ADVANCES IN GEODESY AND GEOINFORMATION
(formerly GEODESY AND CARTOGRAPHY)
Vol. 74, no. 2, article no. e68, 2025

https://doi.org/10.24425/agg.2025.154152

Original paper

Tree species classification in heterogeneous forest with hyperspectral data from unmanned aerial vehicle

Ivana Horáková, Michal Kačmařík*, Lucie Orlíková

Technical University of Ostrava, Ostrava, Czech Republic
e-mail: ivana.horakova1@vsb.cz; ORCID: http://orcid.org/0009-0004-7307-8105
e-mail: michal.kacmarik@vsb.cz; ORCID: http://orcid.org/0000-0002-2352-3483
e-mail: lucie.orlikova@vsb.cz; ORCID: http://orcid.org/0000-0001-9777-7788

*Corresponding author: Michal Kačmařík, e-mail: michal.kacmarik@vsb.cz

Received: 2025-01-09 / Accepted: 2025-05-24

Abstract: Hyperspectral data obtained from unmanned aerial vehicles (UAVs) provide high spectral and spatial resolution, bringing potential for an accurate identification of individual tree species. The aim of this paper was to thoroughly investigate the possibilities of forest tree classification using hyperspectral data taken by the Resonon Pika L camera. Both hyperspectral and ground reference data were collected for a heterogenous forest in the Czech Republic. Standard processing methods (radiometric, atmospheric, and geometric corrections) were applied, followed by testing the methods for reduction (spectral resampling, MNF, PCA) and PPI. The classification phase consists of both unsupervised and supervised approaches, including Maximum Likelihood, Mahalanobis Distance, Spectral Angle Mapper, Minimum Distance, Random Forest, Extra Trees, Support Vector Machines, and Naive Bayes. Within the classical classifiers, the best results were achieved using the Maximum Likelihood classifier. In terms of machine learning algorithms, the best performing classifiers were Random Forest and Extra Trees. The use of Pika L camera in forestry classification is so far minimal, therefore the results can be helpful in potential utilization of this type of camera. The findings of this research not only contribute to a better understanding of UAV-based hyperspectral remote sensing for tree classifications but also provide practical insights and recommendations for improvement.

Keywords: UAV, remote sensing, tree species classification, hyperspectral data, Pika L

1. Introduction

Progress in remote sensing technologies has opened up new possibilities in forest management, as they allow image data acquisition without the need for physical interaction with trees and to perform classification and identification of tree species over large areas without negatively impacting forest ecology (Modzelewska et al., 2020; Hou et al., 2023).

Hyperspectral data acquired from unmanned aerial vehicles (UAVs) provide high spectral and spatial resolution. Obtained tree spectral signatures are useful for classification, identification of tree diseases (Dalponte et al., 2012) etc. The identification of the unique spectral signatures of different tree species is made possible by capturing surface reflectance in hundreds of very narrow bands providing precise spectral information. Effective analytical methods are necessary due to the complex structure and large size of data resulting from numerous spectral bands (Mäyrä et al., 2021). In addition to the need to apply corrections as part of pre-processing, for example, it is highly desirable to apply methods for dimensionality reduction due to the unwelcome correlation and amount of the data (Burger and Gowen, 2011; Cozzolino et al., 2023).

Choosing an appropriate classification algorithm and using data reduction methods is a challenging task and is still being tested, compared, and evaluated. Hycza et al. (2018) tested nine classification algorithms and achieved 90.3% accuracy using three MNF (Minimum Noise Fraction) bands and Maximum Likelihood (ML) classification on the AISA Eagle II data from a Polish forest. Burai et al. (2015) found that Support Vector Machines (SVM) and Random Forest (RF) classifiers performed best with the first nine MNF bands in a Hungarian steppe on AISA Eagle II data, with SVM being optimal due to its training sample size insensitivity. ML provided high accuracy (80.8%), but using a smaller training dataset significantly reduced the classification accuracy (52.6%). For the original bands (without MNF), the SVM and RF classifiers provided high accuracy regardless of the number of training pixels. Richter et al. (2016) found PLS-DA most successful in a German forest on AISA Dual data, outperforming RF and SVM. Nezami et al. (2020) achieved 98.3% accuracy with a 3D-CNN (convolutional neural network) using hyperspectral and RGB bands. Tree species classification using hyperspectral data from UAVs utilizing machine learning and deep learning techniques has also been addressed by other studies, e.g., by Amarasingam et al. (2024) or Xu et al. (2024). A combination of using hyperspectral and LiDAR data for the same purpose has been discussed by several authors (e.g., Sankey et al., 2017; Almeida et al., 2021; Ma et al., 2024).

