Mathematical Model Used in Decision-Making Process with Respect to the Reliability of Geodatabase

Šárka Hošková-Mayerová[1], Václav Talhofer[2], Alois Hofmann[3], and Petr Kubíček[4]

[1] Department of Mathematics and Physics, University of Defence, Faculty of Military Technology, Kounicova 65, 662 10 Brno, Czech Republic
sarka.mayerova@unob.cz
[2,3] Department of Military Geography and Meteorology, University of Defence, Faculty of Military Technology, Kounicova 65, 662 10 Brno, Czech Republic
{vaclav.talhofer,alois.hofmann}@unob.cz
[4] Department of Geography, Masaryk University, Faculty of Science, Kotlarska 2, 611 37 Brno, Czech Republic
kubicek@geogr.muni.cz

Abstract. The first aim of the article is to show how it is possible - thanks to the use of sophisticated analytical tools for evaluation of data quality - to better understand geospatial data. Another aim is to assess the impact of data quality on the results of space analyses that are made of them and that are the basis for such decision-making processes, in which it is necessary to take into account the impact of geographical environment.

Organizations that are engaged in creating geospatial databases usually define the content of these databases (i.e. listing of geographical objects and their features) and quality of the data being saved (e.g. geometric, topological and thematic accuracy, level of standardization etc.). As the area of the land that is described with the use of geospatial data is usually significantly larger than the capacity and technological possibilities of the responsible organization, it is not possible to keep the defined content and its quality in the entire secured area on the same level. When creating the geospatial analysis it is therefore necessary to take into account the immediate quality level of data in the particular area and to have the technologies for finding out the reliability of the result of the particular analysis available. From the real practice a request of commanders is known, that is to have not only the result of their own analysis available as basics for their qualified decision (decision-making process) but also relevant information about its reliability.

The authors of the article have quite good experience from the preparation of digital geospatial data for decision-making processes in the armed forces and within the scope of the described research they have available a large quantity of real geospatial data (current as well as historical), on which they are doing their own research focused on the mathematical modeling and evaluation.

Keywords: reliability, decision making process, mathematical modelling, geospatial data, GIS, quality assessment, utility value.

A.N. Proto et al. (Eds.): Adv. Dynamic Modeling of Economic & Social Systems, SCI 448, pp. 143–162.
springerlink.com

1 Maps and Geospatial Data: The Base for Command and Control

1.1 Command and Control in a Historical Context

Command and control was in the past connected especially to the activities of armed forces. The system of command and operation management in a certain territory in history of mankind changed depending on used weapon technologies. If there were especially cold weapons used, the decisive activities leading to victory or defeat took place in several hours or maximum within a few days. Direct command on the battlefield was then realized with the help of messengers, signals and pre-agreed procedures. At the same time the commander had the overview of the whole battlefield and could make decisions how to proceed.

With the development of weapons and weapons systems and with the creation of big regular armies, changes have been made to command and control systems. The commander then lost the possibility to watch the battlefield directly and to control several-thousand soldiers in the combat units. That is why headquarters staffs who served the commander to support his decisions were formed. The coordination of the staff then necessarily required working on common base. As the military activities very often took place in the landscape, maps, especially *topographic maps*, were as one of the most important base more and more used for command and operation management. They were used for planning the future actions as well as for operational solution of a current situation during a fighting. Their content corresponded to commander and soldier requirements and maps allowed to orientate oneself in the field and to control the fighting activities as well.

Working procedures including the system of using detailed maps for planning and operation management transformed piecemeal also into the civil affairs, especially into fields where it is necessary for a higher number of units to cooperate, e.g. fire brigade, police, etc.

Work of the headquarters and its system developed depending on ways of command as well as on technical and technological development of military technical equipment and used sources. From the point of view of using the information about the territory, digitization of the landscape has slowly started. There is a new expression - *electronic battlefield*. Within the scope of electronization vast infrastructures of space data - that to a certain extent substitute classical paper maps - are currently being created.

