# **Annex 5**

# **Urban Habitat and Naturalness Mapping – Iteration 1 General Method.**

Urban Habitat and Naturalness Mapping is an Earth Observation based approach to generating detailed data on the constituency of the greenness of the urban ecosystem. It uses a new developing approach to Naturalness as a means of trying to understand the broad environmental quality of the constituent elements of the urban ecosystem from the perspective of “apparent degree of management intensity”.

The Urban Habitat and Naturalness maps identify;

* The “Broad Habitats” within the urban areas.
* The “Detailed Habitats” within the urban areas.
* The distribution of “Naturalness” as a measure of broad environmental quality across an urban area.
* The “naturalness” of each Accessible Green Infrastructure space within the urban area using a “Combined Naturalness Factor” based on the mix of urban habitats within any given space and their relative proportions.

Urban Habitat Maps are created using an Earth Observation based approach blending a variety of source data to create maps of the spatial location and extents of a system of Broad and Detailed Urban Habitat Classes.

The data sources used are;

* England Green Infrastructure and Blue Infrastructure Mapping (Open Government Licence). Accessible Green Infrastructure. Please note that all Urban Habitat and Naturalness Mapping done to date (April 2024) uses version 1.2 of the Green Infrastructure Mapping.
* Aerial Photography for Great Britain (Not open data).
* Ordnance Survey British National Grid (Open Government Licence).
* Ordnance Survey Master Map (Not open data).
* National Forest Inventory (Open Government Licence).
* Environment Agency National LiDAR Programme (Open Government Licence).
* OS Open Built-Up Areas (Open Government Licence).
* Priority Habitat Inventory. Coastal Habitats, Wetland Habitats (Open Government Licence).
* Moorland Line (Open Government Licence).

Urban Habitat Map coverage is intended to be urban only. Data outside Built Up Areas is for context only.

The Urban Habitat mapping approach is specifically designed to work within Built Up Areas (BUA). The maps also provide information on the land outside of the BUA, but this is provided very generally and is for context only. Other data sources should be used to understand areas outside of the BUA, such as the Living England maps.

Note that the Naturalness maps are still developmental and whilst the emthoid to their development is included in this section, no maps have yet been published and they do not form part of the V 2.1 release.

## History of the development of Urban Habitat and Naturalness Maps.

Methodological development for Urban Habitat and Naturalness Mapping was conducted during 2021/22 and focussed on the pilot cities: Plymouth, Cambridge and City of Manchester.

Following successful piloting, a second phase of work was undertaken during 2022/23 that further developed the methodology and then developed approaches to upscaling its application to large city conurbations. This phase resulted in the creation of Urban Habitat and Naturalness Maps for Tyneside, Greater Manchester and Greater Birmingham using an amended approach that was more streamlined and more applicable to desk-based application for large scale urban areas.

Further work is being undertaken to expand coverage.

## Purpose and use.

Urban Habitat and Naturalness Mapping data is intended to improve our understanding of;

* The physical composition of the urban ecosystem (Urban Habitats).
* It’s quality using “Naturalness” as a proxy.
* Change in these parameters of the Urban Ecosystem in England over time.

## Overall approach developed as initial mapping for Cambridge, Plymouth and City of Manchester.

The data processing approach to developing Urban Habitat and Naturalness Maps is complex and cannot reasonably be presented in this report.

However, a detailed step by step user guide to developing Urban Habitat and Naturalness Maps using Trimble eCognition software has been developed and is available on request to Natural England.

The overall approach to undertaking the process of developing Urban Habitat and Naturalness maps is set out in figure 8.



Figure 8. Steps in the Urban Habitat and Naturalness Mapping approach.

