Received: 18 August 2025 | Accepted: 4 March 2026

DOI: 10.1002/2688-8319.70234

RESEARCH ARTICLE

BRITISH ECOLOGICAL SOCIETY Ecological Solutions and Evidence

Evaluating the accuracy of the EcoservR toolkit for fine-resolution habitat mapping

Sandra Angers-Blondin[1] | Colm Bowe[1,2] | Luke Kinross Bentley[1] | Alex Owusu Amoakoh[1,2] | Joe Bellis[1] | Lucy Dowdall[1] | Noémie Bonnin[1] | Stamatia Galata[1] | Hannah Branwood[1] | Andrew Clark[3] | Chloe Bellamy[4] | Hanna Partoft[1]

[1]Natural Capital Hub, Department of Biological and Environmental Sciences, Liverpool John Moores University, Liverpool, UK

[2]Liverpool Research Institute for Climate and Sustainability (LiRICS), Liverpool John Moores University, Liverpool, UK

[3]Merseyside Environmental Advisory Service (MEAS), Liverpool, UK

[4]Forest Research, Northern Research Station, Roslin, UK

Correspondence
Alex Owusu Amoakoh
Email: a.o.amoakoh@ljmu.ac.uk

Handling Editor: Ceres Barros

Abstract

1. Accurate and spatially explicit habitat maps are essential for monitoring ecological change, supporting nature-based solutions and informing policy. Despite advances in habitat mapping, classification accuracy varies across landscapes and thematic resolutions, highlighting the need for systematic evaluation of classification workflows.
2. This study evaluates EcoservR, an open-source rule-based toolkit for habitat mapping, using field survey data from two contrasting landscapes in northern England: Merseyside (MER) and the North York Moors (NYM).
3. Habitat classifications were generated by integrating multiple national spatial datasets and were evaluated against ecological survey records and high-resolution aerial imagery across three levels of the Phase 1 habitat classification hierarchy developed by the Joint Nature Conservation Committee (JNCC).
4. Overall accuracy ranged from 0.612 ± 0.015 (Level 3, NYM) to 0.804 ± 0.012 (Level 1, NYM). Broad habitat classes such as woodland and standing water achieved high precision, while grasslands, heathlands and mires showed lower precision. Accuracy declined with increasing thematic detail, reflecting ecological heterogeneity and limitations in the resolution and currency of input datasets. Observed land-cover patterns corresponded with the known ecological structure of each landscape, although classification biases indicate that fine-scale habitat assessments and quantitative ecosystem-service estimates require supplementary data.
5. *Practical implications*. EcoservR provides a reproducible workflow for integrating multiple spatial datasets to generate spatially consistent habitat maps. This approach supports regional planning and nature-recovery initiatives where broad habitat categories are sufficient, and its rule-based architecture allows adaptation to other landscapes where comparable datasets are available.

KEYWORDS
ecosystem services, habitat mapping, modelling, Phase 1 survey, rule-based classification

1 | INTRODUCTION

Global policy frameworks increasingly emphasise the role of nature-based solutions in responding to climate and biodiversity challenges (Mori et al., 2021; Seddon et al., 2020). International agreements such as the Kunming–Montreal Global Biodiversity Framework (CBD, 2022) and the Paris Agreement (UNFCCC, 2016) require countries to produce spatially explicit ecological information to support monitoring, restoration planning and natural capital accounting. This has increased the demand for fine-resolution habitat baselines that represent landscape characteristics in a consistent and policy-relevant form (Ruijs et al., 2019).

Methodologically, habitat mapping spans approaches based on remote sensing of satellite imagery and workflows that derive habitat classes from administrative mapping products (Amoakoh et al., 2021, 2024; Punalekar et al., 2024; Richiardi et al., 2025). Image-based methods perform well where training data are extensive and habitats are spectrally distinct, but their reliability declines in heterogeneous or transitional landscapes and where reference data are limited. Administrative mapping products provide consistent polygon boundaries and national coverage, yet do not inherently represent ecological habitat types and require explicit rule-based translation from descriptive attributes to ecological classes. These datasets also vary in spatial resolution, thematic detail and update frequency, limiting interoperability and complicating harmonisation for biodiversity assessment and natural-capital applications (Department for Environment, Food and Rural Affairs [Defra], 2025a, 2025b; IPBES, 2019; Office for National Statistics, 2024). In this context, a transparent rule-based approach offers a pragmatic means of operationalising existing national datasets, prioritising reproducibility, interpretability and alignment with policy reporting requirements over spectral optimisation.

Against this backdrop, EcoservR was developed as an open-source toolkit to formalise this translation process, making classification logic explicit and enabling systematic evaluation of how administrative land-use descriptors are mapped to ecological habitat classes (Angers-Blondin et al., 2020). The toolkit re-implements the earlier Ecoserv-GIS system developed by Durham Wildlife Trust (Winn et al., 2018) in the R environment to enhance reproducibility and eliminate dependence on proprietary GIS platforms. The workflow extracts polygon-level attributes from OS MasterMap Topography (Ordnance Survey, 2010) and assigns habitats according to the Phase 1 habitat system (JNCC, 2010), a hierarchical field-survey framework comprising broad habitat types, vegetation groups and finer structural categories. Deterministic decision rules based on polygon descriptors are applied initially and subsequently refined using ancillary datasets, including semi-natural habitats of importance, forest inventories, crop map, hedgerow datasets and elevation models. The resulting basemap provides a consistent thematic representation at the spatial resolution of MasterMap parcels and forms the basis for ecosystem-service modelling.

The toolkit has been applied in several operational contexts, including mapping ecosystem-service opportunities for the Liverpool City Region Local Nature Recovery Strategy (Liverpool City Region Combined Authority, 2025) and to explore the delivery of public goods and ecosystem-service gains from agri-environment interventions as part of Defra's Environmental Land Management (ELM) Test and Trial programme (Angers-Blondin et al., 2020). In principle, the resulting basemap products can support Natural Capital accounting frameworks, including the UK Natural Capital Register and Account Tool (Office for National Statistics, 2024), and inform environmental planning, restoration prioritisation and nature-based investment. Despite its growing application, the habitat classification underpinning EcoservR has not been evaluated against an independent ecological survey, limiting understanding of its reliability for decision-making where classification errors may bias estimates of ecosystem extent, condition or service potential.

The workflow also integrates datasets that vary in spatial resolution, survey date and thematic scope, which may influence polygon-level habitat assignment in ways that remain undocumented. This study therefore evaluates the accuracy and limitations of EcoservR habitat basemaps using independent ecological survey data from two contrasting landscapes in northern England: a predominantly urban and peri-urban region (Merseyside) and a largely agricultural and upland environment (North York Moors National Park). Specifically, we assess general and thematic accuracy across the first three levels of the Phase 1 habitat classification hierarchy, identify systematic patterns of misclassification and examine circumstances under which rule-based habitat assignment is most prone to error.

2 | METHODS

2.1 | Study area

The study was conducted in Merseyside (MER) and the North York Moors (NYM) in northern England (Figure 1). MER corresponds to the Liverpool City Region Mayoral Combined Authority and covers approximately 817 km^2 on the west coast around the Mersey Estuary. The area comprises a mix of built-up urban surfaces, agricultural land and habitats of conservation interest, including coastal saltmarshes, mudflats, sand dunes, lowland raised bog and semi-natural woodlands (MEAS and LCRCA, 2023). NYM lies within the North York Moors National Park and covers approximately 413 km^2, encompassing moorland, woodland and farmland. The study area

BRITISH ECOLOGICAL SOCIETY Ecological Solutions and Evidence

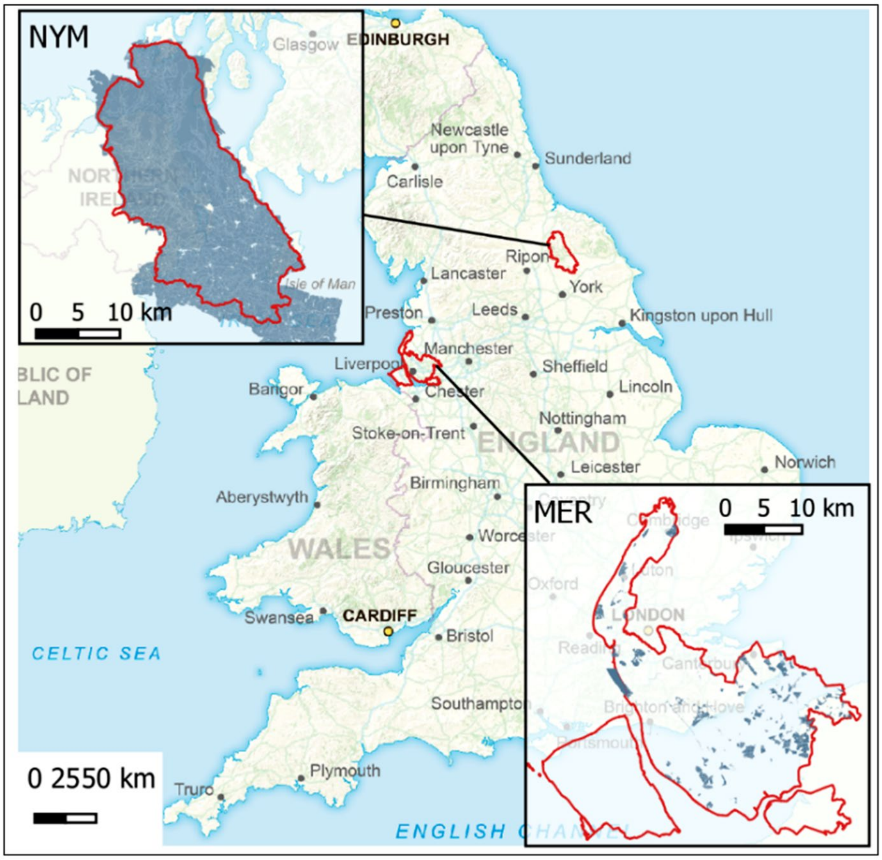

FIGURE 1 Location of the North York Moors (NYM) and Merseyside (MER) in England. Insets show each study area boundary (red outline), with grey shading indicating the spatial coverage of available ecological survey data used for validation.

aligns with the boundaries of the Ryevitalise ecological restoration project (North York Moors National Park Authority, 2018). Together, these areas represent contrasting urban, rural and semi-natural landscapes, providing an appropriate context for evaluating habitat classification performance across different environmental settings, at the spatial scale intended for its application (catchment, administrative region).

