Living England: From satellite imagery to a national-scale habitat probability map

Alex Kilcoyne, Miles Clement, Chris Moore, Amy Woodget, Max Fancourt, Anne Stefaniak, Becky Trippier & Sophie Potter
earth.observation@naturalengland.org.uk

What is Living England?

- An up-to-date, national-scale map of predicted habitat distributions in England (Phase 4: 2021/22).
- Developed by a team with expertise in ecology & field surveying, Earth observation & data science.
- Created using openly-licenced satellite imagery, geospatial data & field-derived habitat data in a machine learning framework.
- Based on broad parcels of land known as 'segments' which sit within wider biogeographic zones (BGZs).
- Accuracy varies between different habitats & regional zones, with average habitat mapping accuracy reported at 88.39% in Phase 4.

What can Living England be used for?

Living England is intended to help inform:

- Environmental policy & decision making (e.g. Environmental Land Management)
- National habitat extent and connectivity assessments
- Natural capital asset assessment
- Ecosystem service modelling
- Plans for nature recovery

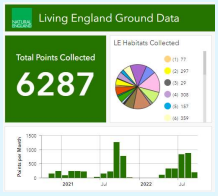
Ecology + Earth Observation + Data Science = Living England

How is the Living England habitat map created?

STEP 1. DATA ACQUISITION, LICENCING & FRAMEWORK

1a. Acquisition of raw geospatial data: Including satellite imagery, OS mapping & data on geology, soils, crops & climate.

1b. Acquisition of ground truth or label data: From existing open-source habitat data, fieldwork and desktop surveys (Esri Field Maps app).



1c. Classification framework: Adapted UKBAP.

1d. Data licensing & permissions: For land access & data.

STEP 2. DATA INGESTION & PREPARATION

2a. Prepare geospatial data: Including Sentinel-1 coherence & backscatter processing & Sentinel-2 mosaicking and cloud/cloud-shadow masking.

2b. Prepare ground truth data: Translate to UKBAP, QC, extract points within segments & statistical validation to remove outliers.

2c. Spatial zonation: Split data into 13 biogeographic zones.

STEP 3. SEGMENTATION

3a. Run standardised segmentation:

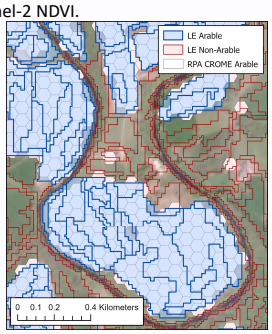
- Uses Trimble eCognition v10.3 software.
- Uses set parameters established through professional judgment to optimise segment size & shape.
- Standardised approach helps maintain consistency for subsequent change analyses.

STEP 4. VECTOR-BASED CLASSIFICATION

4a. Create vector layers:

- Arable & Horticulture: RPA's Crop Map of England.
- Coastal Saltmarsh: Bespoke mapping using S2 and LiDAR data alongside intertidal modelling.
- Water: OS Mastermap.
- Urban Areas: OS Mastermap.
- Bare Ground: bespoke mapping from Sentinel-1 coherence & Sentinel-2 NDVI.

4b. Vector-based classification of above habitats: A hierarchical selection process based on known reliability of underlying datasets and proportion of overlap between vector layer and segment.

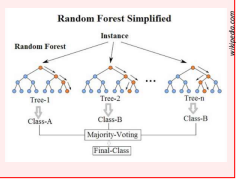


STEP 5. MODEL-BASED CLASSIFICATION

5a. Compute zonal statistics: for geospatial data layers, by segment. Link ground truth data to zonal statistics for spatially concurrent segments.

5b. Model training & parameterisation: train individual random forest models for each BGZ. Select the most significant input variables and model parameters, alongside an iterative desktop validation process.

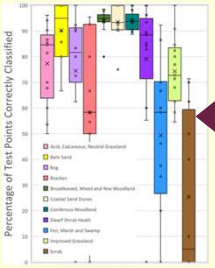
5c. Run models: to predict the top 'most likely' habitat classes for all segments in all BGZs.

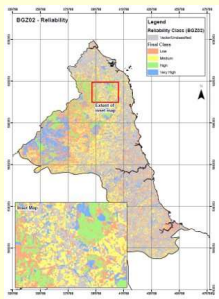


STEP 6. RESULTS & ANALYSIS

6a. Merge vector-based & model-based classifications

6b. Assess model accuracy & confidence: overall measures & broken down by habitat class & BGZ.





6c. Assess model reliability: develop a bespoke reliability scoring system based on known quantitative data & feedback from partners.

STEP 7. OUTPUTS & DELIVERY

Download LE Data

7a. Export Living England Habitat Map: Published under OGL in April '22.

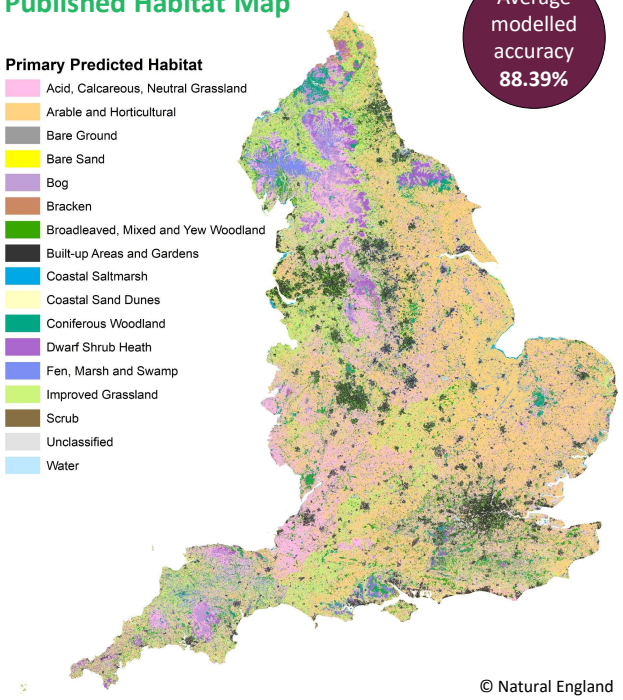
7b. Publish report: Kilcoyne, A.M. *et al.*, (2022) Living England: Technical User Guide. NERR108. Natural England.

7c. Engage with key partners: e.g., ELM, JNCC, EA, NatureScot.

7d. Other outputs: scripts on GitHub.

Living England Phase 4 (2021/22) Published Habitat Map

Average modelled accuracy **88.39%**



© Natural England

Phase 4 known limitations:

- Undermapping of urban areas
- Overmapping of sand dunes
- Overmapping of fen, marsh & swamp in areas of bog & ACN grasslands (esp. in BGZ01)
- Confusion between grasslands
- Particularly low accuracy in BGZ03
- Imbalance in ground truth data across BGZs & habitats

Recent developments:

- Method standardisation
- Addressing known limitations
- Developing an iterative process for modelling & QA
- Further developing the reliability metric
- Exploring new input data (e.g. LiDAR, OS Mastermap)
- Development of change detection workflows
- Commissioning new surveys & gathering data from latest monitoring programmes

What are the next steps?

Living England 2023-24

- Produce an updated Living England habitat map
- Ongoing internal and external field data collection
- Beta version to be released internally March '24, externally Sept '24

Change Detection

- Develop approach for assessing broad habitat change every two years using the Living England datasets
- Identify key stakeholder and end user requirements

NCEA Natural Capital and Ecosystem Assessment

NATURAL ENGLAND