Hindawi
Journal of Sensors
Volume 2023, Article ID 6657171, 14 pages
https://doi.org/10.1155/2023/6657171

Research Article

Google Earth Engine for Advanced Land Cover Analysis from Landsat-8 Data with Spectral and Topographic Insights

Abolfazl Abdollahi,[1,2] **Biswajeet Pradhan**,[2,3] **Abdullah Alamri**,[4] **and Chang-Wook Lee**[5]

[1]*Fenner School of Environment & Society, College of Science, The Australian National University, Canberra, ACT, Australia*
[2]*Centre for Advanced Modelling and Geospatial Information Systems, School of Civil and Environmental Engineering, Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia*
[3]*Earth Observation Centre, Institute of Climate Change, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia*
[4]*Department of Geology & Geophysics, College of Science, King Saud University, P.O. Box, 2455, Riyadh 11451, Saudi Arabia*
[5]*Department of Science Education, Kangwon National University, 1 Gangwondaehak-gil, Chuncheon-si, Gangwon-do 24341, Republic of Korea*

Correspondence should be addressed to Biswajeet Pradhan; biswajeet24@gmail.com

Received 8 June 2023; Revised 18 August 2023; Accepted 2 September 2023; Published 26 October 2023

Academic Editor: Sana Ullah Jan

The primary goal of this research is to see how effective cloud-based computing services such as Google Earth Engine (GEE) platform are at classifying multitemporal satellite images and producing high-quality land cover maps for the target year of 2020, with the possibility of using it on a larger-scale area such as metropolitan Melbourne as a test site. To create high-quality land cover maps, the GEE is utilized to analyze a total of 80 Landsat-8 images. The support vector machine (SVM) approach is used to classify the images. Moreover, we use spectral bands, spectral indices, and topographic parameters to improve classification and address the limitations of existing approaches for classification with restricted input variables. Furthermore, we apply a postprocessing strategy to increase the model's performance by removing the salt-and-pepper noise created by misclassified pixels in supervised classification results. The results demonstrate that given all parameters, the SVM approach achieves an overall accuracy (OA) and kappa accuracy of 88.47% and 85.34%, respectively. Following the implementation of the postprocessing technique, the OA and kappa improve to 92.90% and 90.99%, respectively. The results indicate that Landsat-8 multitemporal data, spectral indices, topographic components, and postprocessing techniques are all important in land cover mapping. Therefore, the use of freely accessible GEE technology and multitemporal Landsat-8 data ensures that decision makers have the resources they need to track land cover throughout the year.

1. Introduction

Land cover information is required for a variety of tasks, such as greenhouse gas emissions monitoring associated with forest degradation and deforestation [1], natural resources assessment [2], regional and urban planning [3], and global environmental modeling [4]. Landsat images have been the dominant source of data for extracting land cover information in cities among the accessible high- and moderate-resolution remote sensing data because of their higher resolution and long-term acquisition capacity [5, 6]. The successful launch of Landsat-8 on February 11, 2013, ensured that the Landsat Earth surveillance mission would continue [7]. The Landsat-8 sensors comprise a thermal infrared sensor with two thermal bands and an operational land imager (OLI) with nine bands, including the high-resolution panchromatic band. Land cover maps are frequently created using remote sensing imagery classification algorithms [8]. However, processing remote sensing imagery in a time-effective manner and producing accurate land cover maps remain a challenge for the remote sensing society.

Machine learning (ML) techniques have been effectively utilized and expanded in various fields due to their remarkable abilities of self-learning and adaptive parallel information processing [9–12]. In land cover classification, ML approaches such as random forest (RF) [13], support vector

machine (SVM) [14], and artificial neural network (ANN) [15] have been commonly used. In the following, more such examples are provided. Using Sentinel-2 and airborne data, Morell-Monzó et al. [16] used RF to effectively quantify and detect discarded agricultural areas and provide a way to map citrus and other in severely fragmented plots. LaRocque et al. [17] applied RF to successfully map 11 types of wetlands in Southern New Brunswick, Canada, by combining multisource remote sensing data. Mao et al. [12] performed different classification approaches such as ANN, SVM, and RF for land cover classification in Hangzhou, China. Their findings revealed that the ANN model performs poorly in critical urban land cover categorization. Chakhar et al. [18] discovered that among 22 nonparametric classification techniques for categorizing irrigated crops in a semiarid environment, SVM and nearest neighbor approaches provided the best balance of resilience and effectiveness. Fragou et al. [19] used the SVM classifier to classify land cover in a Mediterranean area from Landsat Thematic Mapper data for various years, which achieved an overall accuracy (OA) of 90%. Zhao et al. [20] implemented various classification methods such as SVM, RF, and decision tree (DT) to classify land cover in typical mountainous areas. They discovered that SVM categorized quickly but needed detailed feature variables and DT had the best classification performance but the worst consistency.

To classify Sentinel-2 imagery for boreal landscapes, Abdi [21] used eXtreme gradient boosting (XGBoost), RF, and SVM classifiers. With 75.8% OA, the SVM model outperformed the others in the experiments. Jia et al. [22] used maximum likelihood and SVM classifiers to explore the OLI data for land cover categorization in Beijing, China. They discovered that SVM and OLI data showed satisfactory results for the land cover classification. When processing large amounts of satellite data for land cover classification, ML techniques have obtained satisfactory results. However, Adam et al. [23] stated that the ANN technique has a significant degree of complexity in terms of computing processing. They also claimed that RF is vulnerable to noise and overtraining. Prasad et al. [24] reported that DT is very sensitive to modest changes in the training dataset and is often inconsistent. According to Naidoo et al. [25], finding an appropriate value of k for the K-nearest neighbors (KNN) approach is difficult. In addition, land cover information extraction research is now primarily focused on monotemporal Landsat data of cities [26]. However, when monotemporal image data are used to extract land cover information, the accuracy of information extraction is severely affected due to the phenomenon of "the same object with different spectra and different objects with the same spectrum."

Image processing, such as the classification of massive volumes of image data using ML techniques, is made possible by cloud-based computing services [27]. Google Earth Engine (GEE) is a free-to-use cloud-based geospatial analytic platform [28]. Accessing a large volume of RS data and preprocessing with GEE have become easier. Prior studies used single-date Landsat images and classification algorithms. In addition, most existing techniques have been implemented for land cover mapping with a restricted number of input parameters, leading to low accuracy.

Therefore, this study aims to use the ML classification approach, e.g., SVM, to create a land cover map based on multitemporal input variables. SVM provides fast calculation speed and better generalization capacity and accuracy compared with classic learning methods and is frequently utilized in image and land cover classification. One of the main advantages of using SVM in comparison to other ML models, such as RF, is its effectiveness in handling high-dimensional data and situations involving limited training samples. SVM is especially well-suited for classification tasks where the number of features (dimensions) exceeds the number of samples —a scenario frequently encountered in applications like remote sensing and image classification. SVM operates by identifying the optimal hyperplane that maximizes the margin between classes, leading to improved generalization even when dealing with a small training dataset. On the contrary, while RF is also a robust classifier, it may encounter challenges such as overfitting when confronted with high-dimensional data or imbalanced class distributions. SVM's capability to address such challenges positions it as the preferred choice for tasks like land cover mapping and others where data dimensionality and limited training samples are prevalent. The main contribution of this research is to create a land cover map for the year 2020 by utilizing Landsat-8 time-series data for a large-scale area (e.g., the Melbourne metropolitan area in Australia as a test site) with 9,993 km^2 based on the cloud-based GEE platform. We also incorporated a variety of data, including spectral bands (SB), spectral indices (SI), and topographic features (TF), to improve the classification and solve the ML difficulties with few input variables discussed before. To put it another way, we looked at a variety of characteristics to see if a multitemporal composite could aid the model in accurately mapping land cover in the study area. Furthermore, we applied a postprocessing technique to remove noise caused by misclassified pixels. We first used an unsupervised clustering method called simple noniterative clustering (SNIC) to create superpixel clustering. Then, we assigned each classified class to each cluster on the basis of majority voting to remove misclassified pixels and improve the classification results.

