EVALUATION OF FEATURE SELECTION METHODS FOR VEGETATION MAPPING USING MULTITEMPORAL SENTINEL IMAGERY

D. Dobrinić [1] *, M. Gašparović [2], D. Medak [1]

[1] Faculty of Geodesy, Chair of Geoinformatics, University of Zagreb, 10000 Zagreb, Croatia - (ddobrinic, dmedak)@geof.unizg.hr
[2] Faculty of Geodesy, Chair of Photogrammetry and Remote Sensing, University of Zagreb, 10000 Zagreb, Croatia - mgasparovic@geof.unizg.hr

Commission III, WG III/6

KEY WORDS: CORINE, Random Forest, SAR, Sentinel-1, Sentinel-2, Variable Selection, Vegetation Mapping.

ABSTRACT:

With the recent advances in remote sensing technologies for Earth observation (EO), many different remote sensors (e.g., optical, radar) collect data with distinctive properties. EO data have been employed to monitor croplands and forested areas, oceans and seas, urban settlements, and natural hazards. The spectral, spatial, and temporal resolutions of remote sensors have been continuously improving, making geospatial monitoring more accurate and comprehensive than ever before. To tackle this issue, various variable selection methods (e.g., filter, wrapper, and embedded methods) are nowadays used to reduce data complexity, and hence improve classification accuracy. Therefore, the goal of this research was twofold. Firstly, to assess the performance of the random forest (RF) classifier in a large heterogeneous landscape with diverse land-cover categories using multi-seasonal Sentinel imagery (i.e., Sentinel-1; S1 and Sentinel-2; S2) and ancillary data. Secondly, to compare RF variable selection methods to identify a subset of predictor variables that will be included in a final, simpler model. Using mean decrease accuracy (MDA) as a feature selection (FS) method, an original dataset was reduced from 114 to 34 input features, and its classification performance outperformed all-feature (114 features) and band-only (36 features) model with an OA of 90.91%. The most pertinent input features for vegetation mapping were S2 spectral bands (14 features), followed by the spectral indices derived from S2, texture features, and S1 bands. This research improved vegetation mapping by integrating radar and optical imagery, especially after applying FS methods which removed redundant and noisy features from the original dataset. Future research should address additional feature selection methods (i.e., filter, wrapper, or the embedded) for vegetation mapping, combined with advanced deep learning methods.

1. INTRODUCTION

During the last decade, the rapid development of remote sensing (RS) imagery techniques (i.e., optical and radar sensors) has produced a tremendous amount of earth-observation (EO) image data used for, e.g., land-cover or vegetation mapping. Therefore, new prospects for research and applications are possible with the recent arrival of the Sentinel-1 and Sentinel-2 time series imagery (Niculescu et al., 2018; Dobrinić et al., 2021). However, vegetation mapping in heterogeneous landscapes is challenging since land-cover categories are difficult to separate spectrally due to low inter-class separability and high intra-class variability (Rodriguez-Galiano et al., 2012). Although ancillary derived input features (e.g., textural measures, vegetation indices) can be useful for EO image data classifications, they can add up to several hundred, sometimes irrelevant, and redundant features (Georganos et al., 2017). Thus, with an increasing number of input variables, highly correlated or redundant variables create noises in datasets, affecting prediction accuracy (Zhang and Yang 2020).

State-of-the-art machine learning (ML) methods, such as Random Forest (RF), can handle high data dimensionality and multicollinearity since it works with subsets of data. Introduced by Breiman (2001), RF is a supervised ML method that is constructed from a multitude of decision trees (DT). Through bagging, the trees are created by drawing a subset with about two-thirds of the training samples, whereas the remaining one-third of the data (out-of-bag – oob) can conveniently be used as a test set (Belgiu and Drăgut 2016). The final classification decision is made through a majority vote calculated by all produced trees. Therefore, RF is increasingly being applied in vegetation mapping using multispectral (Noi and Kappas 2018; Dobrinić et al., 2020) and radar (Gašparović and Dobrinić 2020) satellite sensor imagery. Furthermore, RF provides variable importance measures using the Gini index and the oob subset. The former measure is based on the average loss of entropy criterion, whereas the latter observes the values of the variables which are randomly permuted in the oob samples (Genuer et al., 2012).

