

The evaluation of water pollution with the help of remote sensing tools

P. Šádek^{1*}, J. Struhár¹

¹ VSB-TUO, Faculty of mining and geology, 70800 Ostrava, Czech Republic - (petr.sadek; juraj.struhar)@vsb.cz

Commission VI, WG VI/4

KEY WORDS: UAV, Landsat 8, Remote Sensing, vegetation indices, regression analysis, water pollution

ABSTRACT:

With the growing population, there is a growing demand for quality drinking water. Especially in developing parts of the world, this is a serious problem. The aim of this work is to test remote sensing methods for water quality monitoring. The presented part of the project is focused on introducing the process of water pollution assessment using vegetation indices, which are derived only using RGB images. Water quality monitoring is based on satellite imagery Landsat 8 and UAV images Phantom 3. As reference data was used in-site measurements in profiles points. In-site measurements were repeated every month in the vegetation period from April to September. Based on regression analysis, the equation for the calculation of the amount of chlorophyll and the statistical evaluation of the quality of these equations is derived for each vegetation index. The best results were achieved using the ratio aquatic vegetation index (RAVI) and ExG (Excess green) indices of 97% and 96.8% respectively.

1. INTRODUCTION

1.1 General Instructions

It isn't an overstatement to say, that water is synonymous with life on Earth. Live organisms – whether plants or animals without exception need water for their life. All organisms have water as a constituent in their body, and for many of them water represents their basic living environment. For humans, it is as important as for other living organisms. The basic requirement for their survival is then the continuous provision of clean and quality water. In Europe and North America, supplying a sufficient amount of water doesn't generally present such a problem, however considering the current development of global climate change, they could be facing such problems in the near future. With the development of civilisation and ever-growing industrialisation, the gradual pollution of surface and partially also underground water is becoming a more and more significant problem.

Current civilisation daily produces an enormous amount of waste. Predominantly in developing countries, a vast amount of this waste ends up in surface waters. A large part of river networks in all corners of the world is therefore exposed to intense pollution and water in these places is far from drinkable, but its possibility for use for other purposes is also problematic. With increasing waste and run-off of chemical fertilizers from agricultural soil in water streams and reservoirs, the amount of various dissolved chemical substances in these waters increases as well.

Let's focus mainly on two of those chemicals – phosphorus and nitrogen. These two substances significantly speed up the process of eutrophication of water, a process where water reservoirs become slowly overgrown with algae or cyanobacteria, is becoming a serious problem.

Identification, analysis, and mainly the search for solutions of problems associated with water quality is posing in these times a significant challenge for the world of research. Water pollution isn't only biological, caused by humans. Other processes play a part in pollution, and therefore for an

interdisciplinary point of view, it's necessary to combine information from various fields of science for evaluation and problem-solving.

Currently, remote sensing tools are being increasingly utilized for the monitoring of the quality of water. There is a vast literature on using Landsat satellite systems for water quality monitoring (Acharya et al. 2018; Andersson 2012; Grimm 2018; Han, Jordan 2007). In the last few years, we have seen a massive expansion of UAVs. UAVs are most often used in conjunction with multispectral cameras (Guimarães et al. 2017; Zhang et al. 2019), but there are approaches using only RGB cameras for monitoring water quality (Cândido et al. 2016).

The aim of this research is to discover the relationship between the dissolved chlorophyll in water and the vegetation indexes, which are calculated both from the satellite images from the Landsat 8 system, and from the UAV. The results of this study are very convincing, and clearly show the suitability of remote sensing tools for the monitoring of water quality.

The goal of this study is to use the UAV images, satellite images and results from laboratory analyses as a basis for calculating the amount of chlorophyll and vegetation indexes, and in this way test the remote sensing tools method for monitoring water quality.