To the best of our knowledge, the presented methodology has not been performed in the literature, and this is the first time that the model has been applied on multitemporal Landsat-8 images with additional input variables and postprocessing to improve land cover classification for the study area. The developed method was tested in a relatively large region of the Melbourne metropolitan area in Australia. The land cover maps' accuracy is assessed by comparing them with reference data. We also compared our classification results with those of other ML methods to show the effectiveness of the proposed technique for land cover classification from multitemporal data. Furthermore, we undertook an examination of the model's transferability and generalizability across different geographic regions, yielding valuable insights into its effectiveness across diverse areas. Through the application of the model to another region, specifically

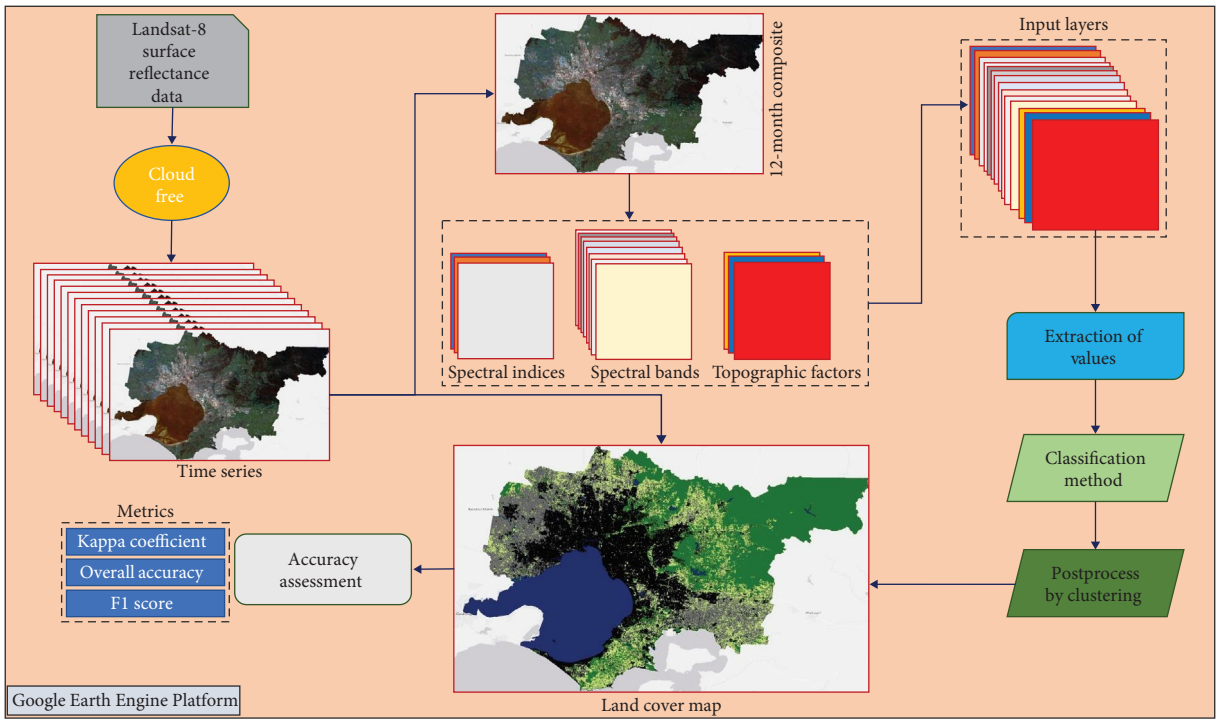

Figure 1: Overall flowchart of the land cover classification process in the GEE platform.

the Australian Capital Territory (ACT), we evaluated its performance under various conditions, showcasing its capacity to deliver satisfactory results in land cover mapping across a range of terrains and landscapes. The results demonstrated that the suggested land cover mapping workflow, which includes a cloud-based image processing method, could provide a reliable way for utilizing multitemporal data and mapping land cover.

The paper introduces several main contributions, which are outlined as follows:

(1) Creation of a regional-scale land cover map using Landsat-8 time-series data and the GEE platform with ML techniques.
(2) Integration of SB, SI, and TF to improve classification and overcome ML difficulties with limited input variables.
(3) Application of a postprocessing technique to remove noise caused by misclassified pixels.
(4) Assessment of land cover map accuracy through comparison with reference data and other ML methods.
(5) Evaluation of the model's transferability and generalizability across diverse geographic regions provides valuable insights into its reliability across different terrains and landscapes.

2. Methodology

The classification procedure consisted of three main steps: data collection, classification and accuracy measurement, and postprocessing step. GEE was first utilized to collect surface reflectance data for the year 2020, which is the target year. Then, SB, SI, and TF were calculated. In the second stage, image classification and accuracy evaluation were performed using the SVM algorithm and a confusion matrix. Finally, to eliminate noise and improve classification performance, we used a postprocessing phase based on an unsupervised clustering technique and majority voting. The general techniques for developing land cover maps are shown in Figure 1, and more information on each step will be presented in the subsequent subsections.

2.1. Data and Study Area. Metropolitan Melbourne, the capital of the state of Victoria, Australia, is the test region, and it is located at 37.8001°S, 145.3143°E of Australia's southeastern part, with a total area of 9,993 km^2. Multitemporal Landsat-8 satellite imagery was gathered to implement the classification, and the collection contains multiple cloud-free images for the year 2020 (12 months). We classified and evaluated the spatial distribution of land cover classes for the year using a total of 80 Landsat-8 images. Table 1 shows the spatial and spectral resolution of Landsat-8 data. Figure 2 depicts the location of the study area.

2.2. Training and Testing Samples. Water body, high vegetation (e.g., trees, shrubs), low vegetation (e.g., grass, irrigated pasture), crops, and built-up areas were used to create the classification scheme in this study. Ground truth samples were taken using Google Earth images, examination of the true- and false-color composites of the Landsat-8 data, and expert knowledge. A total of 4,497 samples were selected at the pixel level and then randomly divided into training and testing sets, with an 80–20 split. The suggested model was

Table 1: Spatial and spectral resolution of Landsat-8 satellite data.

Bands	Wavelength (μm)	Resolution (m)
Coastal aerosol	0.43–0.45	30
Blue	0.45–0.51	30
Green	0.53–0.59	30
Red	0.64–0.67	30
NIR	0.85–0.88	30
SWIR1	1.57–1.65	30
SWIR2	2.11–2.29	30
Panchromatic	0.50–0.68	15
Cirrus	1.36–1.38	30
Thermal infrared 1	10.60–11.19	100
Thermal infrared 2	11.50–12.51	100

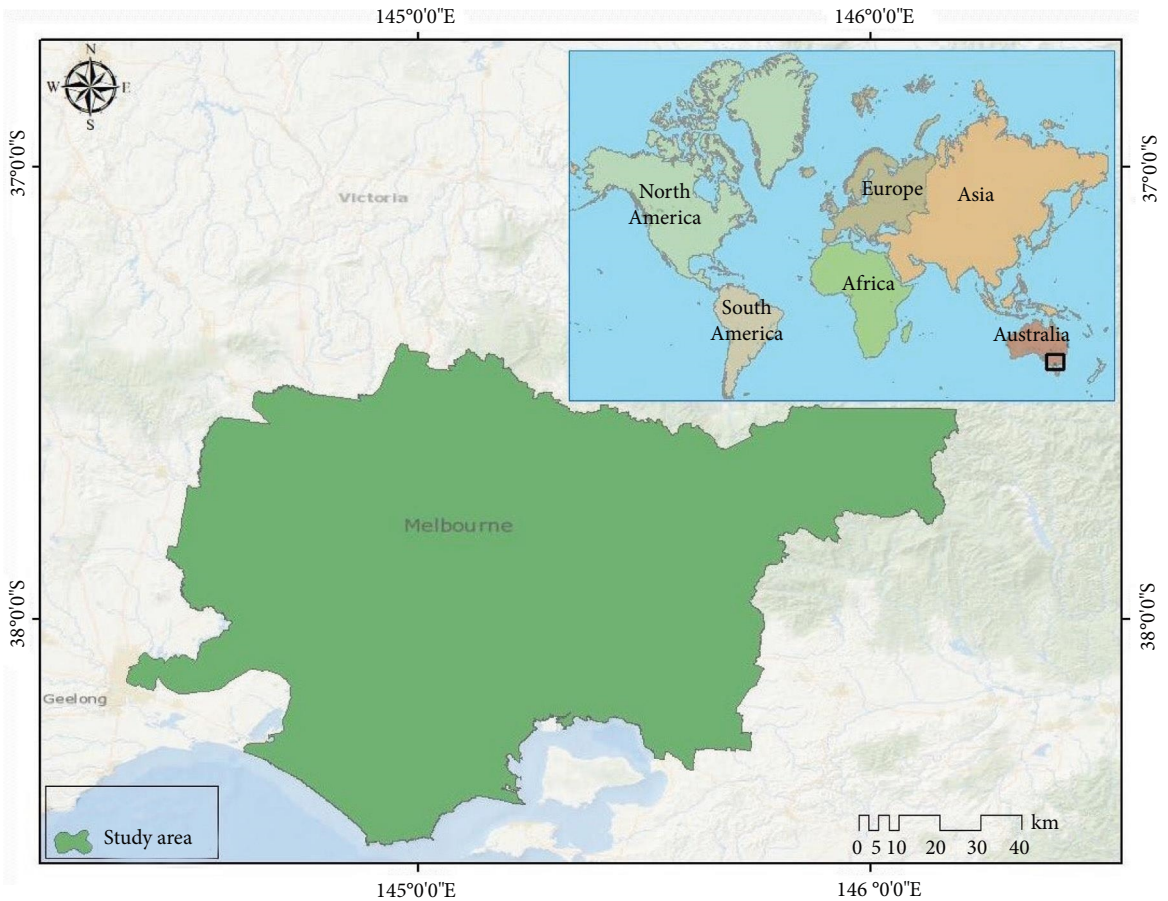

Figure 2: Location of the Melbourne metropolitan region, Australia.

trained with the training set, and the model's performance was evaluated with the testing set.