2.2 | Habitat classification scheme

Habitat classification in this study follows the hierarchical Phase 1 habitat classification system (JNCC, 2010), a standard UK framework for mapping semi-natural habitats used in local and national assessments. The system comprises three nested levels of thematic detail, ranging from broad habitat groups to finer structural and vegetation-based categories (Table 1). As Phase 1 terminology is not widely used outside the United Kingdom, these levels are here referred to as Levels 1–3 to indicate increasing thematic resolution. To ensure consistency between predicted and reference datasets, habitat codes were harmonised across the first three levels of the Phase 1 hierarchy. Where reference survey codes did not correspond directly to EcoservR outputs, they were aggregated to a coarser level (see Table 2). This ensured comparability between datasets and reduced bias arising from mismatches in thematic resolution between predicted and reference classifications.

2.3 | Reference data and sampling design

Field-based ecological habitat surveys conducted using Phase 1 habitat survey methods (JNCC, 2010) were used as reference datasets for validation. In these surveys, trained personnel systematically survey sites on foot, record dominant vegetation and land-cover characteristics, and delineate habitat units that are subsequently digitised and attributed using the Phase 1 classification. For Merseyside, reference data were obtained from the Merseyside Environmental Advisory Service and comprise surveys of local wildlife sites conducted since 2010. For the North York Moors, reference data were provided by the North York Moors National Park Authority and were updated in 2017, drawing on earlier field surveys conducted between 1987 and the 1990s and supplemented by interpretation of aerial imagery by ecological specialists. Although these datasets were not collected specifically for validating EcoservR outputs, they represent expert-led, systematic survey programmes and therefore provide a defensible basis for assuming higher thematic accuracy relative to automatically derived habitat classifications (Stehman & Foody, 2008).

TABLE 1 Hierarchical Phase 1 habitat codes used by EcoservR.

Habitat type	Level 1	Level 2	Level 3	Description
Woodland and scrub	A	A1	A11	Woodland; broadleaved
			A12	Woodland; coniferous
			A13	Woodland; mixed
		A2	A21	Scrub; dense
			A22	Scrub; scattered
		A3	A31	Parkland; broadleaved
			A32	Parkland; coniferous
			A33	Parkland; mixed
Grassland and marsh	B	Bu (B1, B2, B3)	Bu1	Grassland; (unknown soil); unimproved
			Bu2	Grassland; (unknown soil); semi-improved
		B4		Grassland; improved
		B5		Grassland; marshy
		B6		Grassland; poor; semi-improved
Tall herb and fern	C	C1	C11	Bracken; continuous
			C12	Bracken; scattered
		C3	C31	Other tall herb and fern; tall ruderal
			C32	Other tall herb and fern; non-ruderal
			C3u	Other tall herb and fern; (unknown vegetation)
Heath	D	D1	D11	Dry dwarf shrub heath; acidic
			D12	Dry dwarf shrub heath; basic
		D2		Wet dwarf shrub heath
		D3		Lichen/bryophyte heath
		D4		Montane heath/dwarf shrub
		D5		Mosaic of acid grassland and dry heath
		D6		Mosaic of acid grassland and wet heath
Mire	E	E1	E16	Bog; blanket or raised
			E17	Bog; wet modified
			E18	Bog; dry modified
		E2	E21	Flush and spring; acidic or neutral
			E22	Flush and spring; basic
		E3	E31	Fen; valley mire
			E32	Fen; basin mire
			E33	Fen; flood-plain mire
		E4		Bare peat
Swamp	F	F1		Swamp
Open water	G	G1	G16	Standing water; brackish or saline lagoons
		G2	G26	Running water; brackish or tidal
		G3		Sea
Coastal	H	H1	H11	Intertidal; mud/sand
			H12	Intertidal; shingles/cobbles
			H13	Intertidal; boulders/rocks
		H2	H24	Saltmarsh; scattered
			H26	Saltmarsh; dense/continuous
		H3		Shingle above high-tide mark
		H4		Boulders/rocks above high-tide mark

BRITISH ECOLOGICAL SOCIETY Ecological Solutions and Evidence

TABLE 1 (Continued)

Habitat type	Level 1	Level 2	Level 3	Description
		H6	H64	Sand dune; dune slack
			H65	Sand dune; dune grassland
			H66	Sand dune; dune heath
			H67	Sand dune; dune scrub
			H68	Sand dune; open dune
		H8	H81	Maritime cliff and slope; hard cliff
			H82	Maritime cliff and slope; soft cliff
			H84	Maritime cliff and slope; coastal grassland
			H85	Maritime cliff and slope; coastal heathland
Rock exposure and waste	I	I1	I11	Natural rock; inland cliff
			I12	Natural rock; scree
			I13	Natural rock; limestone pavement
			I14	Natural rock; other exposure (e.g. on mountaintops or in riverbeds)
		I2	I21	Artificial exposure and waste; quarry
			I22	Artificial exposure and waste; spoil
			I23	Artificial exposure and waste; mine
			I24	Artificial exposure and waste; refuse-tip
Miscellaneous	J	J1	J11	Cultivated/disturbed land; arable
			J12	Cultivated/disturbed land; amenity grassland
			J13	Cultivated/disturbed land; ephemeral/short perennial
			J14	Cultivated/disturbed land; introduced scrub
		J2	J21	Boundaries; intact hedge
			J22	Boundaries; defunct hedge
			J23	Boundaries; hedgerow with trees
			J24	Boundaries; fence
			J25	Boundaries; wall
			J26	Boundaries; dry ditch
			J28	Boundaries; earth bank
		J3	J34	Built-up area; caravan site
			J35	Built-up area; sea wall
			J36	Built-up area; buildings
			J37	Built-up area; sealed surface
		J4		Bare ground
		J5	J51	Other habitat; road
			J52	Other habitat; roadside/pavement
			J53	Other habitat; railway
			J54	Other habitat; path
			J55	Other habitat; gardens/parks/brownfield sites
			J56	Other habitat; private garden

Sampling points for validation were generated using a systematic grid design, with points placed at regular 100 m intervals across each study area to ensure spatially even coverage and minimise clustering effects. This resulted in 4459 sampling points for MER and 39,578 sampling points for North York Moors, reflecting differences in the spatial extent and completeness of available survey coverage (Figure 1). Accuracy results are therefore reported separately for each study area and interpreted as area-specific assessments rather than as direct comparisons between landscapes. Reference datasets pre-date the EcoservR input layers by approximately 4–11 years; the implications of this temporal mismatch are addressed explicitly through adjudication procedures described in Section 2.5.

TABLE 2 Harmonisation of reference habitat codes to EcoservR output classes.

Reference code(s)	Original description	Harmonised class used in analysis	Rationale
G11, G12, G13, G14, G15	Running water	G1	EcoservR outputs only broad water classes; aggregation ensures comparability
G21, G22, G23, G24, G25, G26	Standing water	G2	Same as above; second-digit distinctions not available in predictions
B11, B21, B31	Acid, neutral, calcareous unimproved grassland	Bu1 (unknown soil, unimproved)	Soil type distinction not produced by EcoservR
B12, B22, B32	Acid, neutral, calcareous semi-improved grassland	Bu2 (unknown soil, semi-improved)	Soil type distinction not produced by EcoservR
Buildings, roads, hardstanding	Built-up surfaces	Excluded	Not recorded in reference surveys; would bias accuracy estimates

2.4 | EcoservR habitat classification

EcoservR was used to generate a harmonised habitat basemap by integrating multiple nationally available spatial datasets (Table 3). The R package implements a deterministic, rule-based workflow in which habitat classes are assigned to Ordnance Survey (OS) MasterMap polygons through an ordered sequence of decision rules. These rules (Table S1) operate primarily on core MasterMap attributes, including *Descriptive Group*, which indicates broad feature type; *Descriptive Term*, which provides more specific information on physical or functional characteristics (e.g. non-coniferous trees, scrub, rough grassland, static water or foreshore); and *Make*, which distinguishes natural from manmade features. Classification decisions also draw on polygon geometry and a set of ancillary datasets, including CORINE Land Cover, the National Forest Inventory, the Priority Habitat Inventory and the Crop Map of England. The rule-based framework provides an explicit and reproducible means of assigning habitat classes without reliance on training data or parameter optimisation.

An initial habitat classification is first produced for all MasterMap polygons and is subsequently refined through the integration of thematic datasets with higher ecological specificity. Priority Habitat Inventory layers identify habitats of conservation significance such as deciduous woodland, lowland meadows, blanket bog and coastal priority habitats. The National Forest Inventory contributes information on woodland extent and structure, including areas of recent felling. The Crop Map of England and CORINE Land Cover support differentiation between arable land, improved grassland and unimproved grassland and correct agricultural polygons misclassified during the initial assignment. Greenspace datasets provide contextual information on amenity land and public accessibility, while elevation and slope derived from a 5 m digital elevation model are used to refine upland grasslands, montane habitats and slope-dependent transitions. Ancillary datasets are imported in their native vector or raster formats for the area of interest. Spatial projection inconsistencies are resolved prior to rasterisation of key attributes (e.g. land cover class) to a 2 m grid. The modal attribute value is then extracted and assigned to each MasterMap polygon. Hedgerow datasets are processed separately to improve identification of linear woody features while avoiding destructive changes to the MasterMap framework. They can be incorporated into subsequent ecosystem service modelling alongside the main habitat map. Decision rules are applied deterministically in a fixed sequence, such that datasets with greater thematic certainty override broader or less specific sources. The resulting habitat map integrates multiple national datasets within a transparent and internally consistent classification framework.