The goal of this study was to thoroughly investigate the possibilities of forest tree classification using hyperspectral data taken by an UAV. To achieve this goal, hyperspectral and ground reference data were collected over a heterogeneous forest in the Czech Republic. Standard preprocessing steps, including radiometric, atmospheric, and geometric corrections, were applied. Several data reduction techniques were tested, including spectral resampling, Minimum Noise Fraction, Principal Component Analysis, along with Pixel Purity Index extraction. The classification phase involved both unsupervised and supervised approaches, utilizing methods such as Maximum Likelihood, Mahalanobis Distance, Spectral Angle Mapper, Minimum Distance, Random Forest, Extra Trees, Support Vector Machines, and Naive Bayes classifiers.

The use of UAV data from hyperspectral sensors can be expected to become more and more significant in a near future and therefore due attention should be paid in the research of optimal procedures for their processing. In our case, capturing a heterogeneous forest with a variety of tree species allows a detailed evaluation of various pre-processing and classification techniques and contributes to the above-mentioned effort. Novelty of the work also lies in utilization of the Resonon Pika L hyperspectral camera. The use of this camera in forestry classification is so far minimal (Shuai et al., 2024), but finds its use for instance in vegetation analysis, disease detection, chlorophyll quantification or segmentation models for crop mapping (Yang et al., 2021; Yu et al., 2021a, 2021b; Niu et al., 2022; Hariharan et al., 2023), so this work can offer valuable information for the case of tree species classification. Most importantly, it can provide insights into the potential of hyperspectral data from UAVs for forest monitoring and management.

2. Materials and methods

2.1. Study site

The area of interest (approximately 3.3 ha) is located in Ostrava-Poruba (Fig. 1), Czech Republic. It is a part of the local forest (approx. 49.837249°N, 18.149631°E to 49.835390°N, 18.151746°E). The area is densely forested with more than ten of deciduous and coniferous tree species, which is desirable in terms of testing classification methods.

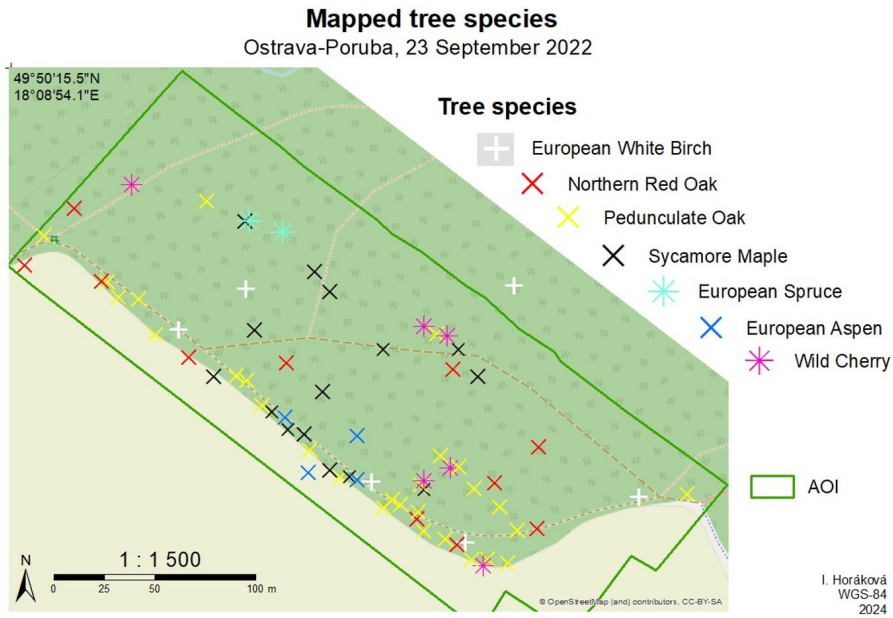

Fig. 1. Area of interest with mapped trees

2.2. Equipment

The DJI Matrice 600 Pro hexacopter and the Resonon Pika L hyperspectral VNIR camera (spectral range: 400–1000 nm; spectral resolution: approx. 3 nm; total of 200 bands) were used for aerial imaging. Field of view (FOV) of the used camera lens is 17.6°. For image stabilization, the camera was placed in a DJI Ronin-MX gimbal. SBG Ellipse-N localization unit was used. The flight planning and realization was performed with DJI Ground Station Pro software.

For tree species mapping, GNSS receiver Trimble R10 receiver with RTK (Real-Time Kinematic) positioning technique was used in the first round. The second mapping was performed using an internal GNSS receiver of smartphone Samsung A52 5G.

2.3. Field data collection

Hyperspectral imagery of the selected area was acquired on June 2, 2022 during the growing season of the trees. Data were captured on a cloudless, windless day around 9:30 A.M. A total of 46 raw data cubes were captured from a flight height of 90 m. The area of interest was covered by five flight lines with a planned side overlap of 20%.

