LAND COVER CHANGE DETECTION NEAR SMALL WATER BODIES BASED ON RGB UAV DATA: CASE STUDY OF THE POND BAROCH, CZECH REPUBLIC

Jakub Jech*, Jitka Komarkova, Pavel Sedlak

Faculty of Economics and Administration, University of Pardubice, Czech Republic, (jakub.jech, jitka.komarkova, pavel.sedlak)@upce.cz

KEY WORDS: UAV, RGB Data, Maximum Likelihood, Support Vector Machine, Deep Learning, Supervised classification.

ABSTRACT:

Monitoring changes of land cover near water bodies and water bodies themselves represents a part of environment protection and management. The management can be done at the global or local level. The local level requires more detailed data, which can be collected i.e. by means of aircraft or UAV. The paper describes a case study focused on the utilization of UAV-based RGB data to monitor land cover near the pond Baroch, which is located in the Czech Republic, near the city of Pardubice. The area is specific – it is a small pond accompanied by several smaller pools and connecting canals and surrounded by meadows (often watered), reeds, bushes and some trees Used data were collected by authors by in advance planned flights in August, September, October, November, and December 2021. Support Vector Machine, Maximum Likelihood, Random Trees, and Deep Learning are used as methods to process data and detect land cover changes. Manually interpreted data are used as reference data. Because of the nature of the data (only R, G, and B bands), classification into bare land, the water, vegetation, dry vegetation, and wet vegetation classes only was used. Very high heterogeneity of the observed area, availability of RGB bands only, and very high spatial resolution (1,9 cm per pixel) led to isolated cells.

1. INTRODUCTION

The human population modifies its environment. The scope and nature of changes have increased along with the growth of the population and used technologies. The Earth surface transformation belongs to the key changes. Agriculture, cropping, forestry, deforestation, and urbanization can be given as examples of such changes (Vitousek et al. 1997).

Land cover mapping and land cover change detection require serious data and information to make evidence-based decisions. Land cover/land use identification and change detection can provide important information to support environment management. Remotely sensed data has been successfully used for land cover/land use identification since the 70s along with the Landsat satellite system existence (Bukata, Harris and Bruton, 1974; Saharan et al., 2018). Potapov et al. (2020) have proposed and impemented algorithms and web portals to provide Landsat Analysis Ready Data. The data are generated automaticly in a consistent way so they can be used as an input for land cover/land use identification and change analyses.

Many various methods can be used for land cover identification. Wang et al (2020) identified the following key categories of methods: a) rule-based approaches (e.g. decision trees); b) data-driven approaches (e.g. k-nearest neighbor – k-NN, support vector machine – SVM, and Deep Learning – DL); c) reinforced learning (e.g. Q-learning); d) ensemble methods based on use of multiple models.

For local management, which needs to focus on smaller areas, e.g. on on small water bodies, very high spatial resolution and very low elevation data can be very important because of the required high level of detail. An unmanned aerial vehicle (UAV), called drone or remotely piloted aircraft system (RPAS) too, is a suitable carrier for such data collection.

UAVs are increasingly used due to their ability to allow a fast area sensing in the required time, i.e. on demand, and to provide data with a very high spatial resolution at the same time. The price of imagery, especially in the case of a long-term imaging of a small area, may be significantly lower than of the traditional data sources (satellite and aircraft imagery). The widespread of UAVs as a data source is another reason for choosing this device. At the same time, legal and weather conditions must be respected too (Salamí, Barrado and Pastor, 2014; Gallop et al., 2015; Pásler, Komárková and Sedlák, 2015; Iizuka et al., 2018).

Low cost (or older) UAVs are usually equipped with RGB (only R, G, and B bands) camera. It means, specific methods and approaches have to be used, as far as other spectral bands are missing. NIR (near-InfraRed) is the very typical example because NIR band is important for usual vegetation spectral indices calculation, e.g. the normalized difference vegetation index (NDVI).

