Optimal Land-cover Classification Feature Selection in Arid Areas based on Sentinel-2 Imagery and Spectral Indices

Mohammed Saeed[1*], Asmala Ahmad[2], Othman Mohd[3]

Faculty of Communication and Information Technology, Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia[1, 2, 3]
Geological Survey and Mineral Resources Authority, Sana'a, Yemen[1]

Abstract—**Adding spectral indices to Sentinel-2 spectral bands to improve land-cover (LC) classification with limited sample size can affect the accuracy due to the curse of dimensionality. In this study, we compared the performance metrics of Random Forest (RF) classifier with three different combinations of features for land cover classification in an urban arid area. The first combination used the ten Sentinel-2 bands with 10 and 20 m spatial resolution. The second combination consisted of the first combination in addition to five common spectral indices (15 features). The third combination represented the best output of features in terms of performance metrics after applying recursive feature elimination (RFE) for the second combination. The results showed that applying RFE reduced the number of features in combination 2 from 15 to 8 and the average F_1-score indicator increased by nearly 8 and 6 percent in comparison with using the other two combinations respectively. The findings of this study confirmed the importance of feature selection in improving LC classification accuracy in arid areas through removing the redundant variable when using limited sample size and using spectral indices with spectral bands, respectively.**

Keywords—Feature selection; land cover; sentinel-2; arid areas; random forest; accuracy

I. INTRODUCTION

Extracted information from satellite imagery regarding LC and its changes are essential in many applications. They are used as inputs for many models as hydrological models [1], ecosystem modelling [2] and land surface modelling [3]. Thus, the accuracy of LC is critical for these products, as it affects the final results of these models [4].

Adding auxiliary features to improve the classification of LC is a common practice in remote sensing community such as in Landsat imagery data [5], [6]. These auxiliary features include different spectral indices, topographic data, texture, and biophysical parameters [7], [8]. Using too many features to improve LC classification requires increasing the training samples size [9] to overcome the curse of dimensionality which negatively affects the classification accuracy and increases processing time [10]. In most cases, increasing training samples is not cost-effective, and this should be met by selecting only the most relevant features to achieve the optimal classification accuracy. Feature selection is very important in LC classification to overcome the high-dimensional data to increase class separation and compensate for the limited samples used for training classification models [9]. Moreover, feature selection removes the irrelevant and redundant variables which eventually helps in reducing the training data, decreasing the processing time and decrease the requirements of the data storage [11], [12]. In addition, feature selection documented to improve prediction performance and making data more interpretable [10].

According to [13], choosing a feature selection method subjects to various consideration such as stability, simplicity, requirements of computation, accuracy and the number of reduced features. In remote sensing applications, various feature selection methods were used in different spatial areas and for different purposes. Over a global LC classification, Relieff and max-min-associated methods were useful in decreasing computation time [14]. With sentinel-2 imagery, a comparative study among various feature selection methods concluded that similarity-based methods are the best in terms of F_1-score and the optimal features number for mapping landscapes infested by the Parthenium weed in South Africa, while wrapper methods were more accurate but with larger number of the selected features [15]. In northern Germany, grouped forward feature selection helped in data interpretation and reduced processing time in crop mapping. In spite of the several methods which were developed for selecting features, Reference [16] recommended using RFE, a wrapper method, in combination with RF for feature selection due to its stability and ability to improve classification. This method also proved to be efficient in improving accuracy in both the regression and classification processes by selecting the most relevant features [17], [18].

Arid regions are different from other spatial environments as they are dominant with less precipitation, dry climate, and scattered vegetation. Thus, ecosystems in urban arid areas are fragile [19]. These systems are not stable, and their change is rapid [20]. All these conditions are being captured by satellite sensors and translated into image pixels; therefore, need special attentions since these tend to affect the accuracy of LC classification to be produced later.

Since Sentinel-2 imagery was released in 2015, it has been widely used in producing LC maps due to its high spatial resolution, its temporal resolution (5 days), and its spectral wavelength ranging from visible to near-infrared, which helped to map and distinguishing LC classes [21].

---

*Corresponding Author.

In urban arid areas, there is no LC classification model based on the Sentinel-2 imagery for the selected study area in the Arabian arid Peninsula. Therefore, the objective of this study was to develop an optimal LC classification model for this area with cost-effective samples. This will include using RFE with RF classifier to choose the optimal relevant features from the combination of spectral bands and the most common indices. In addition, the effect of feature selection on processing time during training the model and prediction was investigated.

