Evaluating modelling decisions and spatial predictions in ecosystem mapping

Authors:
Alys R. Young[1,2], Nicholas J. Murray[3], Jane Elith[2], Brett A. Bryan[1], Hugh F. Davies[4,5], and Emily Nicholson[2]

Affiliations:
[1] School of Life and Environmental Sciences, Faculty of Science, Engineering, and Built Environment, Deakin University, Burwood VIC 3125 Australia

[2] School of Agriculture, Food and Ecosystem Science, Faculty of Science, The University of Melbourne, Parkville VIC 3010 Australia

[3] College of Science and Engineering, James Cook University, Townsville QLD 4811 Australia

[4] Research Institute for the Environment and Livelihoods, Charles Darwin University, Casuarina NT 0909 Australia

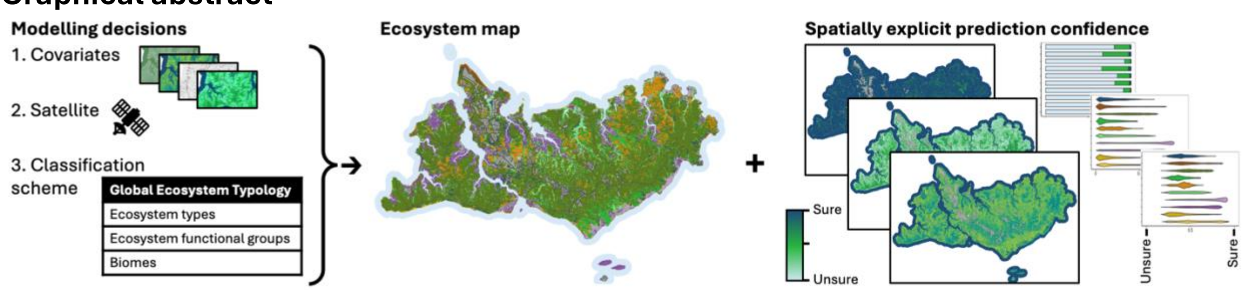
[5] School of Environmental and Rural Science, Faculty of Agriculture, Business and Law, University of New England, Armidale, NSW 2350 Australia

Corresponding author:
Alys Young
a.young@research.deakin.edu.au
alysy.research@gmail.com
ORCID 0000-0002-9562-2253

Abstract

Ecosystem maps support a vast array of applications in conservation, land management and policy. The capacity of an ecosystem map to support these applications is determined by its ability to accurately represent ecosystem distributions, which is heavily influenced by the model used to produce them. Here, we evaluated the influence of key modelling decisions made whilst developing a new and comprehensive ecosystem map using a recently developed ecosystem typology for the remote Tiwi Islands, Australia. We collated a reference set of training points from diverse datasets and employed a pixel-based, random forest model to classify and predict ecosystem distributions. We tested decisions at three stages of the model formulation. First, we tested the number of classes by aggregating ecosystem types (finest scale, n = 11) into functional groups (n = 10) and biomes (coarsest, n = 8) according to the Global Ecosystem Typology. Second, we compared data acquired from the Sentinel-2 satellite using the MSI sensor and Landsat-9 with the OLI-2 sensor. Finally, we tested covariates from satellite image bands only or satellite imagery combined with additional covariates describing other ecological characteristics. We evaluated these decisions using a range of model performance metrics, including overall, by-class and spatially explicit estimates. Our study found that using covariates additional to those from satellite images improved all evaluation metrics for all model decisions. Acquisitions from Landsat-9 tended to improve model performance over Sentinel-2, although the effect was variable. Developing maps at the biome scale (coarsest resolution) slightly improved overall performance but hinders applications that need to differentiate between ecosystem types. Including additional relevant covariates or considering alternative satellites are better options for improving map performance than simplifying the classes. Producing spatially explicit evaluation of ecosystem maps is a rapid and achievable method to communicate limitations and support users to make informed decisions.

Graphical abstract



Key words:

Remote sensing, Earth observation, vegetation mapping, land cover, island ecology, tropical savanna, machine learning, biogeography

1. Introduction

Ecosystem distribution maps form a crucial foundation to understand, monitor, and make decisions about the environment. Applications of ecosystem maps span conservation assessments (Murray *et al.*, 2017; Keith, Ferrer-Paris, *et al.*, 2024), spatial planning (Watson *et al.*, 2023; Keith, Ghoraba, *et al.*, 2024), valuing services (Hein *et al.*, 2020; Xiao *et al.*, 2024) and reporting (Watson *et al.*, 2020; Nicholson *et al.*, 2024). The usefulness of an ecosystem map in these contexts is determined by its ability to accurately model and represent the distributions of ecosystem classes in geographic space.

As ecosystem maps are models of the natural world, decisions made during the modelling process can strongly impact outcomes (Gould *et al.*, 2023). Variations due to modelling decisions, model uncertainty, and errors (henceforth, 'map reliability') propagate through to applications (Burgman, Lindenmayer and Elith, 2005; Jansen *et al.*, 2022), and influence area estimates (Olofsson *et al.*, 2020; Naas *et al.*, 2023), ecosystem accounting (Venter *et al.*, 2024), and assessments (De la Cruz *et al.*, 2017). Therefore, it is important to assess the main decisions influencing reliability and communicate the remaining error and uncertainty to users.

Evaluating modelling decisions is common in other spatial modelling applications, including for landcover which typically focus on structural elements of the landscape, land-use mapping, and species distribution models (Khatami, Mountrakis and Stehman, 2016; Grimmett, Whitsed and Horta, 2020). Fewer studies have examined the impacts of model formulation in ecosystem mapping which presents a unique and challenging case study (Rocchini *et al.*, 2013). Ecosystems are defined by a unique biotic community, the abiotic environment, and driving ecological processes (CBD, 1992). Thus, ecosystem classes can be difficult to visibly distinguish using remotely sensed data. For instance, forest ecosystem types delineated by distinct understories but displaying similar canopy composition and physical structure are indistinguishable with multispectral imagery (Trouvé *et al.*, 2023). Ecosystems also exhibit complex spatiotemporal dynamics because of ecological processes, natural variation, and disturbance (Dryflor *et al.*, 2016; Dorrough *et al.*, 2021; Keith *et al.*, 2022). Finally, the number of ecosystem types are typically higher than in landcover classification. For instance, 98 ecosystem types are described for Italy compared to 66 landcover classes (Capotorti *et al.*, 2023).

Key factors of model formulation known to influence ecosystem maps include the comprehensiveness of the typology (Foody, 2021), the reference data and classification of location into the ecosystem classes (Rocchini *et al.*, 2013; Dorrough *et al.*, 2021; Naas *et al.*, 2023), covariates (Simensen *et al.*, 2020; Trouvé *et al.*, 2023; Naas *et al.*, 2024), and output post-processing (Horvath *et al.*, 2021). Since ecosystem mapping stems from a long history of landcover mapping, the effects of some decisions can be inferred, such as the benefit of covariates, the challenge of many classes, and model types (Yu *et al.*, 2014; Khatami, Mountrakis and Stehman, 2016). Understanding the key factors and the interaction of these factors specific to ecosystems would provide crucial guidance for ecosystem map development, especially important given the given growing focus on developing ecosystem maps at global, national, and regional scales (Galaz García *et al.*, 2023).

In addition to understanding sources of uncertainty in the model formulation, there is a long and growing interest in spatially explicit evaluation metrics to communicate map reliability. Such approaches have emerged in response to current limitations of evaluation assessments (Stehman and Foody, 2019; Foody, 2021) and as a result of modelling advances (Loosvelt *et al.*, 2012; Mitchell, Downie and Diesing, 2018). Spatially explicit evaluation metrics complement confusion matrix-based evaluation by emphasising spatial patterns and facilitate uncertainty propagation into downstream products (Foody, 2002; Jansen *et al.*, 2022). Here, we refer to these spatially explicit evaluation metrics as ‘prediction confidence’ due to their focus on the probability of class membership (McIver and Friedl, 2001; Mitchell, Downie and Diesing, 2018), and acknowledge that high confidence is not synonymous with high accuracy (Stehman and Foody, 2019). Spatially explicit metrics are yet to become standard practice and require further demonstrations in new applications.

In this paper, we sought to evaluate the effects of modelling decisions on ecosystem maps, using the case study of the Tiwi Islands, Australia. On the Indigenous-owned and managed Tiwi Islands, ecosystem maps inform development decisions and management actions (e.g. Richards *et al.*, 2012). We tested the sensitivity of the map reliability to three modelling decisions. Firstly, to represent decisions related to the classification scheme, we used a hierarchical ecosystem typology (Young *et al.*, 2024) that is aligned with the Global Ecosystem Typology (GET, global-ecosystems.org), an internationally accepted classification of ecosystems (UNSD, 2021; Keith *et al.*, 2022). Different levels of a classification hierarchy are ideal for systematically testing the number of classes which change in relation to the thematic resolution (also called thematic scale or class resolution). Secondly, to examine the impact of the choice of satellite, we compared model covariates retrieved from the Landsat-9 satellite with the Operational Land Imager (OLI-2) sensor against the Sentinel-2 satellite with the Multispectral imager (MSI) sensor. The Landsat and Sentinel missions represent two flagship programs providing open-access satellite images (Wulder *et al.*, 2012) and vary in spatial and spectral resolution, length of time series and processing. Thirdly, to assess the implications of model covariates on map reliability, we investigated the use of only satellite image covariates and compared these to models that also include other ecologically meaningful covariates (hereafter named “additional” covariates). Covariates such as elevation and those representing vegetation structure (e.g. canopy height) often improve ecosystems and landcover models (Khatami, Mountrakis and Stehman, 2016; Simensen *et al.*, 2020; Trouvé *et al.*, 2023). We demonstrate three spatially explicit maps of prediction confidences to accompany the ecosystem map which can inform managers of map reliability and improve conservation outcomes.

2. Materials and methods

2.1 Case study location

The Tiwi Islands, including Melville Island (5,788 km^2), Bathurst Island (1,693 km^2) and numerous small islands, are located off the northern coast of the Northern Territory, Australia. The Tiwi Islands are in the Australian "Tiwi-Coburg" bioregion (DCCEEW, 2021) and the global "Arnhem Land tropical savanna" ecoregion (Olson *et al.*, 2001). The lands and waters of the Tiwi Islands are managed by the Indigenous Tiwi peoples. Much of the Islands are remote and challenging to access (Figure 1).

2.2 Classification scheme

To investigate the impact of the number of classes in the classification scheme which change in relation to the thematic resolution, we employed a recent typology of Tiwi Island ecosystem types (Young *et al.*, 2024). This ecosystem typology was developed using the GET and has a known relationship to each GET level. We tested classification schemes for mapping at three levels of the GET hierarchy: the finest thematic resolution level 6 'subglobal ecosystem types' with 11 classes, level 3 'ecosystem functional groups' (EFGs) with 10 classes, and level 2 'biome' as the coarsest resolution with eight classes (Table 1). Here we use the term 'biome' as defined by the GET; biomes represent the subdivision of realms (e.g. fresh water) by similar broad features of ecosystem structure and function (Keith *et al.*, 2022), although recognise other popular definitions (Mucina, 2019).

2.3 Reference points

Reference points (or 'training points') are confirmed occurrences of each ecosystem class in the classification scheme. We employed the reference point collection developed in (Young *et al.*, 2024) but describe the methods in more detail here.

2.3.1 Data collation

We developed reference points from diverse spatial datasets available in a database owned by the Tiwi Land Council, and field visits with Tiwi knowledge authorities (Table 1). The spatial datasets incorporated data collected by numerous academic and industry professionals over 35 years, and included various types of data, such as aerial photographs, high-quality industry maps, and ecological surveys (Figure 1). For the aerial photographs, we labelled each photograph with the ecosystems that were visible and removed uncertain images. Aerial photographs provided essential information in remote areas. We used GPS tracks and PDF maps from Tiwi Plantation Corporation to locate rainforests and removed misclassifications identified in the field notes. These datasets resulted in numerous reference points due to their high spatial accuracy. Consultancy reports and development proposals contained vegetation maps and photographs, and information regarding ecosystem processes (EcOz Environmental Services, 2012; EcOz Environmental Consultants, 2021). Multiple academic datasets were available collected by government and university academics and students. Rainforests were identified using fauna, flora, and threatened species surveys (Russell-Smith, 1991; Menkhorst and Woinarski, 1992; Gambold and Woinarski, 1993; Liddle and Elliott, 2008). *Eucalypt* *savannas* have been surveyed for mammals and threatened fauna (Davies *et al.*, 2018, 2019, 2021; Neave *et al.*, 2024). Vegetation communities of the *treeless plains* ecosystem (Wilson and Fensham, 1994) and *Melaleuca savanna* (Brocklehurst and Lynch, 2001, 2009) have been the focus of previous mapping efforts. However, developing reference points for the *treeless plains* maps was challenged by the low spatial detail in the line drawn maps and land use change since this time. From 2021 to 2023, we undertook on-ground visits with Tiwi knowledge authorities to locations and ecosystem types chosen by the Tiwi knowledge authorities.

2.3.2 Reference point placement

We placed initial reference points on a 30 m x 30 m grid at or near to the locations identified in the collated datasets through visual interpretation with recent Sentinel-2 and Landsat-9 imagery (described in section 2.4) in QGIS (QGIS Development Team, 2018). From the initial reference points, we removed all points closer than 100 m to minimise spatial autocorrelation and inflated evaluation metrics (Stehman, 2009; Stehman and Foody, 2019) using the ‘enmSdmX’ package with R in R-studio (R Core Team, 2018; RStudio Team, 2020; Smith *et al.*, 2023). This process yielded 5,887 reference points for the remainder of the analysis (Table 1, Figure 1). We obtained too few reference points to map *rocky shorelines* as this ecosystem was only identified from visits with Tiwi knowledge authorities. None of the collated datasets distinguished marine and freshwater ecosystems, and hence have been modelled together as *water* in this research. For all software details, see the Supporting Information.

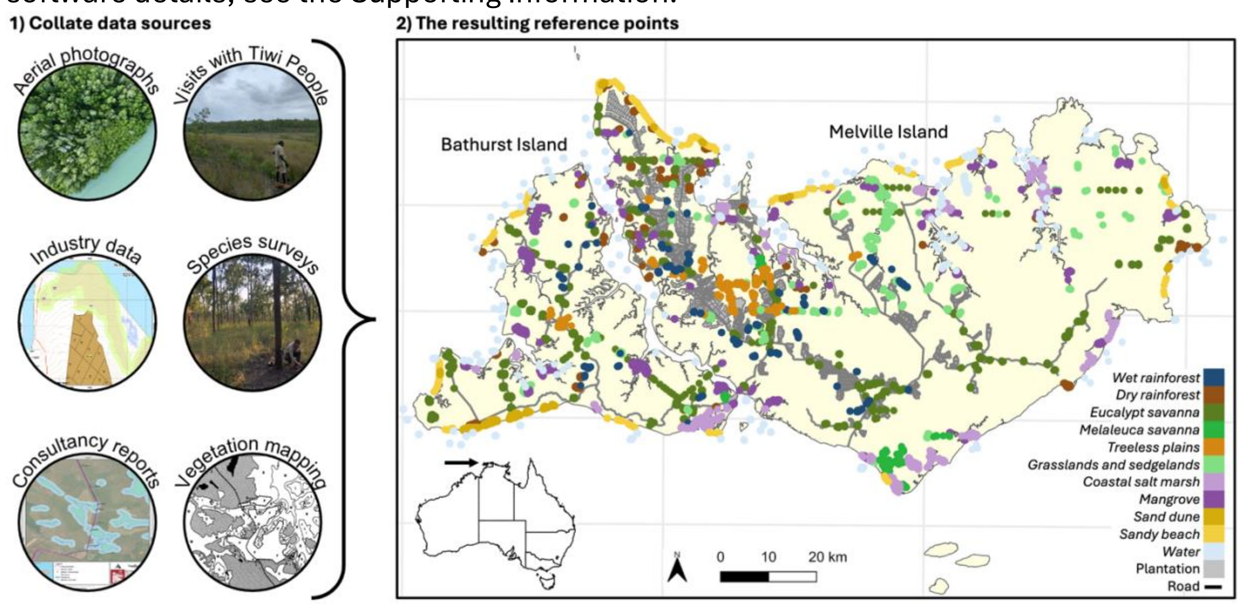


Figure 1. The types of data sources used to develop the reference points and the final reference point locations.

Table 1. Details of how the ecosystem types were grouped into the ecosystem functional group and biome classification schemes according to the Global Ecosystem Typology, and the data sources employed for each class to develop the reference points.