Although RF can handle high dimensional data, various variable selection methods (e.g., filter, wrapper, and the embedded methods) can be used in combination with RF classification, to determine a subset of predictor features to be included in a final model. Filter methods rank the relevance of individual features by their correlation with the dependent variable, wrapper methods use feature subsets and evaluate them based on the classifier performance, whereas embedded methods combine the qualities of filter and wrapper methods (Saeys et al., 2007). A detailed comparison of RF variable selection methods can be found in the paper by Speiser et al. (2019). The authors used more than 300 classification datasets to assess different variable selection techniques to identify preferable methods based on applications in expert and intelligent systems. For land cover mapping in a complex environment (e.g., urbanized coastal area), Zhang and Yang (2020) examined various DT models and feature importance measures in RF. By using Landsat-8 imagery and best fit model (10 features), the overall accuracy

* Corresponding author

(OA) was 89.03% compared to the all-feature model which achieved an OA of 88.21% with 22 features. Stromann et al. (2019) and Orynbaikyzy et al. (2020) used Sentinel-1 (S1) and Sentinel-2 (S2) imagery for land cover and crop type classification, respectively. Both research assessed the impact of feature selection on the classification accuracies, however, the best variable selection methods identified in the research by Speiser et al. (2019) were not investigated for vegetation mapping using S1 and S2 imagery.

Due to the challenges in vegetation mapping in heterogeneous agricultural landscapes using multitemporal or multi-source (i.e., radar and optical) input data, the goal of this study was twofold. Firstly, to assess the performance of the random forest classifier in a large heterogeneous landscape with diverse land-cover categories using multi-seasonal Sentinel imagery and ancillary data. Secondly, to compare RF variable selection methods to identify a subset of predictor variables that will be included in a final, simpler model.

2. STUDY AREA AND DATA

2.1 Study site extent and LC classification system

Eastern part of central Croatia which covers the territory of Bjelovarsko – Bilogorska County was chosen as the study area in this research (Figure 1). The study site covers an area of 2639 km^2, where 1475 km^2 is agricultural land, which mostly includes croplands and gardens (68.7%), grasslands and pastures (27.4%), orchards (2.7%), and vineyards (1.2%) (Tomic 2012). Small-scale farms dominate the County croplands with 0.95 ha of average field size (Tomic 2012). According to the Koppen–Geiger climate classification system (Beck et al., 2018), this region has a temperate oceanic climate (Cfb), characterized by warm summer. The mean annual temperature for 2018 in the study area is 12.8°C with precipitation of 802.8 mm/year (Croatian Statistical Information 2019). Summer crops are sown at the end of March—beginning of April and harvested at the end of August—beginning of September.

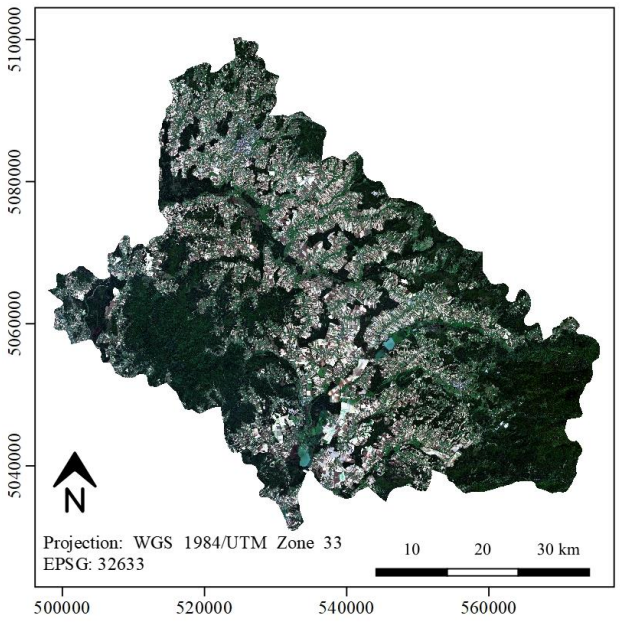

Figure 1. Overview of the Bjelovarsko – Bilogorska County on S2 imagery ("true colour" composite; sensing date: 21-08-2018)

As ground truth data on crop types for the year 2018, a hybrid classification scheme, proposed in the research from Dobrinić and Gašparović (2021), was used. According to the proposed approach, reference data was derived from CORINE, LUCAS, and Land Parcel Identification System (LPIS) land-cover database. Reference points were manually collected by remote sensing experts and confirmed on a time-series of Landsat and Google high-resolution imagery. Afterward, their semantic levels were adjusted according to the classification schemes from the aforementioned land-cover (LC) databases. In general, seven major LC classes were proposed in this research (Table 1). These reference points were uniformly distributed over LC classes (200 samples per class).

ID	Class	CORINE	LUCAS
1	Cropland	2.1, 2.4	B00 (except B70)
2	Forest	3.1	C00
3	Water	4.1, 5.1	G00
4	Built-up	1.1, 1.2	A00
5	Bare soil	3.3	F00
6	Grassland	2.3, 3.2	D00, E00
7	Orchard	2.2.2	B70

Table 1. Overview of the land-cover classes used in this research, including CORINE Level 2/3 and LUCAS classification scheme codes.