2.3. Input Data. A time series of input data was obtained to characterize the training samples and differentiate between diverse land cover types. As shown in Table 2, 13 input variables were employed for land cover mapping in the Melbourne metropolitan region for the year 2020, including SBs, TFs, and SIs. The variables were calculated as a 12-month average. First, we used eight bands (blue, green, red, near-infrared, SWIR1, SWIR2, thermal infrared bands 10 and 11) for the SBs. To maintain spatial consistency among the Landsat-8 data, the thermal infrared bands were downscaled to the spatial resolution SBs. For the TFs, advanced land observing satellite digital elevation model with a horizontal resolution of about 30 m was utilized to determine TFs,

Table 2: Number of input variables for the SVM approach utilized to generate the land cover map.

Category	Description	Input variables number
Spectral bands	Blue, green, red, near-infrared, SWIR1, SWIR2, and thermal infrared bands 10 and 11	8
Spectral indices	NDVI, NDBI, and NDWI	3
Topographic features	Slope and elevation	2
Total variable		13

including slope and elevation (two variables). Then, using Equations (1)–(3), we generated SIs such as the normalized difference water index (NDWI) [29], the normalized difference built-up index (NDBI) [30], and the normalized difference vegetation index (NDVI) [31].

$$\mathrm{NDWI} = \frac{\mathrm{Green} - \mathrm{NIR}}{\mathrm{Green} + \mathrm{NIR}}, \tag{1}$$

$$\mathrm{NDBI} = \frac{\mathrm{MIR(band\,6)} - \mathrm{NIR(band\,5)}}{\mathrm{MIR(band\,6)} + \mathrm{NIR(band\,5)}}, \tag{2}$$

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + \mathrm{Red}}. \tag{3}$$

2.4. Classification Approach. In this work, we used the SVM method for land cover classification. SVM, a supervised learning binary classifier that operates on the structural risk minimization concept, is one of the most prominent ML algorithms [9, 10]. SVM is often used in land cover classification because it has a faster computation speed and better generalization capacity and accuracy than traditional learning approaches [32]. In the categorization task, SVM divides a given training dataset on the basis of a hyperplane, which is known as the maximal margin hyperplane, to maximize the distance between them. SVM's goal is to find an n-dimensional hyperplane that distinguishes between two types on the basis of their maximum gap [33]. The following is the mathematical formula (Equations (4) and (5)):

$$\frac{1}{2}\|w\|^2, \tag{4}$$

$$y_i((w \cdot x_i) + b) \geq 1, \tag{5}$$

where b is denoted as a constant, $\|w\|$ is denoted as the norm of the normal hyperplane, and x and y represent the feature vector and the target. The cost function L can be represented, as shown in Equation (6), after multiplying the Lagrangian coefficient (λ_i):

$$L = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \lambda_i (y_i((w * x_i) + b) - 1). \tag{6}$$

Equation (7) can be changed as follows for the nonseparable case of the slack variable ξ [34]:

$$y_i((w * x_i) + b) \geq 1 - \xi_i. \tag{7}$$

Afterward, Equation (8) can be expressed as follows, where v (0, 1) describes the misclassification [35]:

$$L = \frac{1}{2}\|w\|^2 - \frac{1}{vn}\sum_{i=1}^{n} \xi_i. \tag{8}$$

According to a study by Shi and Yang [36], the radial basis function (RBF) kernel of the SVM method is often employed in land cover classification tasks because it performs well. The RBF kernel of SVM can be expressed, as shown in Equation (9):

$$K(x_i, x_j) = \exp\left(-\gamma \left\| x_i, x_j \right\|^2\right), \gamma > 0, \tag{9}$$

where γ is one of the kernel functions' parameters.

2.5. Postprocessing Step. Salt-and-pepper noise caused by misclassified pixels is common in supervised classification results. To eliminate such noise, the use of postprocessing methods is generally preferable. Thus, we used unsupervised clustering methods to replace classified values in each cluster with the majority value. For the unsupervised clustering method, we applied the SNIC approach to cluster the image. The three key parameters of SNIC are neighborhood size, which avoids tile border artifacts; connectivity, which indicates the contiguity type to combine neighboring clusters (rook: eight or queen: four); and compactness, which impacts the clusters' shape. These parameters were set empirically based on the features of the research area as follows: neighborhood size = 64, connectedness = 8, and compactness = 0.5. After clustering, the classified image was overlaid on the clustered image, every classified class was assigned to each cluster on the basis of majority voting, and the mean pixels were extracted. This postprocessing step removed the salt-and-pepper noise caused by the misclassified pixels and improved the classification results.

2.6. Metrics for Assessing Accuracy. On the basis of OA, recall, precision, F1 score, and kappa coefficient, we calculated the accuracy of the proposed model for land cover classification [37]. The kappa coefficient is a more sophisticated metric that compares random chance with observed accuracy. OA is a straightforward summary evaluation of a case's probability of being accurately classified. The number of true pixels identified in each category is referred to as recall. For each category, precision indicates how many accurate pixels have been detected [38]. In imbalanced training

Table 3: Quantitative results of the suggested SVM approach for land cover classification.

Method	Classes	Precision	Recall	F1 score	OA	Kappa
SVM + SB	Water	0.9839	0.9899	0.9869	0.8412	0.7983
	High vegetation	0.9428	0.885	0.913		
	Low vegetation	0.6773	0.6208	0.6478		
	Crops	0.6733	0.7179	0.6949		
	Built-up	0.8225	0.8944	0.857		
SVM + SB + SI	Water	0.9886	0.9889	0.9887	0.8607	0.8230
	High vegetation	0.9217	0.9261	0.9239		
	Low vegetation	0.6992	0.7026	0.7009		
	Crops	0.7241	0.7374	0.7307		
	Built-up	0.8782	0.8592	0.8686		
SVM + SB + SI + TF	Water	0.9875	0.9932	0.9903	0.8847	0.8534
	High vegetation	0.9546	0.9125	0.9331		
	Low vegetation	0.7514	0.7536	0.7525		
	Crops	0.7866	0.7878	0.7872		
	Built-up	0.8673	0.9107	0.8885		

data, the F1 score is a quantitative statistic that measures the balance between recall and precision. The following metrics (Equations (10)–(14)) can be computed using true positive (TP), false negative (FN), false positive (FP), and true negative (TN):

$$\mathrm{OA} = \frac{\mathrm{TP} + \mathrm{TN}}{N}, \tag{10}$$

$$\kappa = \frac{p_0 - \mathrm{p}_e}{1 - \mathrm{p}_e}, \tag{11}$$

where

$$p_0 = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \text{ and } p_e = \frac{(\mathrm{TP} + \mathrm{FN}) \times (\mathrm{TP} + \mathrm{FP}) \times (\mathrm{FP} + \mathrm{TN}) \times (\mathrm{FN} + \mathrm{TN})}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})^2}, \tag{12}$$

$$\mathrm{F1} = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}, \tag{13}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{14}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}. \tag{15}$$

3. Results

Quantitative and qualitative results yielded by the proposed SVM technique for land cover classification from Landsat-8 data are discussed in this section. Table 3 shows the results of the SVM approach combined with other elements, such as SBs, SIs, and TFs, based on all the abovementioned metrics. On the basis of the table, the SVM model with only SBs (SVM + SB) could achieve 84.12% and 79.83% for OA and kappa. After SIs were added to the model (SVM + SB + SI), the OA and kappa improved to 86.07% and 82.30%, respectively. Also, incorporating TFs into the method (SVM + SB + SI + TF) could even help the model achieve better OA of 88.47% and kappa of 85.34%. In general, the proposed model with different factors could achieve satisfactory results for land cover classification for the year 2020. However, integrating more factors such as SIs and TFs helped the model obtain better OA and kappa. We illustrated the visualization results attained by the proposed model with all the SBs, SIs, and TFs, as shown in Figure 3. The original multitemporal Landsat-8 images, the qualitative results of SVM + SB, the qualitative results of SVM + SB + SI, and the visualization results of SVM + SB + SI + TF are shown in Figure 3(a)–3(d), respectively. On the basis of the figure, after SIs were incorporated into the model (SVM + SB + SI), the visualization outcomes were improved compared with the model with only SBs (SVM + SB), and the method could achieve a smooth land cover map. Furthermore, the SVM approach could obtain better qualitative results when we combined the topographic information with prior factors. In other words, compared with SVM + SB (Figure 3(b)) and SVM + SB + SI (Figure 3(c)), the suggested SVM + SB + SI + TF (Figure 3(d)) could predict fewer FPs and FNs for various land cover classes, thereby producing satisfactory land cover maps.