Classification schemes				Data sources for each ecosystem								
Biome (Level 2)	Ecosystem Functional Group (Level 3)	Tiwi Island mapped ecosystem types (Level 6)	Tiwi Island ecosystem typology (Level 6)	Wildlife aerial survey photos	Visits with Tiwi knowledge authorities	Tiwi Plantation Corporation maps	Threatened species monitoring	Tiwi Plantation Corporation surveys	Consultancy repots	Vegetation mapping aerial photos	Targeted vegetation mapping	Shoreline erosion maps
T1 Tropical-subtropical forests biome (n = 1575)	T1.1 Tropical/Subtropical lowland rainforests (n = 433)	Wet rainforest (n = 433)	Wet rainforest	X	X	X	X	X				
	T1.2 Tropical/Subtropical dry forests and thickets (n = 1142)	Dry rainforest (n = 1142)	Dry rainforest	X	X	X			X			
T3 Shrublands and shrubby woodlands (n = 214)	T3.1 Seasonally dry tropical shrublands (n = 214)	Treeless plains (n = 214)	Treeless plains	X	X				X	X	X	
T4 Savannas and grasslands (n = 1023)	T4.2 Pyric tussock savannas (n = 1023)	Eucalypt savanna (n = 927)	Eucalypt open forest savanna	X	X		X			X		
			Eucalypt and mixed-species savanna									
		Melaleuca savanna (n = 96)	*Melaleuca* savanna	X						X	X	
TF1 Palustrine wetlands biome (n = 704)	TF1.4 Seasonal floodplain marshes (n = 704)	Grasslands and sedgelands (n = 704)	Grasslands and sedgelands	X								
MFT1 Brackish tidal (n = 998)	MFT1.2 Intertidal forests and shrublands (n = 698)	Mangroves (n = 698)	Mangroves	X	X					X		
	MFT1.3 Coastal saltmarshes and reedbeds (n = 300)	Coastal saltmarsh (n = 300)	Coastal saltmarsh	X	X							
MT1 Shorelines biome (n = 428)	MT1.3 Sandy shorelines (n = 428)	Shorelines (n = 428)	Sandy beaches	X	X				X			X
			Rocky shorelines		X							
MT2 Supralittoral coastal biome (n = 531)	MT2.1 Coastal shrublands and grasslands (n = 531)	Sand dunes (n = 531)	Sand dunes		X				X			X
Water (n = 414)	Water (n = 414)	Water (n = 414)	Ocean	X	X							X
			Freshwater									

2.4 Satellite image processing

To test the impact of the choice in sensor and satellite image, we retrieved images acquired by the OLI-2 sensor onboard the Landsat-9 satellite (level 2, collection 2, tier 1), courtesy of the United States Geological Survey, and the MSI sensor on the Sentinel-2 satellite from the surface reflectance harmonised collection (level-2A) with atmospheric correction, courtesy of the European Space Agency. In this paper, we refer to these two data sources as 'Landsat-9' and 'Sentinel-2' for succinctness, recognising that each satellite also represents different sensors, wavelengths measured, return times, and other attributes. We obtained and processed the images using Google Earth Engine via the 'rgee' and 'rgeeExtra' packages in R (Gorelick *et al.*, 2017; Aybar *et al.*, 2020). For more details, see the Supporting Information.

Clouds and smoke are common above the Tiwi islands. We tested multiple approaches for developing cloud-free images suitable for modelling. We compiled image sets based on the starting date (January, February, or March) and ending date (April or May) to capture images prior to prescribed burning, from one year (2023), two years (2022 and 2023), or three years (2021 to 2023). During the 2021 to 2023 period when the images were acquired, there were no known changes in the extent of natural ecosystems and targeted investigations supplementary to this research showed only localised changes in mangroves which is not discussed further in this paper. We filtered the image sets by four cloud cover limits (20%, 30%, 40% and 50%), masked the remaining clouds (see Supporting Information for methods) and then reduced the image sets to a single image by the median value of each pixel. We inspected the resultant 120 images for residual clouds. We selected the method that minimised 1) the residual cloud, 2) the number of years, and 3) the cloud cover limit to include the most images.

The final Landsat-9 composite image was developed as the median of pixels from images acquired over January to May in 2023 with less than 30% cloud cover. The final Sentinel-2 image was a three-year composite (2021 to 2023) of images acquired from January to May each year with less than 20% cloud cover.

2.5 Environmental covariates

To develop covariates for testing, we extracted four bands for the red, green, blue, and near-infrared wavelengths from the two satellite images and calculated the normalised difference vegetation index (NDVI). For the additional covariates we obtained soil composition layers from the Soil and Landscape Grid of Australia (Viscarra Rossel *et al.*, 2015) and calculated a mean for each layer in the top 30 cm and 2 m of soil. We obtained elevation data from the Shuttle Radar Topography Mission (SRTM) 5-m Smoothed Digital Elevation Model (DEM-S) (Gallant *et al.*, 2009) and created the Topographic Roughness Index and slope (in degrees) using the 'terra' package (Hijmans, 2023). We also investigated the height above which 50%, 75% and 95% of the vegetation biomass exists (Scarth *et al.*, 2023). Data sources and detailed descriptions are available in the Supporting Information.

To predict the ecosystem distribution across an area with the model, the covariate rasters for each predictor must be available spatially, in the same resolution, and same projection. We resampled the covariates using bilinear interpolation to the resolution of the visible bands of each satellite (~30 m for OLI-2 sensor on the Landsat-9 satellite and

~10 m for MSI sensor on the Sentinel-2 satellite) and the GDA2020 MGA52S coordinate reference system (EPSG: 7852).

Correlations among predictor covariates are known to bias inference and affect parameter estimates (Dormann *et al.*, 2013). We tested collinearity using Pearson's correlation coefficient, retaining covariates with pairwise correlations of less than 0.7 (Supporting Information). For the satellite image covariate set, we retained red, near-infrared, and NDVI. For the satellite image and additional covariate set, we retained red, near-infrared, NDVI, elevation, slope, height of 50% of the vegetation biomass, and the organic carbon, silt and clay in the top 30 cm of soil.

2.6 Model formulation and fitting

We tested 12 model formulations consisting of combinations of three modelling decisions. For the three classification schemes, two satellites, and two covariate sets (total of 12 formulations), we fitted supervised, pixel-based random forest classification models weighting each class by the number of reference points using the 'ranger' package (Wright and Ziegler, 2017). We parameterised the models by testing the number of trees from 10 to 200 in intervals of 10, the number of covariates options to split the nodes from one to five, and a tree depth of the even numbers from two to 10 as well as one. The optimal parameters were 110 trees, two splitting covariates, and six node depth, and we employed these parameter settings across all models for consistency. After parameterisation, we fitted models for the 12 formulations using a cross-validation procedure. We randomly assigned the reference points to one of five partitions, built the cross-validated models on four of the five partitions and tested on the held-out partition for a total of 60 models.

2.7 Model evaluation

From the cross-validated models, we extracted the variable importance by the permutation and summed the predicted classes for the held-out partition to produce a confusion matrix. From the confusion matrices, we calculated the overall evaluation metrics of the accuracy and kappa, and obtained the out-of-bag error from the model output. We report on kappa because it remains prevalent in the literature (Morales-Barquero *et al.*, 2019), despite known problems (Pontius Jr and Millones, 2011; Foody, 2020). We used the by-class evaluation metrics of sensitivity, specificity, precision, F1, and negative predicted value. All evaluation metrics were calculated using the 'caret' package (Kuhn, 2008) using the equations in the Supporting Information. We tested the sensitivity of the overall model evaluation metrics to the cross-validation procedure by running 10,000 models for each formulation on a random 80% of the data and predicting to the remaining 20%.

2.8 Model prediction

To map the spatial distribution of ecosystems, we predicted the probability of each class for every pixel using the cross-validated models. The per-pixel probability is the proportion of random forest trees that assigned the pixel to the class. The class with the highest probability is the final predicted class for that pixel. We identified the predicted class for each model formulation as the mode of the most probable class from the cross-validated models. When multiple classes were predicted in equal amounts, we selected the class with the highest mean probability. We then visualised the predicted class to map ecosystem distribution and overlayed maps of the modified areas.

2.9 Spatially explicit prediction confidence

To communicate the reliability of the mapped spatial distributions, we demonstrate the use of three spatially explicit evaluation metrics. Across the cross-validated models for each pixel, we calculated the mean probability of the highest class (henceforth, maximum probability; McIver and Friedl, 2001; Loosvelt *et al.*, 2012), the mean difference between the highest and second highest probabilities (henceforth, 'Margin of Victory', MoV; McIver and Friedl, 2001) and the number of unique predicted classes (henceforth, prediction stability; (Grimmett, Whitsed and Horta, 2020). Both the maximum probability and the MoV express the strength of the class assignment compared to the other class options. The prediction stability indicates the repeatability within replicates of the same algorithm.

3. Results

We found that choice of covariates most strongly impacted model output. First, using the satellite image and additional covariates together improved the overall evaluation metrics across all model formulations (Figure 2 and Supporting Information). Most classes also improved in by-class metrics (Figure 3) with few exceptions. The most pronounced improvements were in the *treeless plains*, *Melaleuca savanna*, and the *wet* and *dry rainforest* ecosystems (Figure 3 and Supporting Information). Not all additional covariates contributed equally. On these relatively flat islands, elevation proved the most important additional covariate, while the soil covariates and slope added little explanatory power (Supporting Information).

The satellite from which the satellite image was acquired was the second most influential modelling decision. The models that used the Landsat-9 satellite image achieved higher overall accuracy than those models using the Sentinel-2 image (Figure 2, and Supporting Information). The effect of the satellite was most pronounced when only the satellite-image covariates were used. With additional covariates, the Landsat-9 satellite image still improved model performance, although to a lesser degree (Figure 2). Landsat-9 also produced high by-class accuracies; however, the effect varied (Figure 3). For example, the *dry* and *wet rainforests* showed by-class improvements with images acquired from the Sentinel-2 satellite (Figure 3, and Supporting Information).

The classification scheme was the least impactful modelling decision that we tested on the evaluation metrics. The biome classes (the coarsest grouping) slightly improved the overall evaluation metrics, compared to the ecosystem and ecosystem functional groups (Figure 2). This effect was less pronounced with the combined satellite and additional covariates, and for images acquired from the Sentinel-2 satellite (Figure 2, and Supporting Information). In general, the biome classification scheme did not change the by-class evaluation estimates (Figure 3), the exception being the *wet* and *dry rainforest* ecosystem types which were often misclassified in other classification schemes (Supporting Information).

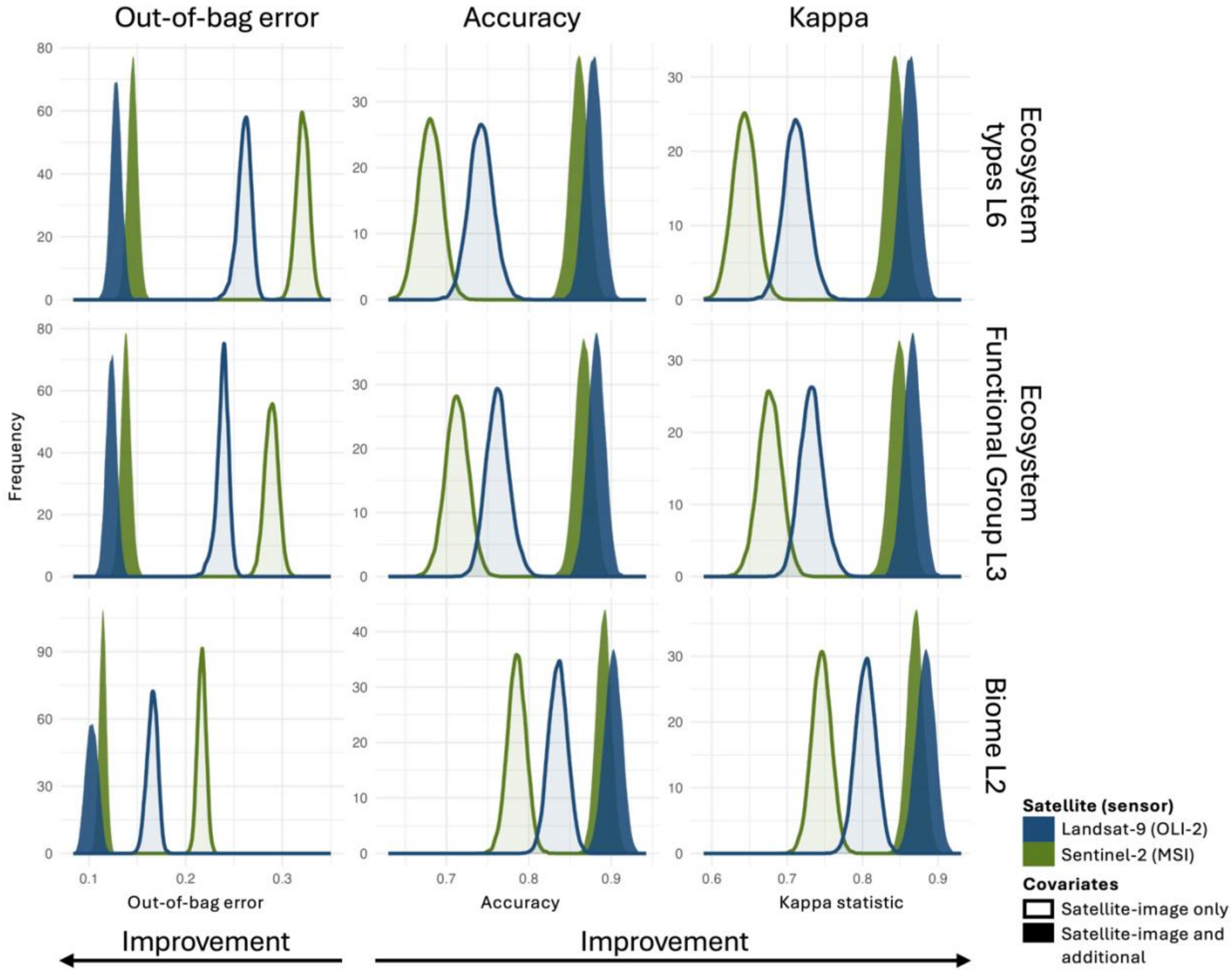


Figure 2. The distribution of the overall evaluation metrics using out-of-bag error (left), accuracy (centre) and kappa statistic (right) from 10,000 random forest models built on 80% of the data, where the model formulations varied by the classification scheme (row), covariates (fill) and satellite (colour).

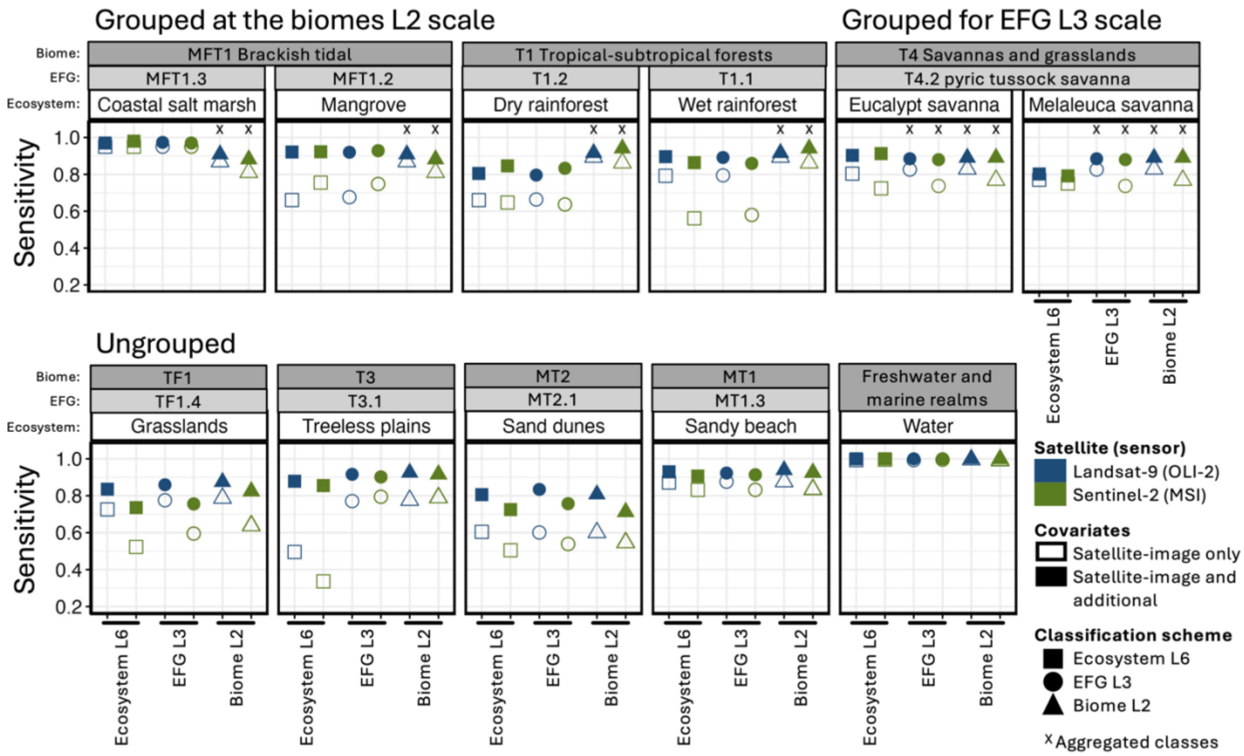


Figure 3. The by-class sensitivity as an exemplar evaluation metric for each class (panels) by the classification schemes (shape), satellite (colour), and covariates (fill). Ecosystem types that were aggregated into ecosystem functional groups (EFGs, data: circle, label: light grey box) and biomes (data: triangle, label: dark grey box above) are identified by an x. Sensitivity is the ability of the model to correctly predict the true class from all those known to be true in the reference points.

The maximum probability and MoV maps imply similar patterns of prediction confidence (Figure 4.B1 and C1). Areas with high confidence occur in a central band and eastern patch on Melville Island, and in isolated areas of Bathurst Island. Low confidence areas, including low stability in the prediction (Figure 4.D1), are scattered across the landscape with an aggregation on the southern coast and far east area of Melville Island. Summarising the prediction confidence across the entire area (Figure 4.B2-D2), the *coastal salt marsh* (light purple) and *mangrove* (dark purple) were predicted with highest confidence (median maximum probability = 75.72% and 64.77%, respectively, and median Mov = 66.15% and 50.46%), indicated by the distribution of the maximum probability and MoV skewed to the right (Figure 4.B2-C2). *Mangroves* were also the most stable ecosystem type with 94.86% of the cells mapped as mangroves only ever predicted to be mangroves, followed by *eucalypt savanna* at 92.57% (Figure 4.D2, light blue boxes). *Sand dunes* were predicted with the low maximum probability values (median of 34.82%) indicated by the distribution skewed to the left (Figure 4.B2, dark yellow), while the MoV distribution is low (median of 11.89%) but comparable to other classes (Figure 4.C2). *Sand dunes* and *sandy beaches* produced unstable predictions with the highest proportion of cells predicted as three different classes (4.65% and 4.03%, Figure 4.D2 dark blue bar), followed by *melaleuca savanna* and *dry rainforests* with the highest proportion of cells with two classes (31.35% and 30.46%, Figure 4.D2 green bar).