2.2 Data preparations

For this research, S1 imagery as a ground range detected (GRD) product was acquired from the ESA's Sentinel HUB. S1 GRD imagery has a spatial resolution of 10 m and consists of vertical-vertical (VV) and vertical-horizontal (VH) polarization bands. Standard pre-processing steps for the GRD imagery are to apply a precise orbit of acquisition, remove thermal noise, calibration, and Range Doppler terrain correction for geometric distortions caused by topography (Filipponi 2019). Additionally, speckle filtering, caused by the interference of waves, was made with a Lee filter with a window size of 5x5 pixels, effectively preserved edges and features, as shown in the research from Gašparović and Dobrinić (2021). Furthermore, very-high-resolution optical imagery (i.e., four spectral bands at 10 m, six bands at 20 m, and three bands at 60 m resolution) acquired from the Sentinel-2 (S2) satellites was also used. The S2 products were downloaded the ESA's Sentinel HUB as Level-2A products, which provides orthorectified Bottom-Of-Atmosphere (BOA) reflectance, with sub-pixel multispectral registration (Mercier et al., 2019). In this research, the S2 bands with 60 m spatial resolution were not used, whereas 20 m spectral bands were resampled to 10 m using the nearest neighbour method to preserve the pixels' original values (Osgouei et al., 2019).

To create classification maps in the highly heterogeneous study area with spectrally similar land covers, it is common to use multitemporal and multi-source remote sensing data to provide phenological changes in vegetation cover states (Jin et al, 2018). Therefore, radar and optical data from April, August, and October have been used for vegetation mapping because these three periods represent the most significant characteristics of the main vegetation types in the study area (Table 2).

	S1	S1 Orbit	S2	S2 Cloud Cover
Date	04-04-2018	ASC	08-04-2018	4.70
	20-08-2018	ASC	21-08-2018	0.75
	01-10-2018	ASC	05-10-2018	0.05

Table 2. Overview of the radar (S1) and optical (S2) satellite imagery used in this research.

In the most common way, land cover classification could be improved by using more input features that represent the feature domain more in detail (Zhang and Yang 2020). Therefore, in this research, textural features and spectral indices were derived from optical and radar satellite imagery, respectively. Therefore, Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI), Pigment Specific Simple Ratio (PSSRa), Normalized Difference Water Index (NDWI), Modified Chlorophyll Absorption in Reflectance (MCARI), Green Normalized Vegetation Index (GNDVI), Modified Soil Adjusted Vegetation Index (MSAVI), Normalized Difference Index 45 (NDI45), and Inverted Red-Edge Chlorophyll Index (IRECI) were calculated from S2 imagery (Dobrinić et al., 2021). Textural features were calculated from the grey-level co-occurrence matrix (GLCM) and used as input features (Haralick et al., 1973), as follows: Mean, Variance, Homogeneity, Contrast, Dissimilarity, Entropy, Second Moment, and Correlation (Gašparović and Dobrinić 2021). Overall, 38 input features were used in this research per date, 114 in total.

3. METHODS

3.1 Hyperparameter optimization

In order to prevent overfitting of the training data, RF involves several hyperparameters, which can be tuned for each classification task. For RF, *ntree* represents the number of trees in the forest, and *mtry* is the number of variables randomly chosen in each split. Whereas the former parameter should be set sufficiently high for optimal performance of the model, the latter is usually set as a square root of the input variables (Probst et al., 2019). Although *ntree* and *mtry* have shown effects on the RF performance stability, they should be carefully chosen so that the variable importance scores can converge to a stable mean (Behnamian et al., 2017). Therefore, a grid search approach with cross-validation was used in this research for hyperparameter optimization.

3.2 Variable importance measurements and selection

Many research avoid the feature selection process before classification, by assuming that various machine learning methods are robust to high dimensional datasets although noisy and correlated features are included (Ma et al., 2017). Therefore, feature selection (FS) can be defined as a process that reduces the number of input variables to reduce the computational cost and improve the model's performance (Zhang and Yang 2020). Depending on the variable selection approach (e.g., filter, wrapper, or the embedded methods), many packages for variable selection in R software have been developed.

Recursive feature elimination (RFE) takes all features firstly in the initial input and iteratively, at each step, the least important predictor is removed. It belongs to a wrapper-type FS method, and when the least important predictor(s) are removed, the model is re-built, and importance scores are computed again. This process is repeated until a specified number of features remains. (Georganos et al., 2018).

Boruta algorithm is a wrapper method that randomly designs shadow features and compares them with the importance of the real predictor variables. The values of those shadow variables are generated by permuting the original values across observations and therefore destroying the relationship with the outcome. A RF is trained on the extended data set, and the variable importance values are collected (Degenhardt et al., 2019).