The ground works to obtain a reference dataset on the location and species of trees was carried out in two days. Information was collected on tree species, approximate height, health status and whether they were single trees or groups. On March 26, 2021, 14 individual trees were recorded with a mean formal horizontal accuracy of 0.8 m. On September 23, 2022, additional 67 trees were mapped with an average formal accuracy of 2.4 m. Efforts focused on mapping solitary trees at the forest edge for a better recognition. The PlantNet app (https://identify.plantnet.org/) verified species identification when uncertain. A few out-of-area points were discarded, so altogether 71 usable points were collected in the field in Table 1, representing the best and largest possible dataset.

Table 1. Number of individual mapped trees according to their specie and number of polygons used for the training and test dataset

Species	Mapped	Training	Test
European white birch (Betula pendula)	5	2	1
Northern red oak (Quercus rubra)	11	3	6
Pedunculate oak (Quercus robur)	28	4	19
Sycamore maple (Acer pseudoplatanus)	15	3	7
European spruce (Picea abies)	2	2	2[a]
European aspen (Populus tremula)	4	3	1
Wild cherry (Prunus avium)	6	3	2
Total:	71[b]	20	38

[a] two validation polygons added via manual labelling based on their typical shape
[b] 15 ambiguous points were not used for training nor test dataset

2.4. Data processing

Data pre-processing was done in the Resonon Spectronon software (ver. 3.0 and 3.1), while data reduction, masking, and classification were performed in ENVI (ver. 4.8 and 5.0). The hyperspectral data were acquired in the so-called radiance mode (Resonon, 2018), which ensured a conversion from raw data to radiance in the camera itself, based on the calibration file supplied by the camera manufacturer. The imagery data obtained during the flight over a calibration tarp of known reflectivity (0.36) was used to convert the radiance to reflectance. Consequently, the data cubes were orthorectified and their mosaic was derived (Fig. 2a).

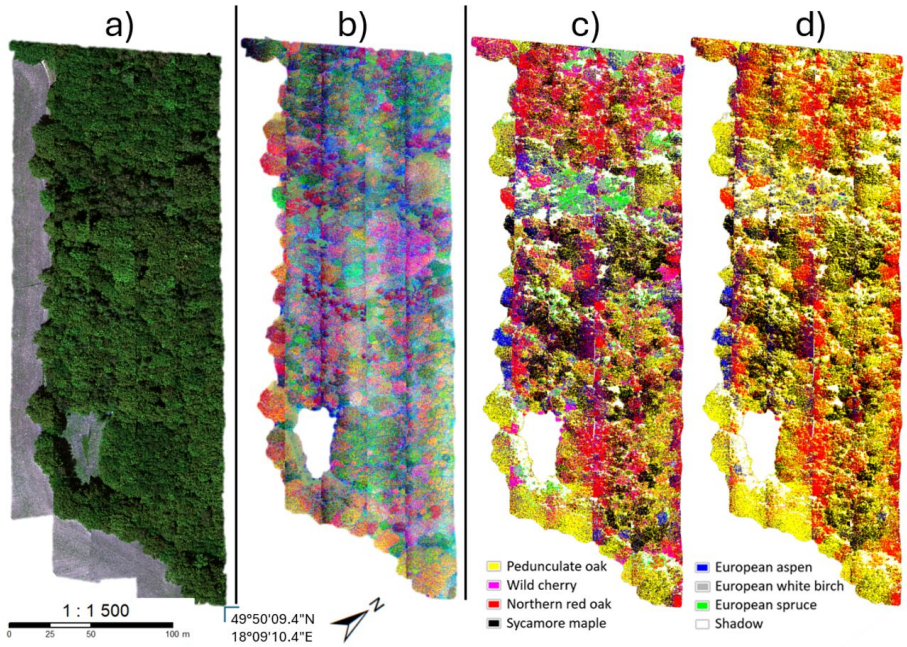

Fig. 2. The resulting forest mosaic in (a) true color, (b) image visualization using the 4th, 6th, and 9th MNF bands (b), and result of supervised ML classification of (c) spectrally resampled data and (d) for unresampled data without smoothing

Georectification was performed using the Czech national digital terrain model (DTM) called DMR 5G derived from aerial laser scanning data. Still, some of the infill rows did not fit together perfectly. Improved DTM or increased flight line overlap might enhance this.

Spectral resampling was performed as part of the data reduction. The original spectral resolution was about 3 nm. Following the example of Cervena et al. (2020), spectral resampling was performed, reducing the original 3 nm spectral resolution. Three options were tested: merging 3, 5 or 7 bands. Merging five bands (16 nm) shown to be optimal band reduction, based on the spectral curves obtained after the spectral reduction, the curves were smooth and still detailed enough.

The MNF method (Green et al., 1988) revealed the first 16 bands with the least noise (eigenvalue > 3) (NV5 GEOSPATIAL SOLUTIONS, 2024; Shawky et al., 2019). Visualization in Figure 2b used the 4th, 6th, and 9th MNF bands.

