

REGIONAL DEVELOPMENT IN THE AGE OF BIG DATA

JÓZSEF ATTILA JANKÓ^{1*}, GYÖRGY SZABÓ²

^{1,2} Budapest University of Technology and Economics, Hungary

*Email: janko.jozsef@epito.bme.hu

Received 20 November 2016, accepted in revised form 14 April 2017



Abstract

Our paper presents a forward looking analytical approach to the territorial development in a region of the Transylvanian Plain situated in the vicinity of Cluj-Napoca, Romania. We outlined the development of this region with the means of landscape architecture supported by a comparable assessment. In the age of Big Data we settled at creative usage of traditional analysis. We extracted yet undetected information from a limited amount of available as yet loosely related data. The key feature of the employed model is the ontological traceability of cause and effect. Although technology is available to collect enormous data, expert knowledge gained by education and professional practice cannot be overlooked. We demonstrate that this method of location based analysis is capable of delivering value added to established principles of spatial planning in the age of trustworthy, large volume, heterogeneous data.

Keywords: Spatial Planning, Location Based Analysis, Heterogeneous Data, Thematic Overlay

1. Introduction

The profession of landscape architecture gained not only substantial technological tools over the 20th century but a pressure from the IT and GIS sectors as well. In the age of Big Data, not all locations abound with current and thematically diversified data. We also identified that scientific and technical derivatives of the advanced technologies (deep learning e.g.) are constrained compared to the employment of functional models. Therefore, we developed a comparable assessment method to transform obtained and collected data into a form of working knowledge (Jankó & Szabó 2013). In this paper, we only highlight some of the techniques where spatial data and heterogeneous statistical data were employed. The aim of the comparable settlement diagnosis is to identify cause and effect relationships in such a way that assessment results can be traced back

to the characteristics of the phenomena examined. If change in assessment results becomes desired, the necessary change in the determining phenomenon can be pointed out. This method can assume the deficiency of available data, while at the same time being capable of including large volume data of heterogeneous type from different sources. This model stands in contrast to an artificial intelligence like neural network analysis or probabilistic approaches where the connections between the input data and descriptive results are chaotic or even hidden. Our approach is best described by the term 'inverse fuzzy logic', where phenomena of the region in the form of pseudo fuzzy sets are transformed into crisp values on a continuous scale. Ultimately, the defuzzification (Skalna et al. 2015) of the assessment outcomes enables design and planning proposals to be based on a controlled interpretation of the landscape utilising the strengths

and addressing weakness of the site while considering opportunities and threats. This method differs from landscape character assessment (Tudor 2014) in the term of character because we also take into account remotely imperceptible factors. Examples of this distinction will be put forward in the later discussion.

A better understating of spatial connections and interactions within the area of interest (the subject of design and planning) can be attained from using thematic overlays represented on maps drawing on statistical data. It is not the same case in our work environment where only raw and weakly connected data is present to incorporate into the design principles. Data availability is limited, and documentation suggests that data reliability might be in question (Advameg 2015). While physical features of a landscape are straightforward to visualise and analyse on thematic layers, maps yet contain only limited and purposefully selected elements from the reality. Our approach complements this practice with such thematic data analysis that connects information from diverse sources in order to construct an elaborated general overview of the current state of the design area. Between 2010 and 2014 we engaged in an extensive data collection effort to overcome these challenges. The team evaluated historical maps, General Urban Plans and current remote sensing sources such as CORINE Land Cover of the European Environment Agency (EEA) and SRTM Digital Elevation Data of CGIAR Consortium for Spatial Information (CGIAR-CSI). The team gathered an abundance of socio-economic information in a variety of forms including statistical data, site visits and interviews with local elders who possess a lifetimes of yet unpublished knowledge. We established a framework for spatial assessment capable of utilizing both historic and current data regardless of its collection techniques. This method provides systematized information of the area of interest.

