

## COMPARISON OF SELECTED VEGETATION INDICES AND DETERMINATION OF SUITABILITY FOR YIELD DESCRIPTION ON AGRICULTURAL FIELD

Katerina Krizova<sup>1</sup>, Jitka Kumhalova<sup>2</sup>  
Czech University of Life Sciences Prague  
krizovak@tf.czu.cz

**Abstract.** Vegetation indices, as a non-destructive and low-cost method, represent an effective approach to vegetation cover evaluation. The simple mathematical formulas used describe wide spectra of conditions (water and nutrition saturation, influence of topography). Remote sensing data are used to derive vegetation indices. Calculations are performed using specialized software when operating with particular bands of electromagnetic spectra. The results contribute to adjustment of current agricultural management that becomes more economically efficient while having lower negative impact on the environment. The paper deals with selected vegetation indices and examines their suitability for yield prediction. Study area was 11.5 ha agricultural plot in Praha-Ruzyně. The experiment was conducted on wheat (2005, 2011) and oat (2006, 2010) having spatially related yield data. The selected indices were derived from LANDSAT 5 imagery with 30 m spatial resolution using SW ENVI. Specific values were obtained using SW ArcGIS. Correlation analysis was conducted to examine the relation between particular VI and the yield data or the Topography Wetness Index. The results indicated relation between vegetation indices and yield in all cases. The Moisture Stress Index reached -0.835 in 2011 as the maximal value of the correlation coefficient, while the minimum was performed by value 0.495 of the Chlorophyll Vegetation Index in 2005. The analysis also indicated relation of the yield to the topographic conditions of the agricultural plot. The Simple Ratio Vegetation Index in 2010 had the strongest correlation with the Topography Wetness Index, the correlation coefficient reached 0,6. Conversely, the minimal value was observed by the Chlorophyll Vegetation Index in 2005, namely 0.19. The Normalized Difference Vegetation Index, as the last of the selected indices, showed average results. Nevertheless, the data were evaluated in terms of the weather conditions as well. The influence of temperature and precipitation in particular growth stages was discussed. At the end, the conclusion was drawn that selected vegetation indices are suitable to describe the yield despite the fact they were derived from 30 m spatial resolution imagery.

**Keywords:** yield, remote sensing, LANDSAT 5, spatial resolution.

### Introduction

The fact about increasing world population is generally known. While there were 7.3 billion citizens in 2015, the latest estimates predict to be 8.5 billion citizens in 2030 and 11.2 billion in 2100 [1]. Related to this the question sustainability becomes even more actual. The demand for quality resources is increasing, while the space for production remains limited [2]. Sustainability of important resources must be therefore solved by changing the attitude and also by using new technologies to make the primary production more effective.

Precision agriculture is an approach that is developing since about the last three decades. Modern technologies and latest knowledge are used to make agricultural management more economically efficient while having lower negative impact on the environment [2]. Satellite imagery contains a set of information describing different characteristics of the displayed area. This information can be analyzed by modern technologies and software. The acquired results contribute to effective agricultural planning in specific area and also to the development of precision agriculture methods in general. Calculating vegetation indices (VI) is such method.

According to the literature there are more than 150 different indices and each of them highlights specific vegetation cover attribute, such as water or biomass content. The acquired values of VI can be compared with topographical and meteorological factors in specific study area. Such comparisons help detect particular regions of agricultural plots with a specific requirement for management [2; 3]. Therefore, the main objective of this study is to describe the capability of selected indices to predict the yield as well as find the relation with topography and weather conditions. The question of utilization of LANDSAT images with 30 m spatial resolution should be also discussed.

### Materials and methods

A 11.5 ha field of the Crop Research Institute in Prague was chosen to be an experimental plot. It is located in Praha-Ruzyně, on 50°05'N and 14°17'30''E (see Figure 1) with the elevation levels from

338.5 to 357.5 meters above the sea level and slope approximately 6 %. The soil was classified as haplic luvisols partly lying on calcareous sandstones and partly on clays.

