

Capturing of Uncertainty by Geographical IS in Decision Support for Selected Aspects of Quality of Life

Jitka Machalová

Faculty of Business and Economics, Mendel University in Brno, Zemedelska 1,
613 00 Brno, Czech Republic
jitka.machalova@mendelu.cz

Abstract. Quality of life has currently been intensively studied and assessed by using a variety of indicators. Information and communication technologies, or more precisely geographical information technologies, prove invaluable in this process. It is clearly difficult to quantify and spatially localize the quality of life. Fuzzy sets allow the capturing of uncertainty. This article deals with the integration of fuzzy sets into a geographical IS to capture, analyse and evaluate selected aspects of the quality of life.

Keywords: Fuzzy Sets, Geographical IS, Quality of Life.

1 Introduction

Information systems are used for decision support at all levels of management. Information systems can be viewed as models of specific parts of reality (storage, human resources etc.). They save information about the modelled reality in the form of data; the data are then restructured, analysed and presented to decision-makers by the systems at appropriate times, in requested quality, form, volume and content. Information systems capable of modelling landscapes and the processes running in landscapes are known as geographic information systems, or geo-information systems (GIS).

There are certain setbacks in using models generated by such systems. A typical disadvantage of every model is its simplification of the modelled reality. One of the major simplifications in models based on GIS, is the use of binary logic, for example, if we need to mark an area as waterlogged or dry. Any decision maker, however, who implements these models uses natural language working with uncertainty, for example: the area is “slightly” or “substantially” waterlogged. The tool to work with such uncertainty is the theory of fuzzy sets. Implementation of fuzzy sets in geographic information systems allows us to model, analyse and evaluate the modelled processes better.

A good example of the application of fuzzy sets could be a request to evaluate the quality of life in a specified area. In scientific literature, the term quality of life (QOL) has been discussed by classic psychologists, such as C. G. Jung. First attempts to measure the quality of life were made in treatment of patients. The findings were published in the field of medicine, especially in psychiatry, oncology and geriatrics.

Later, the term quality of life and attempts to measure it can be seen in social sciences. One approach seeks to specify objective conditions for quality of life. The other approach monitors subjective human perception as satisfaction with one's own life. In the last ten years we can see attempts to measure the quality of life from the perspective of marketing and management. The need for sustainable development forces us to compromise between the need for protecting the environment and its maximal use. The number of tools for measuring the quality of life is growing, as they become more specialized in several domains of this kind of assessment.

This work deals with the integration of fuzzy sets into geographic information systems in order to capture, analyse and evaluate selected aspects of the quality of life. The main ideas on the aspects of the quality of life, collectively known as the "well-being factor" will be presented and applied to a representative area in South Moravia in the Czech Republic.

2 Modelling of Landscape Phenomena

Information about landscape phenomena (elements with a relation to a location on Earth) is saved in the form of geographical data in geographic information systems. Spatial data or geographical data are data which are related to specific places in space with known location at the required level of detail [8].

2.1 Geodata and Map Algebra

Storage of high volume data sets is now managed by geodatabases [11]. The data is usually saved in a vector or raster format as thematic layers. Data in vector format has two parts: a spatial and descriptive i.e. attribute part. The spatial part typically uses the rules of vector graphics; the descriptive part uses rules of relational database systems. In the case of geodata storage in the raster format, the layer is represented by a square network of cells or pixels, where a specific value, a number from a specific domain, is assigned to each cell.

The choice of geodata storage depends on many factors. One of the important factors is the need of spatial analysis, performed on the data. Spatial analysis is a process we use to obtain information from geodata. We can use a map algebra apparatus for data saved in the raster format. The map algebra makes it possible to combine raster layers using different mathematical operations and a map algebra language. The map algebra is a tool which allows us to process the raster representation of reality by using the map algebra language. We can carry out map algebra operations with one or more layers. The operations or functions are divided into local, zonal, focal and global ones [14]. The result of a local function is a change of value of one layer cell. A focal function changes cell values based on values of the neighbouring cells. These functions are mostly used in flow analyses. Zonal functions change values of cells based on values in another layer. A global function concerns the whole layer and is used for example, in distance analyses such as costs or timing.

Functions are built from operators and operands. Operands can be: arithmetic (plus, minus, times etc.), Boolean (true, false), relational (greater than etc.) as well as specialized, such as bit shifts, accumulative, combinatorial assignment operators. The operands are values stored in a cell or cells of layers used in the spatial analysis (Fig. 1).

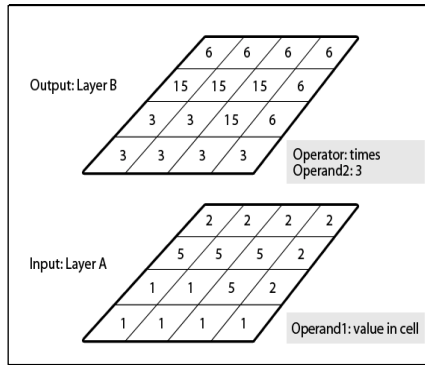


Fig. 1. Example of map algebra used with an arithmetic operator on one layer

2.2 Fuzzy Logic

The basic advantage of fuzzy logic is its ability to mathematically describe information expressed in words. In other words, it enables us to work with ambiguous terms often used in human language. Fuzzy logic permits us to use ICT resources to discern the meanings of “less” or “significantly” based on mutual comparison. The possibility to search for information based on ambiguous or incomplete values with an option to find incorrectly saved information [5] is also an advantage.

