SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12, DOI: 10.2478/sjce-2013-0016

R. ĎURAČIOVÁ

QUERYING UNCERTAIN DATA IN GEOSPATIAL OBJECT-RELATIONAL DATABASES USING SQL AND FUZZY SETS

Renata ĎURAČIOVÁ
Email: renata.duraciova@stuba.sk

Research field: GIS, spatial data management, spatial databases, data quality, uncertainty modeling, spatial analysis

Address: Department of Theoretical Geodesy, Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Radlinského 11, 813 68 Bratislava, Slovak Republic

ABSTRACT

This paper deals with uncertainty modeling in spatial object-relational databases by the use of Structured Query Language (SQL). The fundamental principles of uncertainty modeling by fuzzy sets are applied in the area of geographic information systems (GIS) and spatial databases. A spatial database system includes types of spatial data and implements the spatial extension of SQL. The implementation of the principles of fuzzy logic to spatial databases brings an opportunity for the efficient processing of uncertain data, which is important, especially when using various data sources (e.g., multi-criteria decision making (MCDM) on the basis of heterogeneous spatial data resources). The modeling and data processing of uncertainties are presented in relation to the applicable International Organization for Standardization (ISO) standards (standards of the series 19100 Geographic information) and the relevant specifications of the Open Geospatial Consortium (OGC). The fuzzy spatial query approach is applied and tested on a case study with a fundamental database for GIS in Slovakia.

KEY WORDS

- *Spatial database,*
- *SQL,*
- *GIS,*
- *uncertainty modeling,*
- *fuzzy sets.*

1 INTRODUCTION

Spatial data is inherently uncertain. The main sources of uncertainty include uncertainty in the collection, processing and representation of data (e.g., missing or ambiguous data, measurement uncertainties, raster-to-vector conversion, indeterminate borders of geographical objects, ambiguities in object identification, uncertainty in the interpretation and digitization of data, image classifications, interpolation of values and determination of attributes, etc.). In many situations it is necessary to make important decisions based on an uncertain data analysis. The functions and operations which are performed with the data in information systems are mostly implemented by making use of crisp rules and criteria based on Boolean logic, which often leads to a loss of information resulting from the uncertainty in the input data. It is therefore useful to model the uncertainty by available tools in information systems to reduce the risk of any impact of the adoption of incorrect decisions based on uncertain assumptions.

The goal of this paper is to present a fuzzy spatial query approach, which can be applied in spatial analysis and multi-criteria decision making (MCDM) in geographic information systems (GIS). The paper introduces the possibility of the representation and processing of imprecise and vague spatial data in object-relational databases. It refers to uncertainty modeling by the fuzzy set theory, which does not work with the concepts of randomness and probability, but expresses uncertainty in terms of vagueness or ambiguity. Bivalent (Boolean) logic is then replaced by fuzzy logic in the analysis, resulting, for example, in the introduction of triangular norms (t-norms) as

Vol. XXI, 2013, No. 4, 1 – 12

aggregate functions in the spatial decision making. The principles of modeling uncertainties are applied in the area of GIS and spatial databases, which include spatial data types and which have implemented extensions of Structured Query Language (SQL) for working with them (as spatial functions and operations). A spatial database can be created from heterogeneous data layers of varying quality. For spatial analyses which integrate heterogeneous data layers (spatial tables in the database), we use the fuzzy decision-making criteria expressed by SQL queries. The application of fuzzy spatial SQL was realised in the PostgreSQL object-relational database system.

2 STATE OF THE ART

Uncertainty and spatial data quality modeling is not a new topic of research in GIS (Longley et al., 1999), but now this theme has become quite current because of the frequent handling and processing of heterogeneous geospatial data sources in web environments. Another reason is also the number of activities related to the building of a spatial data infrastructure (Directive, 2007) and data sharing, where the quality of the data plays a very important role for spatial data users. Information about quality (ISO 19113 and ISO 19114) is part of the metadata (ISO 19115), and it can form the basis for modeling uncertainty in spatial databases.

A number of researchers in the GIS area have recently considered spatial data models using fuzzy set approaches (Bosc et al., 2005), as in the modeling of geographic objects with indeterminate boundaries (Burrough and Frank, 1996). The first consistent approach in the use of fuzzy set theory which could be applied in GIS was developed by Robinson (1990) (Bosc et al., 2005). More recently, there have been a number of efforts utilizing fuzzy sets for spatial databases including the definition of fuzzy spatial data types (Schneider, 2008); fuzzy UML modeling (Xiaoxiang, 2005); fuzzy spatial topology (Tang, 2004), (Tang et al., 2006); fuzzy spatial decision making (Morris and Jankowski, 2005), (Petry et al., 2005); fuzzy data modeling for the integration of heterogeneous spatial information (Mukherjee and Ghosh, 2011), etc. Semantic interpretations of fuzzy spatial queries are described in (Wang, 2000), (Cuevas et al., 2008). Fuzzy classes, domains for imprecise vague data, and fuzzy equality operators (FEQ) are defined in (Cuevas et al., 2008). The autors describe the manipulation of fuzzy objects in an object-relational system, but it is applied only for non-spatial data types.

The basis of data querying and manipulation in relational and object-relational databases is SQL. In spatial databases (usually in object-relational databases), it is necessary to use the extensions of SQL for spatial data that are included in the OGC specifications and the series of ISO 19100 Geographic information standards. The processing of uncertain data in relational and object-relational databases is not currently standardized, but it is a hot topic of research in the area of database systems and also in data mining. The first indications and results are presented in many papers, for example in (Kacprzyk et al., 2000), (Hudec, 2009), (Hudec et al., 2012), and (Chen, 2012). The papers introduce possibilities to extend SQL for the processing of uncertain data and vague queries in databases (e.g., Fuzzy SQL (Hudec, 2008), (Hudec et al., 2012) or SQL Uncertainty (USQL) (Chen, 2012)), but they do not work with spatial data. The fuzzy extension of SQL to work with uncertain spatial data is much more difficult than querying in conventional relational databases and will require further research. It is necessary to define a number of new principles and procedures for the processing of uncertain spatial data. The result should be the creation of a standard for working with uncertain spatial data and vague criteria, which would include implementation of the knowledge and principles from two scientific fields, namely, spatial database systems and uncertainty modeling by fuzzy set theory.

3 MATERIALS AND METHODS

3.1 Spatial queries in object-relational database systems

"Spatial data" is a term used to describe data with a reference to a specific location. Spatial data is usually stored in databases as coordinates and topology and is often accessed, manipulated or analyzed through Geographic Information Systems (GIS). A spatial database is a database that is optimized to store and query spatial data. One of the most important directions in the development of GIS is the common storage of attributes and geometric data in the same database. The ability to save spatial objects in relational tables is applied in the case of object-relational databases. A column containing spatial (geometric) data is then called a spatial column. A spatial table is a table that includes one or more spatial columns. Spatial tables have different properties than conventional tables, because they also represent the topological and metric relationships of the objects.

A powerful tool for querying relational databases is the standardized language SQL, which is currently implemented in almost all relational database systems. SQL provides functions for the administration and processing of standard data types (numeric, string, Boolean, etc.), but does not automatically support functions for working with spatial data types. Spatial data handling requires special support for spatial data types, indexes, operations and functions that are integrated directly in a database management system (DBMS). Then the spatial and attribute data can be accessed within the same query. An SQL standard extension called SQL/MultiMedia (SQL/MM) permits storing, retrieving and processing spatial data using SQL in existing database systems. This spatial extension is currently implemented and used in several relational database systems (e.g., PostgreSQL, Oracle Spatial, Microsoft SQL Server (versions of MS SQL Server 2008), and IBM DB2).