The paper is organised as follows: Section II describes the study area, data, and the methodology. In Sections III and IV, results and discussion are presented, respectively. Finally, Section V represents the conclusion of this study.

II. MATERIALS AND METHODS

A. Study Site and Data

The study area was chosen as part of the tile number T38RPN (Fig. 1) from sentinel-2 satellite imagery which covers the metropolitan city of Riyadh, the capital of Kingdom of Saudia Arabia as a urban arid area. The image was selected on 4 July 2022 with ID: L1C_T38RPN_A027817_ 20220704T073152 when the cloud is minimum, and the selected part was considered to represent the variations in LC classes. While Sentinel-2 has 13 bands as shown in, only 10 bands were used in this study and the 60-meter spatial resolution bands related to atmospheric and cloud detection were dropped.
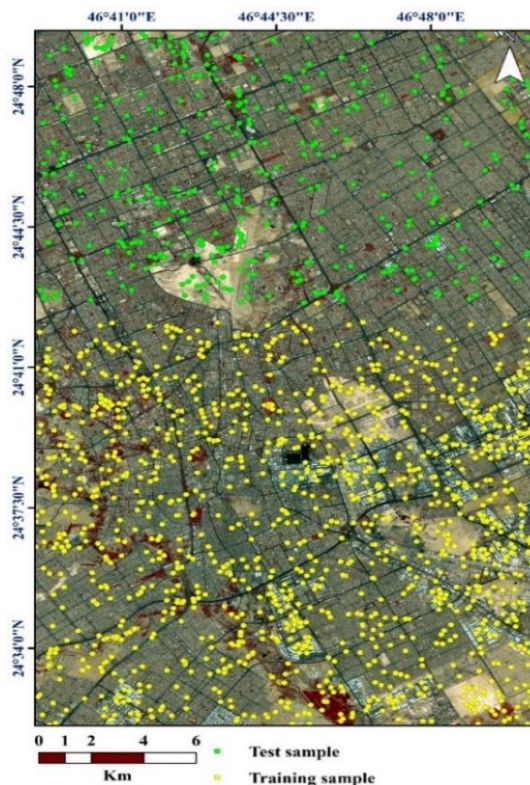

Fig. 1. Location of the study area.

B. Data Preprocessing and Preparation

In order to achieve the optimal accuracy, three steps have been performed on the raw image before clipping to the study area. First, atmospheric and topographic correction were carried out to convert the digital numbers to surface reflectance values using the FORCE algorithm [20] and to remove the effects of shadow, respectively. This preprocess step is initial and proved to contribute in improving LC accuracy. The SRTM digital elevation model (DEM) from EarthExplorer was used for topographic correction. Second, downscaling the 20-meter spatial resolution bands to 10 metre using nearest neighbor technique, which proved to be more accurate than other techniques in terms of producing LC classification accuracy [21]. After that, the image was cropped to the study area shapefile using QGIS software, version 3.

C. Classification System and Sampling

The selection of LC classes was based on the basis that confirms the inclusion of the main land types in the study area with reference to previous studies [22], [23]. In this study, the urban class has been divided into three categories: roads, industrial and building where spectral differences are unique. Table I shows the classes with their representative numbers in the selected study area.

The stratified random sampling method was used to collect training and testing samples. All samples were collected based on the per-pixel as a classification unit to avoid spatial-autocorrelation and reduce redundant data. The choice of samples was based on visual interpretation on the high spatial resolution Google Earth maps with intensive field work for validation. The number of training samples was determined to be in the range of 10–30 times the number of bands used for classification [24]. The test samples were 30% of the total samples and independent of the training samples. Fig. 1 and Table I show the distribution of training, testing samples, and their numbers in the study area, respectively.

D. Spectral Indices

In this urban arid study area, the effect of adding the following spectral indices on accuracy was investigated:

- The normalized difference vegetation index (NDVI) which is used as a monitoring and measuring index for vegetation cover from satellite imagery.

- The normalised difference built-up index (NDBI) is used to distinguish built surfaces, which receive positive values, from bare soils.

- The modified normalised difference water index (MNDWI) was proposed to detect superficial water. However, due to the relation between SWIR and wetness in soils, it can be also used to detect water in surfaces of vegetation or soil.

- The bare soil index (BSI) was proposed to enhance differentiation between bare and built-up lands.

- The soil adjusted vegetation index (SAVI): this index way to fit NDVI index to background average reflectance and minimises shadow effects.