Although the classification results for the MNF bands were not successful, PPI (Pixel Purity Index) (Boardman et al., 1995) calculation (16 bands, 1000 iterations) was applied twice – first to identify shadows, then to clean pixels on illuminated areas. Principal Component Analysis (PCA) (Pearson, 1901; Ma, 2014) was applied to the reflectance image. According to Eigenvalues, the first PCA band contains the most information (99%); adding the remaining bands would add only a small amount of information.

2.5. Classification

Training and test data set

Training polygons for supervised classification were manually created based on the field data. Several different ratios of training and test were prepared and tested, the variant presented in Table 1 was selected since it provided the best classification accuracy. Fewer training points (34:66) were selected rather than standard ratios (Abriha et al., 2023; Elvanidi and Katsoulas, 2023). The training set was chosen to be well balanced (roughly the same number of polygons for each species). Fifteen points from field works were not used at all, as it was not clear to which crown the point belonged. The training set also included a class for shadows. Validation in form of a confusion matrix and overall accuracy was counted using ENVI software.

Classification methods

Unsupervised and supervised classification was performed in ENVI software on both spectrally resampled data (40 bands) and unresampled data (200 reflectance bands). Object-oriented classification to determine tree canopies was ineffective due to dense forest cover, so only species composition was classified. For unsupervised classification, all available algorithms within ENVI were tested. The classic ones were K-means (Jain, 2010) and IsoData (Tou and Gonzalez, 1974). Machine learning algorithms included BIRCH (Zhang et al.,1997) and Mini Batch K-means (Sculley, 2010).

Maximum Likelihood (Aldrich, 1997), Mahalanobis Distance (Mahalanobis, 1936), Spectral Angle Mapper (SAM) (Kruse et al., 1993), Minimum Distance (MD) (Wolfowitz, 1957), and Parallelepiped (Richards and Jia, 1999) algorithms were tested for supervised classification. For machine learning algorithms, Random Forest (Breiman, 2001), Extra Trees (Geurts et al., 2006), Support Vector Machines (Cortes and Vapnik, 1995), Naive Bayes (Wickramasinghe and Kalutarage, 2021) and K-Neighbors (Zhang, 2016) algorithms were used. The classification result was smoothed in all cases (Smooth Kernel Size: 13, Aggregate Minimum Size: 15) and the confusion matrix including the overall classification accuracy was calculated using the test dataset.

3. Results

3.1. Classification results based on derived data products

To assess the applicability of the MNF data, unsupervised classification for the five classes was used. Unfortunately, the first three MNF bands were found to contain undesirable strips along the flight paths. Even when using the most optimal MNF bands, neither supervised nor unsupervised classification yielded good results, so the MNF results were not further utilized.

In case of PPI image, clean pixels (endmembers) were selected for each mapped tree species and SAM classification was performed for both the MNF image and the reflectance image (both with masked shadows). The result was poor in both cases, which could be due to the small number of selected clean pixels. For the original reflectance image, the overall accuracy was 51.39% and the kappa was 0.017. Therefore, the results were not used further.

For PCA image, the first band was used for classification. Interestingly, with a 99.49% match, the result is almost the same as the unsupervised classification result for the original reflectance image. However, the supervised classification (ML) for the single PCA band was not successful, with an overall success rate of only 22.07%. The MD classification was similarly unsuccessful. At least 2 bands would have been required to perform RF and SVM. Due to the unacceptable success rate, the PCA result was not used further.

Issues in all three cases are attributed to a lower side overlap setting (20%) with respect to the camera FOV. The off-nadir angle was too wide and therefore influenced pixels along boundaries of flight lines. The effect was clearly visible in the case of MNF bands (Fig. 2b).

3.2. Unsupervised classification results

The unsupervised classification results were almost identical for the spectrally resampled and unresampled data, e.g. 99.20% match for the five classes. The overall accuracy of the IsoData method after assigning the appropriate tree type for each identified class was 48.66% for 5 classes, and for 6 classes the accuracy dropped to 38.06%. Setting more classes led to an ambiguous classification.

The machine learning algorithms were not successful. Unsupervised classification using the BIRCH algorithm always produced an output of three classes when set to 3, 6 and 7 classes. Mini Batch K-means classification, when set to 6 classes, the output consisted of virtually only two classes, one for the shadows and one for the rest of the image.

3.3. Supervised classification results

The overall accuracy (OA) and kappa coefficients of all the classifiers tested can be seen in Table 2. Confusion matrix for the best performing Maximum Likelihood classification can be seen in Table 3 all the others can be requested from the corresponding author. Within the classical classifiers, the best results were achieved using the ML classifier

(Fig. 2c and Fig. 2d), with an OA of 84.55% (kappa 0.656) for the spectrally unresampled image and 82.70% (kappa 0.631) for the resampled image. The worst performing classifier was the MD classifier, with an OA of around 28% in both cases. Interestingly, we found that the spectrally resampled data provided a lower accuracy by a few percent. For the Parallelepiped classification, the result image was incorrectly covered mostly by one class.