2. STUDY AREA AND MATERIALS

2.1 Study area

The main requirement in selecting the water basin was, that it shouldn't be a source of drinking water, because above these water reservoirs it wouldn't be, without an exception issued by the CAA (Civil Aviation Authority) possible to fly with an UAV. After consulting with the workers of the Odra Basin, there were two water basins chosen, which fulfil these requirements, specifically they are the water basins Těrlicko and Žermanice.

* Corresponding author

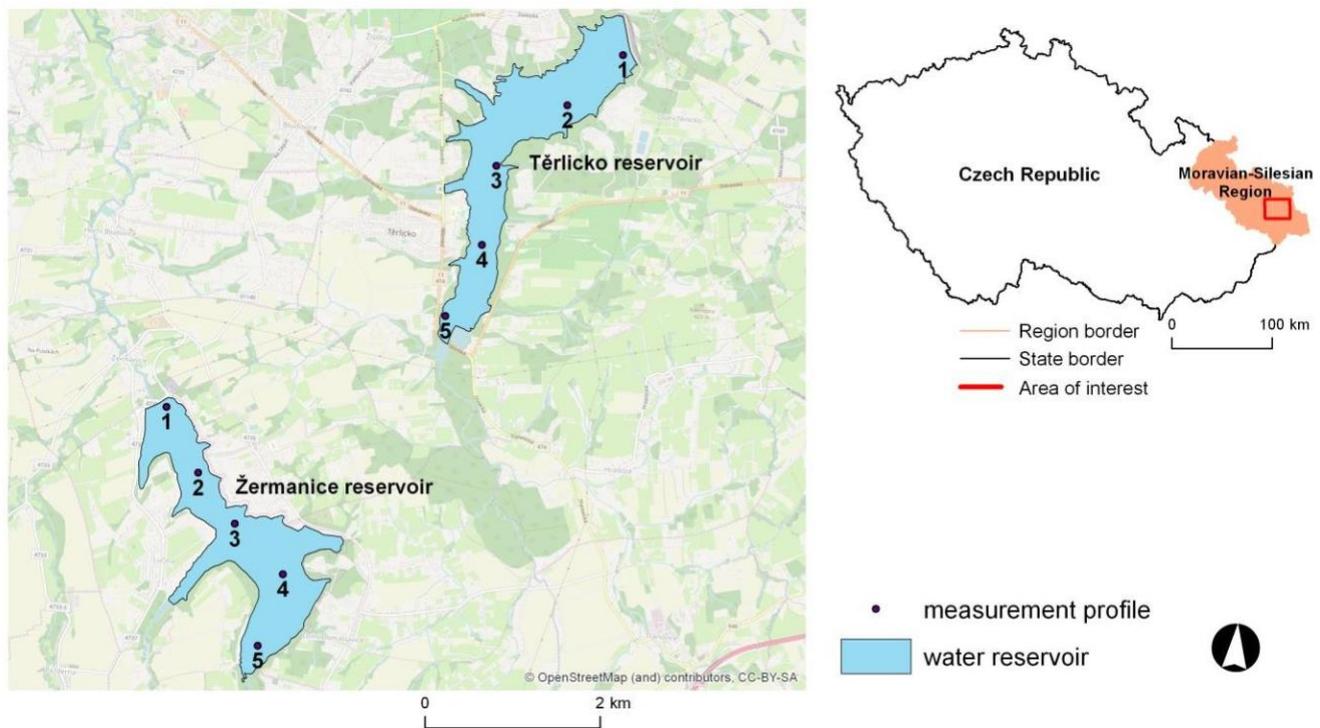


Figure 1 Measurement profile at select water reservoir

The water basin Těrlicko is a valley reservoir on the river Stonávka in Těrlicko in close vicinity of the city Havířov in the Moravian-Silesia region. This basin was built in the years 1955 to 1964 in order to supply the nearby coal mines and Třinec iron works with operation water. It was also used for boosting the low water flow on Stonávka, and mainly as a flood barrier of the lower-lying area. Currently, the basin is abundantly used for recreational purposes. For the purposes of monitoring, there was a network of five measure profiles built on the water basin (Figure 1 **Measurement profile at select water reservoir**), which are numbered in ascending order from no. 1 near the dam up to no. 5 at the end of the swell. The water reservoir Těrlicko falls under the category of eutrophic waters, there is a higher risk of presence of phytoplankton population, especially in summer months. In the past three years, there was a decrease in the amount of cyanobacteria, the amount of water algae has a fluctuating tendency depending on the amount of water in the reservoir and on meteorological conditions. The most significant tributary of the Těrlicko reservoir is the river Stonávka, which is simultaneously the main source of organic and inorganic pollution. In the measure profile, which is located on the river before the water reservoir Těrlicko itself, the quality of water is rated as III. Class out of five. This class represents polluted water.

The measured values of phosphorus concentration on the tributary to the reservoir are relatively stable and exist in the range of limit values of 0.05 mg/l (Skalička 2016), which represents an acceptable annual average of chlorophyll in water.

The second location, where image-capturing took place, is the water reservoir Žermanice. It is a valley reservoir on the river Lučina near the town Žermanice, located 9km away from the above-mentioned Těrlicko reservoir. The reservoir was built between the years 1951 to 1957. Due to insufficient water presence of Lučina, it was necessary to build an artificial tributary from the river Morávka into the Žermanice basin. The purpose and function of this water reservoir is identical to

