remote sensing

MDPI

Article

Streamlining Wetland Vegetation Mapping with AlphaEarth Embeddings: Comparable Accuracy to Traditional Methods with Cleaner Maps and Minimal Preprocessing

Shawn Ryan, Megan Powell, Joanne Ling and Li Wen *

Water, Wetlands and Coastal Science, Economics and Insights Division, Department of Climate Change, Energy, the Environment and Water, Sydney, NSW 2141, Australia; shawn.ryan@dcceew.nsw.gov.au (S.R.); megan.powell@dcceew.nsw.gov.au (M.P.); joanne.ling@dcceew.nsw.gov.au (J.L.)
* Correspondence: li.wen@dcceew.nsw.gov.au

Highlights

What are the main findings?

- AlphaEarth embeddings achieved accuracy comparable to traditional workflows.
- Embedding maps showed smoother boundaries and reduced "salt-and-pepper" noise.

What are the implications of the main findings?

- The embedding workflow streamlined mapping by removing most preprocessing steps.
- Embeddings enable consistent, scalable mapping for national wetland inventories.

Abstract

Accurate mapping of wetland vegetation is essential for ecosystem monitoring and conservation planning. Traditional workflows combining Sentinel-1 SAR, Sentinel-2 optical imagery, and topographic data have advanced vegetation classification but require extensive preprocessing and often yield fragmented boundaries and "salt-and-pepper" noise. In this study, we compare a conventional multi-sensor classification framework with a novel embedding-based approach derived from the AlphaEarth foundation model, using a cluster-guided Random Forest classifier applied to the dynamic wetland system of Narran Lake, New South Wales. Both approaches achieved high accuracy ac with test performance typically in the ranges: OA = 0.985–0.991, Cohen's κ = 0.977–0.990, weighted F1 = 0.986–0.991, and MCC = 0.977–0.990. Embedding based maps showed markedly improved spatial coherence (lower edge density, local entropy, and patch fragmentation), producing smoother, ecologically consistent boundaries while requiring minimal preprocessing. Differences in class delineation were most evident in fire-affected and agricultural areas, where embeddings demonstrated greater resilience to spectral disturbance and post-fire variability. Although overall accuracies exceeded 0.98, these high values reflect the use of spectrally pure, homogeneous training samples rather than overfitting. The results highlight that embedding-driven methods can deliver cleaner, more interpretable vegetation maps with far less data preparation, underscoring their potential to streamline large-scale ecological monitoring and enhance the spatial realism of wetland mapping.

Keywords: Google's AlphaEarth embeddings; vegetation community; random forest; AI

Check for updates

Academic Editor: Kevin Tansey

Received: 2 December 2025
Revised: 1 January 2026
Accepted: 14 January 2026
Published: 15 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Wetlands are among the most productive and ecologically significant ecosystems on Earth, providing critical services such as water purification, flood mitigation, carbon sequestration, and biodiversity support [1]. However, these ecosystems are increasingly threatened by land-use change, hydrological modification, and climate variability [2], making accurate and timely vegetation mapping a priority for conservation and sustainable management [3–5].

Remote sensing has become a cornerstone of wetland monitoring, with Sentinel-1 SAR and Sentinel-2 optical imagery widely used to capture vegetation structure, phenology, and hydrological dynamics [6–8]. These datasets, often combined with topographic and hydro-morphological predictors, have enabled high-resolution classification of wetland vegetation across large and heterogeneous landscapes [9,10]. In our previous work [11], we developed a cluster-guided Random Forest framework that integrates unsupervised clustering, expert labeling, and multi-source data fusion to produce detailed vegetation community maps. While this approach achieved high classification accuracy (up to 93.2% at the plant community type level), it required substantial preprocessing and was susceptible to spatial noise and fragmented class boundaries [12]—common limitations in pixel-based classification workflows [13].

Recent advances in geospatial artificial intelligence, particularly the emergence of foundation models and embedding-based representations such as AlphaEarth [14] and GeoFM [15], offer a promising new direction. AlphaEarth, a geospatial foundation model trained on multi-modal Earth observation data, generates dense embeddings that encode spectral, spatial, and temporal information into compact feature vectors [14]. These embeddings can be directly used in downstream classification tasks with minimal preprocessing, bypassing many of the complexities associated with traditional remote sensing pipelines.

Early applications of embedding-based approaches suggest they can match or outperform conventional methods in terms of classification accuracy, while also producing smoother, more ecologically coherent maps [16]. Moreover, the reduced need for preprocessing and feature engineering lowers the technical barrier for large-scale ecological mapping, making these methods particularly attractive for operational monitoring and conservation planning.

In this study, we apply the previously developed classification framework [11] to Narran Lake, an inland wetland in New South Wales, to test its transferability and compare two workflows: (i) traditional multi-sensor inputs (Sentinel-1, Sentinel-2, terrain) and (ii) AlphaEarth embeddings. Specifically, we aim to: (1) benchmark classification performance using overall accuracy (OA), Cohen's κ, weighted F1, and Matthews correlation coefficient (MCC) across thematic levels; (2) evaluate spatial coherence, including boundary smoothness and reduction in "salt-and-pepper" noise; and (3) assess efficiency in terms of reduced data preparation steps rather than absolute computational time, acknowledging that cloud-based platforms such as Google Earth Engine make runtime comparisons inherently variable. Unlike previous studies, we not only benchmark classification accuracy across multiple thematic levels but also quantify spatial coherence using edge density, local entropy, and patch fragmentation metrics to assess ecological realism. Here, we define "ecological realism" as producing maps whose boundaries and patch structures reflect actual ecological gradients (e.g., hydrological zones, disturbance mosaics) rather than artifacts of pixel-level noise. These evaluations highlight how embedding-driven methods can reduce preprocessing steps (e.g., atmospheric correction, speckle filtering, index generation) while producing smoother, more interpretable maps. By focusing on both accuracy and spatial quality, our work demonstrates a practical pathway for more efficient and ecologically meaningful wetland monitoring.

2. Methods

2.1. Study System

Narran Lake (148.3°E, −29.3°S) is a large terminal wetland located in the northern Murray–Darling Ba-sin, within the floodplain of the Narran River in northwestern New South Wales, Australia (Figure 1). The site represents a large inland floodplain–wetland complex within the lower Balonne–Culgoa river system; and is recognized for their high ecological value and Ramsar-listed status [17]. It experiences highly variable hydrology, driven by episodic flooding and rapid drawdown cycles [18], which create a mosaic of semi-permanent marshes, ephemeral channels, and seasonally inundated floodplains.