Many spectral indices based on RGB bands only have been proposed to allow observation of vegetation, including its health status and volume, e.g. ExG – Excess Green, NExG – Normalized Excess Green (both Woebbecke et al., 1995), ExR – Excess Red (Meyer, Hindman and Laksmi, 1999), GLI – Green Leaf Index (Louhaichi, Borman and Johnson, 2001), VARI – Visible Atmospherically Resistant Index (Gitelson et al., 2002), CIVE – Colour Index of Vegetation Extraction (Kataoka et al., 2003), ExG - ExR difference (Meyer et al., 2004), RGBVI – Red-Green-Blue Vegetation Index (Bendig et al., 2015), and many others.

Classification methods, both supervised and unsupervised, can be used for RGB datasets themselves and for calculated spectral

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-617-2022

indices. Classification of spectral index can lead to better results, as shown e.g. by Komarkova and Jech (2020).

The unsupervised methods cluster pixels according to the similarity of a given classification class. ISODATA and K-means belong to the frequently used methods. The number of output classes is their only input parameter (Bijeesh, 2020).
The supervised classification starts with choice of appropriate training set. The training set represent the correct outputs classes of the classification. The training set is then used to classify the test set (Zhao and Liu, 2007). Choice of the good training set is time demanding, e.g. to prevent choice of overlaps. Maximum Likelihood represents very popular method, which has been successfully used in a long term (Stigler et al., 2007). Support Vector Machine (SVM) has been lately popular too but it is more suitable for smaller data sets (Mountrakis et al., 2011).

The paper is focused on land cover classification of small water bodies and their surroundings using UAV-borne RGB data. Small water bodies are chosen because they are often not monitored with the necessary attention although they represent an important landscape element – as a water source (Heine et al., 2015), and can cause a significant damage in the case of a natural disaster too.

Within our case study we focus on an atypical small water body – a small pond accompanied by several smaller pools and connecting canals. They are surrounded by a meadow (often watered), reeds, bushes and partly trees. Land cover changes have started to be observed in the area along with introduction wild horses as a mean of environment protection and management in this area.

The main aim of the paper is to compare suitability of SVM, Maximum Likelihood, Random Trees, and Deep Learning for UAV-borne RGB data classification for land cover identification near a small water body. We demonstrate results within a case study focused on the pond Baroch.

2. DATA AND METHODS

2.1 Used Procedure and Software

The whole procedure consists of the following steps:

- Area of interest definition
- Repeated data collection by means UAV
- Data processing
- Result interpretation and visualisation

Pix4D mapper (version 4.13.1) software for iOS was used for the flights planning. ArcGIS Pro (version 2.9) was used for data processing and visualisation.

2.2 Area of Interest

The study is focused on a small area during time series of 5 records. Pond Baroch was chosen as the area of interest (see Figure 1). Pond Baroch is a part of nature reservation with evidence number 1926. The pond is situated south to southwest of the village Hrobice near the city of Pardubice, the Czech Republic. The area is managed by the Regional Office of the Pardubice Region. The size of nature reservation is around 30 000 square meters. The reason to protect this nature reservation is a grounded pond, adjacent reeds, forest and meadow communities, ornithological locality.

The pound Baroch consists in fact of several smaller pools, which are connected by water canals. They pools are surrounded by meadow (often watered), reeds, bushes, and mostly individual trees. Wild horses were moved to the are to manage the area and protect rare kinds of vegetation.

