

URBAN VEGETATION CLASSIFICATION WITH HIGH-RESOLUTION PLANETSCOPE AND SKYSAT MULTISPECTRAL IMAGERY

LORÁND SZABÓ^{1*}, DÁVID ABRIHA¹, KWANELE PHINZI¹, SZILÁRD SZABÓ¹

Received 18 March 2021, accepted in revised form 28 April 2021



¹ Department of Physical Geography and Geoinformatics, Institute of Earth Sciences, University of Debrecen, Debrecen, Hungary

*Email: szabo.lorand@science.unideb.hu

Abstract

In this study two high-resolution satellite imagery, the PlanetScope, and SkySat were compared based on their classification capabilities of urban vegetation. During the research, we applied Random Forest and Support Vector Machine classification methods at a study area, center of Rome, Italy. We performed the classifications based on the spectral bands, then we involved the NDVI index, too. We evaluated the classification performance of the classifiers using different sets of input data with ROC curves and AUC values. Additional statistical analyses were applied to reveal the correlation structure of the satellite bands and the NDVI and General Linear Modeling to evaluate the AUC of different models. Although different classification methods did not result in significantly differing outcomes (AUC values between 0.96 and 0.99), SVM's performance was better. The contribution of NDVI resulted in significantly higher AUC values. SkySat's bands provided slightly better input data related to PlanetScope but the difference was minimal (~3%); accordingly, both satellites ensured excellent classification results.

Keywords: vegetation classification, high-resolution, satellite imagery, Planet, SkySat, urban vegetation, NDVI, ROC curve, classification performance, Random Forest, Support Vector Machine

1. Introduction

Remote sensing is one of the fast-growing geospatial technologies, progressively influencing a wide range of areas such as commerce, science, and applied research as well as public policy (Estes et al. 2001). By definition, remote sensing relates to the science and technology of obtaining information about the earth's surface without any direct physical contact (Campbell and Wynne, 2011). The field of remote sensing has evolved considerably. For several decades, aerial photographs, acquired from

airborne vehicles, have been the main source of information, until the early 1970s, satellite images have emerged as an alternative to aerial photography for earth's surface observation (Szabó et al. 2018). One of the advantages of satellite images over aerial images is the spatio-temporal characteristics, which permit for wide-area mapping on a temporal basis. Already, successful application of satellite images in various areas has been demonstrated i.e., land use/land cover monitoring (Kraas, 2007; Singh et al. 2015; Szabó et al. 2016; Phinzi and Ngetar, 2019; Gyenizse et al. 2020), urban

applications (Streutker, 2002; Li et al. 2009; Szabó Z. et al. 2019; Paramita and Matzarakis, 2019), agriculture (Atzberger, 2013), water quality monitoring (Chebud et al. 2012), drought monitoring (Gulácsi and Kovács, 2018) mapping at wetlands (Szabó et al. 2020; Van Leeuwen et al. 2020) and erosion risk assessment (Bakacsi et al. 2019; Phinzi et al. 2020).

Recently, there has been increased interest in the use of remote sensing for urban vegetation mapping. Such increased interest in remote sensing of urban vegetation is a direct response to global climate change, a major burden for cities (Kraas, 2007). The increasing number of urban dwellers, traffic congestion, and the ongoing urban heat island effect, all have a direct bearing on global climate change. In the face of such urban problems and their subsequent contributions to climate change, urban vegetation, trees, in particular, plays a critical role as an ecosystem service, mitigating climate change impacts (Pickett et al. 2011). The availability of accurate information on the spatial distribution of urban vegetation is the first, but important step towards addressing the aforementioned urban problems, ultimately reducing the effects of climate change.

Remote sensing, due to its spatial and temporal characteristic, offers tremendous opportunities for mapping urban vegetation, providing reliable and reproducible information on urban vegetation patterns across large areas (Melesse et al. 2007). However, urban vegetation assessment using remote sensing-based approaches is still faced with challenges, largely related to the spectral and spatial complexity of urban environments. The presence of a variety of vegetation types together with the pronounced 3D structure of urban environments with shadowing and obscured urban objects, as well as rapid temporal changes (Tigges et al. 2013), make it even more difficult to assess urban vegetation using remote sensing. Freely available satellite images like Landsat and Sentinel,

with 30 m and 10 m spatial resolutions, respectively, are probably the most widely used for vegetation mapping over large areas but they are not suitable for detailed urban vegetation mapping because of their relatively low spatial resolutions (Burai et al. 2019).