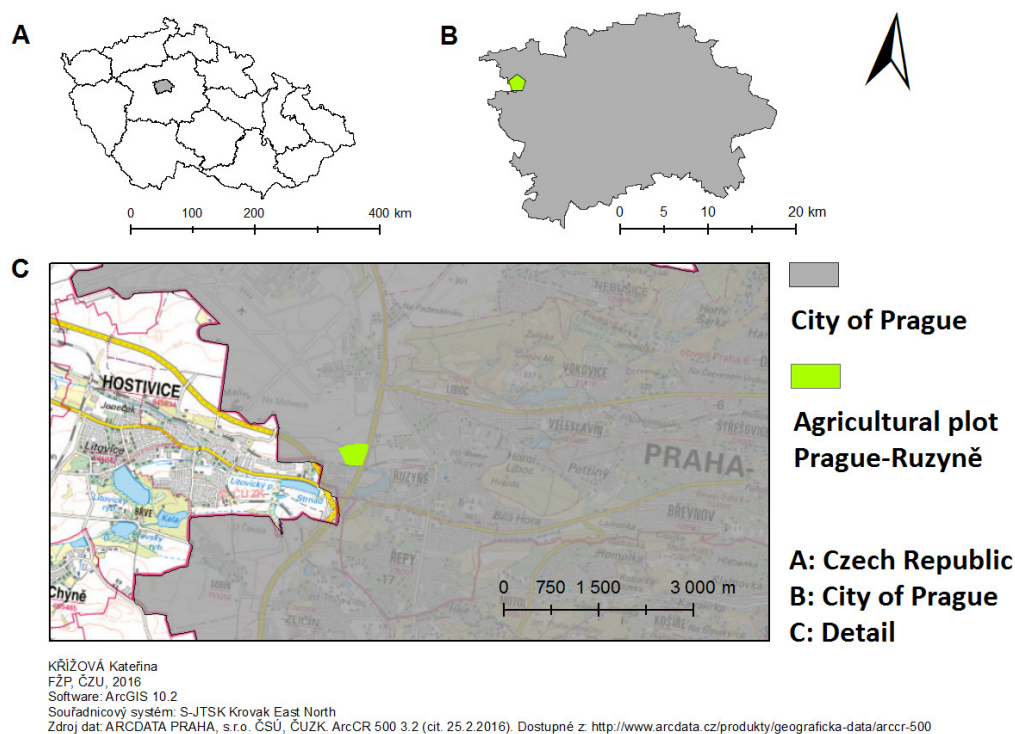


Fig. 1. Experimental plot location.

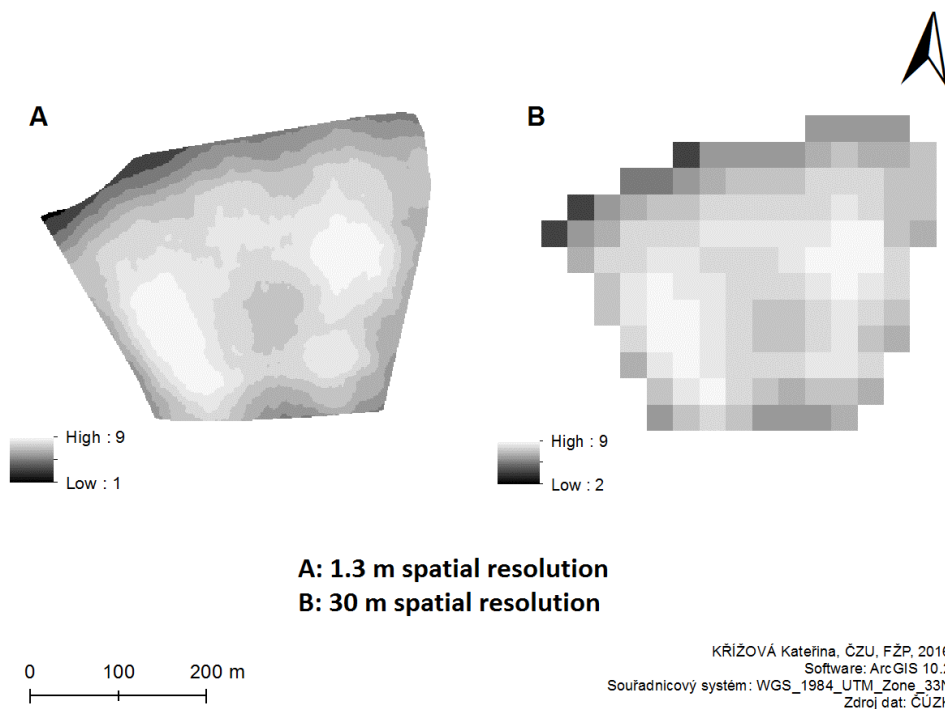
The weather conditions of the studied periods are summarized in detail for specific growth stages and presented as Table 1. The growth stages are represented according to the Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (BBCH) scale [4].