Within RF, two different importance measures are implemented, mean decrease accuracy (MDA) and mean decrease Gini (MDG), that can be used for ranking variables and variable selection. The former measure assesses the importance of a variable by measuring the change in prediction accuracy when the values of the variable are randomly permuted compared to the original observations, whereas the latter is the sum of all decreases in Gini impurity due to a given variable, normalized by the number of trees (Calle and Urea 2010).

The VSURF or variable selection using the RF algorithm implements backward elimination then forward selection based on its importance and removes all the irrelevant features. Therefore, the least important predictors, which increase oob error by a margin larger than a threshold, are removed from the model. At the final step of the procedure, a list of the most discriminant features is obtained (Georganos et al., 2018). To summarize, in this research, the best performance of RF model was determined by evaluating various feature selection methods, which choose an optimal subset of features according to a certain criterion (Figure 2).

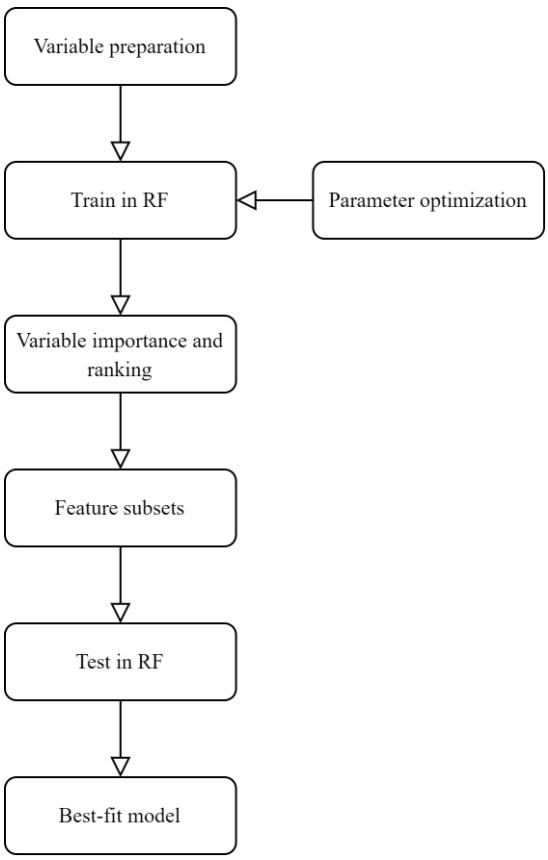

Figure 2. Flow chart of the main steps of the research.

3.3 Accuracy assessment

For this research, stratified random sampling was chosen as a sampling design to produce 200 samples per class, and LC classes were used as strata (Olofsson et al., 2014). Subsequently, a training set ratio of 50% sampling was applied to each stratum to obtain the training objects for constructing the classification model randomly. This number of samples was chosen because it is large enough to be adequately dispersed across the study site.

Confusion or error matrix (Stehman and Foody 2019) was used for accuracy assessment between the reference and predicted data. Overall accuracy (OA) was computed as the map-level accuracy metric, whereas F1 score was used to report class-level accuracy. The F1 score (Equation 1) is defined as the weighted harmonic mean of User's accuracy (UA) and Producer's accuracy (PA) and was calculated as follows:

$$F1=2\times\frac{PA\times UA}{PA+UA} \quad (1)$$

where PA and UA are defined as the complement of the omission and commission error probability, respectively.

Furthermore, as discussed in the research by Pontius and Millones (2011), Kappa coefficient is highly correlated with overall accuracy, and reporting them both is redundant. Therefore, the classification results were also evaluated with two simpler summary parameters: quantity disagreement (QD) and allocation disagreement (AD). Also, a 95% confidence interval (CI) was reported for all accuracy estimates since 10 random trials for each classification scenario were generated.

4. RESULTS AND DISCUSSIONS

4.1 Random Forest hyperparameter optimization results

As outlined in Section Methods, *ntree* and *mtry* hyperparameters of the RF classifier can be tuned to find the optimal LCC values and decrease the runtime needed for prediction (Probst et al., 2019). Grid search, in which all given *ntree* and *mtry* values are evaluated, was used for hyperparameter tuning (Table 3).

		ntree		
		100	500	1000
mtry	4	92.1	92.2	92.5
	7	94.3	93.8	93.1
	10	93.6	93.9	93.5

Table 3. RF hyperparameter tuning of *mtry* and *ntree* according to the overall accuracy (%) using grid search approach.

Since the addition of more trees, neither increases nor decreases the generalization error (Rodriguez-Galiano et al., 2012), for this research the *ntree* parameter was set to 100. Usually, the *mtry* parameter is set to a square root of a total number of input features, since a very high *mtry* may increase the dependency of member trees and thus compromise the model stability. Hence, similar to the research from Zhang and Yang (2020), the *mtry* was set as (n/10 + 1) where n represents the number of input features.