3.1.1 Spatial data types defined by a geometry object model

The Open Geospatial Consortium (OGC) (http://www.opengeospatial.org) develops standards for the definition of spatial and

SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

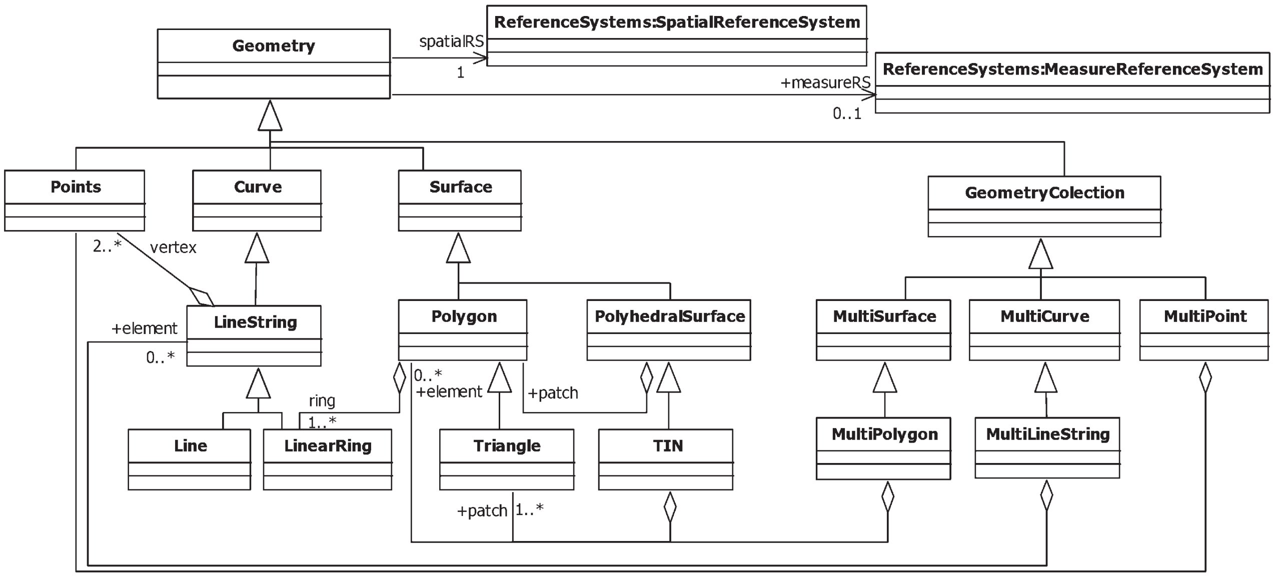

Fig. 1 *Geometry object model - basic spatial data types in UML (according to OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option).*

geographic data. OGC specifications form the basis for ISO standards and are gradually introduced to the most used object-relational database systems. The OGC describes a spatial object by a geometric model based on a hierarchy of classes in the standard Simple Features Access (SFA), which has became the basis for the ISO 19125-1 (ISO 19125-1:2004 Geographic information - Simple feature access - Part 1: Common architecture) and ISO 19125-2 (ISO 19125-1:2004 Geographic information - Simple feature access - Part 2: SQL option) standards. The SFA standard includes a description of the basic geometric vector elements, a specification of any deposits, and access to the object-relational databases. The OGC geometric model is created using the class diagram in Unified Modeling Language (UML) (Fig. 1).

3.1.2 SQL extension for spatial data

The extension of SQL options for processing spatial data (types of spatial data) is the logical result of activities aimed at integrating spatial and non-spatial data. Similarly as in the case of selections by attributes, the result of each spatial selection is a table (relation), and the essential element of the selection is the SQL command SELECT (such as in the case of non-spatial databases). There are types of data and possible operations (functions) specified in an SQL extension for spatial data. Spatial operations supported by the database system can then be included in the basic command structure of SELECT – FROM – WHERE. The functions for spatial data in object-relational databases are:

- functions to define spatial data: *AddGeometryColumn*, *DropGeometryColumn* (the creation or cancellation of a spatial column)
- functions to determine the topological relations of objects (the returned value is "TRUE" if the condition is true, and "false" if the condition is not true): *Equals*, *Intersects*, *Touches*, *Disjoint*, *Within* (one object is located in another one), *Contains* (one object contains another one), etc.
- functions for spatial analysis (they return numeric values or spatial objects as a result of processing): *Distance* (distance between two objects), *Area* (area of the object), *Buffer*, *Intersection*, *Union*, *Difference*, *SymDifference* (symmetric difference), etc.

In addition to these functions, several other functions are defined as well, for example, functions for access to elementary geometry, and functions for format conversion are defined too (Koreň, 2009). All the functions are described in detail in the standard ISO 19125-2 or in the OGC Specification: OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option.

3.2 Modeling of uncertainty of spatial objects using fuzzy sets

Fuzzy sets are a mathematical tool, which provides a description of uncertainty and permits working with it. The concept of a fuzzy set was introduced by Professor Zadeh in 1965: "Let X be a space of points (objects), with a generic element of X denoted by x. Thus, $X = \{x\}$. A fuzzy set (class) A in X is characterized by a membership (characteristic) function $m_A(x)$ which associates with each point in X a real number in the interval [0,1], with the value of $m_A(x)$ at x rep-

SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

resenting the "grade of membership" of x in A." (A fuzzy set A is a set of elements $x \in X$ (X is called a universe), where each of them is assigned by a degree of membership $m_A(x)$ (the membership value of x in A), whose values must be between zero (no membership) and one (definite membership). The degree of membership reflects the rate to which the element belongs to the set. In fuzzy logic, it represents the degree of truth. A membership function for a fuzzy set A on the universe X is defined as m_A: $X \notin [0,1]$. The shape and parameters of the membership functions can be determined based on practical experience or on the known properties of the phenomenon analyzed. A trapezoidal (piecewise linear) membership function is the most commonly used one (Fig. 2 a) in spatial analysis in GIS (Kainz, 2013): if

$$\mu_A(x) = \begin{cases} 0 & \text{if } x < a, \\ \dfrac{x-a}{b-a} & \text{if } a \le x \le b, \\ 1 & \text{if } b < x < c, \\ \dfrac{d-x}{d-c} & \text{if } c \le x \le d, \\ 0 & \text{if } x > d. \end{cases} \tag{1}$$

The special cases of piecewise linear membership functions are presented in Figs. 2 b, c, d. Other frequently used membership function shapes are, e.g., Gaussian, sigmoidal or S-shaped.

3.2.1 Logical operations in fuzzy sets

The basic set operations of intersection, union, and complement are often used in spatial analysis in GIS (e.g., MCDM). In mathematical logic, which is also part of the queries in spatial databases in SQL, these operations correspond to the logical operations of conjunction, disjunction and negation. The logical operations in crisp sets are known from propositional calculus and Boolean algebra. They are implemented in database systems in the same way.