Although Sentinel-2 has the best spatial resolution among the freely available multispectral datasets, the © Planet Labs Inc. (Planet Team, 2017) provides PlanetScope imagery free for educational and research applications. The PlanetScope is a constellation with more than 180 lunchbox multispectral satellites with 3 m spatial resolution and daily revisit time (Shendryk et al. 2019). © Planet Labs Inc. has another ultra-high (1 m multispectral and 0.8 m panchromatic) spatial resolution satellite constellation, the SkySat satellites. SkySat is a commercial database, but it has a few sample scenes ready to use.

The use of PlanetScope and SkySat images has not been widely reported in the literature, especially their capabilities in vegetation classification. The aim of this study is to compare the classification capabilities of the urban vegetation using these high spatial resolution satellite imageries. The main goal is to reveal that a) the images provide accurate maps of vegetation cover; b) which classifier provides the best classification accuracy; c) does the NDVI improve the classification results?

2. Materials and Methods

Study area

The study area is located in the center of Rome, Italy. The area is 30 km² and contains mostly urban built-up areas, green parks, and the River Tevere. There is very little amount of vegetation in the built-up areas, especially in the city center. Most of the vegetation is located in the green parks and along the main roads and the river.

Satellite imagery

In this study, we used two high-resolution satellite imageries, both from the Planet Labs database (Planet Team, 2017) and with the same capture date 28th August 2018. First is PlanetScope satellite imagery. PlanetScope has four spectral bands, blue, green, red, and near-infrared. The spatial resolution is 3 m, and the constellation has a daily revisit time. The second is a SkySat scene. SkySat also has four multispectral bands (blue, green, red, and near-infrared) with 1 m spatial resolution and a panchromatic band with 0.8 m spatial resolution. In this study, we used the pan-sharpened multispectral dataset with 0.8 m spatial resolution.

Reference data

We applied the binary approach; thus, reference data were collected as vegetation and non-vegetation classes. We gathered the data from the same spots in the images. In the vegetation class, we included pixels from trees and herbaceous plants as well. The “non-vegetation” class incorporated more diverse land cover types for e.g. buildings, roads, water surfaces etc.

Classification methods

Analyses were conducted with two supervised classification algorithms: Random Forest (RF) and Support Vector Machine (SVM). RF is an ensemble learning method that uses multiple (i.e. 100-500) decision trees to make predictions. Class labels are assigned based on the majority votes of the decision trees (Belgiu and Drăguț, 2016; Breiman, 2001). The basic idea behind SVM is to find a line (hyperplane) which separates the classes, however, there could be infinite possible lines to do this, so the algorithm’s goal is to find the optimal hyperplane by maximizing the margin between the support vectors (Chapelle et al., 1999).

Evaluation of the classification performance

In this study, we performed ROC curves to evaluate the accuracy of our classifications. The method represents the tradeoff between false-positive and true-positive rates (McClish, D. K., 1989). The Area Under the Curve (AUC) values represent the classification quality.

Statistical analysis

We determined the correlations among the satellite bands and NDVI with the Pearson correlation test; furthermore, as a quantified comparison tool, we used the Cronbach’s alpha, a measure of internal consistency (0 indicates the lack of correlations, and when the shared covariances increase, Cronbach’s alpha approaches 1).

We evaluated the AUC values (as dependent variable) with General Linear Modelling (GLM), and determined the explained variance by the following factors (as independent variables): satellite type, classifier, and the usage of NDVI as an additional variable in the classification. We also determined the statistical interactions among the factors (i.e. to reveal if a factor influences the effect of another). Effect sizes (ω^2) had been calculated to quantify the contribution of the variables in the model (Field, 2013; Rotigliano et al., 2018).

3. Results

Classification by spectral bands and by involving NDVI

Evaluating the classification performances of the Planet data by the ROC curve (Fig. 1.) show that SVM had 0.97 AUC, better than RF (0.96 AUC) classifying with the spectral bands. Involving the NDVI into the classification the RF reached 0.98 AUC (Fig. 3.), better than SVM (0.97 AUC).