Suppose there is an element a . We have to decide, whether it belongs to a given set or not. Each element a in a domain D will be assigned an integer z , $z \in \langle 0;1 \rangle$, which expresses the grade of membership in the set. Value 0 means: completely not included in the set; value 1 means: fully included in the set. If we have a set of elements, where each element is characterized by a grade of membership, we call this a fuzzy set [16]. The traditional set can be understood as a special case of the fuzzy set, $z \in Z$, $Z \equiv \{0;1\}$.

Numbers z define a membership function. The traditional sets with exact borders h_1, h_2 can be defined as follows:

$$MF(a) = 1, \text{ if } a \in \langle h_1; h_2 \rangle \tag{1}$$

$$MF(a) = 0, \text{ if } a \notin \langle h_1; h_2 \rangle \tag{2}$$

Fuzzy sets can be expressed as follows:

$$A = (a, MF(a)), \text{ if } \forall a \in D, \tag{3}$$

where $MF(a)$ is a membership function of element a in domain A with values in closed interval $\langle 0;1 \rangle$. Range of the values of the function MF can be either the whole interval $\langle 0;1 \rangle$ or only few selected values.

The MF function and its range can be defined by an expert in the given field or by using operations research methods. The former is used when we model the semantics of words of a natural language. The meaning of words is a subjective matter. When possible, we can ask a number of people to make a statement. There are two ways of doing this. The first option is to let the subjects to express themselves using a number or a graphic scale in order to determine whether a given element satisfies the

researched expression. Subsequently, the resulting grade of membership is found using a statistical evaluation e.g. calculating the average of the individual observations. The second option is to let the people express whether the given element satisfies the researched expression or not and the resulting grade of membership will be determined as the percentage of positive answers to the total count of answers. In a real situation, a single person (expert) defines a point A which is not included in the studied set, a point D which is another extreme value not belonging to the set, an interval of values $\langle B;C \rangle$ with full membership (grade of membership equals 1), and intervals $\langle A;B \rangle$ and $\langle C;D \rangle$ to be spanned by a suitable function which approaches the extreme grades of membership at the endpoints of these intervals (Fig. 2).

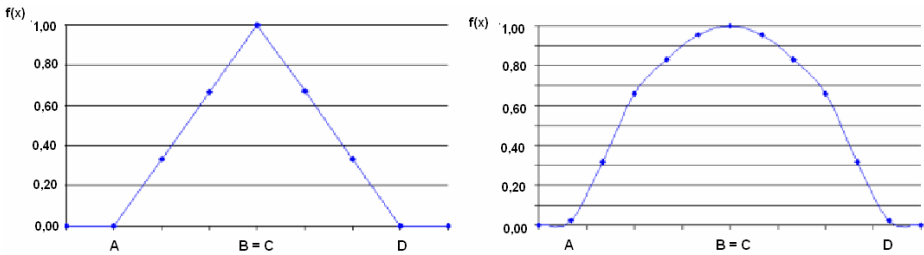


Fig. 2. Linear and parabolic membership function of a fuzzy set

To a great extent, a fuzzy set is determined by the membership function it is associated with. The selection of a membership function influences the borders of the fuzzy set.

3 Principles of Fuzzy Sets Integration into Geographic IS

As mentioned above, GIS are able to model the landscape around us. Some elements of landscapes for example plots, have exact borders. However, it is difficult to delimit the borders of some phenomena, such as waterlogging. It is often unclear, how to correctly define for example the possible limits of land slopes. Fuzzy sets are particularly suitable and useful because of their ability to capture the uncertainty when modeling the aspects of the quality of life. In these cases it appears necessary to introduce uncertainty into the GIS in the form of fuzzy geoelements.

Fuzzy geoelement is a geoelement which in addition to a geometry, topology, attributes and time also has its own membership function. The basic fuzzy geoelements are a fuzzy point, fuzzy line and fuzzy polygon [10].

Each type of the fuzzy geoelement depends on a membership function. Its modeling is based on the concrete membership function. In raster information storage, values of each cell are calculated on the basis of the membership function and a distance of the cell from cells with the value of a membership function equal to one.

Fuzzy point is a fuzzy geoelement given by the ordered triad F, A, M , where F is a membership function, A (with coordinates x, y) is a cell with a value of the membership function equal to one, and M is a set of description (attribute) values of the fuzzy point [10] (Fig. 3).