The logical operations in fuzzy sets are based on fuzzy propositional calculus. The operations of fuzzy conjunction, fuzzy disjunction, and fuzzy complement are generalizations of crisp ones. There is more than one way of generalization, but all of them provide a fulfilment of the conditions in the interval [0,1] instead of the crisp values 0 or 1. The most widely-used operations are standard fuzzy set operations (Zadeh, 1965):

The intersection (Zadeh, 1965) of two fuzzy sets A and B with respective membership functions $m_A(x)$ and $m_B(x)$ is a fuzzy set C, written as $C = A \cap B$, whose membership function is related to those of A and B by

$$\mu_C(x) = \min(\mu_A(x), \mu_B(x)),\ x \in X. \tag{2}$$

The union (Zadeh, 1965) of two fuzzy sets A and B with respective membership functions $m_A(x)$ and $m_B(x)$ is a fuzzy set C, written

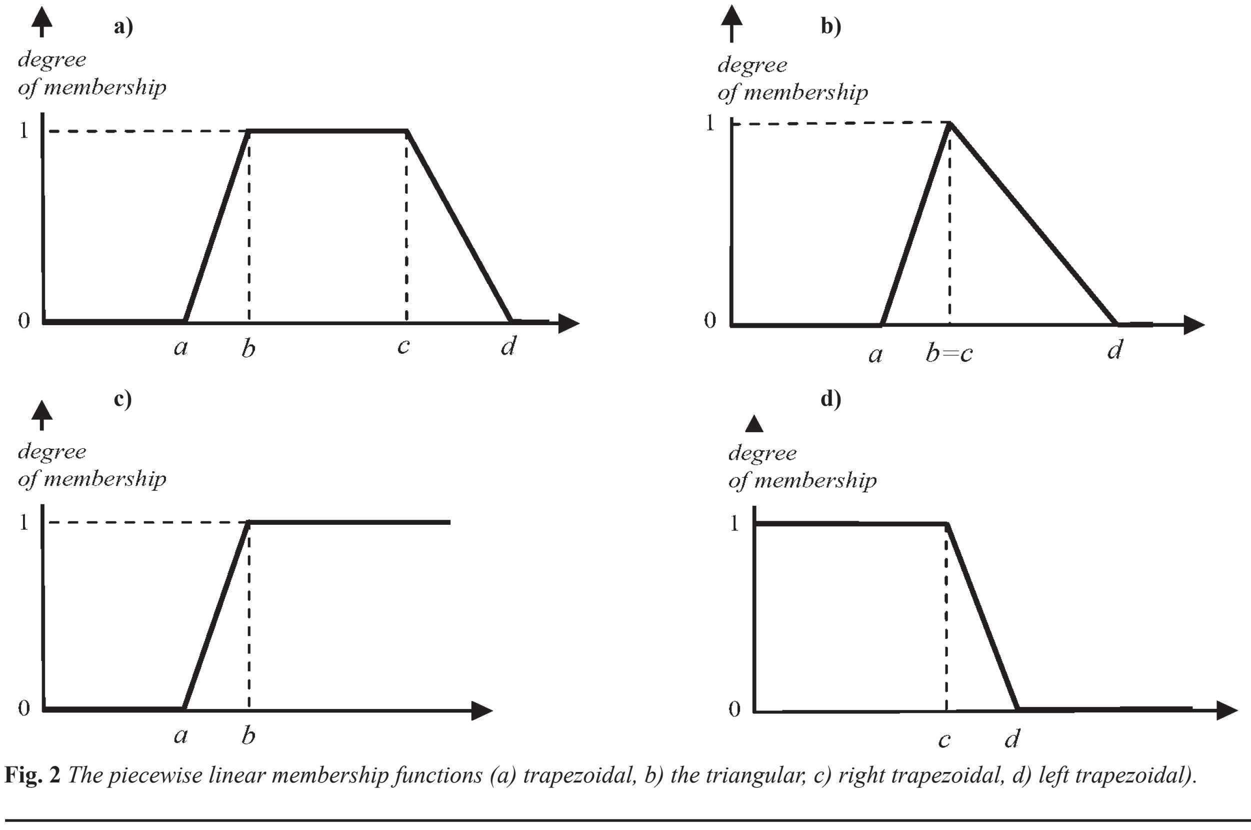

Fig. 2 *The piecewise linear membership functions (a) trapezoidal, b) the triangular, c) right trapezoidal, d) left trapezoidal).*

SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

as $C = A \cup B$, whose membership function is related to those of A and B by

$$\mu_C(x) = \max(\mu_A(x), \mu_B(x)),\ x \in X. \quad (3)$$

The complement (Zadeh, 1965) of a fuzzy set A is denoted by $\overline{A}$ and is defined by $\mu_{\overline{A}} = 1 - \mu_A(x)$.

The standard complement of fuzzy set A is then the fuzzy set $\overline{A}$ with the membership function $\mu_{\overline{A}}$.

The standard method of an creating intersection and union is shown in Fig. 3.

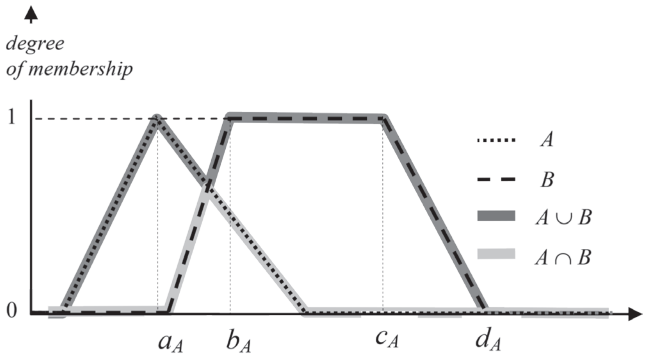

Fig. 3 *Fuzzy conjunction and fuzzy disjunction.*

In fuzzy logic, the other types of conjunction are usually interpreted by triangular norms (t-norms for short), and the other types of disjunction are interpreted by triangular conorms (t-conorms for short). A triangular norm (Mesiar and Navara, 1999) is a commutative, associative, non-decreasing function T: $[0,1]^2 \rightarrow [0,1]$ such that $T(x, 1) = x$ for all $x \in [0,1]$. The basic t-norms are (Grabisch et al., 2009):

the minimum t-norm (greatest conjunctor; it corresponds to the standard conjunction) (Fig. 4a):

$$T_M(x,y) = \min(x,y), \quad (4)$$

the product t-norm (Fig. 4b):

$$T_P(x,y) = xy, \quad (5)$$

the Łukasiewicz t-norm (Fig. 4c):

$$T_L(x, y) = \max(0, x + y - 1), \quad (6)$$

the drastic t-norm (smallest conjunctor) (Fig. 4 d):

$$T_D(x,y) = \begin{cases} \min(x,y), & \text{if } \max(x,y) = 1 \\ 0, & \text{else.} \end{cases} \quad (7)$$

An example of spatial analysis using Boolean logic compared to fuzzy logic using t-norms is presented, for example, in (Lieskovský et al., 2011) and (Ďuračiová et al., 2011).

3.2.2 Many-valued logic in spatial databases

Many-valued logic was introduced by the Polish logician and philosopher Łukasiewicz (1920) (Gottwald, 2010). The first step from bivalent logic to many-valued logic is three-valued logic. SQL has also implemented a three-valued logic. It includes the values TRUE, FALSE and UNKNOWN. SQL also knows the logical operations (conjunction, disjunction and negation), which are included in all programming languages. Their extension to three values is shown in Tab. 1. Three-valued logic is implemented in all currently used database systems.

