03	54	2%	52%
Pineyard to Marsh	3.87	130	2%	54%
Blue Hole to Scrub	3.86	68	2%	56%
Blue Hole to Coastal scrub	3.73	21	2%	58%
Pineyard to Coastal coppice	3.61	97	2%	60%
Dwarf Mangrove to Pond	3.48	219	2%	61%
Pond to Mangrove	3.37	119	2%	63%
Pond to Swamp	3.10	153	2%	65%

The second iteration of the map was refined by Environment Systems using manual Aerial Photo Interpretation (API) following input from relevant ecological experts. Three new habitat classes were created and manually classified: Oakbottom, Salina saltmarsh and Canal. Algal flat was recategorized as Algal flat / Pond as this estuarine habitat is characterised by pools of standing water. The total area changed through manual edits was 70.1km<sup>2</sup>, which is 7% of the total map area. The classification change types that accounted for most of the area changed are shown in Table 6. The classification change which accounted for the largest area changed through manual editing was from Dwarf Swamp to Salina salt marsh, which accounted for 57% of the total area edited.

**Table 6.** Classification change types accounting for 75% of the area of the second iteration map changed through manual edits.

Classification change	Area (km <sup>2</sup> )	Number of objects	Change as % of total changed area	Cumulative %
Dwarf swamp to Salina salt marsh	39.91	2797	57%	57%
Algal flat to Swamp	1.91	671	3%	60%
Tropical dry forest to Oakbottom	1.82	99	3%	62%
Dwarf mangrove to Algal flat	1.62	571	2%	65%
Anthropogenic to Dwarf coastal scrub	1.43	1153	2%	67%
Pineyard to Algal flat	1.35	1109	2%	69%
Algal flat to Dwarf mangrove	0.91	854	1%	70%
Pineyard to Scrub	0.90	28	1%	71%
Algal flat to Mangrove	0.89	815	1%	72%
Dwarf scrub to Dwarf mangrove	0.85	661	1%	74%
Dwarf mangrove to Swamp	0.84	483	1%	75%