1.2 Common Operational Picture and Recognized Environmental Picture

As a basic working environment of staff work in an electronic environment of communication and information systems is a *Common Operational Picture (COP)*, which all authorized personnel of headquarters share and which creates a united platform for solutions of specialized tasks and cooperation of individual units. Basic topographic data are usually put into this system with the help of *Web Mapping Services (WMS*). COP contents visualized picture of landscape which to certain extent corresponds to classic topographic maps.

Geographic units, except providing *basic topographic database (geospatial database)*, create a lot of *geospatial analyses* to support the decision-making process, while using other information and data, that are not a regular part of COP. Results of those analyses are presented as a *Recognized Environmental Picture (REP)*. Examples of these analyses can be for example the calculation of visibility and hidden spaces, calculation of steepness of slopes, analysis of immediate conditions of movement of technical equipment in free space on roads as well as off roads, analysis of climatic conditions with prediction of development of several hours or a day ahead with respect to the possibility of used technical equipment, etc. REPs are created based on standard offer within the frame of geographic support system or they are created based on direct order of the commander. In the next picture there is presentation of one possibility of *Communication and Information Technologies (CIT)* application for COP and REP.

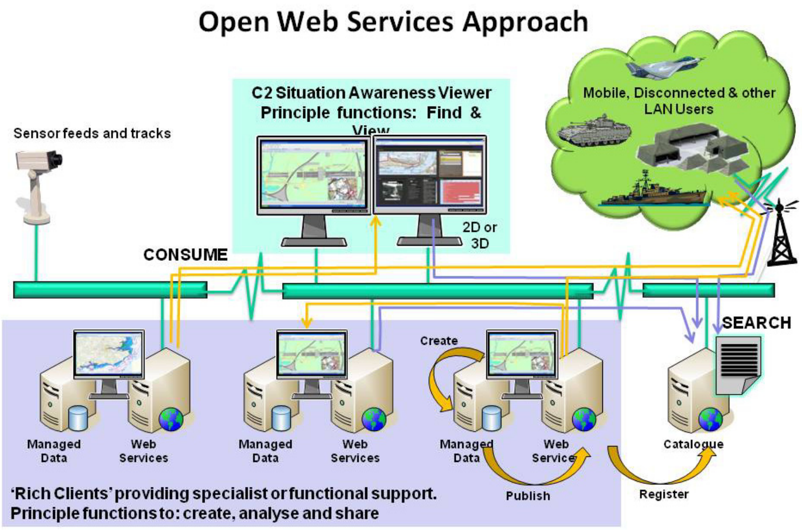

Fig. 1 Application of CIT in a command and control system (source (Tate, 2010))

As source data for REP are used data that are a product of geographic services and about which all necessary information is known, as well as other sources originating from outside the department of defense. In many tasks all the source data are mutually combined and based on mathematically or procedurally described processes new data are created. The user of source and newly created data should always, besides the space information, get also information about its quality. In case of source primary data, this fact is usually provided by the producer. However, if new data are being created, it is necessary to have tools so that it was possible to determine the quality from the relevant quality information of the source data.

2 Quality of Geospatial Data

Systems of data and information evaluation that are positionally determined vary depending if technical parameters of data or technological influences when collecting them are evaluated, or if their final utility value given by the information quality is assessed. According to recommendation of ISO 19113, measurable qualities of provided data are evaluated as components of quality, as it is shown in the following table.