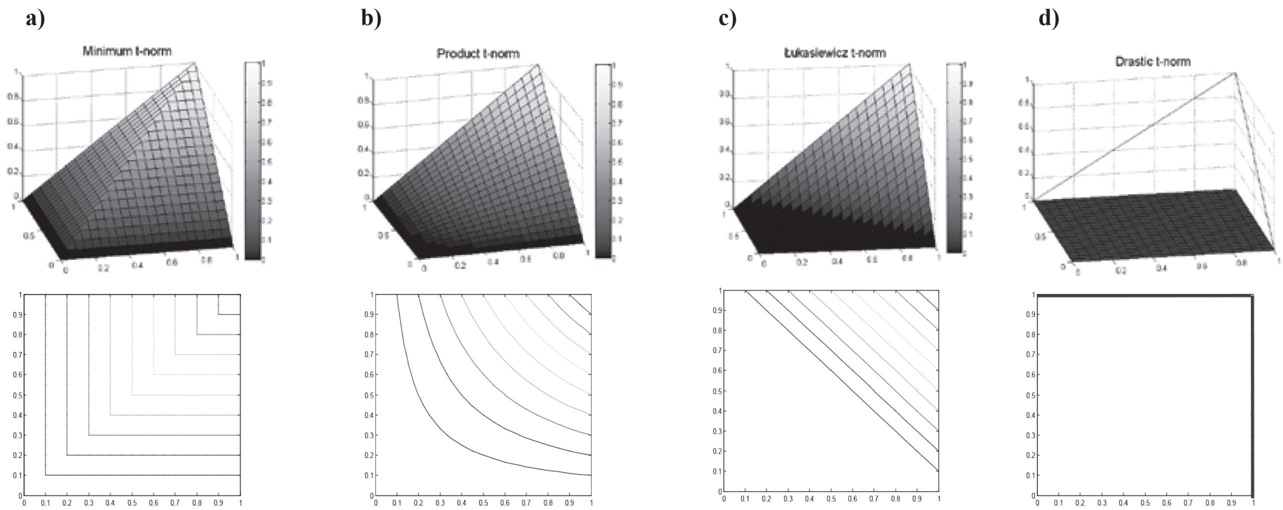

Fig. 4 *The basic t-norms a) minimum, b) product, c) Łukasiewicz, d) drastic.*

SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

An extension of three-valued logic is many-valued logic (e.g., if the truth value of UNKNOWN is 0.5 and, in addition to it, we admit the truth values of the propositions 0.75 and 0.25, in evaluating the truth-values of statements we use five-valued logic). Fuzzy logic is then a generalization of many-valued logic, which models truth values by fuzzy sets.

Tab. 1 *Three-valued logic in database systems.*

A	B	A and B (conjunction)	A or B (disjunction)	not A (negation)
TRUE	TRUE	TRUE	TRUE	FALSE
TRUE	FALSE	FALSE	TRUE	FALSE
TRUE	UNKNOWN	UNKNOWN	TRUE	FALSE
FALSE	TRUE	FALSE	TRUE	TRUE
FALSE	FALSE	FALSE	FALSE	TRUE
FALSE	UNKNOWN	FALSE	UNKNOWN	TRUE
UNKNOWN	TRUE	UNKNOWN	TRUE	UNKNOWN
UNKNOWN	FALSE	FALSE	UNKNOWN	UNKNOWN
UNKNOWN	UNKNOWN	UNKNOWN	UNKNOWN	UNKNOWN

3.2.3 Modeling of spatial data uncertainty in multicriteria decision making

There are two types of uncertainty in MCDM in spatial databases: data uncertainty and criteria uncertainty (Fig. 5).

Both data and criteria can be modelled by fuzzy sets in database systems. Uncertain information can be represented either by a special attribute (e.g., a mean square error) or by a fuzzy set. The spatial uncertainty of an object modeled by a piecewise linear membership function to a fuzzy set is shown in Fig. 6.

Various forms of the uncertainty of objects in a spatial database (e.g., uncertain topology, uncertain geometry and uncertain attributes) are modeled in the class diagram in Fig. 7.

For evaluating queries containing spatial operations such as an Intersection, Union, or Difference, it is necessary to know the topological relations of the spatial objects. For computing the spatial relationships between geometries, the Dimensionally Extended Nine-Intersection Model (DE-9IM) is specified in the OGC "Simple Features for SQL" specification. The DE-9IM considers the two objects' interiors (A^0, B^0), boundaries (∂A, ∂B) and exteriors (A^-, B^-) and analyzes the intersections of the parts of these nine objects for their relationships. The spatial relationships described by the DE-9IM are "*Equals*," "*Disjoint*," "*Intersects*," "*Touches*," "*Crosses*," "*Within*," "*Contains*," and "*Overlaps*." The DE-9IM is based on the Nine-Intersection Model (9IM) (Shekhar and Xiong, 2008). The 9IM is defined by matrix $I_9(A, B)$:

$$I_9(A,B) = \begin{pmatrix} A^o \cap B^o & A^o \cap \partial B & A^- \cap B^- \\ \partial A \cap B^o & \partial A \cap \partial B & \partial A \cap B^- \\ A^- \cap B^o & A^- \cap B^o & A^- \cap B^- \end{pmatrix}. \quad (8)$$