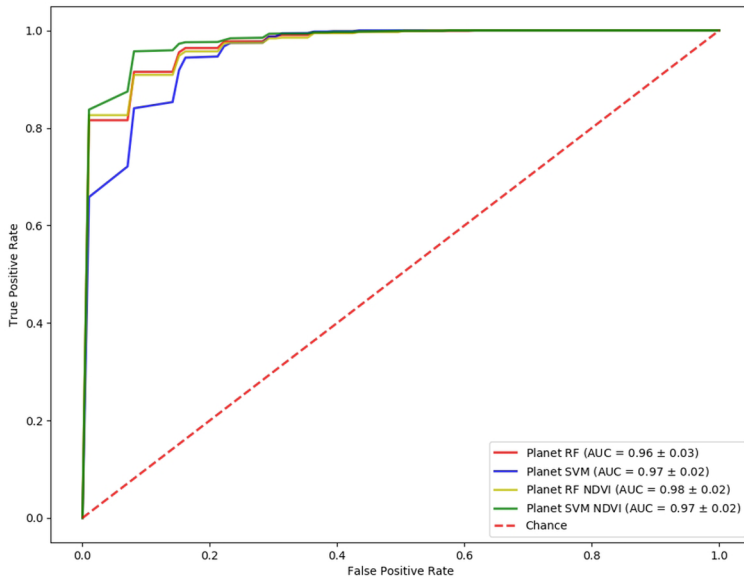


Fig. 1. ROC curves of the Planet classifications

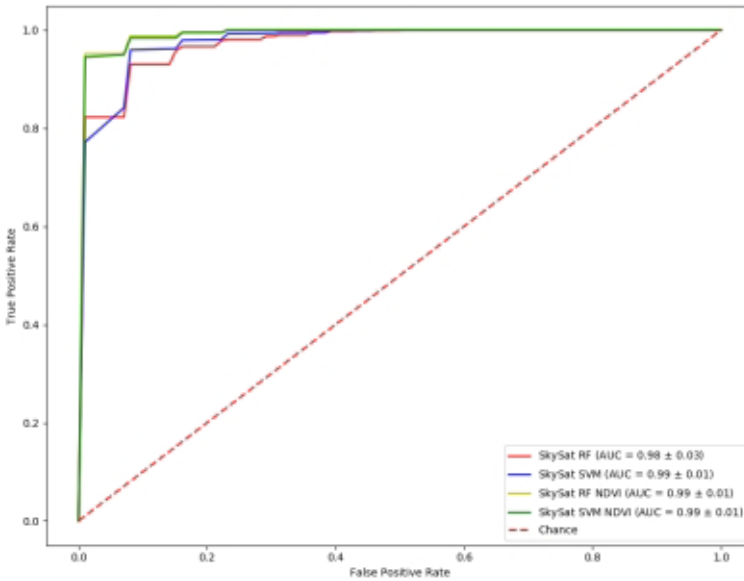


Fig. 2. ROC curves of the SkySat classifications

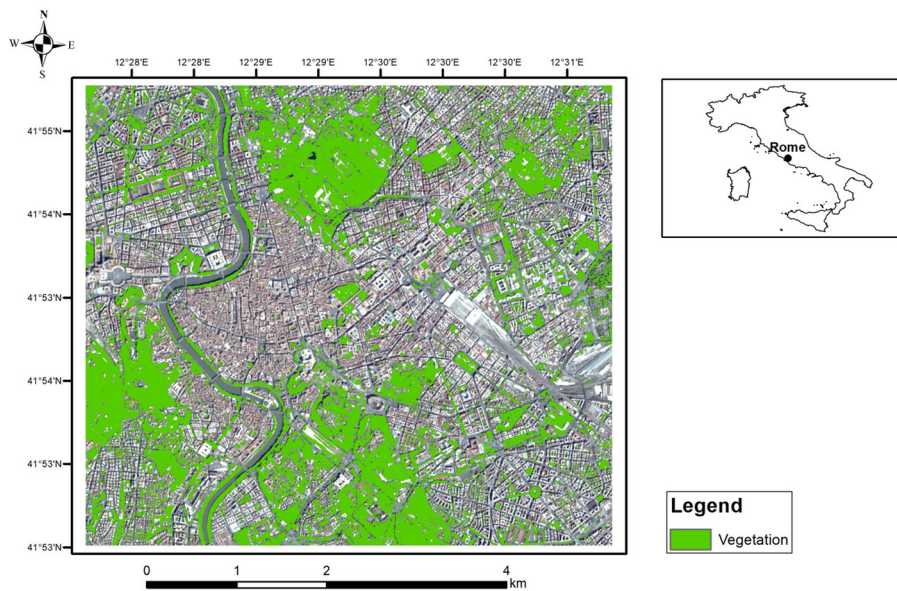


Fig. 3. Extracted vegetation of the Planet (input data: bands+NDVI, classifier: Random Forest)