The Aerial Photography for Great Britain (APGB) imagery was the primary dataset used in the classification workflow. It provided spectral features in the red, green, blue and near-infrared (NIR) parts of the electromagnetic spectrum at a spatial resolution of 50 cm. Additionally, the APGB Digital Terrain Model (DTM) and Digital Surface Model (DSM) were downloaded. Other supplementary datasets were also used, including the Ordnance Survey Master Map (OSMM) to extract building and private garden footprints, the National Forest inventory (NFI) to extract information about woodland types, and the Environment Agency (EA) DTM and point cloud (LAZ). Following a short investigation, the EA LiDAR DTM and point cloud were preferred to the APGB DTM and DSM, due to higher spatial resolution and enhanced details in vegetation mapping. Despite the higher spatial resolution of some of the data sources, the minimum mappable unit for the final classified map was 5 sqm.

For the pilot, the broad and detailed habitat classification maps achieved accuracies ranging from 73% to 87% and 60% to 75% respectively, with Plymouth performing best, closely followed by Manchester City. One of the potential reasons for this was the time of APGB data acquisition – Plymouth was collected in June to July, which is ideal for vegetation mapping as it coincides with peak greening, while Manchester City was collected in April to May and Cambridge in early April during the leaf flushing period.

For the first trial (Greater Manchester, Tyneside and Birmingham, Black Country and Solihul) accuracies achieved ranged from 81 to 91% for Broad Habitats and 59 to 87% for Detailed Habitats. Again, rates were affected by imagery capture dates (with early and later season dates generating lower accuracy rates) but in addition, accuracy rates Detailed Grassland classes proved to be lower than most other classes due to inherent difficulties in grassland differentiation using Earth Observation data. These difficulties are compound where data capture dates were early or late season.

The overall accuracy levels for the pilot and Phase 1 trial mapping are summarised in tables 9 and 10.

**Table 9. Overall accuracy levels for Broad and Detailed Urban Habitats for the Pilot and Phase 1 Trial of the Urban Habitat Mapping.**

|  |  |  |  |
| --- | --- | --- | --- |
| Location | Tile Date | Broad | Detailed |
| City of Cambridge | 05/04/20 | 73 | 60 |
| City of Plymouth | 22/06/19 and 04/07/19 combined | 87 | 75 |
| City of Manchester | 22/04/19 and 22/05/19 combined | 82 | 60 |
| West Midlands | 30/04/2022 | 81 | 61 |
| Tyneside | 26/08/2019 | 91 | 87 |
|  | 21/09/2019 | 85 | 66 |
|  | 25/06/2020 | 91 | 72 |
|  | 19/04/2021 | 84 | 66 |
| Greater Manchester | 22/04/2019 | 80 | 64 |
|  | 22/05/2019 | 81 | 61 |
|  | 23/05/2019 | 81 | 59 |
|  | 30/05/2019 | 88 | 80 |
| Range |  | 73 to 91 | 59 to 87 |
| Mean |  | 84 | 68 |

**Table 10. Range and mean values for accuracy levels for Broad Urban Habitat Classes and usual confusions in identification. See table 11 for full Urban Habitat Classification system used.**

|  |  |  |  |
| --- | --- | --- | --- |
| Urban Habitat | Range | Mean | Confusion |
| A – Grasslands | 60 to 88 | 73 | C, E, F, H sometimes B, D and J and K. |
| B – Woodlands | 63 to 98 | 82 | A, C, E and sometimes F, G, H, J |
| C – Rough, Abandoned and Derelict Land | 10 to 84 | 56 | A, B, E sometimes D, H, J |
| D – Wetlands | 80 to 100 | 96 | A, B, E, H sometimes C, F |
| E – Impervious and non-vegetated | 55 to 88 | 75 | A, B, C, D sometimes F, H, I |
| F – Private Gardens | 83 to 100 | 86 | A, B, C sometimes E,G, H |
| G – Formal Planting | 64 to 100 | 90 | A, B, C sometimes E, F, H |
| H – Parklands | 85 to 95 | 92 | B, D, E sometimes A, C, G, J |
| I - Coastal | 84 to 85 (Small sample) | 84 | H |
| J - Agricultural | 80 to 97 | 88 | A, E sometimes C, F, H |
| K – Upland Habitats | 80 to 85 (Small sample) | 82 | A, D, E, H |

## Classification systems.

Following development during the pilot phase and further testing during the first phase of trialling, an Urban Habitat Classification System was adopted as set out in table 11.