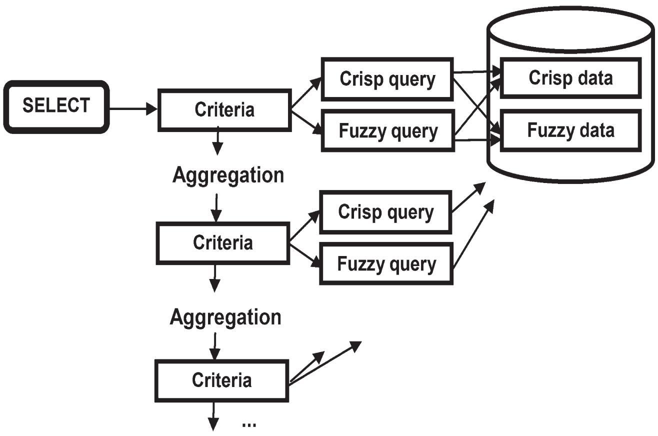

Fig. 5 *Fuzzy spatial queries in MCDM.*

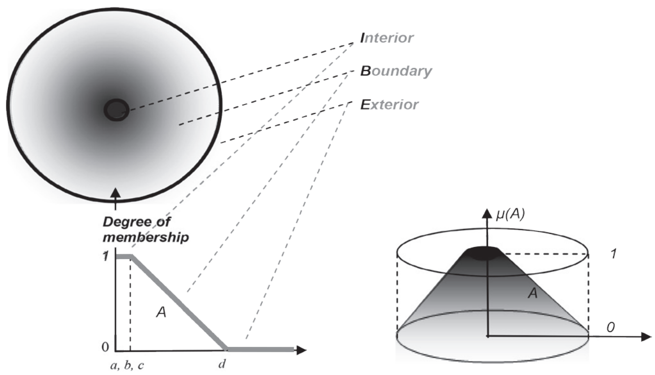

Fig. 6 *Uncertainty of a spatial object.*

SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

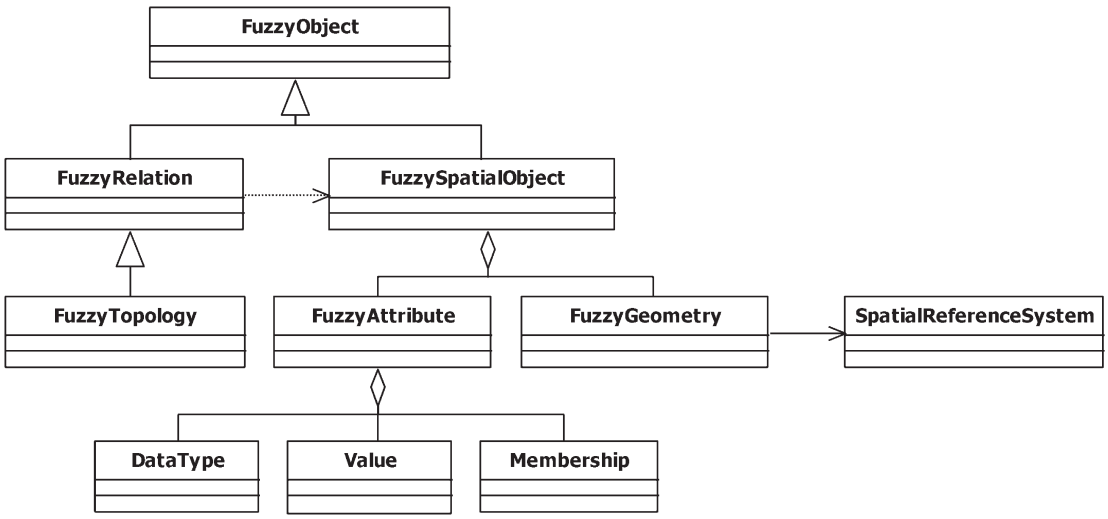

Fig. 7 *Fuzzy object in a spatial database.*

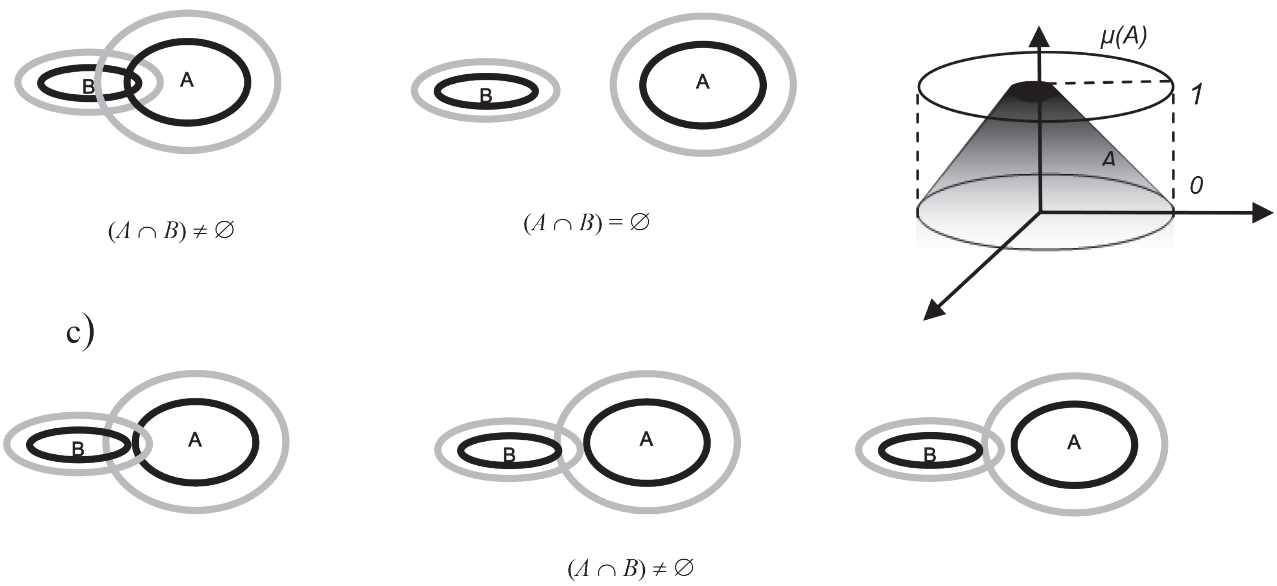

Fig. 8 *Intersection of two spatial objects with an uncertain boundary - a) definitely intersected, b) definitely not intersected, c) possibly intersected.*

If we are working with uncertain spatial objects, their topological relationships will be affected by uncertainty too. The intersection of two objects with a uncertain boundary is shown as an example in Fig. 8. Fuzzy spatial topology is discussed more, e.g., in (Tang, 2004), (Tang et al., 2006).

4 CASE STUDY

A fuzzy spatial query approach is demonstrated in a model situation related to a fundamental database for GIS in Slovakia. The database schema is defined according to the FACC DIGEST (*Digital Geographic Information Exchange Standard*) (ISO/TC 211 Secretariat, 2000) (Tab. 2).

SLOVAK JOURNAL
OF
CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

Tab. 2 *Description of selected features in a spatial database.*

Code	Name	Description
AL015	Building	A relatively permanent structure, roofed and usually walled and designed for some particular use.
AP030	Road	An open way maintained for vehicular use.
BH140	River/Stream	A natural flowing watercourse.
...		

The database schema contains the following attributes:
AL015 Building (EXS, HGT, HWT, BFC, NAM)
BH140 River/Stream (EXS, HOC, HYC, LOC, TUC, NAM, TXT)
AP030 Road (EXS, RDT, RST, LOC, SMC, RTN, WD2, NAM, TXT)
...

The attributes and their value codes are described in (ISO/TC 211 Secretariat, 2000). For example, the BFC code (*Building Function Category*) is defined as "*Type or purpose of the building.*" The HGT attribute (*Height Above Surface Level*) is described as follows: "*Distance measured from the lowest point of the base at ground or water level (downhill side/downstream side) to the tallest point of the feature*."

An MCDM using fuzzy sets in GIS is applied, for example, in this type of analysis:

Find all buildings, which:

- are tall (fuzzy criteria, crisp or fuzzy non-spatial data),
- are close to the river (fuzzy criteria, crisp or fuzzy spatial data),
- their use is: "house" (crisp criteria, crisp non-spatial data).

The appropriate examples of these queries are, e.g.:

1. select all houses from the layer "*Buildings*" (the code value of the BFC attribute for a house is 16 (according to the standard DIGEST)),
2. select all buildings that are within a distance of 35 m from the river,
3. select all moderately tall houses (modeled by the "*Moderately tall house*" fuzzy set),
4. select all buildings that are located near rivers (if the distance is less, the degree of membership in the fuzzy set "*Building near the river*" is higher),
5. select all buildings located on the floodplain of the river (the floodplain boundary is given with a degree of uncertainty).

Example 1 is a standard query without a need to support the spatial data. Example 2 shows the query with a spatial extension of the SQL. The query in example 3 contains uncertain attribute criteria. The query in example 4 describes the uncertain spatial criteria, and the query in example 5 contains uncertain data.

