ENTROPY AND FRACTAL STRUCTURE BASED ANALYSIS IN IMPACT ASSESSEMENT OF BLACK CARBON POLLUTIONS

Veronika Kozma-Bognár[1]*, József Berke[2]

[1] *University of Pannonia Georgikon Faculty, Keszthely, Hungary*

[2] *SFD Information Technology Ltd., Keszthely, Hungary*

*kbv@georgikon.hu

Abstract

Nowadays remote sensing has become a widely used technology collecting information about our environment. Aerial and satellite images are the basic data source in researches where one of the main aims is to detect environmental damaging phenomena, spatial distribution, determine affected areas, identify propagation direction. While assessing the toxicity of heavy metals in many cases, the biological acute toxicity tests are used to determine the effects on remote sensing of heavy metal pollution may also be relevant. Plants can behave as indicator of a certain stress effect caused by environmental pollution. Monitoring the changes in growing and developmental pattern of crop by remote sensing technology the effects of heavy metal pollution caused by traffic can be detected. Several researches confirm that the more information is collected from different spectral ranges of electromagnetic spectrum the more precise results we get in the investigation processes of environmental contamination and its effects.

From 2010 researches of plant-soil-atmosphere system have been carried out at the University of Pannonia, Georgikon Faculty. One of the main objectives is to study and detect the effects of environmental pollution. Impacts of black carbon originated from road traffic on maize growth and development were investigated under field experiment at the Meteorology and Water Department. Exhaust containing black carbon deposits on the leaf surface modifying plant parameters such as radiation and water supply. As a result the radiation balance of the Earth changes thus black carbon soot contributes to global warming.

From the growing season of 2010 parallel with field measurements we took aerial images with the purpose of testing and studying the potentials of different spectral range (visible, near infrared, far infrared ranges) aerial imaging techniques. Subsequently we expanded our research on the examination of information content of time-series images (intensity based, entropy based, spectral fractal based). These tests were to analyse the applicability and to find the best data analysing methodology in the research processes managing road traffic pollutants.

Keywords: environmental pollution, multispectral aerial photography, image processing, entropy, spectral fractal dimension

Összefoglalás

Jelenleg a távérzékelési módszerek egyre szélesebb körben alkalmazott technológiák a környezetünkről gyűjtött információk megszerzésében. A légi- és űrfelvételek nélkülözhetetlen adatforrásként jelennek meg olyan kutatási területeken, ahol a környezetkárosító jelenségek felderítése, térbeli eloszlása, a hatásterületek meghatározása, a terjedési irányok beazonosítása kulcsfontosságú cél. Bár a nehézfémek toxikus hatásának megítélése sok esetben biológiai akut toxikológiai tesztek alapján történik a nehézfém-szennyezés hatásainak távérzékeléssel

történő meghatározása is releváns lehet. A környezetszennyezés által okozott stresszhatásokra a növények indikátorként viselkednek. A növényállomány növekedési és fejlődési ütemében bekövetkezett változások távérzékelés útján történő nyomon követésével kiválóan kimutathatóak a közlekedés okozta nehézfém-szennyezések hatásai. Számos szakirodalmi példa mutatja, hogy az elektromágneses spektrum több különböző spektrális tartományában gyűjtött információ felhasználásával még pontosabb eredmények érhetők el a szennyezések környezetünkre gyakorolt hatásának vizsgálatai során.

A Pannon Egyetem Georgikon Karán 2010 óta folynak növény-talaj-légkör rendszerhez kapcsolódó tudományos kutatások, melyek egyik célja a nehézfém szennyezések várható károsító hatásainak detektálása. A Meteorológiai és Vízgazdálkodás Tanszék szabadföldi kísérletei során a közlekedés okozta környezet-szennyezések közül a korom-szennyezés kukorica növényállományra gyakorolt eredményét vizsgálják. Kiválasztásának fő oka, hogy a kipufogókból származó korom lerakódik a növények felületére, megváltoztatva ezzel annak sugárzási, vízháztartási és egyéb paramétereit. Mindez hatással van a Föld sugárzási egyensúlyára, melynek következtében a korom bizonyos fokig hozzájárul a globális felmelegedéshez is.