The extracted information provided a significant base for our planning process

was divided into four phenomena: the natursphere, the sociosphere, the opussphere and the urbanosphere (Kiss 1996). This four-fold thematic approach contains the state of art in our understanding of the environment, the society, the economy and the built heritage respectively (Jankó & Szabó 2013). The data was aggregated within each local authority, scaled and normalized, providing a comparable relative rank (1-9) of the local authorities in each sphere. The numeric class represents weak, medium or strong quality of the components of each sphere. Strong characteristics are considered to be relatively well developed. The weak characteristics provided the areas that could be addressed by planning proposals. The aggregated indices or ranks are the result of base indices which are the result of detailed indices directly measuring certain phenomenon identified in the region. The goal of the regional development process is to reduce inequalities recognisable from the ranks. Examples of tracing back cause and effects will be provided in later chapter. We also determined opportunities and threats that challenge the achievement of the design targets. Relevant geographical information was aggregated in 1 km² hexagon cells to identify locations with high risks or remarkable potential for the development.

A professional pursuit like landscape architecture is performed within a discipline bounded by interdisciplinary values concerning the development of the built and cultivated environment regarding the beneficiary inhabitants. However, each site requires a unique hierarchy of prioritization. The phenomena within this landscape were evaluated with the real needs of the local society in mind while not passing over ecological and economic considerations. The comparable assessment intended to answer the following types of questions. How much does the current land use support the society? What kind of change may lead to more dependable self-sustaining wellbeing? What sort of intervention could cause the desired change? Where the interventions

could be realized? Institutional and financial means will not be discussed.

2. Methods and Workflow

The assessment and spatial planning proposal concerned ten local authorities (Aiton/Ajton, Așchileu/Esküllő, Borșa/Kolozsborsa, Căianu/Magyarkályán, Cămărașu/Pusztakamarás, Jucu/Zsuk, Mociu/Mocs, Pălăta/Magyarpalatka, Suatu/Magyarszovát, Vultureni/Borsaújfalú) encompassing 49 villages spread over a 600 km² region. The administrative centres listed in alphabetic order are in Romanian (ANCP1 2014) and Hungarian (Szabó & Szabó 2003) respectively. The workflow consisted of three major phases: 1) Investigation of the natural and societal facilities, landscape shaping factors and applicable development strategies; 2) Landscape assessment based on these research findings; 3) Proposals for the enhancement of the quality of life based on this assessment.

Sunshine availability, visual accessibility

Recursive solar irradiation analysis (GRASS function r.sun with sequential moments) was performed to segment and classify the site to gain evidence for agriculture related proposals. Those areas with high irradiation on visually less exposed declivities and located in proximity to settlements provided sites for possible greenhouse culture. Irradiation is the process by which the land is exposed to radiation. We use the term ‘visual accessibility’ to describe the phenomenon in which the land is exposed to visual contact during daily activities.

The visual accessibility analysis determined those areas of potential land use conflicts. Line-of-sight raster analysis (GRASS function r.los with default settings) was executed in nested loops over each digital elevation model (DEM) pixel. Area of viewing positions (number of pixels) having visual contact with target locations in each loop were collected and aggregated over the 1 km² cell size tessellated hexagonal grid