In terms of machine learning algorithms, the best performing classifiers were RF and Extra Trees (OA in the range of 82.04–82.68%). SVM was the only successful algorithm where the accuracy of the spectrally resampled data was higher than that of the unresampled data (81.89% and 73.85%, respectively). The same was observed for the Naive Bayes algorithm, but it was very inaccurate. An error occurred immediately after running the K-Neighbors algorithm.

Table 2. Overall accuracy and kappa coefficients for supervised classification results

	Classification type	Spectrally resampled image (40 bands)		Original spectrally unresampled image (200 bands)	
		Overall accuracy (OA)	Kappa coefficient	Overall accuracy	Kappa coefficient
Classical methods	ML	82.70%	0.631	84.55%	0.656
	Mahalanobis Distance	81.49%	0.599	83.88%	0.662
	SAM	46.30%	0.260	51.12%	0.273
	MD	28.06%	0.067	28.19%	0.068
Machine learning	RF	82.33%	0.551	82.68%	0.564
	Extra Trees	82.04%	0.537	82.34%	0.548
	SVM	81.89%	0.560	73.85%	0.461
	Naive Bayes	24.41%	0.112	15.85%	0.071

ML: Maximum Likelihood, SAM: Spectral Angle Mapper, MD: Minimum Distance, RF: Random Forest, SVM: Support Vector Machine

4. Discussion

The difference between supervised and unsupervised classification using PCA is notable. Unsupervised classification was nearly identical for both PCA and the original reflectance image, but supervised results differed. This likely stems from the single band PCA image's sensitivity to training dataset inconsistencies. While classifying from a single band seems meaningless, a paper by [10]Carr (1996) showed it is possible with bands having drastic value differences, as seen in our PCA band.

Of all the data reduction methods tested (spectral resampling, MNF, and PCA), only the spectral resampling option was suitable. Generally, in most cases, the use of reduced and otherwise modified data leads to an improvement in the data properties and thus to

Table 3. Confusion matrix for Maximum Likelihood classification of spectrally resampled image

	Ground truth (%)							
Class	Quercus rubra	Quercus robur	Prunus avium	Populus tremula	Picea abies	Acer pseudo-platanus	Betula pendula	Total
Quercus rubra	34.16	3.62	0.00	0.00	0.00	14.46	0.00	9.33
Quercus robur	26.17	94.84	21.12	91.63	0.00	9.89	0.00	71.64
Prunus avium	21.76	0.00	58.24	0.00	0.00	0.00	9.46	3.88
Populus tremula	0.00	0.70	0.00	0.00	0.00	0.00	5.41	0.50
Picea abies	0.00	0.43	0.00	0.49	100.00	0.00	0.00	0.60
Acer pseudoplatanus	17.91	0.40	14.40	7.88	0.00	75.65	0.00	13.67
Betula pendula	0.00	0.01	6.24	0.00	0.00	0.00	85.14	0.37
Total:	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

an increase in classification accuracy, but this was not the case for our data, very likely due to the non-sufficient side overlaps with respect to the FOV. For example, [11]Cervena et al. (2020) compared the classification results with and without spectral reduction. A Headwall Nano-Hyperspec camera was used. For band reduction, spectral resampling and PCA were tested. The best results were obtained with the resampled data – five adjacent bands were merged. In case of our work and evaluation of spectral resampling, the result of unsupervised classification is visually identical using both spectrally resampled and unsampled data. A slight difference occurred when using supervised classification, where the overall accuracy of, e.g., ML classifier for the spectrally resampled image was 82.70% and 84.55% for the original spectrally unresampled image. The spectrally unresampled image showed better results for all the classifiers tested, except for the two machine algorithms SVM and Naive Bayes. Nevertheless, the overall results were expected; ML and RF algorithms are popular based on their quality. [22]Kluczek et al. (2022) realized forest data classification from HySpex camera with a broader range (416-2510 nm). They reached OA of 83.4% for RF and 87.9% for SVM. Our results were just slightly lower, more visibly in case of SVM. The reasons behind the varying performance of all the classifiers can be found for example in an extensive review of [40]Tejasree and Agilandeeswari (2024). In any case, it should be pointed out that the number of reference data used, although was the highest possible using available methods, was still very low.

The suboptimal side overlap setting and wide off-nadir angle likely contributed to the issues observed, particularly in the MNF bands. It is advisable to recommend an overlap at the level of 30-40% to reduce inaccurate pixels at the flight line boundaries, which could affect the classification results. When using a reference tarp with high reflectance during data acquisition, we recommend capturing an image of it without a forest or other very dark surface around to avoid oversaturation of pixels with the tarp. To further improve classification accuracy, integrating laser scanning data of a reasonable point density with hyperspectral imaging could be beneficial. Laser scanning derived Digital Terrain

Models and Canopy Height Models (CHMs) can enhance georectification and facilitate more precise object-oriented classification, especially in dense vegetation covers, as demonstrated by [9]Cao et al. (2021). The applied methodology is general; it can be applied in other forest types and regions. Nevertheless, it needs to be tested on other data sources.