Table 1

#### Weather conditions

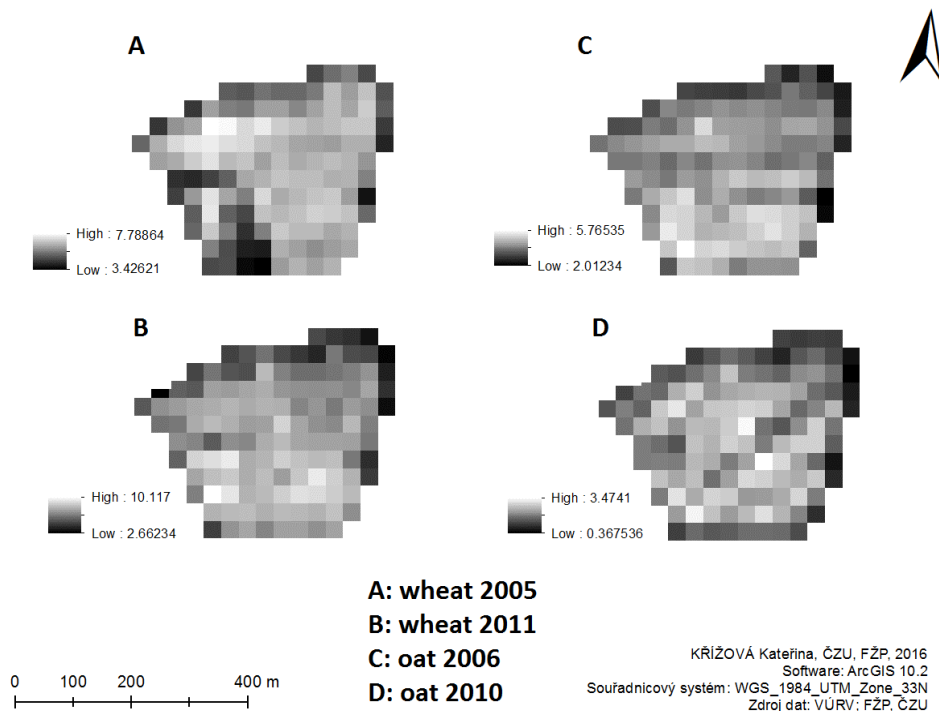
Precipitation, mm				
Scale	wheat		oat	
	2005	2011	2006	2010
BBCH 20-29	83.4	104.4	111.4	93.4
BBCH 30-59	90.4	39.5	48.6	84.7
BBCH >60	207.8	257.4	94.6	142.3
Sum	381.6	401.3	254.6	320.4
Mean	127.2	133.8	84.9	106.8
Temperature, °C				
Scale	wheat		oat	
	2005	2011	2006	2010
BBCH 20-29	4	3.4	14.1	12.3
BBCH 30-59	13.9	14.8	16.6	16.5
BBCH >60	18.4	17.9	22.2	21.1
Mean	12.1	12	17.6	16.6

The topographic data were kindly provided by the Czech Office for Surveying, Mapping and Cadastre. The following data processing has been already done and described in [5]. The result was a raster layer representing the Topography Wetness Index (TWI). This index was utilized to describe the influence of topography on yield variability. The TWI raster is presented by Figure 2.