4.2 Feature selection

In order to select the ideal feature subset for vegetation mapping, five variable importance measurements (i.e., RFE, Boruta, MDA, MDG, and VSURF) were generated. As mentioned in Section 3.2., RFE method firstly creates a model with all variables and then continues until all variables have been removed and consequently, ranked (Georganos et al., 2018). For MDA and MDG, OA was compared depending on the number of the most important predictors (which were incrementally added until a model with all variables was reached), whereas Boruta and VSURF provide a report with information on which features were included or rejected. Therefore, Table 4 shows the OA of each feature selection (FS) method tested and a number of variables included within the reduced models. The top two methods, with similar OA, were MDA and VSURF, whereas MDG performed poorly in comparison to the other methods. Although MDA produced the lowest OA in the research by Georganos et al. (2018), still MDA method is frequently used to pre-select the most important variables (Belgiu and Drăgut 2016; Dobrinić et al., 2021). Furthermore, this research also confirmed that VSURF method offered the lowest number of input variables, as mentioned in the paper from Speiser et al. (2019). This should also be considered as a trade-off between accuracy (i.e., MDA) and complexity of the model (i.e., VSURF).

FS method	OA [%]	Nr. of features
RFE	90.09	80
Boruta	89.95	105
MDA	90.91	34
MDG	89.45	37
VSURF	90.77	19

Table 4. Peak overall accuracy (%) for Random Forest classifier, based on different FS methods.

Furthermore, the variable importance rankings generated from MDA method are shown in Figure 3. The ideal feature subset derived from MDA method consists of 14 S2 spectral bands, 12 spectral indices, six GLCM textural features, and two radar bands. The five most important variables were summer Green and Red Edge-1, spring Green and Red Edge-2 S2 bands, respectively, and summer GLCM Mean derived from S1 VH band.

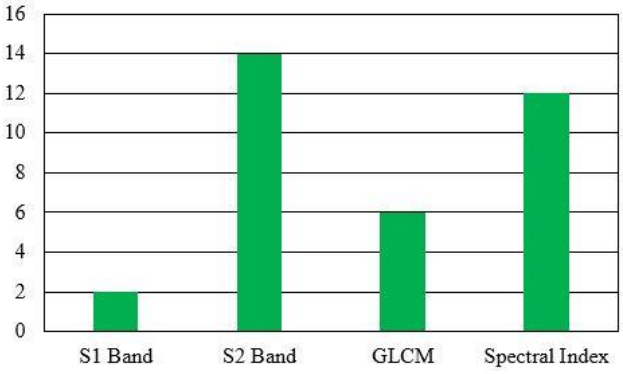

Figure 3. Number of input variables generated by the MDA FS method and grouped by the source of the input feature.

The results above coincide with similar research, e.g., Abdi (2019), where Red Edge bands from spring and summer dates were mostly included in LCC in a boreal landscape using S2 data. Also, the high importance of the Red Edge spectral bands could be associated with the mapping of different crop types (Forkuor et al., 2018). Regarding the Green spectral band, its value has also been proven for vegetation classification (Orynbaikyzy et al., 2020; Zhang and Yang 2020). Furthermore, texture features as a measure of image roughness, especially GLCM Mean proved to increase the classification accuracy for LCC (Balzter et al., 2015; Gašparović and Dobrinić 2021).

4.3 Validation in vegetation mapping scenarios

In order to evaluate the efficiency of the feature selection process, and as shown in Section 4.2, MDA FS method (34 input features) was used for comparison of the model's performance with the band-only model (36 input features) and the all-feature model (114 input features). The LCC produced by the MDA model had the highest OA of 90.91%, followed by the maps from all-feature and band-only models with an OA of 89.65% and 81.55%, respectively (Table 5). A similar trend, in terms of OA, was reported in the research from Zhang and Yang (2020), where the RF model and the ranking of MDA (10 features) outperformed all-feature (22 features) and band-only (13 features).

Class or summary metric	F1 score or model statistics		
	Band-only	All-feature	Best-fit
Cropland	0.81	0.80	0.85
Forest	0.90	0.90	0.95
Water	1.00	0.96	1.00
Built-up	0.74	0.81	0.95
Bare soil	0.69	0.93	0.95
Grassland	0.77	0.79	0.83
Orchard	0.78	0.93	0.89
OA	81.55	89.65	90.91
QD	7.77	4.44	2.39
AD	10.68	5.91	6.70
95% CI	(72, 88)	(85, 93)	(86, 94)

Table 5. Accuracy results for band-only, all-feature, and best-fit (MDA) model.