Table 11. The system of Broad and Detailed Habitat Classes relating the 30 Detailed Urban Habitat Classes to the 11 Broad classes.

| Broad Key | Broad Class Name | Detailed Key | Detailed Class Name |
| --- | --- | --- | --- |
| A | Grasslands | A1 | Amenity Grasslands |
| A | Grasslands | A2 | Undifferentiated Grassland |
| B | Woodlands | B1 | Broadleaved, Mixed and Yew Woodland |
| B | Woodlands | B2 | Conifer-Dominated woodland |
| B | Woodlands | B3 | Isolated and Scattered Trees. |
| C | Rough, Abandoned and Derelict Land | C1 | Habitat Mosaics (Not currently mapped) |
| C | Rough, Abandoned and Derelict Land | C2 | Scrubs |
| D | Wetlands | D1 | Open Water |
| D | Wetlands | D2 | Vegetated Wetlands |
| E | Impervious and Non-vegetated | E1 | Sealed Surfaces and Buildings |
| E | Impervious and Non-vegetated | E2 | Vegetated Building Structures and Green Roofs |
| E | Impervious and Non-vegetated | E3 | Bareground |
| F | Private Gardens | F1 | Non-vegetated Gardens |
| F | Private Gardens | F2 | Vegetated Gardens |
| F | Private Gardens | F3 | Garden Trees |
| F | Private Gardens | F4 | Garden Scrubs |
| G | Formal Planting | G2 | Allotments |
| H | Parklands | H1 | Parkland Amenity Grassland |
| H | Parklands | H2 | Parkland Undifferentiated Grassland |
| H | Parklands | H3 | Parkland Wood Pasture |
| H | Parklands | H4 | Parkland Scrubs |
| I | Coastal | I1 | Coastal Sand |
| I | Coastal | I2 | Coastal Dunes |
| I | Coastal | I3 | Coastal Shingle, Loose and Bare Rocks |
| I | Coastal | I4 | Coastal Mud |
| I | Coastal | I5 | Coastal Saltmarsh |
| I | Coastal | I6 | Coastal Cliffs and Slopes |
| J | Agricultural Land | J1 | Vegetated Fields |
| J | Agricultural Land | J2 | Ploughed Fields |
| K | Upland Habitats | K1 | Upland Habitats |

## Use of Green Infrastructure Contexts.

Urban Habitat Maps are generated with a minimum mappable unit of 5 square metres. All the urban area is mapped (including buildings and manmade surfaces) and the mapping specifically separates and highlights habitat classes that occur within two key contexts of specific interest – Private Gardens Space and Parklands (predominantly Accessible Green Infrastructure spaces). For both Gardens and Parklands, habitats involving grassland, scrubs and trees are mapped as specific Detailed Habitats.

## Urban Habitat Classification detailed overall process.

The urban habitat classification was carried out using “Trimble eCognition 10.2”. This is a commercial software which provides advanced image segmentation tools to perform Object-Based Image Analysis (OBIA), as opposed to pixel-based analysis. OBIA relies on grouping pixels of similar spectral responses together into objects and allows the user to extract object features that could not be acquired using pixel-based techniques alone. These object features include spectral statistics, geometry, texture and context (relations to neighbouring, sub or super-objects in a hierarchical classification system). “eCognition” also allows pre-defined workflows (also called rulesets) to be applied to “scenes” in a batch process, making it an ideal candidate for large scale analysis. The datasets prepared were ingested into “eCognition” using a pre-defined XML file to automatically create a project containing all relevant layers for each city of interest.

The main stages of the ruleset are as follows:

1. Land feature extraction from existing third-party vector datasets.

2. Band calculation to generate a Digital Surface Model (DSM), a Canopy Height Model (CHM), a greyscale image and spectral indices.

