

Article

Pilot-Site Land Cover Mapping Using an Externally-Guided Clustering Framework: A Case Study from Ontario, Canada

Sondos Omar ¹, Reza Shahidi ^{1,*}, Masoud Mahdianpari ^{2,*} and Fariba Mohammadimanesh ³¹ Department of Electrical and Computer Engineering, Memorial University, St. John's, NL A1C 5S7, Canada; somar@mun.ca² C-CORE, Ottawa, ON K2K 2E3, Canada³ National Research Council, Ottawa, ON K1A 0R6, Canada; fariba.mohammadimanesh@nrcan-rncan.gc.ca

* Correspondence: rshahidi@mun.ca (R.S.); masoud.mahdianpari@c-core.ca (M.M.)

Abstract

High-resolution land cover classification is critical for monitoring environmental change and managing natural resources. This study presents an unsupervised framework with externally guided feature prioritization that integrates Sentinel-1 synthetic aperture radar (SAR) and Sentinel-2 optical imagery at 10 m spatial resolution. A cloud-native export protocol in Google Earth Engine (GEE) enables the generation of consistent, cloud-free, and snow-free seasonal composites across Ontario, Canada. A comprehensive feature engineering pipeline combines spectral indices, radar backscatter metrics, terrain derivatives from digital elevation models (DEMs), and temporal statistics to create a rich multi-sensor input space. Dimensionality reduction is performed using Sparse Principal Component Analysis (SparsePCA) and mutual-information-based feature selection. Clustering is conducted using three complementary algorithms: centroid-based K-means, density-based Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), and reachability-based Ordering Points To Identify the Clustering Structure (OPTICS). Final land cover labels are assigned via a majority-voting ensemble, with prediction ties resolved deterministically using OPTICS. OPTICS is particularly effective for modeling heterogeneous landscapes due to its ability to detect clusters of varying density without requiring a global threshold. This study is designed as a pilot-site methodological demonstration using three representative 2 km × 2 km regions in Ontario, rather than a full provincial-scale land cover product. The resulting classification maps are validated against reference land cover data, demonstrating the effectiveness and potential scalability of the proposed external-label guided unsupervised mapping approach.

Keywords: external-label guided unsupervised classification; machine learning; Sentinel-1; Sentinel-2; feature engineering; remote sensing; earth observation; Google Earth Engine



Academic Editor: Alessandro Vitale

Received: 3 May 2026

Revised: 5 July 2026

Accepted: 9 July 2026

Published: 10 July 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Accurate land cover classification is essential for environmental monitoring, urban planning, and resource management. Satellite remote sensing provides a scalable and repeatable means of mapping land cover across large areas [1,2].

The Sentinel-1 and Sentinel-2 missions offer complementary radar and optical datasets at 10 m resolution, enabling frequent and reliable monitoring. Yet unsupervised classification remains challenging due to heterogeneous landscapes and noisy signals. Robust frameworks are needed to effectively fuse multiple data sources, reduce dimensionality, and apply clustering techniques to distinguish land cover types.

Land-cover mapping remains challenging in heterogeneous landscapes because spectral overlap, seasonal variability, cloud contamination, mixed pixels, and limited ground-reference data can reduce classification reliability. Supervised approaches often achieve high accuracy when high-quality training data are available, but such labels are expensive to collect and may not transfer well across regions. In this context, clustering-based approaches offer a label-efficient alternative, particularly when combined with multi-source features from optical, SAR, and terrain data. Sentinel-2 provides rich spectral information related to vegetation, water, and built-up surfaces, while Sentinel-1 SAR contributes structural and moisture-sensitive information that is less affected by cloud cover. DEM-derived variables further help describe topographic controls on land-cover patterns. This study contributes a pilot-site externally guided clustering framework that combines feature engineering, feature selection, sparse dimensionality reduction, and ensemble clustering for land-cover mapping in selected Ontario landscapes.

1.1. Related Work

Prior research highlights the use of multi-sensor data and unsupervised models for land cover classification. The SEN12MS dataset [3] enabled fusion of Sentinel-1/2 satellite data for deep learning. Ghamisi et al. [4] showed kernel-based models perform well even with limited labels.

SAR–optical fusion has proven valuable, as shown by Verma et al. [5] and Xu et al. [6], using clustering and ensemble methods. Alternative strategies include Self-Organizing Maps [7], tile-based weak supervision [8], and distributed fuzzy clustering [9].

Texture-based features and ensemble learning approaches have been shown to significantly enhance the accuracy of land cover classification. For example, Chamundeeswari et al. [10] and Abbas et al. [11] demonstrated the effectiveness of combining principal component analysis (PCA) with clustering algorithms such as K-means and ISODATA to capture spectral variability and reduce dimensionality. In recent years, the emergence of cloud-based platforms such as Google Earth Engine (GEE) has revolutionized large-scale remote sensing analysis by offering access to extensive satellite image archives and geospatial datasets. With its robust server-side processing capabilities and user-friendly APIs, GEE enables scalable and reproducible workflows for environmental monitoring, land cover mapping, and change detection [12]. Additionally, land cover mapping results may be applied to urban land governance, sustainable planning, and built environment management [13], highlighting the important and widespread applications of our work.

1.2. Contributions

We propose a scalable, externally label-guided unsupervised land cover classification framework for the Canadian province of Ontario that fuses multi-source Earth observation data, including Sentinel-1 SAR, Sentinel-2 optical imagery, and DEM-derived terrain features. The pipeline incorporates comprehensive feature engineering where Dynamic World labels are only used on a separate geographic area and acquisition period independent of the regions to be clustered. The pipeline encompasses spectral indices, radar backscatter metrics, and temporal statistics—followed by dimensionality reduction via SparsePCA and mutual information-based feature selection. Land cover structures are inferred using an ensemble of three clustering algorithms, K-means, HDBSCAN, and OPTICS, with the latter also serving as a deterministic fallback to resolve voting ties. This ensemble strategy leverages both partition-based and density-based perspectives, enhancing classification robustness, particularly for minority classes. Validated against Dynamic World reference labels, our approach demonstrates strong reproducibility and improved performance across

heterogeneous landscapes, enabling externally label-guided unsupervised land cover mapping at 10 m resolution.