TABLE I.   LAND COVER CLASSES WITH THEIR TRAINING AND TESTING NUMBER AND DESCRIPTION

Classes	Sampling	
	Training (70%)	*Testing (30%)*
Vegetation	452	138
Roads	478	144
Bare land	474	142
Built-up	534	164
Industrial	530	160
Water	146	44
Total	2614	792

E. Classification Process and Evaluation

In order to investigate the effect of dimensionality reduction on LC classification accuracy in this study, we compared the performance of RF with three different combinations of features. The first combination consisted of the original ten spectral bands. The second combination consisted of the same features of the first combination in addition to the fifth spectral indices already mentioned in the previous subsection 2.5. The third combination represented the subset of the features which achieved the best performance metrics after applying the RFE selection method on the second combination. We referred to these combinations as model-1, model-2 and model-3 in the whole paper.

In each model, RF was used for classification due to its accurate results with less time, less sensitivity to overfitting, and because it requires few internal parameters to be tuned. RF is one of the most supervised ML algorithms widely used in both regression and classification and it can work with continuous and categorical data [26]. It belongs to the family of ensemble learning classifiers which depends on the bagging mechanism. The ntree and mtry are the most two internal parameters in the RF classifier. Each tree in FR model acts as a decision in the classification or regression process and the number of these decision trees is known as ntree and determined by the selected features from the user. In this study, the ntree was set to 500 as recommended by [27]. The mtry parameter refers to the predictors number that are randomly sampled when creating the trees at each split. In this study, mtry was set to the number of square root of the variables used as inputs for classification in each model [28].

The validation process for each model was carried out, using a 10-fold cross validation technique to avoid bias in results and conclusions. In this technique, data set is divided into 10 subsets. Next, a model is trained using a subset formed by combining these nine subsets and tested using the remaining subset. This is done 10 times each using a different subset as a test set and calculating the test set error.

The evaluation for all models was based on the performance metrics represented by the overall accuracy (OA), the user's (UA), and producer's (PA) accuracies, which were calculated from the confusion matrix. In addition, F_1-score was calculated as a balance accuracy measurement [29] and used in this study as the main index for comparing the models.

In order to explore the contribution of each feature to the improvement of the LC accuracy for all models, the built-in variable importance property of the RF classifier (randomForest package) was analysed. Thus, a useful reference can be provided to choose the appropriate features as input variables in other studies in the selected study area.

The last step of evaluation was the computational time analysis which included comparing the average processing time for the three models with their different number of features. The average of 10 running times was used in this study for processing time of the models during training and when the models were used to predict the whole image of the study area. All analyses were carried out using R programming language (version 3.6.1). We used a laptop with Intel® Core™ i7-7700HQ CPU @ 2.80GHz × 8 and 32 GiB memory in Ubuntu 20.04.5 LTS operating system.

III.   RESULTS

A. Model Evaluation by Overall, User's, Producer's Accuracy Matrices and Variable Importance

1) Model-1: Table II shows the confusion matrix when the classification model used only the original 10 spectral bands. The overall accuracy of the model was 81%. Generally speaking, the greatest misclassification has occurred between vegetation and built-up classes, on the one hand, and between industrial and bare classes, on the other hand.

In this model and as shown in Fig. 2, the PA and UA are variated in values among the classes. The PA for the built-up class was the maximum with 98.8 percent, while the water class had the minimum value with 22.7 percent. Regarding the UA, the industrial and built-up classes were the highest and lowest values with 95.8 and 70.4 respectively.

Fig. 3 shows the importance of the features and their contributions in the accuracy. It is clearly that b3, b4, b12, b8 and b2 are the first five bands that contributed to achieving most of classification accuracy improvement while b6 was the less important regarding contribution in classification accuracy.

2) Model-2: Table III shows the confusion matrix when the classification model used the tenth spectral bands with the five spectral indices. The overall accuracy of this model increased by nearly 2 % compared to the previous model.