Moreover, we obtained quantitative and qualitative results for land cover classification from Landsat-8 data after applying the postprocessing step to see the effectiveness of the proposed postprocessing technique in improving the classification results. All the accuracy assessment metrics obtained by the proposed technique for land cover mapping are shown in Table 4. Evidently, after the postprocessing technique was implemented, the results of the SVM method with all the factors were improved. For example, the OA and kappa for SVM + SB + SI + TF increased to 92.90% and 90.99%, respectively, indicating improvements of 4.43% and 5.65%,

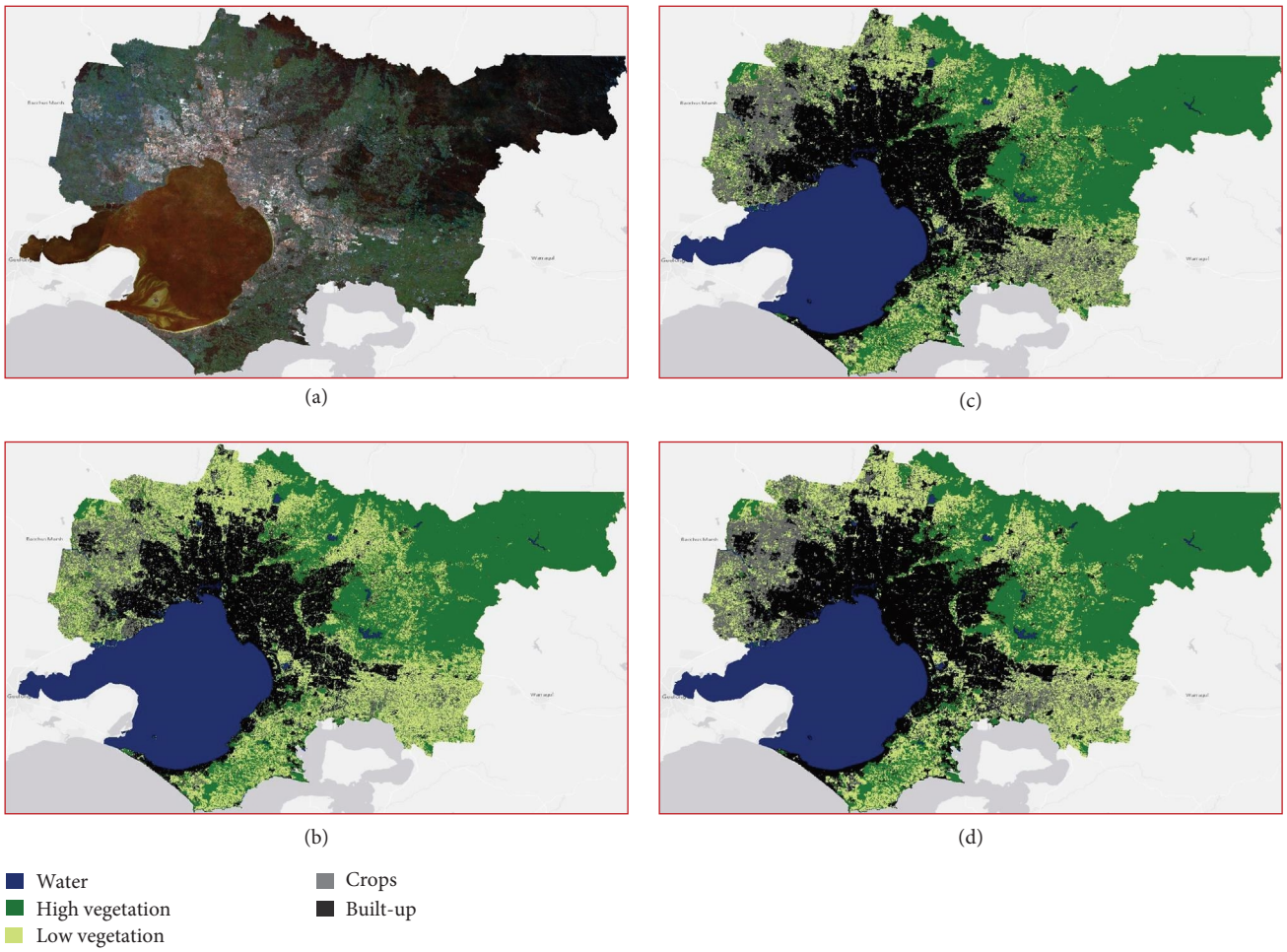

Figure 3: Visualization results of the land cover classification achieved by the proposed model: (a) original multitemporal Landsat-8 image, (b) results of SVM + SB, (c) results of SVM + SB + SI, and (d) results of SVM + SB + SI + TF.

Table 4: Quantitative results after postprocessing step for land cover classification.

Method	Classes	Precision	Recall	F1 score	OA	Kappa
SVM + SB	Water	0.9836	0.9871	0.9853	0.8552	0.8162
	High vegetation	0.9253	0.9284	0.9269		
	Low vegetation	0.6957	0.6971	0.6964		
	Crops	0.7217	0.7103	0.7159		
	Built-up	0.8581	0.8613	0.8597		
SVM + SB + SI	Water	0.9928	0.9844	0.9886	0.9070	0.8817
	High vegetation	0.9733	0.9116	0.9414		
	Low vegetation	0.7956	0.7866	0.7911		
	Crops	0.798	0.8609	0.8282		
	Built-up	0.9029	0.9416	0.9219		
SVM + SB + SI + TF	Water	0.995	0.9925	0.9938	0.9290	0.9099
	High vegetation	0.9753	0.9488	0.9619		
	Low vegetation	0.8881	0.8111	0.8479		
	Crops	0.8314	0.8864	0.858		
	Built-up	0.9159	0.9588	0.9368		

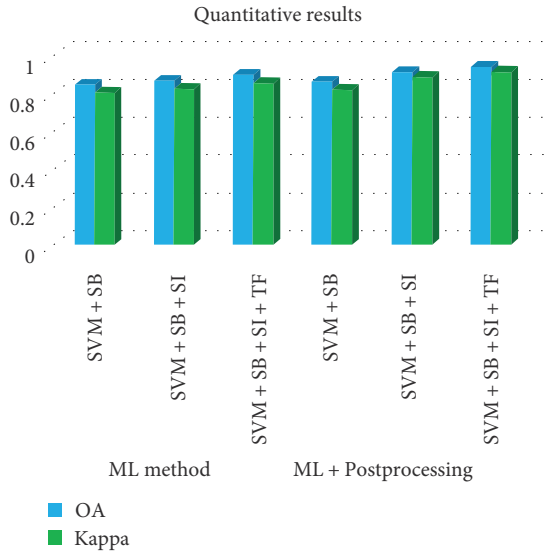

Figure 4: Performance metrics (OA and kappa) achieved by the proposed ML method, including additional factors before and after postprocessing for land cover classification.