Although OA of the best-fit model was higher than an all-feature model, AD was lower for the best-fit model, which means that larger number of the location of a class pixel in the training data was different from the location of the same class in the test data (Pontius and Millones 2011). It should also be noted that every FS method, except MDG, outperformed both all-feature and band-only models. Likewise, Sun et al. (2020) and Schulz et al. (2021) used S1 and S2 time series with RF classifier for crop type mapping over agricultural areas and land use mapping in a heterogeneous landscape in Africa, respectively. Former research used recursive feature increment for FS and achieved an OA of 83.22% for a distribution map of five major crop types, whereas the latter removed features with pairwise correlations higher than 0.8 by calculating the correlation matrix and achieved an OA of 73.3%.

To assess the ability of differentiation between the LC classes, F1 score is presented in Table 5. According to the F1 score of individual LC classes, the best-fit model outperformed other models, except for the orchard due to the high heterogeneity, where the all-feature model performed better. Also, the best-fit model proved to differentiate better several vegetation classes (e.g., cropland, grassland) than the other two models. For the class with the lowest F1 score (i.e., grassland), possible lower classification performance could be associated with the semantic interpretation of LUCAS, CORINE, and LPIS land-cover databases and/or large positional error of the ground truth reference point data (Weigand et al., 2020).

Figure 4 represents the best supervised pixel-based classification scenario (i.e., best-fit model using MDA). Water and forest were in good agreement with testing data, whereas some orchard pixels were classified as cropland or grassland. Also, some confusion occurred between the built-up and forest part in the southeastern part of the study site. Due to the terrain topography or SAR shadowing, this misclassification could be removed using a high-quality digital elevation model (DEM) or additional textural features to enhance the land cover classes (Pesaresi et al., 2016).

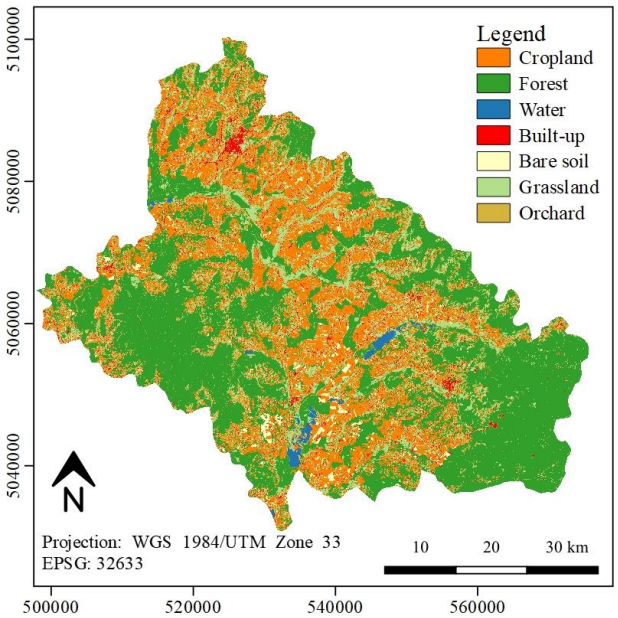

Figure 4. Variable importance of S1 and S2 input features sorted by the decreasing MDA values.

5. CONCLUSIONS

This research aimed to evaluate the performance of various feature selection methods for vegetation mapping using multitemporal Sentinel data and RF classifier.

Using MDA as a FS method, an original dataset was reduced from 114 to 34 input features, and its classification performance outperformed all-feature and band-only model with an OA of 90.91%. The most pertinent input features for vegetation mapping were S2 spectral bands (14 features), followed by spectral indices derived from S2, GLCM features, and S1 bands. However, regarding the classification performance, a trade-off between accuracy (i.e., MDA) and complexity of the model (i.e., VSURF) must be considered for vegetation mapping.

Future research should address additional feature selection methods (i.e., filter, wrapper, or the embedded) for vegetation mapping, combined with advanced deep learning methods.

REFERENCES

Abdi, A. M., 2019. Land Cover and Land Use Classification Performance of Machine Learning Algorithms in a Boreal Landscape Using Sentinel-2 Data. *GIScience and Remote Sensing*, pp: 1–20. doi: 10.1080/15481603.2019.1650447.

Balzter, H., Cole B., Thiel C., and Schmullius, C., 2015. Mapping CORINE Land Cover from Sentinel-1A SAR and SRTM Digital Elevation Model Data Using Random Forests. *Remote Sensing*, 7 (11), pp. 14876–14898. doi: 10.3390/rs71114876.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., and Wood, E.F., 2018. Present and Future Köppen-Geiger Climate Classification Maps at 1-Km Resolution. *Scientific Data*, 5 (1), pp. 1–12. doi: 10.1038/sdata.2018.214.

Behnamian, A., Millard, K., Banks, S.N., White, L., Richardson, M., and Pasher, J., 2017. A Systematic Approach for Variable Selection with Random Forests: Achieving Stable Variable Importance Values. *IEEE Geoscience and Remote Sensing Letters*, 14 (11), pp. 1988–1992. doi: 10.1109/LGRS.2017.2745049.