2. Methodology

2.1. Data and Study Area

The study focuses on three representative 2 km × 2 km regions in Ontario, Canada: Region A (southeast, [75.58° W, 45.23° N]), Region B (central, [79.37° W, 46.09° N]), and Region C (southwest, [83.10° W, 42.26° N]). The locations of these three representative regions are each labeled with the letter corresponding to the region label (A, B, or C) in Figure 1. Each region was processed independently using Sentinel-1 GRD (VV, VH) and Sentinel-2 L2A imagery acquired during the peak growing season (May–September 2023). Cloud-free composites and terrain attributes (elevation, slope, aspect) were derived from the DEM at 10 m resolution in Google Earth Engine. Stratified sampling was then applied to export per-region feature samples for offline clustering. Because the analysis was limited to summer months, the *snow and ice* class was excluded. The *flooded vegetation* class, represented by only 76 samples (0.05% of the data), was merged with *water* due to its scarcity in selected pilot areas.



Figure 1. Map of Ontario showing locations of Regions A, B, and C.

The three pilot regions were selected to capture contrasting land-cover conditions within Ontario. Region A represents a mixed forest and rural landscape, Region B includes a more heterogeneous vegetation–infrastructure setting, and Region C captures a shoreline and built-up environment with water, bare ground, and urban features. These sites were selected to test the methodological behavior of the proposed workflow across different landscape structures. However, they should be interpreted as representative pilot sites rather than a complete ecological or geographic sample of Ontario.

Sentinel-1 GRD images were filtered by acquisition mode, polarization availability, and study period, and the calibrated VV and VH backscatter bands were used after standard GEE preprocessing. Sentinel-2 Level-2A imagery was filtered for the growing season, and cloud-contaminated pixels were removed using the scene classification layer and cloud-probability information where available. Monthly cloud-free composites were generated for May–September 2023 using median aggregation. Temporal statistics, including mean, standard deviation, minimum, maximum, and linear trend, were calculated for each spectral and radar variable across the monthly composites. DEM-derived elevation, slope, and aspect were resampled to 10 m to match the Sentinel grid. All features were standardized before dimensionality reduction and clustering. Stratified samples were then extracted from each pilot region to preserve representation of the available land-cover classes.

Although the clustering stage itself does not use semantic labels during model fitting, the feature-selection stage uses Dynamic World labels from independent geographic areas and acquisition periods. Therefore, the proposed workflow is not a purely unsupervised method; rather, it is an externally label-guided clustering framework. Additionally, the proposed algorithm only outputs cluster IDs and not semantic class labels. However, once cluster IDs have been determined, converting these to semantic class labels is much easier and less intensive than in a supervised framework where each patch has to be independently labeled with a semantic class. Once clustering has been performed, just one sample has to be mapped to a given semantic class label, and then all other samples with the same cluster ID will automatically mapped to the same semantic class label. As a result, we assume this conversion may easily be performed either by a human or by an automatic algorithm.

2.2. Feature Engineering

Optical bands, radar indices (VV, VH, RVI, NDRI), and vegetation/water indices (NDVI, EVI, MNDWI) were combined with DEM features. Monthly statistics (mean, std, min, max, trend) captured temporal patterns. Coordinates and month indicators enriched the spatiotemporal feature space. The resulting dataset included 200+ features. Descriptions of each of these features are included in the third column of Table 1.

Table 1. Descriptions of Features Used in Externally Label-Guided Unsupervised Classification Analysis.

Category	Feature Name	Description/Role
Sentinel-2 (Optical)	B1-B12 (excl. B10) [14]	Surface reflectance bands used in spectral indices and classification
	AOT [15]	Aerosol Optical Thickness—measures atmospheric clarity
	WVP [16]	Water Vapor—influences spectral reflectance
	SCL [17]	Scene Classification Layer—semantic land cover context from Sentinel-2
	NDVI [18], EVI [19], MSAVI [20]	Vegetation indices capturing greenness, canopy density, and soil adjustment
	NDBI [21]	Detects built-up urban areas
	MNDWI [22]	Enhances detection of water bodies
	NDMI [23]	Indicates vegetation/soil moisture levels
	Spectral statistics	Mean, std, min, max, and trend across 5 months for each band
Sentinel-1 (SAR)	VV, VH	Backscatter coefficients in vertical and cross polarizations
	Angle	Radar incidence angle per acquisition
	VV/VH ratio	Differentiates surface structure types
	RVI [24], NDRI	Radar-based indices for vegetation and structural dynamics
	VV_grad, VH_grad	Monthly backscatter gradients to detect surface changes
	SAR statistics	Mean, std, min, max, and trend for SAR bands over time

Table 1. *Cont.*

Category	Feature Name	Description/Role
DEM/Terrain	Elevation	Terrain height from SRTM DEM
	Slope	Measures terrain steepness
	Aspect	Direction of slope—relates to solar exposure
	Slope–aspect interaction	Combines slope and aspect for solar modeling
	DEM trend	Static trend (if multiple temporal DEMs)
Coords & Metadata	Latitude, Longitude	Geospatial reference for each observation
	X, Y, Z	Cartesian transformation of geographic coordinates
	Month	Month of observation—encodes seasonal variation
	Landcover_class	Reference class from Dynamic World for evaluation only

This multi-modal, temporally enriched feature stack forms the input to the clustering algorithms described in the next section. A structured summary of all extracted features is presented in Table 1.