**Fig. 2. Topography Wetness Index (original and resampled)**

The yield data were kindly provided by the Crop Research Institute in Prague. The data processing has been also described more in detail in [5]. The processed dataset was used to create yield maps (see Figure 3) for the following statistical analysis.



**Fig. 3. Yield Maps**

The satellite data were kindly provided by the United States Geological Survey (USGS). There were four LANDSAT 5 images in total, taken on 3.6.2005, 4.6.2011 for wheat and on 13.6.2006 and 17.6.2010 for oat. All the images were taken in 30 m spatial resolution. The satellite imagery represented the basis of the VI calculation. Four VI was chosen to be calculated and then compared with the yield data from four selected seasons. Normalized Difference Vegetation Index (NDVI) as a

commonly used indicator of green vegetation, Simple Ratio Vegetation Index (SRVI), Moisture Stress Index (MSI) and Chlorophyll Vegetation Index (CVI). VI calculations were done using SW ENVI (ENVI 5.3). This software calculates the desired index and presents the result as the raster layer that can be further processed in SW ArcGIS (ArcGIS 10.2). The process was done in four repetitions identically for each of the four satellite images (selected seasons). As the result, there were raster layers with 130 pixels and 30 m spatial resolution. Summary statistics is given by Table 2 for wheat and Table 3 for oat and by Table 4 for the yield and topographic data.

The whole statistical analysis was done using R software. Correlation coefficients ( $r$ ) were calculated. The first level of comparison was yield data and particular VI, second TWI and particular VI and the finally investigated relation was between TWI and the yield.

Table 2

#### Summary statistics for wheat

Statistics	2005				2011			
	NDVI	SRVI	MSI	CVI	NDVI	SRVI	MSI	CVI
Sum	130	130	130	130	130	130	130	130
Mean	0.84	13.38	0.18	6.53	0.83	6.93	0.13	5.82
Median	0.87	14.35	0.16	6.51	0.83	6.35	0.12	5.56
Min	0.49	2.92	0.11	3.38	0.43	2.31	0.05	2.57
Max	0.92	25.53	0.48	10.9	1.3	14.76	0.33	9.81
Stan. dev.	0.08	4.46	0.06	1.58	0.13	2.68	0.06	1.54
Skewness	-2.31	-0.39	2.43	0.35	-0.66	0.63	1.21	0.39

Table 3

#### Summary statistics for oat

Statistics	2006				2010			
	NDVI	SRVI	MSI	CVI	NDVI	SRVI	MSI	CVI
Count	130	130	130	130	130	130	130	130
Mean	0.85	14.98	0.37	7.12	0.88	20.8	0.27	6.53
Median	0.89	17.06	0.34	7.53	0.92	22.78	0.25	6.4
Min	0.41	2.41	0.31	2.96	0.52	3.15	0.20	2.64
Max	0.92	24.12	0.81	10.55	0.95	37.64	0.58	10.42
Stan. dev.	0.1	4.98	0.09	1.44	0.09	7.32	0.08	1.49
Skewness	-2.71	-0.89	2.87	-0.73	-2.55	-0.85	2.73	-0.15

Table 4

#### Summary statistics for yield and topographic data (yield data given in $t \cdot ha^{-1}$ )

Statistics	wheat		oat		TWI
	2005	2011	2006	2010	
Count	130	130	130	130	130
Mean	6.64	6.92	4.72	2.1	6.64
Median	6.29	7.29	4.18	2.11	7.0
Min	3.43	2.66	2.01	0.37	2.0
Max	7.79	10.12	5.77	3.47	9.0
Stan. dev.	0.96	1.54	0.76	0.68	1.67
Skewness	-0.52	-0.7	-0.49	-0.24	-0.52

## Results and discussion

The objective of this study was to calculate four selected VI and describe their capability for yield prediction on particular agricultural plot. Correlation analysis was done on three levels.

At first, the relation between VI and the yield was investigated. As seen in Table 5, a strong correlation was indicated by all selected indices. On wheat, the strongest correlation was found by SRVI in 2005 and by MSI in 2011. Conversely, CVI indicated the weakest relation in both years under

investigation. Only slightly different situation was found on oat. The strongest relation between the yield and VI was indicated by SRVI, the weakest by CVI for both years 2006 and 2010.