Belgiu, M., and Drăgut, L., 2016. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114, pp. 24–31. doi: https://doi.org/10.1016/j.isprsjprs.2016.01.011.

Breiman, L., 2001. Random Forests. *Machine Learning* 45 (1), pp. 5–32. doi: 10.1023/A:1010933404324.

Calle, M.L., Urrea, V., 2011. Stability of Random Forest importance measures. *Briefings in bioinformatics*, 12 (1), pp. 86-89. doi: 10.1093/bib/bbq011.

Degenhardt, F., Seifert, S., and Szymczak, S., 2019. Evaluation of Variable Selection Methods for Random Forests and Omics Data Sets. *Briefings in Bioinformatics*, 20 (2), pp. 492-503. doi: 10.1093/bib/bbx124.

Dobrinić, D., Medak, D., and Gašparović, M., 2020. Integration of Multitemporal Sentinel-1 and Sentinel-2 Imagery for Land-Cover Classification Using Machine Learning Methods. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, 43, pp. 91–98. doi: 10.5194/isprs-archives-XLIII-B1-2020-91-2020.

Dobrinić, D., Gašparović, M., and Medak, D., 2021. Sentinel-1 and 2 Time-Series for Vegetation Mapping Using Random Forest Classification: A Case Study of Northern Croatia. *Remote Sensing*, 13 (12), pp. 1–20. doi: 10.3390/rs13122321.

Filipponi, F., 2019. Sentinel-1 GRD Preprocessing Workflow. *Proceedings*, 18 (1), pp. 11. doi: 10.3390/ecrs-3-06201.

Forkuor, G., Dimobe, K., Serme, I., and Tondoh, J.E., 2018. Landsat-8 vs. Sentinel-2: Examining the Added Value of Sentinel-2's Red-Edge Bands to Land-Use and Land-Cover Mapping in Burkina Faso. *GIScience and Remote Sensing*, 55 (3), pp. 331–354. doi: 10.1080/15481603.2017.1370169.

Gašparović, M., and Dobrinić, D., 2020. Comparative Assessment of Machine Learning Methods for Urban Vegetation Mapping Using Multitemporal Sentinel-1 Imagery. *Remote Sensing*, 12 (12), pp. 1–22. doi: 10.3390/rs12121952.

Gašparović, M., and Dobrinić, D., 2021. Green Infrastructure Mapping in Urban Areas Using Sentinel-1 Imagery. *Croatian Journal of Forest Engineering*, 42 (2), pp. 1–20. doi: 10.5552/crojfe.2021.859

Genuer, R., Poggi, J.M., and Tuleau-Malot, C., 2012. Variable Selection Using Random Forests. *Pattern Recognition Letters*, 31(14), pp. 2225-2236. doi: 10.1016/j.patrec.2010.03.014.

Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., Kalogirou, S., and Wolff, E., 2018. Less Is More: Optimizing Classification Performance through Feature Selection in a Very-High-Resolution Remote Sensing Object-Based Urban Application. *GIScience and Remote Sensing*, 55 (2), pp. 221–242. doi: 10.1080/15481603.2017.1408892.

Haralick, R.M., Shanmugam, K., and Dinstein, I.H., 1973. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3 (6), pp. 610–621. doi: 10.1109/tsmc.1973.4309314.

Jin, Y., Liu, X., Chen, Y., and Liang, X., 2018. Land-Cover Mapping Using Random Forest Classification and Incorporating NDVI Time-Series and Texture: A Case Study of Central Shandong. *International Journal of Remote Sensing*, 39 (23), pp. 8703–8723. doi: 10.1080/01431161.2018.1490976.

Ma, L., Fu, T., Blaschke, T., Li, M., Tiede, D., Zhou Z., Ma, X., and Chen, D., 2017. Evaluation of Feature Selection Methods for Object-Based Land Cover Mapping of Unmanned Aerial Vehicle Imagery Using Random Forest and Support Vector Machine Classifiers. *ISPRS International Journal of Geo-Information*, 6 (2). doi: 10.3390/ijgi6020051.

Mercier, A., Betbeder, J., Rumiano, F., Baudry, J., Gond, V., Blanc, L., Bourgoin, C., et al. 2019. Evaluation of Sentinel-1 and 2 Time Series for Land Cover Classification of Forest–Agriculture Mosaics in Temperate and Tropical Landscapes. *Remote Sensing*, 11 (979), pp. 1–20. doi: 10.3390/rs11080979.

Niculescu, S., Talab Ou Ali, H., and Billey, A., 2018. Random forest classification using Sentinel-1 and Sentinel-2 series for vegetation monitoring in the Pays de Brest (France). In: Remote Sensing for Agriculture, Ecosystems, and Hydrology XX, 10783, pp. 1078305. doi: 10.1117/12.2325546.