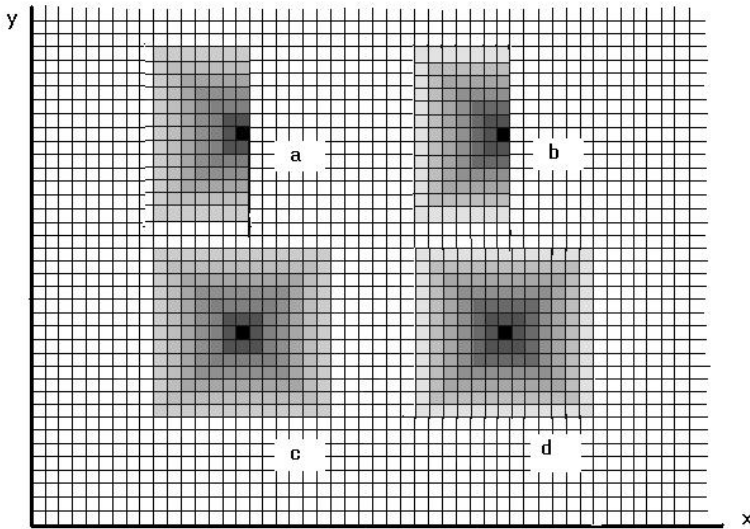


Fig. 3. Fuzzy point in raster representation with different membership function

Fuzzy line is a fuzzy geoelement given by the ordered $n+2$ -tuple $F, A_1, A_2, \dots, A_n, M$, where F is a membership function and A_1 to A_n are neighbouring cells with the membership function equal to one. A_i only has neighbours A_{i-1} and A_{i+1} , A_1 only has a neighbour A_2 , and A_n has only A_{n-1} . M is a set of description values of the fuzzy line [10] (Fig. 4).

Fuzzy polygon is a fuzzy geoelement given by the ordered $n+2$ -tuple $F, A_1, A_2, \dots, A_n, M$, where F is a membership function and A_1 to A_n are neighbouring cells with the membership function equal to one, which fill contiguous space. M is a set of describing values of the fuzzy polygon. Fuzzy polygon, which meets requirements of the previous definition, is a fuzzy line [10] (Fig. 5).

The cell A is a cell of fuzzy geoelement P just if it has the membership function of the fuzzy geoelement P greater than 0.

To make fuzzy geoelements applicable in spatial analyses, it is necessary to define geometrical and topological relations. When Euclidean metrics is used, a position of each cell is defined as the position of its centre, and the distance between two cells is defined as the distance between their centres.

There are three types of distances between fuzzy points:

1. minimal – minimum of distances of the fuzzy geoelement cells with the grade of membership greater than 0,
2. mean – minimum of distances of the fuzzy geoelement cells with the grade of membership equal to 0.5,
3. simple – minimum of distances of the fuzzy geoelement cells with the grade of membership equal to 1.

Two fuzzy lines are continuous only if there is a cell with the membership function value equal to 1 with both lines.

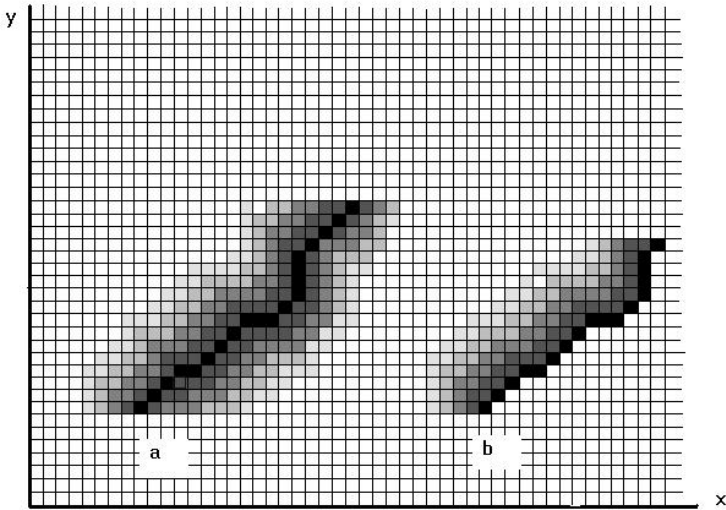


Fig. 4. Fuzzy line in raster representation with a different membership function

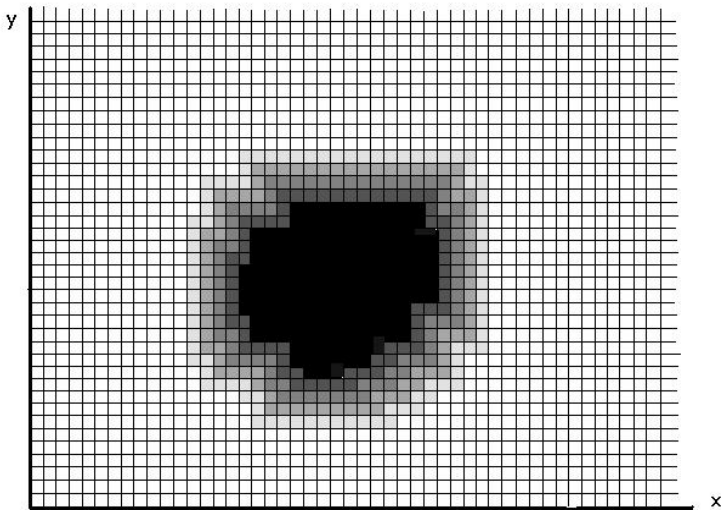


Fig. 5. Fuzzy polygon in raster representation

Two fuzzy polygons are neighbouring only if there is at least one cell which has the membership function value greater than 0 with the first fuzzy polygon, and at the same time, there is a neighbouring cell (with a common border) having the membership function value greater than 0 with the second fuzzy polygon.