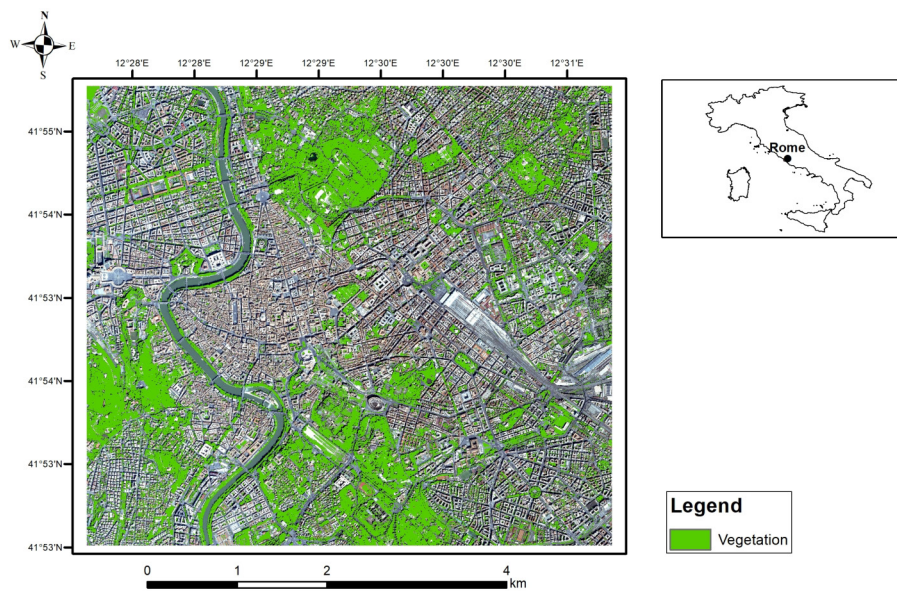


Fig. 4. Extracted vegetation of the SkySat (input data: bands+NDVI, classifier: Random Forest)

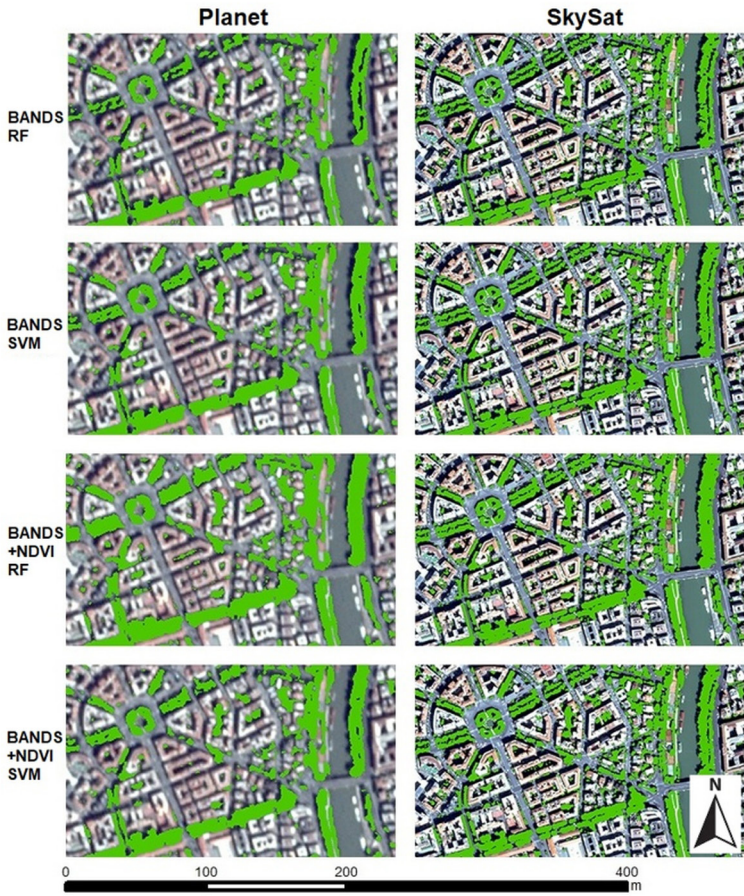


Fig. 5. Samples of each classification. (Left column: Planet, right column: SkySat; input: BANDS or BANDS+NDVI; classifier: RF - Random Forest, SVM - Support Vector Machine)