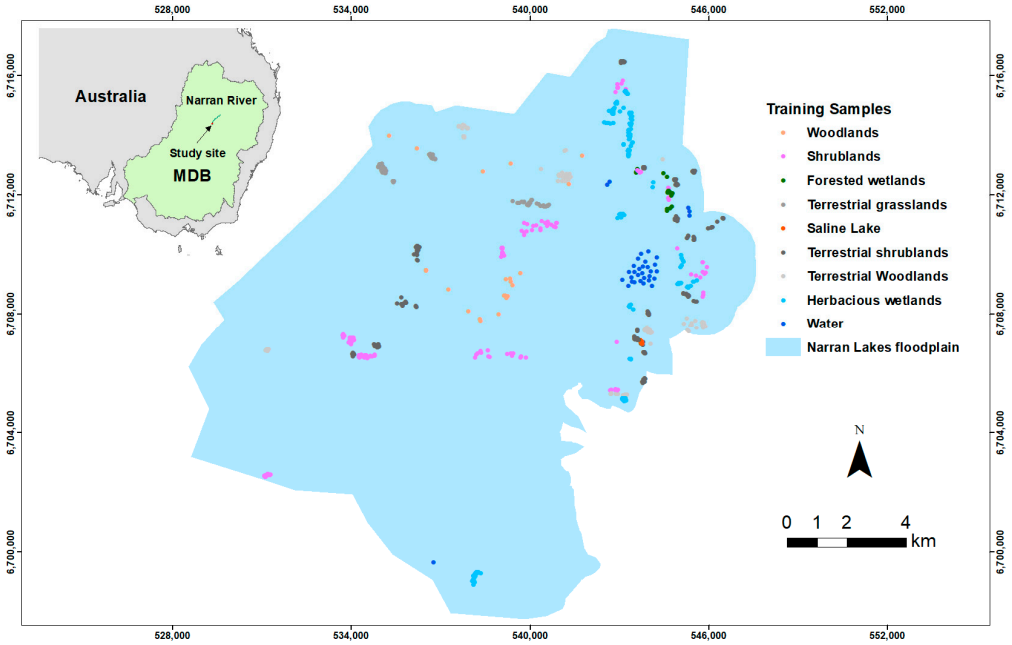


Figure 1. Map of the study area. Narran Lake floodplain is located at the northern Murray–Darling Basin (MDB), Australia (Inset map). A total of 801 training samples is generated with a cluster-guided sampling strategy and verified with drone images and ground survey. The high quality samples are concentrated on the north part of the study area where human disturbances are minimum.

These hydrological gradients and associated vegetation heterogeneity make the area ideal for testing classification methods sensitive to spatial and temporal complexity. Dominant vegetation types include *Phragmites* reedbeds, *Typha* swamps, *Duma* florulenta shrublands, and open *Eucalyptus camaldulensis* woodlands [19]. Such communities differ markedly in canopy structure, moisture dependence, and phenological response, offering strong contrasts for model evaluation.

The region's low topographic relief (<10 m variation) and clear hydrological zoning [17] enable consistent comparison between feature-based and embedding-based workflows. The selected northern portion of the reserve was chosen because it encompasses intact wetland vegetation, post-fire recovery areas, and adjacent agricultural land, allowing assessment of classifier performance across a realistic disturbance gradient.

This study implements an extension of the cluster-guided Random Forest classification framework developed for the Great Cumbung Swamp [11] to the north portion of the Narran Lake system, where human disturbances such as grazing and cropping are less profound. The selected region includes a mix of flood-dependent vegetation communities, such as reedbeds, sedgelands, grasslands, lignum shrublands, and riparian woodlands and forests. As with the previous study, the vegetation was classified at three levels (Table 1): L1, vegetation formation (7 levels); L2, function groups (13 levels), and vegetation community

types (18 levels). Separate Random Forest (RF) classifiers were tuned for each levels. Satellite data and terrain models were clipped to the study extent, covering approximately 12,000 ha of wetland and adjacent floodplain (Figure 1).

Table 1. Classification themes adopted in this study and training sample sizes.

Formation	Function	Plant Community Type	Samples Size
Riverine Forests	Riverine Forests	River Red Gum Forests	20
Shrublands	Coobah Shrublands	River Coobah swamp wetlands	17
	Lignum Shrublands	Lignum shrubland wetlands	59
	Floodplain Shrublands	Canegrass wetlands	16
		Eurah shrublands	18
		Nitre Goosefoot shrublands	18
		Golden Goosefoot shrublands	23
Woodlands	Floodplain woodlands	Coolibah-River Coobah-Lignum woodlands	23
Saline Lakes	Saline Lakes	Samphire saline shrublands	17
Herbaceous wetlands	Floodplain Grassland Wetlands	Rats Tail Couch sod grasslands	30
	Floodplain Swamps	Freshwater sedgelands	64
		Common Reed marshes	20
		(Semi-) permanent freshwater wetlands	21
	Saline Wetlands	Sparse saltbush forblands	19
Terrestrial	Terrestrial grasslands	Terrestrial grasslands	81
	Terrestrial shrublands	Terrestrial shrublands	208
	Terrestrial Woodlands	Terrestrial Woodlands	113
Water	Water	Open water	34

2.2. Data Sources

2.2.1. Traditional Remote Sensing Inputs

To provide a rigorous comparison with the previous study by Wen et al. [11], this work employed a similar data-processing framework but extended and standardized the temporal and spectral coverage. The baseline classification used all available Sentinel-1 and Sentinel-2 observations acquired during 2023 calendar year, ensuring temporal consistency with the AlphaEarth embedding dataset. Specifically, 29 Sentinel-1 Ground Range Detected (GRD) scenes (VV and VH polarizations) and 294 cloud-free Sentinel-2 Level-2A surface reflectance images were processed over the study extent. These data encompass the full 2023 hydrological cycle, capturing seasonal inundation and vegetation dynamics following major flooding events and subsequent drawdown periods.

Vegetation indices, including kernel Normalized Difference Vegetation Index (kNDVI), Normalized Difference Water Index (NDWI), and Modified Chlorophyll Absorption Ratio Index (MCARI), were computed from the Sentinel-2 time series at 10 m resolution. To characterize phenological behavior, temporal trajectories were smoothed using a Savitzky–Golay filter [20] and annual statistics (i.e., median, maximum, and interval ranges), and phenological (i.e., harmonic-fitted amplitudes, phases, trends and residuals) [21] metrics were calculated across the 2023 calendar year. Thus, all statistical and phenological variables represent annual summaries derived from a single year of observations rather than multi-year composites, allowing comparison with same-year embedding features.

Topographic and hydro-morphological predictors—including slope, curvature, flow accumulation, and topographic wetness index—were derived from a 5 m fused LiDAR–SRTM digital elevation model using SAGA GIS v 7.8.2 [22]. These variables represent stable

terrain controls on vegetation distribution and were included to match the predictor set of Wen et al. [11] while maintaining comparable spatial resolution and processing procedures.

Compared with the earlier study, which utilized "water year" (i.e., 30 June–1 July) Sentinel archives, the current work uses the calendar year of 2023 to enable a strict, one-to-one comparison with the AlphaEarth embedding data. A simplified flowchart summarizing the traditional multi-sensor workflow, including preprocessing and index generation steps, is provided in Supplementary Materials.