Two fuzzy polygons are simple neighbours if there is at least one cell having the membership function value equal to 1 with the first fuzzy polygon, and at the same

time, there is a neighbouring cell (with a common border) having the membership function value equal to 1 with the second fuzzy polygon.

Fuzzy polygon P contains a second fuzzy polygon Q if each cell of the fuzzy polygon Q with non-zero membership function value has also non-zero membership function value with the second polygon P [10].

As described above, fuzzy sets provide an option to capture the uncertainty, or vagueness typical for human language. However, the necessity to describe the reality using binary logic causes inaccuracies, which can consequently lead to poor decisions. For example, how can we exactly specify low level of traffic? Fuzzy sets make it possible to capture, analyse and present such “fuzzy” information. The use of fuzzy sets in similar situations is contingent on the following three steps:

1. Correct definition of types of membership functions and the values A, B, C and D (Fig. 2)
2. Assurance of good-quality input data, where the quality is given by precision, authenticity and recency [7].
3. Modelling fuzzy sets defined by experts and execution of spatial analyses over these sets.

Solution of the first two steps is not the main contribution of this article. As noted above, the correct definition of fuzzy sets is up to experts in the given research area. The second step can be critical for the success of the project, because especially with raster data, it requires a compromise between the resolution (number of cells per inch, called DPI) and the size of the file storing the information (in MB, or TB). We usually encounter performance limits of computers at this stage of the analysis, or cluster computers where the analyses are processed.

The article is innovative especially by the approach used in solving the third step. Because the full use of the mentioned methods to provide a comprehensive methodology for assessing the quality of life would be beyond the scope of this article, the article will be focused only on the specific aspects of the well-being factor.

4 Application to the Well-Being Factor

The methodology described above will be applied to one of the quality of life indicators – the well-being factor. The study has been carried out on a sample region in South Moravia in the Czech Republic. The area is 150 km². It is relatively densely and evenly populated. There are several small municipalities present, including a regional centre – a town. The landscape is flat with undulating parts and hills. It is used for agriculture; forests are present as well. We can find factories here and also an area unaffected by industry.

The well-being factor is a summary criterion which means a low density of buildings, traffic, noise, light and smell pollution [13]. In the Czech Republic, the well-being factor is understood as a landscape characteristic comprised of the following factors: visual (the visual contamination of scenery), acoustic (the noise level not reaching or exceeding threshold limits), odor (not reaching hygienic limits, but dominant in a negative sense) and emotional disturbance of well-being factors [12].