2.3. Clustering Methodologies

Mutual information (MI) scores were used to evaluate the relevance of each feature with respect to Dynamic World labels, capturing both linear and non-linear dependencies. These scores were calculated on a separate geographical area and acquisition period, rather than on the clustering evaluation regions themselves. The top 32 features were selected to enhance class separability and minimize redundancy. This maintains the externally label-guided unsupervised aspect of our algorithm since we are only using the Dynamic World labels for filtering the features and not for supervised prediction. To further reduce dimensionality and promote sparsity, SparsePCA was applied. Clustering was then performed using three complementary algorithms: K-means with *k*-means++ initialization [25–27], HDBSCAN for noise-tolerant hierarchical clustering [28,29], and OPTICS for reachability-based density estimation [30]. Each algorithm was tuned to balance accuracy and computational efficiency; the optimal experimental hyperparameters are shown in Table 2. The values of the *min_samples* parameter for the OPTICS algorithm and the *min_cluster_size* parameter for the HDBSCAN algorithm were set to the lowest possible value of 2 in order to ensure that smaller clusters were not missed. Although this led to a larger initial number of clusters, these were subsequently automatically merged to achieve the final number of clusters.

Table 2. Configuration Parameters of Clustering Algorithms.

Algorithm	Parameters
K-means	<i>n_clusters</i> = 200, <i>init</i> = <i>k</i> -means++
OPTICS	<i>min_samples</i> = 2, <i>xi</i> = 0.0001, <i>cluster_method</i> = <i>xi</i> , <i>metric</i> = cosine
HDBSCAN	<i>min_cluster_size</i> = 2, <i>cluster_selection_method</i> = eom

K-means is a partition-based algorithm that iteratively assigns data points to the nearest centroid and updates cluster centers to minimize intra-cluster variance. While computationally efficient, it assumes spherical clusters of comparable size, making it less effective for unbalanced or non-convex structures. To mitigate sensitivity to initialization,

the k -means++ strategy [26] is adopted, which disperses the initial centroids probabilistically to reduce convergence to poor local minima.

HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) extends DBSCAN by constructing a hierarchy of density-based clusters from a mutual reachability graph. Stable clusters are then extracted according to their persistence [28]. Unlike DBSCAN, HDBSCAN does not rely on a global distance threshold, allowing it to adapt to varying densities while robustly identifying noise and outliers [29].

OPTICS (Ordering Points To Identify the Clustering Structure) generalizes DBSCAN by computing a reachability distance for each point rather than applying a fixed density threshold. This enables the discovery of clusters with varying densities and arbitrary shapes, which can be visualized through its reachability plot [30]. OPTICS is particularly useful for detecting subtle boundaries and is employed as a deterministic fallback in the ensemble when ties arise among models.

To consolidate results across algorithms, the clusters were first merged using Euclidean distance, reducing the overall solution to 2000 superclusters. This step harmonizes the varying granularities of the individual methods and ensures that structurally similar clusters are grouped together. Following merging, an ensemble procedure was applied in which each supercluster received votes from the individual algorithms. The final label of each supercluster was determined by majority voting, such that the most frequently predicted class was assigned. This strategy leverages complementary algorithmic strengths while enhancing robustness against biases inherent to any single clustering method.

3. Results

Final labels were assigned via majority voting. If predictions were tied, OPTICS output served as the tie-breaker. Only resolved labels were used in evaluation.

We used several metrics to quantify classification performance, including the confusion matrix, precision, recall, F1-score, overall accuracy (OA), macro-averaged and weighted-averaged scores, as well as the Adjusted Rand Index (ARI) and Normalized Mutual Information (NMI). Reference labels for validation were derived via mode reduction over 5 months. Table 3 lists the number of available samples from Dynamic World for each class.

Table 3. Dynamic World Land Cover Classes (0–7) and Number of Samples Used in Evaluation.

Class ID	Class Label	Sample Count
W	Water	17,418
T	Trees	85,102
G	Grass	10,730
F	Flooded Vegetation	76
C	Crops	13,852
S	Shrub and Scrub	521
Bu	Built Area	38,559
Ba	Bare Ground	2044

Because Dynamic World is itself an algorithm-derived land-cover product, it should not be interpreted as independent ground truth. Therefore, the reported OA, F1-score, ARI, and NMI values represent agreement with Dynamic World labels rather than absolute classification accuracy. Independent validation using manually interpreted high-resolution imagery, field observations, or authoritative land-cover datasets is required to fully assess external validity.

Quantitative and qualitative evaluations underscore the effectiveness of the proposed ensemble-based externally label-guided unsupervised classification framework in delin-

eating land cover across various study sites in Ontario. Table 4 summarizes the clustering performance of the three individual models—K-means, HDBSCAN, and OPTICS—as well as the ensemble strategy. The ensemble model achieved an overall accuracy (OA) of 93.1%, with an Adjusted Rand Index (ARI) of 0.967 and a Normalized Mutual Information (NMI) of 0.922, generally outperforming each individual algorithm, except for overall accuracy, where OPTICS performed slightly better due to slightly improved performance on majority classes. While both HDBSCAN and OPTICS performed competitively (ARI of ~ 0.965 and 0.966 , respectively, and NMI of ~ 0.941 respectively), K-means lagged notably with an ARI of 0.801 and NMI of 0.670. These results highlight the benefits of density-based methods in modeling heterogeneous and non-globular land cover distributions.