The membership function $m_H(x)$ of fuzzy set H "*Moderately tall building*" is, e.g. (Fig. 9 a):

$$\mu_H(x) = \begin{cases} 0 & \text{if } x < 5, \\ \dfrac{x-5}{6} & \text{if } 5 \le x \le 11, \\ 1 & \text{if } x > 11. \end{cases} \tag{9}$$

The membership function $m_C(x)$ of fuzzy set C "*Close to the river*" is (Fig. 9 b):

$$\mu_C(x) = \begin{cases} 1 & \text{if } x < 20, \\ \dfrac{35-x}{15} & \text{if } 20 \le x \le 35, \\ 0 & \text{if } x > 35. \end{cases} \tag{10}$$

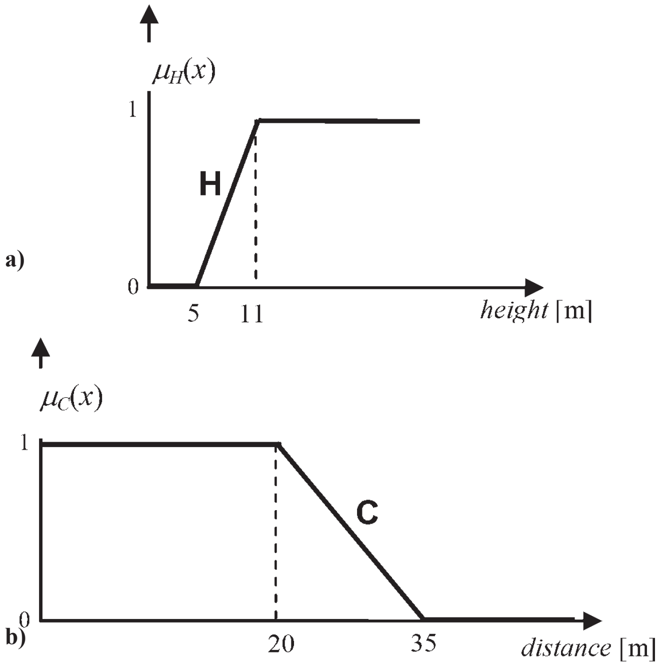

Fig. 9 *a) Fuzzy set H "Moderately tall building", b) Fuzzy set C "Close to the river".*

5 RESULTS

The result of the application of the fuzzy approach is the creation of a fuzzy spatial SQL query to the object-relational database. The resulting query can be composed of a fuzzy query and a crisp query, both of which can be used to select spatial or non-spatial data from the database. The UML activity diagram for creating a simple uncertain spatial query is shown in Fig. 10.

Vol. XXI, 2013, No. 4, 1 – 12

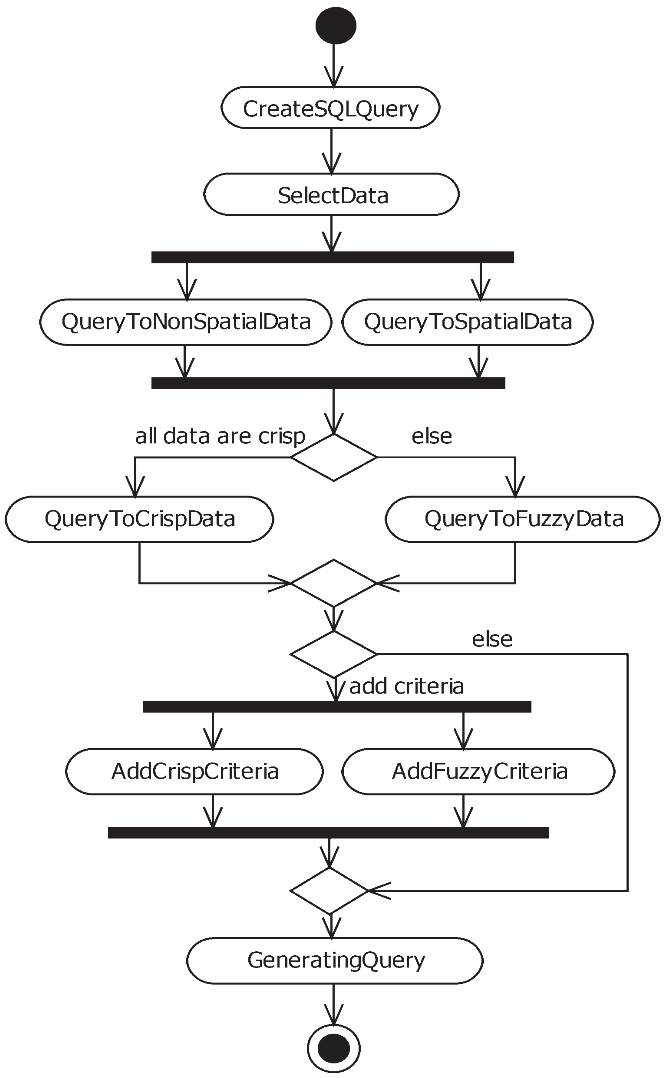

Fig. 10 *UML activity diagram of uncertain spatial querying to object-relational databases.*

For example the SQL query for MCDM considering criteria_0, criteria_1, criteria_2,... criteria_k is as follows:

SELECT attribute_1, attribute_2,..., attribute_n
FROM Table_1, Table_2,...Table_t
WHERE criteria_0 AO_1 criteria_1 AO_2 criteria_2... AO_k criteria_k;

AO_1, AO_2, ... AO_k in the SQL command are aggregation operators for criteria_1, criteria_2,... criteria_k. The most frequently used aggregation operator in this type of problem is AND. This means that all the crisp criteria must be met.

The SQL query for fuzzy MCDM is then:
SELECT attribute_1, attribute_2,... attribute_n, mi_1, mi_2,... mi_m,
FROM Table_1, Table_2,...Table_t
WHERE f_criteria_0 T-NORM f_criteria_1 T-NORM f_criteria_2... T-NORM f_criteria_k;

The attributes mi_1, mi_2,...mi_m are the degrees of membership in fuzzy sets of the at least partially suitable results (for each t-norm). The fuzzy criteria are expressed by f_criteria_0, f_criteria_1,... f_criteria_k, and T-NORM is a selected t-norm used as an aggregation operator. Both the fuzzy sets and t-norms are not defined in the standard SQL. A suitable solution of this absence can be the standard SQL query with the command CASE in the construction SELECT – FROM – WHERE. Then the SQL command for the model situation from the case study above, including criteria 1, 3 and 4 is as follows:

```
SELECT id, mi_1, mi_2, mi_3,
    LEAST(mi_1, mi_2, mi_3) AS t_M,
    (mi_1*mi_2*mi_3) AS t_P,
    CASE
        WHEN (mi_1 + mi_2 + mi_3 - 2) <= 0 THEN 0
        ELSE (mi_1 + mi_2 + mi_3 - 2)
    END AS t_L
FROM
(
SELECT B.id,
    (CASE
        WHEN B.hgt > 11 THEN 1
        WHEN B.hgt < 5 THEN 0
        ELSE (B.hgt - 5)/6 END)
    AS mi_1,
    (CASE
        WHEN distance(B.the_geom, R.the_geom) > 35 THEN 0
        WHEN distance(B.the_geom, R.the_geom) < 20 THEN 1
        ELSE (35 - distance(B.the_geom, R.the_geom))/15 END)
    AS mi_2,
    (CASE
        WHEN B.bfc = 16 THEN 1
        ELSE 0 END)
    AS mi_3
FROM Buildings AS B, Rivers AS R
) AS Fuzzy_Buildings
WHERE t_M > 0;
```

The spatial function *distance* as an extension of the standard SQL is applied to calculate the distance in the spatial criteria 4 (*building close to the river*).

The query above is logically correct and is applicable in all similar situations, but in the case of the lower performance of database systems, a query with the functions GREATEST and LEAST could be a more suitable SQL command:

SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

```
SELECT id, mi_1, mi_2,
    LEAST(mi_1, mi_2) AS t_M,
    (mi_1*mi_2) AS t_P,
    GREATEST((mi_1 + mi_2 - 1),0) AS t_L
FROM
(
SELECT B.id,
    LEAST(GREATEST((B.hgt - 5)/6,0),1) AS mi_1,
    LEAST(GREATEST((35 - distance(B.the_geom,
                              R.the_geom))/15,0),1) AS mi_2
FROM Buildings AS B, Rivers AS R
) AS Fuzzy_Buildings
WHERE B.bfc = 16 AND t_M > 0;
```

(Functions GREATEST (returns the greatest value in a list of expressions) and LEAST (least value) are not part of the SQL standard, but they are very helpful for fuzzy SQL querying.)