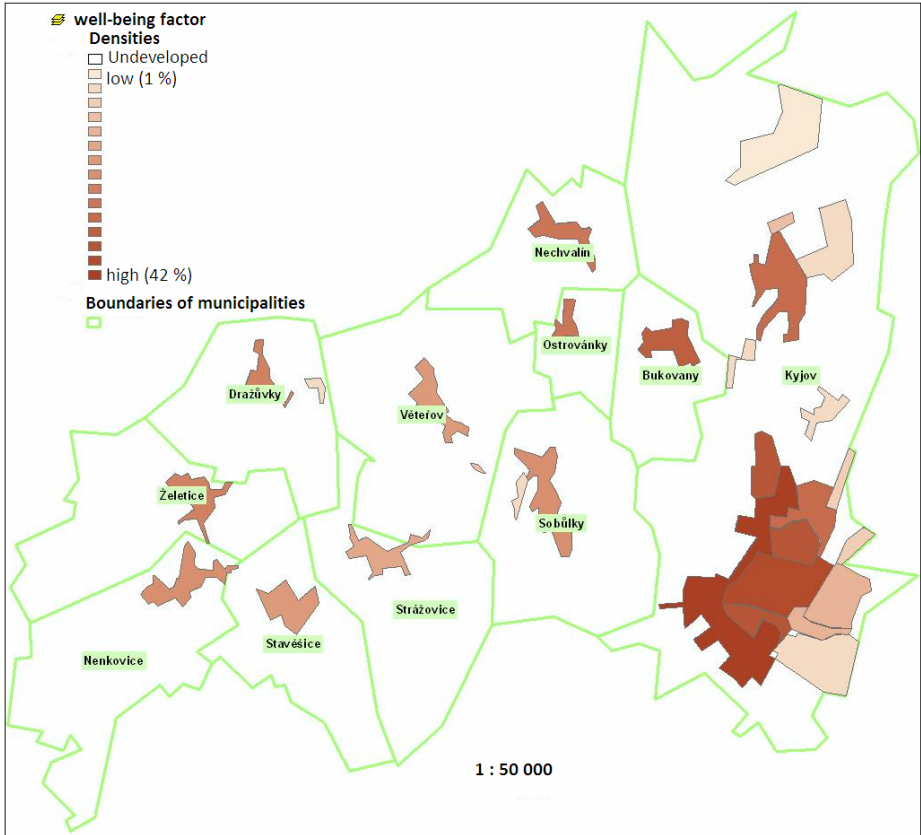


Fig. 6. Fuzzy evaluation of the density of buildings in the monitored area

The density of buildings differs from population density¹. The density of buildings is reported as the percentage of the built-up area to the total area. It is possible to use vector layers of buildings and utility buildings as input data. These layers have to be rasterized, reclassified and merged to a binary raster layer of buildings using map algebra (1 – built-up, 0 – un-built). Using cadastral maps and land plans we will define the built-up area of a municipality and the built-up areas such as country house areas outside the built-up area of the municipality. In these areas, the percentage of the built-up area will be calculated in reference to the total area and expressed using fuzzy sets (Fig. 6) with a linear membership function. Point A = 1 %, B = C = 42 % – the maximal density (Fig. 2). The following geographic representation is not expressed as fuzzy sets, because the monitored areas have fixed boundaries also defined in the land plans.

¹ Population density is measured as the number of inhabitants per square kilometer. In the Czech Republic, it represents 130 inhabitants/km², in the monitored area 150–200 inhabitants/km².

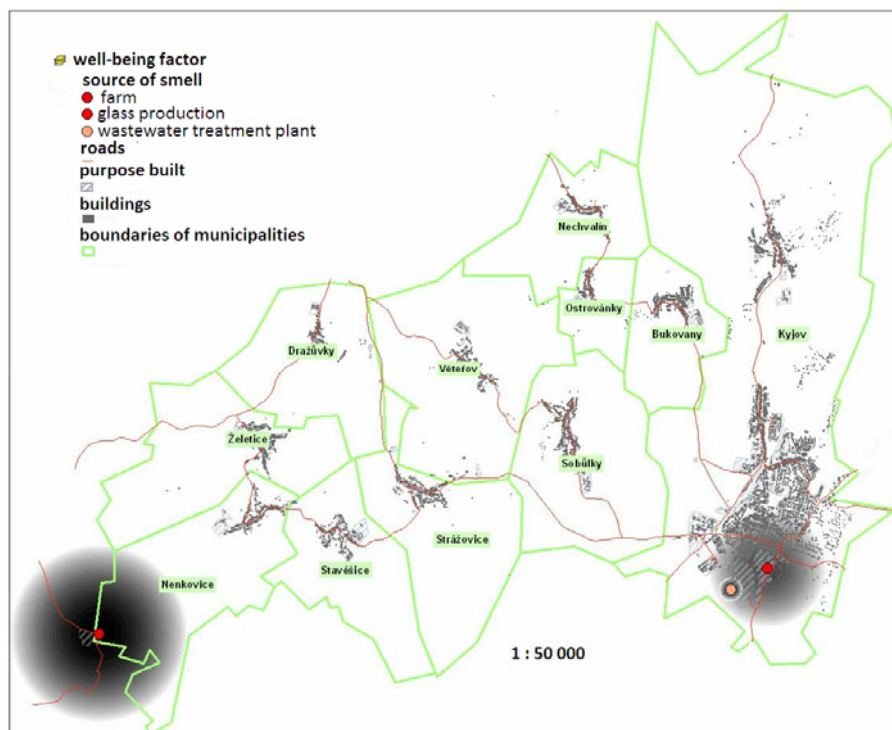


Fig. 7. Fuzzy evaluation of the sources of odor in the monitored area

Odor is a very subjective matter. For the purpose of standardization, relevant legislation regarding odor and its acceptable limits was enacted (Act No. 86/2002 Coll., Decree No. 362/2006). It sets limits for each substance in appropriate units which represent acceptable odor pollution. The well-being factor however, is often affected by values below the legal threshold. There are three sources of odor in the monitored area: a glass factory (arsenic, chromium, cadmium and airborne dust), a sewage water treatment facility (phosphorus) and a cattle and poultry farm (ammonia). The given scenario represents a fuzzy set with a parabolic membership function. Two points, A and D were set, both 1.500 m far from the farm, 500 m far from the sewage water treatment facility and 1.000 m far from the glass factory. The point $B = C$ was set in each source of odor. The figure 7 visualizes the influence on the well-being factor in reference to the character of each source of odor.

There is a personal and public transportation in the monitored area. The transportation takes place on public roads. The quality of life in the area is affected by noise caused by the traffic on the roads and by the quality of roads during their usage. In addition the distance from stops and the time distance from the regional centers while using the public transportation also affects the quality of life.