TABLE II.   CONFUSION MATRIX OF MODEL-1

Classes		Predicted Values						Total
Actual values	Vegetation	82	10	2	42	0	2	138
	Roads	2	136	0	6	0	0	144
	Bare land	0	5	134	0	3	0	142
	Built-up	0	2	0	162	0	0	164
	Industrial	0	0	42	4	114	0	160
	Water	6	10	0	16	2	10	44
	Total	90	163	178	230	119	12	792
Overall accuracy 80.56 %								

TABLE III.   CONFUSION MATRIX OF MODEL-2

Classes		Predicted values						Total
Actual values	Vegetation	90	7	3	38	0	0	138
	Roads	0	134	0	10	0	0	144
	Bare land	0	4	136	0	2	0	142
	Built-up	0	2	2	160	0	0	164
	Industrial	0	0	32	2	126	0	160
	Water	4	11	0	19	0	10	44
	Total	94	158	173	229	128	10	792
Overall accuracy 82.83 %								

Fig. 4.   Producer's and User's accuracy in model-2.
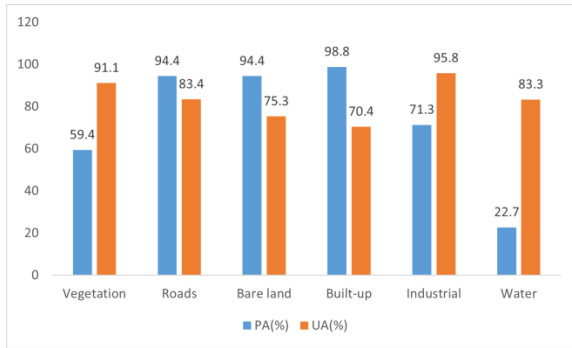
Fig. 2.   Producer's and User's accuracy in model-1.
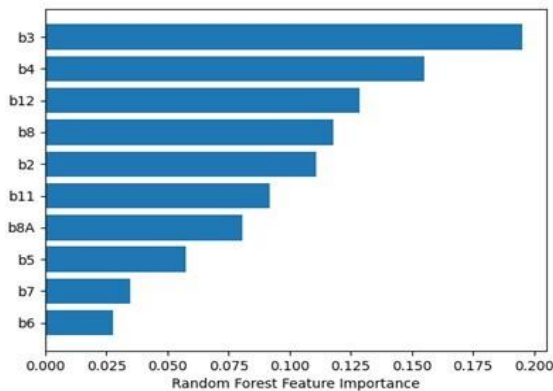
Fig. 3.   Variable importance in model-1.

This improvement in overall accuracy can be interpreted by the contribution of the indices in improving the per-class accuracy.

As shown in Fig. 4, the PA and UA accuracy in this model are variated. The highest and lowest PA values were registered for the built-up and water classes, respectively, while the highest and lowest UA values were registered for water and industrial classes, respectively.

The variable importance in this model is shown in Fig. 5. From the first five features, two spectral indices: SAVI and BSI are the most important in classification accuracy. It is noticeable that NDVI has a medium importance in improving accuracy, while MNDWI and NDBI have less importance in improving classification accuracy.
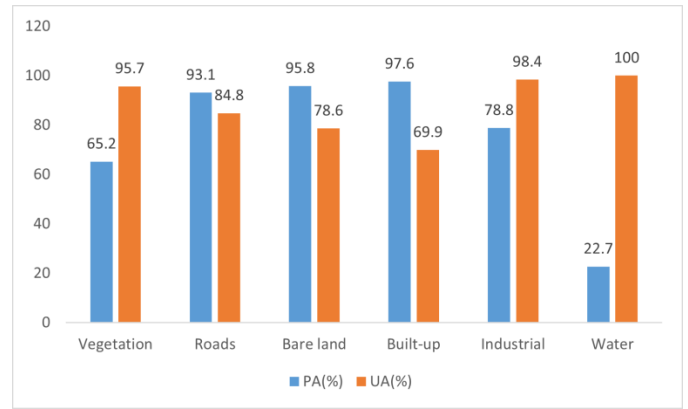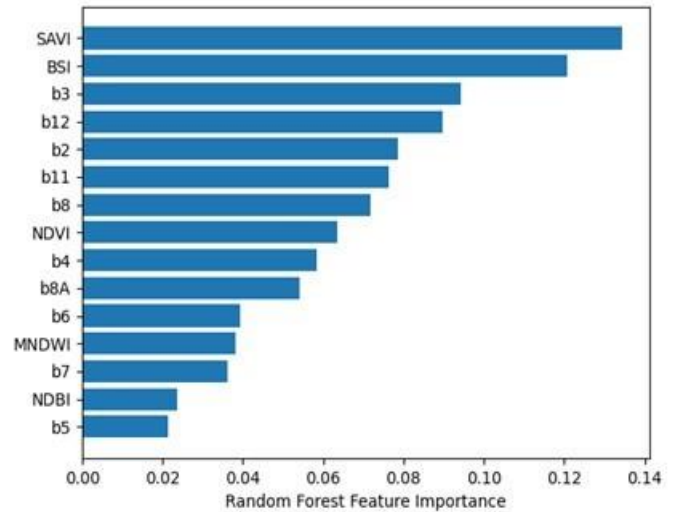

Fig. 5.   Variable importance in model-2.

