JOURNAL OF ENVIRONMENTAL GEOGRAPHY
Journal of Environmental Geography 9 (3–4), 49–53.
DOI: 10.1515/jengeo-2016-0012
ISSN: 2060-467X

IDENTIFICATION AND SPECTRAL EVALUATION OF AGRICULTURAL CROPS ON HYPERSPECTRAL AIRBORNE DATA

Bálint Csendes*, László Mucsi

Department of Physical Geography and Geoinformatics, University of Szeged, Egyetem u. 2-6, H-6722 Szeged, Hungary
*Corresponding author, e-mail: bcsendes@geo.u-szeged.hu

Research article, received 10 October 2016, accepted 15 December 2016

Abstract

Hyperspectral remote sensing combined with advanced image processing techniques is an efficient tool for the identification of agricultural crops. In our study we pursued spectral analysis on a relatively small sample area using low number of training points to examine the potential of high resolution imagery. Spectral separability measurements were applied to reveal spectral overlapping between 4 crop species and for the discrimination we also used statistical comparisons such as plotting the PC values and calculating standard deviation of single band reflectance values on our classes. These statistical results were proven to be good indicators of spectral similarity and potential confusion of data samples. The classification of Spectral Angle Mapper (SAM) had an overall accuracy of 72% for the four species where the poorest results were obtained from the test points of garlic and sugar beet. Comparing the statistical analyses we concluded that spectral homogeneity does not necessarily have influence on the accuracy of mapping, whereas separability scores strongly correlate with classification results, implying also that preliminary statistical assessments can improve the efficiency of training site selection and provide useful information to specify some technical requirements of airborne hyperspectral surveys.

Keywords: hyperspectral remote sensing, Spectral Angle Mapper, spectral separability measurements, agricultural monitoring

INTRODUCTION

High resolution aerial and satellite spectrometers opened new horizons for the computer-assisted analysis of land cover. Hyperspectral remote sensing techniques are not only used for identifying minerals, soils and urban surfaces, but they also constitute a powerful tool for the mapping of vegetation and natural habitats. Chlorophyll and other biochemical components have their very specific spectral characteristics, like absorption bands or the the red edge in the near-infrared wavelength region, thus reflectance curves can be accurately classified if the spectral sampling of the data is subtle enough to detect these features.

Agricultural parcels are widely used for the calibration and testing of image processing tools as the spatially separated and homogeneous blocks of crops can be easily identified both on the images and in the field surveys.

There are a high number of scientific papers that deal with the spectral analysis of different vegetation parameters, natural habitats and agricultural plants (Visi-Rajczi et al., 2012; Burai et al., 2014; Kertész et al., 2014; Lausch et al., 2015). For example, advanced machine learning algorithms were employed for the detection of mixed pixels of weeds by Moshou et al. (2001), and Liu et al. (2010) were also using vector quantization for the mapping of fungal plant infections. The scope of environmental applications includes floodplains and saline soils (Burai andTomor, 2011; Kardeván et al., 2003), laboratory measurements for the estimation of various biochemical parameters (Lausch et al., 2016) and the monitoring of plant diseases (Liu et al., 2010; Visi-Rajczi et al., 2012).

In our study we focused on some basic statistical approaches to estimate the separability and the classification accuracy of crops based on hyperspectral aerial photography. Conclusions derived from our results can be used for the improvement of field surveys, feature and noise reduction and the preliminary planning of aerial photography campaigns.

BACKGROUND

The so called curse of dimensionality is a significant obstacle to hyperspectral data interpretation, the large number of data bands combined with data noise and limited training areas can lead to poorer classification accuracy (Landgrebe, 2003). A possible way to mitigate this risk is either to improve the ground truth data set or to perform feature reduction (i. e. principal component transformation). In order to test the training pixels we can calculate spectral separability indices that indicate the overlappings between classes. Both Jeffries-Matusita and Transformed divergence algorithms were proven to be efficient tools for the evaluation of vegetation, soil and other land cover training sites (Büttner et al., 1988; Metternicht and Zinck, 1998; Ustin et al., 2009).

Another way to present spectral similarities is to plot image data to a 2 dimensional feature space where pixels are usually arranged in a triangle of soils, water surfaces and vegetation (Tobak et al., 2012).

The elaboration of ground survey is crucial for the agricultural applications of remote sensing, some studies show that adding biochemical and soil parameters increases mapping accuracy on high resolution imagery (Burai, 2006). Hence, the limitation of ground truth data is a serious constrain for the classifications which can be mitigated using non-parametric methods that are less sensitive to small deviations, like the SAM (Burai et al., 2010).