A 2010. évi tenyészidőszaktól kezdődően a földi mérésekkel párhuzamosan légifelvételezésekre is sor került, melynek elsődleges célja a különböző spektrális tartományú (látható-, közeli infravörös-, távoli infravörös tartományok) légifelvételezési technikák által kínált lehetőségek tesztelése, tanulmányozása volt. Ezt követően a vizsgálatok kiterjedtek az idősoros felvételek információtartalmú elemzéseire is (intenzitás alapú, entrópia alapú, spektrális fraktáldimenzió alapú), amelyek a kutatásaink szempontjából a legmegbízhatóbb adatkiértékelési eljárások alkalmazhatóságát vizsgálták a közlekedés eredetű szennyezőanyagok kutatásai során.

Introdution

The 21th century developments in information technology have opened new horizons for the processes and phenomena detection of our planet and the surrounding universe. These developments left their mark on the field of remote sensing science, so nowadays the remote sensing technologies are becoming more widely used applications to understand our environment more comprehensive. The database of the aerial and satellite images are extremely important in prevention, detection and monitoring of the problems caused by the pollution. Thanks to technological developments of the last decade are increasingly used the sensors with high spectral and spatial resolution (Nagy 2005) that allow more accurate results in investigation of the effect of environmental pollutions.

There are many examples of literature shows that remote sensing images taken in different spectral ranges further contribute to understand more exact the processes and phenomena our environment. In remote sensing after the appearance of multichannel satellite sensor (e.g. AVIRIS, LANDSAT, SPOT, Quickbird, ASTER, MODIS etc.) the development of airborne detector started (APEX, ARES AISA, DAIS etc.). As a consequence, increased the diversity of applications in industry, geology, meteorology, agriculture, forestry, environmental protection, defence and remediation activities etc. (Sabins 1987). These advanced information technology solution make it possible to solve many practical environmental task for example spread of plat diseases, ragweed detection, prediction of biomass and forest fires, investigation of water quality, air pollution, soil contamination, waste management, effect of remediation, urban ecological changes, habitat mapping etc. (Boregasser et al. 2008). The technical development of the sensors is followed by the significant delay of the processing methods and applications (Nixon and Aguado 2008, Young and Fu 1986). Therefore it is reasonable to develop new methods, in order to get more useful information from measured data.

Materials and Methods

Research area

The testing area was situated at Agro-meteorological Research Station in Keszthely (Hungary), where maize crops were polluted by heavy metals. Our test area was located north of the station (the centre of the test area N: 46°44'08.55", E: 17°14'19.76", H:114 m) where six plots were established (Figure 1.) (Kozma-Bognar and Berke 2012):

- BC = Black Carbon, polluted,
- BC-W = Black carbon polluted and irrigated,
- Cont = Control,
- Cont-W = Control irrigated,
- Cd = Cadmium polluted,
- Cd-W = Cadmium polluted irrigated.

In our study we simulated the effects of black carbon pollution. We applied maize hybrid Sperlona (FAO 340) as test plant with short growing season. The black carbon was dispensed by a motorised sprayer of SP 415 type. During the research we used chemically “pure” black carbon, which was free any contaminants. We sprayed black carbon 3 gm^{-2}/week doses onto the leaf surface to see the effect of the growth of plants.

Figure 1. Aerial image of 2012 in VIS

Research activities also focused on the examination of the potential negative or positive influence of the irrigation to. The irrigated plots received over the natural precipitation additional water supply. The dropping irrigation method was performed, 4-6 mm/hour intensity, depending on the weather (Anda and Illes 2012).

Research Tools

During three growing seasons (2010, 2011, 2012) was collected remote sensing data in different spectral ranges parallel with the field measurements. We used digital sensors in the visible (VIS), Near InfraRed (NIR) and Far InfraRed (FIR) spectral ranges to get more information about the effect of the black carbon pollution. In the growing season we took aerial images in each phonological phases, to follow the changes of the maize plant (Table 1.)

Growing season in 2010	Growing season in 2011	Growing season in 2012
25/04/2010	04/06/2011	02/04/2012
27/06/2010	20/06/2011	29/05/2012
21/07/2010	04/07/2011	10/06/2012
09/08/2010	17/07/2011	21/06/2012
17/08/2010	01/08/2011	23/07/2012
26/08/2010	22/08/2011	24/08/2012
09/09/2010	03/09/2011	07/09/2012

Table 1. Date of the aerial photography

During each aerial flight were mapped the plots more tracks (3-5 track logs) to get enough data to the statistical analysis. The spatial resolution of these images was under 10x10 cm in VIS, NIR spectral range, and 30x30cm in FIR spectral range, which allowed using an accurate plot-level evaluation (Kozma-Bognar et al. 2012). High intensity and spatial resolution data was an important part of the multitemporal imagine sensing. The parameters oh the major collection tools were used for mapping can be seen in Table 2.