3) Model-3: The confusion matrix after applying the RFE to subset the feature with the best accuracy is shown in Table IV. The best accuracy derived from this model was 85.98 % using only eight features: six spectral bands and two spectral indices.

Comparing the previous two models, the accuracy in this model increased by nearly 5 and 3 percent, respectively. Most of the accuracy improvement was in both vegetation, industrial and water classes where the number of the correct instances increased by 8, 4 and 9 respectively in comparison with the same classes in the previous model.

In this model, the PA ranged between 43.2 % for the water class and 97.6 % for the built-up class. In terms of UA, the water class ranked the best, while the built-up class ranked the lowest (Fig. 6).

The first five feature importance in this model included four spectral bands: b3, b12, b2, and b8 and only one spectral index: BSI as shown in Fig. 7. After applying RFE, the feature numbers decreased nearly to half in comparison with the previous model.

TABLE IV.   CONFUSION MATRIX OF MODEL-3

Classes		Predicted values						Total
Actual values	Vegetation	98	7	5	28	0	0	138
	Roads	0	137	5	2	0	0	144
	Bare land	0	3	137	0	2	0	142
	Built-up	2	2	0	160	0	0	164
	Industrial	0	0	23	7	130	0	160
	Water	3	8	0	14	0	19	44
	Total	103	158	169	213	133	19	792
Overall accuracy 85.98 %								
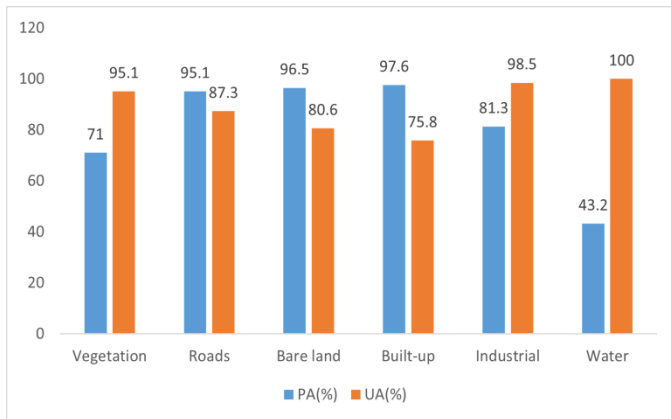
Fig. 6.   Producer's and User's accuracy in model-3.
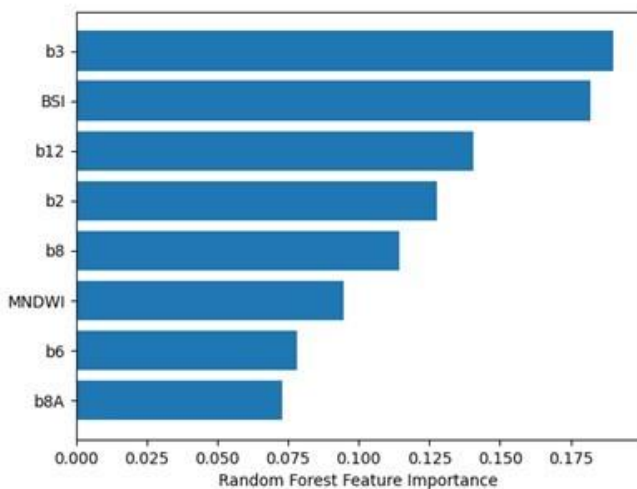

Fig. 7.   Variable importance in model-3.