3. Area of Interest (AOI) delineation and image segmentation to create spectrally distinct objects.

4. Initial object classification based on landscape features rather than land use.

5. Detailed classification to combine features and their context within the GI.

6. Clean-up of the detailed classification.

7. Broad classification derived from the detailed classes.

8. Export of final maps.

9. Accuracy assessment.

### Land Feature Extraction from Third-Party Vector Datasets.

Land features were extracted from existing third-party vector datasets and stored in their own temporary layers within the “eCognition” project. Such features include;

Building footprints, private gardens, paths and natural spaces from the OSMM. Parklands, waterbodies, woodlands and allotments from the GI database These are crucial in supporting the analysis, particularly to provide context (e.g. within vs outside of parks or private gardens). Building footprints also aid in the detection of green roofs, which may otherwise be confused with trees (elevated vegetation), whilst waterbodies help in picking up vegetated wetlands, which may be confused with other low-lying vegetated areas.

### Band Calculation.

Digital Surface Model (DSM) and Canopy Height Model (CHM). A Digital Surface Model (DSM) was created by rasterising the EA LiDAR point cloud with a kernel size of 3. The original point cloud density (before rasterisation) was 1 point density per sqm (1ppsqm) on average. This means that the actual “spatial resolution” of the LIDAR DSM could be assumed to be about 1m.

This number is not exact as point cloud density varies across the scan. During rasterisation, some pixels may contain higher point density than 1, and other pixels may have no data values. Linear interpolation was used in these instances.

This ensured sufficient detail was retained, whilst reducing the size of gaps. Linear interpolation was then carried out to fill the gaps in the DSM, but a waterbody mask was used to prevent artefacts over water surfaces. During analysis, all input datasets were resampled to 50 cm pixel size, which corresponds to the APGB CIR spatial resolution. As a result, the LiDAR point cloud was also rasterised to 50 cm pixel size.

Once the DSM was finalised, a Canopy Height Model (CHM) was calculated by subtracting the Digital Terrain Model (DTM) from the DSM. The spatial resolution of the DTM was 1m per pixel, but it was resampled to 50 cm pixel size during analysis. Similarly, the pixel size of the resulting CHM was 50cm, but the actual spatial resolution is closer to 1m per pixel.

Greyscale Image. The RGB imagery was used to calculate a greyscale image using varying band weights: Greyscale = 0.299 Red + 0.592 Green + 0.114 Blue.

Spectral Indices. The RGB and Colour Infrared (CIR) aerial datasets provided 6 bands from which to generate a series of spectral indices. However, two of these bands were duplicates (Red and Green). Table 12 shows the spectral indices calculated.

When choosing bands to calculate an index, bands from the same source dataset were used. E.g., for NDVI, only bands from the CIR dataset were used. For GRVI, only bands from the RGB dataset were used. This was to reduce potential artefacts from combining data sources with different spatial resolution together when creating these indices.

Spectral ranges for the APGB RGB were not available but ranges for the CIR were as follows.

NIRF18A (Near Infrared – NIR) Spectral range 690 to 1000 nm.

REDF14A (Red) Spectral range 580 to 700 nm.

GRNF16A (Green) Spectral range 480 to 630 nm.

Table 12. Table setting out the Spectral Indices used in Urban Habitat Mapping.

|  |  |  |
| --- | --- | --- |
| Spectral Index. | Acronym. | Calculation. |
| Green-Red Vegetation Index. | GRVI | (RGB Green – RGB Red) / (RGB Green + RGB Red) |
| Normalised Difference Soil Index. | NDSI | (Blue – RGB Red) / (Blue + RGB Red) |
| Normalised Difference Vegetation Index. | NDVI | (NIR – CIR Red) / (NIR + CIR Red) |
| Normalised Difference Water Index. | NDWI | (NIR – CIR Green) / (NIR + CIR Green) |
| Soil-Adjusted Vegetation Index. | SAVI | 1.5 \* (NIR – CIR Red) / (NIR + CIR Red + 0.5) |

### Area of Interest (AOI) Delineation & Image Segmentation.

The target areas Area of Interest (AOI) boundary was first used to delineate the analysis. Multi-resolution segmentation was then applied to the AOI using the CIR Red, CIR Green, NIR and CHM layers and a scale of 10. A higher weight was given to the NIR band. The segmentation was constrained to allotments, parks, waterbodies, building footprints, private gardens and paths to avoid objects overlapping two different contextual features.