STUDY AREA

The geographical region of our study, the flood plain area called Tápai-rét is situated in southeastern Hungary, at the outskirts of the city of Szeged, over the confluence of rivers Tisza and Maros (Fig. 1). The landscape is mainly characterized by cultivated agricultural land and parcels of various sizes and settlements of scattered farmsteads. Due to its small differences of elevation and the remarkable diversity of crops Tápai-rét is an ideal site to examine spectral discrimination techniques.

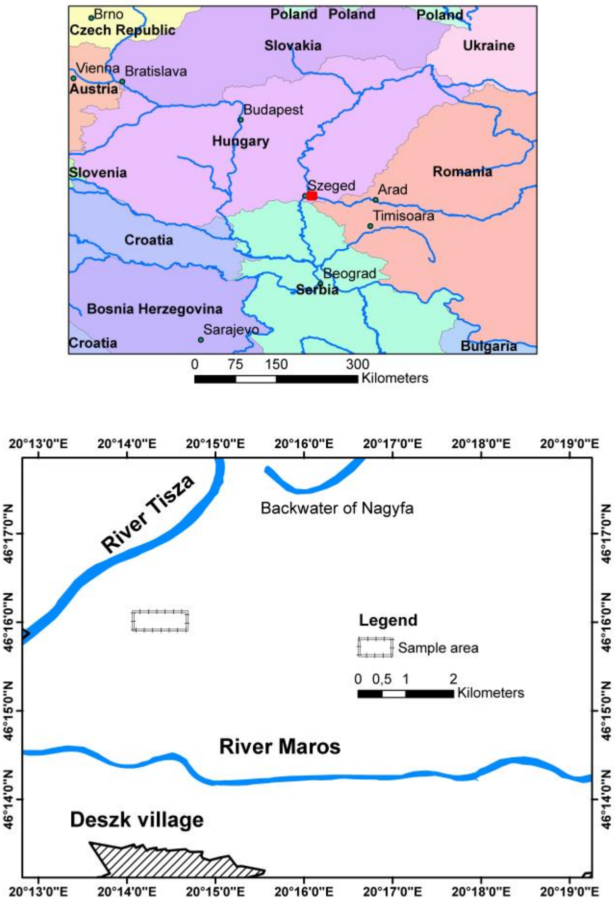

Fig. 1 The geographical location of the study area

Land use in Hungary is usually dominated by smaller agricultural parcels and it is also true for the Tápai-rét area where a great part of the territory belongs to the irregular network of farm. The mosaic of small and diverse parcels imposes an obstacle to field surveys and significantly limits the number of available training points for the crops.

DATA AND METHODS

Hyperspectral imagery were acquired in September 2010 using the airborne spectrometer of AISA. Reflectance values were recorded on 359 spectral bands between the wavelength of 0.4 and 2.4 micrometers with a spatial resolution of 1.5 metres. Some of the bands contained significant amount of noise, the values for these wavelength regions were left blank on the reflectance curve diagram (Fig. 2). Around 120 spectral bands were removed based on their spatial autocorrelation values, however, we kept all the data from the region between 400 and 900 nanometers as these reflectance values are the most informative about vegetation characteristics.

Fig. 2 Average spectra of the four crop species

We used online cadastral maps (Fig. 3) and the expertise of local farmers to collect ground truth data about the crops. The online map contains the 5 digit ID numbers for each parcel that are connected to the ownership information, the database is managed and regularly updated by the Hungarian Cadastral Office.