The result of the queries is the table: Fuzzy_buildings(id, mi_1, mi_2, mi_3, t_M, t_P, t_L, the_geom), where mi_1, mi_2, mi_3 are the degrees of membership in the fuzzy sets of criteria_1 (*moderately tall building*), criteria_2 (*building close to the river*) and criteria_3 (house). The attributes t_M, t_P, t_L are the values of the degrees of membership in the fuzzy set of the suitable results aggregated by the relevant t-norms (t_M - minimum, t_P – product, t_L – Łukasiewicz). If t_M > 0 then t_L > 0 and t_P > 0 because of (Grabisch et al., 2009):

$$T_D(x,y) \le T_L(x,y) \le T_P(x,y) \le T_M(x,y) \tag{11}$$

This implies that the condition *t_M* > 0 in clause WHERE is sufficient.

The values t_M, t_P, t_L are very useful as decision support for users (e.g., a decision on the suitability of an object in the selection).

The fuzzy SQL queries were applied in PostgreSQL (http://www.postgresql.org), which is the most widely-used open source

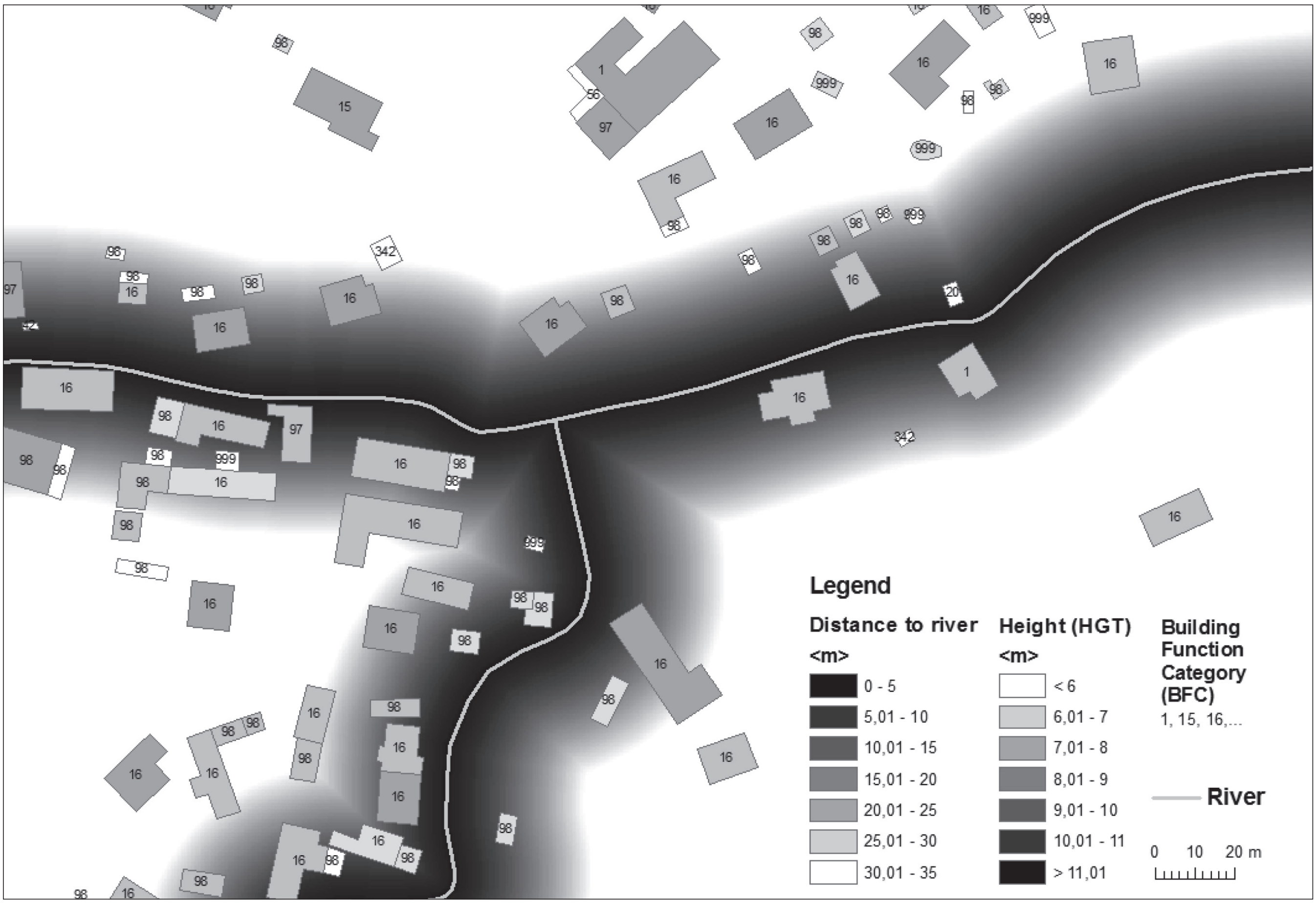

Fig. 11 *Visualization of the model situation (the displayed attribute of each building is a BFC value).*

SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

database system for storing and managing spatial data. A visualization of the case study in the ArcGIS 10.1software environment is illustrated in Fig. 11.

6 CONCLUSIONS

Spatial databases are useful in conventional (non-geographic) information systems (e.g., multimedia systems, building information systems, space systems in astronomy, etc.), but the dominant areas for their use are geographic information systems. In many geographical applications and much spatial analysis, there is a need to use heterogeneous data of various qualities and model their uncertainty. The use and integration of multi-source heterogeneous data is currently becoming more common. This fact is related mainly to the increasing number of Web services (especially the Web Map Service (WMS) and the Web Feature Service (WFS)) and spatial data sources provided in the web environment. The application of fuzzy modeling principles in spatial databases brings the possibility of the efficient modeling of data quality and processing of uncertain data. Fuzzy spatial querying enables a more effective data selection, and it is also a suitable approach for the integration of multi-source data.

The basis of the proposed solution is the query language SQL, which is implemented in all of the most commonly used relational and object-relational database systems. The limitation of SQL is that it makes only crisp selections (only Boolean logic). That means, for example, that the record in a database (or a spatial object) might not be selected even if it is extremely close to the conditions of a query (e.g., the difference is much less than the value of the data's uncertainty). This situation is avoided by the application of the fuzzy logic. The greatest advantage of the fuzzy approach presented in this paper is that it does not require the implementation of any other query language for querying uncertain data. This solution can therefore easily become the basis of functionality in the development of an application for querying uncertain spatial data. The implementation of a user-friendly interface for fuzzy spatial querying and decision-making criteria modeling is the subject of our current research.

REFERENCES

Bosc, P. - Kraft, D. - Petry, F. (2005) *Fuzzy sets in database and information systems: Status and opportunities*. Fuzzy Sets and Systems 156, 2005, pp. 418–426.

Burrough, P. - Frank, A. (Eds.) (1996) *Geographic Object with Indeterminate Boundaries*. GISDATA Series, vol. 2, London, UK, Taylor and Francis, 352 pp.

Cuevas, L. - Marín, N. - Pons, O. - Vila, M. A. (2008) *pg4DB: A fuzzy object-relational system.* Fuzzy Sets and Systems 159, 2008, pp. 1500–1514.

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) 2007 [cit. 2012-12-03]. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF.

