DOI: 10.2478/mgrsd-2006-0035

MISCELLANEA GEOGRAPHICA
Vol. 12 2006 pp. 305-313

Karol Kącki
Chair of Geoinformatic and Remote Sensing
E-mail: karolkacki@wp.pl

THE APPLICATION OF DECISION BINARY TREES TO ASSESS THE USEFULNESS OF THE DIGITAL TERRAIN MODEL IN STUDYING THE RELATIONSHIPS BETWEEN RELIEF AND VEGETATION IN THE POLISH HIGH TATRA

Abstract: The relationships between individual components of the natural environment have long been an object of research (Kostrowicki, Wójcik, 1972; Rączkowska, Kozłowska, 1994; Kozłowska, Rączkowska, 1996).
This paper is an attempt to analyse the relationships between two geocomponents of the natural environment: relief and vegetation, from a perspective contrary to the one currently prevailing in the literature of the subject. This approach assumes that relief, with its dominant role as a component strongly affecting the formation of the remaining factors, can be indicative in character and as such can represent basic factors that help determine and anticipate the occurrences of certain plant communities as well as locations with no vegetation.
Using geoinformation data along with the tools to process them, an attempt was made to assess the usefulness of the DTM (Digital Terrain Model) to identify selected plant communities, rock and water.
The development of a model of the relationships between the relief and the vegetation is an attempt to capture the correspondence between the parameters characterising the relief, calculated using the DTM model and classes of objects, with the use of information obtained from an Ikonos XS image. This model was subsequently used to draw a map of the land cover for a part of the Gąsienicowa Valley in the High Tatra (*Dolina Gąsienicowa*). For the purpose of this exercise, a technique of data classification called DBT (Decision Binary Trees) was used.

Key words: relief – vegetation relationship, Decision Binary Trees (DBT), Digital Terrain Model (DTM), Ikonos XS, image classification, geoinformation.

CONSTRUCTION AND ANALYSIS OF THE DIGITAL TERRAIN MODEL

The Digital Terrain Model (DTM) is a digital, discrete representation of the land surface that makes it possible to reproduce the surface (its shape) for a given area and to interpolate the elevation values for any point with a location identified by its x and y coordinates.

The cartographic technology was used to develop the DTM for a part of the High Tatra range, and included vectorisation and editing of contour lines. In most cases, this involves vectorisation of topographic maps or contour base maps on the computer screen. Next, the hypsometric layers are allocated relevant elevation attributes. The advantages of this method include its simplicity and the fact that it does not require sophisticated and expensive software.

When constructing the DTM, we used data based on lines, points and polygons. The interpolation of elevation data was completed using the ARC INFO 7.1.2 software. With this software, it is possible to conduct a preliminary interpolation using the triangulated irregular network method, and then resample the surface into a regular network of squares. For this study, the selected grid size was 4 metres, which corresponds to the ground resolution of the Ikonos XS satellite imagery.

Terrain analysis involves the processing and graphic simulation of elevation data, usually recorded in the form of a digital matrix using a raster structure. Each raster cell contains information about elevation (z value), which is used to calculate the value of the parameter characterising the land surface for any x, y variable.

In this work, the Digital Terrain Model was applied to create layers to describe the morphometry of a given land structure: slope gradient, aspect, curvature of slopes (map of concave and convex relief forms), SAI (*Slope – Aspect – Index*) value and the terrain roughness (altitude difference). All maps were prepared in raster form.

COMPILATION OF THE VEGETATION DATA

The satellite imagery used in the study was taken by the Ikonos satellite on 22 August 2000. This is a multi-spectral image, recorded in four spectral bands: 0.45 – 0.52 μm (blue), 0.51 – 0.60 μm (green), 0.63 – 0.70 μm (red) and 0.76 – 0.85 μm (near infrared).

High ground resolution (4 metres per pixel) and 11-bit radiometric resolution (2,048 shades of grey) render Ikonos in a similar category to other state-of-the-art satellites, such as Quick Bird, which can identify objects as small as several dozen centimetres.

Owing to the methodical nature of the research and the resultant experimental approach to the issues discussed in this paper, we decided to distinguish only four classes in the satellite images: ponds, stands of dwarf mountain pine (*Pinus montana*), Alpine grass, and bare rock and scree.

These classes were identified using the supervised classification of satellite images, which is an allocation of pixels to a finite, predefined number of classes based on training fields indicated by the user. This assumes that the user has prior knowledge of the data and the researched area, plus the ability to identify objects based on spatial and spectral information.