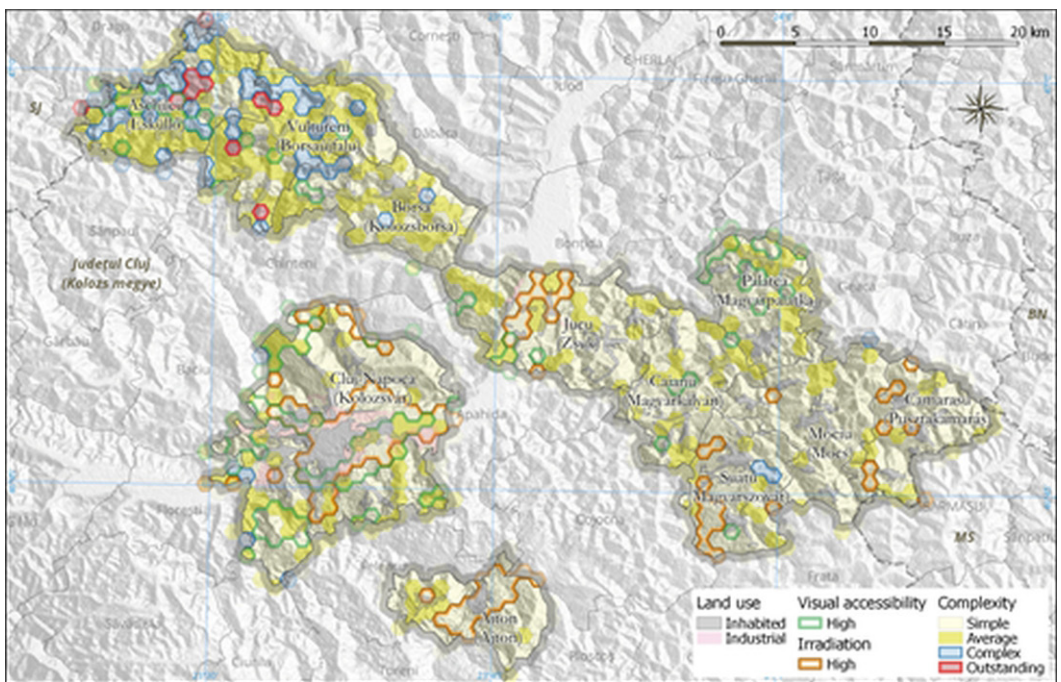


Fig. 1: Recreational complexity classes with area of high visual accessibility and irradiation

utilised for the rest of spatial classification and comparison. Spatial units visible from larger area were considered visually more susceptible on our visual accessibility scale. The algorithmic base of a linear computation along roads was published in the work of Miorelli & Zatelli (2006). Our general-purpose territorial approach would benefit from access to a current digital surface model (DSM) which includes artificial visual obstacles. The significance of visibility (thus the significance of viewpoint) could have been weighted by current land cover/land use classification but the scale

of our assessment did not require such a refinement.

Complexity assessment

We evaluated the recreational potential in the region adapting the Kiemstedt formula (Fuchs 2010), modified to suit our classification needs. We did not consider climatic data, because its constant value did not make difference at this scale. North-east areas with proportional forest edges and frequent watercourses produced complex but also outstanding potential for recreation. The Suatu Lake in eastern south-east region

1. Table. Natursphere aggregated indices

			Civil parish									
			Aiton	Așchileu	Borșa	Căianu	Cămărașu	Jucu	Mociu	Pălatca	Suatu	Vultureni
SIRUTA			55598	55776	56265	56354	56522	58259	58623	58794	59657	60099
Cl	Ref.	Description of base index										
D	N1	Average recreational complexity	1	9	3	3	1	1	1	2	1	7
D	N2	Proportion of area of outstanding recr. complexity	1	9	1	1	1	1	1	1	1	5
I	N3	Proportion of inhabit area over admin. area	4	7	9	8	6	7	9	9	1	8
I	N4	Environmental impact of airport in Cluj	9	9	9	1	9	7	7	9	6	9
I	N5	Environm. impact of overhead electricity lines	5	5	9	3	9	1	9	9	1	9
I	N6	Environmental impact of main roads	9	9	9	3	1	5	1	9	9	9
D	N7	Proportion of forests (CLC 31)	1	9	2	4	1	1	2	4	2	9
D	N8	Prop. of wetlands (41) outside artificial surf. (11)	1	1	1	1	9	1	2	1	1	5
D	N9	Proportion of water bodies (CLC 51)	1	1	1	1	1	9	1	1	7	1
D	N10	Specific length of water courses	1	9	6	3	1	2	1	2	1	8
D	N11	Specific number of knoll weighted with elevation	9	2	1	3	6	4	3	1	5	1
D	N12	Proportion of conservation area	1	1	1	1	1	1	1	1	9	1
D	N13	Proportion of Natura2000 area	2	1	9	1	1	2	1	1	2	7
I	N14	Rate of land use conflicts	9	1	9	5	9	1	9	9	5	9
D	N15	Natural resources of touristic interest	3	1	1	1	1	3	1	1	9	3
Natursphere aggregated indices			4	5	5	3	4	3	3	4	4	6

Note: Class (Cl.) D - directly proportional; I - inversely proportional

(Fig. 1.) is significant because of the primarily agricultural areas in the eastern parishes. We proposed introducing signed hiking paths and the developing a guesthouse network as part of a complex tourism destination management programme. The lake is used for fishing without any proper infrastructure for such an activity. In the meantime, an insufficient communal waste collection scheme resulted in unregulated dumping on the shore adjacent to the children and the pets of adult anglers. The identification of this functional conflict recommended that development proposals consider reviewing communal waste collection scheme effectiveness to have the popular fishing tradition revived (Fig. 3a.).