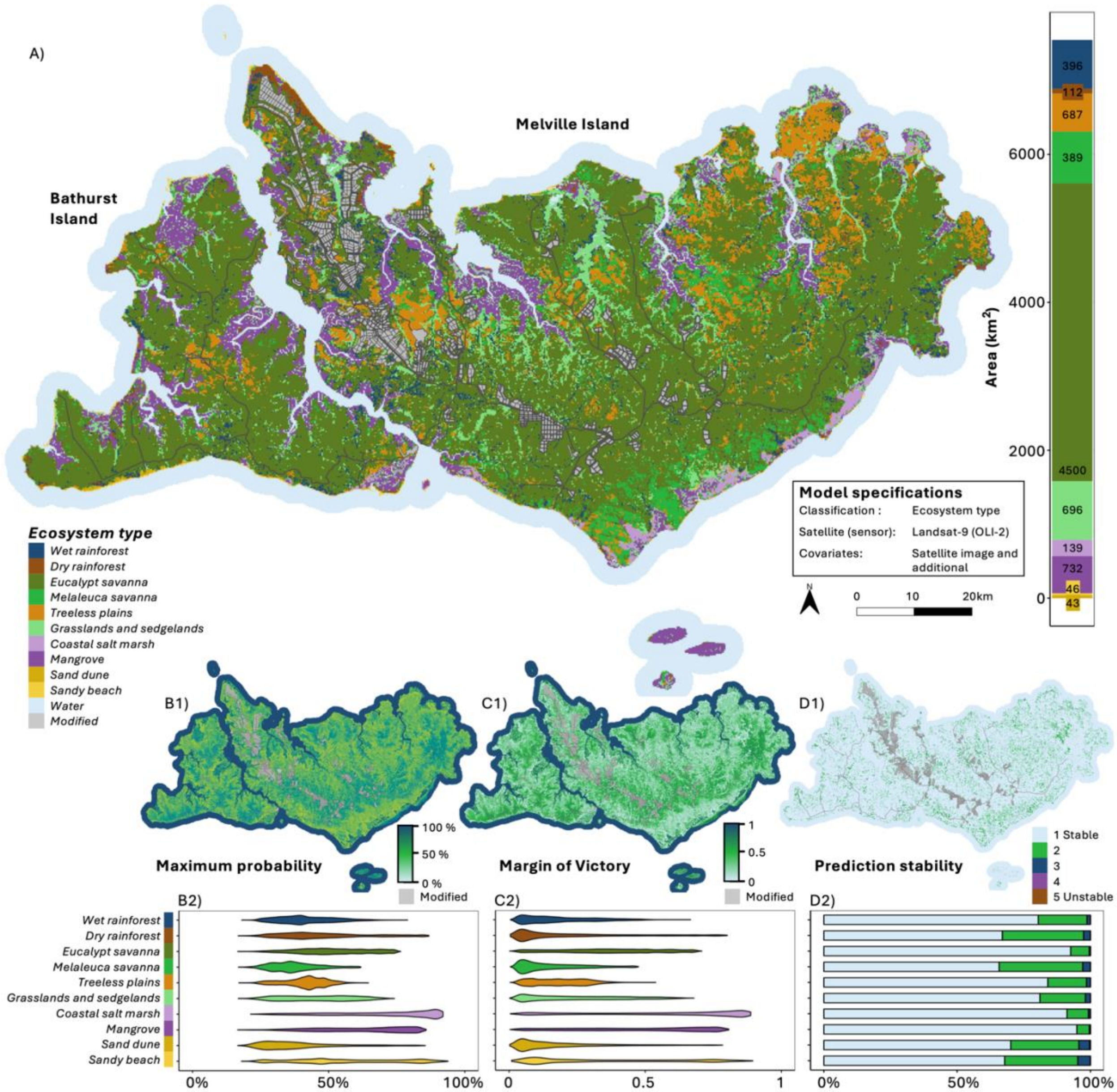


Figure 4. The predicted ecosystem map (A) and spatially explicit evaluation metrics (B-D) for an example model using the ecosystem type as the classification scheme, imagery from the Landsat-9 satellite, and additional covariate alongside those from the satellite image.

4. Discussion

We found that decisions made during the ecosystem mapping procedure strongly impacted model outputs (Figure 2 and Figure 3), consistent with previous studies (Simensen *et al.*, 2020; Trouvé *et al.*, 2023; Naas *et al.*, 2024). The combination of satellite image and additional covariates greatly improved the model performance, supporting previous calls to ensure that ecosystem models capture key attributes of ecosystems and are developed with ecosystem scientists (Xiao *et al.*, 2024). Since we found that the choice of satellite and classification scheme were less influential, additional elements can be considered to guide the decision; Landsat has the advantage of a longer archive, allowing mapping of change through time, while mapping finer ecosystem units has benefits for management of biodiversity.

Ecological theory posits that the distribution of biodiversity is shaped by environmental gradients. Our results showed that the best predictions came from a model including both satellite and additional covariates, aligning with previous research (Simensen *et al.*, 2020; Trouvé *et al.*, 2023; Naas *et al.*, 2024). The elevation covariate added the most explanatory power (Supporting Information), potentially as a proxy for other ecological gradients and processes (Whittaker, 1956). Topographic covariates representing water availability are often valuable to distinguish wet and dry forest types, including rainforests and riparian forests (Trouvé *et al.*, 2023). While the additional covariates contributed useful information, we found that the satellite covariates were still highly informative (Supporting Information), congruent with other studies that suggest ecological or climate covariates are best used alongside covariates from other sources, particularly satellite imagery (Simensen *et al.*, 2020; Trouvé *et al.*, 2023; Naas *et al.*, 2024). Soil covariates were the least informative in this study (Supporting Information), potentially due to underlying data inaccuracies in the available dataset, as noted in other global and national soil maps (Rossiter *et al.*, 2022; Maynard *et al.*, 2023), rather than a lack of ecological importance (Simensen *et al.*, 2020; Keith *et al.*, 2022). Improving the availability, accessibility, and spatiotemporal resolution of ecological covariates would improve both the map reliability and our understanding of the environmental gradients defining their extent. Once such covariates are available, deep learning models, reproducible workflows, and infrastructure are critical to interrogate such large datasets and offer novel insights (Galaz García *et al.*, 2023; Pettorelli *et al.*, 2024).

In addition to the model covariates, the choice of the satellite also influenced model performance. We found that the Landsat-9 satellite imagery generally provided higher overall (Figure 2) and by-class evaluation metrics (Figure 3), although the effect lessens with the inclusion of additional covariates. Exceptions to this were the *wet* and *dry rainforest* ecosystem types, where we detected improvements with the Sentinel-2 satellite image (Figure 3). The Sentinel-2 satellite imagery with the MSI sensor capture finer spatial resolution imagery and may better detect the sharp boundaries that delineate rainforests, reducing the number of pixels containing multiple ecosystem types (i.e. mixed pixels). Mixed pixels are a high source of uncertainty in landcover mapping (Loosvelt *et al.*, 2012) and hamper the reliability of global and national maps (Herold *et al.*, 2008; Congalton *et al.*, 2014).

Alternatively, high spatial resolution sensors may detect structural variability within ecosystem classes, leading to high intra-class variability and noise (Nagendra and

Rocchini, 2008). For instance, savanna ecosystems display highly variable tree occurrence and canopy cover (Keith *et al.*, 2022). In general, satellite spatial resolution has no consistent effect on map reliability (Yu *et al.*, 2014; Morales-Barquero *et al.*, 2019). This suggests that management objectives and ecosystem characteristics should determine the satellite and sensor used (Horvath *et al.*, 2021; Venter *et al.*, 2022; Naas *et al.*, 2024). The additional benefit of the Landsat satellites is the rich archive of images (Wulder *et al.*, 2012) and hence the potential to detect historical changes (Murray *et al.*, 2019; Calderón-Loor, Hadjikakou and Bryan, 2021).

While here we have described the potential effect that spatial resolution may have on ecosystem mapping, we cannot disentangle this effect from the other differences between the Landsat-9 and Sentinel-2 missions. Satellite missions vary in many attributes, including the return time influencing the number of images captured, spectral resolution such as acquiring hyperspectral imagery, and the presence of other instruments with unique data captured such as synthetic aperture radar (Pettorelli *et al.*, 2014). These differences are particularly important in tropical regions where obtaining an image with limited cloud and smoke cover is challenging, as experienced in this research. As we compiled cloud-free composite images from images taken over a period of time, there is a risk of intra- and inter-annual change. In locations with high rates of landscape change, composite images require careful use and would reduce the map reliability.

The least influential modelling decision was the classification scheme, where the overall evaluation metrics were slightly improved with the GET level 2 'biome' scale representing the fewest classes and coarsest scale of biodiversity. Aggregating classes is a common method to improve evaluation metrics (Congalton and Green, 1993; Remmel, 2009) but overall, the benefits are small and variable (Yu *et al.*, 2014). Importantly, modelling biomes presents a direct trade-off with usefulness for future applications where the finer scale classification of ecosystems types is fundamental to management, such as with ecosystem accounting and ecological risk assessments (Hein *et al.*, 2020; Keith, Ferrer-Paris, *et al.*, 2024). The improvements we observed were driven by aggregating specific classes that were often misclassified, namely the *wet* and *dry rainforest*. Such rainforest ecosystems are both represented by the GET level 2 'tropical and subtropical forests biome' but globally these ecosystems differ in threat status (Etter *et al.*, 2017; Murray *et al.*, 2020; Noh *et al.*, 2020) and protection (Wohlfart, Wegmann and Leimgruber, 2014; Rivas, Guerrero-Casado and Navarro-Cerillo, 2021). Aggregating and mapping these ecosystems at the biome scale obscures the urgency and practicality of protecting and managing the world's tropical forests.

Thoughtful model formulation can reduce but never remove error and uncertainty in the model outputs (Rocchini *et al.*, 2013; Foody, 2021). As demonstrated here, spatially explicit prediction confidences are immediate tools that can be readily implemented to communicate spatial patterns of reliability in the maps. Our analysis produced generally low confidence metrics with broadly consistent spatial patterns across the metrics (Figure 4). There are multiple reasons which may lead to the lower confidence predictions (Elith, Burgman and Regan, 2002; Regan, Colyvan and Burgman, 2002). The model may poorly define and predict classes due to a lack of relevant covariates or measurement errors in these covariate layers (Elith, Burgman and Regan, 2002; Barry and Elith, 2006), such as the global soil maps described earlier (Rossiter *et al.*, 2022; Maynard *et al.*, 2023). Natural variation within heterogeneous classes may drive the lower confidence predictions for the *sand dunes* and *melaleuca savanna* which display high variation in grass and tree cover (Young *et al.*, 2024). Ecotones, mixed pixels, or too few reference points often produce poor accuracy (Loosvelt *et al.*, 2012; Rocchini *et al.*, 2013; Foody, 2022). The exact metrics of prediction confidence depend on the model type and warrants research for emerging machine learning models (Pettorelli *et al.*, 2024).

Conclusion

Ecosystem maps tend to be presented without a discussion of the decisions made during the modelling processes nor an evaluation of the implications of these decisions. As new avenues in broad-scale monitoring and change detection of ecosystems arise (Galaz García *et al.*, 2023; Pettorelli *et al.*, 2024), the need to carefully examine the impact of modelling decisions grows. Given the influence of modelling decisions that we identified, both modellers and users must continue to be aware of the role model formulation plays in ecosystem mapping and endeavour to account for map reliability in future applications. Incorporating uncertainty into decision-making is paramount, albeit not always straightforward (Burgman, Lindenmayer and Elith, 2005). The responsibility lies on both the producer of any map to communicate reliability in ways transferable to future applications, and on the user to propagate known uncertainties.

Acknowledgements

We acknowledge the extensive work undertaken by previous researchers, consultants, and industry professionals in developing their respective datasets and products which allowed the creation of the reference points for this research. The Tiwi Islands are the lands of the Tiwi people from eight land-owning clans: Jikilaruwu, Malawu, Mantiyupwi, Marrikawuyanga, Munupi, Wulirankuwu, Wurankuwu, and Yimpinari. We would like to thank Mavis Kerinaiua, Colin Kerinaiua, Simon Munkara, Bernard Tipoloura, Gemma Munkara, John Louis Munkara, Kinjia Munkara-Murray and Marie Munkara for visiting locations on the Tiwi Islands. We thank Mavis Kerinariua, Alana Brekelmans, Margaret Ayre, Michaela Spencer, the Tiwi Land Council, Tiwi Resources, and Tiwi rangers for their involvement and facilitation during on-ground visits. Analysis and writing were undertaken on the lands of the Wurundjeri Woi Wurung peoples of the Kulin Nation.

Conflict of interest statement:

The authors disclose no conflict of interest.

Ethics statement:

Human ethics was approved by The University of Melbourne Human Ethics (#1955248) and Deakin University Human Research Ethics Committee (#2022097). Permission to enter the

Tiwi Islands, to view the satellite imagery, and access the datasets for the training points was granted by the Tiwi Land Council.

Funding statement:

This research was funded by the Australian Research Council (ARC) linkage grant (LP170100305) in partnership with the Tiwi Land Council. The field visits with Tiwi knowledge authorities was funded by the Foundation for National Parks and Wildlife community conservation grant (FNPW028CCG22).

References

Aybar, C., Wu, Q., Bautista, L., Yali, R. and Barja, A. (2020) ‘rgee: An R package for interacting with Google Earth Engine’, *Journal of Open Source Software*, 5(51), p. 2272. Available at: https://doi.org/10.21105/joss.02272.

Barry, S. and Elith, J. (2006) ‘Error and uncertainty in habitat models’, *Journal of Applied Ecology*, 43(3), pp. 413–423. Available at: https://doi.org/10.1111/j.1365-2664.2006.01136.x.

Brocklehurst, P. and Lynch, B. (2009) *Northern Territory Melaleuca forest survey.* Technical Report 25/2009D. Palmerston, Northern Territory: Department of Natural Resources, Environment, The Arts and Sport.

Brocklehurst, P. and Lynch, D. (2001) *Melaleuca Survey of the Top End, Northern Territory.* Technical Report 25/2009D. Palmerston, Northern Territory: Department of Natural Resources, Environment, The Arts and Sport, p. 89.

Burgman, M.A., Lindenmayer, D.B. and Elith, J. (2005) ‘Managing Landscapes for Conservation Under Uncertainty’, *Ecology*, 86(8), pp. 2007–2017. Available at: https://doi.org/10.1890/04-0906.

Calderón-Loor, M., Hadjikakou, M. and Bryan, B.A. (2021) ‘High-resolution wall-to-wall land-cover mapping and land change assessment for Australia from 1985 to 2015’, *Remote Sensing of Environment*, 252, p. 112148. Available at: https://doi.org/10.1016/j.rse.2020.112148.

Capotorti, G., Del Vico, E., Copiz, R., Facioni, L., Zavattero, L., Bonacquisti, S., Paolanti, M. and Blasi, C. (2023) ‘Ecosystems of Italy. Updated mapping and typology for the implementation of national and international biodiversity-related policies’, *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*, 157(6), pp. 1248–1258. Available at: https://doi.org/10.1080/11263504.2023.2284135.

CBD (1992) *Convention on biological diversity*. Rio de Janeiro: United Nations. Available at: https://treaties.un.org/doc/Treaties/1992/06/19920605%2008-44%20PM/Ch_XXVII_08p.pdf.

Congalton, R.G. and Green, K. (1993) ‘A practical look at the sources of confusion in error matrix generation.’, *Photogrammetric engineering and remote sensing*, 59(5), pp. 641–644.

Congalton, R.G., Gu, J., Yadav, K., Thenkabail, P. and Ozdogan, M. (2014) ‘Global Land Cover Mapping: A Review and Uncertainty Analysis’, *Remote Sensing*, 6(12), pp. 12070–12093. Available at: https://doi.org/10.3390/rs61212070.

Davies, H.F., McCarthy, M.A., Firth, R.S.C., Woinarski, J.C.Z., Gillespie, G.R., Andersen, A.N., Rioli, W., Puruntatameri, J., Roberts, W., Kerinaiua, C., Kerinauia, V., Womatakimi, K.B. and Murphy, B.P. (2018) ‘Declining populations in one of the last refuges for threatened mammal species in northern Australia’, *Austral Ecology*, 43(5), pp. 602–612. Available at: https://doi.org/10.1111/aec.12596.

Davies, H.F., Rangers, T.L., Rees, M.W., Stokeld, D., Miller, A.C., Gillespie, G.R. and Murphy, B.P. (2021) ‘Variation in feral cat density between two large adjacent islands in Australia’s monsoon tropics’, *Pacific Conservation Biology*, 28(1), pp. 18–24. Available at: https://doi.org/10.1071/PC20088.

Davies, H.F., Rioli, W., Puruntatameri, J., Roberts, W., Kerinaiua, C., Kerinauia, V., Womatakimi, K.B., Gillespie, G.R. and Murphy, B.P. (2019) ‘Estimating site occupancy and detectability of the threatened partridge pigeon (*Geophaps smithii*) using camera traps’, *Austral Ecology*, 44(5), pp. 868–879. Available at: https://doi.org/10.1111/aec.12755.

DCCEEW (2021) ‘Interim Biogeographic Regionalisation for Australia (IBRA)’.

De la Cruz, M., Quintana-Ascencio, P.F., Cayuela, L., Espinosa, C.I. and Escudero, A. (2017) ‘Comment on “The extent of forest in dryland biomes”’, *Science* [Preprint]. Available at: https://doi.org/10.1126/science.aao0369.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. and Lautenbach, S. (2013) ‘Collinearity: a review of methods to deal with it and a simulation study evaluating their performance’, *Ecography*, 36(1), pp. 27–46. Available at: https://doi.org/10.1111/j.1600-0587.2012.07348.x.

Dorrough, J., Tozer, M., Armstrong, R., Summerell, G. and Scott, M.L. (2021) ‘Quantifying uncertainty in the identification of endangered ecological communities’, *Conservation Science and Practice*, 3(11), p. e537. Available at: https://doi.org/10.1111/csp2.537.