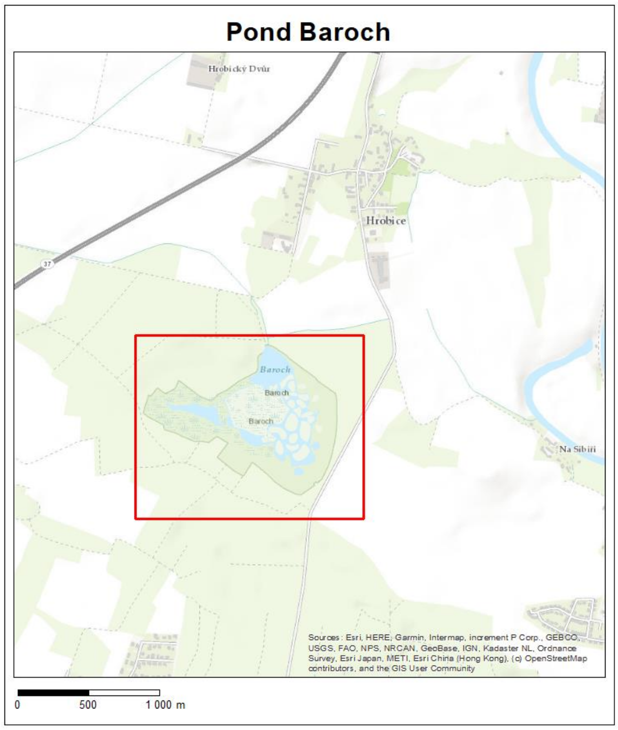

Figure 1. Localization of pond Baroch (Esri, 2022)

2.3 Data Collection

DJI Mavic 2 DUAL ENTERPRISE was used for data collection. The key characteristics of this UAV are as follows (DJI, 2022):

- Weight 899 g
- 4 motors
- Max. flight time approximately 31 minutes

The drone contains build-in ultra-HD and thermal camera. The visible camera is equipped with an F2.8 lens with a 24mm focal length and a viewing angle of 85 degrees. The build-in camera is assigned to drone by 3-axis gimbal (DJI, 2022).

The planned flight was used for comprehensive data collection. The planned flight ensures low data redundancy, selects the optimal points for capturing images (waypoints) and ensures a smooth flight of the desired area. Additionally, it ensures that the same area is monitored. The flight was planned in Pix4D mapper software for iOS and planned according to the rules for obtaining high quality data (Pix4D, 2022). Data from the planned flight led to individual images with 60 % overlay with each other. The flight altitude was 60 m. Also, the flight itself is subject to the legislation on the operation of unmanned aerial vehicles. The rules on flying are administered by the Civil Aviation Authority – CAA (CAA CZ, 2022).

Data were obtained in the same day time, around 11 AM, during August, September, October, November and December, year 2021. The date of data acquisition was around 25 of each month. To ensure the same quality data, the flight was planned in advance with respect to light condition. Similar light conditions during the data collection support quality of data (Buyukdemircioglu and Kocaman, 2021). Additionally, data collection around noon allows to minimize shadows. No atmospheric corrections were done because of the low flight level. Figure 2 shows one of the output mosaics. The area of interest was limited to the south part (see Figure 3).

Figure 2. Mosaic of the Pond Baroch obtained by UAV

2.4 Processing Methods

This paper is focused on supervised methods for image classification. Maximum Likelihood, SVM, and Deep Learning were used. Maximum Likelihood represents the most common supervised (Stigler et al., 2007) method and SVM provided good results in previous study (Komarkova and Jech, 2020).

Maximum likelihood (Stigler et al., 2007) classification assumes that the statistics for each class in each band are normally distributed and calculates the probability that a given pixel belongs to a specific class. Unless a probability threshold is chosen, all pixels are classified. Each pixel is assigned to the class that has the highest probability (that is, the Maximum Likelihood). If the highest probability is smaller than a specified threshold, the pixel remains unclassified. ArcGIS Pro maximum likelihood method using input data of raster image and signature file, which contained training sets (2 samples per each class), was used. The default ArcGIS Pro settings were used.

Support Vector Machine is a supervised learning model with associated learning algorithms that analyse data for classification and regression analysis (Lazar and Shellito, 2009). SVM is one of the most robust prediction methods, being based on statistical learning frameworks or Vapnik–Chervonenkis theory (Vapnik and Chervonenkis, 1974). Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. SVM tool (ArcGIS PRO, 2016) maps training examples to points in space so as to maximise the width of the gap between the two categories. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall into. ArcGIS Pro default settings were used.