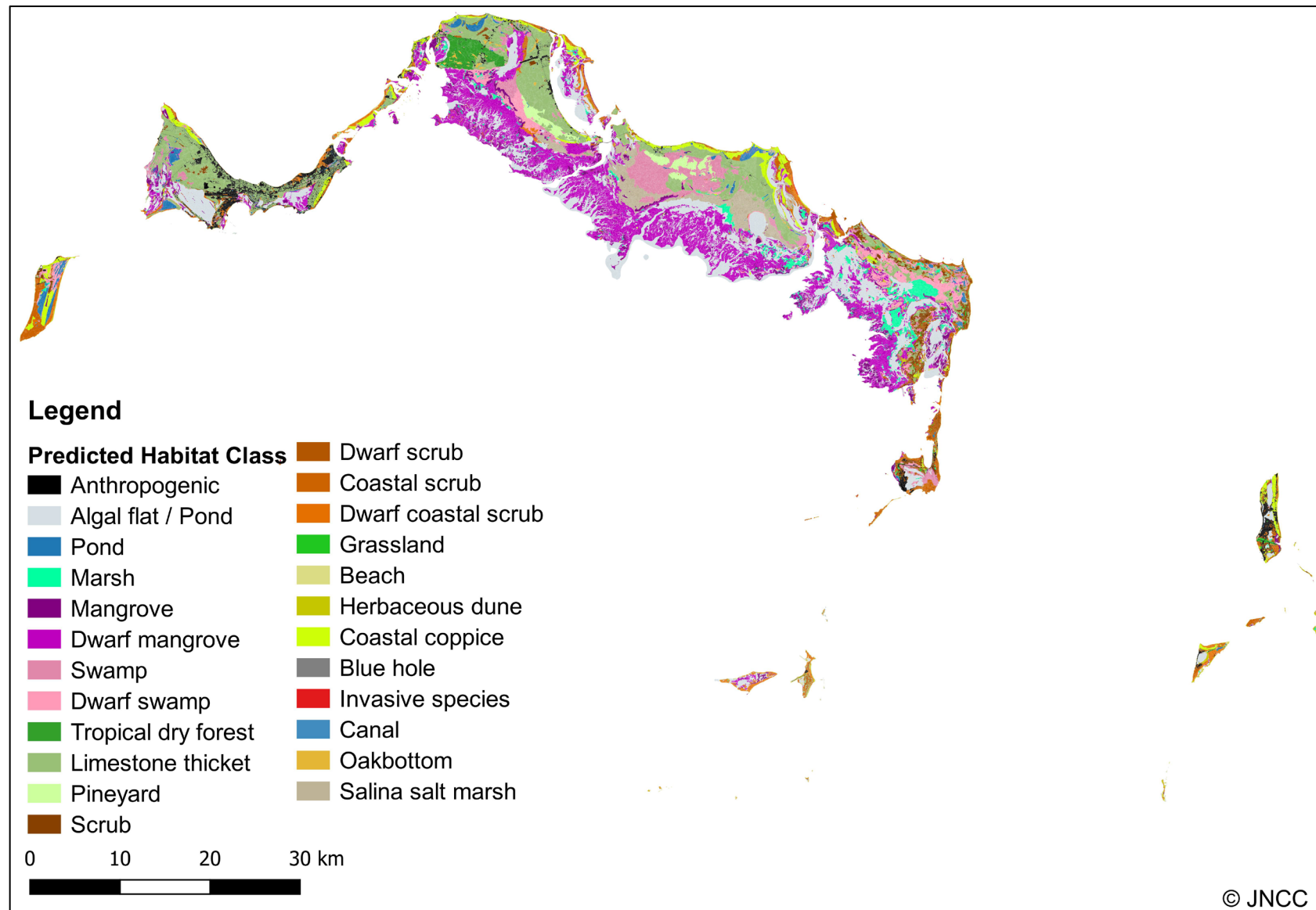
## 5 Project outputs

### 5.1 Habitat map

The habitat map of the TCI produced by the workflows described above is shown in Figure 5 and the predicted extent of each habitat class as a proportion of total map area is shown in Figure 6. The predominant habitats in terms of predicted extent are:

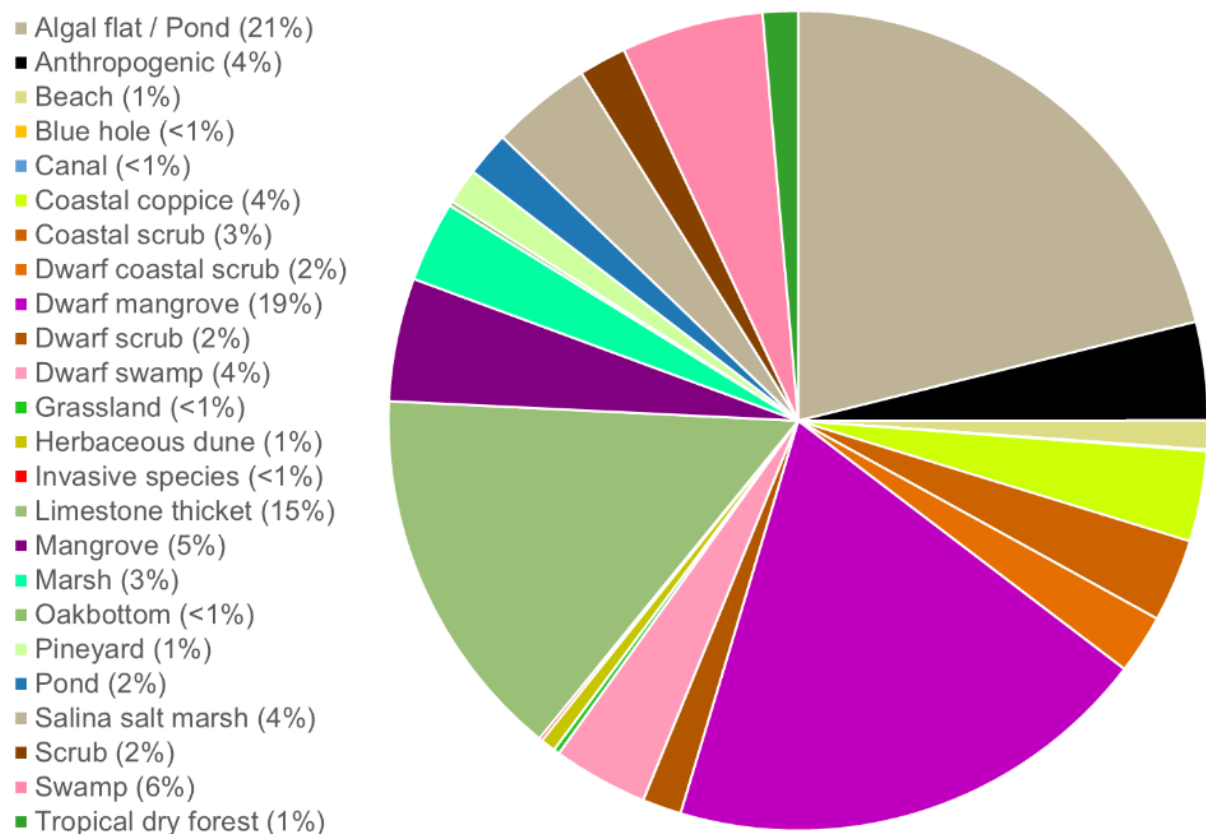
- **Mangrove** and **Dwarf mangrove** which together account for just under a quarter of total land area, concentrated on the southern coasts of the North, Middle and East Caicos Islands.
- **Algal flat / Pond**, an extensive network of estuarine habitats which occupies low-lying areas throughout the southern regions of the Caicos islands, supporting both marine and terrestrial biodiversity and providing important ecosystem services.
- **Limestone thicket**, which occupies the upland areas in the northern parts of Providenciales, North and Middle Caicos. Together with the far less extensive **Oakbottom** and **Tropical Dry Forest**, these upland wooded habitats provide important nesting and foraging areas for bird populations and support many threatened and endemic plant, invertebrate, and reptile species.