At the time of analysis, EcoservR was under active development by the Natural Capital Hub at Liverpool John Moores University and had no formally versioned release. The development version used in this study is publicly accessible via GitHub (https://github.com/ecoservR/ecoserv_tool) and is updated iteratively to incorporate user feedback and functional improvements. Habitat maps for each study area were generated using R version 4.1.2 (R Core Team, Vienna, Austria) by clipping national datasets to study boundaries, harmonising coordinate reference systems and resolving spatial overlaps prior to executing the full EcoservR rule set to produce final habitat maps (Figure 2). The workflow is designed to accommodate large and heterogeneous spatial datasets while maintaining consistency with nationally used land-cover and habitat classifications relevant to environmental policy applications.

2.5 | Accuracy assessment and adjudication

Habitat map accuracy was evaluated using the systematic sampling framework described in Section 2.3, consistent with best practice for land-cover accuracy assessment (Olofsson et al., 2014).

BRITISH ECOLOGICAL SOCIETY Ecological Solutions and Evidence

TABLE 3 Spatial datasets used for land cover and habitat analysis, including their sources, year represented, spatial resolution and update frequency.

Datasets	Source	Year represented	Spatial resolution/type	Update frequency/ temporal notes	Role in classification workflow
OS MasterMap Topography Layer*	Ordnance Survey	2021	Vector polygons	Updated approximately every 6 weeks	Provides baseline geometry and primary attributes (Descriptive Group, Descriptive Term, Make) that underpin the first-round rule-based habitat classification
OS MasterMap Greenspace Layer*	Ordnance Survey	2021	Vector polygons	Updated approximately every 6 months	Identifies parks, gardens, amenity greenspace and private gardens; used to refine classifications and determine access types
CORINE Land Cover	European Environment Agency	2018	100 m raster	Updated every 6 years; coarse resolution for UK applications	Distinguishes broad land-cover types (e.g. pastures vs. arable); used jointly with CROME where available
Crop Map of England (CROME)	Rural Payments Agency	2019	2 m raster	Represents a single agricultural year	Differentiates arable crops from improved grasslands; overrides CORINE where classifications conflict
Priority Habitat Inventory (PHI)	Natural England	2022	Vector polygons	Compiled from diverse survey years; not consistently updated	Assigns statutory conservation habitats (e.g. lowland meadows, blanket bog, coastal saltmarsh), overriding earlier rules
National Forest Inventory (NFI)	Forestry Commission	2020	Vector polygons	Multi-temporal, periodically updated	Adds woodland structure information, particularly identifying felled stands
BlueSky/GetMapping DEM*	BlueSky/ GetMapping	2022	5 m raster	High-resolution elevation; regularly updated regionally	Used to refine upland versus lowland habitats, slope-based grassland rules and montane thresholds
Hedgerow Data	Rural Payments Agency	2022	Vector lines	Derived from agricultural submissions; inconsistent years	Used to assign hedgerow habitat classes along linear woody features

Note: Datasets marked with an asterisk (*) were accessed under the Digimap educational licence or the Aerial Photography Great Britain agreement, while others were obtained through government open data portals. National hedgerow data were provided on request by the Rural Payments Agency.

Sampling points were intersected with both predicted habitat maps and reference survey polygons to construct confusion matrices summarising agreement and disagreement between classifications. Spatial processing was undertaken using the *sf* package (version 1.0–5) (Pebesma, 2018) in R, and confusion matrices and accuracy metrics were calculated using the *caret* package (version 6.0-94) (Kuhn, 2008). Accuracy was assessed separately for Merseyside and the North York Moors and reported for the first three levels of the Phase 1 habitat classification hierarchy, reflecting increasing thematic resolution. To account for differences between sample counts and mapped area proportions, habitat class fractions derived from the predicted maps were used to convert raw confusion-matrix counts into area-weighted proportions following Olofsson et al. (2014). Built-up classes absent from the reference data were excluded from this step. From the area-weighted confusion matrices, overall accuracy (OA), class-level precision and recall, F1 score and weighted F1 score were calculated (Equations 1–5). These metrics capture complementary aspects of classification performance, including overall agreement, class-specific reliability and sensitivity, and the influence of class imbalance on aggregate accuracy.

$$\text{Precision}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i}, \tag{1}$$

$$\text{Recall}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i}, \tag{2}$$

$$\text{F1}_i = 2 \times \frac{\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}, \tag{3}$$

$$\text{OA} = \frac{\text{TP}_i + \text{TN}_i}{\text{TP}_i + \text{FP}_i + \text{TN}_i + \text{FN}_i}, \tag{4}$$

$$\text{Weight F1} = \sum_{i=1}^{k} w_i \times F1_i, \quad w_i = \frac{n_i}{N}, \tag{5}$$

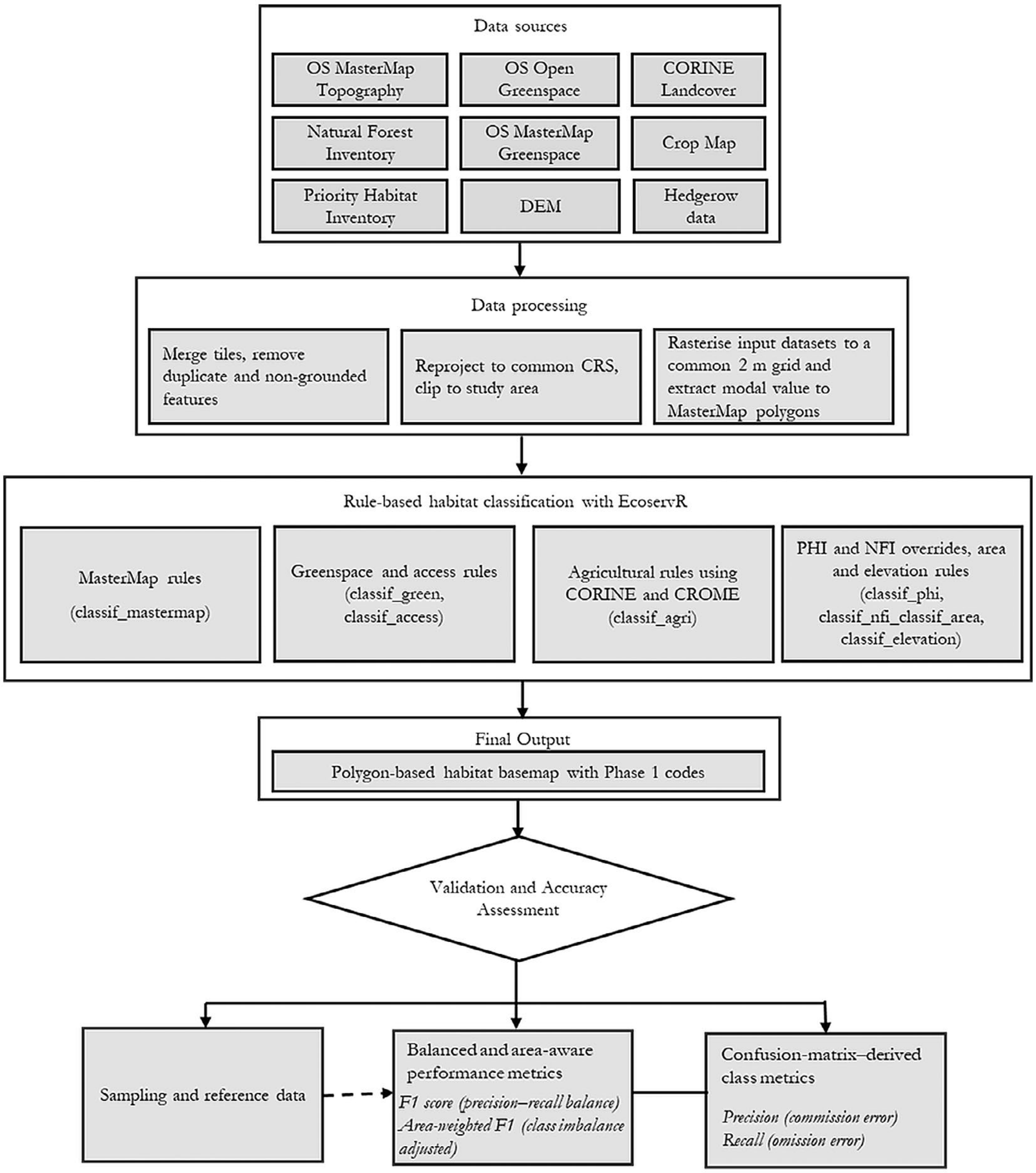

FIGURE 2 Workflow for data integration and rule-based habitat classification using EcoservR.

where TP_i represents the true positives, that is, samples correctly predicted as class i, FP_i represents false positives, that is, samples incorrectly predicted as class i, FN_i represents false negatives, that is, samples belonging to class i but predicted as another class and TN_i represents true negatives, that is, samples correctly predicted as not belonging to class i. k is the total number of classes, N is the total number of samples w_i is the weight for class i, calculated as the proportion of samples in class i and n_i is the number of samples in class i. The F1 score summarises the balance between precision and recall for each habitat class, while the weighted F1 score accounts for class imbalance by aggregating class-specific F1 values in proportion to their representation in the sample. These metrics ensured a robust assessment of classification accuracy across multiple habitat classes.

A standard assumption in map validation is that reference data provide a more accurate representation of land cover than predicted outputs (Stehman & Foody, 2008). In this study, reference datasets were assumed to offer a suitable benchmark but were not collected specifically for validating EcoservR and pre-dated the predicted habitat maps by approximately 4–11 years. To address potential discrepancies arising from temporal landscape change or limitations in the reference data, disagreements between predicted and reference classifications were examined through a structured adjudication process. For each study area, a random subset of 200 sampling points where classifications disagreed was selected for detailed review.