Deep Learning has been increasingly used in remote sensing for land cover identification (Ma et al., 2019). ArcGIS Pro default settings were used.

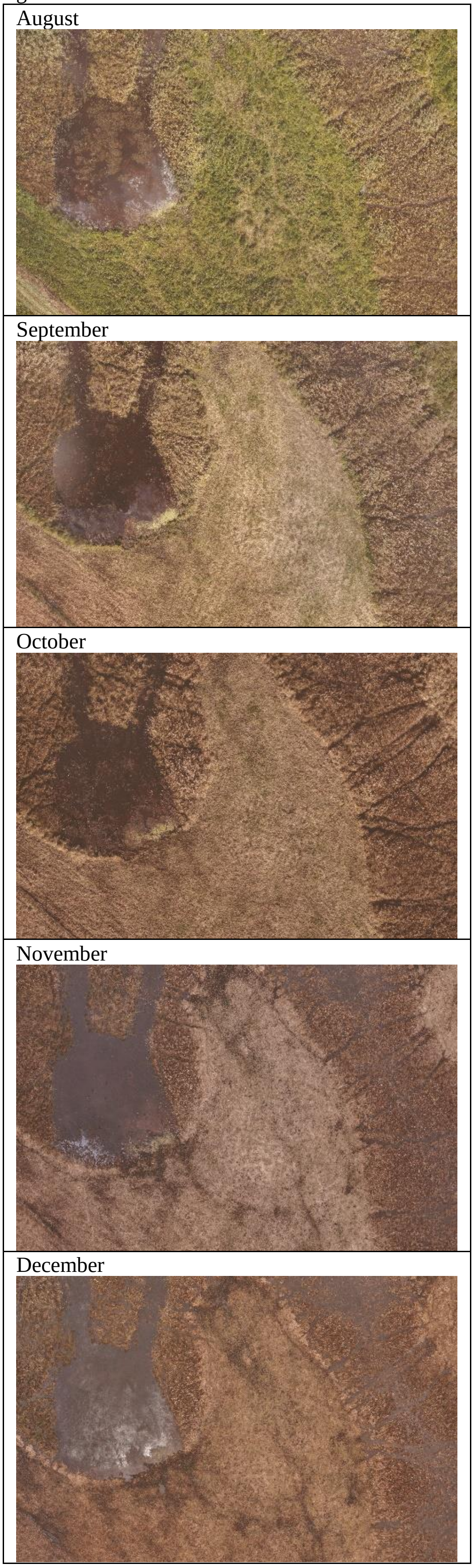

Figure 3. Data time series of Baroch

Random Trees: ArcGIS Pro default settings were used; the only input was the training set.

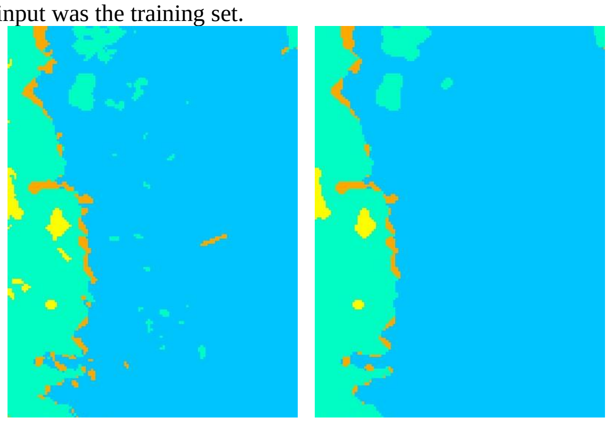

Figure 4. Comparison of Majority filter use

The target classes were set to: water (clear water surface), green vegetation, wet vegetation (vegetation in flooded areas), bareland, dry vegetation, freshly cutted vegetation (withered, i.e. its colour was different).