There are three basic methods of supervised classification: the parallelpiped method, the minimum distance method and the maximum likelihood method. All of these methods involve comparison of the radiation reflection values for the different ranges of each pixel with a predefined pattern, i.e. the training sample.

In this study, the supervised classification was carried out using the maximum likelihood method, which is considered to be the most effective method of image classification, and hence is the most popular method, used in many software applications. It is based on an analysis of the probability distribution of the attribution of a given pixel to a given class. Including the dispersion parameters for random variables means that for classes which are inherently characterised by a high dispersion of spectral reflection as compared to the mean value, further spectral distances will be more acceptable than for uniform classes. Classification was conducted using the ERDAS IMAGINE 8.3.1 software. The total accuracy value of the classification was 90.4%, with a Kappa coefficient of 0.87 for the proportional reduction of errors occurring during the classification process as compared to the random classification error. The highest accuracy was reached when classifying water (100%) and bare rock and scree (91.5%), while the lowest for Alpine grass (84.7%), with the accuracy for stands of dwarf mountain pine at 88.4%.

MODELLING THE RELATIONSHIP BETWEEN DTM-COMPUTED PARAMETERS AND CLASSES DISTINGUISHED IN THE IKONOS XS SATELLITE IMAGE

A two-stage analysis was used to identify the relationship between the parameters calculated using the DTM model and the classes identified in satellite imaging. First the correlation coefficient and covariance values were computed in order to obtain information about the relief and the classes distinguished in the satellite image. Next, a relief–vegetation relationship model was constructed for a part of the Gąsienicowa Valley.

Covariance is a measure of tendency, of mutual changes in the values of the same pixel found in two different files (data ranges). It is an average product of the differences in relevant values in two ranges from the corresponding average. Similar to variance, it is expressed in squared units.

The correlation coefficient is a quotient of the covariance and the product of standard variations of pixels from two data ranges. The value of the coefficient remains within the limits of -1 to 1; where 0 means there is no relationship between the categorised classes and the parameters computed using the DTM model, while a value close to 1 or -1 indicates a very strong relationship.

The correlation coefficient and covariance values were calculated between the parameters computed using the DTM (altitude in metres above sea-level, slope gradient, aspect, curvature of slopes, SAI and terrain roughness calculated in windows of 50, 75, 100 and 125 pixels) and the classes distinguished in the satellite image (ponds, stands of dwarf mountain pine, Alpine grass, bare rock and scree, see Table 1). Both were computed using the S-PLUS software.

Table 1.

Correlation coefficient and covariance between DTM-computed parameters and classes distinguished in the Ikonos XS satellite image.

	Classes distinguished in the satellite image								
	Altitude above sea-level (m)	**Slope gradient**	**Aspect**	**Curvature**	**SAI**	**Roughness 50**	**Roughness 75**	**Roughness 100**	**Roughness 125**
Correlation coefficient	0.63	0.49	-0.45	-0.08	-0.32	0.28	0.33	0.40	0.47
Covariance	27.5	5.2	-0.42	-0.07	0.01	3.31	4.75	6.53	8.12

Constructing the relationship model for the Gąsienicowa Valley area was based on the Decision Binary Tree (DBT) method, which allows classifying data on the basis of a hierarchical division of the input data into increasingly homogenous subsets (*S-PLUS 2000 Guide to Statistical and Mathematical Analysis*, 1999) to identify certain classes of objects as predefined by the user.

This results in a binary tree, which is a graphic representation of the relationships between the analysed data, with a distinct differentiation of two classes in its upper part and more detailed differentiations of the remaining classes in the lower part. Using this method makes it possible to assess the relevance of a given input channel (e.g. DTM-computed parameter), to discriminate between the specific classes of objects.

In the analysis, the DTM parameters used had the highest correlation coefficient and covariance values, as compared to the classes distinguished in the satellite image. These included: altitude in metres above sea level, slope gradient, slope aspect and terrain roughness computed for a window sized at 125×125 pixels. The data describing the land relief (independent variables) were used as variables determining the occurrence of the classes distinguished in the satellite image. These dependent variables were: stands of dwarf mountain pine, Alpine grass, bare rock and scree. As the slope gradient and aspect parameter allows the class of ponds to be easily distinguished (for aspect they are ascribed the value of 361, with 0 as the value of gradient), it was not necessary to take this class into account in the process of computing binary trees.

In an analysis of the existing relationships for a large research area, such as the entire Gąsienicowa Valley, the obtained result (binary tree) forms an elaborated graph (with many subordinate branches). The analysis of such a tree in such a case would be extremely difficult.