**Figure 5.** Terrestrial habitat map of the Turks and Caicos Islands produced through object-based image analysis of multi-source optical satellite imagery.





**Figure 6.** Predicted habitat extents as proportion of total map area.

Quantitative accuracy assessment of the habitat map was not possible due to the lack of ground truth data. Qualitative evaluation was carried out through visual assessment by terrestrial ecology and environmental policy experts at DECR. The map was deemed fit for the purpose of informing policy and helping to direct conservation monitoring and management efforts. It provides reliable insight into the distribution and extent of habitat based on current data but should be considered a 'living' map to be updated as new data and information become available. The update process would need to be managed by DECR, who would determine the appropriate frequency of updates. Future updates will likely include developing more detailed definitions for the habitats including predicted/expected species. The latest version of the map will be available from DECR.

## 5.2 Data, code and knowledge transfer

The final iteration of the habitat map, the input data and the code used to create the map have been shared with the TCI Government. The mapping methods used and outputs have been discussed at several workshops over the last two years (funded by the Conflict Security and Stability Fund's (CSSF) 'Natural capital in the Caribbean and South Atlantic Overseas Territories: Valuation, vulnerability and monitoring change project'; in collaboration with the Darwin Plus project 'Mapping for evidence-based policy, recovery and environmental resilience' with key staff from the TCI Government. The map is considered an output of the Darwin Plus project.

The workshops highlighted that the 2010 habitat map (SWA Ltd *et al.* 2010) was delivered to DECR in a format that the Department had neither the software, equipment or capacity to edit and update at the time. This mapping exercise and the Darwin Plus project created a useable product and built staff capacity to manage the data so they can continue to work with a living habitat map of the TCI.

The twin satellites of the Sentinel-2 mission provide new imagery every five days at the equator and every two-three days at mid latitudes. The rapid revisit times and high spectral and spatial resolution of Sentinel-2 provides unprecedented opportunities for ecological monitoring and change detection. Knowledge-sharing and capacity building delivered through the Darwin Plus project will enable the relevant authorities in the TCI to take full advantage of these new opportunities offered by EO data to improve environmental decision-making and meet operational evidence needs.