Images obtained by UAV have very high spatial resolution, specifically 1,9 cm per pixel. For this reason, filtration was applied to classified images to remove individuals classify pixel to others group. Majority filter was performed on classified images, but it had to be run at least 12 times to provide the first visible results, see Figure 4. Because of very high spatial resolution we have decided to not use majority filter in the end.

3. RESULTS AND DISCUSSION

Results from supervised classification methods (Maximum Likelihood and Support Vector Machine) are shown by Figure 5; Figure 6 shows legend for the classes.

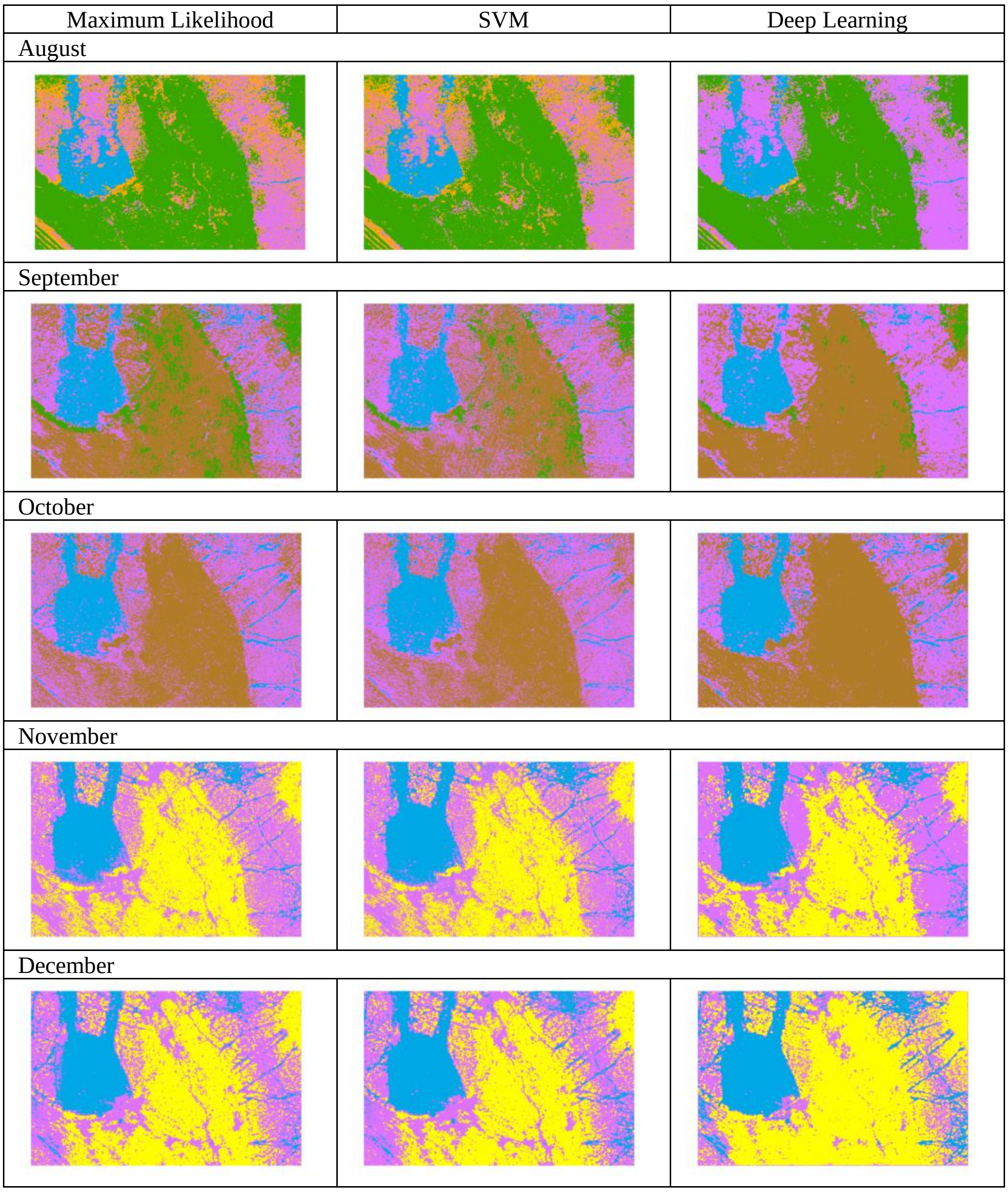

Figure 5. Results from supervised methods

Figure 6. Legend of classes

Kappa coefficient was calculated for each output. The best classification according to the Kappa was for SVM in August, with the value 0,56; for Maximum Likelihood in August with the value 0,55; for Deep Learning in September with the value 0,74. Reason for lower Kappa can be very high spatial resolution and heterogeneity of the area, which can lead to many individual pixels. Comparison of identified classes by count and manual identification shows Figure 7. It shows that classification into water, vegetation, dry vegetation, and wet vegetation classes can be done for RGB data with acceptable results.