Dryflor, Banda-R, K., Delgado-Salinas, A., Dexter, K.G., Linares-Palomino, R., Oliveira-Filho, A., Prado, D., Pullan, M., Quintana, C., Riina, R., Rodríguez M., G.M., Weintritt, J., Acevedo-Rodríguez, P., Adarve, J., Álvarez, E., Aranguren B., A., Arteaga, J.C., Aymard, G., Castaño, A., Ceballos-Mago, N., Cogollo, Á., Cuadros, H., Delgado, F., Devia, W., Dueñas, H., Fajardo, L., Fernández, Á., Fernández, M.Á., Franklin, J., Freid, E.H., Galetti, L.A., Gonto, R., González-M., R., Graveson, R., Helmer, E.H., Idárraga, Á., López, R., Marcano-Vega, H., Martínez, O.G., Maturo, H.M., McDonald, M., McLaren, K., Melo, O., Mijares, F., Mogni, V., Molina, D., Moreno, N.D.P., Nassar, J.M., Neves, D.M., Oakley, L.J., Oatham, M., Olvera-Luna, A.R., Pezzini, F.F., Dominguez, O.J.R., Ríos, M.E., Rivera, O., Rodríguez, N., Rojas, A., Särkinen, T., Sánchez, R., Smith, M., Vargas, C., Villanueva, B. and Pennington, R.T. (2016) 'Plant diversity patterns in neotropical dry forests and their conservation implications', *Science*, 353(6306), pp. 1383–1387. Available at: https://doi.org/10.1126/science.aaf5080.

EcOz Environmental Consultants (2021) *Terrestrial Ecology Report - Tiwi Islands H2 Project. Provaris Energy,*. Report for Provaris Energy. Darwin, Northern Territory.

EcOz Environmental Services (2012) *Kilimiraka Notice of Intent: Kilimiraka Mineral Sands Project, Bathurst Island, N.T.* Report for Matilda Zircon.

Elith, J., Burgman, M.A. and Regan, H.M. (2002) 'Mapping epistemic uncertainties and vague concepts in predictions of species distribution', *Ecological Modelling*, 157(2), pp. 313–329. Available at: https://doi.org/10.1016/S0304-3800(02)00202-8.

Etter, A., Andrade, Á., Saavedra, K., Amaya, P. and Arévalo, P. (2017) *Estado de los Ecosistemas Colombianos: una aplicación de la metodología de la Lista Roja de Ecosistemas (Vers2.0)*. Bogota, Colombia: Pontificia Universidad Javeriana y Conservación InternacionalColombia, p. 138. Available at: https://www.conservation.org.co/media/A7.LRE-Colombia_INFORME%20FINAL_%202017.pdf.

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

Foody, G.M. (2020) 'Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification', *Remote Sensing of Environment*, 239, p. 111630. Available at: https://doi.org/10.1016/j.rse.2019.111630.

Foody, G.M. (2021) 'Impacts of ignorance on the accuracy of image classification and thematic mapping', *Remote Sensing of Environment*, 259, p. 112367. Available at: https://doi.org/10.1016/j.rse.2021.112367.

Foody, G.M. (2022) 'Global and Local Assessment of Image Classification Quality on an Overall and Per-Class Basis without Ground Reference Data', *Remote Sensing*, 14(21), p. 5380. Available at: https://doi.org/10.3390/rs14215380.

Galaz García, C., Bagstad, K.J., Brun, J., Chaplin-Kramer, R., Dhu, T., Murray, N.J., Nolan, C.J., Ricketts, T.H., Sosik, H.M., Sousa, D., Willard, G. and Halpern, B.S. (2023) 'The future of ecosystem assessments is automation, collaboration, and artificial intelligence', *Environmental Research Letters*, 18(1), p. 011003. Available at: https://doi.org/10.1088/1748-9326/acab19.

Gallant, J., Wilson, N., Tickle, P.K., Downling, T. and Read, A. (2009) '3 second SRTM Derived Digital Elevation Model (DEM) Version 1.0.' Canberra: Geoscience Australia. Available at: http://pid.geoscience.gov.au/dataset/ga/69888.

Gambold, N. and Woinarski, J.C.Z. (1993) 'Distributional patterns of herpetofauna in monsoon rainforests of the Northern Territory, Australia', *Australian Journal of Ecology*, 18(4), pp. 431–449. Available at: https://doi.org/10.1111/j.1442-9993.1993.tb00470.x.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. (2017) 'Google Earth Engine: Planetary-scale geospatial analysis for everyone', *Remote Sensing of Environment*, 202, pp. 18–27. Available at: https://doi.org/10.1016/j.rse.2017.06.031.

Gould, E., Fraser, H.S., Parker, T.H., Nakagawa, S., Griffith, S.C., Vesk, P.A., Fidler, F., Hamilton, D.G., Abbey-Lee, R.N., Abbott, J.K., Aguirre, L.A., Alcaraz, C., Aloni, I., Altschul, D., Arekar, K., Atkins, J.W., Atkinson, J., Baker, C., Barrett, M., Bell, K., Bello, S.K., Beltrán, I., Berauer, B.J., Bertram, M.G., Billman, P.D., Blake, C.K., Blake, S., Bliard, L., Bonisoli-Alquati, A., Bonnet, T., Bordes, C.N.M., Bose, A.P.H., Botterill-James, T., Boyd, M.A., Boyle, S.A., Bradfer-Lawrence, T., Bradham, J., Brand, J.A., Brengdahl, M.I., Bulla, M., Bussière, L., Camerlenghi, E., Campbell, S.E., Campos, L.L.F., Caravaggi, A., Cardoso, P., Carroll, C.J.W., Catanach, T.A., Chen, X., Chik, H.Y.J., Choy, E.S., Christie, A.P., Chuang, A., Chunco, A.J., Clark, B.L., Contina, A., Covernton, G.A., Cox, M.P., Cressman, K.A., Crotti, M., Crouch, C.D., D'Amelio, P.B., Sousa, A.A. de, Döbert, T.F., Dobler, R., Dobson, A.J., Doherty, T.S., Drobniak, S.M., Duffy, A.G., Duncan, A.B., Dunn, R.P., Dunning, J., Dutta, T., Eberhart-Hertel, L., Elmore, J.A., Elsherif, M.M., English, H.M., Ensminger, D.C., Ernst, U.R., Ferguson, S.M., Fernández-Juricic, E., Ferreira-Arruda, T., Fieberg, J., Finch, E.A., Fiorenza, E.A., Fisher, D.N., Fontaine, A., Forstmeier, W., Fourcade, Y., Frank, G.S., Freund, C.A., Fuentes-Lillo, E., Gandy, S.L., Gannon, D.G., García-Cervigón, A.I., Garretson, A.C., Ge, X., Geary, W.L., Géron, C., Gilles, M., Girndt, A., Gliksman, D., Goldspiel, H.B., Gomes, D.G.E., Good, M.K., Goslee, S.C., Gosnell, J.S., Grames, E.M., Gratton, P., Grebe, N.M., Greenler, S.M., Griffioen, M., Griffith, D.M., Griffith, F.J., Grossman, J.J., Güncan, A., Haesen, S., Hagan, J.G., Hager, H.A., Harris, J.P., Harrison, N.D., Hasnain, S.S., Havird, J.C., Heaton, A.J., Herrera-Chaustre, M.L., Howard, T.J., Hsu, B.-Y., Iannarilli, F., Iranzo, E.C., Iverson, E.N.K., Jimoh, S.O., Johnson, D.H., Johnsson, M., Jorna, J., Jucker, T., Jung, M., Kačergytė, I., Kaltz, O., Ke, A., Kelly, C.D., Keogan, K., Keppeler, F.W., Killion, A.K., Kim, D., Kochan, D.P., Korsten, P., Kothari, S., Kuppler, J., Kusch, J.M., Lagisz, M., Lalla, K.M., Larkin, D.J., Larson, C.L., Lauck, K.S., Lauterbur, M.E., Law, A., Léandri-Breton, D.-J., Lembrechts, J.J., L'Herpiniere, K., Lievens, E.J.P., Lima, D.O. de, Lindsay, S., Luquet, M., MacLeod, R., Macphie, K.H., Magellan, K., Mair, M.M., Malm, L.E., Mammola, S., Mandeville, C.P., Manhart, M., Manrique-Garzon, L.M., Mäntylä, E., Marchand, P., Marshall, B.M., Martin, C.A., Martin, D.A., Martin, J.M., Martinig, A.R., McCallum, E.S., McCauley, M., McNew, S.M., Meiners, S.J., Merkling, T., Michelangeli, M., Moiron, M., Moreira, B., Mortensen, J., Mos, B., Muraina, T.O., Murphy, P.W., Nelli, L., Niemelä, P., Nightingale, J., Nilsonne, G., Nolazco, S., Nooten, S.S., Novotny, J.L., Olin, A.B., Organ, C.L., Ostevik, K.L., Palacio, F.X., Paquet, M., Parker, D.J., Pascall, D.J., Pasquarella, V.J., Paterson, J.H., Payo-Payo, A., Pedersen, K.M., Perez, G., Perry, K.I., Pottier, P., Proulx, M.J., Proulx, R., Pruett, J.L., Ramananjato, V., Randimbiarison, F.T., Razafindratsima, O.H., Rennison, D.J., Riva, F., Riyahi, S., Roast, M.J., Rocha, F.P., Roche, D.G., Román-Palacios, C., Rosenberg, M.S., Ross, J., Rowland, F.E., Rugemalila, D., Russell, A.L., Ruuskanen, S., Saccone, P., Sadeh, A., Salazar, S.M., Sales, K., Salmón, P., Sánchez-Tójar, A., Santos, L.P., Santostefano, F., Schilling, H.T., Schmidt, M., Schmoll, T., Schneider, A.C., Schrock, A.E., Schroeder, J., Schtickzelle, N., Schultz, N.L., Scott, D.A., Scroggie, M.P., Shapiro, J.T., Sharma, N., Shearer, C.L., Simón, D., Sitvarin, M.I., Skupien, F.L., Slinn, H.L., Smith, G.P., Smith, J.A., Sollmann, R., Whitney, K.S., Still, S.M., Stuber, E.F., Sutton, G.F., Swallow, B., Taff, C.C., Takola, E., Tanentzap, A.J., Tarjuelo, R., Telford, R.J., Thawley, C.J., Thierry, H., Thomson, J., Tidau, S., Tompkins, E.M., Tortorelli, C.M., Trlica, A., Turnell, B.R., Urban, L., Vondel, S.V. de, Wal, J.E.M. van der, Eeckhoven, J.V., Oordt, F. van, Vanderwel, K.M., Vanderwel, M.C., Vanderwolf, K.J., Vélez, J., Vergara-Florez, D.C., Verrelli, B.C., Vieira, M.V., Villamil, N., Vitali, V., Vollering, J., Walker, J., Walker, X.J., Walter, J.A., Waryszak, P., Weaver, R.J., Wedegärtner, R.E.M., Weller, D.L., Whelan, S., White, R.L., Wolfson, D.W., Wood, A., Yanco, S.W., Yen, J.D.L., Youngflesh, C., Zilio, G., Zimmer, C., Zimmerman, G.M. and Zitomer, R.A. (2023) 'Same data, different analysts: variation in effect sizes due to analytical decisions in ecology and evolutionary biology'. Available at: https://ecoevorxiv.org/repository/view/6000/?utm_source=miragenews&utm_medium=miragenews&utm_campaign=news (Accessed: 1 August 2024).

Grimmett, L., Whitsed, R. and Horta, A. (2020) 'Presence-only species distribution models are sensitive to sample prevalence: Evaluating models using spatial prediction stability and accuracy metrics', *Ecological Modelling*, 431, p. 109194. Available at: https://doi.org/10.1016/j.ecolmodel.2020.109194.

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

Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A. and Schmullius, C. (2008) 'Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets', *Remote Sensing of Environment*, 112(5), pp. 2538–2556. Available at: https://doi.org/10.1016/j.rse.2007.11.013.

Hijmans, R.J. (2023) 'terra: Spatial Data Analysis'. Available at: https://CRAN.R-project.org/package=terra.

Horvath, P., Halvorsen, R., Simensen, T. and Bryn, A. (2021) 'A comparison of three ways to assemble wall-to-wall maps from distribution models of vegetation types', *GIScience &*

Remote Sensing, 58(8), pp. 1458–1476. Available at: https://doi.org/10.1080/15481603.2021.1996313.

Jansen, J., Woolley, S.N.C., Dunstan, P.K., Foster, S.D., Hill, N.A., Haward, M. and Johnson, C.R. (2022) ‘Stop ignoring map uncertainty in biodiversity science and conservation policy’, *Nature Ecology & Evolution*, 6(7), pp. 828–829. Available at: https://doi.org/10.1038/s41559-022-01778-z.

Keith, D.A., Ferrer-Paris, J.R., Ghoraba, S.M.M., Henriksen, S., Monyeki, M., Murray, N.J., Nicholson, E., Rowland, J.A., Skowno, A., Slingsby, J.A., Storeng, A.B., Valderrábano, M. and Zager, I. (eds) (2024) *Guidelines for the application of IUCN Red List of ecosystems categories and criteria version 2*. IUCN International Union for Conservation of Nature. Available at: https://doi.org/10.2305/IUCN.CH.2016.RLE.1.en.

Keith, D.A., Ferrer-Paris, J.R., Nicholson, E., Bishop, M.J., Polidoro, B.A., Ramirez-Llodra, E., Tozer, M.G., Nel, J.L., Mac Nally, R., Gregr, E.J., Watermeyer, K.E., Essl, F., Faber-Langendoen, D., Franklin, J., Lehmann, C.E.R., Etter, A., Roux, D.J., Stark, J.S., Rowland, J.A., Brummitt, N.A., Fernandez-Arcaya, U.C., Suthers, I.M., Wiser, S.K., Donohue, I., Jackson, L.J., Pennington, R.T., Iliffe, T.M., Gerovasileiou, V., Giller, P., Robson, B.J., Pettorelli, N., Andrade, A., Lindgaard, A., Tahvanainen, T., Terauds, A., Chadwick, M.A., Murray, N.J., Moat, J., Pliscoff, P., Zager, I. and Kingsford, R.T. (2022) ‘A function-based typology for Earth’s ecosystems’, *Nature*, 610(7932), pp. 513–518. Available at: https://doi.org/10.1038/s41586-022-05318-4.

Keith, D.A., Ghoraba, S.M.M., Kaly, E., Jones, K.R., Oosthuizen, A., Obura, D., Costa, H.M., Daniels, F., Duarte, E., Grantham, H., Gudka, M., Norman, J., Shannon, L.J., Skowno, A. and Ferrer-Paris, J.R. (2024) ‘Contributions of the IUCN Red List of Ecosystems to risk-based design and management of protected and conserved areas in Africa’, *Conservation Biology*, 38(3), p. e14169. Available at: https://doi.org/10.1111/cobi.14169.

Khatami, R., Mountrakis, G. and Stehman, S.V. (2016) ‘A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: General guidelines for practitioners and future research’, *Remote Sensing of Environment*, 177, pp. 89–100. Available at: https://doi.org/10.1016/j.rse.2016.02.028.

Kuhn, M. (2008) ‘Building Predictive Models in R Using the caret Package’, *Journal of Statistical Software*, 28(5), pp. 1–26. Available at: https://doi.org/10.18637/jss.v028.i05.

Liddle, D.T. and Elliott, L.P. (2008) ‘Tiwi Island threatened plants 2006 to 2008: field survey, population monitoring including establishment of a program to investigate the impact of pigs, and weed control.’, p. 50.

Loosvelt, L., Peters, J., Skriver, H., Lievens, H., Van Coillie, F.M.B., De Baets, B. and Verhoest, N.E.C. (2012) ‘Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification', *International Journal of Applied Earth Observation and Geoinformation*, 19, pp. 173–184. Available at: https://doi.org/10.1016/j.jag.2012.05.011.

Maynard, J.J., Yeboah, E., Owusu, S., Buenemann, M., Neff, J.C. and Herrick, J.E. (2023) 'Accuracy of regional-to-global soil maps for on-farm decision-making: are soil maps "good enough"?', *SOIL*, 9(1), pp. 277–300. Available at: https://doi.org/10.5194/soil-9-277-2023.

McIver, D.K. and Friedl, M.A. (2001) 'Estimating pixel-scale land cover classification confidence using nonparametric machine learning methods', *IEEE Transactions on Geoscience and Remote Sensing*, 39(9), pp. 1959–1968. Available at: https://doi.org/10.1109/36.951086.

Menkhorst, K.A. and Woinarski, J.C.Z. (1992) 'Distribution of mammals in monsoon rainforests of the Northern Territory', *Wildlife Research*, 19, pp. 295–316. Available at: https://doi.org/10.1071/WR9920295.

Mitchell, P.J., Downie, A.-L. and Diesing, M. (2018) 'How good is my map? A tool for semi-automated thematic mapping and spatially explicit confidence assessment', *Environmental Modelling & Software*, 108, pp. 111–122. Available at: https://doi.org/10.1016/j.envsoft.2018.07.014.

Morales-Barquero, L., Lyons, M.B., Phinn, S.R. and Roelfsema, C.M. (2019) 'Trends in Remote Sensing Accuracy Assessment Approaches in the Context of Natural Resources', *Remote Sensing*, 11(19), p. 2305. Available at: https://doi.org/10.3390/rs11192305.

Mucina, L. (2019) 'Biome: evolution of a crucial ecological and biogeographical concept', *New Phytologist*, 222(1), pp. 97–114. Available at: https://doi.org/10.1111/nph.15609.

Murray, N.J., Keith, D.A., Bland, L.M., Nicholson, E., Regan, T.J., Rodríguez, J.P. and Bedward, M. (2017) 'The use of range size to assess risks to biodiversity from stochastic threats', *Diversity and Distributions*, 23(5), pp. 474–483. Available at: https://doi.org/10.1111/ddi.12533.

Murray, N.J., Keith, D.A., Duncan, A., Tizard, R., Ferrer-Paris, J.R., Worthington, T.A., Armstrong, K., Nyan Hlaing, Win Thuya Htut, Aung Htat Oo, Kyaw Zay Ya and Grantham, H. (2020) 'Myanmar's terrestrial ecosystems: Status, threats and conservation opportunities', *Biological Conservation*, 252, p. 108834. Available at: https://doi.org/10.1016/j.biocon.2020.108834.

Murray, N.J., Phinn, S.R., DeWitt, M., Ferrari, R., Johnston, R., Lyons, M.B., Clinton, N., Thau, D. and Fuller, R.A. (2019) 'The global distribution and trajectory of tidal flats', *Nature*, 565(7738), pp. 222–225. Available at: https://doi.org/10.1038/s41586-018-0805-8.