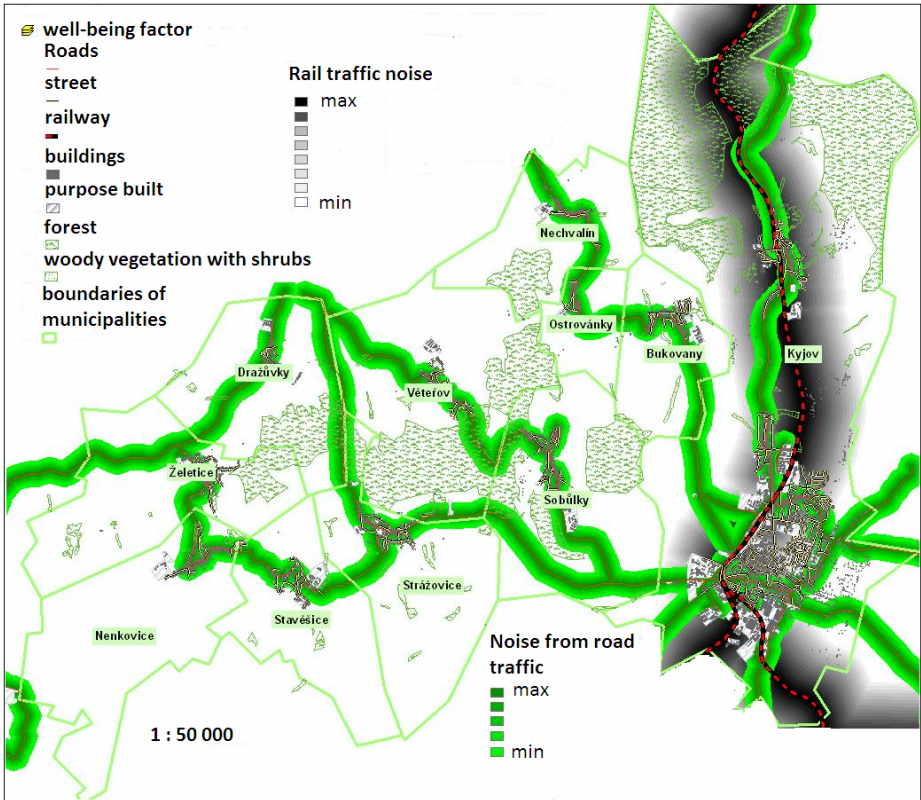


Fig. 8. Fuzzy evaluation of the source of noise in the monitored area

According to the data published by the World Health Organization, a long-term noise burden above 55 dB is harmful and can cause a variety of illnesses. Noise above 65 dB is reported by physicians as in the long-term unbearable and demonstrably harmful to health [1]. Strategic noise maps have been published on the basis of the obligation to create and disclose noise maps. This obligation was stipulated by the European Directive No. 2002/49/EC relating to the assessment and management of environmental noise. However this is done only for those urban areas and road systems, where more than 6 million cars pass annually. Road traffic has a 95 percent share in the proven annoyance by noise [1].

The noise map of the Czech Republic classifies areas into 7 zones according to the noise level. This classification however, was not made with fuzzy sets. Our monitored area is not included in the noise map. In 2005, the Road and Motorway Directorate of the Czech Republic prepared a summary map of the number of cars passing measure points in 24 hours. According to the results, the intensity of traffic is relatively low and stable on the public roads, with values from 1 500 to 3 000 cars in 24 hours [3]. There were also local measurements carried out in the streets under the auspices of the municipal authorities. The result is that most streets are used for private transportation of the locals. Another source of noise is a non-electrified state-owned railway.

There are no other important sources of noise caused by industry present in the monitored area.

Buildings, vegetation and terrain relief are important in the elimination of the spread of noise. The landscape is relatively open and flat with one major hill. The input data for analyses are the layers of buildings, utility buildings, forests, greenery with scrubs and slopes. In order to get a qualified analysis, the layers must be rasterized first and then reclassified in accordance with the influence on the noise elimination.

Furthermore, a summary layer of the elements eliminating noise must be created using map algebra. The layer then inputs together with the membership function of the spread of noise into the function of spreading. The result is a pair of fuzzy layers presenting the noise caused by the railway and road (public roads + streets) traffic (Fig. 8).