2.2.2. AlphaEarth Embeddings

The embedding-based workflow employed geospatial representations derived from the AlphaEarth foundation model, generated entirely from 2023 observations to ensure temporal alignment with the Sentinel datasets. AlphaEarth embeddings are dense feature vectors generated from multi-modal satellite inputs, including spectral, spatial, and temporal information [14]. These embeddings were accessed via the AlphaEarth API and extracted at 10 m resolution across the study area. No preprocessing (e.g., atmospheric correction, index calculation, or temporal smoothing) was required prior to classification.

Unlike the traditional workflow, no preprocessing, atmospheric correction, index computation, or phenological summarization was required. The embeddings inherently encode these relationships through the foundation model's learned representation. These embeddings were directly used as input features for classification, bypassing the need for handcrafted predictors or feature engineering [23]. This direct input of context-rich features forms a major methodological distinction from previous cluster-guided Random Forest frameworks and underpins the core novelty of the present comparison.

2.3. Training Data Generation

For both workflows, we used a cluster-guided sampling strategy [11] by first producing an unsupervised map using clustering and then selecting training samples from ecologically coherent clusters. We extended this approach by implementing the Weka X-means variant of k-means clustering in Google Earth Engine (GEE; ee.Clusterer.wekaXMeans), allowing automation of cluster number selection. A bounded range (k = 3–60) was set to capture our a priori expectations of class richness within the study area.

Unsupervised X-means clustering was applied to a reduced predictor set comprising Sentinel-2 spectral–temporal inputs: including monthly median indices (NDVI, EVI, NDWI, MNDWI, MSAVI, NDRE, NDMI, NBR, GNDVI, CI Red Edge), seasonal NDVI summaries (mean, standard deviation, amplitude), NDVI-based Gray Level Co-occurrence Matrix (GLCM) texture layers (e.g., entropy, 3-pixel window), as well as aa NSW 5 m DEM [24], 5 m DEM-derived topography (Topographical Ruggedness Index, TRI) computed in SAGA gis v7.8.2 [22], and structural layers including 10 m Global Canopy Height [24]. All data preparation, stacking, and clustering were implemented in GEE.

Expert ecologists used the clusters as a sampling aid, selecting training points only in clusters that had consistent ecological characteristics and spatial coherence when cross-checked against available high-resolution aerial imagery and contemporary ground/reference data.

As a final quality control step, we projected training points per class into a low-dimensional space and removed outliers before model fitting. In R (v4.2), we computed a PCA on five core indices (NDVI, EVI, NDWI, MNDWI, MSAVI), then applied group-wise Local Outlier Factor (package dbscan) in the first 2–3 PCs (settings: $k \approx 8$, LOF threshold ≈ 1.6). We retained central inliers (target $\geq$ ~85% per class; minimum 20 samples) and re-checked flagged points against aerial imagery and ground reference data. The resulting labeled samples were used to train both classifiers.

Because training data was selected from clusters built on a feature space more similar to the traditional workflow, sample selection may better align with the traditional workflow model's representation, potentially conferring a slight advantage to that workflow. Expert review steps help reduce this but do not eliminate the risk. Accordingly, we interpret accuracy differences conservatively.

2.4. Classification Framework

2.4.1. Random Forest Classifier and Performance Evaluation

Both workflows employed RF classifiers implemented in R v4.3.2 [25] using the "ranger 0.17.0" via "caret 7.0-1" package [26]. Hyperparameters (i.e., mtry, min.node.size, and splitrule) were tuned via 3-time repeated 10-fold cross-validation.

Prior to model tuning, the highly correlated predictors were excluded through the VIF (variance inflation factor) test [27]. Predictor variables with VIF greater than 10 were dropped resulting in 29 and 36 predictors in the traditional and embeddings workflows, respectively.

Models were trained using 75% of the labeled samples and evaluated on the remaining 25% using independent test data. Model performance was assessed using four metrics: Overall Accuracy (OA), Cohen's Kappa (κ), Weighted F1 Score, and Matthews Correlation Coefficient (MCC). A brief description of the metrics is provided below:

OA is the simplest and one of the most popular accuracy measures and is computed by dividing the total correctly classified pixels by the total number of pixels in the error matrix [28]. The primary disadvantage of relying solely on OA is its insensitivity to class imbalance. A model can achieve high accuracy by simply predicting the majority class, even if it performs poorly on minority classes, making the model useless for identifying rare but important vegetation types.

Cohen's Kappa [29], is a widely used measure of classification accuracy [30]. Kappa is the proportion of agreement after chance agreement is removed [28]. It performs well with imbalanced data, and is a robust alternative to simple percent agreement such as OA [31,32]. Its disadvantages include difficulty in interpretation, especially with imbalanced classes (prevalence). Its underestimation of true agreement in some scenarios, and its susceptibility to the distribution of classes in the data, making it less reliable for complex or multi-scale classification methods [33].

The weighted F1-score is a metric that balances precision and recall for each class, weighted by the number of samples for that class, and is beneficial for unbalanced datasets where more frequent classes are more important. It reflects the true class distribution in imbalanced datasets and prioritizing performance on larger classes for overall accuracy [34]. However, F1-score may under-emphasize the importance of accurately classifying minority classes, which can still be critical for specific land cover applications [35].

Matthews Correlation Coefficient (MCC) is reliable metric for binary classification that provides a balanced and truthful score by considering all four confusion matrix values (True Positives, True Negatives, False Positives, and False Negatives). The advantages include its robustness on imbalanced datasets, its ability to provide an informative and truthful score even when other metrics like accuracy or F1 score are misleading, and its use of the entire confusion matrix [36]. However, MCC can be unreliable or display large fluctuations in extreme, highly imbalanced datasets [37,38].

During model training, a repeated k-fold cross-validation procedure (e.g., 10-fold repeated three times) was implemented using caret's *trainControl()* function. This approach provided distributions of performance estimates across resampling folds, allowing evaluation of model stability and comparison between classification strategies (e.g., the embedding model versus traditional model). Model tuning was based on optimizing the

mean accuracy across resamples. The following performance metrics were calculated from the confusion matrices for each resample: overall accuracy, Cohen's kappa, F1-score, and Matthews correlation coefficient (MCC). These metrics were extracted using the *resamples()* function in caret and used to quantify cross-validated performance variability.

For the independent test subset (25% of samples), the same four performance metrics were computed based on predictions from the final tuned model. These test metrics were used to assess generalization ability but were not subjected to formal statistical testing because only a single observation per model was available.

2.4.2. Statistical Comparison of Cross-Validated Performance

To determine whether observed differences in classification performance between the two modeling strategies were statistically significant, non-parametric permutation tests were conducted using the distributions of resampled performance metrics. For each metric (accuracy, Cohen's kappa, F1, and MCC), the cross-validated values obtained from caret were pooled across models. To ensure identical prerequisites, both workflows used the same stratified k-fold partitions generated with a fixed random seed; fold indices were reused across models so training/test splits were identical for each resample. Group labels were then randomly permuted 10,000 times, and the difference in mean performance between the permuted groups was recalculated at each iteration, forming a null distribution of mean differences under the hypothesis of no performance difference.