Naas, A.E., Halvorsen, R., Horvath, P., Wollan, A.K., Bratli, H., Brynildsrud, K., Finne, E.A., Keetz, L.T., Lieungh, E., Olson, C., Simensen, T., Skarpaas, O., Tandstad, H.R., Torma, M., Værland, E.S. and Bryn, A. (2023) ‘What explains inconsistencies in field-based ecosystem mapping?’, *Applied Vegetation Science*, 26(1), p. e12715. Available at: https://doi.org/10.1111/avsc.12715.

Naas, A.E., Keetz, L.T., Halvorsen, R., Horvath, P., Mienna, I.M., Simensen, T. and Bryn, A. (2024) ‘Choice of predictors and complexity for ecosystem distribution models: effects on performance and transferability’, *Ecography*, n/a(n/a), p. e07269. Available at: https://doi.org/10.1111/ecog.07269.

Nagendra, H. and Rocchini, D. (2008) ‘High resolution satellite imagery for tropical biodiversity studies: the devil is in the detail’, *Biodiversity and Conservation*, 17(14), pp. 3431–3442. Available at: https://doi.org/10.1007/s10531-008-9479-0.

Neave, G., Murphy, B.P., Tiwi Rangers, Andersen, A. and Davies, H.F. (2024) ‘The intact and the imperilled: contrasting mammal population trajectories between two large adjacent islands’, *Wildlife Research* [Preprint].

Nicholson, E., Andrade, A., Brooks, T.M., Driver, A., Ferrer-Paris, J.R., Grantham, H.S., Gudka, M.S., Keith, D.A., Kontula, T., Lindgaard, A., Londono-Murcia, M.C., Murray, N.J., Raunio, A., Rowland, J.A., Sievers, M., Skowno, A.L., Stevenson, S.L., Valderrabano, M., Vernon, C.M., Zager, I. and Obura, D. (2024) ‘Roles of the Red List of Ecosystems in the Kunming-Montreal Global Biodiversity Framework’, *Nature Ecology & Evolution* [Preprint]. Available at: https://doi.org/10.1038/s41559-023-02320-5.

Noh, J.K., Echeverria, C., Kleemann, J., Koo, H., Fürst, C. and Cuenca, P. (2020) ‘Warning about conservation status of forest ecosystems in tropical Andes: National assessment based on IUCN criteria’, *PLOS ONE*. Edited by R. Nóbrega, 15(8), p. e0237877. Available at: https://doi.org/10.1371/journal.pone.0237877.

Olofsson, P., Arévalo, P., Espejo, A.B., Green, C., Lindquist, E., McRoberts, R.E. and Sanz, M.J. (2020) ‘Mitigating the effects of omission errors on area and area change estimates’, *Remote Sensing of Environment*, 236, p. 111492. Available at: https://doi.org/10.1016/j.rse.2019.111492.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D’amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. and Kassem, K.R. (2001) ‘Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity’, *BioScience*, 51(11), pp. 933–938. Available at: https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Pettorelli, N., Laurance, W.F., O’Brien, T.G., Wegmann, M., Nagendra, H. and Turner, W. (2014) ‘Satellite remote sensing for applied ecologists: opportunities and challenges’, *Journal of Applied Ecology*. Edited by E.J. Milner-Gulland, 51(4), pp. 839–848. Available at: https://doi.org/10.1111/1365-2664.12261.

Pettorelli, N., Williams, J., Schulte to Bühne, H. and Crowson, M. (2024) 'Deep learning and satellite remote sensing for biodiversity monitoring and conservation', *Remote Sensing in Ecology and Conservation* [Preprint]. Available at: https://doi.org/10.1002/rse2.415.

Pontius Jr, R.G. and Millones, M. (2011) 'Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment', *International Journal of Remote Sensing*, 32(15), pp. 4407–4429. Available at: https://doi.org/10.1080/01431161.2011.552923.

QGIS Development Team (2018) 'QGIS Geographic Information System'.

R Core Team (2018) 'R: A language and environment for statistical computing'. Vienna, Austria: R Foundation for Statistical Computing.

Regan, H.M., Colyvan, M. and Burgman, M.A. (2002) 'A taxonomy and treatment of uncertainty for ecology and conservation biology', *Ecological Applications*, 12(2), pp. 618–628. Available at: https://doi.org/10.1890/1051-0761(2002)012[0618:ATATOU]2.0.CO;2.

Remmel, T.K. (2009) 'Investigating Global and Local Categorical Map Configuration Comparisons Based on Coincidence Matrices', *Geographical Analysis*, 41(2), pp. 144–157. Available at: https://doi.org/10.1111/j.1538-4632.2009.00738.x.

Richards, A.E., Andersen, A.N., Schatz, J., Eager, R., Dawes, T.Z., Hadden, K., Scheepers, K. and Van Der Geest, M. (2012) 'Savanna burning, greenhouse gas emissions and indigenous livelihoods: Introducing the Tiwi Carbon Study: THE TIWI CARBON STUDY', *Austral Ecology*, 37(6), pp. 712–723. Available at: https://doi.org/10.1111/j.1442-9993.2012.02395.x.

Rivas, C.A., Guerrero-Casado, J. and Navarro-Cerillo, R.M. (2021) 'Deforestation and fragmentation trends of seasonal dry tropical forest in Ecuador: impact on conservation', *Forest Ecosystems*, 8(1), p. 46. Available at: https://doi.org/10.1186/s40663-021-00329-5.

Rocchini, D., Foody, G.M., Nagendra, H., Ricotta, C., Anand, M., He, K.S., Amici, V., Kleinschmit, B., Förster, M., Schmidtlein, S., Feilhauer, H., Ghisla, A., Metz, M. and Neteler, M. (2013) 'Uncertainty in ecosystem mapping by remote sensing', *Computers & Geosciences*, 50, pp. 128–135. Available at: https://doi.org/10.1016/j.cageo.2012.05.022.

Rossiter, D.G., Poggio, L., Beaudette, D. and Libohova, Z. (2022) 'How well does digital soil mapping represent soil geography? An investigation from the USA', *SOIL*, 8(2), pp. 559–586. Available at: https://doi.org/10.5194/soil-8-559-2022.

RStudio Team (2020) 'RStudio: Integrated Development for R'. Boston, MA: RStudio, PBC. Available at: http://www.rstudio.com.

Russell-Smith, J. (1991) ‘Classification, species richness, and environmental relations of monsoon rain forest in northern Australia’, *Journal of Vegetation Science*, 2(2), pp. 259–278. Available at: https://doi.org/10.2307/3235959.

Scarth, P., Armston, J., Lucas, R. and Bunting, P. (2023) ‘Vegetation Height and Structure - Derived from ALOS-1 PALSAR, Landsat and ICESat/GLAS, Australia Coverage.’ Available at: https://portal.tern.org.au/metadata/TERN/de1c2fef-b129-485e-9042-8b22ee616e66.

Simensen, T., Horvath, P., Vollering, J., Erikstad, L., Halvorsen, R. and Bryn, A. (2020) ‘Composite landscape predictors improve distribution models of ecosystem types’, *Diversity and Distributions*, 26(8), pp. 928–943. Available at: https://doi.org/10.1111/ddi.13060.

Smith, A., Murphy, S., Herderson, D. and Erickson, K. (2023) ‘Including imprecisely georeferenced specimens improves accuracy of species distribution models and estimates of niche breadth’, *Global Ecology & Biogeography*, 32(3), pp. 342–355. Available at: https://doi.org/doi:10.1111/geb.13628.

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

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

Trouvé, R., Jiang, R., Fedrigo, M., White, M.D., Kasel, S., Baker, P.J. and Nitschke, C.R. (2023) ‘Combining Environmental, Multispectral, and LiDAR Data Improves Forest Type Classification: A Case Study on Mapping Cool Temperate Rainforests and Mixed Forests’, *Remote Sensing*, 15(1), p. 60. Available at: https://doi.org/10.3390/rs15010060.

UNSD (2021) *System of Environmental-Economic Accounting—Ecosystem Accounting: Final Draft version 5*. Department of Economic and Social Affairs, Statistical Division, United Nations, pp. 1–350. Available at: https://unstats.un.org/unsd/envaccounting/seeaRev/ SEEA_CF_Final_en.pdf (Accessed: 24 January 2024).

Venter, Z.S., Barton, D.N., Chakraborty, T., Simensen, T. and Singh, G. (2022) ‘Global 10 m Land Use Land Cover Datasets: A Comparison of Dynamic World, World Cover and Esri Land Cover’, *Remote Sensing*, 14(16), p. 4101. Available at: https://doi.org/10.3390/rs14164101.

Venter, Z.S., Czúcz, B., Stange, E., Nowell, M.S., Simensen, T., Immerzeel, B. and Barton, D.N. (2024) ‘“Uncertainty audit” for ecosystem accounting: Satellite-based ecosystem extent is biased without design-based area estimation and accuracy assessment’, *Ecosystem Services*, 66, p. 101599. Available at: https://doi.org/10.1016/j.ecoser.2024.101599.

Viscarra Rossel, R.A., Chen, C., Grundy, M.J., Searle, R., Clifford, D. and Campbell, P.H. (2015) ‘The Australian three-dimensional soil grid: Australia’s contribution to the GlobalSoilMap project’, *Soil Research*, 53(8), p. 845. Available at: https://doi.org/10.1071/SR14366.

Watson, J.E.M., Keith, D.A., Strassburg, B.B.N., Venter, O., Williams, B. and Nicholson, E. (2020) ‘Set a global target for ecosystems’, *Nature*, 578(7795), pp. 360–362. Available at: https://doi.org/10.1038/d41586-020-00446-1.

Watson, J.E.M., Venegas-Li, R., Grantham, H., Dudley, N., Stolton, S., Rao, M., Woodley, S., Hockings, M., Burkart, K., Simmonds, J.S., Sonter, L.J., Sreekar, R., Possingham, H.P. and Ward, M. (2023) ‘Priorities for protected area expansion so nations can meet their Kunming-Montreal Global Biodiversity Framework commitments’, *Integrative Conservation*, 2(3), pp. 140–155. Available at: https://doi.org/10.1002/inc3.24.

Whittaker, R.H. (1956) ‘Vegetation of the Great Smoky Mountains’, *Ecological Monographs*, 26(1), pp. 2–80. Available at: https://doi.org/10.2307/1943577.

Wilson, B.A. and Fensham, R.J. (1994) ‘A comparison of classification systems for the conservation of sparsely wooded plains on Melville Island, Northern Australia’, *Australian Geographer*, 25(1), pp. 18–31. Available at: https://doi.org/10.1080/00049189408703095.

Wohlfart, C., Wegmann, M. and Leimgruber, P. (2014) ‘Mapping Threatened Dry Deciduous Dipterocarp Forest in South-East Asia for Conservation Management’, *Tropical Conservation Science*, 7(4), pp. 597–613. Available at: https://doi.org/10.1177/194008291400700402.

Wright, M.N. and Ziegler, A. (2017) ‘Ranger: a fast implementation of random forests for high dimensional data in C++ and R.’, *Journal of Statistical Software*, 77(1), pp. 1–17. Available at: https://doi.org/doi:10.18637/jss.v077.i01.

Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R. and Woodcock, C.E. (2012) ‘Opening the archive: How free data has enabled the science and monitoring promise of Landsat’, *Remote Sensing of Environment*, 122, pp. 2–10. Available at: https://doi.org/10.1016/j.rse.2012.01.010.

Xiao, H., Driver, A., Etter, A., Keith, D.A., Obst, C., Traurig, M.J. and Nicholson, E. (2024) ‘Synergies and complementarities between ecosystem accounting and the Red List of Ecosystems’, *Nature Ecology & Evolution*, 8, pp. 1794–1803. Available at: https://doi.org/10.1038/s41559-024-02494-6.

Young, A.R., Davies, H.F., Ayre, M.L., Brekelmans, A., Bryan, B.A., Elith, J., Hadden, K., Kerinaiua, M., Keith, D.A., Lewis, D.L., Munkara-Murray, K.M., Ryan, S., Spencer, M. and

Nicholson, E. (2024) ‘Applying the Global Ecosystem Typology to classify, describe, and map ecosystems from regional data and Indigenous knowledge’. EcoEvoRxiv. Available at: https://doi.org/10.32942/X20P75.

Yu, Le, Liang, L., Wang, J., Zhao, Y., Cheng, Q., Hu, L., Liu, S., Yu, Liang, Wang, X., Zhu, P., Li, Xueyan, Xu, Y., Li, C., Fu, W., Li, Xuecao, Li, W., Liu, C., Cong, N., Zhang, H., Sun, F., Bi, X., Xin, Q., Li, D., Yan, D., Zhu, Z., Goodchild, M.F. and Gong, P. (2014) ‘Meta-discoveries from a synthesis of satellite-based land-cover mapping research’, *International Journal of Remote Sensing*, 35(13), pp. 4573–4588. Available at: https://doi.org/10.1080/01431161.2014.930206.

Supplementary material

Appendix 1 – Modelling methodology

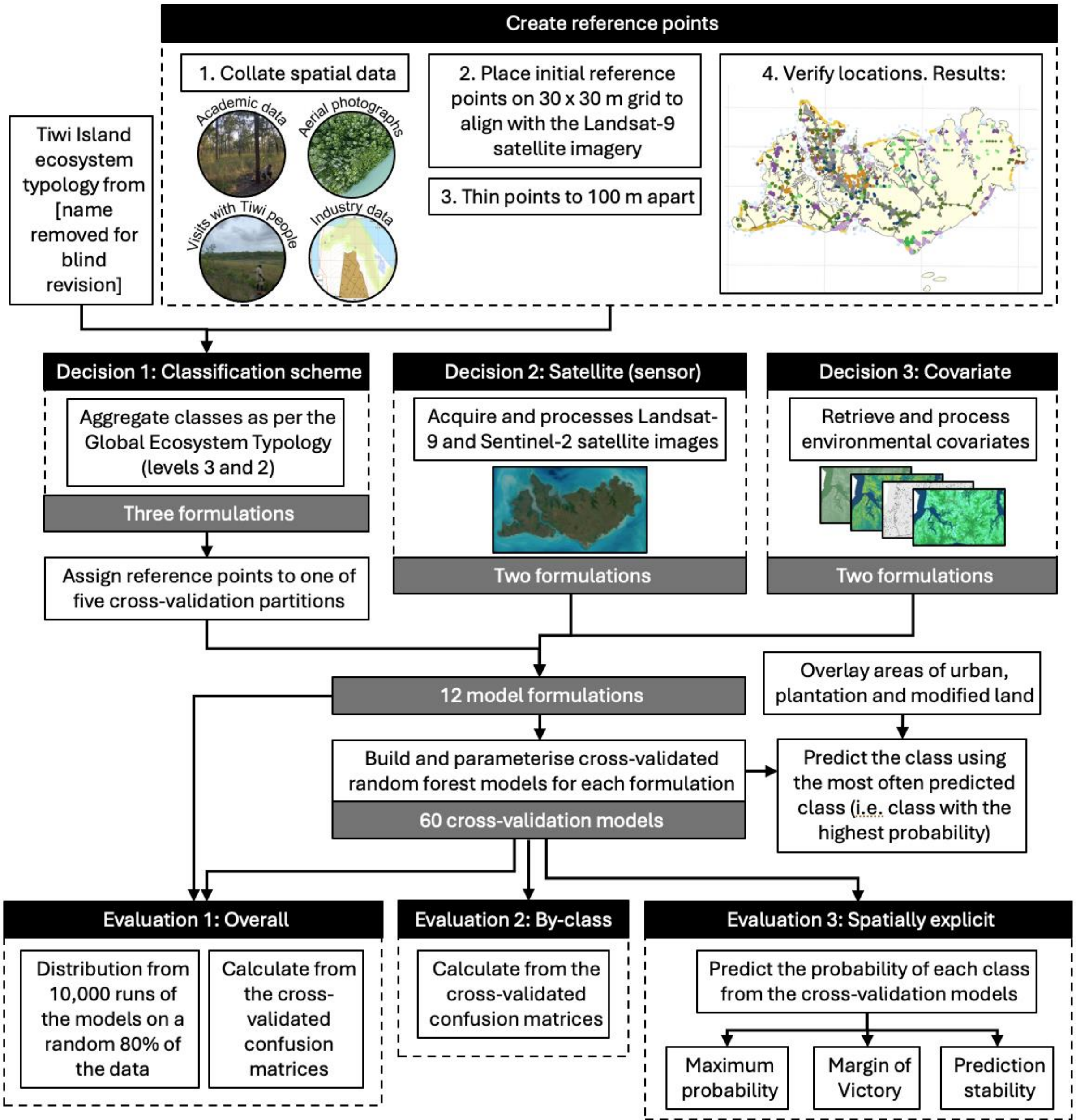


Figure 5. Flow chart of the methods to test three modelling decisions on mapping the extent of ecosystems and assess the decisions with three assessment metrics.

Appendix 2 – Software

Software

QGIS (version 3.22.12)

Google Earth Engine (Gorelick et al., 2017)

R (version 4.3.0) (R Core Team, 2018)
R-studio (version 2023.09.1+949) (RStudio Team, 2020)

R packages

Satellite imagery and environmental covariates:

- 'rgee' (version 1.1.6.9999) (Aybar *et al.*, 2020)
- 'rgeeExtra' (version 0.0.1) (Aybar *et al.*, 2020)

Data cleaning and manipulation:

- 'enmSdmX' package (version 1.1.2) (Smith *et al.*, 2023)
- 'dplyr' (version 1.1.2) (Wickham *et al.*, 2023)
- 'tidyr' (version 1.3.0) (Wickham, Vaughan and Girlich, 2023)
- 'stringr' (version 1.5.0) (Wickham, 2022)

Spatial data handling:

- 'sf' (version 1.0-16) (Pebesma, 2018)
- 'terra' (version 1.7-29) (Hijmans, 2023)

Model fitting, evaluation and prediction:

- 'ranger' (version 0.15.1) (Wright and Ziegler, 2017)
- 'vip' (version 0.3.2) (Greenwell and Boehmke, 2020)
- 'caret' (version 6.0-94) (Kuhn, 2008)

Visualisations:

- 'tidyterra' (version 0.4.0) (Hernangomez, 2024)
- 'ggplot2' (version 3.4.3) (Wickham, 2016)
- 'ggspatial' (version 1.1.8) (Dunnington, 2023)
- 'ggh4x' (version 0.2.8) (van den Brand, 2024)
- 'ggnewscale' (version 0.4.9) (Campitelli, 2023)
- 'ggstance' (version 0.3.7) (Henry, Wickham and Chang, 2024)

Appendix 3 – Satellite imager processing

We applied scaling factors to the satellite images obtained from the Landsat-9 satellite with the OLI-2 sensor and from the Sentinel-2 satellite with the MSI sensor. For the optical bands (i.e. the name begins with SR) of Landsat-9 OLI images, the band was first multiplied by 2.75*e-5 then minuses 0.2. For the thermal bands (i.e. the name begins with ST) of Landsat-9, the band was first multiplied by 3.41802*e-3 then added 149. The Sentinel-2 images were scaled by 0.0001 to reverse the scaling factor applied for efficient data storage.