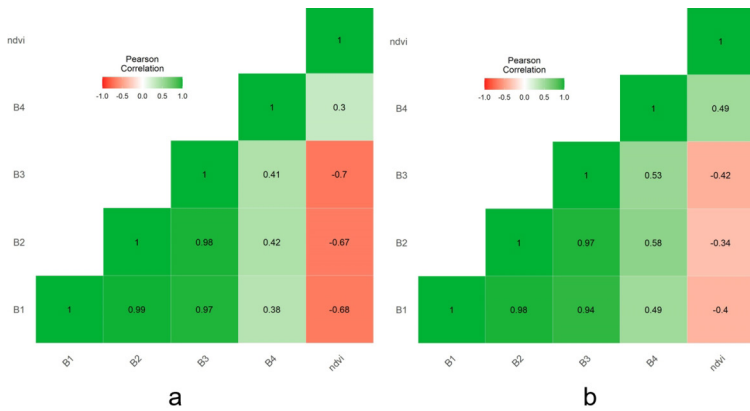


Fig. 6. Correlation matrices of Planet (a) and SkySat (b) bands and the derived NDVIs

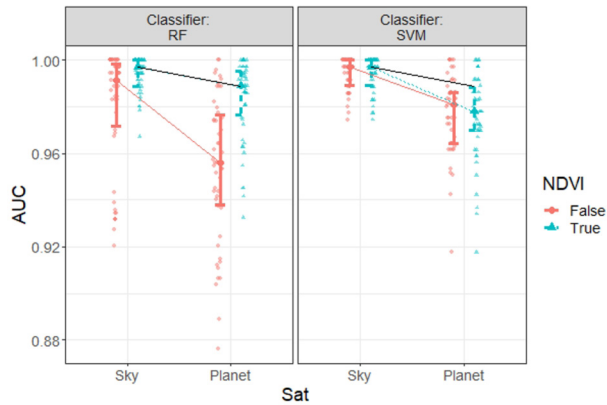


Fig. 7. AUC values of 50 cross-validated samples by satellites (SAT) classification algorithms, (RF or SVM) and involved data (NDVI, False: original bands, True: original bands + NDVI)

Table 1. GLM of 50 models of cross-validation (SS: Sum of Squares, df: degree of freedom, F: F-statistic, p: significance, ω^2 : effect size; SAT: satellite type)

	SS	df	F	p	ω^2
Model	0.06410	7	27.39161	<.001	0.316
Classifier	0.00539	1	16.11659	<.001	0.026
Sat	0.03266	1	97.69124	<.001	0.165
NDVI	0.01113	1	33.29940	<.001	0.055
Classifier × Sat	2.15e-6	1	0.00642	0.936	-0.002
Classifier × NDVI	0.01193	1	35.67160	<.001	0.059
Sat × NDVI	0.00124	1	3.72352	0.054	0.005
Classifier × Sat × NDVI	0.00175	1	5.23251	0.023	0.007
Residuals	0.13105	392			
Total	0.19515	399			

In the case of the SkySat, the RF classification with the spectral bands had 0.98 AUC. The same classification using SVM, and involving NDVI using both classifiers reached 0.99 AUC as well (Fig. 2, 4.). Figure 5. shows samples of the classified maps.

Correlations among the bands and NDVI

NDVI differently correlated with the original bands of the two satellites (Fig. 6.). While the correlation differences were minimal (0.01) in the case of B1-B3 bands, these were ~0.1 in the case of B4 and 0.2-0.3 between NDVIs. Cronbach’s alpha was 0.819 for the Planet and 0.827 for the SkySat

images. Accordingly, both values indicated high consistency, in the case of SkySat it was slightly higher.