To evaluate per-class performance, Table 5 presents the classification report comparing ensemble predictions against Dynamic World reference labels. Dominant land cover classes such as *Trees*, *Water*, and *Built Area* achieved precision, recall, and F1-scores consistently above 0.90, underscoring the robustness of the ensemble approach in mapping major landscape features. In contrast, minority classes exhibited lower performance due to limited representation in the dataset. *Shrub* and *Scrub*, which represents 0.3% of the data, achieved a moderate F1-score of 0.475, which remains noteworthy given its rarity. These findings highlight the ensemble model’s ability to retain sensitivity to ecologically significant but underrepresented categories. The confusion matrices for each individual algorithm and for the ensemble (Figure 2) further reinforce these results, illustrating the ensemble’s superior capacity to resolve the boundaries between classes and to reduce misclassification between spectrally similar classes such as *Crops* and *Grass*.

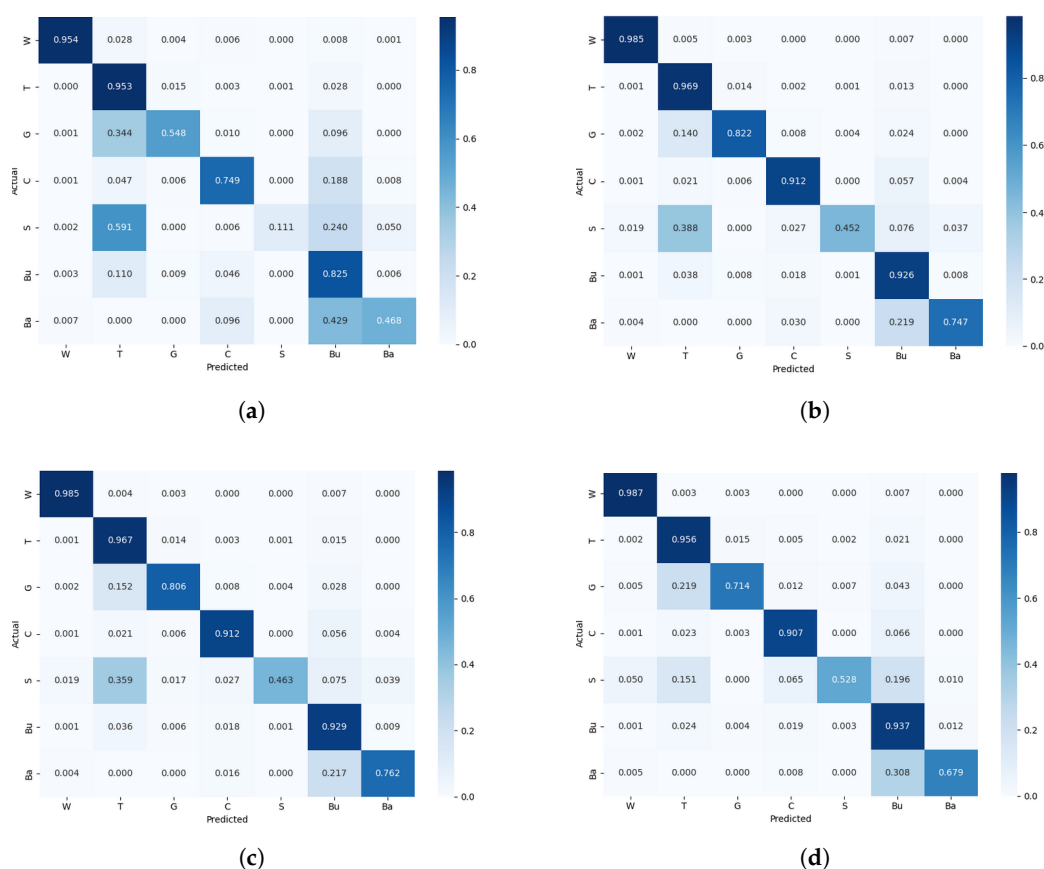


Figure 2. Confusion matrices for externally label-guided unsupervised clustering results using (a) K-means, (b) HDBSCAN, (c) OPTICS, and (d) the proposed ensemble method.

Table 4. Clustering Performance Metrics.

Model	ARI	NMI	OA
HDBSCAN	0.965	0.918	0.941
K-means	0.801	0.670	0.873
OPTICS	0.966	0.921	0.941
Ensemble	0.967	0.922	0.931

Table 5. Ensemble Classification Report Compared to Dynamic World Labels.

Class	Precision	Recall	F1-Score
W	0.987	0.969	0.978
T	0.956	0.956	0.956
G	0.714	0.834	0.769
C	0.907	0.904	0.905
S	0.528	0.432	0.475
Bu	0.937	0.900	0.918
Ba	0.679	0.748	0.712
Accuracy			0.931
Macro Avg	0.815	0.820	0.816
Weighted Avg	0.931	0.928	0.932

4. Discussion

This section presents a detailed evaluation of externally label-guided unsupervised land cover classification results across three ecologically diverse regions in Ontario using K-means, HDBSCAN, OPTICS, and a voting-based ensemble approach. The assessment integrates quantitative metrics, including Adjusted Rand Index (ARI), Normalized Mutual Information (NMI), and class-wise performance measures (precision, recall, F1-score), alongside qualitative spatial comparisons with Dynamic World reference labels.

4.1. Performance of K-Means Clustering

Across all three regions, K-means clustering demonstrated acceptable overall accuracy (87.3%), primarily driven by high performance on the dominant *Trees* class (class 1), which constitutes the majority of the dataset (85,102 samples). For instance, over all three regions, the weighted F1-score was 0.867. However, the macro-averaged F1-scores remained significantly lower (e.g., 0.694 overall), indicating substantial underperformance on minority classes. Notably, K-means performed poorly on *Shrub and Scrub* (class 5) and *Bare Ground* (class 7). Its limitations are further reflected in the overall ARI (0.801) and NMI (0.670) scores, suggesting weak agreement with the reference labels, especially in regions with spatially complex or underrepresented land cover types.