For the above reasons, the analysis was conducted using 1,000 randomly selected pixels. It should be noted that in such situations a certain imprecision of the captured relationships must be allowed for, with the obtained picture being "schematic" because only a proportion of the pixels from the studied area are taken into account.

Figures 1, 2, 3 and 4 show binary trees as a graphic representation of the relationships between the selected DTM parameters and the classes identified in the Ikonos image.

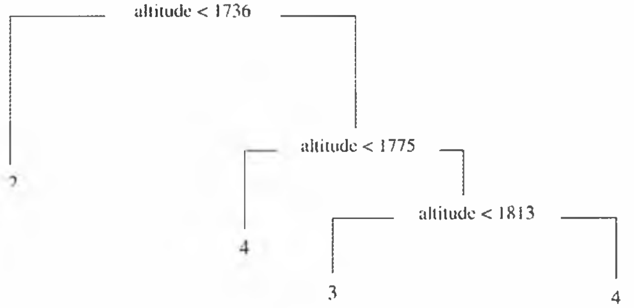

Fig. 1. Graphic representation of the relationship between elevation (altitude in metres above sea-level) and classes distinguished in the Ikonos XS satellite image, where: 2 – stands of dwarf mountain pine, 3 – Alpine grass, 4 – bare rock and scree, prepared using the S–PLUS software.

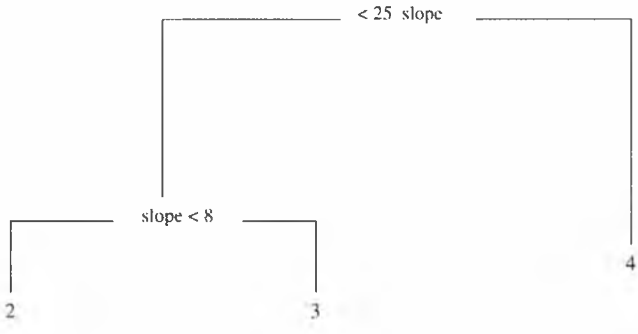

Fig. 2. Graphic representation of the relationship between slope gradient and classes distinguished in the Ikonos XS satellite image, where: 2 – stands of dwarf mountain pine, 3 – Alpine grass, 4 – bare rock and scree, prepared using the S–PLUS software

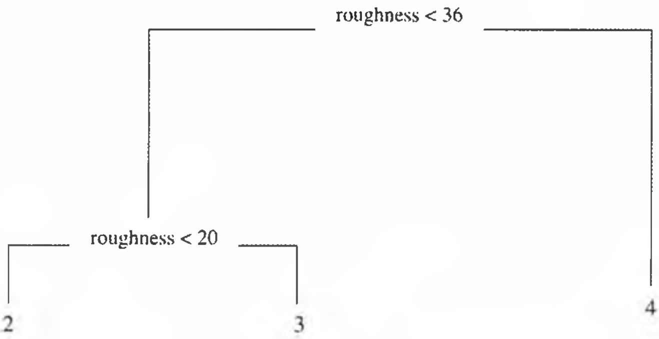

Fig. 3. Graphic representation of the relationship between terrain roughness (altitude difference) and classes distinguished in the Ikonos XS satellite image, where: 2 – stands of dwarf mountain pine, 3 – Alpine grass, 4 – bare rock and scree, prepared using the S–PLUS software

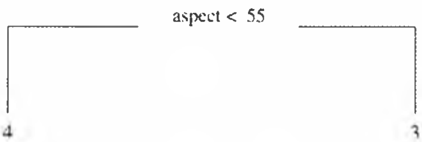

Fig. 4. Graphic representation of the relationship between slope aspect and classes distinguished in the Ikonos XS satellite image, where: 3 – Alpine grass, 4 – bare rock and scree, prepared using the S–PLUS software

Figures 1, 2, 3 and 4 indicate that such DTM parameters as altitude above sea level, slope gradient and terrain roughness helped to distinguish all the classes identified in the Ikonos image, whereas the slope aspect parameter proved useful only in distinguishing the Alpine grass class from that of bare rock and scree. Figure 1 shows that the elevation parameter makes it possible to distinguish the class of stands of dwarf mountain pine (2) for altitudes below 1736 metres a.s.l, while the altitudes between 1736 – 1775 metres and those above 1813 metres help to distinguish the class of bare rock and scree (4), with the class of Alpine grass (3) distinguishable for the altitudes between 1775 – 1813 metres. The slope gradient parameter (Fig. 2) made it possible to discriminate class 2 for the gradient values below 8 degrees and class 3 – for the values between 3 – 25 degrees, whereas the slope gradient higher than 25 degrees allowed to distinguish class 4. We used a similar interpretation process for the remaining binary trees. The altitude difference values below 20 metres made it possible to distinguish class 2 and the range between 20 – 36 metres – class 3, while the altitude difference in excess of 36 metres allowed us to distinguish class 4, as shown in Figure 3. The aspect values below 55, describing the northern and north-eastern slope aspect, made it possible to distinguish between class 3 and class 4, as shown in Figure 4.