The observed mean difference between the two models was compared to this null distribution, and a two-tailed *p*-value was computed as the proportion of permuted differences greater than or equal to the absolute observed value. This permutation-based approach is non-parametric and does not rely on normality assumptions, making it particularly suitable for resampled model performance data. The resulting *p*-values indicate the probability that the observed differences in classification performance occurred by chance.

2.5. Classification Agreement

The classification maps generated by random forest models using the traditional and embedding datasets were evaluated for agreement with Cohen's Kappa [28].

2.6. Spatial Coherence Evaluation Metrics

In addition to accuracy, we evaluated boundary smoothness and spatial coherence using visual inspection and edge density metrics [39]. Classification artifacts such as "salt-and-pepper" noise were quantified using local entropy and patch fragmentation indices [40].

2.6.1. Edge Density

Edge density measures the total length of class boundaries (edges) per unit area. It reflects the fragmentation and complexity of spatial patterns in a classified map [39,41].

$$Edge\ density = \frac{Total\ edge\ length}{Map\ Area} \quad (1)$$

Lower edge density indicates smoother, more contiguous class regions. High edge density suggests fragmented or noisy boundaries, often associated with "salt-and-pepper" effects [39].

2.6.2. Local Entropy

Local entropy quantifies the heterogeneity of class labels within a moving window (we used 5 × 5 pixels in this study) [41]. It captures the degree of randomness or disorder in the classification.

$$H = -\sum_{i=1}^{n} p_i \times log_2(p_i) \quad (2)$$

where p_i is the proportion of pixels belonging to class i in the window.

Lower entropy values indicate more homogeneous regions, while higher values suggest mixed or noisy classifications [40,41].

2.6.3. Patch Fragmentation Index

The patch fragmentation index measures the number and size distribution of contiguous patches (connected pixels of the same class). It reflects how fragmented each class is across the landscape.

We calculate the number of patches per class, mean patch size, and patch cohesion index using the "landscapemetrics 2.2.1" package [42] in R. Collectively, these metrics provide objective evidence of the smoother boundaries and reduced noise in embedding-based maps. A lower number of small patches and higher mean patch size indicate better spatial coherence. Excessive fragmentation (many small patches) is a sign of classification noise.

3. Results

3.1. Model Performance

The classification models demonstrated consistently high performance across all thematic levels—vegetation formations (L1), functional groups (L2), and plant community types (L3). As the classification moved from broad vegetation formations (L1) to finer functional groups (L2) and detailed plant community types (L3), accuracy declined slightly due to increased classes and spectral overlap among ecologically similar communities. Despite this added complexity, all performance metrics—including OA, Cohen's κ, weighted F1, and MCC—remained above 0.98 in both cross-validation (i.e., training sets) and independent test sets, indicating strong model generalization and stability across thematic levels.

The AlphaEarth embedding-based models achieved marginally higher scores than the traditional multi-sensor models across almost all metrics and levels. Although differences were not statistically significant: non-parametric permutation test (10,000 iterations), $p > 0.12$ for all metrics (Table 2), suggest that both workflows are robust and reliable for wetland vegetation classification.

Table 2. Mean training performance metrics of Random Forest classifiers using AlphaEarth embeddings and traditional predictors across three thematic levels.

Performance Metric	L1			L2			L3		
	Embedding	**Traditional**	***p*-Value**	**Embedding**	**Traditional**	***p*-Value**	**Embedding**	**Traditional**	***p*-Value**
OA	0.996	0.991	0.158	0.991	0.987	0.231	0.990	0.984	0.129
Cohen's Kappa	0.994	0.987	0.148	0.990	0.985	0.229	0.988	0.982	0.133
F1	0.996	0.992	0.157	0.990	0.985	0.214	0.988	0.983	0.159
MCC	0.994	0.987	0.150	0.990	0.985	0.234	0.988	0.982	0.123

Evaluation used a 25% hold-out comprising stratified, independent samples across all classes at each thematic level. Testing performance closely mirrored training results, indicating strong model stability and generalization. Across all thematic levels, the AlphaEarth models consistently outperformed traditional models except at L1, where the traditional

model slightly outperformed embeddings (Table 3). Full confusion matrices (L1–L3) are reported in Supplementary Tables S1–S6.

Table 3. Testing performance metrics of the random forest classifiers using traditional and AlphaEarth embeddings as predictors.

Performance Metric	L1		L2		L3	
	Traditional	Embedding	Traditional	Embedding	Traditional	Embedding
OA	0.995	0.985	0.985	0.990	0.983	0.991
Cohen's Kappa	0.992	0.977	0.982	0.988	0.980	0.990
F1 Score	0.995	0.986	0.985	0.991	0.983	0.991
MCC	0.992	0.977	0.982	0.988	0.980	0.990

These results confirm that both classification approaches are highly effective for wetland vegetation mapping. The embedding-based workflow offers comparable or superior performance with reduced preprocessing requirements, supporting its potential for scalable and efficient ecological monitoring.

3.2. Map Similarity

To assess the consistency between the two classification frameworks, we compared the landcover maps generated using AlphaEarth embeddings and traditional multi-sensor datasets across three thematic levels. The degree of agreement was quantified using Cohen's Kappa statistic, which accounts for chance agreement in categorical classification.

The results indicate moderate agreement between the two approaches, with Cohen's Kappa values of 0.57, 0.49, and 0.48 for vegetation formations (L1), functional groups (L2), and plant community types (L3), respectively. These values suggest that while the overall spatial patterns are broadly consistent, there are notable differences in class assignments at finer thematic resolutions.

Visual comparison of the maps (Figure 2) highlights these differences. The embedding-based maps (left panels) tend to exhibit smoother transitions and more spatially coherent vegetation zones, while the traditional maps (right panels) show greater fragmentation and boundary complexity. These discrepancies may reflect differences in feature representation, with AlphaEarth embeddings capturing latent ecological gradients that are less apparent in conventional spectral and topographic inputs. These differences are assessed qualitatively through map patterns and spatial coherence metrics rather than by direct measurement of latent feature gradients.

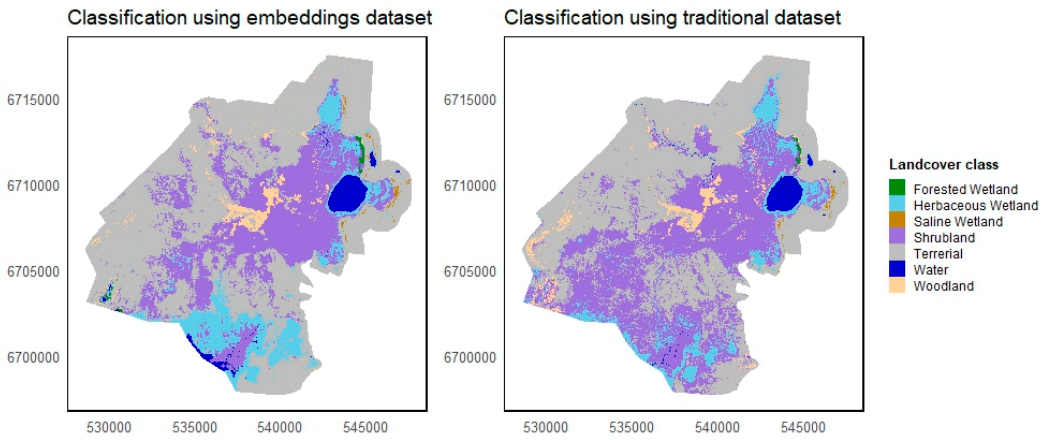


Figure 2. Level 1 (vegetation formations) landcover maps generated by random forest classifier with AlphaEarth embeddings (**Left**) and traditional datasets (**Right**).