In all three regions, K-means exhibited particularly strong performance on the majority class (class 2, *Trees*), yielding an F1-score of 0.924. Classes 1 (*Water*) and 6 (*Built Area*) also demonstrated reasonable accuracies, with F1-scores of 0.924 and 0.820, respectively. However, performance dropped notably for the minority classes. Class 5 (*Shrub and Scrub*), in particular, had a very low recall of 0.111 and an F1-score of only 0.174, indicating a significant misclassification rate. Similarly, class 7 (*Bare Ground*) was only moderately identified, with an F1-score of 0.566. The macro-averaged F1-score of 0.694 highlights a disparity in model performance across classes, whereas the higher weighted average of 0.867 reflects the model's bias toward well-represented classes. These results indicate that while K-means performs reliably for dominant and moderately distributed classes in Region C, its performance deteriorates for rare and spectrally overlapping land cover types.

4.2. Effectiveness of Density-Based Clustering (HDBSCAN and OPTICS)

Both HDBSCAN and OPTICS exhibited significantly higher classification quality across all evaluation dimensions. The overall performance of HDBSCAN yielded an ARI of 0.965 and NMI of 0.918, while OPTICS achieved an ARI of 0.966 and NMI of 0.921. These values reflect strong agreement with Dynamic World labels. In Region A, both algorithms achieved $\text{ARI} \approx 0.951$ and $\text{NMI} \approx 0.89$, indicating their ability to recover complex spatial structures and small land cover patches.

Critically, these models maintained higher F1-scores across minority classes than K-means. For instance, the *Shrub and Scrub* class, often missed by K-means, was detected with an F1-score exceeding 0.64 over the three regions, much higher than the F1-score from K-Means for the same class (only 0.174). OPTICS and HDBSCAN performed similarly in Region C ($\text{ARI} = 0.951$).

These improvements stem from the density-based models' ability to handle arbitrarily shaped clusters and noise, enabling them to adapt more naturally to the varying spatial distribution of land covers in the dataset. Their success across both homogeneous and heterogeneous regions highlights their robustness in geospatial applications.

4.3. Ensemble Clustering and Voting Strategy

The ensemble clustering strategy, which aggregates the predictions of K-means, HDBSCAN, and OPTICS using majority voting with OPTICS as a tie-breaker, consistently produced the most balanced and accurate results across all regions. The overall ARI and NMI peaked at 0.967 and 0.922, respectively—the highest among all models evaluated. This improvement was found to be consistent in all three regions.

The ensemble approach also exhibited high classification accuracy (93.1%), as well as strong class-wise performance across dominant land covers, and acceptable-to-strong class-wise performance across rare class covers. The macro-averaged F1-score remained stable at 0.816, and precision and recall were uniformly high across all categories. Importantly, the ensemble successfully identified rare classes and preserved spatial coherence in built-up and vegetated areas even in densely mixed regions.

Visual comparisons in Figure 3 confirm the ensemble's superior spatial fidelity. The predicted maps closely follow the Dynamic World reference, effectively delineating boundaries between crops, trees, water, and urban patches. In Region C, the ensemble captured the high-contrast distribution of built-up and bare ground with greater sharpness than individual models.

4.4. Diagnostic Interpretation of Class Confusion

The main confusion observed in the classification results can be explained by a combination of spectral, structural, temporal, and reference-label factors. Grass and crop areas may exhibit similar NDVI, EVI, and MSAVI responses during the growing season, particularly when monthly composites smooth short-term phenological differences. Similarly, shrub and scrub areas may overlap spectrally with young forest, sparse vegetation, or unmanaged grassland. SAR features can help distinguish structural differences, but VV and VH backscatter are also sensitive to vegetation moisture, canopy geometry, and surface roughness, which may vary within the same semantic class. Confusion between bare ground and built-up areas is also expected because exposed soil, gravel surfaces, construction sites, and impervious surfaces can show similar optical brightness and SAR responses. In addition, because Dynamic World is used as the reference product, part of the observed disagreement may reflect uncertainty or temporal mismatch in the reference labels rather than only model error. These factors indicate that class-level performance should be in-

terpreted as agreement with the reference product and should be further validated using independent samples in future work.