These points were compiled into a blind adjudication dataset in which habitat codes were anonymised and presented as two alternatives without indicating their source. Each point was examined using high-resolution aerial imagery (EDINA, n.d.) and assigned one of five outcomes: (1) predicted class correct; (2) reference class correct; (3) neither correct, indicating an alternative habitat; (4) both appropriate, reflecting valid but different representations of the landscape or classification at different hierarchical levels; or (5) inconclusive,

BRITISH ECOLOGICAL SOCIETY Ecological Solutions and Evidence

where imagery quality or landscape context prevented a clear decision. Ambiguous cases were reviewed internally to reach consensus, consistent with recommended practice for interpreting classification disagreement and uncertainty (Olofsson et al., 2014; Stehman & Foody, 2019). This adjudication enabled differentiation between errors attributable to classifier logic and those arising from temporal change or legitimate alternative interpretations, providing essential context for interpreting quantitative accuracy metrics.

3 | RESULTS

3.1 | Classification accuracy

Classification accuracy declined consistently with increasing thematic detail across both study areas (Table 4). Overall accuracy was highest at the broadest level of the Phase 1 hierarchy (76%–80%) and lowest at Level 3 (61%–67%), with weighted F1 scores showing the same trend. At the class level, woodland and tree habitats (A) and water (G) showed consistently high precision and recall across hierarchical levels, whereas tall herb and fern (C) and mire (E) were not detected in one or both study areas (Table 5).

Grassland-related classes and cultivated land accounted for the largest share of misclassifications. In particular, grasslands (B codes) were frequently confused with arable land (J11), while ruderal vegetation (C), scrub (A2) and parkland (A3) were often classified as amenity land (J12) (Figure 3). Heathland habitats (D codes) were primarily misclassified within the same broad group, such as confusion between dry and wet heath, and were rarely assigned to unrelated habitat classes (Figure 3).

3.2 | Hierarchical accuracy based on adjudicated mismatches

Expert adjudication of randomly sampled points where EcoservR and the Phase 1 reference produced discordant habitat labels revealed systematic differences across hierarchical levels (Table 6). At Level 1, the Phase 1 reference label was more frequently accepted than the EcoservR classification. At the intermediate level, the two sources were accepted at similar rates. At Level 3, EcoservR was marginally more often preferred, and the proportion of cases judged valid under both classification schemes increased. As habitat categories became more detailed, it was increasingly common for EcoservR and the reference map to assign different habitat labels that were each reasonable for the same location. In such cases, the disagreement did not reflect a clear error by either method, but instead reflected the fact that complex or transitional habitats can be legitimately described in more than one way at finer levels of classification.

Several habitat groups showed relatively strong performance by EcoservR at Level 3, with the model being accepted in at least 40% of adjudicated mismatches (Table 7). These included coastal habitats, standing and running water, swamp, and certain grassland and flush communities.

3.3 | Spatial distribution of predicted habitats

Figure 4 presents the EcoservR habitat maps for the two study areas, showing the spatial distribution of predicted Phase 1 habitat classes across contrasting landscape contexts. In Merseyside, the mapped landscape is dominated by cultivated and disturbed land, coastal habitats and urban features, including gardens, parks and built-up areas. In contrast, the North York Moors are characterised by extensive semi-natural habitats, with large contiguous areas of grasslands and marshes, heathlands and woodland. Coastal habitats are confined to Merseyside, while heathlands and mires are prominent in the North York Moors and largely absent from Merseyside. These spatial patterns reflect the known urban–rural contrast between the two regions and demonstrate how the rule-based workflow translates heterogeneous national datasets into coherent habitat baselines. Class fractions derived from the predicted maps are summarised in Table 8 to quantify these spatial patterns.

4 | DISCUSSION

4.1 | Classification accuracy across hierarchical levels

The decline in classification accuracy with increasing thematic detail is consistent with long-standing observations in hierarchical habitat mapping, where finer class definitions introduce greater ecological variability and ambiguity (Foody, 2002; Hazeu, 2014). In practice, this reflects a shift from classes defined primarily by broad structure

TABLE 4 Error adjusted accuracy of the EcoservR basemap in Merseyside and the North York Moors based on the first three levels of the Phase 1 hierarchy.

	Merseyside		North York Moors	
	OA	Weighted F1	OA	Weighted F1
Level 1	0.759 ± 0.014	0.731 ± 0.016	0.804 ± 0.012	0.790 ± 0.013
Level 2	0.703 ± 0.014	0.667 ± 0.016	0.643 ± 0.015	0.648 ± 0.015
Level 3	0.670 ± 0.015	0.624 ± 0.016	0.612 ± 0.015	0.620 ± 0.015

TABLE 5 Thematic accuracy metrics across the three Phase 1 hierarchy levels (see Table 1 for codes). Dashes indicate absent or undetected classes. Classes with an underscore suffix (e.g. B4_, B6_) indicate cases where predictions and reference data differed in hierarchical depth but were retained in the confusion matrix following recoding.

	Merseyside				North York Moors		
	Habitat	Precision	Recall	F1-score	Precision	Recall	F1-score
Level 1	A	0.93	0.54	0.69	0.93	0.92	0.93
	B	0.47	0.22	0.30	0.65	0.71	0.68
	D	—	—	—	0.86	0.99	0.92
	G	0.93	1.00	0.97	0.75	1.00	0.86
	H	0.96	0.93	0.95	—	—	—
	I	—	—	—	1.00	1.00	1.00
	J	0.68	0.94	0.79	0.82	0.79	0.81
Level 2	A1	0.87	0.60	0.71	0.93	0.92	0.93
	A2	0.25	0.05	0.08	—	—	—
	B4	0.46	0.29	0.36	—	—	—
	B6	—	—	—	0.62	0.56	0.59
	Bu	0.20	0.04	0.07	—	—	—
	D1	—	—	—	0.04	0.5	0.08
	E1	—	—	—	1.00	1.00	1.00
	E2	—	—	—	0.67	1.00	0.80
	G1	1.00	1.00	1.00	0.15	0.46	0.23
	G2	—	—	—	1.00	1.00	1.00
	H1	0.98	1.00	0.99	—	—	—
	H6	0.95	0.56	0.70	—	—	—
	I2	—	—	—	0.83	0.78	0.80
	J1	0.70	0.84	0.76	0.00	0.00	0.00
	J5	0.48	1.00	0.64	-	-	-
Level 3	A11	0.80	0.66	0.73	0.70	0.78	0.74
	A12	0.43	0.38	0.40	0.90	0.74	0.81
	A13	0.35	0.46	0.40	0.27	0.41	0.32
	B4_	0.46	0.48	0.47	—	—	—
	B6_	—	—	—	0.61	0.61	0.61
	Bu2	0.15	0.03	0.06	—	—	—
	C3_	—	—	—	0.04	0.50	0.07
	D6_	—	—	—	1.00	1.00	1.00
	E18	—	—	—	0.75	1.00	0.86
	G1_	0.95	1.00	0.97	0.25	0.08	0.13
	H11	0.84	1.00	0.91	—	—	—
	H65	0.48	0.44	0.46	—	—	—
	H67	1.00	0.14	0.25	—	—	—
	I21	—	—	—	0.83	0.78	0.80
	I22	—	—	—	0.25	0.50	0.33
	J11	0.82	0.86	0.84	0.08	0.31	0.13
	J12	0.45	0.95	0.61	—	—	—

and land cover to those that depend on vegetation composition, condition or management, where boundaries are less discrete and more context-dependent (Stehman, 2009; Tuanmu & Jetz, 2014). EcoservR followed this expected pattern, performing most reliably at the broadest level of the Phase 1 hierarchy and showing reduced accuracy at finer levels, particularly for grassland and heathland subclasses. Contrasts between the two study areas further illustrate these constraints. The map for the North York Moors, dominated by extensive natural and semi-natural habitats, showed higher accuracy at broader classification levels, whereas Merseyside's heterogeneous mosaic of urban land, disturbed ground and coastal environments posed greater challenges for consistent rule-based