## 6 References

- BENZ, U.C. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information', *ISPRS Journal of Photogrammetry and Remote Sensing*, **58**, pp.239-258.
- BLASCHKE, T., HAY, G.J., KELLY, M., LANG, S., HOFMANN, P., ADDINK, E., FEITOSA, R.Q., VAN DER MEER, F., VAN DER WERFF, H., VAN COILLIE, F. & TIEDE, D. (2014) 'Geographic Object-Based Image Analysis - Towards a new paradigm', *ISPRS Journal of Photogrammetry and Remote Sensing*, **87**, pp. 180-191.
- BELGIU, M. & DRAGUT, L. (2016) 'Random forest in remote sensing: A review of applications and future directions', *ISPRS Journal of Photogrammetry and Remote Sensing*, **114**, pp. 2431.
- BREIMAN, L. (2001) 'Random forests', *Machine Learning*, **45**(1), pp. 5-32.
- JONES, T., WICKS, D., AGASS, S. & BUNTING, P. (2017). Developing standards and automated production for Sentinel-2 Analysis Ready Data. Evidence Project SD1707
- KILCOYNE, A.M., ALEXANDER, R., COX, P. & BROWNETT, J. (2017). Living Maps: Satellite based Habitat Classification. Evidence Project SD1705
- SWA LTD, BLUE DOLPHIN RESEARCH AND CONSULTING INC. & EDSA (2010). Turks and Caicos Islands National Vegetation and Mapping Project: National Standardized Vegetation Classification.
- WOOD, K.M. (2010). Turks and Caicos Islands National Vegetation and Mapping Project: Mapping of the Terrestrial Habitats Report.

## Appendix 1 – R script

### #### Living Maps - automated object-based classification using random forests ####

# Code used to create habitat map of Turks and Caicos Islands in August 2019

```
library(rgdal)
library(raster)
library(rgeos)
library(randomForest)
library(impute)
# install zip file from: www.bioconductor.org/packages/release/bioc/html/impute.html
library(sp)
library(RSQLite)
library(sf)
```

# set working directory

```
setwd("D:/TCI_Habitat_Mapping")
```

# source the required Living Maps scripts

```
source("Scripts/user_producer_accuracy.r")
```

```
source("Scripts/zonal_stats_velox.r")
```

# both can be downloaded from <https://github.com/NE-EEOS/Living-England>

# read in the training data point shapefile.

```
training.data.habitat.shp <- st_read("Training_Data/TCI_training_from2010.shp",
"TCI_training_from2010")
```