Table 5

### Statistical analysis of relation between selected VI and yield

			NDVI	SRVI	MSI	CVI
wheat	2005	yield	0.625***	0.684***	-0.655***	0.495***
	2011	yield	0.787***	0.702***	-0.835***	0.65***
			NDVI	SRVI	MSI	CVI
oat	2006	yield	0.699***	0.768***	-0.655***	0.567***
	2010	yield	0.656***	0.694***	-0.615***	0.514***

Correlation coefficients with levels of statistical significance: \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$

The second analysis aimed to determine the relation of VI and topography represented by TWI. The results (see Table 6) on this level essentially followed the same trend as shown in Table 2. MSI was most correlated with TWI on wheat in both years, while it was SRVI by oat in both years. The lowest values were found by CVI in all cases.

Table 6

### Statistical analysis of relation between selected VI and TWI

			NDVI	SRVI	MSI	CVI
wheat	2005	TWI	0.44***	0.41***	-0.48***	0.19*
	2011	TWI	0.41***	0.320***	-0.49***	0.29***
			NDVI	SRVI	MSI	CVI
oat	2006	TWI	0.5***	0.59***	-0.51***	0.38***
	2010	TWI	0.5***	0.6***	-0.53***	0.3***

Correlation coefficients with levels of statistical significance: \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$

The last comparison of TWI and the yield in particular years described the influence of topography on the final yield. There was no difference found by oat, however, by wheat certain difference was indicated (see Table 7).

Table 7

### Influence of topography on yield

	wheat		oat	
	2005	2011	2006	2010
yield	6.04	6.92	4.72	2.1
TWI	6.64			
r	0.28**	0.61***	0.53***	0.53***

Correlation coefficients with levels of statistical significance: \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$

Considering the results, strong correlation was indicated between selected VI and the yield or topography. This positive outcome is mainly caused by the selection of crops to be investigated. VI are suitable to explain sufficiently yield variability of grains, but not so well of other agricultural crops, e.g., rape. The main difference in spectral characteristics is caused by flowering.

When evaluating the capability of selected VIs to predict the yield, SRVI and MSI tend to be most correlated with the yield data. However, NDVI showed also high values of correlation coefficients. The CVI results were also significant in all four selected seasons, but the correlation coefficients reached the lowest values. Unlike the previous indices, CVI is calculated with the second power of reflectance in green band. When canopy gets darker, the value of reflectance decreases. That may have certain influence on the resulting values of this kind of index. Nevertheless, the crop yield variability is influenced by many factors and therefore it tends to differ year by year. The most influencing factors are topography and the weather conditions [3]. Considering the results regarding the precipitation and average temperature in particular growth stages is therefore indispensable. The highest amount of precipitation was observed in 2011 (see Table 1) that resulted in very low MSI values ( $r = 0.13$ ) by wheat (see Table 2). On the contrary, the lowest precipitations were indicated in 2006, where MSI reached the highest mean value ( $r = 0.37$ ) (see Table 3). Since MSI is the type of

index that describes the crop stress caused by lack of water, the summary statistics results of the correlation analysis uphold this statement very well. Above that, TWI is also associated with the water supply. As stated by [3], convex parts of plot tend to lose water, while the concave parts accumulate water with nutrients. Table 7 gives the results of TWI and yield relation. There are considerable different results of 2005 and 2011, when wheat was grown. The yield was related to topography much less ( $r = 0.280$ ) in 2005 than in 2011 ( $r = 0.610$ ). This is probably caused by erratic precipitation during the particular growth stages in 2011 season (see Table 1). Nevertheless, there is the same value of correlation coefficient ( $r = 0.530$ ) indicated by oat in both seasons. Unlike wheat, oat is a crop with a shallow root system, so it is more dependent on immediate precipitation [3]. Summary statistics of the selected indices is presented by Table 2 for wheat and Table 3 for oat. Table 4 presents the summary statistics of the yield data and TWI. By wheat, the average yield is very similar in both selected seasons ( $6.04$  and  $6.92 \text{ t}\cdot\text{ha}^{-1}$ ), although there is a wide span between the minimal and maximal value in 2011. Repeatedly, this may be caused by erratic rainfall in this season. Summary statistics of four selected VIs showed no substantial difference in mean values. The oat yield average in 2006 was more than double than in 2010 ( $4.72$  and  $2.10 \text{ t}\cdot\text{ha}^{-1}$ ). However, spans between the minimum and maximum and standard deviations seemed to be very similar in both years. As mentioned above, oat is dependent on immediate precipitation. The difference in the yield in 2006 and 2010 is probably caused by substantially higher precipitation in the 2006 season.