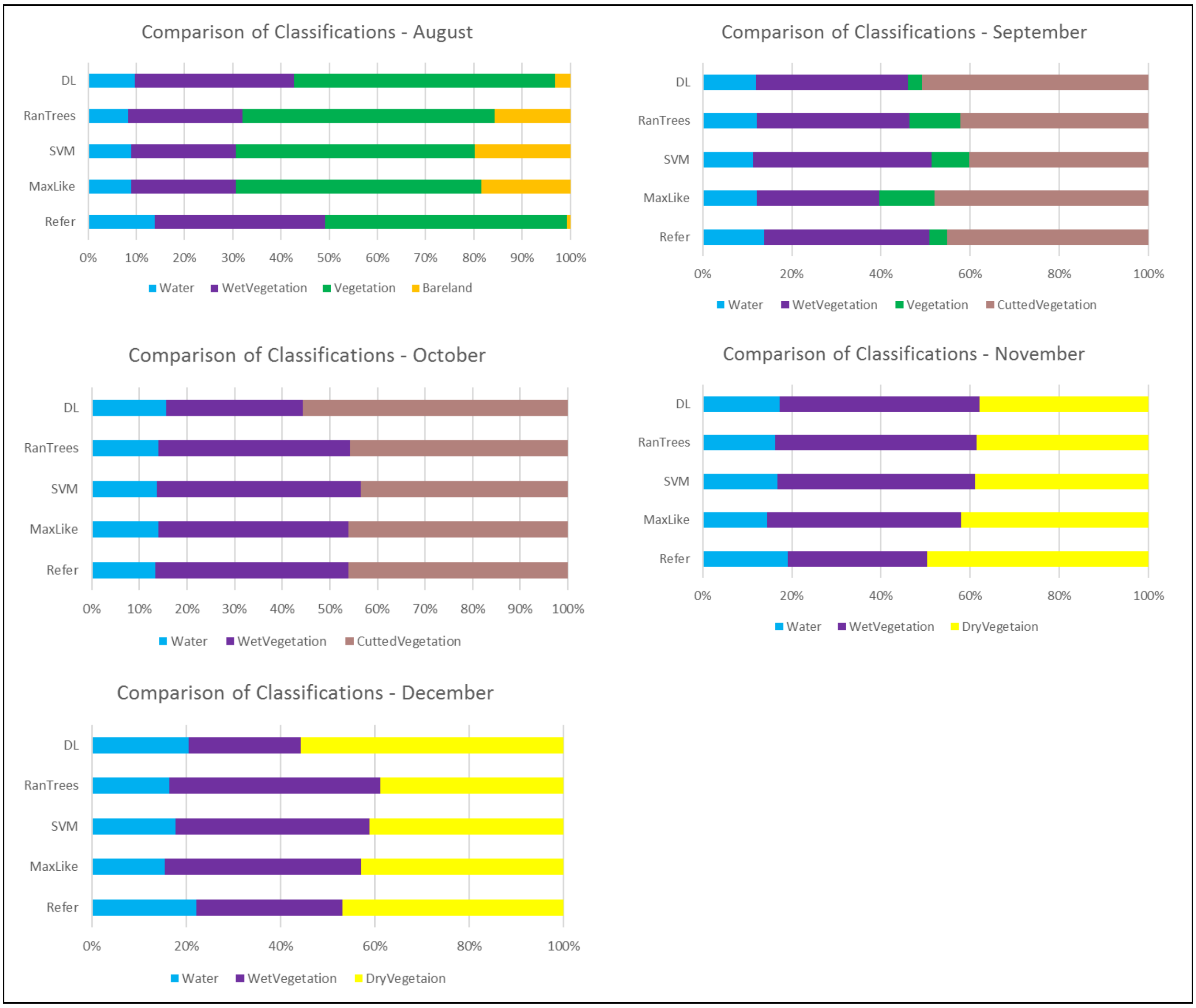

Figure 7. Comparison of classifications results

4. CONCLUSIONS

Environment protection and management has an increasing importance. This applies to the water management too. Small water bodies belong to important local water sources but it is quite complicated to identify them well (i.e. precisely) from freely available satellite imagery like Landsat.

Data obtained by an UAV provide very high spatial resolution data. That data has big advantage which is detail. This high detail can be also disadvantage because of the high detail itself. Particular land covers cannot be easily manually identified from images with a very high spatial resolution. On the other hand, it is just more complex data set for machine learning. Individually classified pixels represent a problem occurring during supervised classification of very highly detailed images. Majority filter can help with that; but our study shown that using majority filter 16 times provided some acceptable results.

Supervised methods for RGB images classification can quite precisely classify into water, vegetation, dry vegetation and cutted vegetation. These classes are spectrally distinguishable and can be easily classified.

Deep Learning is a modern approach of imagery classification. It was the best classifier in our case study, it provided the highest precision of classification. Even its worst result was better than results of other used methods. There is one key limitation – Deep Learning is a time demanding method, which requires good training of the model.

https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-617-2022 | © Author(s) 2022. CC BY 4.0 License.

ACKNOWLEDGEMENTS

This paper was supported by SGS University of Pardubice project No SGS_2022_008.