Assessment of the natural environment

This paper presents with excerpts only one part of the four phenomena, the diagnosis of the natural environment. Aggregated indices derived from base indices part of our comparative settlement assessment (Table 1). Some base index will be explained

to give better understanding of the core calculation principles. The detailed index is generally a specific value that characterise the region. Some of these values are directly proportional with the rank (1-9), some of them are inversely proportional based on expert decision. For example natural aspect of an area less available where artificial elements are dominating, thus Proportion of inhabit area over administrative area (N3) is inversely proportional base index while Proportion of forest area (N7) is considered a directly proportional base index.

Average recreational complexity (N1) is computed from the results of 2.2 Complexity assessment above. Complexity assessment was computed on each 1 km² hexagon cell providing base data. The base data were aggregated over each administrative are (civil parishes) providing an average value. The base index is gained after scaling and normalizing these average values. Proportion of area of outstanding recreational complexity (N2) provides a deeper understanding of how significant the recreational complexity

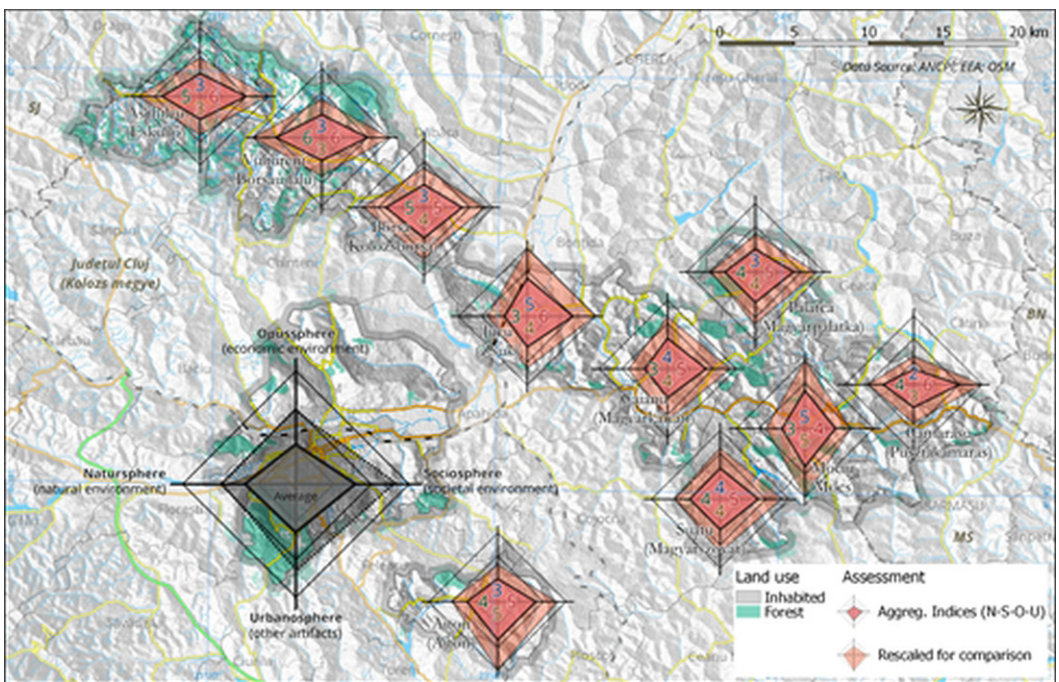


Fig. 2. Comparative settlement diagnosis (naturesphere, sociosphere, opusphere, urbano-sphere) with aggregated indices

in each civil parish. This computation used location based information from remote sensing services by applying traditional GIS methods.