B. Model Evaluation by F_1-score Accuracy Metric

Fig. 8 shows the value of F_1-score accuracy for the three models compared in this study. It is very clear that the F_1-score value for most of the classes was increased from model-1 to model-3. This increase was noticeable in most of the classes after adding the spectral indices in model-2 with reference to the initial model-1 where spectral bands were only used. For instance, F_1-score in vegetation and industrial classes increased by nearly 6 % in model-2 when compared with model-1.
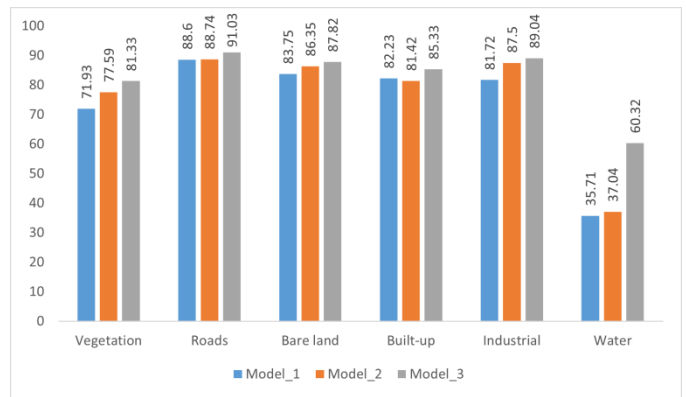

Fig. 8. F_1-score accuracy in the three models.

After applying the RFE to choose the optimal features, the F_1-score increased with different percentage between classes when compared to the values in model-2. The water class achieved the highest increase by 23 %, while the bare land class achieved the lowest increase by 1.67 %.

C. Computation Time Evaluation

Fig. 9 compares the average processing time spent for training and predict in each model aligned with the number of features used. The model-3 achieved the best rank in terms of the training and prediction processing time, where its average reached 1.56 and 1.83 mins respectively. This model used the lowest number of features (6 spectral bands plus 2 spectral indices). On the other hand, the highest training and prediction processing time was associated with model-2, which has the largest number of features (10 spectral bands plus 5 spectral indices).
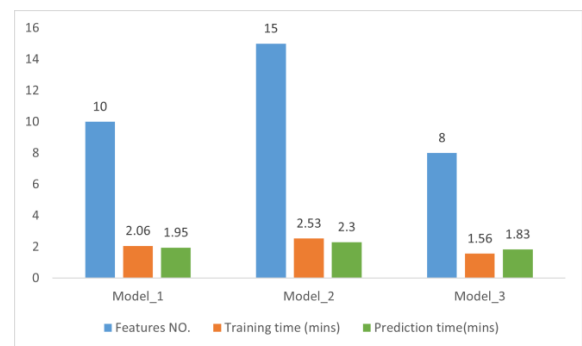

Fig. 9.   Feature number, training, and prediction time in the models.

IV.   DISCUSSION

This study aimed to explore the effect of feature selection on LC classification accuracy and processing time in arid areas aligned with limited sample size. In terms of the overall accuracy, average PA, UA and F_1-score as shown in Fig. 10, there was an increase in accuracy after adding the spectral indices to the spectral bands. In addition, there is a noticeable increase in the accuracy of all these metrics in model-3 after applying the RFE feature selection technique.
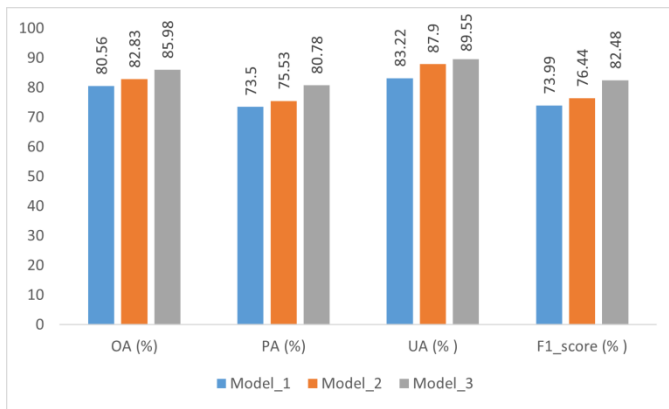
Fig. 10. Average of accuracy indicators for the three models.

Adding spectral indices proved to be effective in improving LC classification accuracy in model-2 in this study by increasing the separability between the individual classes in comparison with model-1. In this study, the SAVI and BSI indices were the two most important features in terms of improving classification accuracy in model-2. The fact of improving accuracy through adding indices is a common practice in many other studies such as in [25]–[27].

Despite the added value of indices in improving LC classification accuracy in model-2, the application of feature selection in model-3 proved to be more effective in improving all the performance metrics in model-3 and in decreasing the processing time without the need for increasing sample size. This could be interpreted by the importance of applying feature selection technique in removing redundant features which affect both the accuracy and processing time [28]. Many studies in terms of improving LC classification recommended applying feature selection methods to select the optimal relevant features and reduce the processing time [29].