REFERENCES

ArcGIS PRO, 2016: Train Support Vector Machine Classifier, https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/train-support-vector-machine-classifier.htm (accessed 11 January 2022).

Bendig, J. *et al.,* 2015: Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation* 39, 79-87. doi.org/10.1016/j.jag.2015.02.012.

Bijeesh, T.V., Narasimhamurthy, K.N., 2020: Surface water detection and delineation using remote sensing images: a review of methods and algorithms. *Sustainable Water Resources Management* 6(4), article No 68. doi.org/10.1007/s40899-020-00425-4.

Bukata, R.P., Harris, G.P. and Bruton, J.E., 1974: The detection of suspended solids and chlorophyll-a utilizing digital multispectral ERTS-1 data. *Second Canadian Symposium on Remote Sensing*, 551-564.

Buyukdemircioglu, M., Kocaman, S., 2021: Geovisualization of aerial photogrammetric flights for data quality assessment. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 43(B4-2021), 333-338. doi.org/10.5194/isprs-archives-XLIII-B4-2021-333-2021.

CAA, 2022: Unmanned Aircraft. Civil Aviation Authority Czech Republic. https://www.caa.cz/en/flight-operations/unmanned-aircraft/.

DJI, 2022: Mavic 2 Enterprise Series. https://www.dji.com/cz/ mavic-2-enterprise?site=brandsite&from=landing_page.

Esri, 2022: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community.

Gallop, S.L., Bosserelle, C., Haigh, I.D. et al., 2015: The impact of temperate reefs on 34 years of shoreline and vegetation line stability at Yanchep, southwestern Australia and implications for coastal setback. *Marine Geology* 2015(369), 224-232.

Gitelson, A.A. et al., 2002: Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment* 80(1), 76–87. doi.org/10.1016/S0034-4257(01)00289-9.