respectively, compared with SVM + SB + SI + TF before postprocessing. Figure 4 demonstrates the difference in accuracy assessment metrics (OA and kappa) attained before and after postprocessing using the suggested ML method with extra factors for land cover mapping. The presented SVM method plus additional factors and the metrics percentage are represented on the horizontal and vertical axes, respectively. The results confirmed the efficacy of incorporating more features such as SIs and TFs into the model, as well as the proposed postprocessing technique in land cover classification. After performing the postprocessing approach, we depicted the qualitative results yielded by the proposed SVM approach and all the factors. Figure 5 shows the visualization results, in which Figure 5(a) shows the original multitemporal Landsat-8 images, Figure 5(b) provides the qualitative results of SVM + SB, Figure 5(c) shows the qualitative results of SVM + SB + SI, and Figure 5(d) presents the visualization results of SVM + SB + SI + TF. As can be seen, the proposed model can attain higher-quality land cover maps for all the factors compared with the visualization results achieved by the method before the postprocessing step. Also, on the basis of the visualization results after postprocessing, the proposed SVM model with all the factors could identify land cover classes more accurately and produce a better land cover map. This result confirms the effectiveness of adding the postprocessing technique and additional features to the model in classification. In addition, to scrutinize whether land cover mapping benefits from incorporating temporal information, we compared the results of the proposed method with the monotemporal Landsat-8 data. Table 5 and Figure 6 depict the quantitative and qualitative results of the technique for the monotemporal data, respectively. As it is evident, the model obtained less accurate results for monotemporal data than multitemporal data with OA of 84.92% and kappa of 80.89%. Also, the model misclassified pixels that present similar spectral values when we only used single-date imagery, which leads to less accuracy for land cover mapping according to the visualization results. Thus, when land uses include time–variant features, incorporating multitemporal information into the classification framework has improved categorization [39]. Our proposed framework for land cover mapping holds significant practical relevance due to its broad range of applications across various domains. For instance, in urban planning, the model's capability to offer detailed and up-to-date land cover information can play a pivotal role in making informed decisions related to urban expansion, infrastructure development, and allocation of green spaces. Moreover, environmental monitoring agencies can leverage the model's outputs to evaluate land use changes, monitor vegetation dynamics, and assess the impact of land management practices on ecosystems. Our proposed methodology not only promotes sustainable land use and resource management practices but also equips decision makers and stakeholders with vital information to tackle pressing challenges in diverse fields.

4. Discussion

To better provide insights into the land cover classification performance of the proposed SVM model, we also applied the classification and regression trees (CART) method to similar multitemporal Landsat-8 data and compared the results. Table 6 shows the qualitative results of calculating the accuracy of all assessment indicators. The results demonstrate that adding additional factors to the CART model such as SIs and TFs could help the model improve the results similar to the SVM method. For example, integrating TFs with the previous factors (SB and SI) helped the CART method improve OA and kappa to 85.24% and 81.23%, respectively. However, the CART approach generally achieves lower accuracy than SVM for land cover classification. For instance, the OA and kappa for CART + SB decreased to 4.7% and 6.04%, respectively, compared with the SVM + SB technique. These findings prove that the SVM method is more generalizable and superior to the CART method for land cover classification from multitemporal Landsat-8 images. Figure 7 illustrates the visualization results of land cover classification achieved using the CART approach with different parameters. Figure 7(a)–7(d) presents the original multitemporal Landsat-8 image, CART + SB results, CART + SB + SI results, and CART + SB + SI + TF results, respectively. Evidently, adding SIs and TFs to the model could help it categorize land cover classes more effectively and generate better visualization outcomes than a CART classifier that solely uses SBs. In practice, using simply SBs, the CART model was unable to reliably identify the pixels of land cover classes, particularly in more complicated regions with more obstructions, thereby generating low-resolution maps. Furthermore, with or without other components, the CART technique predicted more FPs and FNs than the SVM method, particularly for the crops and low vegetation classes, which leads to less accurate land cover maps than the SVM model. Both CART and SVM results demonstrated that adding SIs and TFs to the SVM classifier and postprocessing techniques could enhance the classification performance of

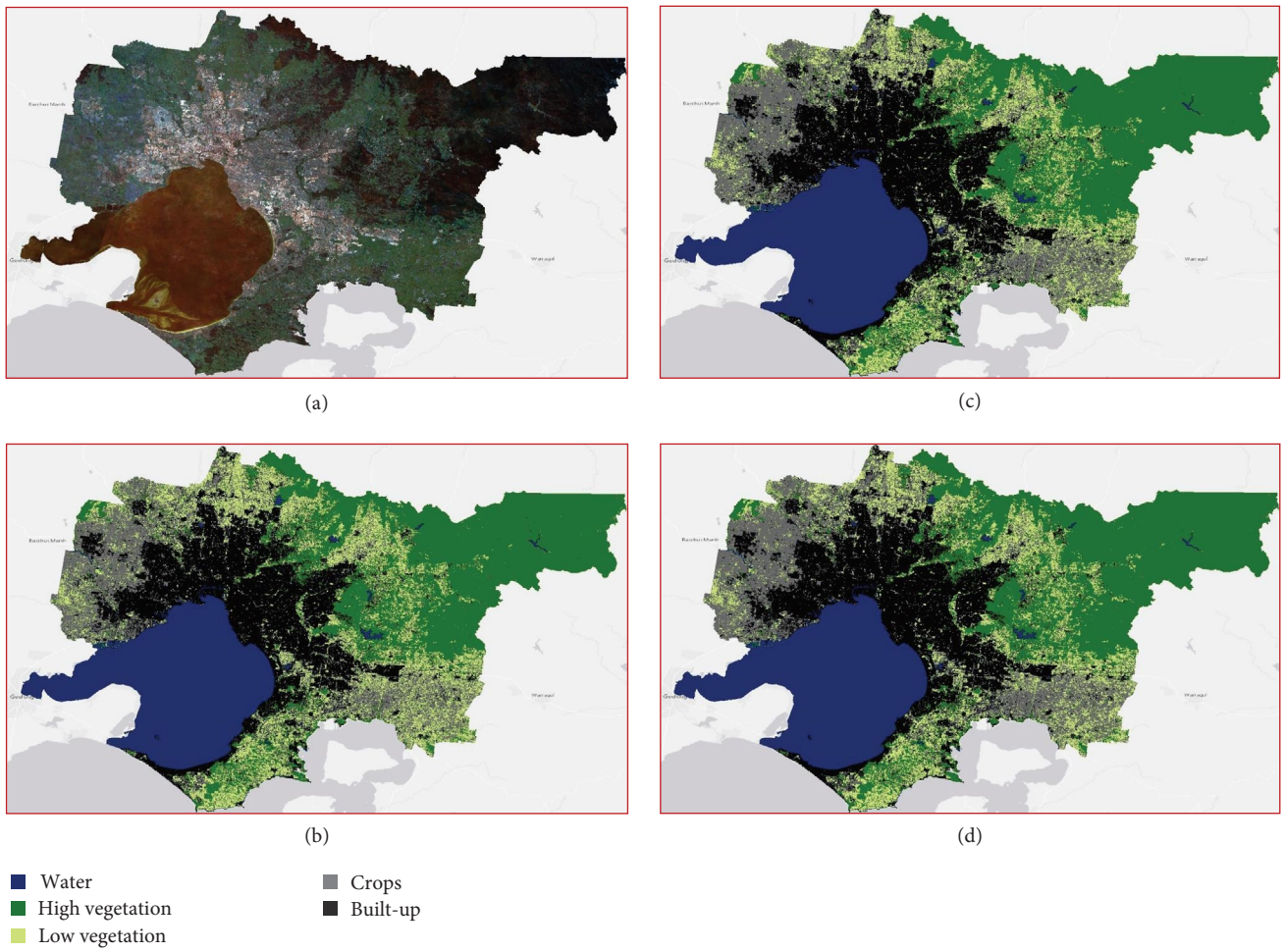

Figure 5: Visualization results of the land cover classification achieved by the proposed model after postprocessing: (a) original multitemporal Landsat-8 image, (b) results of SVM + SB, (c) results of SVM + SB + SI, and (d) results of SVM + SB + SI + TF.

Table 5: Quantitative results for land cover classification based on monotemporal data.

Method		Precision	Recall	F1 score	OA	Kappa
SVM + SB + SI + TF	Water	0.9936	0.9563	0.9746	0.8492	0.8089
	High vegetation	0.9671	0.8697	0.9158		
	Low vegetation	0.8255	0.6385	0.7200		
	Crops	0.6895	0.7792	0.7316		
	Built-up	0.7229	0.9640	0.8262		

multitemporal Landsat-8 images and produce a high-quality land cover map for large-scale areas such as the Melbourne metropolitan region. The results were compared with those of other studies to further investigate the advantage of the suggested method for land cover categorization from multitemporal Landsat-8 data. The results of the other studies were taken from the main published articles, but the proposed technique was built using an experimental dataset. For example, Jia et al. [22] applied SVM and maximum likelihood methods for land cover classification from Landsat-8 OLI images in Beijing, China. They used OA and kappa metrics to calculate the accuracy of the proposed methods for land cover mapping. They obtained OA of 91.3% and kappa of 0.89% for SVM and OA of 90.4% and kappa of 0.88% for the maximum likelihood method. Man et al. [40] used time-series Landsat-8 images to classify land cover in Hanoi, Vietnam, based on the ensemble method, which combines the best of some ML classifiers such as multilayer perceptron (MLP), logistic regression (LR), SVM, and XGBoost. They achieved OA and kappa of 83.2% and 0.77% for XGBoost, 82.6% and 0.77% for LR, 82.9% and 0.78% for SVM, 83.1% and 0.78% for MLP, and 84% and 0.79% for ensemble technique, respectively. In contrast, our proposed method, which includes additional input variables and postprocessing technique, achieved higher OA