Comparative settlement assessment

The study evaluated eighty characteristics of the design area in addition to the outcomes of the assessments mentioned above in four-fold comparative settlement assessment. The comparable indices suggest that the natural and economic environments are moderately developed, proposal should focus on nature related economically viable solutions.

Kertész et al. (2014) that provided an ecological assessment on the region of research comparing land cover classification patches with their calculated ecological potential classification patches. This process found that in 2006 46% of the land of their research area was used according to the land's ecological potential. Central to our efforts is to measure how the real needs of the society are supported by the capabilities of the land's natural, cultivated and artificial capabilities. This comparison is made by summing up the quantitative ratio of the phenomena in the same ontological domain (1).

$$I_n = \sum_{a=1}^b \frac{N_a p_a}{\sum_{x=1}^y C_{ax} w_{ax}}; \quad (1)$$

where the I index of sufficiency is the ratio of N (p portion of) phenomenon that needs phenomenon C (with w weight) by natural, civil or common sense law. Presumed relevancy determine portion and weight in each case. The results are aggregated, scaled and normalised (Fig. 2.). For example, the area of crop land is measured against number of mills which in turn is measured against the number of bakeries and these are measured against the number of inhabitants, bearing in mind the portion of bread for export.

3. Results

Aşchileu and Borşa both ranked 5 in the assessment of its natural environment with different sets of base indices. For example, it can be traced back that the Rate of land use conflicts (N14) base index is weak for Aşchileu (compared to Borşa). Tracing back the base data (not discussed in this paper) can suggest whether this rank could be improved. As a result of this comparison, design proposals could be formulated about where and how the natural environment would achieve a more desirable state.

The past is apparently present on this rural landscape. Abandoned workshops and machineries out of blast indicate the agricultural activity of the previous generation. Such result of an assessment can hardly be gained without human perception and common sense. The fact that this visual experience stands for decades implies the comprehension of stagnant socioeconomic conditions. The Romanian Industrial Heritage Route provides access to those artefacts considered to carry great importance. Crop growing and processing is vital to the local economy but some of the agricultural and industrial facilities look unattended. Natural spectacles like river banks and shores were generally neglected, tourist infrastructure was absent even though there are plenty of heritage and isolated civil initiatives by individuals who are solicitous about their homeland. Traditional research methods reveal national, regional and local initiatives and support the establishment of partners of cooperation in development.

New hiking routes were proposed to avoid significant habitats and wetland. The site contains a portion of the Mary's Way walking pilgrimage. We proposed some change to this pilgrims' way based on our visual accessibility analysis to improve visual access throughout the plain from the ridges. The village of Juc-Herghele (Zsukiménes) carries in its name ('Stud of Juc') the historic tradition of animal husbandry whose breeding studs provided stock for both beasts of burden and hunting

horses. The study proposed revitalizing the historical practice and the development of new horse-back riding trails to connect the spectacular destinations (Fig. 3b.). We proposed reforestation at several locations to generate a better microclimate on arable land. Past agricultural practices resulted in overfertilized soil. The increasing demand for organic farming products requires spatially and technically appropriate land. While we could propose locations, we realized that not all obstacles to organic farming can be eliminated by straight-forward numeric computations. Nevertheless, professional and governmental cooperation will be necessary if organic farming is to succeed. Herb collection points are proposed in villages surrounded by rich meadows. Orchards are proposed at accessible places in terms both of logistics and plant demands. A register of fruit-trees became vital in recreating a diversity of genus where current plant material is identified similarly to the national Register of Vine Varieties established in the European Union member states (Dir. 68/193/EEC n.d.). Fruit and wild-berry processing facilities can once again produce area-specific gastronomic delicacies.