### Initial Object Classification.

The initial classification was focused on features of the landscapes, regardless of land use, such as; water, buildings, other impervious surfaces, bare ground, trees, scrubs and low-lying vegetation. All objects were classified using a threshold-based approach and a combination of spectral features, relational features, height, geometry and third-party data information. Shadows were quite extensive in the imagery and were also separated at this stage using brightness values. Brightness is automatically calculated in eCognition using all available input layers. However, these values are not normalised and can drastically change from scene to scene.

The thresholds were obtained through trial and error and thoroughly tested during the trial.

Table 13 shows the different features of the landscape in order of classification, as well as the data source and conditions used to separate them for Cambridge specifically. A summary of differing threshold values for both Cambridge and Plymouth is given in Table 13. This approach was essentially adopted for all further Urban Habitat Mapping.

Table 13. Summary of initial object classification focusing on features of the landscape rather than land use (where source information is not required, cells in the table are left blank).

| Feature | Source | Condition |
| --- | --- | --- |
| Shadow | Spectral | All objects with mean brightness values less than 60. |
| Water | GI | All objects (including shadows) that overlap GI waterbodies. |
| Buildings | OSMM | Remaining objects (including shadows) that overlap building footprints. |
| Impervious | OSMM, GI, Spectral | Remaining objects that either (1) overlap paths and parks, (2) overlap natural spaces by less than 15% and have mean NDVI and NDSI values smaller than 0.3 and 0.1 respectively, (3) overlap natural spaces by smaller than 15% and have a mean NDVI value smaller than 0.1, (4) overlap natural spaces by more than 15% and have mean NDVI and NDSI values smaller than 0.2 and 0 respectively. |
| Bareground | GI, Spectral | Remaining objects with mean SAVI value smaller than 0.2. Impervious objects that overlap allotments and have a mean SAVI value smaller than 0.3. |
| Trees | Height, Spectral, Relational, OSMM | Remaining objects (including shadows) with a mean CHM value more than 2m.Water objects with a mean CHM value greater than 2m and neighbouring another tree object (allows some tree growth over water areas).Impervious objects that overlap paths and have mean CHM and NDVI values greater than 2m and 0.3 respectively (Allows tree growth over paths).Building objects with mean brightness and NDVI values greater than 70 and 0.2 respectively and a mean CHM difference with other building objects greater than 0 and neighbouring another tree object (Allows tree growth over buildings).Tree objects with a mean NDVI value smaller than 0.2 and a mean CHM difference with other building objects smaller than 0 and neighbouring another building object (Removes false positives along building edges).Remaining building objects that share 100% of their border with trees and have a mean NDVI value greater than 0.25 (Fills holes in trees overhanging buildings). |
| Scrubs | Height, Relational, Geometry | Remaining objects with a mean CHM value between 1m and 2m.Water objects with a mean CHM value between 1m and 2m and neighbouring another scrub object (Allows scrubs over water).Remaining objects (including water) with a mean CHM value greater than 0.5 and neighbouring another scrub object.Remaining objects with a mean CHM value greater than 0.5 and a roundness smaller than 0.3 (Helps to locate isolated scrubs). |
| Coastal | Spectral, relational | Remaining objects (including bareground) that are within 20 of tidal waters and have mean values for SAVI, NDVI and greyscale of -0.1, -0.1 and 100 respectively. |
| Low Vegetation | - | All remaining objects. |

### Detailed Classification.

The detailed classification built on the initial object classification but added context, most often by looking for the presence of certain GI or OSMM features and adjusting the classes accordingly. For example, a tree object found within a private garden would be labelled as a garden tree (F3), but one found in a park would be labelled as park wood pasture (H3) and one found on a street would be labelled as an isolated/scattered tree (B3). Thresholds in combination with spectral features, relational features, geometry and third-party data were used, with the exception of grassland classification, which relied on training a machine learning model.

Table 14 summarises the detailed classes in order of classification, as well as the data source and conditions used to separate them. This stage is mostly contextual, and so requires very little adjustments for other cities as long as the initial classification is refined and accurate.