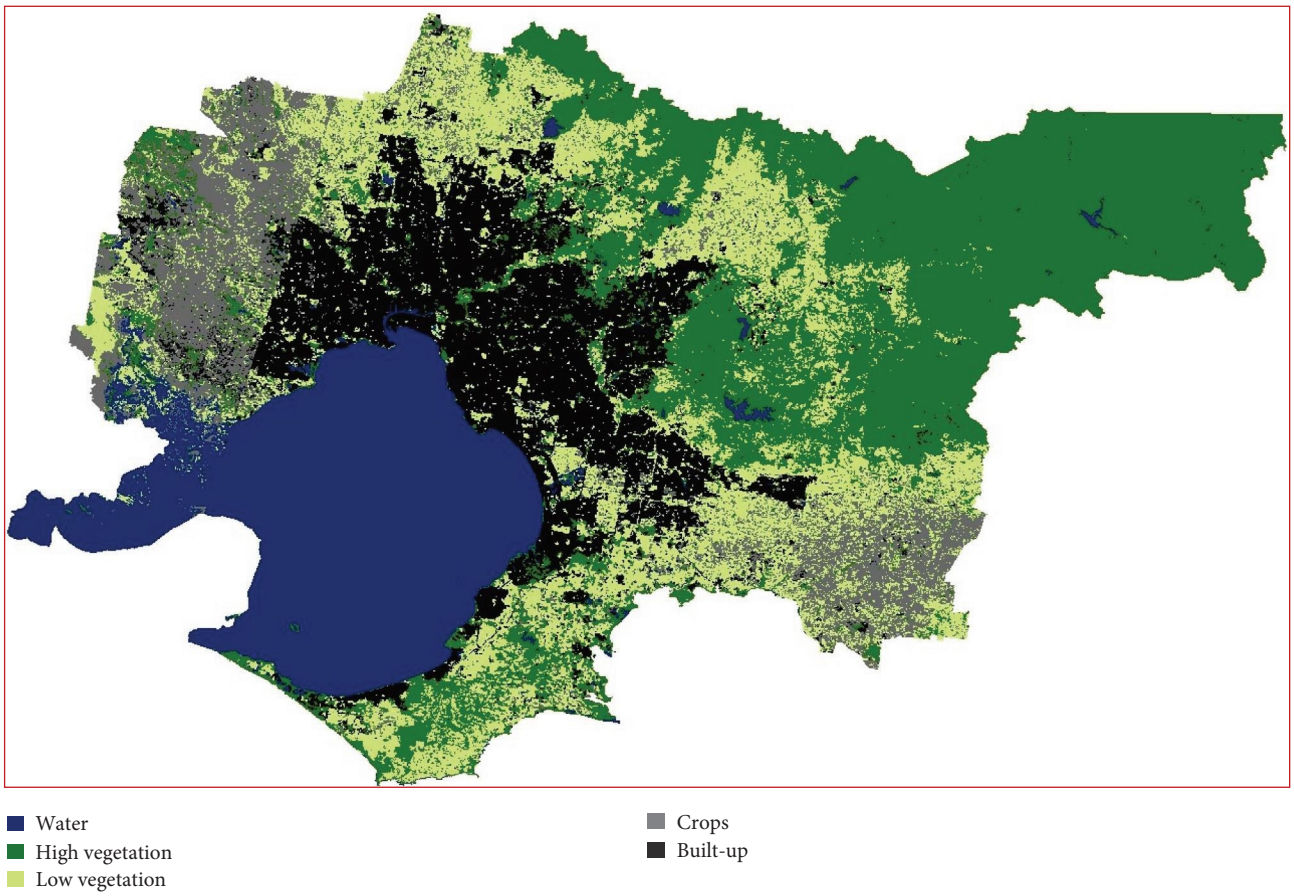

Figure 6: Visualization results of the land cover classification achieved by the proposed model for monotemporal data.

Table 6: Quantitative results of the CART technique for land cover classification.

Method	Classes	Precision	Recall	F1 score	OA	Kappa
CART + SB	Water	0.985	0.9822	0.9836	0.8082	0.7558
	High vegetation	0.9467	0.8311	0.8851		
	Low vegetation	0.5942	0.5467	0.5695		
	Crops	0.581	0.6824	0.6276		
	Built-up	0.7904	0.8883	0.8365		
CART + SB + SI	Water	0.9839	0.9885	0.9862	0.8278	0.7814
	High vegetation	0.9402	0.8733	0.9055		
	Low vegetation	0.6415	0.5795	0.6089		
	Crops	0.6576	0.6928	0.6747		
	Built-up	0.8004	0.9019	0.8481		
CART + SB + SI + TF	Water	0.9853	0.9914	0.9883	0.8524	0.8123
	High vegetation	0.9407	0.8978	0.9187		
	Low vegetation	0.6514	0.6819	0.6663		
	Crops	0.7323	0.7146	0.7234		
	Built-up	0.8441	0.8834	0.8633		

and kappa accuracy of 92.90% and 90.99%, respectively, thereby confirming the effectiveness of the proposed method for land cover classification from multitemporal Landsat-8 images for large-scale areas compared with other studies.

We also applied the model to different regions, such as the ACT, to analyze its performance under various conditions and evaluate its generalizability and transferability across diverse geographic areas (Figure 8). The model achieved an

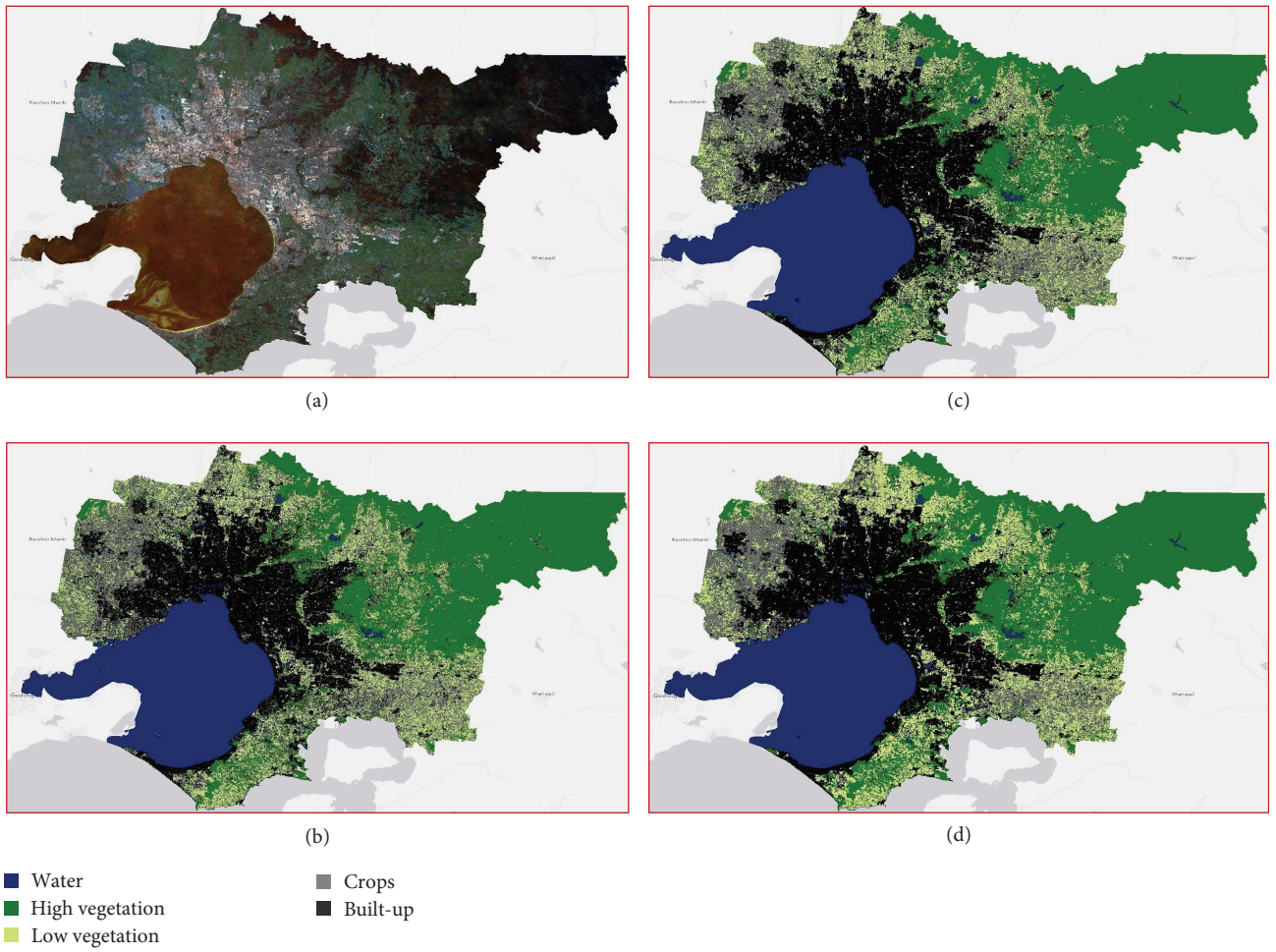

FIGURE 7: Visualization results of the land cover classification achieved by the CART model: (a) original multitemporal Landsat-8 image, (b) results of CART + SB, (c) results of CART + SB + SI, and (d) results of CART + SB + SI + TF.