4. Discussion

The study focused primarily on the geographically appropriate development of cultivated land, but we also evaluated the civic environment. The village Chesău (Mezőkeszű) developed distinct settlement pattern not present in the wider geographical region. The unique spatial character of forecourt farming qualifies it as primary example of cultural heritage worthy of preservation. For over a century the regional style of house was built directly onto the street sometimes featuring a forecourt rather than the domestic structure. While, it is not unique to place a free standing house behind a large forecourt, in Chesău the arrangement is both functional and sanitary. The streets are generally covered with excrement due to the daily passing of cattle. The forecourts contain the cattleshed, stable, pigpen, hen house and muck-heap, typically hidden from the street by the stunning family home in other communities. Chesău maintains the forecourt farming tradition by separating family life (the house and the well) from the dirty public place. It seems contradictory to assess chicken coups on a regional scale while the comparable assessment resulted in making this distinctive attention to the

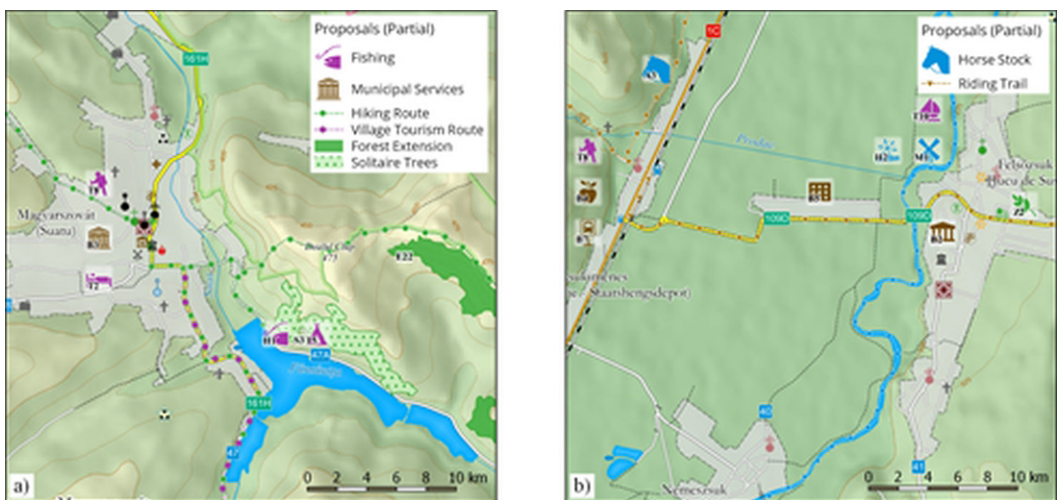


Fig. 3. Design proposals in civil parishes a) Suatu and b) Jucu

conscious arrangement of this village.

Based on our extensive site visit, we found several issues relating to the transportation infrastructure. For example, roads are classified by the level of the authority responsible for maintenance. Roads belonging to the European transportation network are kept in good condition (from the point of view of the drivers) by national authorities. Some roads are maintained by county authorities, others by local authorities. There is an obvious correlation between the level of the responsible authority and the condition of the roads. Sidewalks are constructed in the villages only beside surfaced main roads. Most of the villages have cart-roads strewn with crushed stone. In dry weather, the stone pieces in the dry mud make it very difficult for the elderly population to walk. These individuals rely on the local roads to practice backyard farming and access the other households scattered throughout the village. The purely digital data collection of remote sensing techniques is inadequate to capture such scattered ownership and its human consequences. The household economies are kept alive by this aging population (mostly 60-70 age women), and some of them are in their eighties. Their personal difficulties remain imperceptible by remote sensing applications. The current farming practice, performed by this significant portion of the population, characterise this landscape. It seems contradictory to professionally assess landscape character while the inhabitants are only concerned with stone particles along their daily walk.

Preliminary public participation initiatives revealed another driving force that causes land cover change and hence landscape character change. 40-60 years ago, orchard owners grew fruit varieties of many species. Since then, their children have migrated to big cities at home and abroad, working in large factories. Fruit picking thus became less feasible for the aging orchard owners because family members had migrated away and hence there was a decreased in man-power in the rural region. The new owners of the

orchards took an interest in the land because of the granted state premium based on land area. Thus, the orchards became arable land requiring less labour-intensive cultivation, usually growing the most profitable crop regardless of ecological consequences. 10-15 years ago, large factories in the towns were closed due to planned privatization, part of a process of pulling out of global economic engagement, that had been unforeseen by then 20 year old countrymen living in a peripheral region. Moving home remained the survival strategy for those in their middle ages, approaching retirement or not having any qualification in the booming sectors, regardless of their lack of any proper knowledge of farming, strong social and business networks and without the previous rich orchards.

