# Educational Applications of Entropy and Fractal Structure Analysis in Impact Assessment of Environmental Pollutions

V. Kozma-Bognár* and J. Berke**
* University of Pannonia Georgikon Faculty, Keszthely, Hungary
** SFD Information Technology Ltd., Keszthely, Hungary
kbv@georgikon.hu , berke64@gmail.com

***Abstract*** **— The central tasks of higher education is to improve the range of theoretical knowledge of the students and to coach to find the effective answer to the problems in the practice. Nowadays in the institutions of higher learning beyond the practice-oriented trainings is promoted to take the students in the scientific researches. With this end in view the students in the academic level need study materials, which provides a suitable basis for the implementation of the research. Last years there was more projects at the Department of Meteorology and Water Management (University of Pannonia, Georgikon Faculty) in which have taken part also PhD candidate, PhD., MSc. and BSc. students and beside the researchers and tutors. In this situation the students can get experiences in the accurate and precise solutions for individual tasks and opportunities for effective collaboration with the team. From 2010 have been carried out researches of plant-soil-atmosphere system. One of the main objectives is to study and detect the effects of environmental pollution. Prallel with field measurements we took aerial images with the purpose of testing and studying the potentials of different spectral 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. In this paper are presented the results of the impact assessment in the topic of the particulate pollution from road traffic and the use of this research results in the education.**

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

## I. Introduction

Nowadays remote sensing has become a widely used technology collecting information about our environment. 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 more and more 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 [13] 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. LANDSAT, SPOT, Quickbird, 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, defense and remediation activities etc. [16]. 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, red mud detection, etc. [4, 5, 10]. The technical development of the sensors is followed by the significant delay of the processing methods and applications [14, 20]. Therefore it is reasonable to develop new methods, in order to get more useful information from measured data.

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 and cadmium originated from road traffic on maize growth and development were investigated under field experiment at the Meteorology and Water Department. 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.

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. In this project PhD. candidate, PhD., MSc. and BSc. students have taken part beside the researchers and tutors.

## II. Materials and Methods

### A. 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 center of the test area N: 46°44'08.55", E: 17°14'19.76", H:114 m) where six plots were established (Figure 1) [8, 9]:

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. During the research we used chemically “pure” black carbon and cadmium nitrate. We sprayed black carbon 3 $gm^{-2}$/week and cadmium $10^{-5}$M/week doses onto the leaf surface to see the effect of the growth of plants.

Figure 1. Six plots of 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 [1].

### B. Research tools

During three growing seasons (2010, 2011, 2012 and 2013) was collected remote sensing data in different spectral ranges parallel with the field measurements (Figure 2).


Figure 2. Different spectral ranges of aerial image (13/09/2013): Top - Visible, Middle - Near Infrared, Bottom - Far Infrared

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)

| Date of the aerial photography | | | |
|---|---|---|---|
| Growing season in 2010 | Growing season in 2011 | Growing season in 2012 | Growing season in 2013 |
| 25/04/2010 | 04/06/2011 | 02/04/2012 | 24/04/2013 |
| 27/06/2010 | 20/06/2011 | 29/05/2012 | 19/05/2013 |
| 21/07/2010 | 04/07/2011 | 10/06/2012 | 23/06/2013 |
| 09/08/2010 | 17/07/2011 | 21/06/2012 | 12/07/2013 |
| 17/08/2010 | 01/08/2011 | 23/07/2012 | 25/07/2013 |
| 26/08/2010 | 22/08/2011 | 24/08/2012 | 16/08/2013 |
| 09/09/2010 | 03/09/2011 | 07/09/2012 | 13/09/2013 |

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 [9]. 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 /m/ | 400 | 400 | 400 |
| Spectral band /nm/ | 400 - 700 | 720 - 1150 | 8000 - 14000 |
| Geometrical resolution /m/ | 0.1 | 0.1 | 0.3 |
| 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.