the above-mentioned water reservoir Těrlicko, it's use is foremost the supply of operation water to the Ostravian industry, further then it serves as a flood barrier. On the water reservoir Žermanice, there is a total of five measure profiles, which are numbered similarly to the Těrlicko reservoir. The quality of the water in the reservoir has been in recent years very good, and generally it can be said, that the water in the Žermanice reservoir is better in terms of quality than the water reservoir Těrlicko. The concentration of dissolved phosphorus in the river Lučina has been for a long time very low, the limit value goes above 0,05 mg/l very rarely (Skalička 2016).

2.2 Data

There are two types of data used for research, image data from the UAV and from the Landsat 8 system, and then the data from individual quality analysis of water, which was provided by the Odra Basin. The data was collected regularly every 30 days between April and September, that is in the vegetation period. Individual readings were collected with a multiparameter device, which measures a vertical profile every 1m of depth. From this device we obtained not only physicochemical parameters, but even the amounts of chlorophyll and phycocyanin. Next, the individual samples were transported for processing to the Odra Basin laboratory. Measuring was always conducted on measure profiles of the monitoring network, therefore in points.

The satellite images used in this study were from the Landsat 8 for the period from 2014 to 2017. There are requirements in image-capturing, and that is that there should be the lowest possible cloudiness, and the lowest difference between the date of image-capturing and collection. It wasn't always possible to fulfil both requirements simultaneously. The greatest deviation between image-capturing and collection was 14 days.

Landsat 8 is able to manifest 9 spectral and 2 thermal zones. The images from Landsat 8 were made in individual zones. For the purpose of this study, there were designated

zones no. 2, 3, 4, which correspond to blue, green and red. From the UAV we only gather one resulting image containing three zones. For further processing of such images, it is necessary to divide the respective zones. A great disadvantage in using a UAV for monitoring the quality of water is the fact, that we are reliant on the evaluation of individual, single images. It isn't possible to combine individual images into a mosaic. There is a method of creating such mosaics, where photogrammetric software (Agisoft Photoscan, Pix4D...) in processing such data tries to find common denominator points in individual images and then combines these images into a greater whole. On a water surface it isn't possible to find any common points (it is a uniform surface area), that's why always one image was taken on the first measure profile.