Noi, P.T., and Kappas, M., 2018. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors*, 18 (1), pp. 18. doi: 10.3390/s18010018.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C.E., and Wulder, M.A., 2014. Good Practices for Estimating Area and Assessing Accuracy of Land Change. *Remote Sensing of Environment*, 148, pp. 42–57. doi: 10.1016/j.rse.2014.02.015.

Orynbaikyzy, A., Gessner, U., Mack, B., and Conrad, C., 2020. Crop Type Classification Using Fusion of Sentinel-1 and Sentinel-2 Data: Assessing the Impact of Feature Selection, Optical Data Availability, and Parcel Sizes on the Accuracies. *Remote Sensing*, 12 (17), pp. 2779. Doi: 10.3390/RS12172779.

Osgouei, P. E., Kaya, S., Sertel, E., and Alganci, U., 2019. Separating Built-up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote Sensing*, 11 (3), pp. 1–24. doi: 10.3390/rs11030345.

Pesaresi, M., Corbane, C., Julea, A., Florczyk, A., Syrris, V., Soille, P., Pesaresi, M., et al. 2016. Assessment of the Added-Value of Sentinel-2 for Detecting Built-up Areas. *Remote Sensing*, 8 (4), pp. 299. doi: 10.3390/rs8040299.

Pontius, R.G., and Millones, M., 2011. Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment. *International Journal of Remote Sensing*, 32 (15), pp. 4407–4429. doi: 10.1080/01431161.2011.552923.

Probst, P., Wright, M.N., and Boulesteix, A.L., 2019. Hyperparameters and Tuning Strategies for Random Forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9 (3), pp. 1–15. doi: 10.1002/widm.1301.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., and Rigol-Sanchez, J.P., 2012. An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67 (1), pp. 93–104. doi: 10.1016/j.isprsjprs.2011.11.002

Saeys, Y., Inza, I., and Larrañaga, P., 2007. A Review of Feature Selection Techniques in Bioinformatics. *Bioinformatics*, 23 (19), pp. 2507–2517. doi: 10.1093/bioinformatics/btm344.

Schulz, D., Bernhard Tischbein, H.Y., Verleysdonk, S., Adamou, R., and Kumar, N., 2021. Land Use Mapping Using Sentinel-1 and Sentinel-2 Time Series in a Heterogeneous Landscape in Niger , Sahel. *ISPRS Journal of Photogrammetry and Remote Sensing*, 178, pp. 97–111. doi: 10.1016/j.isprsjprs.2021.06.005

Speiser, J.L., Miller, M.E., Tooze, J., and Ip, E., 2019. A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling. *Expert Systems with Applications*, 134, pp. 93–101. doi: 10.1016/j.eswa.2019.05.028.

Stehman, S.V., and Foody, G.M., 2019. Key Issues in Rigorous Accuracy Assessment of Land Cover Products. *Remote Sensing of Environment*, 231, pp. 111199. doi: 10.1016/j.rse.2019.05.018.

Stromann, O., Nascetti, A., Yousif, O., and Ban, Y., 2020. Dimensionality Reduction and Feature Selection for Object-Based Land Cover Classification Based on Sentinel-1 and Sentinel-2 Time Series Using Google Earth Engine. *Remote Sensing*, 12 (1), pp. 76. doi: 10.3390/RS12010076.

Sun, L., Chen, J., Guo, S., Deng, X., and Han, Y., 2020. Integration of Time Series Sentinel-1 and Sentinel-2 Imagery for Crop Type Mapping over Oasis Agricultural Areas. *Remote Sensing*, 12 (158), pp. 1–27. doi: 10.3390/rs12010158.

Tomić, Franjo. 2012. Razvoj Poljoprivrede Primjenom Navodnjavanja u Bjelovarsko-Bilogorskoj Županiji. *Radovi Zavoda Za Znanstvenoistraživački i Umjetnički Rad u Bjelovaru*, 6 (6), pp. 1–15. URL: https://hrcak.srce.hr/91557.

Weigand, M., Staab, J., Wurm, M., and Taubenböck, H., 2020. Spatial and Semantic Effects of LUCAS Samples on Fully Automated Land Use/Land Cover Classification in High-Resolution Sentinel-2 Data. *International Journal of Applied Earth Observation and Geoinformation*, 88 , pp. 102065. doi: 10.1016/j.jag.2020.102065.

Zhang, F., and Yang, X., 2020. Improving Land Cover Classification in an Urbanized Coastal Area by Random Forests: The Role of Variable Selection. *Remote Sensing of Environment* 251, pp. 112105. doi: 10.1016/j.rse.2020.112105.