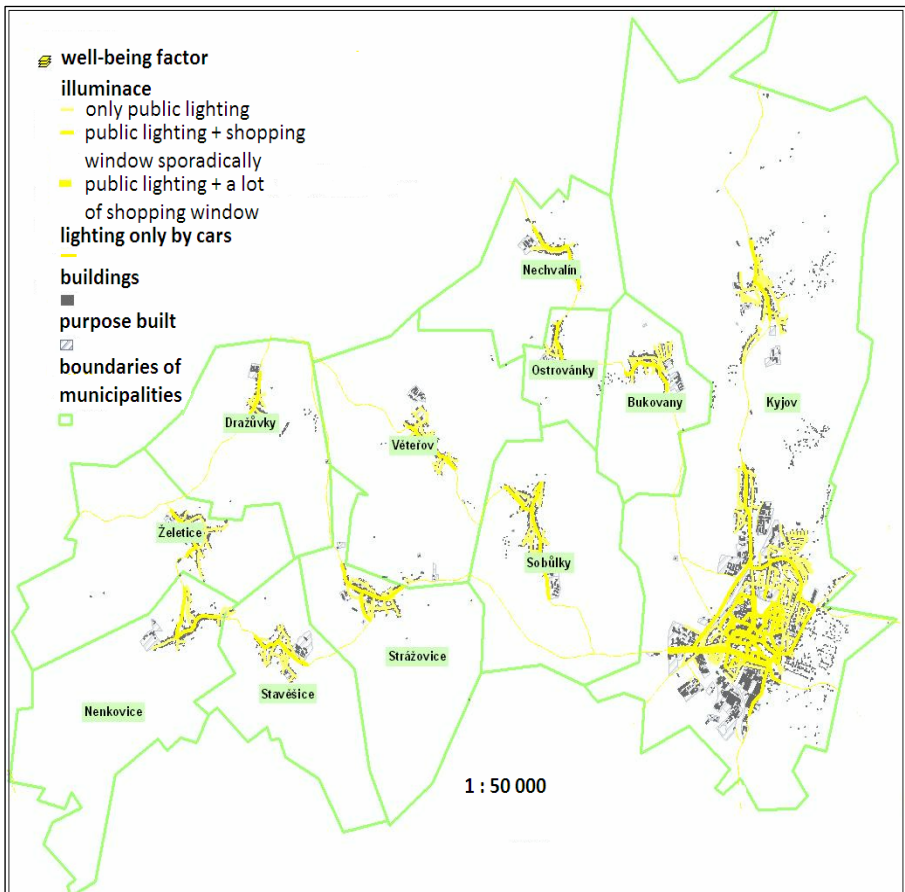


Fig. 9. Fuzzy evaluation of the light pollution in the monitored area

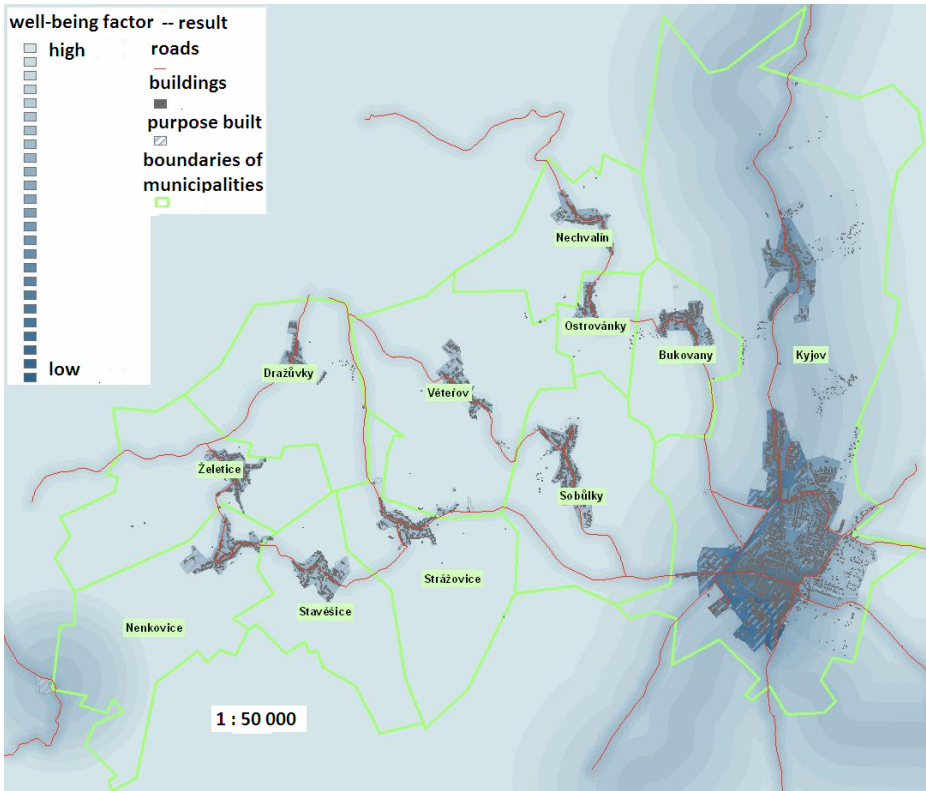


Fig. 10. Fuzzy evaluation of the well-being factor in the monitored area

Light pollution is caused by artificial light, operated without a reason in places, which do not need to be illuminated. This way it disturbs the natural darkness and often glares dangerously. It also disturbs the natural night and biorhythms of the living organisms. The Institute of Science and Technology of Light Pollution (Insituto di Scienza e Tecnologia dell'Inquinamento Luminoso) created a map of Europe classifying the area in Europe into 4 categories, in the range from 0.33 to 27, according to the proportion of the natural and artificially illuminated sky. In our monitored area, the town of Kyjov has the range of values from 3 to 9 and the other areas between 1 and 3. The light well-being is relative present. The light pollution is mainly caused by unsuitable designs of street lamps, illuminated window displays and advertising lasers. In the monitored area, light pollution is caused by passing cars, street lamps and illuminated window displays. The layers of streets and roads with the rating of light intensity are inputs of the analysis. The result are presented in Figure 9.