2.3 Image-capturing and processing

Image-capturing was done with the UAV DJI Phantom 3 Advances on two water reservoirs Těrlicko and Žermanice. Image capturing was done with a maximum one-week delay after sample collection. This delay allowed us to plan image capture with respect to the best possible meteorological conditions. The data was taken altogether in four months of the year 2017, and due to this reason, that in April and in July it wasn't possible to take images due to unfavourable wind conditions. Image-capturing took place always above the first measure profile, that is closest to the dam, and was done from a 90 m height with a perpendicular view down. Semi-automatic exposure was used, and the following settings: ISO 100 and shutter F number 2,8. Exposure was always automatically set according to current light conditions.

The processing of individual images took place in the environment Esri ArcMap 10.3.1. Due to the fact, that processing a larger number of images requires the same repetitive actions, the program ModelBuilder was utilized to process individual images, where every step in the method process there was one model designed, which performs these steps for each image.

3. METHOD

3.1 Processing images from the UAV

The first step in processing is correcting the individual images, mainly to remove sun reflections and clouds of the surface of the water (Figure 2 **First measurement profile at Žermanice reservoir**). This part of processing is a fundamental step in the whole process and its significance can be compared to atmospheric corrections of satellite images. The individual reflections significantly affect the appearance of individual images, and subsequently affect the vegetation index values. At



Figure 2 First measurement profile at Žermanice reservoir

this point of processing, a significant problem arose.

There wasn't any suitable method found for removing individual reflections. All methods for removing reflections from the surface of the water come from the assumption, that in using images in the close infrared spectrum and in the sufficient optical depth of water, it is possible to overlook bed reflection.

Therefore, all the information that the images carry, are represented in the intensity of the reflections on the surface of the water. Based on these assumptions, there are methods created, thanks to which it is possible to remove these reflections (Overstreet, Legleiter 2017). Due to the fact, that there weren't any images in the near infrared spectrum, it is necessary to mention, that it wasn't possible to remove these reflections. One option is to cut out the images according to the outline of individual reflections, or to mask these reflections. The individual reflections from the surface of the water were left, due to the fact, that the monitoring of the water quality was done by point measuring, and so even the vegetation indexes, were taken as point measurements.

After grouping all the images from the UAV into individual zones, it was possible to calculate the individual vegetation indexes. Due to the RGB camera used, there were those of the vegetation indexes selected, which were created based on RGB, specifically RAVI (Ratio Aquatic Vegetation Index), GRVI (Green-red vegetation index) and ExG (Excess green).

$$\text{RAVI} = G/R$$

$$\text{GRVI} = G - R / G + R$$

$$\text{ExG} = 2G - R - B \text{ (Motohka et al., 2010)}$$

3.2 Processing the satellite images

After downloading all the required images, the method was exactly according to the stated scheme. The first step was the creation of an outline polygon layer containing the two dams of interest. All the satellite images were outline-cut according to this layer. The following step was correcting the individual images. Due to the fact, that the product was downloaded, geometric correction was applied already by the provider, it was easy to leave out correction and to focus only radiometric corrections. In this study, the DN (Digital number) correction of values to TOA (Top of atmosphere) was done, where it is necessary to note, that the difference in values between TOA and the reflectivity of Earth's surface is very small, which makes it easy to document (Congedo 2016). Sun-angle correction was done simultaneously with conversion to reflection values. This correction was necessary for comparing time changes in the same zone. DN values were recalculated based on imaginary radiation from the zenith. The last step was calculating the vegetation indexes. The same vegetation indexes were used for satellite images, as were used for UAV images.