# read in the segmentation shapefile

```
segmentation <- st_read("Segmentation/?Seg300_1908.shp", "Seg300_1908")
```

# load and list the raster layers to be used to generate zonal statistics

```
VHR_Mosaic <- "VHR_Mosaic/TCI_Mosaic_1908.tif"
```

```
NDVI <- "VHR_Mosaic/TCI_NDVI_1908.tif"
```

```
NDWI <- "VHR_Mosaic/TCI_NDWI_1908.tif"
```

```
height <- "Topography/TCI_DEM_1908.tif"
```

```
slope <- "Topography/TCI_Slope_1908.tif"
```

```
aspect <- "Topography/TCI_Aspect_1908.tif"
```

```
hillshade <- "Topography/TCI_Hillshade_1908.tif"
```

```
Buildings <- "Urban/Buildings_prox.tif"
```

```
Roads <- "Urban/Roads_prox.tif"
```

```
list.rasters <- list(VHR_Mosaic=c(VHR_Mosaic, "mean", 1,2,3,4),
```

```
NDVI=c(NDVI, "mean", 1),
```

```
NDWI=c(NDWI, "mean", 1),
```

```
height=c(height, "mean", 1),
```

```
slope=c(slope, "mean", 1),
```

```
aspect=c(aspect, "mean", 1),
```

```
hillshade=c(hillshade, "mean", 1),
```

```
Buildings=c(Buildings, "mean", 1),
```

```
Roads=c(Roads, "mean", 1),
```

```
VHR_Mosaic=c(VHR_Mosaic, "median", 1,2,3,4),
```

```
NDVI=c(NDVI, "median", 1),
```

```
NDWI=c(NDWI, "median", 1),
```

```

VHR_Mosaic=c(VHR_Mosaic, "sd", 1,2,3,4),
NDVI=c(NDVI, "sd", 1),
NDWI=c(NDWI, "sd", 1))

# calculate the zonal statistics for each segmented polygon.
start <- proc.time()
zonal_stats_seg <- zonal.stats.velox(segmentation, list.rasters, tiles=10)
proc.time()-start
warnings()

# save the results as an intermediate file.
zonal_stats_seg <- write.table(zonal_stats_seg,
"Zonal_Stats/zonal_stats_seg_20190524.txt", sep="\t")
zonal_stats_seg <- read.table("Zonal_Stats/zonal_stats_seg_20190524.txt", sep="\t",
header=T)

# append area and perimeter from shapefile if not already calculated
if (!"area_ratio1" %in% names(zonal_stats_seg))
{
  zonal_stats_seg <- merge(zonal_stats_seg, segmentation, by="ID")
  zonal_stats_seg$area_ratio1 <- with(zonal_stats_seg, Shape_Area/Shape_Leng)
  zonal_stats_seg$area_ratio2 <- with(zonal_stats_seg, Shape_Leng/sqrt(Shape_Area))
}

# Ensure that categorical data doesn't having any missing or infinite values.
zonal_stats_seg[is.na(zonal_stats_seg)] <- 0
zonal_stats_seg[is.infinite(zonal_stats_seg)] <- 0

# Calculate band ratios to create extra features for use in classification
zonal_stats_seg$Red_Blue_mean_ri <- with(zonal_stats_seg, (VHR_Mosaic_mean_band3 /
VHR_Mosaic_mean_band1))
zonal_stats_seg$Green_Blue_mean_ri <- with(zonal_stats_seg,
(VHR_Mosaic_mean_band2 / VHR_Mosaic_mean_band1))
zonal_stats_seg$NIR_Blue_mean_ri <- with(zonal_stats_seg, (VHR_Mosaic_mean_band4 /
VHR_Mosaic_mean_band1))
zonal_stats_seg$Green_Red_mean_ri <- with(zonal_stats_seg,
(VHR_Mosaic_mean_band2 / VHR_Mosaic_mean_band3))

# Ensure that categorical data doesn't having any missing or infinite values
zonal_stats_seg[is.na(zonal_stats_seg)] <- 0
zonal_stats_seg[sapply(zonal_stats_seg, is.infinite)] <- 0

colnames(zonal_stats_seg)[colnames(zonal_stats_seg)==""] <- "ID"

write.table(zonal_stats_seg, "Zonal_Stats/zonal_stats_TCI_indices.txt", sep="\t")
zonal_stats_seg <- read.table("Zonal_Stats/zonal_stats_TCI_indices.txt", sep="\t")

# set number of training points per class to be used in random forest classification
nmax <- 30

# load the zonal statistics intermediate file
if (!exists("zonal_stats_seg"))
{