Another objective of this study was to evaluate utilization of satellite images with 30 m spatial resolution when describing canopy characteristics on a small agricultural plot. In this study LANDSAT 5 images were used to derive VI for an 11.5 ha field. The results were represented by raster layers with 130 pixels. In source [6] it has been stated that LANDSAT image utilization is limited exactly for its low spatial resolution. However, there are many studies found in scientific literature that worked with LANDSAT images and got significant results. Authors in [7] evaluated the cotton yield on a 50 ha field using these images. There is also a study that describes the corn and soy yield on 100x50 km plot [8]. Authors in [9] evaluated the yield of different crops of the area corresponding with only two LANDSAT images. Nevertheless, the disadvantage was found when resampling other data (yield, topography) to 30 m spatial resolution. Resampling basically calculates average values for particular pixel. That causes a loss of certain percentage of an original dataset. As [3] state, there are still a number of advantages of LANDSAT image utilization. Besides, it is a free source of satellite data, it provides different bands of electromagnetic spectrum for VI calculation. Thus, 30 m spatial resolution seems to be sufficient for 11.5 ha plot evaluation.

## Conclusions

1. Four VIs were derived for each season from LANDSAT 5 images.
2. VI were statistically analyzed and the relation with the yield and topography was described.
3. Based on strong correlation found between VI and the yield data, the conclusion was drawn that all selected VI are suitable for yield prediction.
4. All selected VI were also very strongly correlated with topography of the examined plot represented by TWI.
5. Strong correlation was observed also between the yields in particular seasons and TWI.
6. The results were discussed in terms of the weather conditions as well. The relation varies in the selected crops.
7. Significant results were likely caused by crop selection – spectral characteristics of grains are very suitable to be utilized for crop state evaluation and yield prediction.
8. 30 m spatial resolution of the used satellite images was sufficient for evaluating a 11.5 ha agricultural plot.

## References

1. United Nations Department of Economic and Social Affairs/Population Division. World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. New York, United Nations, 2015. 59 p.
2. Zhang Q. Precision Agriculture Technology for Crop Farming. CRC Press, 2015. 360 p.

3. Kumhálová J., Zemek F., Novák P., Brovkina O. et Mayerová M. Use of Landsat images for yield evaluation within a small plot. *Plant Soil Environment*, vol. 60(11), 2014, pp. 501-506.
4. Schlegel R.H.J. *Rye: Genetics, Breeding and Cultivation*. CRC Press, 2013. 387 p.
5. Kumhálová J., Novák P., Zemek F. et Madaras M. Evaluation of oat and winter wheat variability using Landsat and ALI satellite time series within a small plot. Submitted to *Precision Agriculture*, 2015.
6. Zhang Q. et Pierce J.F. *Agricultural Automation: Fundamentals and Practices*. CRC Press, 2013. 411 p.
7. Guo W., Maas S.J. et Bronson K.F. Relationship between cotton yield and soil electrical conductivity, topography and Landsat imagery. *Precision Agriculture*, vol. 13, 2012, pp. 678-692.
8. Doraiswamy P.C., Hatfield J.L., Jackson T.J., Akhmedov B., Prueger J. et Stern A. Crop condition and yield simulations using Landsat and MODIS. *Remote sensing of Environment*, vol. 92, 2004, pp. 548-559.
9. Julien Y., Sobrino J.A. et Jimenéz-Munoz J.C. Land use classification from multitemporal Landsat imagery using the Yearly Land Cover Dynamics (YLCD) method. *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, 2011, pp. 711-720.