The low accuracy of model-2 in comparison with model-3 indicates that the curse of dimensionality can affect the classification accuracy. Previous studies showed that increasing the number of features can lead to complexity by increasing the processing time and decreasing the potential accuracy of the model [30].

The application of RFE in combination with the property of RF variable importance in this study helped in determining the input features and their contribution in producing the optimal classification accuracy in the study area. This subset of the relevant features is a common appropriate approach for building robust learning models [31].

V. CONCLUSION

With a limited sample size for LC classification, adding spectral indices to improve the classification accuracy is not an ideal solution, as shown in this study. The feature selection techniques proved to overcome the limited size of samples by choosing the relevant features that increase the classes separability. In urban arid areas, the RFE technique decreased the features from 15 to 8 with best F_1-score average accuracy (82.48%) in comparison with the case when only spectral features were used in model-1 (73.99%) or when the spectral bands and indices were used in model-2 (76.44%).

Furthermore, less training and prediction processing time was achieved after applying RFE (1.56 and 1.83 min) when comparing with values of model-1 (2.06 and 1.95 min) and with values in model-2 (2.53 and 2.3 min). The combination of the spectral bands: b2, b3, b6, b8, b8a, and b12 with the spectral indices: BSI and MNDWI represent the optimal variables for LC classification in terms of accuracy and computation time in this geographic study area.

The results of this study showed that feature selection is useful in reducing the dimensionality of spectral bands of Sentinel-2 and the spectral indices as well. This refers that not all indices can contribute to improving classification accuracy when sample size is limited.

Other feature selection techniques are recommended to be explored and compared in alignment with the other machine learning classifiers in urban arid areas. In addition, more multitemporal images for different seasons can be investigated to overcome the single image used in this study.

REFERENCES

[1] C. Chirachawala, S. Shrestha, M. S. Babel, S. G. Virdis, and S. Wichakul, "Evaluation of global land use/land cover products for hydrologic simulation in the Upper Yom River Basin, Thailand," Sci. Total Environ., vol. 708, p. 135148, 2020.

[2] W. Chen, H. Zhao, J. Li, L. Zhu, Z. Wang, and J. Zeng, "Land use transitions and the associated impacts on ecosystem services in the Middle Reaches of the Yangtze River Economic Belt in China based on the geo-informatic Tupu method," Sci. Total Environ., vol. 701, p. 134690, 2020.

[3] K. Hibbard et al., "Research priorities in land use and land-cover change for the Earth system and integrated assessment modelling," Int. J. Climatol., vol. 30, no. 13, pp. 2118–2128, 2010.

[4] C. Xu, "Issues influencing accuracy of hydrological modeling in a changing environment," Water Sci. Eng., vol. 14, no. 2, pp. 167–170, 2021.

[5] P. Hurskainen, H. Adhikari, M. Siljander, P. Pellikka, and A. Hemp, "Auxiliary datasets improve accuracy of object-based land use/land cover classification in heterogeneous savanna landscapes," Remote Sens. Environ., vol. 233, p. 111354, 2019.

[6] L. Qu, Z. Chen, M. Li, J. Zhi, and H. Wang, "Accuracy improvements to pixel-based and object-based lulc classification with auxiliary datasets from Google Earth engine," Remote Sens., vol. 13, no. 3, p. 453, 2021.

[7] C. Hong, X. Jin, J. Ren, Z. Gu, and Y. Zhou, "Satellite data indicates multidimensional variation of agricultural production in land consolidation area," Sci. Total Environ., vol. 653, pp. 735–747, 2019.

[8] I. Klein, U. Gessner, A. J. Dietz, and C. Kuenzer, "Global WaterPack–A 250 m resolution dataset revealing the daily dynamics of global inland water bodies," Remote Sens. Environ., vol. 198, pp. 345–362, 2017.

[9] Y. Huang, C. Zhao, H. Yang, X. Song, J. Chen, and Z. Li, "Feature selection solution with high dimensionality and low-sample size for land cover classification in object-based image analysis," Remote Sens., vol. 9, no. 9, p. 939, 2017.

[10] S. Georganos et al., "Less is more: Optimizing classification performance through feature selection in a very-high-resolution remote sensing object-based urban application," GIScience Remote Sens., vol. 55, no. 2, pp. 221–242, 2018.

[11] C. Xu et al., "A Comprehensive Comparison of Machine Learning and Feature Selection Methods for Maize Biomass Estimation Using

Sentinel-1 SAR, Sentinel-2 Vegetation Indices, and Biophysical Variables," Remote Sens., vol. 14, no. 16, p. 4083, 2022.