```

```
zonal_stats_seg <- read.table("Zonal_Stats/zonal_stats_TCI_indices.txt", sep="\t",
header=T, as.is=T)
}

# set field in training data to define habitat class for classification
training.data.shp <- training.data.habitat.shp[c("JNCC_class")]

# identify the segmented polygons that the training points fall within and extract the zonal
statistics for these
training.data.ids <- as.numeric(st_within(training.data.shp, segmentation))
training.data <- cbind(as.data.frame(training.data.shp), zonal_stats_seg[training.data.ids,])
training.data <- training.data[names(training.data)!="geometry"]
training.data.all <- training.data[c(1:ncol(training.data))]

# remove rows with missing values.
training.data.all <- training.data.all[complete.cases(training.data.all),]

# split into stratified training and test datasets
training.data <- NULL
training.data.test <- NULL

# loop through all the classes, first setting a random seed to ensure consistent training and
test datasets
set.seed(1)
for(c in unique(training.data.all$JNCC_class))
{
  # select the subset of rows for the current class
  training.data.sub <- subset(training.data.all, JNCC_class==c)

  # split the data using a random sample
  subset <- random.subset(training.data.sub, 0.8)
  training.data <- rbind(training.data, training.data.sub[subset,])
  training.data.test <- rbind(training.data.test, training.data.sub[-subset,])
  rownames(training.data.test) <- NULL
}

# remove duplicates from test dataset
training.data.test <- training.data.test[!duplicated(training.data.test),]

# write training data to text file
write.table(training.data, "Training_Data/training_data_subset.txt", sep="\t")
write.table(training.data.test, "Training_Data/training_data_test_subset.txt", sep="\t")

# read in training and test datasets
training.data <- read.table("Training_Data/training_data_subset.txt", sep="\t", header=T)
training.data.test <- read.table("Training_Data/training_data_test_subset.txt", sep="\t",
header=T)

# ensure that any categorical data are converted to factors.
training.data$JNCC_class <- as.factor(as.character(training.data$JNCC_class))

# predict habitats using random forest, run for only the top 42 most important variables
M.rf.habitats.all <- randomForest(JNCC_class ~ .,
data=training.data[c(1,2:ncol(training.data))], na.action=na.omit)
```

```

i <- colnames(training.data) %in%
c(rownames(M.rf.habitats.all$importance)[order(M.rf.habitats.all$importance,
decreasing=T)][1:42], "JNCC_class")
M.rf.habitats <- randomForest(JNCC_class ~ ., data=training.data[i], na.action=na.omit)

# calculate confusion matrix
p <- predict(M.rf.habitats, training.data.test, type="response")
confusion.matrix(training.data.test$JNCC_class, p)

# predict classes for all polygons
results.habitats.probs <- predict(M.rf.habitats, zonal_stats_seg,
                                type="vote", norm.votes=TRUE,
                                progress="text")

responseNFromProbs <- function(df, n=1) {
  columns <- colnames(df)
  response <- apply(df, MARGIN=1, FUN=function(x) {columns[order(x,
decreasing=TRUE)[n]]})
  return (response)
}

probNFromProbs <- function(df, n=1) {
  response <- apply(df, MARGIN=1, FUN=function(x) {sort(x, decreasing=TRUE)[n]})
  return (response)
}

results.habitats.response1 <- responseNFromProbs(results.habitats.probs, n=1)
results.habitats.prob1 <- probNFromProbs(results.habitats.probs, n=1)*100
results.habitats.response2 <- responseNFromProbs(results.habitats.probs, n=2)
results.habitats.prob2 <- probNFromProbs(results.habitats.probs, n=2)*100

# combine results with segmentation polygons and save to new shapefile
results.rf <- data.frame(ID=zonal_stats_seg$ID,
                        A_pred=results.habitats.response1,
                        A_prob=results.habitats.prob1,
                        B_pred=results.habitats.response2,
                        B_prob=results.habitats.prob2)
segmentation.p <- merge(segmentation, results.rf, by="ID")

st_write(segmentation.p,
        "Outputs/TCI_Map_201908_v1.shp",
        "TCI_Map_201908_v1",
        driver="ESRI Shapefile",
        delete_layer = TRUE)

rm(segmentation.p)

# load libraries to create confusion matrix and graphs
library(reshape2)
library(ggplot2)
theme_set(theme_bw())

# create a variable called confusion matrix - broad classes
cm1 <- broadclass.confusion.matrix(training.data.test$JNCC_class,
training.data.test$JNCC_class, p)

```

```
C_graph1 <- barplot.confusion.matrix(cm1)
```

```
# adjust margins and plot graph
```

```
C_graph1 <- C_graph1 + theme(plot.margin = unit(c(1,1,1,1), "cm"))
```

```
C_graph1
```

```
# create a simple percentage plot - broad classes
```

```
C_graph3 <- barplot.percent(cm1)
```

```
# adjust margins and plot graph
```

```
C_graph3 <- C_graph3 + theme(plot.margin = unit(c(1,1,1,1), "cm"))
```

```
C_graph3
```

```
# plot habitat map shaded to show probability of predicted habitat class
```

```
output.shapefile <- st_read("Outputs/TCI_Map_20190905.shp", "TCI_Map_20190905")
```

```
output.shapefile$A_prob<-as.factor(output.shapefile$A_prob)
```

```
output.shapefile$Colour <- "#FFFFFF"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 0 ] <-  
"#FA0000"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 10 ] <-  
"#FA3200"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 20 ] <-  
"#FA7D00"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 30 ] <-  
"#FAAF00"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 40 ] <-  
"#FAD100"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 50 ] <-  
"#AACA09"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 60 ] <-  
"#94A637"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 70 ] <-  
"#67A637"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 80 ] <-  
"#70CD6A"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) > 90 ] <-  
"#5CFA00"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) == 100 ] <-  
"#11FF00"
```

```
output.shapefile$Colour[(as.numeric(as.character(output.shapefile$A_prob))) == 0] <- "grey"
```

```
plot(output.shapefile, border=NA, main="Output coloured by probability per segment",  
col=output.shapefile$Colour)
```