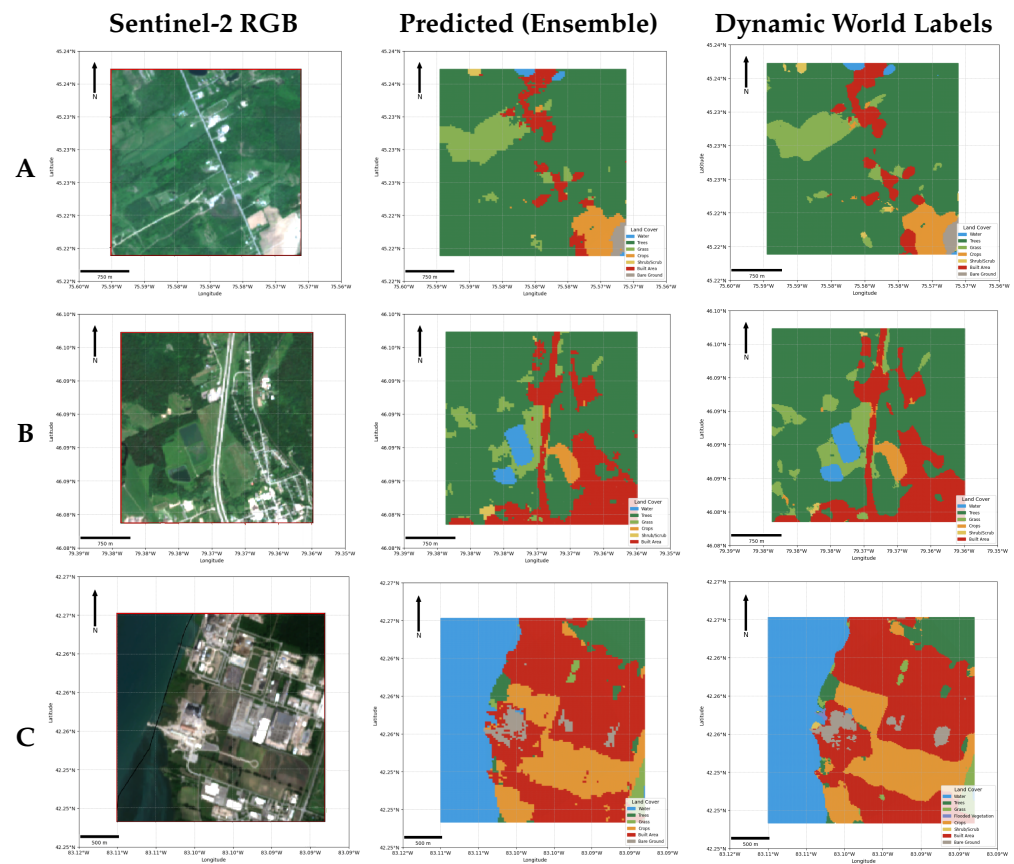


Figure 3. Visual comparison of land cover classification results across Regions A–C. Each row corresponds to a region (labeled (A)–(C)), with columns showing the Sentinel-2 RGB imagery (left), predicted labels from the ensemble method (middle), and Dynamic World reference labels (right).

4.5. Inclusion of Latitude, Longitude and Positional Information as Features

The inclusion of latitude, longitude, and transformed X/Y/Z coordinates may introduce location-specific bias. In the future, we intend to perform an ablation experiment to investigate the impact of inclusion of this positional information and to determine whether we may exclude these features and still achieve good performance from our proposed framework.

4.6. Implications for Externally Label-Guided Unsupervised Land Cover Classification

These findings demonstrate that **density-based clustering methods**—particularly when integrated through ensemble strategies—offer significant advantages for unsupervised land cover classification using multisource remote sensing data. Unlike K-means, which is sensitive to initialization and geometric assumptions, HDBSCAN and OPTICS adaptively capture spatial variance and small-scale heterogeneity. The ensemble further enhances this capacity by combining different clustering perspectives, mitigating individual weaknesses while reinforcing shared structures.

Moreover, the ensemble method’s consistent success in detecting **minority classes**, maintaining high **macro-averaged F1-scores**, and achieving **visual alignment with reference maps** reinforces its potential as a scalable framework for large-scale land cover mapping, especially in regions where annotated training data are scarce or costly to obtain.

4.7. Scalability of the Process

A practical consideration for operational use is the scalability of the proposed framework to regional or provincial extents. The algorithms evaluated in this study exhibit differing computational characteristics that influence their applicability at larger scales. KMeans, with its iterative centroid updates, scales efficiently to millions of samples when combined with mini-batch or distributed implementations. OPTICS and HDBSCAN, while more computationally intensive, provide advantages in capturing non-linear cluster structures and noise, but may require careful parameter tuning and partitioned processing when applied at provincial scales. The integration with Google Earth Engine for feature preparation and grid-based sampling provides a viable pathway to extend the analysis across Ontario, as the data handling and preprocessing are distributed across cloud resources. Future work could investigate hybrid strategies, such as combining fast global algorithms (e.g., KMeans) with localized density-based refinements (e.g., HDBSCAN), to balance scalability with accuracy when applied to large-scale land cover mapping.

Scaling the proposed framework to provincial or national extents would require partitioning the study area into manageable spatial tiles. The optimal tile size is inherently dependent on the available computational resources: smaller tiles reduce memory and runtime demands but increase the number of models to be trained, whereas larger tiles minimize the number of clustering runs at the expense of higher computational loads. These tiles can be processed either sequentially or in parallel, depending on the processing environment. Cloud-based platforms such as Google Earth Engine and distributed computing frameworks allow for substantial parallelization, thereby enabling the simultaneous processing of multiple tiles. This tiling approach ensures that the framework remains computationally feasible at regional scales while maintaining flexibility to adapt to varying hardware and software capacities.

4.8. Comparison with Existing Literature

Positioning our findings within prior work clarifies external validity and the incremental contribution of the framework. Fully controlled cross-study comparisons are inherently difficult due to differences in regions, taxonomies, and label sources. To mitigate this, we (i) focus on commonly reported metrics (OA, F1, ARI, NMI), (ii) harmonize class groupings where feasible, and (iii) interpret gaps only when sensors and spatial resolutions are comparable.

Alignment with prior unsupervised S1/S2 studies.

Multiple studies report that fusing Sentinel-1 SAR with Sentinel-2 optical imagery improves separability of water, wetlands, and mixed vegetation relative to optical-only pipelines [5,31–33]. Our results are consistent with these observations: SAR features (e.g., VH/VV ratio and temporal trends) sharpen water–wetland boundaries, while vegetation indices (e.g., NDVI/MSAVI) aid discrimination among vegetated classes at 10 m.

Comparison to baseline clustering pipelines.

Several land-cover pipelines rely primarily on centroid-based K-Means for partitioning [8,11,34]. In contrast, our ensemble (KMeans + HDBSCAN + OPTICS) aims to handle density heterogeneity and boundary ambiguity. OPTICS, in particular, is well-suited for revealing hierarchical density structure [30].

Feature representations and dimensionality reduction.

Prior work highlights that careful feature design and compact representations can stabilize clustering in high-dimensional remote-sensing spaces [35–38]. Consistent with

this, concentrating on MI-ranked features and applying SparsePCA reduces collinearity and noise before clustering; the effect is reflected in our stability trends.