Table 1 Elements and subelements of data quality according to ISO 19113

Element of quality	Subelement of quality
Completeness	Commission
	Omission
Logical consistency	Conceptual consistency
	Domain consistency
	Format consistency
	Topological consistency
Positional accuracy	Absolute or external accuracy
	Relative or internal accuracy
	Gridded data position accuracy
Temporal accuracy	Accuracy of a time measurement
	Temporal consistency
	Temporal validity
Thematic accuracy	Classification correctness
	Non-quantitative attribute correctness
	Quantitative attribute accuracy

As technical parameters of data, the accuracy of positional information can be evaluated, it is often given as a standard positional error or a standard error in individual coordinate axes. The accuracy of thematic information is also evaluated this way. In this respect, however, the evaluation itself can be more complicated because thematic information may be various (e.g. characteristics of construction material given by the selection of one value from offered options in contrast to setting number of inhabitants of a certain settlement, etc.) and also because not all evaluated parameters must always be completed with the given object.

Generally, when formulating a problem of quality evaluation of spatial data and resulting geospatial information it is necessary to follow the recommendation of international organizations, such as ISO, OGC and DIGIWG, which pursue the development of geoinformatics, and recommendation of direction INSPIRE in the long term (Konečný, Kubíček, & Staněk, 1998). These organizations and consortiums develop systems for quality evaluation. For instance, according to Guidebook for Implementation of Standards ISO 19100 for Quality Evaluation of Geographical Information (Jacobsson & Giversen, 2007) it is necessary to evaluate quality as a complex production as well as user's problem (see Fig. 2)

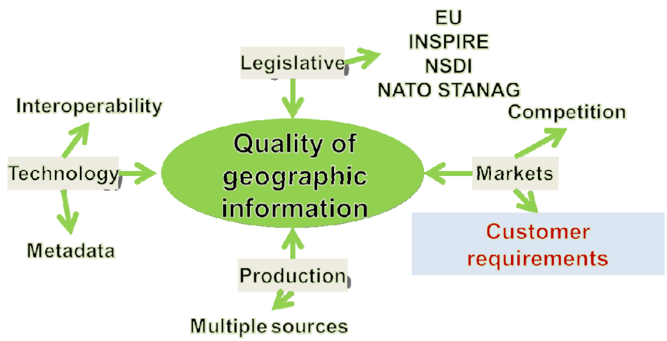

Fig. 2 Reasons for introduction of standards for quality evaluation of geospace information - according to (Jacobsson & Giversen, 2007)

The quality evaluation itself comes from general scheme components of quality, where production-technological aspects are evaluated, as well as operational and safety aspects, and in relation to a specific use of the product or service also aspects of reliability (see Fig. 3).

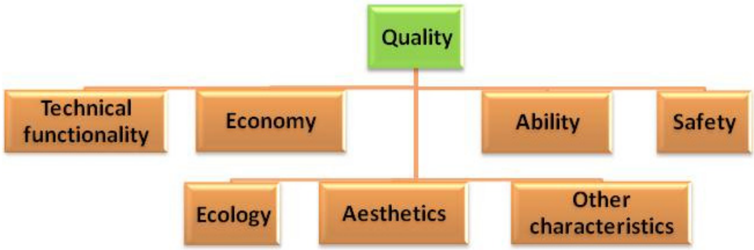

Fig. 3 General components of quality evaluation

While technical and technological parameters are possible to be evaluated generally without the necessity to know a given task, procedure or use of spatial information, reliability is necessary to be evaluated - with respect to the specific use - in a concrete process. The following text briefly explains how is possible to work with geospatial data quality in geospatial analysis.

3 Geospatial Analyses

Spatial geographic analyses (geospatial analyses) are mostly created based on a demand of the commander who wants to know specific current impacts of geographical environment and current weather conditions on his intended or planned activity, perhaps also stated impacts on the current activity in progress. Geographers and meteorologists are responsible for preparing these kinds of analyses. Most analyses that are created are usually provided to commanders, without telling them any information about the quality of the source data and their impact on reliability of the resulting analysis. This situation is documented in the following picture.