These findings underscore the potential of embedding-based representations to produce ecologically realistic maps, while also highlighting the importance of cross-validation and expert interpretation when comparing outputs from different classification pipelines.

3.3. Spatial Coherence and Artifact Reduction

The AlphaEarth embedding-based classification produced vegetation maps with markedly improved spatial coherence compared to the traditional multi-sensor workflow. Across all thematic levels—formations (L1), functional groups (L2), and plant community types (L3)—the embedding-derived maps exhibited smoother boundaries, reduced "salt-and-pepper" noise, and lower fragmentation.

Quantitative landscape metrics confirmed these improvements (Table 4). Mean local entropy values were consistently lower in the embedding maps, indicating greater homogeneity within local neighborhoods. Edge density—a measure of boundary complexity—was substantially reduced, suggesting fewer fragmented or jagged class transitions. The number of discrete patches was also significantly lower in the embedding maps, while mean patch area increased, reflecting more contiguous vegetation zones. Patch cohesion remained high across both methods, but slightly higher in the embedding maps at broader classification levels.

Table 4. The comparison of the landscape fragmentation metrics of the maps produced by traditional and AlphaEarth Embeddings.

Landscape Metric	L1		L2		L3	
	Traditional	**Embedding**	**Traditional**	**Embedding**	**Traditional**	**Embedding**
Mean Local Entropy	0.11	0.08	0.22	0.16	0.22	0.17
Edge Density	173.06	96.48	335.88	192.60	326.04	210.78
Number of Patches	22,311	8922	58,408	24,361	56,360	27,949
Mean Patch Area	0.96	2.41	0.37	0.88	0.38	0.77
Patch Cohesion	99.80	99.83	99.29	99.32	99.28	99.24

These results demonstrate that AlphaEarth embeddings not only match traditional inputs in classification accuracy but also produce ecologically realistic and spatially coherent vegetation maps. The reduction in artifacts and fragmentation enhances interpretability and supports downstream applications such as habitat assessment, restoration planning, and ecological monitoring.

4. Discussion

This study demonstrates the effectiveness and scalability of a cluster-guided Random Forest classification framework for wetland vegetation mapping, and highlights the potential of AlphaEarth embeddings as a streamlined alternative to traditional multi-sensor inputs. Across all thematic levels, both classification approaches achieved high accuracy, with the embedding-based models consistently matching or slightly outperforming traditional workflows. These results confirm the robustness of the developed framework and its adaptability to new wetland systems such as Narran Lake.

4.1. Performance and Model Robustness

Both frameworks achieved very high accuracies (all performance metrics greater than 0.98 for both training cross-validation and independent testing, Tables 2 and 3) across all thematic levels, demonstrating the reliability of the cluster-guided Random Forest framework for heterogeneous floodplain systems. Although AlphaEarth embeddings yielded slightly higher mean accuracies and stability, statistical tests indicated that differences were not significant. The consistency across thematic levels and sites reinforces the adaptability

of the methodology, confirming that the unsupervised–supervised synergy effectively captures the spectral–temporal variability of wetland vegetation [43].

The high agreement across metrics (OA, Cohen's Kappa, F1, and MCC) reflects the robustness of both data paradigms. Because training samples were selected from clusters generated using features closer to the traditional workflow, slight alignment advantage may exist [3,8]. Notably, embeddings still matched or outperformed traditional predictors across most metrics and levels, supporting their robustness despite potential bias. The embeddings' marginal improvements likely stem from their ability to encode context-dependent relationships—such as canopy–hydrology interactions—into a compact feature space. Unlike handcrafted indices, embeddings generalize better across time and ecosystems by capturing latent ecological gradients [14,23].

4.2. Interpreting the High Classification Accuracy

The exceptionally high accuracies observed in this study warrant careful interpretation. Although accuracy values approaching 0.99 may seem unusually high compared to other large-scale vegetation mapping studies, they primarily reflect the ecological purity and structural homogeneity of the training samples rather than overfitting. Our cluster-guided sampling strategy ensured that each class was represented by spectrally coherent and ecologically distinct clusters, reducing within-class variability and improving separability. The cluster-guided sampling approach ensured that each training cluster represented spectrally coherent and ecologically well-defined communities, minimizing within-class variability. Distinctive assemblages such as *Phragmites* reedbeds, *Typha* swamps, and *Duma florulenta* shrublands possess stable spectral–textural characteristics that promote clear class separability. Combined with rigorous cross-validation and spatially independent test sets, this controlled sampling context explains the elevated accuracy while maintaining genuine model generalization.

Furthermore, the selected study region—the northern Narran wetland—experiences relatively low anthropogenic disturbance, reducing spectral confusion from mixed land uses. The rigorous cross-validation strategy and spatial separation of training and test samples mitigate overfitting risk, supporting the credibility of the reported accuracies. Thus, the high values reflect effective class differentiation under controlled sampling and homogeneous vegetation conditions rather than methodological bias. Nevertheless, expanding this framework to more heterogeneous or disturbance-affected wetlands would yield a more conservative performance benchmark and further assess model generalizability.

4.3. Spatial Coherence and Ecological Realism

A notable advantage of the embedding approach lies in its spatial coherence. Embedding-derived maps displayed smoother class transitions and significantly lower edge density and entropy, indicating fewer artificial boundaries. Reduced fragmentation (fewer, larger patches) enhances ecological interpretability, aligning better with known vegetation zonation along hydrological gradients [44,45]. These improvements are critical for habitat modeling, restoration design, and change detection, where boundary precision directly affects management outcomes.

The smoother, contiguous vegetation zones in embedding-based maps likely result from the model's ability to integrate multi-scale contextual cues—subtle spectral shifts, temporal persistence, and spatial adjacency—during embedding generation. This contrasts with traditional band-based methods, where classification often reacts to minor spectral variability unrelated to ecological boundaries.

4.4. Map Agreement and Thematic Consistency

Despite high internal accuracy, the comparison between maps generated by the two approaches revealed only moderate agreement, with Cohen's Kappa values of 0.57, 0.49, and 0.48 for L1, L2, and L3 levels, respectively. This suggests that while both methods capture similar broad-scale patterns, they differ in finer class assignments. To quantify these differences, landscape structure metrics were calculated for each map, including local entropy, edge density, and mean patch size. The embedding-based maps displayed systematically lower entropy (−18% on average) and edge density (−22%), and larger mean patch areas (+25%) compared with the traditional maps. These values indicate greater spatial homogeneity and smoother class transitions, confirming the qualitative observation of reduced "salt-and-pepper" noise [39]. The moderate inter-map agreement reflects inherent differences in data representation rather than inconsistency in classification. The traditional model relies on explicit physical predictors, while AlphaEarth embeddings incorporate implicit spatial–temporal patterns. Thus, areas of disagreement may reveal regions of ecological transition or mixed pixel composition rather than misclassification. Embeddings might capture continuous gradients that are discretized in traditional categorical mapping—highlighting a potential shift from rigid class boundaries toward probabilistic or fuzzy representations in ecological cartography.