Effects of factorial variables on the AUC

GLM revealed that type of satellites, classifiers and the involvement/omitting of NDVI explained 31.6% (adjusted $R^2 = 0.316$) of the variance. Although most of the factors and their interactions were significant (except the interaction of satellite type and the classifier, and the interaction of satellite type and the inclusion of NDVI), effect sizes indicated large effect only in case of satellite type, and all other factors had a small effect

(Table 1.). Comparison by factors (Fig. 7.) showed that on average Skysat had 0.018 higher AUC values than Planet ($t = -9.88$, $df = 392$, $p < 0.001$), inclusion of NDVI improved AUC by 0.011 ($t = -5.77$, $df = 392$, $p < 0.001$), and SVM provided 0.007 better AUC ($t = -4.01$, $df = 392$, $p < 0.001$).

4. Discussion

Classifying the Planet and SkySat data showed that SVM provided slightly better AUC values than RF in case of the classification with the spectral bands. When we involved the NDVI into the classification RF provided slightly better results in case of the Planet. SkySat, except the RF classification of the spectral bands provided 0.99 AUC, the best outcomes in this evaluation method. Involving NDVI improves the classification performance in almost all cases. The SkySat performed better than the Planet, although the RF classification of the Planet involving NDVI also reached the same AUC than the SkySat classified the spectral bands with RF.

Several studies proved that RF outperforms, at least slightly, SVM (Schlosser et al. 2020, Szabó L. et al. 2019, Liu et al. 2013), in this case SVM provided better accuracy. Although the difference of AUC was only small (0.007), but significant. This is a special case of statistical analysis when the significant differences are not completely useful: the difference is true, but the magnitude between the groups is rather small; thus, effect size can be a more useful metric than the p-value. The ω^2 indicated a small effect; accordingly, the two classifiers can be characterized with almost similar accuracies.

These satellites are rather new ones, thus, the literature is not wide of them, and there is not much experience with their applicability. Zeng et al. (2018) applied the Planet satellites in South Asia to reveal the expansion of croplands against forests; however, the Planet images were used as auxiliary data. Olthof and Svacina (2020) applied Planet images in flood mapping, and they used

these data to determine the maximum flood extent to serve as a validation dataset in their flood simulations. Shendryk et al. (2019) also found Planet images efficient to filter out clouds and shadows in land cover mapping. Planet images were successfully applied in oil spill detection (Park et al. 2019), and in vegetation mapping (Gašparović et al. 2018), in this latter case, Planet performed ~5% better than Sentinel-2 in vegetation mapping. Terra Bella's SkySat (a previous generation of this type of satellites with 2 m resolution) was used in smallholder (<0.3 ha) plots to predict crop yield and served appropriate data (Jain et al. 2016). Generally, similarly to our findings, all previous studies found microsatellites useful in environmental mapping.

The contribution of NDVI resulted in a significant difference in AUC values; nevertheless, this was only ~1%. In this case the ω^2 was 0.05, which indicated medium effect. NDVI had different correlations with the original bands, and in case of Skysat correlations were only moderate. However, the interaction between satellite type and the involvement of NDVI into the set of predictors had only a very small effect ($\omega^2=0.005$). It means that in spite of the differences in the correlation matrices (i.e. NDVI provided different predictors by satellites), the AUC did not result in significantly differing outcomes with high magnitude of differences; differences were insignificant ($p=0.054$).

5. Conclusions

In this study the PlanetScope and SkySat satellite imageries were compared based on their classification capabilities of urban vegetation. Study area was in the city centre of Rome, Italy. We classified the imageries using RF and SVM classification methods using the original bands of the satellites and the NDVI indices. We found that satellite bands resulted in better outcomes using SVM than RF, and adding NDVI provided higher AUC values. AUC values were slightly

better (0.04) by using the SkySat imagery than Planet which was also high (AUC=0.96). While the different classification methods did not result in significantly differing outcomes (AUC values between 0.96 and 0.99), the contribution of NDVI caused significantly higher AUC.

Acknowledgement

The publication is supported by the EFOP-3.6.1-16-2016-00022 project. The project is co-financed by the European Union and the European Social Fund.

Abriha, D. was supported by the Doctoral Student Scholarship Program of the Co-operative Doctoral Program of the Ministry of Innovation and Technology financed from the National Research, Development and Innovation Fund.