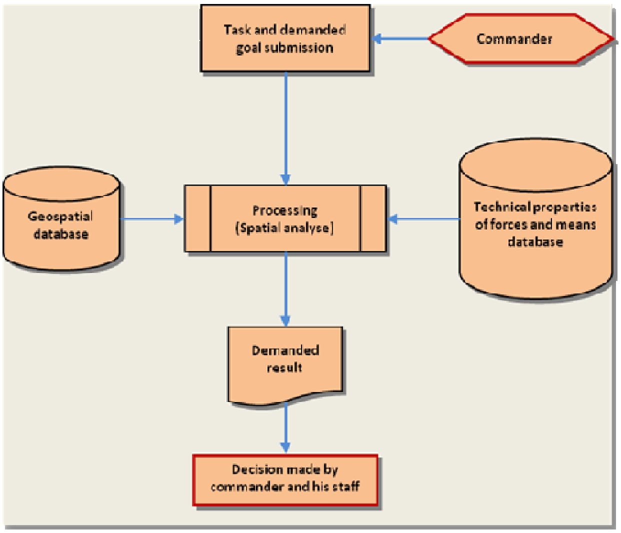

Fig. 4 Usual process of spatial analyses for the decision making process

This usual procedure has several advantages and disadvantages. The advantages are:

- relatively simple solution of the task,
- usually unambiguous result,
- commander does not have to think about the reliability characteristics of the obtained result.

Geospatial analyses are always created according to mathematical models or created schemes and given spatial data are used for them. Parameters of data quality can significantly affect the obtained results and that is why the stated procedure of geospatial analyses has certain disadvantages:

- total dependence on geospatial data,
- quality of the analysis result is not known,
- without additional information about the quality of required result the commander is left to use only one solution and has no possibility of choice among various options.

So that the above-stated disadvantages were minimized, we suggest involving also quality evaluation of source data into the geospatial analyses results.

If also level of geospatial data quality is evaluated in the process of geospatial analysis, it is possible to create several solutions. Geographer-analyst can prepare more options of solutions according to how he/she deals with quality characteristics. It is possible to create reliable, less reliable or very risky options. The commander then gets not only the result itself but also other information of riskiness of obtained solutions and it is then his responsibility what steps to take. He is either satisfied with all information he has received or he asks the geographer to increase the quality of spatial data and to prepare a new version of the space analysis. Certainly the appropriate conditions must be created to be able to increase the quality - time, personal as well as materialistic. The following picture illustrates this procedure.

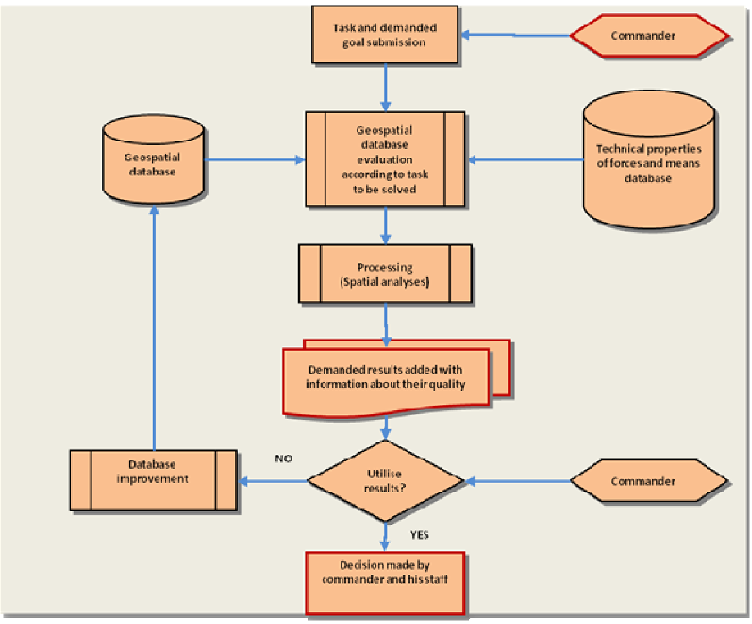

Fig. 5 Process of geospatial analyses for decision making process with the data quality consideration

4 Geospatial Data Quality Control

The process illustrated in the Fig. 5 assumes to have any complex system for level of spatial data quality evaluation which enables us not only determinate level of data quality but also to express a possibility how to increase the quality and what costs should be expect for this improvement.