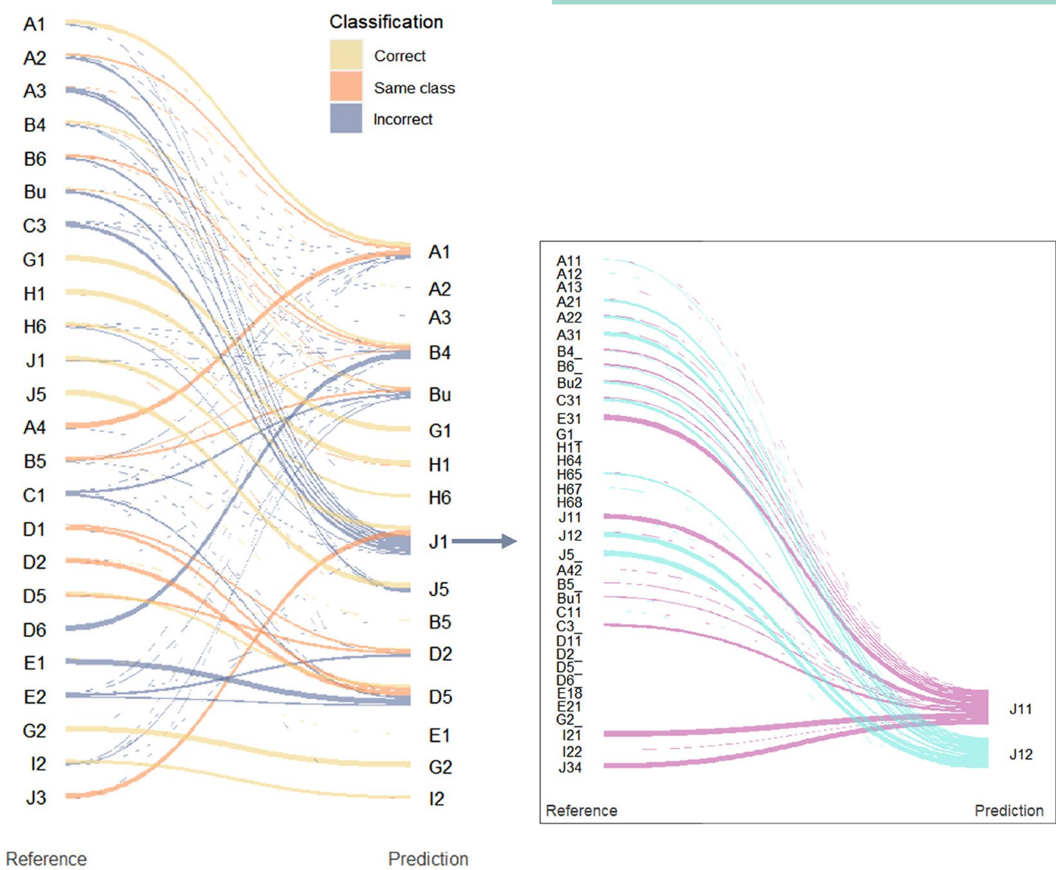

FIGURE 3 Sankey diagrams illustrating agreement and disagreement between reference habitat classes and EcoservR predictions. Flows show transitions from reference classes (left) to predicted classes (right), with colours indicating correct classification, misclassification within the same broad habitat group, and incorrect classification. The right panel presents a detailed example of misclassification patterns for class J1 (disturbed or cultivated land), highlighting the dominant confusion pathways between specific reference and predicted subclasses.

assignment. Landscapes characterised by sharp spatial transitions and mixed land uses are known to complicate categorical classification, particularly where deterministic rules must reconcile competing signals from multiple datasets (Herold et al., 2004; Weng, 2012). The parallel decline in weighted F1 scores across hierarchical levels indicates that reduced performance at finer resolution was not driven solely by class imbalance, but by genuine difficulty in resolving closely related habitat types.

These results highlight inherent limits of a deterministic, rule-based workflow when applied at fine thematic resolution. EcoservR can only discriminate habitat classes to the extent supported by the thematic and temporal detail of its input datasets. Where information on vegetation condition, soil characteristics or recent management is absent or inconsistent, ecologically similar subclasses cannot be reliably separated. This limitation is widely recognised in ecosystem mapping and natural capital accounting, where nationally consistent datasets tend to represent extent more robustly than condition or function (Hein et al., 2020; Keith et al., 2017). Although statistical and machine-learning approaches can sometimes improve discrimination by exploiting subtle multivariate patterns (Amoakoh et al., 2024; Belgiu & Drăguţ, 2016; Maxwell et al., 2018), EcoservR deliberately prioritises transparency, reproducibility and alignment with policy datasets over predictive optimisation. Such trade-offs reflect the requirements of policy-facing applications, where interpretability and auditability are often valued over marginal gains in classification accuracy (Bateman & Balmford, 2018; Kenter et al., 2019).

Given the limited spatial extent and contrasting character of the study areas, the findings here should be interpreted as site-specific rather than indicative of national-scale performance. Nonetheless, they demonstrate that EcoservR is well suited to broad-scale habitat mapping for policy reporting and monitoring, while underscoring the need for caution where outputs are used at finer thematic resolution without additional ecological or contextual data.

4.2 | Habitat-specific classification challenges

Classification performance varied substantially across habitat types. EcoservR achieved consistently high precision for woodland (A1), standing water (G1) and intertidal habitats (H1) (Table 5). These habitats share relatively clear physical structure and spatial boundaries that are consistently represented in national datasets such as OS MasterMap and the Priority Habitat Inventory, reducing ambiguity

TABLE 6 Proportion of points where the habitat label assigned by EcoservR, the reference data (Ph1) or both was judged correct by the assessor, disaggregated by thematic resolution. 'Unsure' denotes cases where the imagery was inconclusive, and 'None' denotes cases where neither label was accepted.

Hierarchy	EcoservR %	Ph1%	Both %	Unsure %	None %	Key message
Level 1	34.9	55.9	5.3	3.2	0.8	At the coarsest level, the reference data is chosen as the correct label in a clear majority of disagreements; EcoservR is second
Level 2	40.3	40.4	10.6	8.0	0.6	At the second level, the two classifiers are virtually tied; 'Both correct' and 'Unsure' rise to almost one fifth of cases
Level 3	36.6	35.4	20.4	7.2	0.5	At the finest level, the balance tips slightly towards EcoservR, and roughly one fifth of the mismatches were classified as 'both correct'

TABLE 7 It lists habitat groups for which EcoservR was accepted in ≥40% of the H3 mismatches.

Habitat group (Phase 1 code)	Discordant points (*n*)	Points resolved in favour of EcoservR (%)
Coastal shingle/strandline (H1_, H3_)	62	100
Coastal dune systems (H6_)	47	60
Running water (G2_)	24	75
Standing-water fringes (G16)	18	67
Improved grassland (B4_)	126	48
Marshy grassland (B5_)	79	46
Swamp (F1_)	36	50
Acid/neutral flush (E16)	30	50

in rule-based assignment. Similar patterns have been reported in United Kingdom and European habitat mapping exercises, where structurally distinct habitats tend to be classified more reliably than those defined primarily by species composition or ecological condition (Büttner, 2014; Hazeu, 2014). Woodland habitats, in particular, benefit from well-defined polygon boundaries across both vector and raster products, limiting opportunities for misclassification.

In contrast, grasslands (B codes) exhibited consistently lower and more variable precision and recall across study areas. Heathlands (D codes) and mires (E codes) were generally well identified at broader thematic levels where present but showed evidence of within-group confusion and reduced discrimination at finer levels, particularly where class distinctions depend on subtle differences in vegetation structure or moisture regime (e.g. dry vs. wet heath). These patterns reflect the inherent heterogeneity of grassland and wetland habitats, which often occur along ecological gradients or in transitional settings and are therefore difficult to resolve using nationally consistent datasets that provide limited information on vegetation condition, management intensity or hydrological status. Similar challenges are widely recognised in ecosystem service mapping and habitat-condition assessment, where grasslands and wetlands exhibit high within-class variability and weak correspondence with coarse thematic land-cover classes (Eigenbrod et al., 2010; Lavorel et al., 2017). Given that these habitats constitute a large proportion of semi-natural environments in temperate regions, uncertainty in their classification has direct implications for estimates of habitat extent and for downstream applications such as conservation prioritisation and ecosystem service assessment. For instance, in the UK context, where grasslands, heaths and bogs underpin many nature-recovery and land-use planning initiatives, such uncertainty may propagate into assessments of habitat condition and ecosystem function.

Grassland misclassifications were particularly evident in Merseyside, where managed and semi-natural grasslands were frequently confused with arable land (J11) or amenity grassland (J12). This pattern reflects a broader tendency for disturbed or intensively managed classes to be over-assigned in heterogeneous urban and peri-urban landscapes, where managed grasslands share contextual and structural characteristics with both arable and amenity land. Comparable limitations have been documented in large-scale land-cover datasets, particularly where urban form, land management and vegetation structure interact at fine spatial scales (Amoakoh et al., 2024; Horning et al., 2010; Wang et al., 2024). Heathlands and mires in the North York Moors also posed challenges, especially where distinctions depend on subtle differences in vegetation structure or moisture regime, such as between wet and dry heath. These distinctions are poorly resolved by coarse-resolution inputs such as CORINE Land Cover, a limitation that has been noted in European wetland and peatland inventories relying on harmonised land-cover products (Davidson et al., 2018; Rebelo et al., 2018).

Generally, these patterns point to areas where targeted refinement could improve EcoservR performance. Incorporating