To mask the clouds in the Landsat-9 images, we used the quality assessment bands for the cloud and cloud shadow (bits 3 and 5). For the Sentinel-2 images, we used the Scene Classification Layer and removed the pixels classified as no data (SCL = 0), saturated (SCL = 1), medium or high cloud probability (SCL = 8 and 9), high cirrus cloud (SCL = 10), snow and ice (SCL = 11).

Appendix 4 – Environmental covariates

We tested correlation in the variables using the absolute value of the Pearson's correlation coefficient with a cut-off of 0.7 (Figure 2). The red, green and blue bands were all highly correlated. NDVI was least correlated to the red band for both satellites.
Each of the soil variables were correlated between the two depths. We retained the top 30 cm variables to reflect the root zone of more of the plant species. Nitrogen, phosphorus and soil sand content were highly correlated to near infrared and NDVI for the Landsat-9 variables and hence removed. We retained slope instead of the correlated TRI to represent rainfall run-off and easier interpretation of the results. Each of the vegetation biomass height variables were correlated. We retained the height of 50% of the biomass as it was least correlated to all the other covariates.

Table 2. Details of the environmental covariates.

Layer	Description	Rational	Source
Satellite image covariates			
Red	The red, green, blue and near infrared bands.	Spectral characteristics represent physical and chemical attributes of the ecosystem.	Landsat-9 satellite atmospherically corrected surface reflectance (level 2, collection 2, tier 1) courtesy of the United States Geological Survey (USGS). For Landsat-9, the red band is B4, green is B3, blue is B2 and near infrared is B5. Sentinel-2 surface reflectance harmonised collection (level-2A) with atmospheric correction from the Copernicus Sentinel missions are by the European Space Agency (ESA). For Sentinel-2, the red band is B4, green is B3, blue is B2 and near infrared is B8.
Green			
Blue			
NIR			
NDVI	Normalised difference vegetation index.	Greenness of the canopy which is correlated to primary productivity.	Calculated from the satellite image using the red and near infrared bands where: $NDVI = \frac{NIR - Red}{NIR + Red}$
Additional covariates			
Height_50	The height where 50, 75 and 95% of the plant cover has been intercepted.	The height of the vegetation biomass relates to the vegetation structure.	Terrestrial Ecosystem Research Network https://portal.tern.org.au/metadata/TERN/de1c2fef-b129-485e-9042-8b22ee616e66
Height_75			
Height_95			
Elev	Elevation in meters.	The elevation is a proxy for range of environmental relationships including access to groundwater, influence of floods, exposure to wind on hilltops, and exposure to wave disturbances on coastal ecosystem. The topographic measures of the slope, position and roughness also relate to soil moisture and run off which strongly drive ecosystem functioning.	The Smoothed Digital Elevation model (DEM-S) at a 5 m resolution from the Shuttle Radar Topography Mission (SRTM) by from Geoscience Australia in 2000 https://developers.google.com/earth-engine/datasets/catalog/AU_GA_DEM_1SEC_v10_DEM-S
Slp	Slope in degrees.		Created using the 'terrain' function from the 'terra' package in R on the elevation model. Slope was computed with the four neighbouring cells and measured in degrees.
TRI	Topographic roughness index.		
Clay	Percentage of 1) clay, 2) silt, 3) sand, 4) soil organic carbon, 5) nitrogen or 6) phosphorus in the top 30 cm and 2 m of the soil.	The soil composition influences many aspects of plant growth and soil moisture, including nutrient availability and drainage.	Soil and Landscape Grid of Australia. Averaged by the depth over which the attribute was measured (depth-weighted average). https://dx.doi.org/10.1071/SR14366
Silt			
Sand			
SOC			
NTO			
PTO			

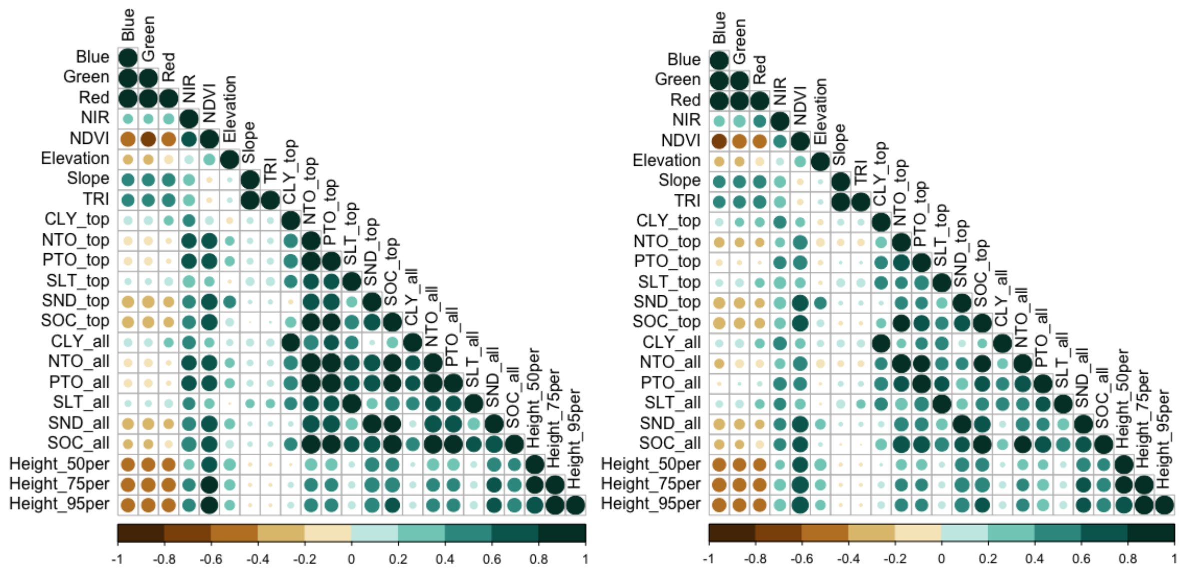


Figure 6. Correlation of the environmental predictors at a 30 m resolution with Landsat-9 satellite imagery using the OLI-2 sensor (left) and a 10 m resolution with Sentinel-2 satellite imagery using the MSI sensor (right).

Appendix 5 – Model evaluation

For the confusion matrix

		Reference	
		1	0
Predicted	1	a	b
	0	c	d

a represents the number of true positive values, b the false positives, c the false negatives and d the true negatives. This confusion matrix is used to calculate the evaluation metrics in Table 3.

Table 3. Descriptions of the overall and by-class evaluation metrics.

Evaluation metric	Other names	Equation	Description
Overall metrics			
Accuracy		$\frac{a+d}{a+b+c+d}$	A measure of agreement between the predicted and true values, such that 1 indicates perfect agreements and 0 indicates no agreement.
Kappa statistic	Cohen's kappa	$N = a+b+c+d$ $p_0 = \text{Accuracy}$ $p_e = \frac{a+c}{N} \times \frac{a+b}{N} + \frac{b+d}{N} \times \frac{c+d}{N}$ $Kappa = \frac{p_0 - p_e}{1 - p_e}$	A measure of agreement between the predicted and true values, such that 1 indicates perfect agreements and 0 indicates no more agreement than expected by chance.
Out-of-bag error (OOB)	Out-of-bag score		The average error for the random forest trees using bootstrap aggregation and calculated on the out-of-bag samples.
By-class metrics			
Sensitivity	Producer's accuracy, recall, true positive rate	$\frac{a}{a+c}$	The ability of the model to correctly identify all the true cases from those known to be true.
Specificity	True negative rate	$\frac{d}{b+d}$	The ability of the model to correctly identify all the false cases from those known to be false.
Precision	User's accuracy, positive predicted value	$\frac{a}{a+b}$	The ability of the model to correctly identify all the true cases from those predicted to the class.
F1		$2 \times \frac{Sensitivity \times Precision}{Sensitivity + Precision}$	A balance of the models ability to predict the true cases from thoses known to be true (i.e. sensitivity) and the correctly true from all those predicted to be true (i.e. precision).
Negative predicted value		$\frac{d}{c+d}$	The ability of the model to correctly identify all the false cases from those predicted to be false.

Appendix 6 – Additional model results

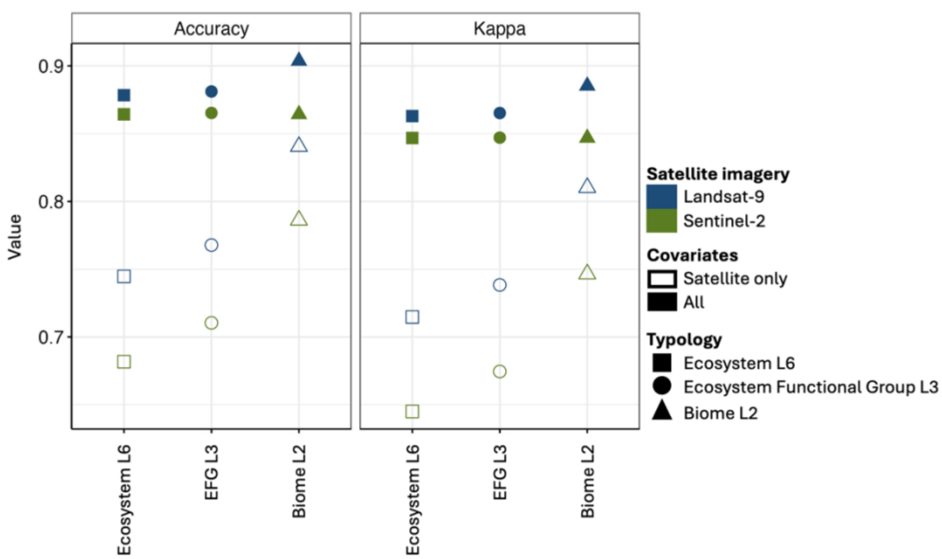
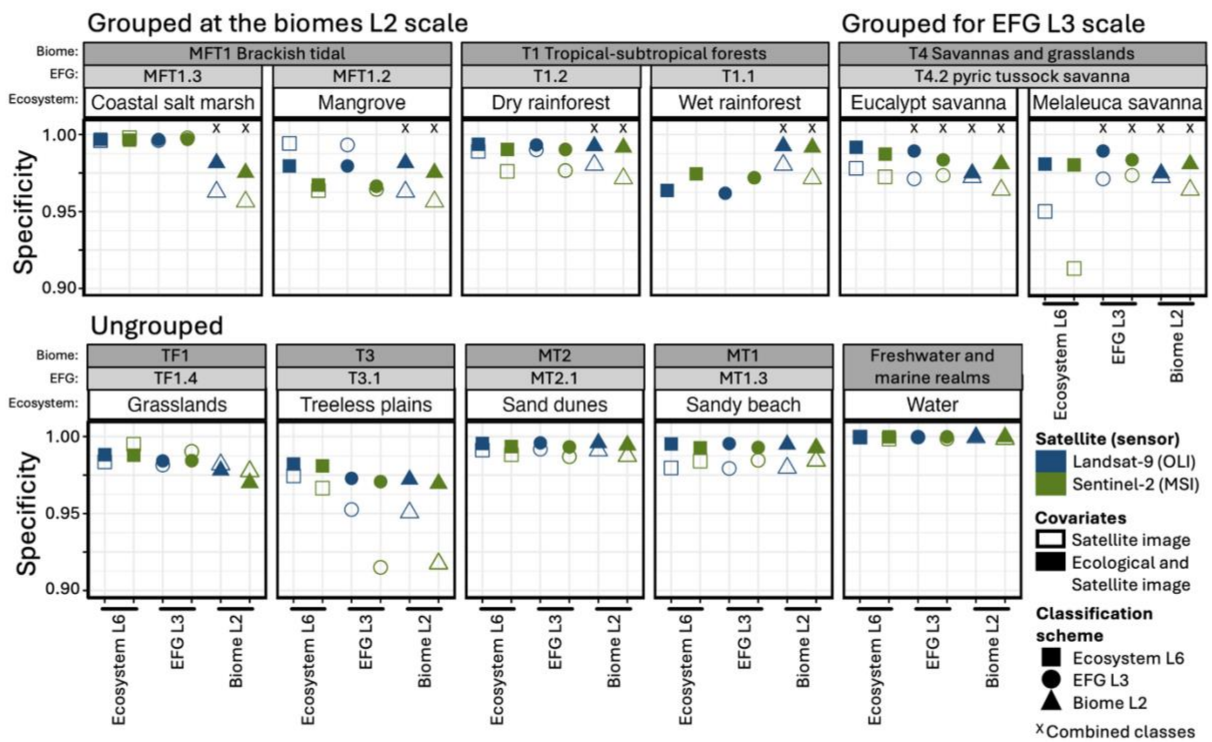
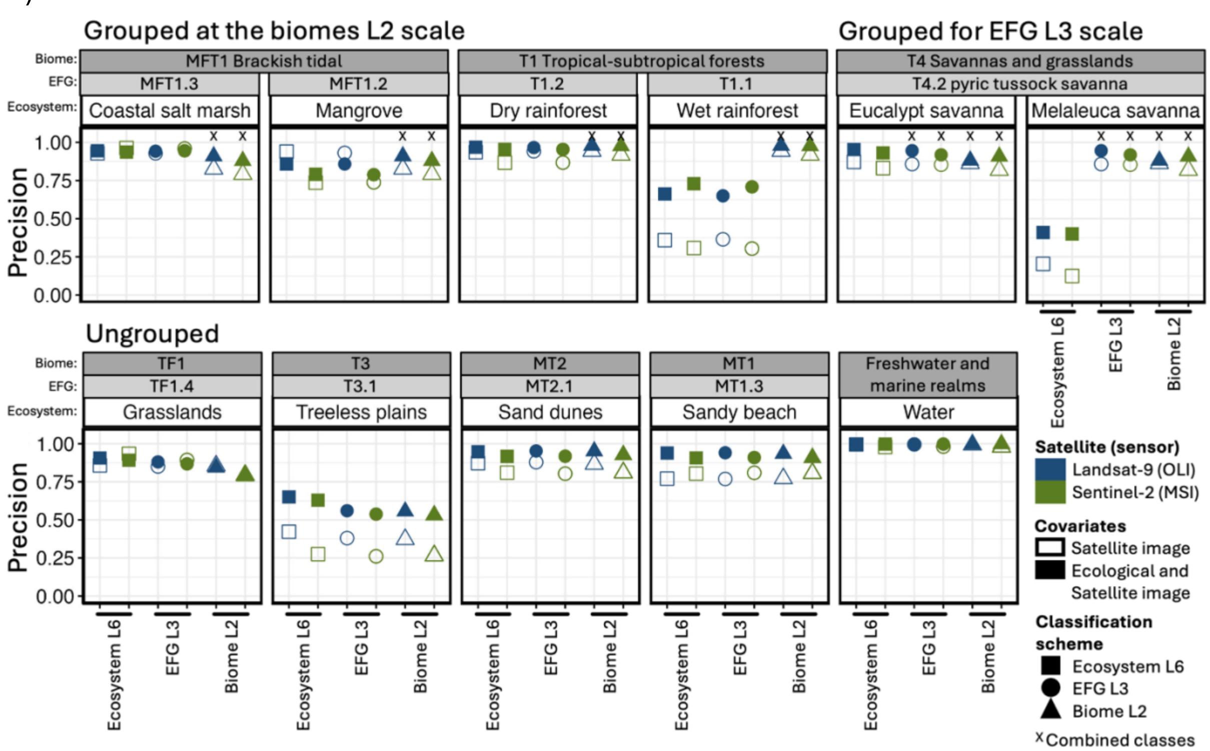


Figure 7. The mean accuracy and kappa statistics calculated from the confusion matrix of 12 model formulations varying at three modelling decisions and each run with five cross-validated models. The modelling decisions were the typology (shape), covariates (fill) and satellite imagery (colours).