4.1 Geospatial Database Utility Value Evaluation

Five essential criteria imply from geospatial data quality review (Talhofer & Hofmann, 2009). Their assessment gives the baseline for relatively reliable determination of each product utility value:

- *Database content* expresses mostly compliance of its definition and users' needs, i.e. concord of the "real modelled world" and its model represented by objects and phenomena stored in the database.
- *Database quality* defines the quality of stored data.
- *Database timeliness* explains how frequently the entire database or its elements updated is.
- *Importance of the area* is determined by users' needs so that it meets the requirements of processed or supported area range.
- *User friendliness.* This criterion defines data usability in various software environment types of GIS nature reflected mostly in compliance to standard principles. By the help of this criterion data standardization, independence, and security is revised.

Each of the criteria is mathematically assessable through independent tests and can be described as a quality parameter. In the next table there is a list of all used criteria.

4.2 General Assessment of Geospatial Data Utility

The product or a part of the product resultant function utility value may be assessed based on the above mentioned criteria using a suitable aggregation function (Talhofer, Hoskova, Hofmann, & Kratochvil, 2009):

$$F = p_3k_3p_4k_4(p_1k_1 + p_2k_2 + p_5k_5) \quad (1)$$

The value of F function expresses the degree of usability of geospatial database for a given task. The chosen form of the aggregation function concerns also the case if the user gets to obsolete data or data are from an area beyond his interest so that their usage could seriously affect or even disable the geospatial data functions. Therefore k_3 and k_4 are multiple terms in the formula (1). The weight of each criterion is marked as p_i, where $i = 1,\ldots,5$. The mentioned aggregation function proves the product status at the questioned instant and its utility rate. It is applicable also to experiments to find the ways of how to increase product utility at minimum cost increment.

Table 2 List of criteria for the geospatial geodatabase utility value evaluation

<table>
<tr><th>Main characteristics - main criteria</th><th>Sub-criteria characteristics</th><th>Definition</th><th>Quality parameter</th></tr>
<tr><td rowspan="2">Data model content – k_1</td><td>Complexity of conceptual landscape model</td><td>Concord of conceptual model and user requirements.</td><td>Percentage of incomplete information - k_{11}</td></tr>
<tr><td>Compliance of required resolution of geometric and thematic data – k_{12}</td><td>Concord of required geometric and thematic resolution.</td><td>Percentage of objects without required level of geometric and thematic resolution – k_{121}, k_{122}</td></tr>
<tr><td rowspan="12">Technical functionality – k_2</td><td rowspan="2">Transparency of data sources and methods for secondary data derivation - k_{21}</td><td>Transparency of source materials on primary data collection</td><td>Level of availability of information about used sources - k_{211}</td></tr>
<tr><td>Transparency of used methods and model for secondary data derivation</td><td>Level of availability of information about used methods – k_{212}</td></tr>
<tr><td rowspan="2">Position accuracy – k_{22}</td><td>Compliance with declared horizontal accuracy</td><td>Percentage of objects with unsatisfied conditions of declared horizontal accuracy – k_{221}</td></tr>
<tr><td>Compliance with declared vertical accuracy</td><td>Percentage of objects with unsatisfied conditions of declared vertical accuracy – k_{222}</td></tr>
<tr><td>Thematic accuracy – k_{23}</td><td>Compliance with declared accuracy of thematic data</td><td>Percentage of objects with unsatisfied conditions of declared thematic accuracy – k_{23}</td></tr>
<tr><td rowspan="3">Logical consistency – k_{24}</td><td rowspan="3">Degree of adherence of geographic data (data structure, their features, attributes and relationships) to the models and schemas (conceptual model, conceptual schema, application schema and data model)</td><td>Percentage of objects with topological inconsistence – k_{241}</td></tr>
<tr><td>Percentage of objects with thematic inconsistence – k_{242}</td></tr>
<tr><td>Percentage of objects with time inconsistence – k_{243}</td></tr>
<tr><td rowspan="2">Data completeness – k_{25}</td><td rowspan="2">Degree of adherence of the entirety of geographic data (features, their attributes and relationships) to the entirety of the modelled part of landscape</td><td>Percentage of missing objects or objects there are surplus – k_{251}</td></tr>
<tr><td>Percentage of incomplete thematic properties of objects – k_{252}</td></tr>
<tr><td rowspan="3">Database timeliness – k_3</td><td rowspan="3">Degree of adherence geographic data to the time changes in the landscape – k_3</td><td rowspan="3">Value of the time function describing process of the landscape changing</td><td>Number of changes</td></tr>
<tr><td>Time since the last up-date</td></tr>
<tr><td>…</td></tr>
<tr><td>Landscape importance –</td><td>Value of inverse distance to objects</td><td>Landscape importance for subserved task or</td><td>Position of evaluated objects</td></tr>
</table>