## III. RESEARCH METHODS

### A. Entropy

Information-theoretic concept of entropy is used today in 1948, Claude E. Shannon [17, 18] respectively, and illustrated through a practical example [19], 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:

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

(..)

where

H - the information-theoretic entropy

$p_i$ - the ith message occurrence probability

General definition of entropy mathematical sense by Alfred Rényi [15] that the

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

where

$$\propto \geq 0 \text{ and } \propto \neq 1$$

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

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

$$0 \leq H \leq \log_2 n \tag{3}$$

where n is the number of possible messages.

The entropy is a minimum ($H_{min}=0$) if the source is still sending the same message.

The entropy is taken to be the maximum value $H_{max}=\log_2 n$, if all the messages with equal probability ($p_i = -\log_2 n$).

### B. Spectral Fractal Structure

The SFD is a general fractional dimension derived from the structure of the processing [12], 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 [2, 3] 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 [3]:

$$SFD_{measured} = \frac{n \times \sum_{j=1}^{S-1} \frac{\log(BM_j)}{\log(BT_j)}}{S-1} \tag{4}$$

Where,

n – the number of image layers or image channels

S – the spectral resolution (in bits)

BMj - valued spectral pixel boxes containing j number of bits

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

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

$$BT_j = (2^S)^n \tag{5}$$

The SFD formula metric [3], the evaluations of both hyper- and multispectral images can be used for exact measurements.

## IV. RESULTS

The 2010th, 2011th 2012th and 2013th 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 analyzed 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
- treatment types
- temporal variations in the vegetation period
- changes in the yearly analyses
- geometric resolution
- water supply.

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).

| Parameters | Treatments | Date of aerial photography in 2013 | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 24/04 | 19/05 | 23/06 | 12/07 | 25/07 | 16/08 | 13/09 |
| Entropy values | Black carbon | 13.2877 | 13.2877 | 13.2877 | 13.5626 | 12.9758 | 13.2877 | 13.2877 |
| | Control | 13.2877 | 13.2877 | 13.2877 | 13.5627 | 12.9817 | 13.2877 | 13.2877 |
| | Cadmium | 13.2877 | 13.2877 | 13.2876 | 13.5627 | 12.9824 | 13.2877 | 13.2877 |
| SFD values | Black carbon | 1.0363 | 1.2098 | 1.3394 | 1.2840 | 1.3082 | 1.3241 | 1.3121 |
| | Control | 1.0285 | 1.1825 | 1.3239 | 1.2741 | 1.2971 | 1.3257 | 1.2958 |
| | Cadmium | 1.0375 | 1.1731 | 1.3227 | 1.2811 | 1.3057 | 1.3237 | 1.2849 |

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 [9, 10, 11].

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 [6]. Was considered optimal for spectral bands used to select and validate the fractal dimension (SFD) values. Reviews of the images containing the tested crop areas were analyzed, 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 (Figure 3), as further investigations of these images gave the most reliable results (hit accuracy as well as other measured parameters and correlation studies).

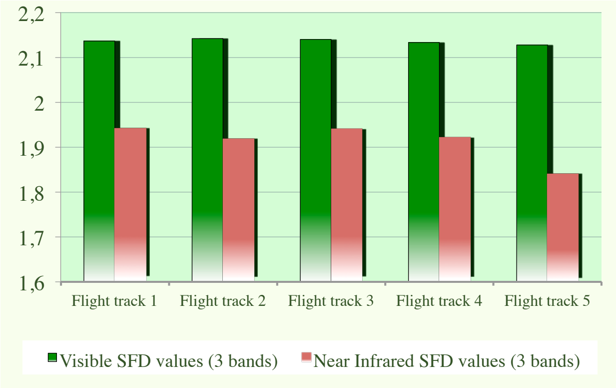


Figure 3. Maximum SFD values of the multispectral images

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 [7].

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 (Figure 4). The fourth image can be seen from the offset correction about the structural value of the original image data.