[12] D. Dobrinić, M. Gašparović, and D. Medak, "Evaluation of Feature Selection Methods for Vegetation Mapping Using Multitemporal Sentinel Imagery," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., vol. 43, pp. 485–491, 2022.

[13] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," Comput. Electr. Eng., vol. 40, no. 1, pp. 16–28, 2014.

[14] L. Yu, H. Fu, B. Wu, N. Clinton, and P. Gong, "Exploring the potential role of feature selection in global land-cover mapping," Int. J. Remote Sens., vol. 37, no. 23, pp. 5491–5504, 2016.

[15] Z. Kiala, O. Mutanga, J. Odindi, and K. Peerbhay, "Feature selection on sentinel-2 multispectral imagery for mapping a landscape infested by parthenium weed," Remote Sens., vol. 11, no. 16, p. 1892, 2019.

[16] X. Fan et al., "Sentinel-2 Images Based Modeling of Grassland Above-Ground Biomass Using Random Forest Algorithm: A Case Study on the Tibetan Plateau," Remote Sens., vol. 14, no. 21, p. 5321, 2022.

[17] R. R. Pullanagari, G. Kereszturi, and I. Yule, "Integrating airborne hyperspectral, topographic, and soil data for estimating pasture quality using recursive feature elimination with random forest regression," Remote Sens., vol. 10, no. 7, p. 1117, 2018.

[18] G. Ewa, F. David, and O. Katarzyna, "Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the Polish Carpathians," Remote Sens. Environ., vol. 251, no. 112103, p. 7, 2020.

[19] H. Xie, Y. Zhang, Z. Wu, and T. Lv, "A bibliometric analysis on land degradation: Current status, development, and future directions," Land, vol. 9, no. 1, p. 28, 2020.

[20] M. Lin and C. Chen, "Application of fuzzy models for the monitoring of ecologically sensitive ecosystems in a dynamic semi-arid landscape from satellite imagery," Eng. Comput., vol. 27, no. 1, pp. 5–19, Jan. 2010, doi: 10.1108/02644401011008504.

[21] M. Drusch et al., "Sentinel-2: ESA's optical high-resolution mission for GMES operational services," Remote Sens. Environ., vol. 120, pp. 25–36, 2012.

[22] A. F. Alqurashi, L. Kumar, and P. Sinha, "Urban land cover change modelling using time-series satellite images: A case study of urban growth in five cities of Saudi Arabia," Remote Sens., vol. 8, no. 10, p. 838, 2016.

[23] M. Rahman and R. Planning, "Land use and land cover changes and urban sprawl in Riyadh, Saudi Arabia: An analysis using multi-temporal Landsat data and Shannon's Entropy Index," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., vol. 41, pp. 1017–1021, 2016.

[24] C. Li, J. Wang, L. Wang, L. Hu, and P. Gong, "Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery," Remote Sens., vol. 6, no. 2, pp. 964–983, 2014.

[25] P. Ettehadi Osgouei, S. Kaya, E. Sertel, and U. Alganci, "Separating built-up areas from bare land in mediterranean cities using Sentinel-2A imagery," Remote Sens., vol. 11, no. 3, p. 345, 2019.

[26] G. Kuc and J. Chormański, "Sentinel-2 imagery for mapping and monitoring imperviousness in urban areas," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., vol. 42, pp. 43–47, 2019.

[27] B. Kavhu, Z. E. Mashimbye, and L. Luvuno, "Climate-based regionalization and inclusion of spectral indices for enhancing transboundary land-use/cover classification using deep learning and machine learning," Remote Sens., vol. 13, no. 24, p. 5054, 2021.

[28] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," presented at the 2014 science and information conference, 2014, pp. 372–378.

[29] M. Kganyago, J. Odindi, C. Adjorlolo, and P. Mhangara, "Selecting a subset of spectral bands for mapping invasive alien plants: a case of discriminating Parthenium hysterophorus using field spectroscopy data," Int. J. Remote Sens., vol. 38, no. 20, pp. 5608–5625, 2017.

[30] P. Gong and P. Howarth, "An assessment of some factors influencing multispectral land-cover classification," Photogramm. Eng. Remote Sens., vol. 56, no. 5, pp. 597–603, 1990.

[31] H. Fei et al., "Cotton classification method at the county scale based on multi-features and Random Forest feature selection algorithm and classifier," Remote Sens., vol. 14, no. 4, p. 829, 2022.