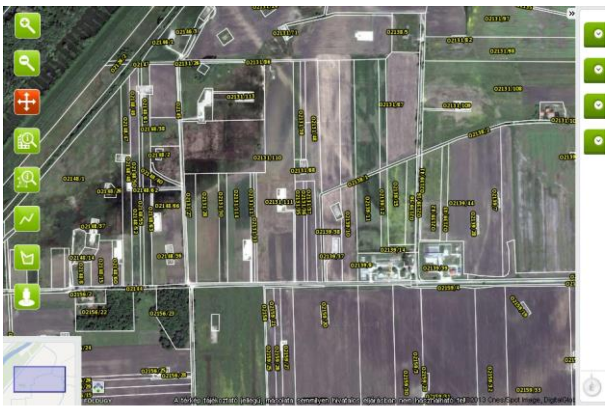

Fig.3 The online cadastrial map of the study area

An accurate method to verify the selection of training points is the calculation of spectral separability. Formulas like the Jeffries-Matusita or the Transformed Divergence distance show index values on a scale of 0 to 2 where 0 refers to complete overlapping and 2 indicates perfect separability. These calculations are based on the comparison of reflectance or other data values within a certain range. However, it is essential to have more training points than the dimensionality, or in other words, the number of the spectral bands. In order to meet this criteria, feature reduction of the hyperspectral imagery is required if the set of data samples cannot be extended. For this purpose, in our study we performed a principal component transformation, where the first few transformed data bands contain most of the spectral information with a reduced degree of noise.

Another important aspect of the training site selection is the spectral homogeneity of the pixels. It can be affected by some possible spatial autocorrelation, the presence of noise, and also the heterogeneity of the examined land cover features. We selected spatially diverse training areas for the better characterization of our classes and to avoid the use of pixels with similar traces of noise. Standard deviation was calculated for single spectral bands both from the visible (550 nm) and the near-infrared region (900 nm).

The Spectral Angle Mapper (SAM) algorithm was applied for the classification which is considered as a relatively simple spectral statistical method as it calculates only the average of the sample spectra and the vector angle deviations of the individual test points, measured by the milliradian. Our goal was to compare the results of this non-parametric classifier with the values obtained from the training site image statistics.

RESULTS

Spectral separability measurements

To examine the possible spectral overlappings between the classes we performed spectral separability measurements. Jeffries-Matusita (JM) and Transformed Divergence (TD) indices have proven to be powerful tools for the evaluation of training areas (Metternicht and Zinck, 1998; Ustin et al., 2009; Tobak et al., 2013). Since these calculations require more spectral samples than the number of the input bands, we used a principal component transformation to reduce data dimensionality. Separability values were measured on the first two PC bands for the four crop classes (Table 1).

The TD index happened to be less sensitive to the overlappings than the JM, however, the results are in agreement as the ranking of the classes is the same, medick is completely separable from the rest of the species, while a high extent of spectral similarity occurs between the pixels of the other 3 classes, especially in the case of those of garlic and sugar beet.

Table 1 Jeffries-Matusita (JM) and Transformed Divergence (TD) indices based on the values of the first two Principal Component bands

	Medick	Corn	Garlic	Sugar beet	TD
Medick	0	2	2	2	Medick
Corn	2	0	1.08	0.58	Corn
Garlic	2	0.78	0	0.27	Garlic
Sugar beet	2	0.53	0.26	0	Sugar beet
JM	Medick	Corn	Garlic	Sugar beet	

Scatter plot visualization

Spectral classes can be visualized by plotting the pixel values of certain image bands, in our study we used the first two principal component bands to illustrate the relative positions of the pixel groups in the spectral space and indicated denser areas with lighter colour shades (Fig. 4). The elements of 2 dimensional plots usually form the shape of a triangle when studying imagery of a natural landscape where the three endmembers are vegetation, water and soil surfaces (Mucsi and Henits, 2011; Tobak et al., 2012). The scatter plot of our study area has the pixels of medick in the upper left corner, separated from the other crops which are closer to the brighter central region where most of the vegetation data points are located, thus confirming the findings of the separability measurements. In comparison, grassland pixels from the external regions of the original aerial photography can be found scattered below the main line of vegetation, water and some shaded surfaces are situated in the upper right corner of the triangle, while soils and concrete are in the bottom (Fig. 4).