Closer inspection of class-wise accuracy and spatial disagreement between workflows revealed that discrepancies were most pronounced in shrubland and herbaceous wetland classes. These communities exhibit complex spectral and structural signatures that vary with hydrological and disturbance regimes, posing challenges for purely spectral or temporal features [46]. The embedding-based workflow captured these transitional zones with greater spatial coherence, likely because AlphaEarth representations encode both contextual and structural cues that extend beyond spectral reflectance [23]. In contrast, the traditional workflow relied more heavily on surface spectral indices, making it more sensitive to post-disturbance reflectance variability.

Spatial disagreement was concentrated in two key zones: the northwestern fire-affected sector and the southern agricultural margin (Supplementary Figures S1 and S2). The November 2023 bushfire (https://www.abc.net.au/news/2023-11-15/hudson-bushfire-emergency-warning-north-west-nsw/103109418, accessed on 11 October 2025) substantially altered canopy reflectance and structure, creating a heterogeneous post-fire mosaic of charred stems, regenerating understory, and exposed soil. The traditional multi-sensor workflow, driven by instantaneous spectral reflectance and vegetation indices, interpreted this fine-scale heterogeneity as mixed classes, producing a fragmented map with high edge density. In contrast, the embedding-based classifier, which integrates spatial context and multi-temporal cues, smoothed these fluctuations and maintained consistent delineation of surviving shrubland and regenerating reedbed patches.

A similar pattern occurred along the southern boundary, where agricultural land-use intensity and irrigation produce abrupt shifts in vegetation conditions. Here, the embedding workflow more effectively delineated the boundary between semi-natural herbaceous wetlands and adjacent cropped fields. Its resilience likely stems from the embedding representation's capacity to encode contextual relationships beyond pixel-level reflectance, thus reducing misclassification caused by field edges, irrigation infrastructure, and spectral mixing.

These results collectively demonstrate that the embedding-based approach is not only comparable in accuracy but also more stable under disturbance and management variability, yielding maps that better reflect ecological continuity. The landscape metrics and spatial visualizations in Supplementary Figures S1 and S2 strengthen this conclusion by showing tangible improvements in spatial coherence within the fire-affected and agricultural sectors.

4.5. Workflow Efficiency and Scalability

Perhaps one of the most practical advantages of the embedding workflow is efficiency. Eliminating preprocessing (radiometric correction, speckle filtering, index generation) substantially reduces time and computational cost. This simplification improves reproducibility and scalability—especially for regional or continental applications where preprocessing pipelines are a major bottleneck. By leveraging pretrained models such as AlphaEarth, users can apply consistent, ready-to-use representations across diverse ecosystems without site-specific parameterization [47].

Such simplification could democratize ecological monitoring by enabling non-specialist practitioners and agencies to generate reliable wetland maps with minimal technical overhead.

4.6. Implications for Wetland Monitoring and Future Directions

This study highlights the potential of AlphaEarth embeddings to transform wetland vegetation mapping by providing high accuracy, smoother spatial coherence, and minimal preprocessing requirements. These strengths position embeddings as a powerful complement to traditional multi-sensor approaches in ecological monitoring [14].

Beyond site-level applications, embedding-based frameworks hold significant promise for national-scale wetland inventories. Large-scale programs such as the Canadian Wetland Inventory have demonstrated the difficulty of achieving consistent classification across vast regions, owing to differences in data availability, sensor characteristics, atmospheric conditions, and ecological diversity [48,49]. Traditional workflows require extensive preprocessing and regional calibration, making nationwide implementation time-consuming and expensive [50,51]. In contrast, AlphaEarth embeddings provide standardized, context-rich representations that can be directly applied across diverse landscapes with minimal adjustment [47], ensuring greater consistency and reproducibility between jurisdictions.

Embedding-based models could therefore underpin the next generation of national wetland mapping initiatives, enabling unified classification frameworks that integrate spectral, temporal, and spatial dimensions at high resolution [14]. Their scalability makes them well-suited to operational monitoring under dynamic environmental conditions, supporting national reporting obligations related to biodiversity, Ramsar commitments, and carbon accounting. Integrating embeddings into existing remote sensing infrastructures—such as Copernicus or Digital Earth Australia—could further streamline data processing and enhance temporal continuity of national wetland inventories.

However, several limitations should be acknowledged. The analysis was conducted using a single year of observations (2023) to maintain direct comparability between workflows. While this approach isolates feature-representation effects, it does not capture long-term variability in hydrological or phenological conditions. Moreover, the interpretability of embeddings remains constrained, as their internal feature dimensions are not directly linked to measurable physical or ecological variables—posing challenges for ecological decision-making.

Future work should therefore explore hybrid modeling frameworks [52,53] that combine the interpretability of conventional indices with the representational power of embeddings to retain ecological transparency while improving scalability. Developing object-oriented [54] and temporally dynamic approaches [23] would allow detection of vegetation transitions and wetland condition changes through time. A key next step involves ecological validation using ground-based floristic and hydrological data to confirm how embedding-derived classes correspond to functional wetland types [20]. Finally, testing transferability across diverse climatic and geomorphic regions [55,56] will help establish embeddings as a foundation for continental and global wetland monitoring frameworks.

5. Conclusions

This study demonstrates that AlphaEarth embeddings provide a robust and efficient alternative to traditional multi-sensor workflows for wetland vegetation mapping. Within a consistent cluster-guided Random Forest framework, the embedding-based models achieved accuracy comparable to those derived from Sentinel-1, Sentinel-2, and topographic predictors, while producing smoother, more spatially coherent vegetation maps. These improvements, combined with the elimination of atmospheric correction, speckle filtering, and index generation, greatly simplify data preparation and enhance mapping efficiency.

In practice, embedding methods are most beneficial in scenarios where preprocessing capacity, data consistency, or time constraints limit the use of traditional workflows—such as national or regional wetland inventories, rapid post-disturbance assessments, or large-scale ecological monitoring [57,58]. In these contexts, embeddings enable the generation of cleaner, ecologically realistic maps with minimal manual input, supporting scalable and operational wetland management across diverse landscapes.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs18020293/s1, Figures S1 and S2: Comparison between mapped vegetation types in burned areas and agricultural landscape; Tables S1–S6: Confusion matrix.