6. References

- Atzberger, C. (2013): Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote sensing*, 5(2), 949-981.
- Bakacsi, Z. – Tóth, T. – Makó, A. – Barna, G. – Laborczi, A. – Szabó, J. – Pásztor, L. (2019): National level assessment of soil salinization and structural degradation risks under irrigation. *Hungarian Geographical Bulletin*, 68(2), 141-156.
- Belgiu, M., – Drăguț, L. (2016): Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24-31.
- Breiman, L. (2001): Random forests. *Machine learning*, 45(1), 5-32.
- Burai, P., Beko, L., Lenart, C., Tomor, T., Kovacs, Z. (2019). Individual Tree Species Classification Using Airborne Hyperspectral Imagery and Lidar Data. In 2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS) (pp. 1-4). IEEE.
- Campbell, J. B. – Wynne, R. H. (2011): Introduction to remote sensing. Guilford Press.
- Chapelle, O. – Haffner, P. – Vapnik, V. N. (1999): Support vector machines for histogram-based image classification. *IEEE transactions on Neural Networks*, 10(5), 1055-1064.
- Chebud, Y. – Naja, G. M. – Rivero, R. G. – Melesse, A. M. (2012): Water quality monitoring using remote sensing and an artificial neural network. *Water, Air, & Soil Pollution*, 223(8), 4875-4887.
- Estes, J. – Kline, K. – Collins, E. (2001): Remote Sensing. In *International Encyclopedia of the Social & Behavioral Sciences*; Neil J. Smelser, Paul B. Baltes., Eds.; Pergamon, pp. 13144-13150.
- Field, A. (2013): *Discovering statistics using IBM SPSS statistics*. sage.
- Gašparović, M. – Medak, D. – Pilaš, I. – Jurjević, L. – Balenović, I. (2018): Fusion of Sentinel-2 and PlanetScope Imagery for Vegetation Detection and Monitoring. In *Volumes ISPRS TC I Mid-term Symposium Innovative Sensing-From Sensors to Methods and Applications*.
- Gulácsy, A. – Kovács, F. (2018): Drought monitoring of forest vegetation using MODIS-based normalized difference drought index in Hungary. *Hungarian Geographical Bulletin*, 67(1), 29-42.
- Gyenezse, P. – Lóczy, D. – Dezső, J. – Pirkhoffer, E. – Słowik, M. (2020): Adaptation of land use based on the assessment of inundation risk in the Kapos Valley, Southwest Hungary. *Hungarian Geographical Bulletin*, 69(3), 299-316.
- Jain, M. – Srivastava, A. K. – Joon, R. K. – McDonald, A. – Royal, K. – Lisaius, M. C. – Lobell, D. B. (2016): Mapping smallholder wheat yields and sowing dates using micro-satellite data. *Remote sensing*, 8(10), 860.
- Kraas, F. (2007): Megacities and global change: key priorities. *The Geographical Journal*, 173(1), 79-82.
- Li, J. J. – Wang, X. R. – Wang, X. J. – Ma, W. C. – Zhang, H. (2009): Remote sensing evaluation of urban heat island and its spatial pattern of the Shanghai metropolitan area, China. *Ecological Complexity*, 6(4), 413-420.
- Liu, M. – Wang, M. – Wang, J. – Li, D. (2013): Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical*, 177, 970-980.
- McClish, D. K. (1989): Analyzing a portion of the ROC curve. *Medical Decision Making*, 9(3), 190-195.
- Melesse, A. M. – Weng, Q. – Thenkabail, P. S. – Senay, G. B. (2007): Remote sensing sensors and applications in environmental resources mapping and modelling. *Sensors*, 7(12), 3209-3241.