Table 14. Summary of detailed classification in order of classification in the Urban Habitat Mapping assessment process (where source information is not required, cells in the table are left blank).

| Class | Source | Condition |
| --- | --- | --- |
| Conifer Woodland (B2) | Geometry, NFI | Merged tree objects larger than 0.5 ha and that overlap NFI conifers. |
| Mixed Broadleaved Woodland (B1) | Geometry | Remaining merged tree objects larger than 0.5 ha |
| Isolated and Scattered Trees (B3) | Geometry | Remaining merged tree objects. |
| Garden Trees (F3) | OSMM | Isolated/Scattered Trees (B3) that overlap private gardens. |
| Parklands Wood Pasture (H3) | OSMM | Remaining isolated/Scattered Trees (B3) that overlap parks. |
| Non-vegetated Gardens (F1) | OSMM | Bareground and Impervious objects that overlap private gardens. |
| Vegetated Gardens (F2) | OSMM | Low Vegetation objects that overlap private gardens. |
| Garden Scrubs (F4) | OSMM | Scrub objects that overlap private gardens. |
| Parkland Scrubs (H4) | GI | Remaining scrub objects that overlap parks. |
| Scrubs (C2) | - | Remaining scrub objects. |
| Green Roofs (E2) | Spectral | Building objects that have mean NDVI and NDSI and brightness values greater than 0.2 and 0.05 and 70 respectively. |
| Sealed Surfaces and Buildings (E1) | - | Remaining building and impervious objects. |
| Bareground (E3) | - | All bareground objects. |
| Coastal Sand (I1) | Spectral | Coastal objects that have a mean GRVI value of - 0.03. |
| Coastal Shingle, Loose and Bare Rocks (I3) | Spectral | Remaining coastal objects that have a mean GRVI value between -0.3 and 0. |
| Vegetated Wetlands (D2) | Spectral | Water objects that have mean NDVI and NDSI values greater than 0.2 and 0.05 respectively. |
| Open Water (D1) | - | Remaining water objects. |
| Allotments (G2) | GI | All objects (including shadows but excluding woodlands) that overlap allotments. |

Agricultural Land and Upland Habitats were added as contextual classes outside of the Built Up Areas during the first trial project.

### Machine Learning for Grassland Classification.

Amenity grasslands are heavily maintained – they are kept relatively short and tend to be species poor, resulting in a homogeneous landscape. In contrast, undifferentiated grasslands often contain a mix of species of different heights and spectral signatures. In an attempt to separate the two habitat classes based on their homogeneity, Grey Level Co-occurrence Matrix (GLCM) texture layers were generated using the greyscale image and the CHM for low-lying vegetated areas derived from the initial object classification:

* Greyscale and CHM contrast layers – higher values indicate higher contrast between neighbouring pixels in the object.
* Greyscale and CHM entropy layers – higher values indicate lower orderliness and a higher level of randomness in the object.
* Greyscale correlation layer – higher values indicate that neighbouring pixels in the object have predictable and linear relationships between them.

Amenity and undifferentiated grasslands have proved difficult to separate using thresholds alone. As a results, a Random Forest machine learning model was trained using the samples collected during the desk-based survey. The samples were first simplified to combine all grasslands of the same type together, whether outside or within parks (H1 combined with A1 and H2 combined with A2), thus increasing the pool of data. The samples for each grassland type were then split in half in a random manner, with 50% used in training the model, and the remaining 50% reserved to generate a confusion matrix and assess the accuracy of the model.

The mean object features fed into the model were as follows:

* Spectral bands: RGB Red, RGB Green, Blue, CIR Red, CIR Green, NIR
* Spectral indices: GRVI, NDSI, NDVI, NDWI, SAVI
* LiDAR: CHM
* GLCM textures: CHM contrast, CHM entropy, greyscale contrast, greyscale correlation, greyscale entropy

RGB and CIR spectral bands were used because spectral information about the APGB RGB dataset was lacking. Because of this, it couldn’t be assumed that the bands were exactly the same. In addition, the original datasets have different spatial resolutions, yielding different average pixel values when resampling. As a result, all bands were retained for mapping grasslands using Machine Learning (Random Forest). This allowed for testing of the usefulness of each.