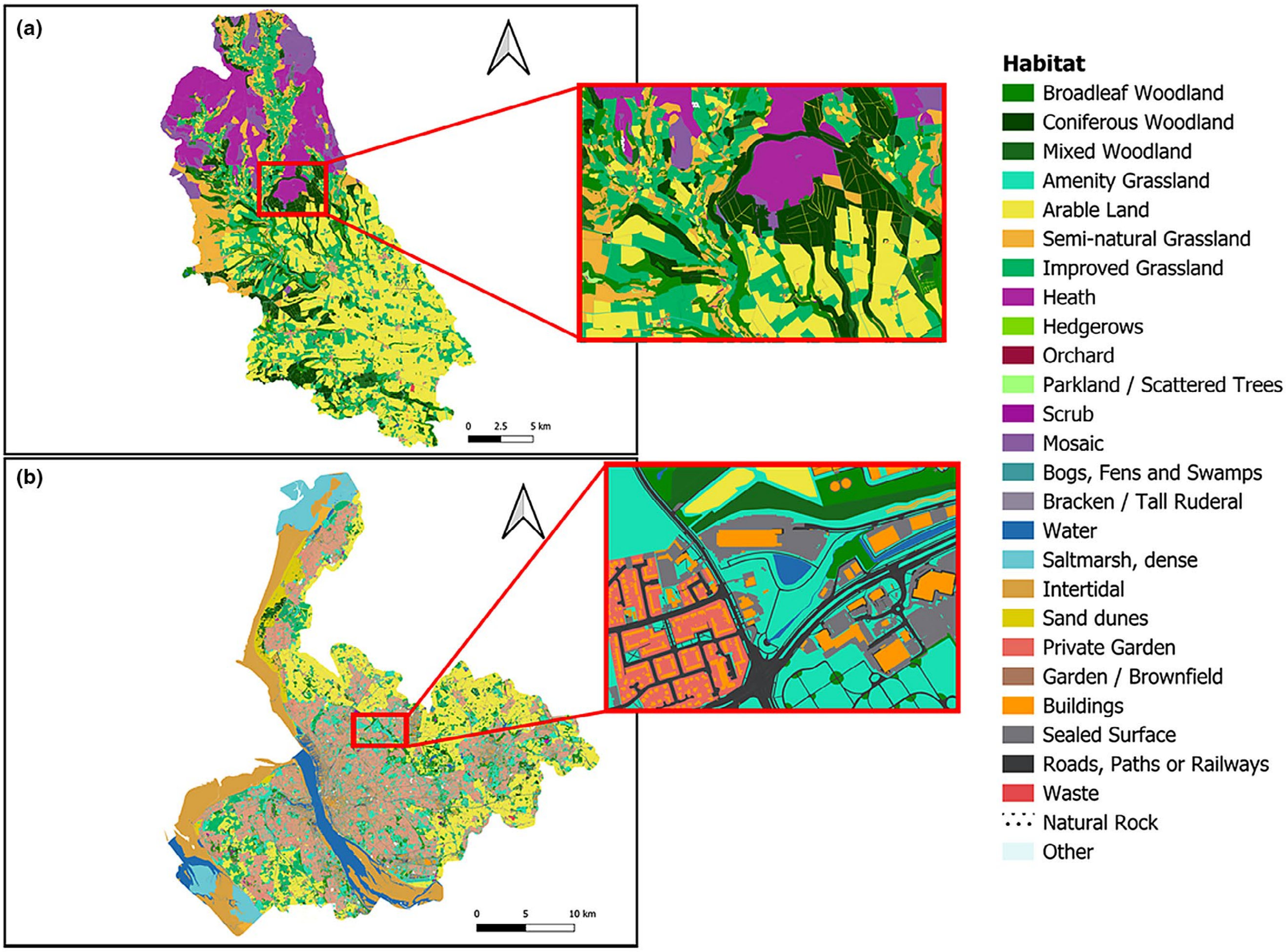

FIGURE 4 Predicted Phase 1 habitat classes produced by the EcoservR toolkit for (a) North York Moors (NYM) and (b) Merseyside (MER). Insets show fine-scale spatial detail. © Crown copyright and database rights 2021 Ordnance Survey.

finer scale datasets, such as high-resolution tree cover layers or LiDAR-derived vegetation metrics, would enhance detection of linear woody features, transitional habitats and grassland subtypes. Similarly, integrating indicators of habitat condition, including vegetation height, soil properties (e.g. moisture, organic matter content or drainage class) and grazing intensity, could improve discrimination among ecologically variable habitat classes that are currently difficult to resolve using rule-based classification alone. Such developments are increasingly emphasised in ecosystem accounting approaches that distinguish between ecosystem extent and condition as separate but complementary dimensions of assessment, central to policy-relevant environmental reporting (Office for National Statistics, 2024; UNSC, 2021).

4.3 | EcoservR performance relative to reference data

Ground-reference datasets are commonly treated as definitive benchmarks in land-cover and habitat validation, despite long-standing recognition that they are subject to observer bias, temporal mismatch and survey-specific constraints (Carlotto, 2009; Foody, 2010). The quality of reference data must be interrogated explicitly (Olofsson et al., 2014; Stehman & Foody, 2019), particularly where ecological surveys are reused outside their original design. Against this background, the comparison of EcoservR outputs with both Phase 1 survey data and high-resolution aerial imagery provides a useful lens through which to interpret disagreements. Across all hierarchical levels, EcoservR was judged to provide the more accurate habitat label in approximately one-third of adjudicated mismatches when evaluated against aerial imagery. This suggests that disagreement cannot be attributed solely to classifier error. In part, it likely reflects ecological change occurring after the original Phase 1 surveys were conducted, which EcoservR may capture more effectively due to closer temporal alignment with the imagery (Foody, 2010; Olofsson et al., 2014). In addition, the reference datasets were compiled over extended periods by multiple surveyors, introducing variability in habitat delineation and interpretation. Such observer-related uncertainty is well documented in Phase 1 surveys, particularly for habitats with diffuse boundaries, mixed vegetation or transitional characteristics (Cherrill, 2016; Cherrill & McClean, 1999).

TABLE 8 Overview of land cover for Merseyside (MER) and North York Moors (NYM) as mapped by EcoservR.

Level 1	Level 2	Habitat class	Fraction cover (%) MER	Fraction cover (%) NYM
A		Woodland and scrub	7.1	18.1
B		Grassland and marsh	5.3	26.3
D		Heathland	0.1	19.9
E		Mire	<0.1	0.0
G		Water	1.2	0.4
H		Coastal	17.7	0.0
I		Rock, exposure and waste	0.1	0.1
J	J1	Cultivated/disturbed land	30.5	32.4
	J3	Urban/built-up/infrastructure	23.0	2.0
	J5	Gardens/parks/brownfield	14.4	0.6
		Other/uncertain	0.6	<0.1

Note: These proportions are not equal to the weights used for the area-weighted accuracy (Table 3 row totals) as they include manmade features and built-up areas which were not present in the reference data. The habitat classes have been aggregated to illustrate the type of land use summary that can be calculated from EcoservR's outputs.

Validation outcomes were also shaped by the thematic scope of the reference data. Several urban land-cover classes, including roads, buildings and other built-up features, were excluded from Level 2 and Level 3 accuracy assessments because they are not recorded in Phase 1 surveys. These classes represent a substantial proportion of land cover in Merseyside (over 25%; Table 8) and a smaller proportion in the North York Moors. Consequently, reported accuracy metrics may underestimate overall agreement between EcoservR and the contemporary landscape, while avoiding inflation of accuracy through inclusion of classes with limited ecological relevance. This trade-off reflects broader debates in ecosystem-service and habitat mapping regarding whether validation should prioritise ecological relevance or statistical completeness, depending on the intended application of the outputs (Eigenbrod et al., 2010; Schröter et al., 2014). For example, applications oriented towards spatial planning and policy appraisal may place greater emphasis on accurate representation of built and managed land uses, consistent with natural capital approaches used in land-use decision-making and economic appraisal, whereas conservation-focused applications may prioritise discrimination among semi-natural habitats (Bateman et al., 2013; HM Treasury, 2003; Office for National Statistics, 2024).

4.4 | Land-cover patterns and implications for ecosystem services

Land-cover patterns derived from the EcoservR basemaps reflect the contrasting ecological and socio-environmental contexts of the two study areas. Merseyside is dominated by cultivated or disturbed land, coastal habitats and urban features, whereas the North York Moors comprise extensive grasslands, heathlands and mires alongside agricultural land. These broad patterns are consistent with expectations for peri-urban versus upland landscapes and align with national land-cover products such as the UKCEH Land Cover Map (Marston et al., 2023; Rowland et al., 2017). Comparable convergence between rule-based basemaps and national land-cover datasets has been observed in other policy-facing mapping exercises, where consistency at aggregated thematic levels is a key requirement for strategic assessment (Büttner, 2014; Keith et al., 2017).

At this level of aggregation, the EcoservR outputs provide a defensible spatial framework for high-level ecosystem-service assessments. Broad land-cover patterns are sufficient to support indicative analyses of services that are strongly linked to extent and location, such as carbon storage potential, recreation opportunities and flood regulation capacity. However, the habitat-specific classification biases identified earlier introduce uncertainty when outputs are applied to fine-scale habitat assessment or quantitative ecosystem-service estimation. Difficulties in differentiating grasslands, heathlands and bogs are likely to affect metrics that depend on precise habitat characterisation, including carbon accounting based on vegetation structure or habitat-condition indicators and biodiversity-focused conservation and management measures.

Similar limitations have been reported across a range of ecosystem-service mapping approaches, including machine-learning and expert-led systems, where semi-natural habitats with high internal variability are consistently challenging to represent using nationally harmonised datasets (Eigenbrod et al., 2010; Schirpke et al., 2019). This indicates that the uncertainties observed here are not specific to EcoservR, but reflect broader constraints associated with balancing national coverage against ecological detail. As a result, ecosystem-service assessments that rely on fine thematic discrimination are best supported through hybrid approaches that combine national basemaps with targeted local data or field evidence in decision-critical contexts (Kenter et al., 2019; Lavorel et al., 2017).

Despite these constraints, the spatial patterns identified by EcoservR remain informative for regional planning and strategic prioritisation. In Merseyside, the dominance of urban and coastal environments highlights issues related to urban climate resilience, green-infrastructure provision and coastal protection. In the North York Moors, the extent of heathlands and mires points to opportunities for carbon storage, peatland restoration and biodiversity recovery. While finer differentiation of semi-natural habitats would improve confidence for site-level interventions, the basemaps provide a useful evidence base for orienting strategic investment and policy attention at landscape scale.

4.5 | Outlook for UK mapping and policy

This study points to a clear role for automated, rule-based habitat mapping in the UK nature recovery and natural capital policy sector, where analysis is conducted at broad thematic resolution. EcoservR is best suited to landscape and regional scale applications, and could therefore support nature recovery strategies, strategic green-infrastructure planning and indicative ecosystem-service assessment. In these contexts, consistency, transparency and reproducibility are often more critical than fine thematic precision. Rule-based approaches provide an auditable means of integrating heterogeneous national datasets and generating comparable baselines that can support coordination across planning authorities and policy instruments. At the same time, the results indicate that deployment in more detailed regulatory or site-specific contexts would require additional information on habitat condition, management and vegetation structure. Improvements in national datasets, particularly those capturing ecological heterogeneity within grasslands, heaths and bogs, would enhance the policy relevance of such workflows beyond strategic assessment. Although the current implementation relies on datasets specific to England, ongoing development for Scotland and Wales demonstrates how the approach can be adapted to different data environments and governance settings. In Scotland, for example, the framework has been incorporated into NatureScot's Natural Capital Tool to support opportunity mapping and ecosystem service change assessment (scenario mapping) (NatureScot, 2025). More broadly, while the Phase 1 classification system is UK-specific, the underlying logic of transparent, rule-based data harmonisation has wider relevance for national habitat mapping and natural-capital assessment in contexts where comparable spatial data infrastructures exist.