- Olthof, I. – Svacina, N. (2020): Testing Urban Flood Mapping Approaches from Satellite and In-Situ Data Collected during 2017 and 2019 Events in Eastern Canada. *Remote Sensing*, 12(19), 3141.
- Paramita, B. – Matzarakis, A. (2019): Urban morphology aspects on microclimate in a hot and humid climate. *Geographica Pannonica*, 23(4), 398-410.
- Park, S. H. – Jung, H. S. – Lee, M. J. – Lee, W. J. – Choi, M. J. (2019): Oil spill detection from PlanetScope satellite image: application to Oil spill Accident near Ras Al Zour area, Kuwait in august 2017. *Journal of Coastal Research*, 90(SI), 251-260.
- Phinzi, K. – Ngetar, N. S. (2019): Land use/land cover dynamics and soil erosion in the Umzintlava catchment (T32E), Eastern Cape, South Africa. *Transactions of the Royal Society of South Africa*, 74(3), 223-237.
- Phinzi, K. – Ngetar, N. S. – Ebhuoma, O. (2020): Soil erosion risk assessment in the Umzintlava catchment (T32E), Eastern Cape, South Africa, using RUSLE and random forest algorithm. *South African Geographical Journal*, 1-24.
- Pickett, S. T. – Cadenasso, M. L. – Grove, J. M. – Boone, C. G. – Groffman, P. M. – Irwin, E., – Pouyat, R. V. (2011): Urban ecological systems: Scientific foundations and a decade of progress. *Journal of environmental management*, 92(3), 331-362.
- Planet Team (2017): Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. <https://api.planet.com>.
- Rotigliano, E. – Martinello, C. – Agnesi, V. – Conoscenti, C. (2018): Evaluation of debris flow susceptibility in El Salvador (CA): a comparison between Multivariate Adaptive Regression Splines (MARS) and Binary Logistic Regression (BLR). *Hungarian Geographical Bulletin*, 67(4), 361-373.
- Schlosser, A.D. – Szabó, G. – Bertalan, L. – Varga, Z. – Enyedi, P. – Szabó, S. (2020): Building Extraction Using Orthophotos and Dense Point Cloud Derived from Visual Band Aerial Imagery Based on Machine Learning and Segmentation. *Remote Sens.* 12, 2397.
- Shendryk, Y. – Rist, Y. – Ticehurst, C. – Thorburn, P. (2019): Deep learning for multi-modal classification of cloud, shadow and land cover scenes in PlanetScope and Sentinel-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 157, 124-136.
- Singh, S. K. – Mustak, S. – Srivastava, P. K. – Szabó, S. – Islam, T. (2015): Predicting spatial and decadal LULC changes through cellular automata Markov chain models using earth observation datasets and geo-information. *Environmental Processes*, 2(1), 61-78.
- Streutker, D. R. (2002): A remote sensing study of the urban heat island of Houston, Texas. *International Journal of Remote Sensing*, 23(13), 2595-2608.
- Szabó, G. – Bertalan, L. – Barkóczy, N. – Kovács, Z. – Burai, P. – Lénárt, C. (2018): Zooming on aerial survey. In *Small Flying Drones* (pp. 91-126). Springer, Cham.
- Szabó, L. – Burai, P. – Deák, B. – Dyke, G. – Szabó, Sz. (2019): Assessing the efficiency of multispectral satellite and airborne hyperspectral images for land cover mapping in an aquatic environment with emphasis on the water caltrop (*Trapa natans*). *INTERNATIONAL JOURNAL OF REMOTE SENSING*, 40 (13). pp. 5192-5215. ISSN 0143-1161
- Szabó, L. – Deák, B. – Bíró, T. – Dyke, G. J. – Szabó, S. (2020): NDVI as a Proxy for Estimating Sedimentation and Vegetation Spread in Artificial Lakes—Monitoring of Spatial and Temporal Changes by Using Satellite Images Overarching Three Decades. *Remote Sensing*, 12(9), 1468.
- Szabó, L. – Deák, M. – Szabó, Sz. (2016): Comparative analysis of Landsat TM, ETM+, OLI and EO-1 ALI satellite images at the Tisza-tó area, Hungary. *Acta Geographica Debrecina Landscape & Environment*, 10, 2, 53
- Szabó, Z. – Schlosser, A. – Túri, Z. – Szabó, S. (2019): A review of climatic and vegetation surveys in urban environment with laser scanning: A literature-based analysis. *Geographica Pannonica*, 23(4), 411-421.
- Tigges, J. – Lakes, T. – Hostert, P. (2013): Urban vegetation classification: Benefits of multitemporal RapidEye satellite data. *Remote Sensing of Environment*, 136, 66-75.
- Van Leeuwen, B. – Tobak, Z. – Kovács, F. (2020): Sentinel-1 and-2 Based near Real Time Inland Excess Water Mapping for Optimized Water Management. *Sustainability*, 12(7), 2854.
- Zeng, Z. – Estes, L. – Ziegler, A. D. – Chen, A. – Searchinger, T. – Hua, F. – Wood, E. F. (2018): Highland cropland expansion and forest loss in Southeast Asia in the twenty-first century. *Nature Geoscience*, 11(8), 556-562.