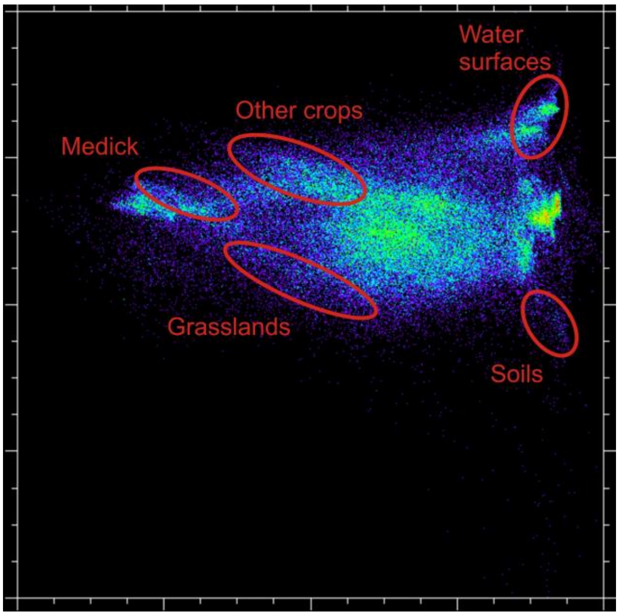

Fig.4 Crops featured on the scatter plot of the first two2 PC bands

Homogeneity analysis

The outcome of the supervised classification is affected by the homogeneity of the training areas, therefore we analysed the standard deviation of reflectance values per single spectral bands (Table 2), where data from the visible range (550 nanometres), from the near infrared (900 nm), and one band from the short-wavelength infrared (SWIR) region (2100 nm) were used.

Reflectance values were found more homogeneous in the visible and in the SWIR wavelength regions, while there is a significant increase of the standard deviation at the near infrared light range that can be explained with the individual characteristics of the plants' spectral red edge. Corn shows relatively low heterogeneity compared to the other three species, however, these values depend not only from the spectral features of the vegetation, but also from the spatial distribution and the noise content of pixels.

Table 2 Standard deviation of reflectance values per single bands

Wavelength	Medick	Corn	Garlic	Sugar beet
550 nm	0.0090	0.0031	0.0067	0.0037
900 nm	0.0394	0.0235	0.0357	0.0337
2100 nm	0.0068	0.0041	0.0068	0.0062

SAM classification

The SAM classification was performed without any spectral angle threshold specified to have more information on misclassifications between our crop classes. Table 3 shows the confusion matrix of the result, where columns represent the ground truth points (15 items per class) and rows display the classified pixels.

The two approaches of classification accuracy measurement (Users' and Producer's Accuracy) show more or less the same results on the reliability of identification. The overall accuracy of the mapping is on an acceptable level (72%), however, significant differences can be observed between the results of certain classes. The worst performance was in the category of sugar beet, where the slight majority of control pixels where misclassified. Also, the highest extent of confusion was registered between sugar beet and garlic.

Figure 5 presents the spatial distribution of the 4 classes on the true colour hyperspectral image. As it can be seen on the picture, the mapping categories of the crops rarely extend beyond the parcels, the SAM classification rather omitted to detect certain pixels of agricultural vegetation. The parcels of corn and medick are easily recognizable, their mapping classes designate more or less homogenous bocks. On the other hand, in the case of the two western parcels there is a high level of misclassification, especially between the categories sugar beet, garlic, and corn, as it was predicted by the separability measurements.

Table 3 Confusion matrix of the SAM classification, indicating Users' Accuracy (U. A.) and Producer's Accuracy (P. A.)

	Medick	Corn	Garlic	Sugar beet	U. A.
Medick	15				100%
Corn		12	1	3	75%
Garlic			9	5	64%
Sugar beet		3	5	7	47%
P. A.	100%	80%	60%	47%	72%

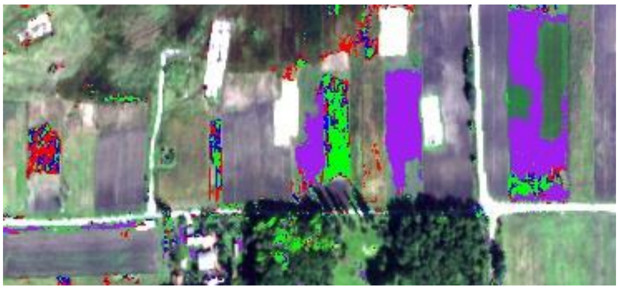

Fig. 5 The four SAM classes placed over the true colour image of the hyperspectral data: sugar beet (blue), garlic (red), corn (green), and medick (purple)

DISCUSSION

SAM classifier was proven to be an accurate technique for the spectral discrimination of most of the crop species, however, in the cases of spectrally less separable plants (sugar beet and garlic) classification results were significantly poorer. The calculation of spectral angles has also the advantage that the outcome of the analysis is insensitive to the level of illumination of the surface objects, thus shaded pixels will have a lower rate of misclassification (van der Meer, 2004; Lillesand et al., 2004; Kruse et al., 1993). Some papers also suggest that SAM provides acceptable results on hyperspectral data when the number of training points is significantly limited (Burai et al., 2010; Tobak et al., 2012), although an increasing number of authors prefers machine learning algorithms for advanced classifications and land cover mapping (Huang et al., 2002; Lary et al., 2015).