The features of greatest importance in the model related to the CHM which provides height and structural information. The overall accuracy for grassland classification in Cambridge consistently exceeded 75%. Amenity and undifferentiated grasslands within parks (as defined in section 2.2) were then converted to the correct detailed class (H1 or H2).

### Clean-Up.

This stage aims to clean-up and simplify the outputs before the final broad classification. All neighbouring objects of the same detailed class are first merged together. Shadows and objects smaller than 5 sqm are then combined with neighbouring objects based on spectral similarity.

### Broad Classification.

The broad classification is performed on a hierarchical level above the detailed classification, allowing broad classes to be inherited directly from the detailed level below. For example, if an object is classified as a garden tree (F3) on the detailed level, it automatically inherits the private garden (F) class at the broad level. All neighbouring objects of the same broad class are then merged together to create larger parcels.

### Accuracy assessments.

For the pilot project, accuracy checks included some field verification. However, one of the objectives of the first phase of trialling was to amend the process so that it can be run for large scale whole city regions. Accuracy assessment was restricted to desk-based assessment performed directly on the final maps, rather than through digitisation prior to classification. “eCognition” objects for each urban habitat class at both detailed and broad levels were assessed using stratified sampling. If the assessor disagreed with the assigned class, a corrective class was suggested. This process was done on a selected sample of map output tiles. Confusion matrices were then generated to assess the Producer Accuracy (PA) and User Accuracy (UA) of each class, as well as the Overall Accuracy (OA) for each tile.

As a result of the inherent identification confusions resulting in variable accuracy levels, the resultant Urban Habitat Maps must be considered as “probability maps”.

An example Confusion Matrix is given below. Actual accuracy rates vary by habitat class and date tile for the aerial imagery so that a confusion matrix has to be generated for each date tile. Accuracy can also be affected by date mismatches between the Aerial tiles and the LiDAR data. This matrix in table 15 is purely illustrative of the approach for the Broad Habitat Class system.

The table shows where classification samples agree between initial output (producer) and accuracy check (user) and where they do not, and which habitat class confusions have occurred.

An overall accuracy is also calculated.

Table 15. Accuracy Assessment Confusion Matrix detailing accuracy rates for each habitat class, incidence, and type of confusion with other classes and an overall Aerial data tile accuracy. For example, in this illustrative case, Habitat Class A (Grasslands) had 28 samples with 3 having confusions with Scrubs and agricultural land. This gave producer accuracy of 89% and User Accuracy of 63%. The overall tile accuracy (calculated across all classes) was 91%.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Classification** | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** | **J** |
| **A** | 25  | 0  | 1  | 0  | 11  | 1  | 0  | 0  | 2  |
| **B** | 0  | 42  | 1  | 0  | 1  | 0  | 0  | 0  | 0  |
| **C** | 2  | 0  | 15  | 0  | 2  | 0  | 0  | 1  | 0  |
| **D** | 0  | 0  | 0  | 26  | 0  | 0  | 0  | 0  | 0  |
| **E** | 0  | 0  | 0  | 0  | 52  | 0  | 0  | 1  | 4  |
| **F** | 0  | 0  | 0  | 0  | 0  | 78  | 0  | 0  | 0  |
| **G** | 0  | 0  | 0  | 0  | 0  | 0  | 20  | 0  | 0  |
| **H** | 0  | 0  | 0  | 0  | 6  | 1  | 0  | 71  | 0  |
| **J** | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 30  |
| **Total samples** | **28** | **42** | **17** | **26** | **72** | **80** | **20** | **73** | **36** |
| **Producer accuracy** | **89** | **100** | **88** | **100** | **90** | **100** | **100** | **90** | **97** |
| **User accuracy** | **63** | **95** | **75** | **100** | **90** | **100** | **100** | **90** | **96** |
| **Overall accuracy.** | **91** |  |  |  |  |  |  |  |  |