5. Conclusions

The aim of this study was to thoroughly investigate the possibilities of forest tree classification using hyperspectral data taken by an UAV. The selection of a heterogenous forest with a high number of tree species and the use of the Resonon Pika L camera, whose application in forestry is so far minimal, is a significant advantage over other works. Individual trees were mapped in the area of interest to obtain training and test dataset. A total of 71 reference points entered the processing.

After data pre-processing in the native Spectronon software, masking and testing of PPI and methods for data reduction were performed in ENVI. Of all the data reduction methods tested, only the spectral resampling option was suitable. Both spectrally resampled and unresampled data (40 and 200 bands) were used for classification of forest tree species. Within the supervised classification, four classical algorithms were used: ML, Mahalanobis Distance, MD and SAM. The best results were achieved using the Maximum Likelihood classifier: the overall accuracy for the spectrally unresampled image was 84.55% (kappa 0.656) and for the resampled image 82.70% (kappa 0.631). Interestingly, we found that the spectrally resampled data provided a few percentage points less accuracy. The machine learning algorithms tested were RF, Extra Trees, SVM, and Naive Bayes. RF and Extra Trees were the best performing ones (OA ranging from 82.04% to 82.68%). The study concludes that both classical classifiers (particularly ML) and selected machine learning algorithms performed well. The quality of their results could be improved by extending the training dataset, especially in the case of machine learning.

Author contributions

Conceptualization: I.H., L.O., M.K.; funding acquisition: I.H., M.K.; methodology: I.H., L.O.; resources: M.K.; mapping: M.K., I.H., L.O.; formal analysis: I.H.; software: I.H.; supervision: L.O.; validation: I.H.; writing – original draft: I.H., L.O.; writing – review and editing: M.K., I.H.

Data availability statement

The data can be requested from the corresponding author.

Acknowledgements

This work was partially supported by Grant of SGS No. SP2024/057, Faculty of Mining and Geology, VSB – Technical University Ostrava.

References

Abriha, D., Srivastava, P.K., and Szabó, S. (2023). Smaller Is Better? Unduly Nice Accuracy Assessments In Roof Detection Using Remote Sensing Data With Machine Learning And K-Fold Cross-Validation. *Heliyon*, 9(3). DOI: 10.1016/j.heliyon.2023.e14045.

Aldrich, J. (1997). R.a. Fisher And The Making Of Maximum Likelihood 1912-1922. *Statistical Science*, 12(3). DOI: 10.1214/ss/1030037906.

Almeida, D.R.A., Broadbent E.N., Ferreira M.P. et al. (2021). Monitoring Restored Tropical Forest Diversity And Structure Through UAV-Borne Hyperspectral And LiDAR Fusion. R*emote Sens. Environ.*, 264. DOI: 10.1016/j.rse.2021.112582.

Amarasingam, N., Kelly J.E., Sandino J. et al. (2024). Bitou Bush Detection And Mapping Using UAV-Based Multispectral And Hyperspectral Imagery And Artificial Intelligence. *Remote Sens. Appl.: Soc. Environ.*, 34. DOI: 10.1016/j.rsase.2024.101151.

Boardman, J.W., Kruse, F.A., and Green, R.O. (1955). *Mapping Target Signatures Via Partial Unmixing Of AVIRIS Data*. NASA: Pasadena, CA, USA, 3–6.

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

Burai, P., Deák B., Valkó O. et al. (2015). Classification Of Herbaceous Vegetation Using Airborne Hyperspectral Imagery. *Remote Sens.*, 7(2), 2046–2066. DOI: 10.3390/rs70202046.

Burger, J., and Gowen A. (2011). Data Handling In Hyperspectral Image Analysis. *Chemom. Intell. Lab. Syst.*, 108(1), 13–22. DOI: 10.1016/j.chemolab.2011.04.001.

Cao, J., Liu, K., Zhuo, L. et al. (2021). Combining UAV-Based Hyperspectral And LiDAR Data For Mangrove Species Classification Using The Rotation Forest Algorithm. *Int. J. Appl. Earth Obs. Geoinf.*, 102. DOI: 10.1016/j.jag.2021.102414.

Carr, J.R. (1996). Spectral And Textural Classification Of Single And Multiple Band Digital Images. *Comp. Geosci.*, 22(8), 849–865. DOI: 10.1016/S0098-3004(96)00025-8.

Cervena, L., Lysak, J., Potuckova, M. et al. (2020). Zkušenosti Se Zpracováním Hyperspektrálních Dat Pořízených UAV. *GIS Ostrava*. DOI: 10.31490/9788024843988-4.