OA of 87.89% and a kappa of 81.37%, signifying its effectiveness in land cover mapping for this new region. Nevertheless, we observed variations in accuracy among different land cover classes. The model faced challenges in accurately classifying low-vegetation areas (e.g., grasslands) and crops, resulting in lower F1 scores for these categories compared to others. One potential reason for this discrepancy could be related to the phenology of grass and crops, as their spectral characteristics may exhibit significant variations throughout the year.

Further investigation into the optimal timing of data acquisition for distinguishing between these classes could potentially enhance their accuracy. Moreover, the spatial heterogeneity in agricultural practices and growth stages of crops might have contributed to classification difficulties. Incorporating more diverse datasets, such as synthetic aperture radar (SAR) images, could offer complementary information and improve the model's ability to discriminate between land cover classes, especially in challenging environmental conditions. Additionally, testing robust models like deep learning algorithms could prove beneficial in addressing the challenges we observed and enhancing the accuracy of different land cover classes on the maps. Deep learning models, renowned for their capacity to automatically learn intricate patterns and feature representations from data, have exhibited promising results in image classification tasks. Despite these challenges, the model's overall satisfactory performance underscores its generalizability and transferability to diverse geographic areas. As we continue to advance in the field of land cover mapping, future research efforts can focus on investigating the effectiveness of these cutting-edge techniques and addressing temporal- and dataset-related considerations. This endeavor aims to ultimately achieve comprehensive and reliable results for land cover classification across diverse regions.

5. Conclusion

This work provides an effective approach for mapping land cover for a large-scale area like the Melbourne metropolitan and ACT regions, Australia, by using cloud-based image processing technologies such as GEE. On the basis of the SVM model, high-quality land cover classification maps were created from multitemporal Landsat-8 data. To improve the classification results, we added additional variables such as topographic parameters and SIs to the model. Moreover, we utilized a postprocessing technique to remove the salt-and-

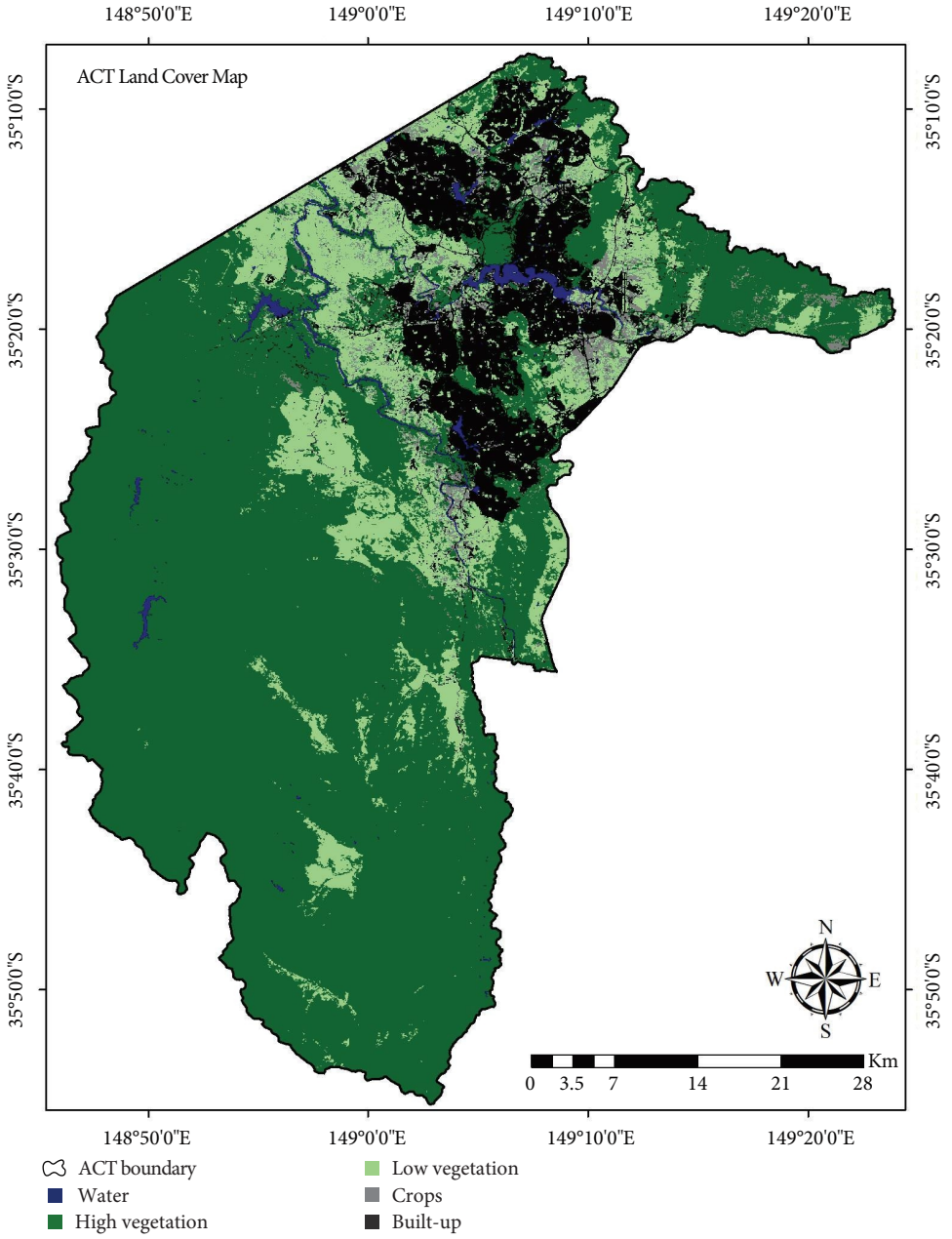

Figure 8: Visualization results of the land cover classification achieved by the model for ACT region.

pepper noise caused by misclassified pixels in supervised classification results and to improve the model's performance. The accuracy of the land cover classification map was assessed visually and numerically. The SVM method could achieve an OA and kappa accuracy of 88.47% and 85.34%, respectively, with all the factors. The OA and kappa also improved to 92.90% and 90.99%, respectively, after the postprocessing technique was implemented. The visualization results showed that the suggested model was capable of producing high-quality land cover maps. The findings showed that multitemporal Landsat-8 data, SBs, SIs, topographic parameters, and postprocessing method were all important in differentiating classes and were helpful in land cover mapping. Furthermore, the SVM model demonstrated satisfactory performance overall, indicating its ability to generalize and transfer successfully to diverse geographic areas. The findings underscore the potential of our method to achieve satisfactory land cover mapping in a wide range of terrestrial settings. Furthermore, GEE, multitemporal data, ML, and modern computer technologies have paved the way for developing a real-time land cover mapping platform.

Data Availability

Open access data have been used in this research.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The study is supported by the Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), University of Technology Sydney, Ministry of Education of the Republic of Korea, the National Research Foundation of Korea (NRF-2023R1A2C1007742), and in part by the Researchers Supporting Project number RSP-2023/14, King Saud University. Open access publishing facilitated by University of Technology Sydney, as part of the Wiley-University of Technology Sydney agreement via the Council of Australian University Librarians.

References

[1] J. Haarpaintner, D. de la Fuente Blanco, F. Enßle et al., "Tropical forest remote sensing services for the democratic republic of congo inside the EU FP7 Recover Project (final results 2000–2012)," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-7/W3, no. 7, pp. 397–402, 2015.

[2] J. Reynolds, K. Wesson, A. L. J. Desbiez, J. M. Ochoa-Quintero, and P. Leimgruber, "Using remote sensing and random forest to assess the conservation status of critical Cerrado habitats in Mato Grosso do Sul, Brazil," *Land*, vol. 5, no. 2, Article ID 12, 2016.