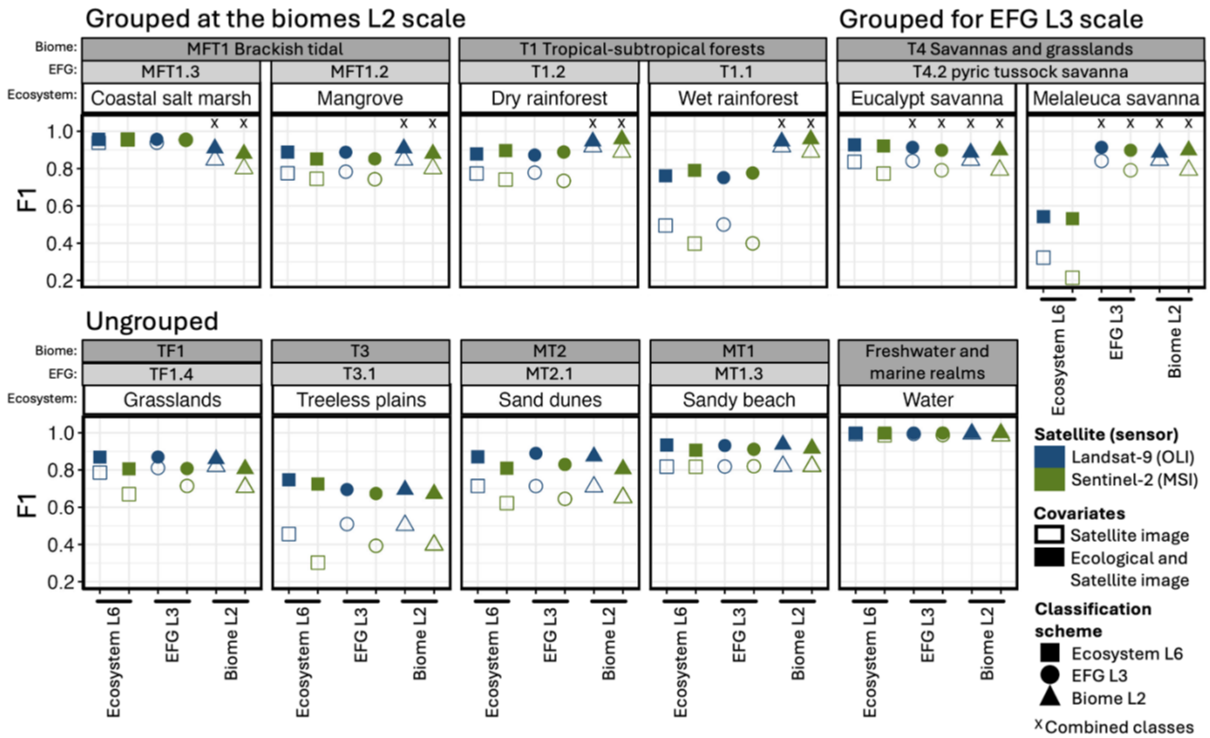
A)


Grouped at the biomes L2 scale
Grouped for EFG L3 scale
Biome:
EFG:
Ecosystem:
MFT1 Brackish tidal
MFT1.3
MFT1.2
Coastal salt marsh
Mangrove
T1 Tropical-subtropical forests
T1.2
T1.1
Dry rainforest
Wet rainforest
T4 Savannas and grasslands
T4.2 pyric tussock savanna
Eucalypt savanna
Melaleuca savanna
Specificity
1.00
0.95
0.90
Ungrouped
TF1
TF1.4
Grasslands
T3
T3.1
Treeless plains
MT2
MT2.1
Sand dunes
MT1
MT1.3
Sandy beach
Freshwater and marine realms
Water
Ecosystem L6
EFG L3
Biome L2
Satellite (sensor)
Landsat-9 (OLI)
Sentinel-2 (MSI)
Covariates
Satellite image
Ecological and Satellite image
Classification scheme
Ecosystem L6
EFG L3
Biome L2
x Combined classes

B)

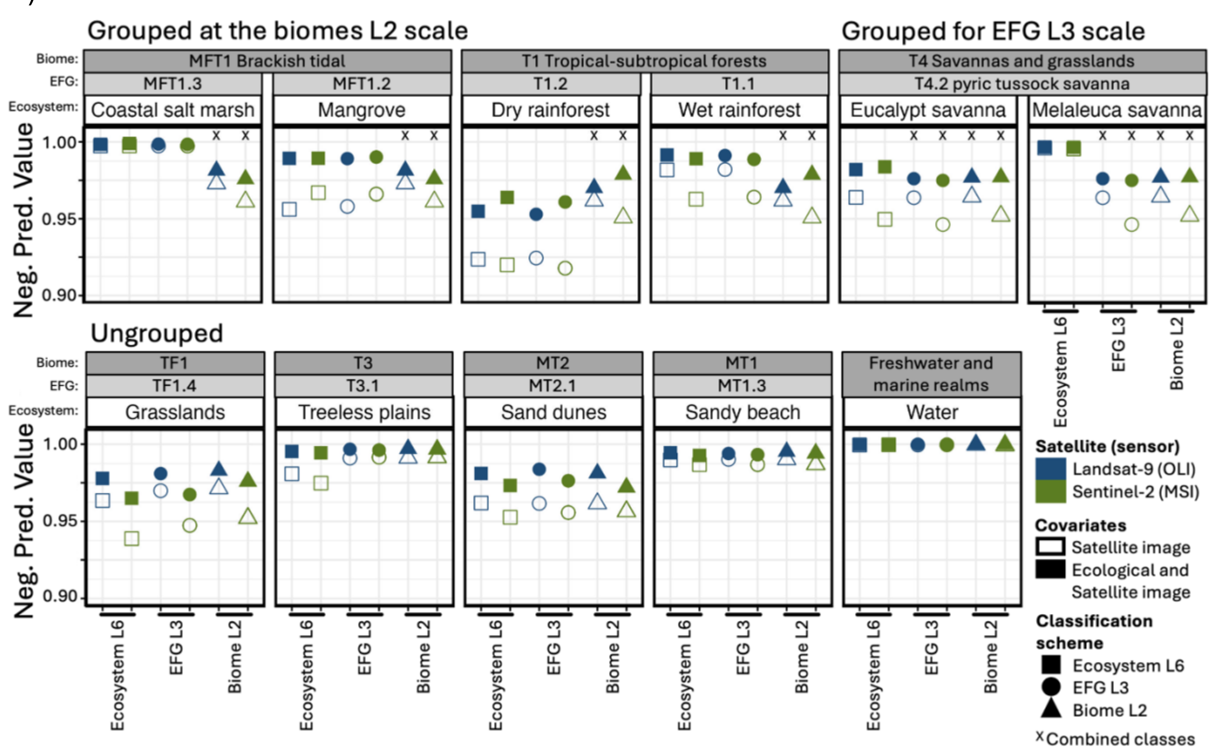

Grouped at the biomes L2 scale
Grouped for EFG L3 scale
Biome:
EFG:
Ecosystem:
MFT1 Brackish tidal
MFT1.3
MFT1.2
Coastal salt marsh
Mangrove
T1 Tropical-subtropical forests
T1.2
T1.1
Dry rainforest
Wet rainforest
T4 Savannas and grasslands
T4.2 pyric tussock savanna
Eucalypt savanna
Melaleuca savanna
Precision
1.00
0.75
0.50
0.25
0.00
Ungrouped
TF1
TF1.4
Grasslands
T3
T3.1
Treeless plains
MT2
MT2.1
Sand dunes
MT1
MT1.3
Sandy beach
Freshwater and marine realms
Water
Ecosystem L6
EFG L3
Biome L2
Satellite (sensor)
Landsat-9 (OLI)
Sentinel-2 (MSI)
Covariates
Satellite image
Ecological and Satellite image
Classification scheme
Ecosystem L6
EFG L3
Biome L2
x Combined classes

C)



Grouped at the biomes L2 scale
Grouped for EFG L3 scale
Biome:
EFG:
Ecosystem:
MFT1 Brackish tidal
MFT1.3
MFT1.2
Coastal salt marsh
Mangrove
T1 Tropical-subtropical forests
T1.2
T1.1
Dry rainforest
Wet rainforest
T4 Savannas and grasslands
T4.2 pyric tussock savanna
Eucalypt savanna
Melaleuca savanna
F1
1.0
0.8
0.6
0.4
0.2
Ungrouped
TF1
TF1.4
Grasslands
T3
T3.1
Treeless plains
MT2
MT2.1
Sand dunes
MT1
MT1.3
Sandy beach
Freshwater and marine realms
Water
Ecosystem L6
EFG L3
Biome L2
Satellite (sensor)
Landsat-9 (OLI)
Sentinel-2 (MSI)
Covariates
Satellite image
Ecological and Satellite image
Classification scheme
Ecosystem L6
EFG L3
Biome L2
x Combined classes

D)



Grouped at the biomes L2 scale
Grouped for EFG L3 scale
Biome:
EFG:
Ecosystem:
MFT1 Brackish tidal
MFT1.3
MFT1.2
Coastal salt marsh
Mangrove
T1 Tropical-subtropical forests
T1.2
T1.1
Dry rainforest
Wet rainforest
T4 Savannas and grasslands
T4.2 pyric tussock savanna
Eucalypt savanna
Melaleuca savanna
Neg. Pred. Value
1.00
0.95
0.90
Ungrouped
TF1
TF1.4
Grasslands
T3
T3.1
Treeless plains
MT2
MT2.1
Sand dunes
MT1
MT1.3
Sandy beach
Freshwater and marine realms
Water
Ecosystem L6
EFG L3
Biome L2
Satellite (sensor)
Landsat-9 (OLI)
Sentinel-2 (MSI)
Covariates
Satellite image
Ecological and Satellite image
Classification scheme
Ecosystem L6
EFG L3
Biome L2
x Combined classes

Figure 8. The four by-class evaluation metrics specificity (A), precision (B), F1 (C), and negative predicted value (D) measured for three classification schemes (shape), two satellite/sensors (colour) and two covariate sets (fill). When multiple ecosystems (shape: square, label above: white) were aggregated into an ecosystem functional group (shape: circle, label: light grey) or into a biome (shape: triangle, label: dark grey), the class is indicated by an x.

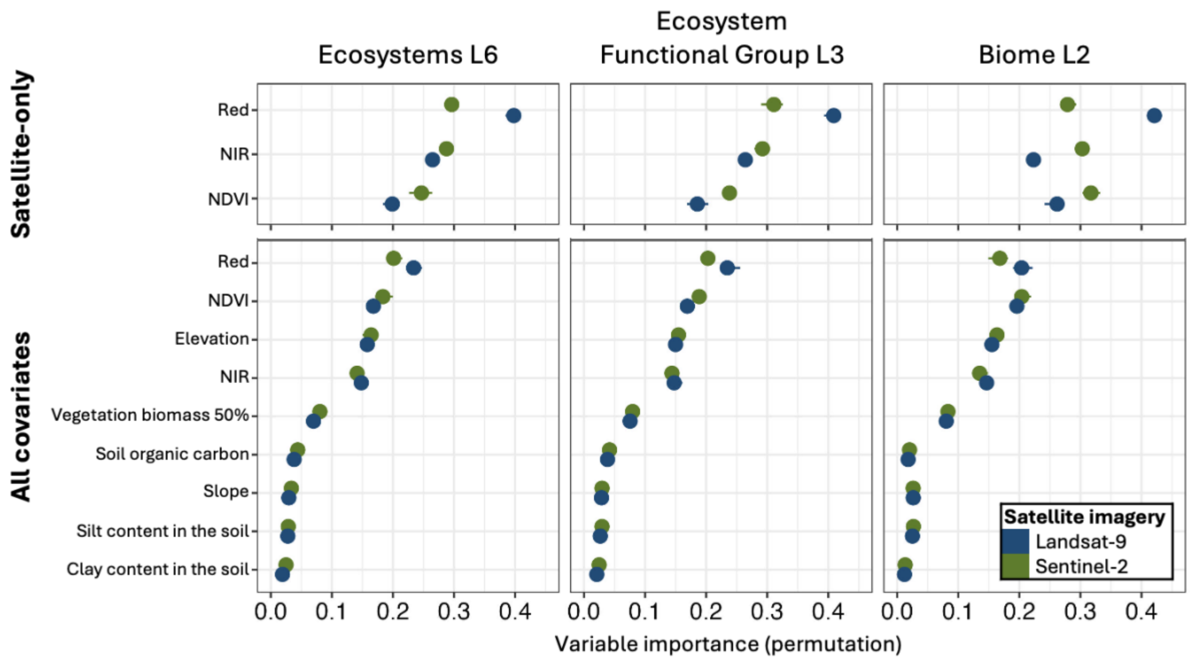


Figure 9. Importance of the environmental covariates in the ecosystem classification model across three classification schemes (columns), two options for the covariates (row) and two satellite (colours). NDVI is for the normalised difference vegetation index and NIR is for the near-infrared band from the satellite image.

Appendix 7 – confusion matrices

Table 4. Confusion matrix for the ecosystem classification model using Landsat-9 satellite imagery from the OLI-2 sensor as the only covariates.

Predicted type	Training points													
	Coastal salt marsh	Dry rainforest	Eucalypt savanna	Grassland and sedgeland	Mangrove	*Melaleuca* savanna	Treeless plains	Sand dunes	Sandy beach	Water	Wet rainforest	Total	UA	CE
Coastal salt marsh	285	0	0	4	1	0	0	1	13	3	0	307	0.93	0.07
Dry rainforest	0	754	0	0	2	0	0	0	0	0	50	806	0.94	0.06
Eucalypt savanna	0	7	745	42	9	9	20	14	0	0	8	854	0.87	0.13
Grassland and sedgeland	7	0	13	511	43	1	4	2	0	0	15	596	0.86	0.14
Mangrove	0	3	0	10	461	0	0	0	0	0	17	491	0.94	0.06
Melaleuca savanna	0	1	111	76	0	74	79	22	0	0	0	363	0.20	0.80
Treeless plains	0	1	39	26	0	12	106	67	0	0	0	251	0.42	0.58
Sand dunes	1	0	0	0	0	0	5	321	41	0	0	368	0.87	0.13
Sandy beach	7	0	0	0	0	0	0	104	373	0	0	484	0.77	0.23
Water	0	0	0	1	0	0	0	0	1	411	0	413	1.00	0.00
Wet rainforest	0	376	19	34	182	0	0	0	0	0	343	954	0.36	0.64
Total	300	1142	927	704	698	96	214	531	428	414	433	5887		
PA	0.95	0.66	0.80	0.73	0.66	0.77	0.50	0.60	0.87	0.99	0.79			
OE	0.05	0.34	0.20	0.27	0.34	0.23	0.50	0.40	0.13	0.01	0.21			

Model formulation

Classification scheme: Ecosystem (level 6 of the Global Ecosystem typology)
Satellite/sensor: Landsat-9/OLI-2
Covariate set: Satellite image covariates only

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Table 5. Confusion matrix for the ecosystem classification model using Landsat-9 satellite imagery from the OLI-2 sensor and using satellite image and additional covariates.

Predicted type	Coastal salt marsh	Dry rainforest	Eucalypt savanna	Grassland and sedgeland	Mangrove	*Melaleuca* savanna	Treeless plains	Sand dunes	Sandy beach	Water	Wet rainforest	Total	UA	CE
	Training points													
Coastal salt marsh	291	0	0	8	1	0	0	1	7	0	0	308	0.94	0.06
Dry rainforest	0	919	0	0	6	0	0	0	0	0	24	949	0.97	0.03
Eucalypt savanna	0	3	837	18	5	1	1	5	0	0	8	878	0.95	0.05
Grassland and sedgeland	3	0	6	588	36	1	4	5	0	0	6	649	0.91	0.09
Mangrove	0	45	1	52	643	0	0	1	0	0	7	749	0.86	0.14
Melaleuca savanna	0	1	37	18	2	77	21	32	0	0	0	188	0.41	0.59
Treeless plains	0	0	27	19	0	17	188	38	0	0	0	289	0.65	0.35
Sand dunes	1	1	0	0	0	0	0	428	22	0	0	452	0.95	0.05
Sandy beach	5	0	0	0	0	0	0	21	398	0	0	424	0.94	0.06
Water	0	0	0	0	0	0	0	0	1	414	0	415	1.00	0.00
Wet rainforest	0	173	19	1	5	96	0	0	0	0	388	586	0.66	0.34
Total	300	1142	927	704	698	96	214	513	428	414	433	5887		
PA	0.97	0.80	0.90	0.84	0.92	0.80	0.88	0.81	0.93	1.00	0.9			
OE	0.03	0.20	0.10	0.16	0.08	0.20	0.12	0.19	0.07	0.00	0.1			

Model formulation

Classification scheme: Ecosystem (level 6 of the Global Ecosystem typology)
Satellite/sensor: Landsat-9/OLI-2
Covariate set: Satellite image and additional covariates

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Table 6. Confusion matrix for the ecosystem classification model using Sentinel-2 satellite imagery from the MSI sensor as the only covariates.

	Training points													
Predicted type	Coastal salt marsh	Dry rainforest	Eucalypt savanna	Grassland and sedgeland	Mangrove	*Melaleuca* savanna	Treeless plains	Sand dunes	Sandy beach	Water	Wet rainforest	Total	UA	CE
Coastal salt marsh	285	0	0	3	1	0	0	1	5	1	0	296	0.96	0.04
Dry rainforest	0	739	0	2	17	0	0	1	0	0	93	852	0.87	0.13
Eucalypt savanna	0	2	671	45	16	16	23	12	0	0	22	807	0.83	0.17
Grassland and sedgeland	3	0	1	368	11	3	3	3	0	0	2	394	0.93	0.07
Mangrove	0	55	0	61	527	0	0	1	0	0	72	716	0.74	0.26
Melaleuca savanna	0	0	217	137	3	72	113	33	0	0	1	576	0.13	0.88
Treeless plains	0	0	19	31	0	5	72	131	4	0	0	262	0.27	0.73
Sand dunes	2	0	0	0	0	0	3	268	58	0	0	331	0.81	0.19
Sandy beach	7	0	0	0	0	0	0	79	356	1	0	443	0.80	0.20
Water	3	0	0	1	0	0	0	0	5	412	0	421	0.98	0.02
Wet rainforest	0	346	19	56	123	0	0	2	0	0	243	789	0.31	0.69
Total	300	1142	927	704	698	96	214	531	428	414	433	5887		
PA	0.95	0.65	0.72	0.52	0.76	0.75	0.34	0.5	0.83	1.00	0.56			
OE	0.05	0.35	0.28	0.48	0.24	0.25	0.66	0.5	0.17	0.00	0.44			

Model formulation

Classification scheme: Ecosystem (level 6 of the Global Ecosystem typology)
Satellite/sensor: Sentinel-2/MSI
Covariate set: Satellite image covariates only

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Table 7 . Confusion matrix for the ecosystem classification model using Sentinel-2 satellite imagery from the MSI sensor and using satellite image and additional covariates.