Cortes, C., and Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3), 273–297. DOI: 10.1007/BF00994018.

Cozzolino, D., Williams P.J., and Hoffman, L.C. (2023). An Overview Of Pre-Processing Methods Available For Hyperspectral Imaging Applications. *Microchem. J.*, 193. DOI: 10.1016/j.microc.2023.109129.

Dalponte, M., Bruzzone, L., and Gianelle, D. (2012). Tree Species Classification In The Southern Alps Based On The Fusion Of Very High Geometrical Resolution Multispectral/Hyperspectral Images And LiDAR Data. Online. *Remote Sens. Environ.*, 123, 258–270. DOI: 10.1016/j.rse.2012.03.013.

Elvanidi, A., and Katsoulas, N. (2023). Artificial Neural Network Based On Multilayer Perceptron Algorithm As A Tool For Tomato Stress Identification In Soilless Cultivation. *Acta Horticult.*, 1377, 447–454. DOI: 10.17660/ActaHortic.2023.1377.54.

Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely Randomized Trees. *Machine Learning*, 63(1), 3–42. DOI: 10.1007/s10994-006-6226-1.

Green, A.A., Berman, M., Switzer, P. et al. (1988). A Transformation For Ordering Multispectral Data In Terms Of Image Quality With Implications For Noise Removal. *IEEE Trans. Geosci. Remote Sens.*, 26(1), 65–74. DOI: 10.1109/36.3001.

Hariharan, J., Ampatzidis, Y., Abdulridha, J. et al. (2023). An AI-Based Spectral Data Analysis Process For Recognizing Unique Plant Biomarkers And Disease Features. *Comp. Electr. Agricult.*, 204. DOI: 10.1016/j.compag.2022.107574.

Hou, Ch., Liu, Z., Chen, Y. et al. (2023). Tree Species Classification From Airborne Hyperspectral Images Using Spatial–Spectral Network. *Remote Sens*., 15(24). DOI: 10.3390/rs15245679.

Hycza, T., Sterenczak, K., and Balazy, R. (2018). Potential Use Of Hyperspectral Data To Classify Forest Tree Species. *New Zealand J. Forestry Sci.*, 48(1). DOI: 10.1186/s40490-018-0123-9.

Jain, A.K. (2010). Data Clustering: 50 Years Beyond K-Means. *Pattern Recognition Lett.*, 31(8), 651–666. DOI: 10.1016/j.patrec.2009.09.011.

Kluczek, M., Zagajewski, B., and Kycko, M. (2022). Airborne HySpex Hyperspectral Versus Multitemporal Sentinel-2 Images for Mountain Plant Communities Mapping. *Remote Sens.*, 14(5). DOI: 10.3390/rs14051209.

Kruse, F.A., Lefkoff, A.B., Boardman, J.B. et al. (1993). The Spectral Image Processing System (SIPS) – Interactive Visualization And Analysis Of Imaging Spectrometer Data. *Remote Sens. Environ.*, 44, 145–163. DOI: 10.1016/0034-4257(93)90013-N.

Ma, Y.Z. (2014). *A Tutorial On Principal Component Analysis*. DOI: 10.13140/2.1.1593.1684.

Ma, Y., Zhao, Y., Im, J. et al. (2024). A Deep-Learning-Based Tree Species Classification For Natural Secondary Forests Using Unmanned Aerial Vehicle Hyperspectral Images And LiDAR. *Ecol. Indic.*, 159. DOI: 10.1016/j.ecolind.2024.111608.

Mäyrä, J., Keski-Saari, S., Kivinen, S. et al. (2021). Tree Species Classification From Airborne Hyperspectral And LiDAR Data Using 3D Convolutional Neural Networks. *Remote Sens. Environ.,* 256. DOI: 10.1016/j.rse.2021.112322.

Mahalanobis, P.C. (1936). On The Generalized Distance In Statistics. *Proceedings Of The National Institute Of Science Of India*, 49–55.

Modzelewska, A., Fassnacht F.E., and Sterenczak, K. (2020). Tree Species Identification Within An Extensive Forest Area With Diverse Management Regimes Using Airborne Hyperspectral Data. Online. *Int. J. Appl. Earth Obs. Geoinf.*, 84. DOI: 10.1016/j.jag.2019.101960.

Nezami, S., Khoramshahi, E., Nevalainen, O. et al. (2020). Tree Species Classification Of Drone Hyperspectral And RGB Imagery With Deep Learning Convolutional Neural Networks. *Remote Sens.*, 12(7). DOI: 10.3390/rs12071070.

Niu, B., Feng, Q., Chen, B. et al. (2022). HSI-Transunet: A Transformer Based Semantic Segmentation Model For Crop Mapping From UAV Hyperspectral Imagery. *Comp. Elect. Agricult.*, 201. DOI: 10.1016/j.compag.2022.107297.