Ďuračiová, R. - Lieskovský, T. - Kročková, K. - Sabo, M. (2011) *Multikriteriálne rozhodovanie pomocou fuzzy množín v prostredí GIS a jeho využitie v archeologickej predikcii (Multicriteria decision making using fuzzy sets in a GIS environment and its application in archeological predictions)*. Geodetický a kartografický obzor 57, 9, pp. 205-215 (in Slovak).

Grabisch, M. - Marichal, J.-L. Mesiar, R. - Pap, E. (2009) *Aggregation Functions. Encyclopedia of Mathematics and its Applications, No. 127.* Cambridge: Cambridge University Press, 460 pp. ISBN:978-0-521-51926-7

Gottwald, S. (2010) *Many-Valued Logic*, **In: Zalta, E. N.** (ed.). *The Stanford Encyclopedia of Philosophy*. Springer 2010 Edition. [online] [cit. 2013-01-14]. Available at: http://plato.stanford.edu/archives/spr2010/entries/logic-manyvalued/.

Hudec, M. (2008) *Fuzzy Structured Query Language (SQL) for Statistical Database.* **In: Meeting on the Management of Statistical Information Systems** (MSIS2008). Luxembourg, 7-9 April 2008. [online] [cit. 2013-01-03]. Available at: http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.50/2008/wp.12.e.pdf.

Hudec, M. (2009) *An Approach to Fuzzy Database Querying Analysis and Realisation*. ComSIS, Vol. 6, No. 2, 2009, pp. 127–140.

Hudec, M. - Büchler, P. - Ružovičová, Z. (2012) *Výskum, aplikovanie a šírenie fuzzy prístupu pre spracovanie a vyhodnocovanie štatistických dát (Research, dissemination and application of fuzzy approach for processing and evaluation of statistical data)*. [online] [cit. 2012-12-03]. Available at: http://www.infostat.sk/cevavstat/fuzzy/prispevky/MH-T-CEVAVSTAT0401.pdf. (in Slovak).

Chen, D. W. (2012) *Supporting Uncertainty in Standard Database Management Systems. Master*'s *thesis*, Montreal, Quebec: Concordia University, 101 pp.

ISO/TC 211 Secretariat (2002) *ISO 19113:2002 Geographic information – Quality Principles*. Oslo: ISO.

ISO/TC 211 Secretariat (2003a) ISO *19114:2005 Geographic information - Quality evaluation procedures*. Oslo: ISO.

SLOVAK JOURNAL OF CIVIL ENGINEERING

Vol. XXI, 2013, No. 4, 1 – 12

ISO/TC 211 Secretariat (2003b) *ISO19115:2003 Geographic information – Metadata*. Oslo: ISO.

ISO/TC 211 Secretariat (2000) *New work item proposal: Geographic information - Profile - FACC DataDictionary*. Oslo: ISO.

Kacprzyk, J. - Pasi, G. - Vojtáš, P. - Zadrożny, S. (2000) *Fuzzy Querying: Issues and perspectives*. Kybernetika, Vol. 36, No. 6, pp. 605-616.

Kainz, W. (2013) *Fuzzy Logic and GIS*. [online] [cit. 2013-01-14]. Available at: http://www.scribd.com/doc/44096805/Fuzzy-in-GIS-Basic.

Koreň, M. (2009) *Databázové systémy (Database systems)*. Zvolen: Technická univerzita vo Zvolene Lesnícka fakulta, 90 pp. ISBN 978-80-228-2084-4 (in Slovak).

Liang, H. - Hu, Sh. - Qiao, Ch. (2013) *Study of Fuzzy Uncertainty of GIS Product.* [online] [cit. 2013-01-14]. Available at: http://www.commission4.isprs.org/workshop_hangzhou/papers/225-230%20Hongyou%20LIANG-A087.pdf.

Lieskovský, T. - Faixová Chalachanová, J. - Ďuračiová, R. - Blažová, E. (2011) *Archeologické predikčné modelovanie z pohľadu geoinformatiky (Archaeological predictive modeling from the perspective of geoinformatics)*. Bratislava: REMPrint, s.r.o., 187 pp. ISBN 978-80-970741-0-4 (in Slovak).

Longley, P. A. et al. (1999) *Geographical Information Systems. Vols. 1, 2. 2nd Ed*. New York: John Wiley and Sons, 1170 pp. ISBN 0-471-32182-6

Mesiar, R. - Navara, M. (1999) *Diagonals of continuous triangular norms*. Fuzzy Sets and Systems, Vol. 104, No 1, pp. 35–41.

Morris, A. - Jankowski, P. (2005) *Spatial Decision Making Using Fuzzy GIS*. **In: Petry, F. E., – Robinson, V. B. – Cobb, M. A.** (eds.) *Fuzzy Modeling with Spatial Information for Geographic Problems*. Berlin, Heidelberg: Springer-Verlag, 344 pp. ISBN 3-540-23713-5

Mukherjee, I. - Ghosh, S. K. (2011) *Fuzzy Spatial Data Modeling for Integration of Heterogeneous Geospatial Information.* [online] [cit. 2013-03-13]. Available at: http://www.geospatialworldforum.org/2011/proceeding/pdf/GhoshFullPaper.pdf.

OGC Inc. (2011) *OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture*. [online] [cit. 2012-11-06]. Available at: http://www.opengeospatial.org/standards/sfa.

OGC Inc. (2010) *OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option*. [online] [cit. 2012-11-06]. Available at: http://www.opengeospatial.org/standards/sfs.

Petry, F. E. - Robinson, V. B. - Cobb, M. A. (2005) *Fuzzy Modeling with Spatial Information for Geographic Problems*. Berlin, Heidelberg: Springer-Verlag, 344 pp. ISBN 3-540-23713-5

Robinson, V. (1990) *Interactive machine acquisition of fuzzy spatial relation*. Computers and Geosciences 6, pp. 857–872.

Shekhar, S. - Xiong, H. (2008) *Encyclopedia of GIS*. Springer, 1370 pp. ISBN 978-0-387-30858-6

Schneider, M. (2008) *Fuzzy Spatial Data Types for Spatial Uncertainty Management in Databases. Handbook of Research on Fuzzy Information Processing in Databases*. Information Science Reference, pp. 490-515. [online] [cit. 2013-03-14]. Available at: http://www.cise.ufl.edu/~mschneid/Research/papers/Sch08BoCh.pdf.

Tang, X. (2004) *Spatial object modeling in fuzzy topological spaces with Applications to Land Cover Change. ITC Dissertation No. 108.* Enschede: ITC Printing Department, 219 pp. ISBN 90-6164-220-5

Tang, X. -Fang, Y. - Kainz. W. (2006) *Fuzzy Topological Relations Between Fuzzy Spatial Objects.* **In: Wang, L. et al.** (eds.): *FSKD 2006 LNAI 4223*, Berlin, Heidelberg: Springer-Verlag. pp. 324-333.

Xiaoxiang, Z. - Jing, Y. - Manchun, L. (2005) *Mechanism and Methods Of Fuzzy Geographical Object Modeling*. ISPRS Archives – Vol. XXXVI-2/W25, 2005, WG II/1 Proceedings of International Symposium on Spatio-temporal Modeling, Spatial Reasoning, Analysis, Data Mining and Data Fusion. 27-29 Aug, Beijing, China. 6 pp.

Zadeh, L. (1965) *Fuzzy Sets*. Information and Control, No. 8, pp. 338-353.

Internet sources:

OGC – http://www.geospatial.org

PostgreSQL – http://www.postgresql.org/