Our finding, that the Jeffries-Matusita distance is more sensitive to spectral similarities than Transformed Divergence has confirmed the same conclusions of Jensen 1986.

In his study Burai (2006) examined the spectral features of similar plants (medick, corn, sugar beet, etc.) in a much larger study area, obtaining an overall mapping accuracy of 85,5%. He argued that the use of an extended and more detailed ground truth database including soil parameters can significantly improve the reliability of classifications, which is in line with our conclusions.

CONCLUSION

Our main goal was to compare classification results with some spectral statistical analyses we performed on the PC transformed data and on single spectral bands. Using spectral separability measurements we found two classes that show significant overlapping (sugar beet and garlic) what was also confirmed by the SAM accuracy results. Also, it was proven that spectral separability distances correlate with classification results, where a higher degree of spectral overlapping on the PC transformed data can lead to poorer accuracy even in the case of a high resolution hyperspectral dataset. We also examined the spectral homogeneity of training points and we concluded that the standard deviation values of the classes do not show a strong correlation with the classification's outcome.

As the SAM results resembled those of the separability distance calculations, we drew the conclusions that despite the possible data loss, principal component transformation is an applicable tool for the identification and comparison of crops on high resolution imagery even when spectral differences are very subtle and the size of training areas is limited.

Acknowledgement

This research was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program'.

References

Burai, P. 2006. Földhasználat-elemzés és növény-monitoring különböző adattartalmú és térbeli felbontású távérzékelt felvételek alapján (Land use analysis and vegetation monitoring using remote sensed images of different data content and spatial resolution). *Agrártudományi Közlemények* 2006/22 Különszám, 7–12. (In Hungarian)

Burai, P., Tomor, T. 2011. Hiperspektrális felmérés eredményei az Ipoly Balassagyarmat és Ipolytarnóc közötti szakaszán (The results of the hyperspectral survey along river Ipoly between Balassagyarmat and Ipolytarnóc. In: A Bükki Nemzeti Park Igazgatóság természeti értékeinek kutatása I.: „Az Ipoly–vízgyűjtő vizes élőhelyeinek komplex felmérése, közösségi jegyzékeinek kidolgozása" Felsőtárkány, 2011 (Conference book, in Hungarian). Online at: https://bnpi.hu/file/45/.

Burai, P., Lövei, G. Zs., Lénárt, Cs., Nagy, I., Enyedi, P. 2010. Mapping aquatic vegetation of the Rakamaz-Tiszanagyfalui Nagy-morotva using hyperspectral imagery. *Acta Geographica Debrecina* Landscape & Environment, 4 (1), 1–10.

Burai, P., Deák, B., Orolya, V., Lénárt, Cs. 2014. Mapping of Grass Species Using Airborne Hyperspectral Data. In: Pfeifer, N., Zlinszky, A. (eds) Proceedings of the International Workshop on Remote Sensing and GIS for Monitoring of Habitat Quality, Vienna, 87–88.

Büttner, Gy., Csillag, F., Mather, P. M. 1988. Spectral And Spatial Information Content Of Spot Data. In: Geoscience and Remote Sensing Symposium, 1988. IGARSS '88. Remote Sensing: Moving Toward the 21st Century. International; 10/1988 DOI: 10.1109/IGARSS.1988.570172

Huang, C., Davis, L. S., Townshend, J. R. G. 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23, 725–749. DOI: 10.1080/01431160110040323

Jensen, J. R. 1986. Introductory digital image processing: A remote sensing perspective. Englewood Cliffs, NJ: Prentice-Hall.

Kardeván, P., Vekerdy, Z., Róth, L., Sommer St.–Kemper, Th., Jordan, Gy., Tamás, J., Pechmann, I., Kovács, E., Hargitai, H., László, F. 2003. Outline of scientific aims and data processing status of the first Hungarian hyperspectral data acquisition flight campaign, 3rd EARSeL Workshop on Imaging Spectroscopy in Oberpfaffenhofen, 13-16 May 2003

Kertész, P., Király, G.,Burai, P. 2014. Tree Species Mapping Using Airborne Hyperspectral Remote Sensing. In: Pfeifer, N., Zlinszky, A. (eds) Proceedings of the International Workshop on Remote Sensing and GIS for Monitoring of Habitat Quality, Vienna, 60–62.