## Summary tables for definitions and thresholds.

Table 16. Summary table setting out the definitions and any thresholds used to identify Broad Class Urban Habitats.

|  |  |
| --- | --- |
| Broad Urban Habitat Class | Identification and thresholds |
| Grassland | Spectral thresholds, vegetation below 1M. |
| Woodlands | Spectral thresholds, brightness and CHM <2M. |
| Rough, Abandoned and Derelict Land | Currently limited to “Scrubs”. CHM 1 to 2 m. Roundness < 0.5 m.  |
| Wetlands | Green Infrastructure Database – Blue Infrastructure Network. |
| Impervious and non-vegetated surfaces. | OSMM Buildings and sealed surfaces and spectral thresholds. |
| Private Gardens | OSMM |
| Formal Planting | OSMM Allotments. |
| Parklands | OSMM / Green Infrastructure data - the following GI assets were considered as Parklands. Access Land (CRoW), Activity Spaces Provision, Cemeteries and Religious Grounds, Golf Courses, Other Sports Facilities, Play Space Provision, Playing Fields, Country Parks, General Public Parks, Millennium or Doorstep Greens. |
| Coastal | OSMM and PHI. |
| Agricultural Land | Agricultural land outside Built Up Areas. |
| Upland Habitats | All areas above the Moorland Line – undifferentiated. |

Table 17. Summary table setting out the definitions and any thresholds used to identify Detailed Urban Habitat Classes.

| Detailed Urban Habitat Classes. | Identification and thresholds. |
| --- | --- |
| Amenity Grassland. | Grasslands detected via machine learning (random samples 50:50 split) as Amenity. |
| Undifferentiated Grassland. | Grasslands detected via machine learning (random samples 50:50 split) as Undifferentiated. |
| Broadleaved, mixed and Yew Woodlands. | NFI broadleaved woodland <2M and greater than 0.5 ha not overlapping Parklands or Gardens. |
| Conifer Dominated Woodlands. | NFI conifer dominated <2M and greater than 0.5 ha not overlapping Parklands or Gardens. |
| Isolated and Scattered Trees. | Greater than 2M not overlapping Parklands or Gardens. |
| Habitat Mosaics. | Not identified in current method. |
| Scrubs. | All vegetation between 1 and 2M outside Parklands and Gardens. |
| Open Water. | Green Infrastructure BI Network data. |
| Vegetated Wetlands. | Open water with spectral signature for vegetation and PHI data. |
| Sealed surfaces and buildings. | OSMM and spectral thresholds. |
| Vegetated building surfaces and Green Roofs. | Buildings above required spectral threshold for vegetation. |
| Bareground. | Spectral thresholds outside Allotments. |
| Non-vegetated Gardens. | Sealed surface and bare ground within gardens. |
| Vegetated Gardens. | Vegetation below 1M in Gardens. |
| Garden Trees. | Greater than 2M and overlapping Gardens. |
| Garden Scrubs. | Vegetation between 1 and 2 M in a garden (Cannot differentiate Scrub, from Shrub from Hedge from scattered bush). |
| Allotments. | Vegetation below 1M and Bareground within Allotments (derived from GI Mapping and OSMM). |
| Parkland Amenity Grasslands. | Grasslands detected via machine learning (random samples 50:50 split) as Amenity and in Parklands. |
| Parkland Undifferentiated Grassland. | Grasslands detected via machine learning (random samples 50:50 split) as Undifferentiated and within Parklands. |
| Parkland Wood Pasture. | Greater than 2M, Smaller than 0.5 Ha and overlapping parklands. |
| Parkland Scrubs. | Vegetation between 1 and 2M within Parklands. |
| Coastal Habitats. | OSMM and PHI data. |
| Vegetated Fields. | Areas of low vegetation (below 1 m) outside Built Up Areas that overlap “Agricultural land” class in OSMM. |
| Ploughed Fields. | Bareground outside of Built Up Areas that overlap “Agricultural land” class in OSMM. |
| Upland Habitats. | All areas above the Moorland Land – undifferentiated. |