5 | CONCLUSIONS

This study provides an evidence-based evaluation of the EcoservR rule-based workflow for habitat mapping across two contrasting landscapes. The results show that the framework performs most reliably for broad habitat classes, offering a consistent and transparent approach for generating baseline habitat maps suitable for regional planning and high-level natural capital assessments. Classification accuracy declined with increasing thematic detail, particularly for grasslands, heaths and bogs, reflecting both ecological heterogeneity and limitations in the national datasets underpinning the workflow. Given the restricted spatial extent of the validation, these findings should be interpreted as area-specific rather than representative of EcoservR performance nationally. Discrepancies between EcoservR outputs and reference data were not attributable solely to classifier error. Temporal mismatches, variation in survey focus and known uncertainties within Phase 1 habitat data contributed to disagreement, emphasising the need for caution when validating contemporary maps against historical field surveys. Adjudication of mismatched points indicated that EcoservR sometimes captured land-cover changes absent from older survey datasets, highlighting the value of automated workflows that integrate up-to-date national data for monitoring dynamic landscapes at little or no cost. Future development should prioritise improved representation of ecologically variable habitats through integration of finer-resolution spatial data, vegetation structure metrics and habitat-condition indicators. Extending validation to a wider range of landscapes would strengthen understanding of the workflow's general performance and its suitability for supporting statutory habitat reporting and biodiversity assessment. Within these bounds, EcoservR provides a transparent and reproducible framework with clear potential to inform nature-recovery policy, ecosystem-service analysis and strategic land-use planning.

AUTHOR CONTRIBUTIONS

Conceptualisation: Sandra Angers-Blondin, Colm Bowe, Luke Kinross Bentley, Hannah Branwood and Alex Owusu Amoakoh. Data curation and formal analysis: Sandra Angers-Blondin, Luke Kinross Bentley and Alex Owusu Amoakoh. Software and visualisation: Sandra Angers-Blondin, Luke Kinross Bentley, Alex Owusu Amoakoh and Noémie Bonnin. Validation: Sandra Angers-Blondin, Luke Kinross Bentley and Alex Owusu Amoakoh. Investigation (field data collection and pre-processing): Sandra Angers-Blondin, Luke Kinross Bentley, Stamatia Galata and Hannah Branwood. Writing—original draft preparation: Sandra Angers-Blondin, Joe Bellis and Alex Owusu Amoakoh. Writing—review and editing: Andrew Clark, Hanna Partoft, Hannah Branwood, Colm Bowe and Chloe Bellamy. Supervision: Colm Bowe, Chloe Bellamy and Hanna Partoft. All authors reviewed and approved the final manuscript.

ACKNOWLEDGEMENTS

The authors thank Liverpool City Region Combined Authority and North York Moors National Park for providing access to ecological habitat survey data.

FUNDING INFORMATION

No funding information applies.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest related to this study.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/2688-8319.70234.

DATA AVAILABILITY STATEMENT

The validation workflow supporting this study is openly available at https://doi.org/10.24377/LJMU.d.00000256 (Angers-Blondin et al., 2026). Due to third-party licensing restrictions associated with Ordnance Survey MasterMap and related datasets, the EcoservR habitat basemap cannot be shared publicly. However, all scripts used for sampling, adjudication, accuracy assessment and visualisation are publicly available within the archived workflow.

ETHICS STATEMENT

This research did not involve human participants or animals and therefore did not require ethical approval.

ORCID

Sandra Angers-Blondin https://orcid.org/0000-0001-8907-3502
Colm Bowe https://orcid.org/0000-0001-7302-3906
Luke Kinross Bentley https://orcid.org/0009-0004-5468-0563
Alex Owusu Amoakoh https://orcid.org/0000-0001-8394-1241
Joe Bellis https://orcid.org/0000-0003-2787-3736
Lucy Dowdall https://orcid.org/0000-0002-5227-2548
Noémie Bonnin https://orcid.org/0000-0001-5529-7452
Stamatia Galata https://orcid.org/0000-0002-9016-6308
Chloe Bellamy https://orcid.org/0000-0002-3830-0995
Hanna Partoft https://orcid.org/0000-0002-3162-4540

REFERENCES

Amoakoh, A. O., Aplin, P., Awuah, K. T., Delgado-Fernandez, I., Moses, C., Alonso, C. P., Kankam, S., & Mensah, J. C. (2021). Testing the contribution of multi-source remote sensing features for random forest classification of the greater Amanzule tropical peatland. *Sensors*, *21*(10), 3399. https://doi.org/10.3390/s21103399

Amoakoh, A. O., Aplin, P., Rodríguez-Veiga, P., Moses, C., Alonso, C. P., Cortés, J. A., Delgado-Fernandez, I., Kankam, S., Mensah, J. C., & Nortey, D. D. N. (2024). Predictive modelling of land cover changes in the greater Amanzule peatlands using multi-source remote sensing and machine learning techniques. *Remote Sensing*, *16*(21), 4013. https://doi.org/10.3390/rs16214013

Angers-Blondin, S., Bowe, C., Bentley, L., Amoakoh, A. O., Bellis, J., Dowdall, L., Bonnin, N., Galata, S., Branwood, H., Clark, A., Bellamy, C., & Partoft, H. (2026). Validation workflow and derived comparison data for: Evaluating the accuracy of the EcoservR toolkit for fine-resolution habitat mapping. *Liverpool John Moores University Research Data Repository*. https://doi.org/10.24377/LJMU.d.00000256

Angers-Blondin, S., Pimblett, J., Bellamy, C., Rouquette, J., Holt, A., Varley, M., & Bowe, C. (2020). *EcoservR: A natural capital mapping tool for measuring public goods*. ELM Test and Trial 074. Final report presented to Defra, September 2020. https://ecosystemsknowledge.net/resources/tool-assessor/ecoservr/

Bateman, I. J., & Balmford, B. (2018). Public funding for public goods: A post-Brexit perspective on principles for agricultural policy. *Land Use Policy*, *79*, 293–300. https://doi.org/10.1016/j.landusepol.2018.08.022

Bateman, I. J., Harwood, A. R., Mace, G. M., Watson, R. T., Abson, D. J., Andrews, B., Binner, A., Crowe, A., Day, B. H., Dugdale, S., & Fezzi, C. (2013). Bringing ecosystem services into economic decision-making: Land use in the United Kingdom. *Science*, *341*(6141), 45–50. https://doi.org/10.1126/science.1234379

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Büttner, G. (2014). CORINE land cover and land cover change products. In I. Manakos & M. Braun (Eds.), *Land use and land cover mapping in Europe. Remote sensing and digital image processing* (Vol. *18*). Springer. https://doi.org/10.1007/978-94-007-7969-3_5

Carlotto, M. J. (2009). Effect of errors in ground truth on classification accuracy. *International Journal of Remote Sensing*, *30*(18), 4831–4849. https://doi.org/10.1080/01431160802672864

CBD. (2022). *Kunming–Montreal Global Biodiversity Framework*. Secretariat of the Convention on Biological Diversity. https://www.cbd.int/gbf

Cherrill, A. (2016). Inter-observer variation in habitat survey data: Investigating the consequences for professional practice. *Journal of Environmental Planning and Management*, *59*(10), 1813–1832. https://doi.org/10.1080/09640568.2015.1090961

Cherrill, A., & McClean, C. (1999). The reliability of "phase 1" habitat mapping in the UK: The extent and types of observer bias. *Landscape and Urban Planning*, *45*(2–3), 131–143. https://doi.org/10.1016/S0169-2046(99)00027-4

Davidson, N. C., Fluet-Chouinard, E., & Finlayson, C. M. (2018). Global extent and distribution of wetlands: Trends and issues. *Marine and Freshwater Research*, *69*(4), 620–627. https://doi.org/10.1071/MF17019

Department for Environment, Food and Rural Affairs (Defra). (2025a). *Natural Capital and Ecosystem Assessment (NCEA) programme*. Defra. https://www.gov.uk/government/publications/natural-capital-and-ecosystem-assessment-ncea-programme

Department for Environment, Food and Rural Affairs (Defra). (2025b). *Enabling a Natural Capital Approach (ENCA): Guidance*. Defra. https://www.gov.uk/government/publications/enabling-a-natural-capital-approach-enca-guidance

EDINA. (n.d.). *Digimap*. University of Edinburgh. Available at:. https://digimap.edina.ac.uk Accessed: 20 March 2024.

Eigenbrod, F., Armsworth, P. R., Anderson, B. J., Heinemeyer, A., Gillings, S., Roy, D. B., Thomas, C. D., & Gaston, K. J. (2010). The impact of proxy-based methods on mapping the distribution of ecosystem services. *Journal of Applied Ecology*, *47*(2), 377–385. https://doi.org/10.1111/j.1365-2664.2010.01777.x

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*(1), 185–201. https://doi.org/10.1016/S0034-4257(01)00295-4

Foody, G. M. (2010). Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, *114*(10), 2271–2285. https://doi.org/10.1016/j.rse.2010.05.003

Hazeu, G. W. (2014). Operational land cover and land use mapping in the Netherlands. In I. Manakos & M. Braun (Eds.), *Land use and land cover mapping in Europe. Remote sensing and digital image processing* (Vol. *18*). Springer. https://doi.org/10.1007/978-94-007-7969-3_18

Hein, L., Bagstad, K. J., Obst, C., Edens, B., Schenau, S., Castillo, G., Soulard, F., Brown, C., Driver, A., Bordt, M., & Steurer, A. (2020). Progress in natural capital accounting for ecosystems. *Science*, *367*(6477), 514–515. https://doi.org/10.1126/science.aaz8901

Herold, M., Roberts, D. A., Gardner, M. E., & Dennison, P. E. (2004). Spectrometry for urban area remote sensing—Development and analysis of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment*, *91*(3–4), 304–319. https://doi.org/10.1016/j.rse.2004.02.013

BRITISH ECOLOGICAL SOCIETY Ecological Solutions and Evidence

HM Treasury. (2003). *The green book: Appraisal and evaluation in central government*. TSO. https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government

Horning, N., Robinson, J. A., Sterling, E. J., Turner, W., & Spector, S. (2010). *Remote sensing for ecology and conservation: A handbook of techniques*. Oxford University Press.