The model of relationships, constructed using the Spatial Modeller tool of the ERDAS IMAGINE software, was used to create a map showing the following classes of objects: ponds, stands of dwarf mountain pine, Alpine grass, and bare rock and scree. To enhance the accuracy of the preliminary map, the model of relationships (the value ranges of individual DTM parameters) was improved by visually comparing the identified classes on the newly prepared map with the DTM parameters shown in the geographically correlated windows.

Based on the tested model of relationships, the following DTM parameters were obtained and which were then used to identify selected classes in the investigated fragment of the Gąsienicowa Valley:

- Ponds (1); slope gradient = 0, aspect = 361
- Stands of dwarf mountain pine (2); altitude<1736, slope gradient <10, roughness<26
- Alpine grass (3); 1774< altitude<1813, 9< slope gradient <21, 25< roughness<31, 54< aspect <305
- Bare rock and scree (4); 1735< altitude<1775 and altitude>1813, slope gradient >20, roughness>30, 55> aspect >304

Figure 5 below shows the part of the Gąsienicowa Valley for which the model of relationships between relief and vegetation was constructed and tested. The picture on the left shows a satellite image using a RGB composition (4, 2, 1) in true colours, the one in the middle shows the result of the image's supervised classification, and the one on the right shows a map of plant communities.

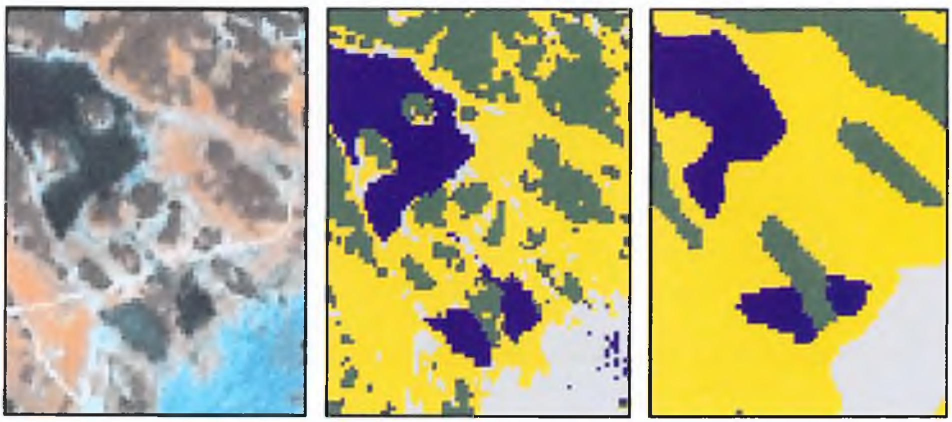

Fig. 5. Part of Hala Gąsienicowa near the Kurtkowiec Pond

Finally, the DTM-computed parameters and our improved model of relationships were used to prepare a map (Fig. 6) for the part of the Gąsienicowa Valley near the Kurtkowiec Pond. The total accuracy of the classification for the above objects in this map was 66%, while the Kappa coefficient was only 0.47. The highest accuracy of the classification was obtained for the Alpine grass class (78%) and ponds (75%), with significantly worse results for the class of bare rock and scree (62%) and stands of dwarf mountain pine (only 49%).