Resolution and label-source considerations.

Comparisons across resolutions and label sources require caution. At 10 m, Sentinel-2 is widely reported to improve built-up delineation and water-edge detection relative to 30 m optical products [39]. Where studies benchmark against global products (e.g., ESA WorldCover), thematic differences can confound metric comparisons [40]. Our use of Dynamic World as an external reference introduces its own assumptions [41]; we therefore pair quantitative scores with qualitative overlays to assess boundary realism.

Takeaways

In line with existing literature, three patterns emerge: (1) multi-sensor fusion consistently aids water and wetland delineation [5,31,32]; (2) density-aware clustering and reachability-based extraction (OPTICS) mitigate failure modes observed with centroid-only pipelines [30]; and (3) moderate feature pruning and sparse projections improve stability without sacrificing separability [35,36]. Open challenges remain in standardizing cross-dataset protocols and harmonizing taxonomies for apples-to-apples benchmarking.

5. Conclusions

This work highlights the efficacy of integrating multisource remote sensing data, advanced feature engineering, and ensemble-based externally label-guided unsupervised classification to enhance land cover differentiation in the absence of ground-truth labels. The synergistic use of Sentinel-1 SAR and Sentinel-2 optical data, combined with multi-temporal data aggregation, enriches the feature space with complementary spatial and spectral information. Augmenting the input with DEM derivatives and transformed geolocation coordinates further improves the model's ability to distinguish spectrally similar classes. An ensemble of three diverse clustering algorithms, K-means, HDBSCAN, and OPTICS, was employed, with majority voting yielding more robust and accurate land cover delineation than individual models. These findings underscore the potential of ensemble-driven externally label-guided unsupervised methods for scalable and label-efficient land cover mapping. Although the proposed framework shows promising performance at the pilot-site scale, the current results should not be interpreted as evidence of province-wide mapping accuracy. Future work will extend the evaluation to larger spatial extents and more diverse ecological and urban–rural gradients.

Author Contributions: Conceptualization, M.M. and F.M.; methodology, R.S., M.M. and F.M.; software, S.O., R.S., M.M. and F.M.; validation, S.O., R.S., M.M. and F.M.; formal analysis, S.O., R.S., M.M. and F.M.; investigation, S.O., R.S., M.M. and F.M.; resources, R.S. and M.M.; data curation, S.O.; writing—original draft preparation, S.O.; writing—review and editing, S.O., R.S., F.M. and M.M.; visualization, S.O., R.S., M.M. and F.M.; supervision, R.S., F.M. and M.M.; project administration, R.S., F.M. and M.M.; funding acquisition, R.S. and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NSERC Discovery Grant RGPIN-2020-07155 to Reza Shahidi and NSERC Discovery Grant RGPIN-2022-04766 to Masoud Mahdianpari.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

Abbreviations

The following abbreviations are used in this manuscript:

HDBSCAN	Hierarchical Density-Based Spatial Clustering of Applications with Noise
OPTICS	Ordering Points to Identify the Clustering Structure
ARI	Adjusted Rand Index
NMI	Normalized Mutual Information

References

1. Fayaz, M.; Dang, L.M.; Moon, H. DLAN: A dual attention network for effective land cover classification in remote sensing. *Knowl.-Based Syst.* **2025**, *319*, 113620.
2. Treitz, P.; Rogan, J. Remote sensing for mapping and monitoring land-cover and land-use change—An introduction. *Prog. Plan.* **2004**, *61*, 269–279. [https://doi.org/10.1016/S0305-9006\(03\)00064-3](https://doi.org/10.1016/S0305-9006(03)00064-3).
3. Schmitt, M.; Hughes, L.H.; Qiu, C.; Zhu, X.X. SEN12MS—A curated dataset of georeferenced multi-spectral sentinel-1/2 imagery for deep learning and data fusion. *arXiv* **2019**, arXiv:1906.07789.
4. Ghamisi, P.; Plaza, J.; Chen, Y.; Li, J.; Plaza, A.J. Advanced spectral classifiers for hyperspectral images: A review. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–32.
5. Verma, A.; Kumar, S.; Bhattacharya, A. A cloud-based global land cover clustering framework using Sentinel-1 and Sentinel-2 data. In Proceedings of the 2023 International Symposium on Signals, Circuits and Systems (ISSCS), Iasi, Romania, 13–14 July 2023; pp. 1–4.
6. Xu, S.; Xiao, W.; Ruan, L.; Chen, W.; Du, J. Assessment of ensemble learning for object-based land cover mapping using multi-temporal Sentinel-1/2 images. *Geocarto Int.* **2023**, *38*, 2195832.
7. Gonçalves, M.; Costa, J.; Netto, M.; Mwasiagi, J. Land-cover classification using self-organizing maps clustered with spectral and spatial information. *Self-Organ. Maps Appl. Nov. Algorithm Des.* **2011**, *1*, 299–322.
8. Paris, C.; Gasparella, L.; Bruzzone, L. A scalable high-performance unsupervised system for producing large-scale hr land cover maps: The italian country case study. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 9146–9159.
9. Mai, D.S.; Nguyen, T.K.; Le, C.H.; Trinh, L.H. Applying Semi-Supervised Fuzzy C-Means Clustering Algorithm Based on Collaborative Clustering Model for Landcover Classification from Landsat-7 Imagery. *J. Sci. Tech.* **2024**, *7*, 171–184.
10. Chamundeeswari, V.V.; Singh, D.; Singh, K. An analysis of texture measures in PCA-based unsupervised classification of SAR images. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 214–218.
11. Abbas, A.W.; Minallh, N.; Ahmad, N.; Abid, S.A.R.; Khan, M.A.A. K-Means and ISODATA clustering algorithms for landcover classification using remote sensing. *Sindh Univ. Res. J.-SURJ (Sci. Ser.)* **2016**, *48*, 315–318.
12. Tamiminia, H.; Salehi, B.; Mahdianpari, M.; Quackenbush, L.; Adeli, S.; Brisco, B. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS J. Photogramm. Remote Sens.* **2020**, *164*, 152–170.
13. Ding, D.; Zhang, H.; Cao, Y.; Wang, D.; Sun, S. Multi-Source Data-Driven Brownfield Identification: Methodology and Spatial Characteristics of Third Line Construction Cities. *Urban Build. Sci.* **2026**, *2*, 5. <https://doi.org/10.53941/ubs.2026.100011>.
14. European Space Agency (ESA). Sentinel-1 SAR GRD: C-Band Synthetic Aperture Radar Ground Range Detected, IW Mode. 2015. Available online: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD (accessed on 1 May 2026).
15. Kaufman, Y.J.; Tanré, D. Algorithm for remote sensing of tropospheric aerosol from MODIS. *NASA MODIS Algorithm Theor. Basis Doc. Goddard Space Flight Cent.* **1998**, *85*, 3–68.
16. Gao, B.C.; Kauffman, Y.J. Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels: SAFARI 2000-Southern African Regional Science Initiative. *J. Geophys. Res.* **2003**, *108*, 4389.
17. Louis, J.; Debaecker, V.; Pflug, B.; Main-Knorn, M.; Bieniarz, J.; Mueller-Wilm, U.; Cadau, E.; Gascon, F. Sentinel-2 Sen2Cor: L2A processor for users. In Proceedings of the Living Planet Symposium 2016, Prague, Czech Republic, 9–13 May 2016; pp. 1–8.
18. Rouse, J.W., Jr.; Haas, R.H.; Schell, J.; Deering, D. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; Technical Report; NASA: Washington, DC, USA, 1973.
19. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213.
20. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126.
21. Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594.

22. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033.
23. Gao, B.C. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266.
24. Szigarski, C.; Jagdhuber, T.; Baur, M.; Thiel, C.; Parrens, M.; Wigneron, J.P.; Piles, M.; Entekhabi, D. Analysis of the radar vegetation index and potential improvements. *Remote Sens.* **2018**, *10*, 1776.
25. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–137.
26. Arthur, D.; Vassilvitskii, S. *K-Means++: The Advantages of Careful Seeding*; Technical Report; Stanford University: Stanford, CA, USA, 2006.
27. Rodríguez-Corominas, G.; Blesa, M.J.; Blum, C. FastKPP: An exact accelerated k-means++ algorithm. *Softw. Impacts* **2026**, *28*, 100830.
28. Campello, R.J.; Moulavi, D.; Sander, J. Density-based clustering based on hierarchical density estimates. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 160–172.
29. McInnes, L.; Healy, J.; Astels, S. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2017**, *2*, 205.
30. Ankerst, M.; Breunig, M.M.; Kriegel, H.P.; Sander, J. OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod Rec.* **1999**, *28*, 49–60.
31. Zhao, L.; Huang, H.; Lu, D. Fusion of Sentinel-1, Sentinel-2, and DEM Data for Land Cover Classification Using Hybrid Features. *Remote Sens.* **2021**, *13*, 734. <https://doi.org/10.3390/rs13040734>.
32. Ienco, D.; Interdonato, R.; Gaetano, R.; Minh, D.H.T. Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 11–22.
33. Razzano, F.; Iandolo, M.R.; Zarro, C.; Yogesh, G.; Ullo, S.L. Integration of Sentinel-1 and Sentinel-2 data for Earth surface classification using Machine Learning algorithms implemented on Google Earth Engine. In *Proceedings of the 2023 IEEE India Geoscience and Remote Sensing Symposium (InGARSS)*, Bangalore, India, 10–13 December 2023; pp. 1–4.
34. Kharat, V.; Khatdeo, S.; Kothe, H.; Kshirsagar, R.; Dixit, M.; Balan, M.S. Land Cover Clustering and Classification of Satellite Images Using K-Means and Random Forest with SLIC Segmentation. *ITM Web Conf.* **2023**, *56*, 02004. <https://doi.org/10.1051/itmconf/20235602004>.
35. Chakraborty, S. Clustering Techniques for Land Use Land Cover Classification of Remotely Sensed Images. *Int. J. Remote Sens. GIS* **2019**, *8*, 17–23.
36. Romero, A.; Gatta, C.; Camps-Valls, G. Unsupervised deep feature extraction for remote sensing image classification. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 1349–1362.
37. Xie, J.; Girshick, R.; Farhadi, A. Unsupervised deep embedding for clustering analysis. In *Proceedings of the International Conference on Machine Learning*, New York, NY, USA, 20–22 June 2016; pp. 478–487.
38. Yang, Y.; Pei, Y.; Cao, Y. Feature-driven compact representation model for analysis and visualization of large-scale multivariate SAMR data. *Comput. Graph.* **2025**, *113*, 104331.
39. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.; Murayama, Y.; Ranagalage, M. Sentinel-2 data for land cover/use mapping: A review. *Remote Sens.* **2020**, *12*, 2291.
40. Duarte, D.; Fonte, C.; Costa, H.; Caetano, M. Thematic comparison between ESA WorldCover 2020 land cover product and a national land use land cover map. *Land* **2023**, *12*, 490.
41. Brown, C.F.; Brumby, S.P.; Guzder-Williams, B.; Birch, T.; Hyde, S.B.; Mazzariello, J.; Czerwinski, W.; Pasquarella, V.J.; Haertel, R.; Ilyushchenko, S.; et al. Dynamic World, Near real-time global 10 m land use land cover mapping. *Sci. Data* **2022**, *9*, 251.

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.