4. RESULTS AND DISCUSSION

Regression analysis was used for finding out the relationship between the vegetation index values and the amount of chlorophyll. Before utilizing the regression analysis, it was necessary to perform a few steps first. The first one was the preparation of data for evaluation. One source of data was the images and the vegetation index calculated from them, the second was the laboratory analyses from the Odra basin. The analyses were made along the whole height of the water column, and at the same time the vegetation index values were calculated only for a very thin layer near the surface of the water. It was necessary to pool all chlorophyll values for the entire water column into one value, for this step there were

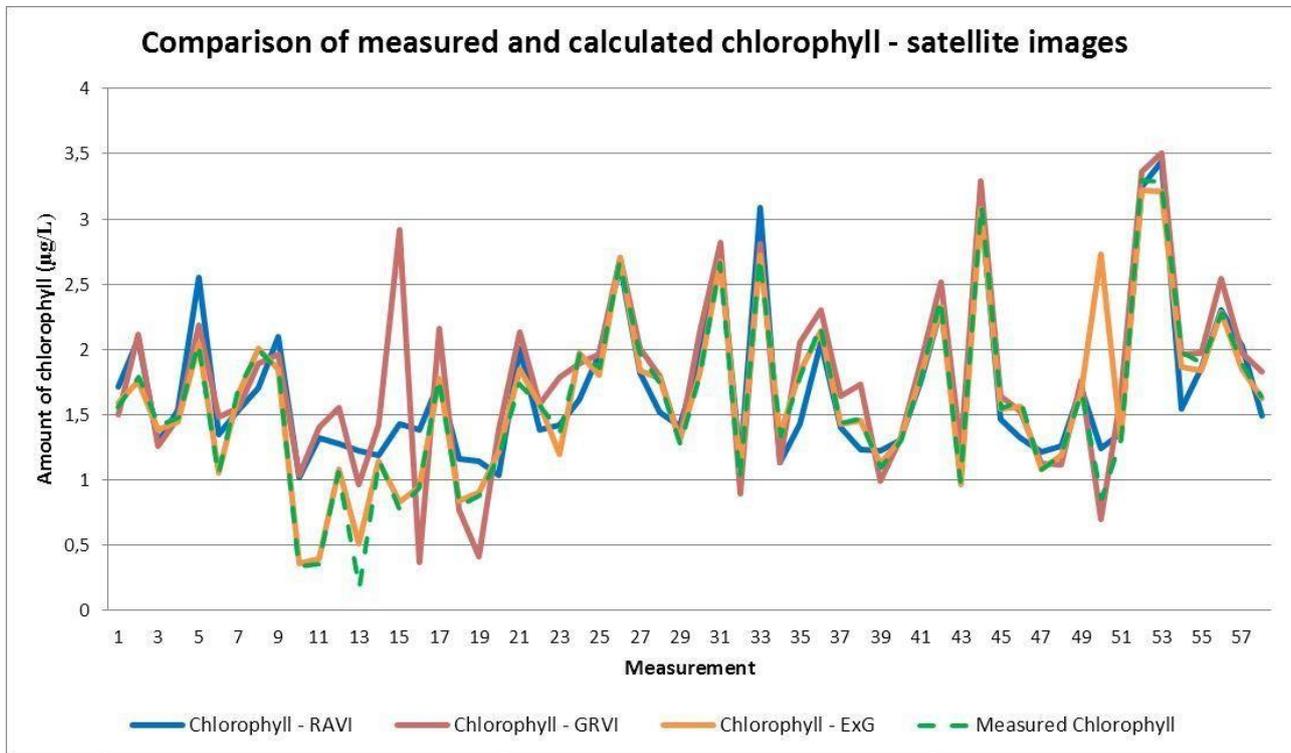


Figure 3 Comparison of measured and calculated chlorophyll – satellite images

weighted averages calculated, where the chlorophyll values were mostly an inverse value of depth.

The then data file was divided into two subsets, the first subset was the training set. The data set contained both data from the satellite images, and data from the UAV, a total of 238 records. The reason for this step was the fact, that the images from the UAV were made only for one vegetation period, and individually didn't form a representative data set. The training set contains 75 % of all data, the selection was made based on random selection. The number of records in the training set was 179. The purpose of the training set was to find the relationship between the amount of chlorophyll and vegetation indexes.