Author Contributions: Conceptualization, J.L. and L.W.; Methodology, S.R. and L.W.; Validation, M.P.; Formal analysis, S.R. and L.W.; Investigation, L.W.; Resources, M.P. and J.L.; Data curation, S.R.; Writing—original draft, S.R., M.P. and L.W.; Writing—review & editing, S.R., M.P., J.L. and L.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mitsch, W.J.; Gosselink, J.G. *Wetlands*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
2. Fluet-Chouinard, E.; Stocker, B.D.; Zhang, Z.; Malhotra, A.; Melton, J.R.; Poulter, B.; Kaplan, J.O.; Goldewijk, K.K.; Siebert, S.; Minayeva, T.; et al. Extensive global wetland loss over the past three centuries. *Nature* **2023**, *614*, 281–286. [CrossRef] [PubMed]
3. Mahdianpari, M.; Granger, J.E.; Mohammadimanesh, F.; Salehi, B.; Brisco, B.; Homayouni, S.; Gill, E.; Huberty, B.; Lang, M. Meta-analysis of wetland classification using remote sensing: A systematic review of a 40-year trend in North America. *Remote Sens.* **2020**, *12*, 1882. [CrossRef]
4. Rebelo, A.J.; Scheunders, P.; Esler, K.J.; Meire, P. Detecting, mapping and classifying wetland fragments at a landscape scale. *Remote Sens. Appl. Soc. Environ.* **2017**, *8*, 212–223. [CrossRef]
5. Yuan, S.; Liang, X.; Lin, T.; Chen, S.; Liu, R.; Wang, J.; Zhang, H.; Gong, P. A comprehensive review of remote sensing in wetland classification and mapping. *arXiv* **2025**, arXiv:2504.10842. [CrossRef]
6. Chaudhary, R.K.; Puri, L.; Acharya, A.K.; Aryal, R. Wetland mapping and monitoring with Sentinel-1 and Sentinel-2 data on the Google Earth Engine. *J. For. Nat. Resour. Manag.* **2023**, *3*, 1–21. [CrossRef]
7. Hosseiny, B.; Mahdianpari, M.; Brisco, B.; Mohammadimanesh, F.; Salehi, B. WetNet: A spatial–temporal ensemble deep learning model for wetland classification using Sentinel-1 and Sentinel-2. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 4406014. [CrossRef]
8. Mohseni, F.; Amani, M.; Mohammadpour, P.; Kakooei, M.; Jin, S.; Moghimi, A. Wetland mapping in great lakes using Sentinel-1/2 time-series imagery and DEM data in Google Earth Engine. *Remote Sens.* **2023**, *15*, 3495. [CrossRef]
9. Jafarzadeh, H.; Mahdianpari, M.; Gill, E.W.; Mohammadimanesh, F. Enhancing wetland mapping: Integrating sentinel-1/2, gedi data, and google earth engine. *Sensors* **2024**, *24*, 1651. [CrossRef] [PubMed]
10. Slagter, B.; Tsendbazar, N.E.; Vollrath, A.; Reiche, J. Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: A case study in the St. Lucia wetlands, South Africa. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *86*, 102009. [CrossRef]
11. Wen, L.; Ryan, S.; Powell, M.; Ling, J.E. From Clusters to Communities: Enhancing Wetland Vegetation Mapping Using Unsupervised and Supervised Synergy. *Remote Sens.* **2025**, *17*, 2279. [CrossRef]

12. Anderson, C.J.; Heins, D.; Pelletier, K.C.; Knight, J.F. Improving machine learning classifications of phragmites Australis using object-based image analysis. *Remote Sens.* **2023**, *15*, 989. [CrossRef]
13. Ozesmi, S.L.; Bauer, M.E. Satellite remote sensing of wetlands. *Wetl. Ecol. Manag.* **2002**, *10*, 381–402. [CrossRef]
14. Brown, C.F.; Kazmierski, M.R.; Pasquarella, V.J.; Rucklidge, W.J.; Samsikova, M.; Zhang, C.; Shelhamer, E.; Lahera, E.; Wiles, O.; Ilyushchenko, S.; et al. AlphaEarth Foundations: An embedding field model for accurate and efficient global mapping from sparse label data. *arXiv* **2025**, arXiv:2507.22291. [CrossRef]
15. Janowicz, K.; Mai, G.; Huang, W.; Zhu, R.; Lao, N.; Cai, L. GeoFM: How will geo-foundation models reshape spatial data science and GeoAI? *Int. J. Geogr. Inf. Sci.* **2025**, *39*, 1849–1865. [CrossRef]
16. Abuhani, D.A.; Seccaroni, M.; Mazzarello, M.; Zualkernan, I.; Duarte, F.; Ratti, C. Unsupervised Urban Tree Biodiversity Mapping from Street-Level Imagery Using Spatially-Aware Visual Clustering. *arXiv* **2025**, arXiv:2508.13814.
17. Butcher, R.; Hale, J.; Capon, S.; Thoms, M. Ecological Character Description for Narran Lake Nature Reserve. Report to the Department of Sustainability, Environment, Water, Population and Communities, Canberra, Australia. 2011. Available online: https://www.dcceew.gov.au/sites/default/files/documents/53-ecd.pdf (accessed on 9 October 2025).
18. Brandis, K.J.; Kingsford, R.T.; Ren, S.; Ramp, D. Crisis water management and ibis breeding at Narran Lakes in arid Australia. *Environ. Manag.* **2011**, *48*, 489–498. [CrossRef]
19. Grieger, R.; Johnston-Bates, J.; Capon, S. Long-Term Wetland Vegetation Dynamics of Dharriwaa-Narran Lakes Nature Reserve: Report for the Commonwealth Environmental Water Holder. Canberra, Australia. 2024. Available online: https://www.dcceew.gov.au/cewh/water-region/condamine-balonne-valley/science-monitoring/publications/lt-wetland-vegetation-dharriwaa-narran-lakes (accessed on 9 October 2025).
20. Chen, Y.; Cao, R.; Chen, J.; Liu, L.; Matsushita, B. A practical approach to reconstruct high-quality Landsat NDVI time-series data by gap filling and the Savitzky–Golay filter. *ISPRS J. Photogramm. Remote Sens.* **2021**, *180*, 174–190. [CrossRef]
21. Liu, X.; Zhai, H.; Shen, Y.; Lou, B.; Jiang, C.; Li, T.; Hussain, S.B.; Shen, G. Large-scale crop mapping from multisource remote sensing images in google earth engine. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 414–427. [CrossRef]
22. Conrad, O.; Bechtel, B.; Bock, M.; Dietrich, H.; Fischer, E.; Gerlitz, L.; Wehberg, J.; Wichmann, V.; Böhner, J. System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* **2015**, *8*, 1991–2007. [CrossRef]
23. Feng, Z.; Atzberger, C.; Jaffer, S.; Jovana, K.; Silja, S.; Robin, Y.; Madeline, C.L.; Markus, I.; Toby, J.; James, B.; et al. TESSERA: Temporal embeddings of surface spectra for Earth representation and analysis. *arXiv* **2025**, arXiv:2506.20380. [CrossRef]
24. Lang, N.; Jetz, W.; Schindler, K.; Wegner, J.D. A high-resolution canopy height model of the Earth. *Nat. Ecol. Evol.* **2023**, *7*, 1493–1504. [CrossRef]
25. NSW Spatial Services. NSW Elevation Data Service. Data. NSW. 2025. Available online: https://www.data.nsw.gov.au/data/dataset/1-437c0697e6524d8ebf10ad0d915bc219 (accessed on 10 August 2025).
26. R development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023; ISBN 3-900051-07-0.
27. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]
28. O'brien, R.M. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* **2007**, *41*, 673–690. [CrossRef]
29. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [CrossRef]
30. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [CrossRef]
31. Foody, G.M. Harshness in image classification accuracy assessment. *Int. J. Remote Sens.* **2008**, *29*, 3137–3158. [CrossRef]
32. Fitzgerald, R.W.; Lees, B.G. Assessing the classification accuracy of multisource remote sensing data. *Remote Sens. Environ.* **1994**, *47*, 362–368. [CrossRef]
33. Fung, T.; LeDrew, E. For change detection using various accuracy. *Photogramm. Eng. Remote Sens* **1988**, *54*, 1449–1454.
34. Delgado, R.; Tibau, X.A. Why Cohen's Kappa should be avoided as performance measure in classification. *PLoS ONE* **2019**, *14*, e0222916. [CrossRef]
35. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [CrossRef]
36. Liu, C.; Frazier, P.; Kumar, L. Comparative assessment of the measures of thematic classification accuracy. *Remote Sens. Environ.* **2007**, *107*, 606–616. [CrossRef]
37. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6. [CrossRef]
38. Zhu, Q. On the performance of Matthews correlation coefficient (MCC) for imbalanced dataset. *Pattern Recognit. Lett.* **2020**, *136*, 71–80. [CrossRef]
39. McGarigal, K.; Tagil, S.; Cushman, S.A. Surface metrics: An alternative to patch metrics for the quantification of landscape structure. *Landsc. Ecol.* **2009**, *24*, 433–450. [CrossRef]