Parameters	Visible data (VIS)	Near Infrared data (NIR)	Far Infrared data (FIR)
Type of sensor	Canon 30D	Canon 30DIR	HX-IDS-M 110
Height of flight	400 m	400 m	400 m
Spectral band	400 - 700 nm	720 - 1150 nm	8000 - 14000 nm
Geometrical resolution	0.1 m	0.1 m	0.3 m
Data recording	14 bit/pixel	14 bit/pixel	14 bit/pixel

Table 2. The main parameters of remote sensing data collection

Time-series analyses were carried out based on the remote sensing data. To perform the analysis we used various image processing techniques. With these pre-processing methods were selected the optimal images for our research. After data pre-processing intensity, entropy and spectral fractal dimension measurement evaluation methods were applied to examine black carbon polluted and control maize canopy.

Research methods

Entropy

Information-theoretic concept of entropy is used today in 1948, Claude E. Shannon (Shannon 1948a, 1948b) respectively, and illustrated through a practical example (Shannon 1951), which was called the proposal by John von Neumann entropy function. Accordingly, the average information content of the message (in the case of independent messages) - entropy can be defined as follows:

$$H = \sum_{i=1}^{m} p_i \, ld\left(\frac{1}{p_i}\right) \quad (1)$$

(1

Where H - the information-theoretic entropy

p_i - the ith message occurrence probability

General definition of entropy mathematical sense by Alfred Rényi (Rényi 1961) that the

$$H_{\propto}(X) = \frac{1}{1-\propto}\log\left(\sum_{i=1}^{n} p_i^{\propto}\right) \quad (2)$$

Where $\propto \geq 0$ and $\propto \neq 1$

Should also take into account the following when calculating the entropy in many practical situations:

1. A closed system can be the following values for the entropy of information theory (1) as:

$$0 \leq H \leq \log_2 n$$

Where n is the number of possible messages.

2. The entropy (1) is a minimum (H_{min}=0) if the source is still sending the same message.
3. The entropy (1) is taken to be the maximum value $H_{max}=\log_2 n$, if all the messages with equal probability ($p_i = -\log_2 n$).

Spectral Fractal Structure

The SFD is a general fractional dimension derived from the structure of the processing (Mandelbrot 1983), which is a novel application of fractals. The spatial dimensional structure of the SFD outside of the spectral bands is also suitable for measuring the colour structure (Berke 2006, Berke 2007) and provides sufficient information to colours, shades and other fractal properties as well. Calculating the values of the SFD (two or more band images), the defined fractal dimension is applied to the measured data as a function (the total number of boxes valued spectral function of the spectral box) simple mathematical averages are calculated as follows (Berke 2007):

$$SFD_{measured} = \frac{n \times \sum_{j=1}^{S-1} \frac{\log(BM_j)}{\log(BT_j)}}{S-1} \quad (3)$$

Where,

n – the number of image layers or image channels

S – the spectral resolution (in bits)

BM_j - valued spectral pixel boxes containing j number of bits

BT_j – all possible spectral boxes for the number of j-bits

The possible number of boxes spectral j bits is calculated as follows:

$$BT_j = (2^S)^n \quad (4)$$

The formula (3) metric (Berke 2007), the evaluations of both hyper- and multispectral images can be used for exact measurements.

Results

The 2010th, 2011th and 2012th based on the year is near and thermal infrared aerial photographs of multi-temporal analyses were performed with carbon black and cadmium contaminated and control for stocks. Our research included the analysis of the possible positive or negative effects of the irrigation in a way that also examined the differences between the irrigated and non-irrigated stocks.

The multispectral images and the typical starting reference data began to entropy-based and SFD-based different plant populations measurements. All this was done, as most mathematical evaluation methods for average information content (mean, standard deviation, etc.) or structure (PCA, ICA, SFD, etc.) work on the basis based data. During the evaluation of test shots with masked areas of crops were analysed and measured entropy and SFD values only in respect of these. The results are summarized became possible to determine which method results mainly reflect the impact of plants on the carbon and cadmium, as well as the application of which parameters are to be distinguished from the most unpolluted and heavy metal polluted plots. The six-parameter studies were included:

- typical recordings spectral range (λ)
- treatment types (p)
- temporal variations in the vegetation period (t)
- changes in the yearly analyses (T)
- geometric resolution (r)
- water supply (w).