A marketing research aimed at finding the importance of the researched factors for local inhabitants was recently conducted in the monitored area. 506 respondents living in the Kyjov area took part in the research. 300 respondents were women of age between 22-66 years with the prevailing secondary education. In addition, controlled interviews were conducted with all mayors of surrounding villages. The survey

showed clearly that respondents attached the lowest importance which influences the subjective well-being to light pollution and the highest importance to traffic noise. The respondents found the density of buildings as fair; an excessive density of buildings would negatively affect their well-being factor. Odor pollution was found as important as the noise pollution by the respondents of areas affected by odor.

The resulting well-being factor was created using map algebra where the input operands were the layers of the density of buildings, smell, noise and light pollution; the operator is addition. The operands are multiplied by a coefficient which emerged from the marketing research (Fig. 10).

$$\text{Well-being factor} = 2 \times \text{the density of buildings} + 4 \times \text{the smell pollution} + 5 \times \text{the noise pollution} + \text{the light pollution}$$

5 Conclusions

The result of the analyses is the spatial presentation of the well-being factor. The presentation assigns a well-being factor value of the resulting fuzzy layer to each place of the monitored area. According to the known facts that the well-being factor is mostly influenced by the built-up area, noise, odor and light pollution, the lowest value of the well-being factor can be found in the built-up area and the areas in the neighbourhood of roads or the railway. The highest values of the well-being factor in built-up areas are in the village of Věteřov which is off the main roads. The lowest value of the well-being factor can be found in the regional centre – the town of Kyjov. This can be credited to railway transport, several main streets, two sources of odor and the highest values of the light pollution. The well-being factor is only one of the factors influencing the quality of life. It is completely in accordance with current trends of people moving from the urban areas to rural calm, but within the grasp of the city. Comprehensive modelling of the spatial aspects of the quality of life is a subject of further research.

The use of fuzzy layers and map algebra can better capture the uncertainty natural to human thinking expressed in language. Unlike the traditional geoelements, the fuzzy geoelements are able to model selected spatial phenomena more accurately. Therefore the results reflect the modelled reality better.

References

1. Bernard, M.: Analysis of results of the strategic noise maps, http://aa.ecn.cz/img_upload/analyza.doc
2. Czech Astronomical Society: Light Pollution, <http://www.astro.cz/zncistenii/>
3. Directorate of Roads and Highways: Traffic intensity in 2005 in the Czech Republic (2005), <http://www.rsd.cz/Silnicni-a-dalnicni-sit/Intenzita-dopravy>
4. Hrebicek, J., Soukopova, J., Kutova, E.: Standardization of Key Performance Indicators for Environmental Management and Reporting in the Czech Republic. *International Journal of Energy and Environment* 4, 169–176 (2010)

5. Klir, G.J., Yuan, B.: *Fuzzy Sets and Fuzzy Logic*. Prentice Hall, New Jersey (1995)
6. Konecny, M., Zlatanova, S., Bandrova, T.L. (eds.): *Geographic Information and Cartography for Risk and Crisis Management*. Springer, Heidelberg (2010)
7. Lake, R.: *Virtual Globes as Essential Services?*,
<http://www.galdosinc.com/archives/533>
8. Longley, P., Goodchild, M.F., Maguire, D.J., Rhind, D.W.: *Geographic Information Systems and Science*. John Wiley & Sons Ltd., Vancouver (1981)
9. Machalova, J.: Modeling of chosen selectable factors of the develop of tourism with geographic IT and Fuzzy sets using. In: *Acta univ.agric.et silvic Mendel., Brun., vol. LVIX(1)*, pp. 189–198 (2011)
10. Machalova, J.: Definition of fuzzygeoelements in raster representation for decision-making. In: *International Symposium Digital Earth*, pp. 242–244. Masaryk University, Brno (2004)
11. Prochazka, D.: On Development of Search Engine for Geodata. In: *Acta univ.agric.et silvic Mendel., Brun., vol. LVIII(6)*, pp. 389–398 (2010)
12. Sklenička, P.: Protection of Landscape Character in Process EIA. In: *Zpravodaj EIA, Ministry of the Environment of the Czech Republic, Praha* (2002)
13. Swanwick, C.: *Land use consultants. Landscape Character Assessment – Guidance for England and Scotland*. The Countryside Agency – Scottish Natural Heritage, Edinburgh (2002)
14. Tomlin, C.D.: *Geographic Information Systems and Cartographic Modeling*. Prentice Hall, New Jersey (1990)
15. Wilson, J.P., Gallant, J.C.: *Terrain Analysis: principles and applications*. John Wiley & Sons Ltd., Vancouver (2000)
16. Zadeh, L.A.: *Fuzzy sets. Inf & Control* (1965)