IPBES. (2019). *Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* (E. S. Brondizio, J. Settele, S. Díaz, & H. T. Ngo, Eds.). IPBES Secretariat. https://doi.org/10.5281/zenodo.3831673

JNCC. (2010). *Handbook for Phase 1 habitat survey—A technique for environmental audit*. https://data.jncc.gov.uk/data/9578d07b-e018-4c66-9c1b-47110f14df2a/Handbook-Phase1-HabitatSurvey-Revised-2016.pdf

Keith, H., Vardon, M., Stein, J. A., Stein, J. L., & Lindenmayer, D. (2017). Ecosystem accounts define explicit and spatial trade-offs for managing natural resources. *Nature Ecology & Evolution*, *1*(11), 1683–1692. https://doi.org/10.1038/s41559-017-0309-1

Kenter, J. O., Raymond, C. M., van Riper, C. J., Azzopardi, E., Brear, M. R., Calcagni, F., Christie, M., Christie, P., Fordham, A., Gould, R. K., Ives, C. D., Hejnowicz, A. P., Gunton, R., Horcea-Milcu, A. I., Kendal, D., Kronenberg, J., Massenberg, J., O'Connor, S., Ravenscroft, N., … Verbruggen, A. (2019). Loving the mess: Navigating diversity and conflict in social values for sustainability. *Sustainability Science*, *14*, 1439–1461. https://doi.org/10.1007/s11625-019-00726-4

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, *28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05

Lavorel, S., Bayer, A., Bondeau, A., Lautenbach, S., Ruiz-Frau, A., Schulp, N., Seppelt, R., Verburg, P., Van Teeffelen, A., Vannier, C., & Arneth, A. (2017). Pathways to bridge the biophysical realism gap in ecosystem services mapping approaches. *Ecological Indicators*, *74*, 241–260. https://doi.org/10.1016/j.ecolind.2016.11.015

Liverpool City Region Combined Authority. (2025). *Local Nature recovery strategy: opportunity mapping method report*. LCRCA. https://api.liverpoolcityregion-ca.gov.uk/wp-content/uploads/Supporting-Document_LCR-LNRS-Opportunity-Mapping-Method-Report.pdf

Marston, C. G., O'Neil, A. W., Morton, R. D., Wood, C. M., & Rowland, C. S. (2023). LCM2021—The UK land cover map 2021. *Earth System Science Data*, *15*, 4631–4649. https://doi.org/10.5194/essd-15-4631-2023

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817.

Merseyside Environmental Advisory Service (MEAS) and Liverpool City Region Combined Authority (LCRCA). (2023). *Pathways to nature recovery: Investing in nature recovery for the Liverpool City Region*. DEFRA Natural Environment Investment Readiness Fund. Available at: https://investinginnaturelcr.com/wp-content/uploads/2025/04/LCR-NIERF-Final-report-for-NEIRF.pdf (Accessed: 23 March 2026).

Mori, A. S., Dee, L. E., Gonzalez, A., Ohashi, H., Cowles, J., Wright, A. J., Loreau, M., Hautier, Y., Newbold, T., Reich, P. B., Matsui, T., Takeuchi, W., Okada, K. i., Seidl, R., & Isbell, F. (2021). Biodiversity–productivity relationships are key to nature-based climate solutions. *Nature Climate Change*, *11*(6), 543–550. https://doi.org/10.1038/s41558-021-01062-1

NatureScot. (2025). *Natural capital tool*. NatureScot. https://www.nature.scot/doc/natural-capital-tool

North York Moors National Park Authority. (2018). *Ryevitalise project area—Feature layer*. ArcGIS. https://www.arcgis.com/home/item.html?id=a9c22e8ce5b54b05ad65d015e58d2e61

Office for National Statistics. (2024). *UK natural capital accounts: 2024*. Newport: ONS. https://www.ons.gov.uk/economy/environmentalaccounts/bulletins/uknaturalcapitalaccounts/2024

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57. https://doi.org/10.1016/j.rse.2014.02.015

Ordnance Survey. (2010). *OS MasterMap Topography Layer: User guide and technical specification (D05300_27, v1.9, December 2010)*. Ordnance Survey Southampton. https://digimap.edina.ac.uk/help/files/resource-hub/downloads/osmanuals/os-mastermap-topography-layer-user-guide_v1_9.pdf

Pebesma, E. J. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, *10*(1). https://digitalcommons.unl.edu/r-journal/626/

Punalekar, S. M., Hurford, C., Lucas, R. M., Planque, C., & Chognard, S. (2024). Hierarchical-modular framework for habitat mapping through systematic and informed integration of remote sensing data with contextual information. *Ecological Informatics*, *82*, 102714. https://doi.org/10.1016/j.ecoinf.2024.102714

Rebelo, L. M., Finlayson, C. M., Strauch, A., Rosenqvist, A., Perennou, C., Tottrup, C., Hilarides, L., Paganini, M., Wielaard, N., Siegert, F., Ballhorn, U., Navratil, P., Franke, J., & Davidson, N. (2018). *The use of Earth Observation for wetland inventory, assessment and monitoring: An information source for the Ramsar Convention on Wetlands. Ramsar technical report no. 10*. Ramsar Convention Secretariat.

Richiardi, C., Blonda, P., Rana, F. M., Santoro, M., Tarantino, C., Vicario, S., & Adamo, M. (2025). Unravelling decades of habitat dynamics in protected areas: Methodological advances and practical tools for large-scale habitat mapping. *Environmental Monitoring and Assessment*, *197*, 1216. https://doi.org/10.1007/s10661-025-14669-0

Rowland, C. S., Morton, R. D., Carrasco, L., McShane, G., O'Neil, A. W., & Wood, C. M. (2017). *Land Cover Map 2015 (vector, GB)*. NERC Environmental Information Data Centre. https://doi.org/10.5285/6c6c9203-7333-4d96-88ab-78925e7a4e73

Ruijs, A., Vardon, M., Bass, S., & Ahlroth, S. (2019). Natural capital accounting for better policy. *Ambio*, *48*(7), 714–725. https://doi.org/10.1007/s13280-018-1107-y

Schirpke, U., Candiago, S., Vigl, L. E., Jäger, H., Labadini, A., Marsoner, T., Meisch, C., Tasser, E., & Tappeiner, U. (2019). Integrating supply, flow and demand to enhance the understanding of interactions among multiple ecosystem services. *Science of the Total Environment*, *651*, 928–941. https://doi.org/10.1016/j.scitotenv.2018.09.235

Schröter, M., Van der Zanden, E. H., van Oudenhoven, A. P., Remme, R. P., Serna-Chavez, H. M., De Groot, R. S., & Opdam, P. (2014). Ecosystem services as a contested concept: A synthesis of critique and counter-arguments. *Conservation Letters*, *7*(6), 514–523. https://doi.org/10.1111/conl.12091

Seddon, N., Daniels, E., Davis, R., Chausson, A., Harris, R., Hou-Jones, X., Huq, S., Kapos, V., Mace, G. M., Rizvi, A. R., Reid, H., Roe, D., Turner, B., & Wicander, S. (2020). Global recognition of the importance of nature-based solutions to the impacts of climate change. *Global Sustainability*, *3*, e15. https://doi.org/10.1017/sus.2020.8

Stehman, S. V. (2009). Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, *30*(20), 5243–5272. https://doi.org/10.1080/01431160903131000

Stehman, S. V., & Foody, G. M. (2008). Accuracy assessment. In *The SAGE handbook of remote sensing* (pp. 297–314). https://doi.org/10.4135/9780857021052.n21

Stehman, S. V., & Foody, G. M. (2019). Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, *231*, 111199. https://doi.org/10.1016/j.rse.2019.05.018

Tuanmu, M. N., & Jetz, W. (2014). A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography*, *23*(9), 1031–1045. https://doi.org/10.1111/geb.12182

UNFCCC. (2016). *The Paris Agreement—Publication*. United Nations Framework Convention on Climate Change Secretariat. https://unfccc.int/sites/default/files/resource/parisagreement_publication.pdf

UNSC. (2021). *System of environmental-economic accounting—743 ecosystem accounting (seea-ea)*. https://seea.un.org/

Wang, Y., Xu, Y., Xu, X., Jiang, X., Mo, Y., Cui, H., Zhu, S., & Wu, H. (2024). Evaluation of six global high-resolution land cover products over China. *International Journal of Digital Earth*, *17*(1). https://doi.org/10.1080/17538947.2023.2301673

Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment*, *117*, 34–49. https://doi.org/10.1016/j.rse.2011.02.030

Winn, J. P., Bellamy, C. C., & Fisher, T. (2018). *EcoServ-GIS: A toolkit for mapping ecosystem services*. Scottish Natural Heritage Research Report No. 954. https://www.nature.scot

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Confusion matrices for habitat classification accuracy.

Table S1. Deterministic rule-based workflow implemented in EcoservR.

How to cite this article: Angers-Blondin, S., Bowe, C., Bentley, L. K., Amoakoh, A. O., Bellis, J., Dowdall, L., Bonnin, N., Galata, S., Branwood, H., Clark, A., Bellamy, C., & Partoft, H. (2026). Evaluating the accuracy of the EcoservR toolkit for fine-resolution habitat mapping. *Ecological Solutions and Evidence*, *7*, e70234. https://doi.org/10.1002/2688-8319.70234