Heine I. et al., 2015: Reconstruction of Lake Level Changes of Groundwater Fed Lakes in Northeastern Germany Using RapidEye Time Series. *Water* 7(8), 4175-4199.

Iizuka, K., et. al., 2018: Advantages of unmanned aerial vehicle (UAV) photogrammetry for landscape analysis compared with satellite data: A case study of postmining sites in Indonesia. *COGENT Geosci.* 4 (1). doi.org/10.1080/23312041.2018.1498180.

Kataoka, T. et al., 2003: Crop growth estimation system using machine vision, in IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 1079–1083. doi.org/10.1109/AIM.2003.1225492.

Komarkova, J., Jech, J., 2020: Processing UAV Based RGB Data to Identify Land Cover with Focus on Small Water Body Comparison of Methods. *In 15th Iberian Conference on Information Systems and Technologies (CISTI)*, 1-6. doi.org/10.23919/CISTI49556.2020.9141170.

Lazar A., Shellito B.A., 2009: Classification in GIS using support vector machines. In: Handbook of research on geoinformatics. IGI Global, 106-112.

Louhaichi, M., Borman, M.M. and Johnson, D.E., 2001: Spatially located platform and aerial photography for documentation of grazing impacts on wheat. *Geocarto International*, 16(1), 65-70. doi.org/10.1080/10106040108542184.

Ma, L., Liu, Y., Zhang, X.L., Ye, Y.X., Yin, G.F., Johnson, B.A., 2019: Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 166–77.

Meyer, G. E., Hindman, T. and Laksmi, K., 1999: Machine vision detection parameters for plant species identification, in *Proceedings of SPIE - The International Society for Optical Engineering*, 3543, 327-335.

Meyer, G.E. *et al.*, 2004: Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Computers and Electronics in Agriculture*, 42(3), 161-180. doi.org/10.1016/j.compag.2003.08.002.

Mountrakis G., Im, J., Ogole, C., 2011: Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(3), 247-259.

Pásler, M., Komárková, J. and Sedlák, P., 2015: Comparison of possibilities of UAV and Landsat in observation of small inland water bodies. In *International Conference on Information Society, i-Society 2015*, 45-49. doi.org/10.1109/i-Society.2015.7366855.

Potapov, P. et al. 2020: Landsat Analysis Ready Data for Global Land Cover and Land Cover Change Mapping. Remote Sensing 12(3), Article No 426. doi.org/10.3390/rs12030426,

Saharan, M. A. *et al.,* 2018: Classification and assessment of the land use - Land cover changes in Jodhpur city using remote sensing technologies. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 767-771.

Salamí, E., Barrado, C., Pastor, E., 2014: UAV flight experiments applied to the remote sensing of vegetated areas. *Remote Sensing* 6(11), 11051–11081.
X1 Stigler S.M., et al., 2007: The epic story of maximum likelihood. *Statistical Science* 22(4), 598-620.

Stigler S.M., et al., 2007: The epic story of maximum likelihood. *Statistical Science* 22(4). 598-620.

Vapnik, V., Chervonenkis, A., 1974: Theory of Pattern Recognition. Nauka, Moscow (in Russian).

Vitousek, P.M, Mooney, H.A., Lubchenco, J., Melillo, J.M., 1997: Human Domination of Earth's Ecosystems. Science 277(5325), 494–99.

Wang, L., Yan, J., Mu, L., & Huang, L., 2020: Knowledge discovery from remote sensing images: A review. *WIREs Data Mining and Knowledge Discovery*. 10:e1371. doi.org/10.1002/widm.1371.

Woebbecke, D.M. et al., 1995: Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the American Society of Agricultural Engineers* 38(1), 259-269.

Zhao, Z., Liu, H., 2007: Spectral feature selection for supervised and unsupervised learning. In: *Proceedings of the 24th international conference on Machine learning*, 1151-1157. doi: doi.org/10.1145/1273496.1273641.