With this prepared data, it was possible to conduct various statistical operations. Before using regression analysis, it was necessary to check the normality of the data sets, which is one of the preconditions of regression analysis. Data normality was tested for the variable's chlorophyll and for the three calculated vegetation indexes. Based on statistical tests, it was found, that none of the variables fulfilled the requirements for normal grouping. Due to not fulfilling the data normality requirement, it was necessary to convert the variables into normal grouping with the help of root conversion. After normalizing the data, it was possible to begin regression analysis. The purpose of regression analysis was to find the relationship between the amount of dissolved chlorophyll in water and the vegetation index values.

Table 1 Derived regression relationships and R²

Vegetation index	Derived relationship	R ²
RAVI	$Y = 0,107 + 1,013X$	0,971
GRVI	$Y = -0,156 + 3,665X$	0,791
ExG	$Y = 0,053 + 3,125X$	0,968

From the results of the regression analysis it is possible to conclude, that the coefficient of determination values is very high (Table 1 Derived regression relationships and R²), which indicates, that the derived relationships are suitable for calculating the amount of chlorophyll. The quality of derived relationships was checked by applying a testing data set, which was created before the statistic evaluation. In the test set, there were chlorophyll values calculated from the derived relationships, which were compared with the real, recorded values of chlorophyll (Figure 3 **Comparison of measured and calculated chlorophyll**). The test data set contained all the vegetation value indexes, that were calculated from the images taken by the UAV.

For evaluating the quality of derived relationships of the satellite images, there were basic statistical indicators calculated for each vegetation index (Table 2 Evaluating the quality of derived relationships – Landsat 8), it is interesting to note, that the minimum in RAVI is 1,02, which is significantly more, than the minimum recorded chlorophyll value.

From this fact we can conclude, that the derived relationship for RAVI can't correctly create a model with values lower than 1, and at the same time it is slightly higher, than the maximum measured value. The maximum of all vegetation indexes is very close to the real maximum, and only two relationships measured slightly higher, and those are RAVI and GRVI. The minimums of calculated values using GRVI and ExG are very similar to the minimum measured values of chlorophyll.

Table 2 Evaluating the quality of derived relationships – Landsat 8

	Chlorophyll 1 (µg/L)	RAVI	GRVI	ExG
Standard deviation	0,66	0,58	0,67	0,64
Minimum	0,17	1,02	0,37	0,36
Maximum	3,29	3,45	3,51	3,22
Average deviation		0,21	0,22	0,08
Spread	0,43	0,33	0,45	0,41
Median	1,59	1,48	1,74	1,61

The average deviations between the calculated and real values are relatively low, no value is greater than 0,25. The deviation of calculated values is very similar to the spread of measured values and median values are very similar to each other. The best resulting value of this comparison is the ExG vegetation index, where the average deviation was 0,08, and where even in terms of R^2 , this relationship is among the best. The worst calculated relationship was then shown to be the GRVI index, which was the worst out of all three indexes, even in terms of the R^2 comparison.

The same statistical indicators were calculated for the images from the UAV. A value of note is the standard deviation of the ExG index (Table 3 Evaluating the quality of derived relationships – UAV), which is almost half of the standard deviation of the real, recorded values. The minima of all values are very similar to each other, the maxima in the first two vegetation indexes are like the real maximum, which is however, slightly higher. The maximum of the ExG index is slightly higher than the real maximum, and therefore it's possible to say, that thanks to the relationship it isn't possible to create a model with higher values of chlorophyll, than values closer to 2. The average deviations are relatively low, again an exception is the ExG index, which if we consider the median value, cannot present the variability of the amount of

chlorophyll very well. The best suitable vegetation index for evaluating the quality of water using a UAV is RAVI and GRVI (Figure 4 **Comparison of measured and calculated chlorophyll – UAV**). These vegetation indexes provide almost identical results. The vegetation index ExG didn't achieve such results, as in connection with satellite images, and therefore isn't suitable for using in connection with a UAV.