The final goal was to determine the dependence for each parameter of the entropy and SFD based data structure.

According to the values in the table 3nd clearly establish that the average information content (entropy) based NIR range data show no significant difference between the carbon black, cadmium, and the control data. However, data obtained from the SFD structural studies were significant differences that are capable of detecting differences between the treatments (Table 3.). In our view, the 07/09/2012 nearly the same SFD values can be explained by the complete cessation of photosynthesis, the plants withered state.

Parameters	Treatments	Date of aerial photography in 2012						
		02/04	29/05	10/06	21/06	23/07	14/08	07/09
Entropy values	Black carbon	13.2877	13.2877	13.1397	13.1397	13.2877	13.2877	13.2877
	Control	13.2877	13.2877	13.1396	13.1397	13.2876	13.2877	13.2877
SFD values	Black carbon	1.0407	1.1771	1.2627	1.2591	1.3048	1.3971	1.3692
	Control	1.0648	1.1705	1.2545	1.2339	1.2618	1.3802	1.3693

Table 3. An examination of the differences between the black carbon and the control sample from entropy and SFD NIR range 2012th year's growing season

During our research we aimed to develop a method that allows for multiband aircraft images, which selects maximum information to be considered in the production of optimal images. Another consideration is that when different methods are used together sensors produced high spectral resolution images and high spatial resolution data, which is workable and useable (Kozma-Bognár and Berke 2012).

The starting point of the multispectral images, using the characteristics of the data reference, began when the different plant populations, spectral fractal dimension measurements were taken (Table 4.). The measurement program developed by SFD Information Technology Ltd (Internet site of SFD). Was considered optimal for spectral bands used to select and validate the fractal dimension (SFD) values. A review of the images containing the tested crop areas were analysed, and those values only measured SFD.

We have developed the channel selection procedure based on the SFD values of the multispectral images used to optimize the selection of images. In investigation, the SFD values were deduced from the size of the actual information content of images. Airborne imaging of one workflow (e.g. flight track) recording usually occurs in the Keszthely sample area where they were looking to provide the best and most reliable information for the object-finding study. The multispectral images of the pre-processing, post-processing only recorded SFD maximum purchasing values per sensor and per flight (Table 4.), as further investigations of these images gave the most reliable results (hit accuracy as well as other measured parameters and correlation studies).

Number of the flight track	Visible SFD values 3 bands	Near Infrared SFD values 3 bands
Flight track 1	2.1361	1.9418
Flight track 2	2.1416	1.9187
Flight track 3	2.1398	1.9400
Flight track 4	2.1328	1.9215
Flight track 5	2.1273	1.8406

Table 4. Maximum SFD values of the multispectral images

In our research the professional digital photography is the practice of using trade-specific software, including widespread noise reduction and image enhancement features, the camera type and the optics are aligned corrective actions (chromatic aberration, vignetting, geometric distortion, noise by sensors, etc.) were performed in order to examine what influence the operation of the fractal structure correction in the spectral images of the human visual system is adapted (Internet site of DxO Optics Pro 8 software).

	3 bands visible image		3 bands near infrared image	
Number of the flight track	**Basic data**	**After DxO v8.0 correction**	**Basic data**	**After DxO v8.0 correction**
Flight track 1	2,1361	2,1610	1,9418	1,9412
Flight track 2	2,1416	2,1733	1,9187	1,9471
Flight track 3	2,1398	2,1665	1,9400	1,9470
Flight track 4	2,1328	2,1648	1,9215	1,9328
Flight track 5	2,1273	2,1716	1,8406	1,8973
Average	**2,1355**	**2,1674**	**1,9125**	**1,9331**

Table 5. The multispectral images SFD values before and after correction, flight per track

The results unambiguously confirmed by an increase in the fractal dimension of the structure (diversity) after the corrections. For the near-infrared range, however, it is more random images, no systematic basis of the measured data (Table 5). The fourth image can be seen from the offset correction about the structural value of the original image data.

We were able to isolate four major groups based on the combined (PCA) test data measured on the ground and extracted on the basis of aerial photographs (Table 6). The four parameters show a temperature, reflectance, Chlorophyll and structure-based grouping.