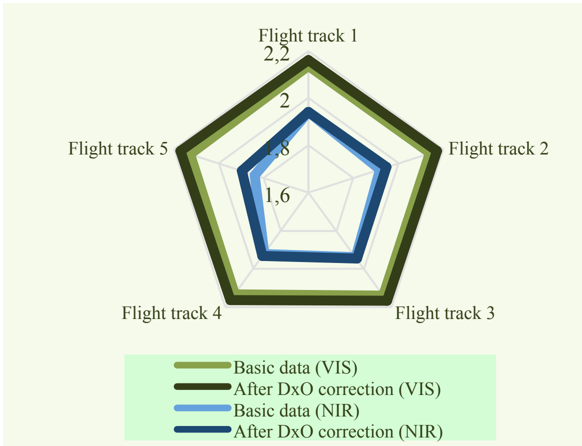


Figure 4. The multispectral images SFD values before and after correction, flight per track

The results of this project were included in a curriculum too. The Power Point presentation related to the topic of remote sensing application and was prepared in Hungarian. The presentation has nearly 100 slides and four themes:

I. Basic concept and process of remote sensing
II. Remote sensing images and sensors
III. Image processing of remote sensing data
IV. Case study

In the part of the case study are the main results of the remote sensing data analyses (Figure 5).

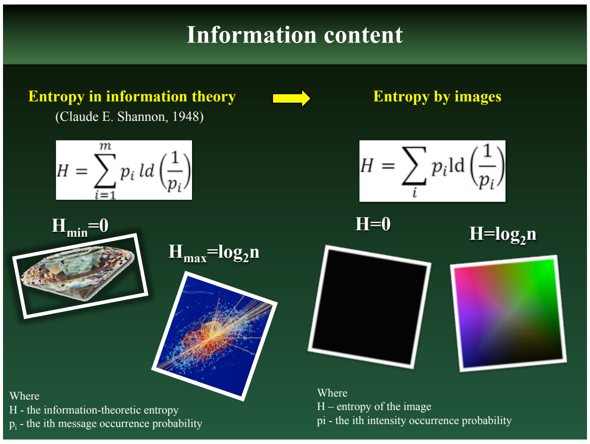


Figure 5. The slide of the information content (entropy) in the PowerPoint presentation

This document and some other information of the projects can we find on the website of the author: http://www.digkep.hu/remoteimages/Kozma-Bognar_Veronika/Kozma-Bognar_Veronika.html

## V. CONCLUSION

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 (Table 7).

| Type of Investigation | Using Entropy | Using SFD |
|---|---|---|
| Spectral range | partially /under appropriate conditions/ | Yes |
| Treatments | No | Yes |
| Within the growing season | Yes | Yes |
| According to analyzes by year | partially /under appropriate conditions/ | Yes |
| Geometric resolution | Yes | Yes |
| Irrigation | partially /under appropriate conditions/ | Yes |

Table 7. Utility of the parameters of spectral fractal dimension and entropy in investigations

The research results in many areas have been directly applied in education. Represented by the authors related disciplines (waste management, pest control, remote sensing, applied IT) BSc, MSc and PhD courses already been introduced. Experts, engineers and research workers in addition to the areas of database information can be used to solve the actual results achieved during daily tasks.


## *ACKNOWLEDGMENT

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.


## REFERENCES


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

[2] 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.

[3] 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.

[4] Berke, J., Kozma-Bognár, V., Nagy, T., Kováts, L.D., Tomor, T., 2013. Applications of termovision in environmental protection and Agriculture. 18th International THERMO Conference. pp. 15-22. Budapest, 2013. július 02-05.

[5] 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.

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

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

[8] 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.

[9] 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.

[10] Kozma-Bognár, V., Berke, J. 2013. Entropy and fractal structure based analysis in impact assessement of black carbon pollutions, Georgikon for Agriculture. 17: (2). pp. 53-68. ISSN: 0239-1260.

[11] Kozma-Bognar V, Berke J, Martin G. 2013. Remote sensing applications in evaluation of cadmium pollution effects In: Abstracts volumes of the European Geosciences Union General Assembly 2013. p. 15.Vienna, Austria, April 7-12 2013.

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

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

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

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

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

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

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

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

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