NV5 GEOSPATIAL SOLUTIONS, INC. (2024). Minimum Noise Fraction Transform. Retrieved September 1, 2025, from https://www.nv5geospatialsoftware.com/docs/MinimumNoiseFractionTransform.html.

Pearson, K. (1901). On Lines And Planes Of Closest Fit To Systems Of Points In Space. *Philos. Mag.*, 2 (6), 559–572.

Resonon. (2018). Data Modes. Airborne User Manual, 1–2. Retrieved September 1, 2025, from https://resonon.com/content/files/AirborneUserManual-4.1.pdf.

Richards, J.A., and Jia, X. (1999). Parallelepiped Classification. *Remote Sens. Digital Image Anal.*, 192. DOI: 10.1007/978-3-662-03978-6.

Richter, R., Reu, B., Wirth, Ch. et al. (2016). The Use Of Airborne Hyperspectral Data For Tree Species Classification In A Species-Rich Central European Forest Area. *Int. J. Appl. Earth Obs. Geoinf.i,* 52, 464–474. DOI: 10.1016/j.jag.2016.07.018.

Sankey, T., Donager, J., McVay, J. et al. (2017). UAV LiDAR And Hyperspectral Fusion For Forest Monitoring In The Southwestern USA. *Remote Sens. Environ.*, 195, 30–43. DOI: 10.1016/j.rse.2017.04.007.

Sculley, D. (2010). Web-Scale K-Means Clustering. In Proceedings Of The 19Th International Conference On World Wide Web, April, 1177–1178. DOI: 10.1145/1772690.1772862.

Shawky, M.M., El-Arafy, R.A., Zalaky, M.A.E. et al. (2019). Validating (MNF) Transform To Determine The Least Inherent Dimensionality Of Aster Image Data Of Some Uranium Localities At Central Eastern Desert, Egypt. *J. Afr. Earth Sci.,* 149: 441–450. DOI: 10.1016/j.jafrearsci.2018.08.022.

Shuai, L., Li, Z., Chen, Z. et al. (2024). A Research Review On Deep Learning Combined With Hyperspectral Imaging In Multiscale Agricultural Sensing. *Comp. Electr. Agricult.,* 217. DOI: 10.1016/j.compag.2023.108577.

Tejasree, G., and Agilandeeswari, L. (2024). An extensive review of hyperspectral image classification and prediction: techniques and challenges. *Multim. Tools Appl.* DOI: 10.1007/s11042-024-18562-9.

Tou, J.T., and Gonzalez, R.C. (1974). *Pattern Recognition Principles*. Addison: Wesley Publishing Company.

Wickramasinghe, I., and Kalutarage, H. (2021). Naive Bayes: Applications, Variations And Vulnerabilities: A Review Of Literature With Code Snippets For Implementation. *Soft Comp.*, 25(3), 2277–2293. DOI: 10.1007/s00500-020-05297-6.

Wolfowitz, J. (1957). The Minimum Distance Method. *The Ann. Math. Stat.*, 28(1), 75–88. DOI: 10.1214/aoms/1177707038.

Xu, L., Lu Ch., Zhou T. et al. (2024). A 3D-2DCNN-CA Approach For Enhanced Classification Of Hickory Tree Species Using UAV-Based Hyperspectral Imaging. *Microche. J.l* 199. DOI: 10.1016/j.microc.2024.109981.

Yang, Z., Tian, J., Feng, K. et al. (2021). Application Of A Hyperspectral Imaging System To Quantify Leaf-Scale Chlorophyll, Nitrogen And Chlorophyll Fluorescence Parameters In Grapevine. *Plant Phys. Biochem.*, 166, 723–737. DOI: 10.1016/j.plaphy.2021.06.015.

Yu, R., Luo, Y., Zhou, Q. et al. (2021a). A Machine Learning Algorithm To Detect Pine Wilt Disease Using UAV-Based Hyperspectral Imagery And LiDAR Data At The Tree Level. *Int. J. Appl. Earth Obs. Geoinf.*, 101. DOI: 10.1016/j.jag.2021.102363.

Yu, R., Ren, L., and Luo, Y. (2021b). Early Detection Of Pine Wilt Disease In Pinus Tabuliformis In North China Using A Field Portable Spectrometer And UAV-Based Hyperspectral Imagery. *Forest Ecosyst.,* 8. DOI: 10.1186/s40663-021-00328-6.

Zhang, T., Ramakrishnan, R. and Livny, M. (1997). Birch: A New Data Clustering Algorithm And Its Applications. *Data Mining And Knowledge Discovery*, 1(2), 141–182. DOI: 10.1023/A:1009783824328.

Zhang, Z. (2016). Introduction To Machine Learning: K-Nearest Neighbors. *Ann. Trans. Medic.*, 4(11), 218–218. DOI: 10.21037/atm.2016.03.37.