5. Conclusion and Outlook

Any data (quantitative or qualitative) is only a selected representation of reality perceived with some prior preconception. It is culturally determined how data is comprehended and what sort of knowledge is acquired in order to make meaningful decisions regarding development and conservation of resources and cultural heritage. Remote sensing techniques and big data collection are apparently helpful. But they map only certain aspects of a territory. Unsupervised classification of remote sensing images does not recognize the unique settlement patterns and significant architectural styles. Understanding the spatial and cultural significance is more complex than identifying land cover classes. We did not intend to predict future trends, but search for spatial planning solutions to plan the future using location-based information collected from many sources with varied methods within the limits of data accessibility in the Romanian regional development culture.

This method of comparative settlement diagnosis utilizing heterogeneous, constantly incrementing data served as a framework

in our technical evaluation. However, this method can also equip authorities to evaluate the impact of accomplishments sought by the design proposals themselves. We used our method in order to assist our landscape design process. But derivatives of this method are capable of supporting more specific or more general fields of development like tourism destination management, distribution logistics, public transportation, utility development, education, senior care or military operations for example. Additional base indices could be incorporated into the assessment when more base data becomes available. A new thinking about spatial design is underway, fuelled by big data.

6. References

- Advameg, I. (2015), Romania - Ethnic groups. Encyclopedia of the Nations. <http://www.nationsencyclopedia.com/Europe/Romania-ETHNIC-GROUPS.html>
- ANCPI (2014), Agenției Naționale de Cadastru și Publicitate Imobiliară. Website of the National Cadastral and Real Estate Agency. <http://ancpi.ro>
- Dir. 68/193/EEC (n.d.), Council Directive 68/193/EEC of 9 April 1968 on the marketing of material for the vegetative propagation of the vine. <http://eur-lex.europa.eu/eli/dir/1968/193/oj>
- Fuchs, C. (2010), Potentialbewertung, Nutzungskonflikte und Lösungsansätze. http://www.d-w-v.de/drupal/cms/sites/default/files/beitraege_upload/WSM_Facharbeit_Potentialbewertung_Nutzungskonflikte_Dr_Fuchs_web.pdf
- Jankó, J. A. & Szabó, Gy. (2013), Comparative Settlement Diagnosis. In: Lóki, J. (Ed.). Az elmélet és a gyakorlat találkozása a térinformatikában IV.: Térinformatika Konferencia és Szakkiállítás. Debreceni Egyetemi Kiadó, Debrecen. 237–243. pp.
- Kertész, Á., Őrsi, A. & Tóth, A. (2014), Tájéértékelés térinformatikai módszerekkel. In: Márkus, B. (Ed.). Térinformatika 2014. Székesfehérvár. 351–362. pp. <http://real.mtak.hu/id/eprint/20226>
- Kiss, I. (1996), A település(együttes)ek vizsgálata és egybevetési módszertana, Polgári Tanácsadó Szolgálat, Budapest, pp. 1–215.
- Miorelli, E. & Zatelli, P. (2006), Landscape relevance analysis with GRASS. Geomatics Work-books. 7 http://geomatica.como.polimi.it/workbooks/n6/articoli/zatelli_visibilita.pdf.
- Skalna, I., Rebiasz, B., Gaweł, B., et al. (2015), Advances in Fuzzy Decision Making, Studies in Fuzziness and Soft Computing, Springer International Publishing, Cham, <http://link.springer.com/10.1007/978-3-319-26494-3>.
- Szabó, M. A. & Szabó, M. E. (2003), Erdélyi helységnevszótár; 2. átd. és jav. kiad. Kriterion, Kolozsvár, <http://szabo.adatbank.transindex.ro/index.php>.
- Tudor, C. (2014), An Approach to Landscape Character Assessment. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/396192/landscape-character-assessment.pdf