Predicted type \ Training points	Coastal salt marsh	Dry rainforest	Eucalypt savanna	Grassland and sedgeland	Mangrove	*Melaleuca* savanna	Treeless plains	Sand dunes	Sandy beach	Water	Wet rainforest	Total	UA	CE
Coastal salt marsh	294	0	0	7	1	0	0	5	7	0	0	314	0.94	0.06
Dry rainforest	0	966	1	1	16	0	0	1	0	0	27	1012	0.95	0.05
Eucalypt savanna	0	1	846	30	5	0	2	4	0	0	21	909	0.93	0.07
Grassland and sedgeland	1	1	4	518	26	3	6	19	1	0	2	581	0.89	0.11
Mangrove	0	51	1	107	644	0	0	3	0	0	8	814	0.79	0.21
Melaleuca savanna	0	0	34	21	2	76	21	35	0	0	1	190	0.40	0.60
Treeless plains	0	0	30	17	0	16	183	44	1	0	0	291	0.63	0.37
Sand dunes	0	0	1	2	0	0	2	385	30	0	0	420	0.92	0.08
Sandy beach	5	0	0	0	0	0	0	35	388	0	0	428	0.91	0.09
Water	0	0	0	0	0	0	0	0	1	414	0	415	1.00	0;00
Wet rainforest	0	123	10	1	4	1	0	0	0	0	374	513	0.73	0.27
Total	300	1142	927	704	698	96	214	531	428	414	433	5887		
PA	0.98	0.85	0.91	0.74	0.92	0.79	0.86	0.73	0.91	1.00	0.86			
OE	0.02	0.15	0.09	0.26	0.08	0.21	0.14	0.27	0.09	0.00	0.14			

Model formulation

Classification scheme: Ecosystem (level 6 of the Global Ecosystem typology)
Satellite/sensor: Sentinel-2/MSI
Covariate set: Satellite image and additional covariates

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Table 8. Confusion matrix for the ecosystem functional group classification model using Landsat-9 satellite imagery from the OLI-2 sensor as the only covariates.

		Training points												
		Dry rainforest	Grassland and sedgeland	Mangrove	Coastal salt marsh	Sandy beach	Sand dunes	Savanna	Treeless plains	Water	Wet rainforest	Total	UA	CE
Predicted type	Dry rainforest	758	0	2	0	0	0	0	0	0	45	805	0.94	0.06
	Grassland and sedgeland	1	546	41	7	0	6	22	5	0	14	642	0.85	0.15
	Mangrove	3	10	472	0	0	0	0	0	0	22	507	0.93	0.07
	Coastal salt marsh	0	4	1	285	13	1	0	0	3	0	307	0.93	0.07
	Sandy beach	0	0	0	7	375	106	0	0	0	0	488	0.77	0.23
	Sand dunes	0	0	0	1	39	319	0	4	0	0	363	0.88	0.12
	Savanna	7	60	9	0	0	16	845	40	0	8	985	0.86	0.14
	Treeless plains	0	49	0	0	0	83	137	165	0	0	434	0.38	0.62
	Water	0	1	0	0	1	0	0	0	411	0	413	1.00	0.00
	Wet rainforest	373	34	173	0	0	0	19	0	0	344	943	0.36	0.64
	Total	1142	704	698	300	428	531	1023	214	414	433	5887		
	PA	0.66	0.78	0.68	0.95	0.88	0.60	0.83	0.77	0.99	0.79			
	OE	0.34	0.22	0.32	0.05	0.12	0.40	0.17	0.23	0.01	0.21			

Model formulation

Classification scheme: Ecosystem Functional Group (level 3 of the Global Ecosystem typology)
Satellite/sensor: Landsat-9/OLI-2
Covariate set: Satellite image covariates only

PA: Producer’s accuracy
UA: User’s accuracy
OE: Omission error
CE: Commission error

Table 9. Confusion matrix for the ecosystem functional group classification model using Landsat-9 satellite imagery from the OLI-2 sensor and using the satellite image and additional covariates.

		Training points											
		Dry rainforest	Grassland and sedgeland	Mangrove	Coastal salt marsh	Sandy beach	Sand dunes	Savanna	Treeless plains	Water	Wet rainforest	Total	UA	CE
Predicted type	Dry rainforest	909	1	6	0	0	0	0	0	0	25	941	0.97	0.03
	Grassland and sedgeland	0	605	39	2	0	12	13	10	0	6	687	0.88	0.12
	Mangrove	46	50	642	0	0	1	1	0	0	8	748	0.86	0.14
	Coastal salt marsh	0	6	1	292	11	0	0	0	0	0	310	0.94	0.06
	Sandy beach	0	0	0	5	395	20	0	0	0	0	420	0.94	0.06
	Sand dunes	0	0	0	1	21	443	0	0	0	0	465	0.95	0.05
	Savanna	4	22	5	0	0	5	905	8	0	8	957	0.95	0.05
	Treeless plains	0	19	0	0	0	50	85	196	0	0	350	0.56	0.44
	Water	0	0	0	0	1	0	0	0	414	0	415	1	0
	Wet rainforest	183	1	5	0	0	0	19	0	0	386	594	0.65	0.35
	Total	1142	704	698	300	428	531	1023	214	414	433	5887		
	PA	0.8	0.86	0.92	0.97	0.92	0.83	0.88	0.92	1	0.89			
	OE	0.2	0.14	0.08	0.03	0.08	0.17	0.12	0.08	0	0.11			

Model formulation

Classification scheme: Ecosystem Functional Group (level 3 of the Global Ecosystem typology)
Satellite/sensor: Landsat-9/OLI-2
Covariate set: Satellite image and additional covariates

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Table 10. Confusion matrix for the ecosystem functional group classification model using Sentinel-2 satellite imagery from the MSI sensor as the only covariates.

		Training points												
		Dry rainforest	Grassland and sedgeland	Mangrove	Coastal salt marsh	Sandy beach	Sand dunes	Savanna	Treeless plains	Water	Wet rainforest	Total	UA	CE
Predicted type	Dry rainforest	727	2	17	0	0	2	0	0	0	90	838	0.87	0.13
	Grassland and sedgeland	0	419	14	3	0	5	18	8	0	2	469	0.89	0.11
	Mangrove	55	60	522	0	0	1	0	0	0	69	707	0.74	0.26
	Coastal salt marsh	0	3	1	285	5	1	0	0	2	0	297	0.96	0.04
	Sandy beach	0	0	0	7	356	77	0	0	0	0	440	0.81	0.19
	Sand dunes	0	0	0	2	61	286	0	7	0	0	356	0.80	0.20
	Savanna	2	50	15	0	0	13	754	29	0	20	883	0.85	0.15
	Treeless plains	0	111	3	0	2	145	220	170	0	1	652	0.26	0.74
	Water	0	1	0	3	4	0	0	0	412	0	420	0.98	0.02
	Wet rainforest	358	58	126	0	0	1	31	0	0	251	825	0.30	0.70
	Total	1142	704	698	300	428	531	1023	214	414	433	5887		
	PA	0.64	0.60	0.75	0.95	0.83	0.54	0.74	0.79	1.00	0.58			
	OE	0.36	0.40	0.25	0.05	0.17	0.46	0.26	0.21	0.00	0.42			

Model formulation

Classification scheme: Ecosystem Functional Group (level 3 of the Global Ecosystem typology)
Satellite/sensor: Sentinel-2/MSI
Covariate set: Satellite image covariates only

PA: Producer’s accuracy
UA: User’s accuracy
OE: Omission error
CE: Commission error

Table 11. Confusion matrix for the ecosystem functional group classification model using Sentinel-2 satellite imagery from the MSI sensor and using satellite image and additional covariates.

Predicted type	Dry rainforest	Grassland and sedgeland	Mangrove	Coastal salt marsh	Sandy beach	Sand dunes	Savanna	Treeless plains	Water	Wet rainforest	Total	UA	CE
	Training points												
Dry rainforest	951	1	11	0	0	1	2	0	0	31	997	0.95	0.05
Grassland and sedgeland	1	532	28	3	0	20	18	9	0	2	613	0.87	0.13
Mangrove	57	106	648	0	0	3	1	0	0	7	822	0.79	0.21
Coastal salt marsh	0	6	1	291	6	4	0	0	0	0	308	0.94	0.06
Sandy beach	0	0	0	5	391	34	0	0	0	0	430	0.91	0.09
Sand dunes	0	0	0	1	30	402	2	3	0	0	438	0.92	0.08
Savanna	1	37	5	0	0	7	900	9	0	21	980	0.92	0.08
Treeless plains	0	20	0	0	0	60	86	193	0	0	359	0.54	0.46
Water	0	0	0	0	1	0	0	0	414	0	415	1.00	0.00
Wet rainforest	132	2	5	0	0	0	14	0	0	372	525	0.71	0.29
Total	1142	704	698	300	428	531	1023	214	414	433	5887		
PA	0.83	0.76	0.93	0.97	0.91	0.76	0.88	0.90	1.00	0.86			
OE	0.17	0.24	0.07	0.03	0.09	0.24	0.12	0.10	0.00	0.14			

Model formulation

Classification scheme: Ecosystem Functional Group (level 3 of the Global Ecosystem typology)
Satellite/sensor: Sentinel-2/MSI
Covariate set: Satellite image and additional covariates

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Table 12. Confusion matrix for the biome classification model using Landsat-9 satellite imagery from the OLI-2 sensor as the only covariates.

		Training points										
		Grassland and sedgeland	Mangrove	Rainforest	Sandy beach	Sand dune	Savanna	Treeless plains	Water	Total	UA	CE
Predicted type	Grassland and sedgeland	554	48	16	0	3	22	5	0	648	0.85	0.15
	Mangrove	38	867	130	9	0	3	0	2	1049	0.83	0.17
	Rainforest	10	60	1407	0	0	15	0	0	1492	0.94	0.06
	Sandy beach	0	11	0	375	100	0	0	0	486	0.77	0.23
	Sand dunes	1	2	0	42	319	0	4	0	368	0.87	0.13
	Savanna	49	10	22	0	15	848	39	0	983	0.86	0.14
	Treeless plains	51	0	0	0	94	135	166	0	446	0.37	0.63
	Water	1	0	0	2	0	0	0	412	415	0.99	0.01
	Total	704	998	1575	428	531	1023	214	414	5887		
	PA	0.79	0.87	0.89	0.88	0.60	0.83	0.78	1.00			
	OE	0.21	0.13	0.11	0.12	0.40	0.17	0.22	0.00			

Model formulation

Classification scheme: Biome (level 2 of the Global Ecosystem typology)
Satellite/sensor: Landsat-9/OLI-2
Covariate set: Satellite image covariates only

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Table 13. Confusion matrix for the biome classification model using Landsat-9 satellite imagery from the OLI-2 sensor and using satellite image and additional covariates.

Predicted type	Training points:	Grassland and sedgeland	Mangrove	Rainforest	Sandy beach	Sand dune	Savanna	Treeless plains	Water	Total	UA	CE
	Grassland and sedgeland	616	60	17	0	15	14	8	0	730	0.84	0.16
	Mangrove	43	907	31	6	9	1	0	0	997	0.91	0.09
	Rainforest	1	16	1443	0	2	13	0	0	1475	0.98	0.02
	Sandy beach	0	8	0	402	20	0	0	0	430	0.93	0.07
	Sand dunes	0	2	1	19	429	0	0	0	451	0.95	0.05
	Savanna	21	5	83	0	5	911	8	0	1033	0.88	0.12
	Treeless plains	23	0	0	0	51	84	198	0	356	0.56	0.44
	Water	0	0	0	1	0	0	0	414	415	1.00	0.00
	Total	704	998	1575	428	531	1023	214	414	5887		
	PA	0.88	0.91	0.92	0.94	0.81	0.89	0.93	1.00			
	OE	0.12	0.09	0.08	0.06	0.19	0.11	0.07	0.00			

Model formulation

Classification scheme: Biome (level 2 of the Global Ecosystem typology)
Satellite/sensor: Landsat-9/OLI-2
Covariate set: Satellite image and additional covariates

PA: Producer’s accuracy
UA: User’s accuracy
OE: Omission error
CE: Commission error

Table 14. Confusion matrix for the biome classification model using Sentinel-2 satellite imagery from the MSI sensor as the only covariates.

		Training points										
		Grassland and sedgeland	Mangrove	Rainforest	Sandy beach	Sand dune	Savanna	Treeless plains	Water	Total	UA	CE
Predicted type	Grassland and sedgeland	449	58	24	0	6	20	9	0	566	0.79	0.21
	Mangrove	64	809	143	4	2	0	0	0	1022	0.79	0.21
	Rainforest	23	92	1358	0	3	5	0	0	1481	0.92	0.08
	Sandy beach	0	8	0	357	74	0	0	4	443	0.81	0.19
	Sand dunes	0	2	0	60	290	0	6	0	358	0.81	0.19
	Savanna	59	22	50	0	14	786	30	0	961	0.82	0.18
	Treeless plains	107	4	0	3	142	212	169	0	637	0.27	0.73
	Water	2	3	0	4	0	0	0	410	419	0.98	0.02
	Total	704	998	1575	428	531	1023	214	414	5887		
	PA	0.64	0.81	0.86	0.83	0.55	0.77	0.79	0.99			
	OE	0.36	0.19	0.14	0.17	0.45	0.23	0.21	0.01			

Model formulation

Classification scheme: Biome (level 2 of the Global Ecosystem typology)
Satellite/sensor: Sentinel-2/MSI
Covariate set: Satellite image covariates only

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Table 15. Confusion matrix for the biome classification model using Sentinel-2 satellite imagery from the MSI sensor and using satellite image and additional covariates.

		Training points										
		Grassland and sedgeland	Mangrove	Rainforest	Sandy beach	Sand dune	Savanna	Treeless plains	Water	Total	UA	CE
Predicted type	Grassland and sedgeland	580	80	20	1	32	14	9	0	736	0.79	0.21
	Mangrove	65	880	33	4	19	0	0	0	1001	0.88	0.12
	Rainforest	0	27	1482	0	1	8	0	0	1518	0.98	0.02
	Sandy beach	0	7	0	396	33	0	0	0	436	0.91	0.09
	Sand dunes	1	1	0	26	378	1	1	0	408	0.93	0.07
	Savanna	36	3	40	0	6	911	8	0	1004	0.91	0.09
	Treeless plains	22	0	0	0	62	89	196	0	369	0.53	0.47
	Water	0	0	0	1	0	0	0	414	415	1.00	0.00
	Total	704	998	1575	428	531	1023	214	414	5887		
	PA	0.82	0.88	0.94	0.93	0.71	0.89	0.92	1.00			
	OE	0.18	0.12	0.06	0.07	0.29	0.11	0.08	0.00			

Model formulation

Classification scheme: Biome (level 2 of the Global Ecosystem typology)
Satellite/sensor: Sentinel-2/MSI
Covariate set: Satellite image and additional covariates

PA: Producer's accuracy
UA: User's accuracy
OE: Omission error
CE: Commission error

Additional references

Aybar, C., Wu, Q., Bautista, L., Yali, R. and Barja, A. (2020) 'rgee: An R package for interacting with Google Earth Engine', *Journal of Open Source Software*, 5(51), p. 2272. Available at: https://doi.org/10.21105/joss.02272.

van den Brand, T. (2024) 'ggh4x: Hacks for "ggplot2"'. Available at: https://CRAN.R-project.org/package=ggh4x.

Campitelli, E. (2023) 'ggnewscale: Multiple Fill and Colour Scales in "ggplot2"'. Available at: https://CRAN.R-project.org/package=ggnewscale.

Dunnington, D. (2023) 'ggspatial: Spatial Data Framework for ggplot2'. Available at: https://CRAN.R-project.org/package=ggspatial.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. (2017) 'Google Earth Engine: Planetary-scale geospatial analysis for everyone', *Remote Sensing of Environment*, 202, pp. 18–27. Available at: https://doi.org/10.1016/j.rse.2017.06.031.

Greenwell, B.M. and Boehmke, B.C. (2020) 'Variable Importance Plots—An Introduction to the vip Package', *The R Journal*, 12(1), pp. 343–366. Available at: https://doi.org/10.32614/RJ-2020-013.

Henry, L., Wickham, H. and Chang, W. (2024) ‘ggstance: Horizontal “ggplot2” Components’. Available at: https://CRAN.R-project.org/package=ggstance.

Hernangomez, D. (2024) ‘tidyterra: tidyverse Methods and ggplot2 Helpers for terra Objects’. Available at: https://dieghernan.github.io/tidyterra/.

Hijmans, R.J. (2023) ‘terra: Spatial Data Analysis’. Available at: https://CRAN.R-project.org/package=terra.

Kuhn, M. (2008) ‘Building Predictive Models in R Using the caret Package’, *Journal of Statistical Software*, 28(5), pp. 1–26. Available at: https://doi.org/10.18637/jss.v028.i05.

Pebesma, E. (2018) ‘Simple Features for R: Standardized Support for Spatial Vector Data’, *The R Journal*, 10(1), pp. 439–446. Available at: https://doi.org/10.32614/RJ-2018-009.

R Core Team (2018) ‘R: A language and environment for statistical computing’. Vienna, Austria: R Foundation for Statistical Computing.

RStudio Team (2020) ‘RStudio: Integrated Development for R’. Boston, MA: RStudio, PBC. Available at: http://www.rstudio.com.

Smith, A., Murphy, S., Herderson, D. and Erickson, K. (2023) ‘Including imprecisely georeferenced specimens improves accuracy of species distribution models and estimates of niche breadth’, *Global Ecology & Biogeography*, 32(3), pp. 342–355. Available at: https://doi.org/doi:10.1111/geb.13628.

Wickham, H. (2016) ‘ggplot2: Elegant Graphics for Data Analysis.’ New York: Springer-Verlag.

Wickham, H. (2022) ‘stringr: Simple, Consistent Wrappers for Common String Operations’. Available at: https://CRAN.R-project.org/package=stringr.

Wickham, H., François, R., Harry, L., Müller, K. and Vaughan, D. (2023) ‘dplyr: A Grammar of Data Manipulation’. Available at: https://CRAN.R-project.org/package=dplyr.

Wickham, H., Vaughan, D. and Girlich, M. (2023) ‘Tidyr: tidy messy data’. Available at: https://CRAN.R-project.org/package=tidyr.

Wright, M.N. and Ziegler, A. (2017) ‘Ranger: a fast implementation of random forests for high dimensional data in C++ and R.’, *Journal of Statistical Software*, 77(1), pp. 1–17. Available at: https://doi.org/doi:10.18637/jss.v077.i01.