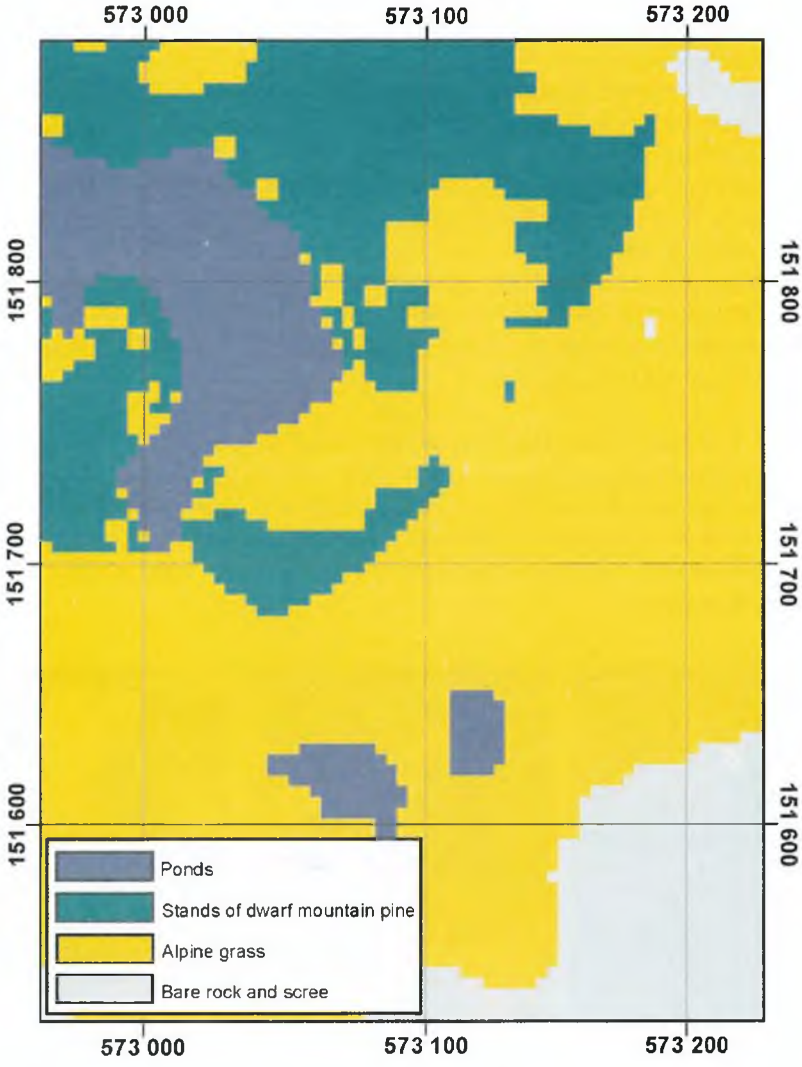

Fig. 6. Part of Hala Gąsienicowa near the Kurtkowiec Pond. The map was created on the basis of DTM-computed parameters and the model of relationships between these parameters and classes distinguished in the Ikonos XS satellite image

CONCLUSIONS

This paper attempts to assess the usefulness of the Digital Terrain Model and the Decision Binary Trees to foresee locations of selected plant communities, rock and water; making it possible to evaluate the relevance of the DTM for this type of research.

The 66% accuracy of the maps created for a part of the Gąsienicowa Valley proves that examination of the correspondences between relief and vegetation, using this solely computer-based method, allows these relationships to be captured. Although the interdependencies between the DTM-computed parameters and the classes discriminated in the Ikonos XS image are not very strong, they justify a positive opinion on the usefulness of the method as applied in the study of high-mountain vegetation. This is corroborated by a high degree of accuracy (78%) in the identification of Alpine grass in the Gąsienicowa Valley using the Digital Terrain Model.

A more extensive discussion of the results and a comprehensive analysis of this issue is presented in the author's MA dissertation entitled *Application of Geoinformation Methods in Studying the Impact of Relief in the Distribution of Vegetation in the Polish High Tatra Mountains*, written at the Remote Sensing of Environment Laboratory of the Faculty of Geography and Regional Studies, Warsaw University.

REFERENCES

Kostrowicki A. S., Wójcik Z., 1972, Podstawy teoretyczne i metodyczne oceny warunków przyrodniczych przy pomocy wskaźników roślinnych [Theoretical and Methodological Foundations for Assessing Natural Conditions Using Plant Indicators, in Polish], [in]: *Metody oceny warunków przyrodniczych produkcji rolniczej* [Methods of Assessing Natural Conditions for Agricultural Production], *Biuletyn KPZK PAN*, 71, pp. 3–63.

Kozłowska A. B., Rączkowska Z., 1996, The spatial relations of relief and vegetation, *Studia Geomorph. Carpatho-Balcanica* Vol. 30, pp. 117–128.

Miller C. L., Laflamme R. A., 1958, The Digital Terrain Model: Theory and Application, *Photogrammetric Engineering*, Vol. 25, pp. 433–442.

S-PLUS 2000 Guide to Statistical and Mathematical Analysis, 1999, Volume 1, Data Analysis Products Division, MathSoft Inc., Seattle, Washington.