Kruse, F. A., Lefkoff, A. B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A. T., Barloon, J. P., Goetz, A. F. H. 1993. The spectral image processing system (SIPS) - Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment* 44, 145–163. DOI: 10.1016/0034-4257(93)90013-n

Landgrebe, D. A. 2003. Signal Theory Methods in Multispectral Remote Sensing. John Wiley & Sons, Inc.

Lary, D. J., Alavi, A. H., Gandomi, A. H.,Walker, A. L. 2015. Machine learning in geosciences and remote sensing. *Geoscience Frontiers* 7, 3–10. DOI: 10.1016/j.gsf.2015.07.003

Lausch, A., Salbach, C., Doktor, D., Schmidt, A., Merbach, I.,Pause, M. 2015. Deriving phenology of barley with imaging hyperspectral remote sensing. *Ecological Modelling* 295, 123–135. DOI: 10.1016/j.ecolmodel.2014.10.001

Lausch, A., Bannehr, L., Beckmann, M., Boehm, C., Feilhauer, H., Hacker, J. M., Heurich, M., Jung, A., Klenke, R., Neumann, C., Pause, M., Rocchini, D., Schaepman, M. E., Schmidtlein, S., Schulz, K., Selsam, P., Settele, J., Skidmore, A. K., Cord, A. F. 2016. Linking Earth Observation and taxonomic, structural and functional biodiversity: Local to ecosystem perspectives. Ecological indicators. DOI: 10.1016/j.ecolind.2016.06.22 (in press)

Lillesand, T. M., Kiefer, R. W., Chipman, J. W. 2004. Remote Sensing and Image Interpretation (Fifth Edition). John Wiley & Sons, Inc.

Liu, Zh-Y., Wu, H-F., Huang, J-F. 2010. Application of neural networks to discriminate fungal infection levels in rice panicles using hyperspectral reflectance and principal components analysis. *Computers and Electronics in Agriculture* 72, 99–106. DOI: 10.1016/j.compag.2010.03.003

van der Meer, F. D. 2004 Pixel-Based, Stratified and Contextual Analysis of Hyperspectral Imagery. In: de Jong, S. M., van der Meer, F. D. (eds) Remote Sensing Image Analysis: Including the Spatial Domain, Dordrecht, Springer Academic Publishers, 153–180.

Metternicht, G. I., Zinck, J. A. 1998. Evaluating the information content of JERS-1 SAR and Landsat TM data for discrimination of soil erosion features. *ISPRS Journal of Photogrammetry & Remote Sensing*, 53, 143–153. DOI: 10.1016/s0924-2716(98)00004-5

Moshou, D., Vrindts, E., De Ketelaere, B., De Baerdemaeker, J., Ramon, H. 2001. A neural network based plant classifier. *Computers and Electronics in Agriculture* 31, 5–16. DOI: 10.1016/s0168-1699(00)00170-8

Mucsi, L. Henits, L. 2011. Belvízelöntési térképek készítése közepes felbontású űrfelvételek szubpixel alapú osztályozásával (Mapping inland excess water using subpixel classification of middle resolution satellite images) *Földrajzi közlemények* 135(4), 365–378. (In Hungarian).

Tobak, Z. 2013. Urban surface analyses using high temporal and spectral resolution aerial imagery. PhD theses, University of Szeged, Hungary.

Tobak, Z., Csendes, B., Henits, L., van Leeuwen, B., Szatmári, J., Mucsi, L. 2012. Városi felszínek spektrális tulajdonságainak vizsgálata légifelvételek alapján (Spectral analysis of urban surfaces using aerial photography) In: Lóki, J. (ed) Az elmélet és gyakorlat találkozása a térinformatikában III, Debrecen, 413-420. ISBN:978-963-318-218-5 (In Hungarian:).

Ustin, L. S., Valko, P. G., Kefauver, S. C., Santos, M. J., Zimpfer, J. F., Smith, S. D. 2009. Remote sensing of biological soil crust under simulated climate change manipulations in the Mojave Desert. *Remote Sensing of Environment* 113, 317–328. DOI: 10.1016/j.rse.2008.09.013

Visi-Rajczi, E., Burai, P., Király, G., Albert, L. 2012. Ecological Characterization of the green areas in Sopron by Plant Chemical Analysis and Hyperspectral Recording. In: The Impact of Urbanization, Industrial and Agricultural Technologies on the Natural Environment: International Scientific Conference on Sustainable Development and Ecological Footprint. University of West Hungary, Sopron. ISBN 978-963-334-047-9