Table 3 Evaluating the quality of derived relationships – UAV

	Chlorophyll (µg/L)	RAVI	GRVI	ExG
Standard deviation	0,8	0,77	0,88	0,42
Minimum	0,72	0,75	0,65	0,76
Maximum	3,03	3,09	3,19	2,28
Average deviation		0,14	0,07	0,48
Spread	0,64	0,59	0,77	0,18
Median	2,68	2,51	2,85	1,69

At the end of the statistical evaluation, it was verified with the help of the Mann – Whitney test, that the calculated chlorophyll values are comparable with the recorded values.

Cândido et al. (2016) suggested that the use of UAV and RGB vegetation indices have good results for monitoring small water bodies. This case study demonstrates that using RGB vegetation indices is efficient and it's possible to use it for estimate the chlorophyll. In our study we show that using RGB vegetation indices are suitable even for monitoring larger water bodies, such as water reservoir or lakes.

Torres-Sánchez et al. (2014) who mapped wheat field via UAV using conventional aerial photography for computing vegetation indices. Among the vegetation indices which are computed from UAVs camera, the GRVI and ExG shown the best results which partly corresponding with our case study.

5. CONCLUSIONS

The aim of this work was to create a processing method for

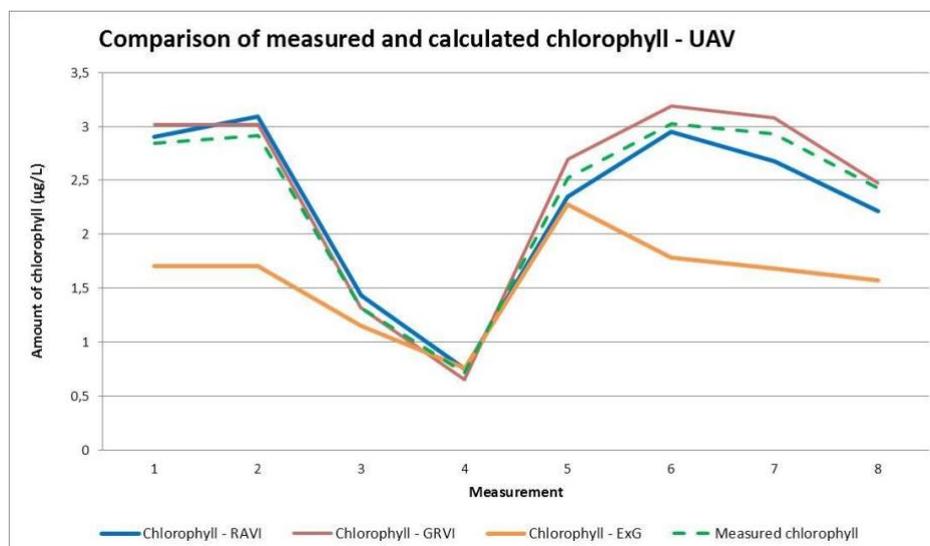


Figure 4 Comparison of measured and calculated chlorophyll – UAV

determining the amount of chlorophyll based on UAV and satellite images and laboratory analyses and deriving the relationship between the amount of chlorophyll based on vegetation indexes. Regression analysis was used for deriving the relationship between the measured chlorophyll values and the individual vegetation indexes. The evaluation of UAV images was limited by their number, due to the fact, that image-capturing took place only within one vegetation period. It was therefore necessary to evaluate this image together with satellite images. Based on regression analysis, it was found, that RAVI and ExG vegetation indexes reach the highest R^2 value (97 and 96,8 %). In comparing the recorded and calculated chlorophyll values it was discovered, that the best suitable index for monitoring the quality of water with satellite images is ExG, which modelled the recorded values very accurately.