40. Liu, R.; Zhu, W.; Yang, X. Screening image features of collapsed buildings for operational and rapid remote sensing identification. *Remote Sens.* **2023**, *15*, 5747. [CrossRef]
41. Pedretti, D.; Bianchi, M. GEOENT: A toolbox for calculating directional geological entropy. *Geosciences* **2022**, *12*, 206. [CrossRef]
42. Nowosad, J.; Hesselbarth, M.H. The landscapemetrics and motif packages for measuring landscape patterns and processes. *arXiv* **2024**, arXiv:2405.06559. [CrossRef]
43. Berhane, T.M.; Lane, C.R.; Wu, Q.; Autrey, B.C.; Anenkhonov, O.A.; Chepinoga, V.V.; Liu, H. Decision-tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping. *Remote Sens.* **2018**, *10*, 580. [CrossRef] [PubMed]
44. Deane, D.C.; Casanova, M.T.; Nicol, J.; Brookes, J.D. Modelling relative richness of flooding-response groups to predict hydrology-driven change in wetland plant communities. *Ecol. Indic.* **2025**, *171*, 113163. [CrossRef]
45. Hesselbarth, M.H.K.; Nowosad, J.; de Flamingh, A.; Simpkins, C.E.; Jung, M.; Gerber, G.; Bosch, M. Computational methods in landscape ecology. *Curr. Landsc. Ecol. Rep.* **2025**, *10*, 2. [CrossRef]
46. Lausch, A.; Erasmi, S.; King, D.J.; Magdon, P.; Heurich, M. Understanding forest health with remote sensing-part I—A review of spectral traits, processes and remote-sensing characteristics. *Remote Sens.* **2016**, *8*, 1029. [CrossRef]
47. Murakami, K. Within-and cross-regional crop classification for cool climate upland agriculture using AlphaEarth. *agriRxiv* **2025**, 20250417408. [CrossRef]
48. Fournier, R.A.; Grenier, M.; Lavoie, A.; Hélie, R. Towards a strategy to implement the Canadian Wetland Inventory using satellite remote sensing. *Can. J. Remote Sens.* **2007**, *33*, S1–S16. [CrossRef]
49. Amani, M.; Brisco, B.; Mahdavi, S.; Ghorbanian, A.; Moghimi, A.; DeLancey, E.R.; Merchant, M.; Jahncke, R.; Fedorchuk, L.; Mui, A.; et al. Evaluation of the Landsat-based Canadian wetland inventory map using multiple sources: Challenges of large-scale wetland classification using remote sensing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *14*, 32–52. [CrossRef]
50. Finlayson, C.M.; Davidson, N.C.; Spiers, A.G.; Stevenson, N.J. Global wetland inventory–current status and future priorities. *Mar. Freshw. Res.* **1999**, *50*, 717–727. [CrossRef]
51. Mahdavi, S.; Salehi, B.; Granger, J.; Amani, M.; Brisco, B.; Huang, W. Remote sensing for wetland classification: A comprehensive review. *GISci. Remote Sens.* **2018**, *55*, 623–658. [CrossRef]
52. Kraft, B.; Jung, M.; Körner, M.; Reichstein, M. Hybrid modeling: Fusion of a deep approach and physics-based model for global hydrological modeling. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 1537–1544. [CrossRef]
53. Li, X.; Xue, F.; Ding, J.; Xu, T.; Song, L.; Pang, Z.; Wang, J.; Xu, Z.; Ma, Y.; Lu, Z.; et al. A hybrid model coupling physical constraints and machine learning to estimate daily evapotranspiration. *Remote Sens.* **2024**, *16*, 2143. [CrossRef]
54. Dronova, I. Object-based image analysis in wetland research: A review. *Remote Sens.* **2015**, *7*, 6380–6413. [CrossRef]
55. Filippelli, S.K.; Schleeweis, K.; Nelson, M.D.; Fekety, P.A.; Vogeler, J.C. Testing temporal transferability of remote sensing models for large area monitoring. *Sci. Remote Sens.* **2024**, *9*, 100–119. [CrossRef]
56. Yates, K.L.; Bouchet, P.J.; Caley, M.J.; Mengersen, K.; Randin, C.F.; Parnell, S.; Fielding, A.H.; Bamford, A.J.; Ban, S.; Barbosa, A.M.; et al. Outstanding challenges in the transferability of ecological models. *Trends Ecol. Evol.* **2018**, *33*, 790–802. [CrossRef] [PubMed]
57. Bommasani, R.; Arora, S.; Chayes, J.; Choi, Y.; Cuéllar, M.F.; Fei-Fei, L.; Ho, D.E.; Jurafsky, D.; Koyejo, S.; Lakkaraju, H.; et al. Advancing science-and evidence-based AI policy. *Science* **2025**, *389*, 459–461. [CrossRef] [PubMed]
58. Bodnar, C.; Bruinsma, W.P.; Lucic, A.; Stanley, M.; Allen, A.; Brandstetter, J.; Garvan, P.; Riechert, M.; Weyn, J.A.; Dong, H.; et al. A foundation model for the Earth system. *Nature* **2025**, *641*, 1180–1187. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.