	Component			
	1	2	3	4
Land average temperature	,965	,026	-,045	-,237
Micro climate	,926	-,082	-,255	-,110
SFD FIR	,914	,099	-,003	-,081
FIR temperature	,900	,271	,093	,310
Albedo	,083	,973	-,002	-,201
NIR Reflectance	-,229	,863	,123	-,403
VIS Reflectance	,387	,851	,060	,342
Relative humidity	-,591	-,601	-,442	,294
Chlorophyll A	-,024	,225	,964	,126
Chlorophyll B	-,188	-,076	,945	-,106
Leaf Area Index	,095	,095	,817	-,537
SFD VIS	-,109	-,122	-,082	,982
SFD NIR	-,165	-,630	-,183	,730

Table 6. The measured on the ground and in aerial photos taken from the PCA Rotated Component Matrix of parameters

Conclusions

As conclusion it could be determined which examination methods represent the real information content of aerial images. The entropy values of the images of traffic pollutant plants showed a significant difference only in part by spectral range, yearly analyses and water supply. From the entropy values of treatments types are there any differences, so that in our view, the entropy-based analysis does not work in this case. The average information content gives appropriate results in the investigation of vegetation period and geometric resolution. If we evaluate the fractal structure of the black carbon and control plant images we get positive results for each of the six different types. Consequently, it can be said that the spectral fractal dimension parameter is well used to determine the actual information content of the aerial images in the spectral range, treatment types, vegetation period, yearly analyses, geometric resolution, water supply examinations.

Acknowledgement

This research was realized in the frames of TÁMOP 4.2.4. A/1-11-1-2012-0001 „National Excellence Program – Elaborating and operating an inland student and researcher personal support system" The project was subsidized by the European Union and co-financed by the European Social Fund. This article was made partly under the project TÁMOP-4.2.2/B-10/1-2010-0025. This project is supported by the European Union and co-financed by the European Social Fund.

References

Anda, A., Illés, B. 2012. Impact of simulated airborne soot on maize growth and development. J Environ Prot Vol. 3. 8: 773-781.

Berke, J. 2006. Measuring of Spectral Fractal Dimension, Advances in Systems Computing Sciences and Software Engineering. Springer-Verlag pp. 397-402. ISBN 10 1-4020-5262-6.

Berke, J. 2007. Measuring of Spectral Fractal Dimension, Journal of New Mathematics and Natural Computation. vol. 3/3. pp. 409–418. DOI: 10.1142/S1793005707000872.

Borengasser, M., Hungate, W.S., Watkins R. 2008. Hyperspectral Remote Sensing: Principles and Applications. Taylor and Francis Group New York. ISBN 978-1-56670-654-4.

Internet site of parameter Spectral Fractal Dimension, http://www.digkep.hu/sfd/index.htm.

Internet site of DxO Optics Pro 8 software, www.dxo.com.

Kozma-Bognár, V., Berke, J. 2012. Determination of Optimal Hyper- and Multispectral Image Channels by Spectral Fractal Structure, International Joint Conferences on Computer, Information, and Systems Sciences, and Engineering. December 7-9, 2012.

Kozma-Bognár, V., Berke, J., Martin, G. 2012. Application possibilities of aerial and terrain data evaluation in particulate pollution effects, European Geosciences Union General Assembly, EGU2012-3063, 22-27 April, 2012, Wien.

Mandelbrot, B.B. 1983. The fractal geometry of nature. W.H. Freeman and Company, New York.

Nagy,T. 2005. Resurrection of Thermovision, Magyar Elektronika 1-2.

Nixon, M., Aguado, A. 2008. Feature Extraction and Image Processing. Academic Press. ISBN: 978-0-1237-2538-7.

Rényi, A. 1961. On measures of information and entropy, Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability. 1960:547–561.

Sabins, F.F. 1987. Remote Sensing. Principles and Interpretation. W. H. Freeman and Co. Los Angeles. ISBN 0-7167-1793-X.

Shannon, C. E. 1948a. A Mathematical Theory of Communication. The Bell System Technical Journal. 27:379–423.

Shannon, C. E. 1948b. A Mathematical Theory of Communication. The Bell System Technical Journal. 28:623–656.

Shannon, C. E. 1951. Prediction and entropy of printed English. The Bell System Technical Journal. 30:50–64.

Young, T.Y. - Fu, K.-S. 1986. Handbook of Pattern Recognition and Image Processing. Academic Press, Inc., London.