As the comparison of UAV measured and calculated chlorophyll values showed, the vegetation index ExG wasn't able to model the course of the values correctly. A very suitable vegetation index of these images showed to be RAVI and GRVI, which based on regression analysis seemed to be the worst with $R^2 = 0,791$. In terms of processing of individual images, some problems arose, which were not solved to satisfaction, and for the practical utilization of this method, they require further attention. Overall, it was shown, that using the data from remote sensing tools is suitable for evaluating water quality.

ACKNOWLEDGEMENTS

Petr Šádek and Juraj Struhár were kindly supported by VSB – Technical University of Ostrava project SV511998/SP2019/64.

The authors also thank the Odra Basin for providing data of measured water quality, without which it wouldn't have been possible to make this study.

REFERENCES

ACHARYA, T., A. SUBEDI a D. LEE. Evaluation of Water Indices for Surface Water Extraction in a Landsat 8 Scene of Nepal. *Sensors* [online]. 2018, 18(8). DOI: 10.3390/s18082580.

ANDERSSON, M. (2012). Estimating Phosphorus in rivers of Central Sweden using Landsat TM data. Stockholm. Disertace. Stockholm University, Faculty of Science, Department of Physical Geography and Quaternary Geology.

CÂNDIDO, A. K. A. A., A. C. P. FILHO, M. R. HAUPENTHAL, N. M. DA SILVA, J. DE SOUSA CORREA a M. L. RIBEIRO (2016). Water Quality and Chlorophyll Measurement Through Vegetation Indices Generated from Orbital and Suborbital Images. *Water, Air, & Soil Pollution* [online]. 227(7), DOI: 10.1007/s11270-016-2919-7.

CONGEDO L. (2016). Semi-Automatic Classification Plugin Documentation. DOI: 10.13140/RG.2.2.29474.02242/1

GRIMM, A. Mapping Submerged Aquatic Vegetation in the Great Lakes Using Satellite image-capturing. State of Michigan [online]. Michigan, 2018, 2015 [cit. 2018-03-04]. Dostupné z: http://www.michigan.gov/documents/deq/GLBA1029-GrimmAmanda-MappingAquaticVegetation_506955_7.pdf

GUIMARÃES, T., M. VERONEZ, E. KOSTE, et al. An Alternative Method of Spatial Autocorrelation for Chlorophyll

Detection in Water Bodies Using Remote Sensing. Sustainability [online]. 2017, 9(3). DOI: 10.3390/su9030416.

HAN, L. a K. J. JORDAN. Estimating and mapping chlorophyll- a concentration in Pensacola Bay, Florida using Landsat ETM+ data. *International Journal of Remote Sensing* [online]. 2007, 26(23), 5245-5254. DOI: 10.1080/01431160500219182.

MOTOHKA, T., K. N. NASAHARA, H. OGUMA a S. TSUCHIDA (2010). Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. *Remote Sensing* [online]. 2(12), 2369-2387. DOI: 10.3390/rs2102369.

OVERSTREET, B. T. a C. J. LEGLEITER (2017). Removing sun glint from optical remote sensing images of shallow rivers. *Earth Surface Processes and Landforms* [online]. 42(2), 318-333. DOI: 10.1002/esp.4063.

SKALIČKA, M., M. ŠTRAJT a Č. ONDRUŠÁK (2016). JAKOST VODY V ÚDOLNÍCH NÁDRŽÍCH ZA ROK 2016. Ostrava.

TORRES-SÁNCHEZ, J., J.M. PEÑA, A.I. DE CASTRO a F. LÓPEZ-GRANADOS. Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. *Computers and Electronics in Agriculture* [online]. 2014, 103, 104-113. DOI: 10.1016/j.compag.2014.02.009.

ZHANG, L., H. ZHANG, Y. NIU a W. HAN. Mapping Maize Water Stress Based on UAV Multispectral Remote Sensing. *Remote Sensing* [online]. 2019, 11(6). DOI: 10.3390/rs11060605