[3] P. Gamba and M. Aldrighi, "SAR data classification of urban areas by means of segmentation techniques and ancillary optical data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 4, pp. 1140–1148, 2012.

[4] A. Braun and V. Hochschild, "A SAR-based index for landscape changes in African savannas," *Remote Sensing*, vol. 9, no. 4, Article ID 359, 2017.

[5] L. L. Coulter, D. A. Stow, Y.-H. Tsai et al., "Classification and assessment of land cover and land use change in southern Ghana using dense stacks of Landsat 7 ETM+ imagery," *Remote Sensing of Environment*, vol. 184, pp. 396–409, 2016.

[6] B. Mack, P. Leinenkugel, C. Kuenzer, and S. Dech, "A semi-automated approach for the generation of a new land use and land cover product for Germany based on Landsat time-series and Lucas in-situ data," *Remote Sensing Letters*, vol. 8, no. 3, pp. 244–253, 2017.

[7] J. Tollefson, "Landsat 8 to the rescue," *Nature News*, vol. 494, no. 7435, Article ID 13, 2013.

[8] C. Gómez, J. C. White, and M. A. Wulder, "Optical remotely sensed time series data for land cover classification: a review," *ISPRS Journal of Photogrammetry Remote Sensing*, vol. 116, pp. 55–72, 2016.

[9] A. Abdollahi and B. Pradhan, "Urban vegetation mapping from aerial imagery using explainable AI (XAI)," *Sensors*, vol. 21, no. 14, Article ID 4738, 2021.

[10] A. Abdollahi and B. Pradhan, "Integrated technique of segmentation and classification methods with connected components analysis for road extraction from orthophoto images," *Expert Systems with Applications*, vol. 176, Article ID 114908, 2021.

[11] A. Abdollahi and M. Yebra, "Forest fuel type classification: review of remote sensing techniques, constraints and future trends," *Journal of Environmental Management*, vol. 342, Article ID 118315, 2023.

[12] W. Mao, D. Lu, L. Hou, X. Liu, and W. Yue, "Comparison of machine-learning methods for urban land-use mapping in Hangzhou City, China," *Remote Sensing*, vol. 12, no. 17, Article ID 2817, 2020.

[13] W. Meng, Z. Xinchang, W. Jiayao, S. Ying, and J. Ge, "Forest resource classification based on random forest and object oriented method," *Acta Geodaetica et Cartographica Sinica*, vol. 49, no. 2, Article ID 235, 2020.

[14] C. Yangbo, Z. Tao, D. Peng, D. Liming, and C. Hua, "Error sources and post processing method for land use/cover change estimation of Dongguan city based on Landsat remote sensing imagery with SVM," *Remote Sensing Technology and Application*, vol. 32, no. 5, pp. 893–903, 2017.

[15] H. Haiyan, H. Jinliang, H. Chunlin, and W. Yunchen, "Retrieve snow depth of North of Xinjiang region from ARMS 2 data based on artificial neural network technology," *Remote Sensing Technology and Application*, vol. 33, no. 2, pp. 241–251, 2018.

[16] S. Morell-Monzó, J. Estornell, and M.-T. Sebastiá-Frasquet, "Comparison of Sentinel-2 and high-resolution imagery for mapping land abandonment in fragmented areas," *Remote Sensing*, vol. 12, no. 12, Article ID 2062, 2020.

[17] A. LaRocque, C. Phiri, B. Leblon, F. Pirotti, K. Connor, and A. Hanson, "Wetland mapping with Landsat 8 OLI, Sentinel-1, ALOS-1 PALSAR, and LiDAR Data in Southern New Brunswick, Canada," *Remote Sensing*, vol. 12, no. 13, Article ID 2095, 2020.

[18] A. Chakhar, D. Ortega-Terol, D. Hernández-López, R. D. Ballesteros, J. F. Ortega, and M. A. Moreno, "Assessing the accuracy of multiple classification algorithms for crop classification using Landsat-8 and Sentinel-2 data," *Remote Sensing*, vol. 12, no. 11, Article ID 1735, 2020.

[19] S. Fragou, K. Kalogeropoulos, N. Stathopoulos et al., "Quantifying land cover changes in a mediterranean environment using Landsat TM and support vector machines," *Forests*, vol. 11, no. 7, p. 750, 2020.

[20] D. Zhao, H. Gu, and Y. Jia, "Comparison of machine learning method in object-based image classification," *Remote Sensing*, vol. 41, pp. 181–186, 2016.

[21] A. M. Abdi, "Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data," *GIScience Remote Sensing*, vol. 57, no. 1, pp. 1–20, 2020.

[22] K. Jia, X. Wei, X. Gu, Y. Yao, X. Xie, and B. Li, "Land cover classification using Landsat 8 operational land imager data in Beijing, China," *Geocarto International*, vol. 29, no. 8, pp. 941–951, 2014.

[23] E. Adam, O. Mutanga, J. Odindi, and E. M. Abdel-Rahman, "Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers," *International Journal of Remote Sensing*, vol. 35, no. 10, pp. 3440–3458, 2014.

[24] A. M. Prasad, L. R. Iverson, and A. Liaw, "Newer classification and regression tree techniques: bagging and random forests for ecological prediction," *Ecosystems*, vol. 9, no. 2, pp. 181–199, 2006.

[25] L. Naidoo, M. A. Cho, R. Mathieu, and G. Asner, "Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a random forest data mining environment," *ISPRS journal of Photogrammetry Remote Sensing*, vol. 69, pp. 167–179, 2012.

[26] A. R. Shahtahmassebi, Y. Lin, L. Lin et al., "Reconstructing historical land cover type and complexity by synergistic use of

Landsat Multispectral Scanner and CORONA," *Remote Sensing*, vol. 9, no. 7, Article ID 682, 2017.

[27] H. Zhang, A. Eziz, J. Xiao et al., "High-resolution vegetation mapping using eXtreme gradient boosting based on extensive features," *Remote Sensing*, vol. 11, no. 12, Article ID 1505, 2019.

[28] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google earth engine: planetary-scale geospatial analysis for everyone," *Remote Sensing of Environment*, vol. 202, pp. 18–27, 2017.

[29] Y. Du, Y. Zhang, F. Ling, Q. Wang, W. Li, and X. Li, "Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band," *Remote Sensing*, vol. 8, no. 4, Article ID 354, 2016.

[30] R. Xu, J. Liu, and J. Xu, "Extraction of high-precision urban impervious surfaces from sentinel-2 multispectral imagery via modified linear spectral mixture analysis," *Sensors (Basel, Switzerland)*, vol. 18, no. 9, Article ID 2873, 2018.

[31] P. D'Odorico, A. Gonsamo, A. Damm, and M. E. Schaepman, "Experimental evaluation of Sentinel-2 spectral response functions for NDVI time-series continuity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 3, pp. 1336–1348, 2013.

[32] P. Gong, B. Chen, X. Li et al., "Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018," *Science Bulletin*, vol. 65, no. 3, pp. 182–187, 2020.

[33] W. Yao-Nan and Y. Xiao-Fang, "SVM approximate-based internal model control strategy," *Acta Automatica Sinica*, vol. 34, no. 2, pp. 172–179, 2008.

[34] C. Xu, F. Dai, X. Xu, and Y. H. Lee, "GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China," *Geomorphology*, vol. 145, pp. 70–80, 2012.

[35] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 1999.

[36] D. Shi and X. Yang, "Support vector machines for land cover mapping from remote sensor imagery," in *Monitoring and Modeling of Global Changes: A Geomatics Perspective*, J. Li and X. Yang, Eds., pp. 265–279, Springer, Dordrecht, 2015.

[37] N. Ghasemkhani, S. S. Vayghan, A. Abdollahi, B. Pradhan, and A. Alamri, "Urban development modeling using integrated fuzzy systems, ordered weighted averaging (OWA), and geospatial techniques," *Sustainability*, vol. 12, no. 3, Article ID 809, 2020.

[38] A. Abdollahi and B. Pradhan, "Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model," *The Science of the Total Environment*, vol. 879, Article ID 163004, 2023.

[39] L. El Mendili, A. Puissant, and M. Chougrad, "Towards a multi-temporal deep learning approach for mapping urban fabric using sentinel 2 images," *Remote Sensing*, vol. 12, no. 3, Article ID 423, 2020.

[40] C. D. Man, T. T. Nguyen, H. Q. Bui, K. Lasko, and T. N. T. Nguyen, "Improvement of land-cover classification over frequently cloud-covered areas using Landsat 8 time-series composites and an ensemble of supervised classifiers," *International Journal of Remote Sensing*, vol. 39, no. 4, pp. 1243–1255, 2018.