Table 2 (*continued*)

Main characteristics - main criteria	Sub-criteria characteristics	Definition	Quality parameter
k_4	of interest – k_4	functions	
Techniques of application and safety – k_5	Data standardization – k_{51}	Declared standards adherence	Percentage of non-compliance objects with declared standard
	Independency on application software– k_{52}	Degree of independency on application software	Independency/dependency on software used for data editing
	Data protection – k_{53}	Degree of the data protection system and its level	Level of protection of user access right– k_{531}
			Level of protection of copyright– k_{532}
			Level of physical data protection– k_{533}

4.3 Individual Benefit Cost Assessment Structure

The organisation, such as the Geographic Service of the Army of the Czech Republic or the Czech Office for Surveying, Mapping, and Cadastre, are usually responsible for *geospatial databases (GDB)* development continuously covering all the Czech Republic area or some parts of the World. Digital Landscape Model (DLM25 or DMU25 in the Czech language), Multinational Geospatial Co-Production Program (MGCP) or Vector Map Level 1 (VMap1) can be mentioned as examples from military branch.

The GDB are usually developed and maintained by individual partial components of the complete database, such as save units, measurement units, map sheets etc. Therefore, it is quite a good idea to assess their utility value in the above-described system within the established the storing units introducing *individual utility value*. Similarly the individual utility value can be applied for the selected part of master databases from given *area of interest* which is used for certain task.

When assessing database utility, it is useful to define *ideal quality level* at first. The ideal level is used as a *comparison standard* to express each criterion compliance level. Using the comparison standard the individual criteria compliance level and consequently aggregate utility may be assessed.

The compliance level of each individual criterion $u_{n,s}$ is given as follows:

$$u_{n,s} = \frac{k_s}{k_s^*} \tag{2}$$

where

- k_s is for the value of s^{th} criterion compliance,
- k_s^* is for the level of compliance of s^{th} criterion or its group criterion of the comparison standard.

Than the aggregate individual utility value (*individual functionality – U_n)* of the n^{th} save unit is defined by the aggregation function of the some type as (1). Therefore:

$$U_n = p_3 u_{n,3} p_4 u_{n,4} (p_1 u_{n,1} + p_2 u_{n,2} + p_5 u_{n,5}) \tag{3}$$

Particular criteria usually consist of several sub-criteria (see Table 2). The authors took 20 criteria into their consideration; hence the equation for calculation the aggregate individual utility value is therefore a function of 20 variables that characterise the levels of compliance for each individual criterion.

Any modification of selected criterion has an impact on the value of U_n. Individual variables are independent one to another, so the derivation of the function can model the changed utility values or individual utility values.

$$dU = \frac{dU_n}{du_{n,i}} \tag{4}$$

where $i = 1,\ldots, 5$, $n = 1,\ldots, N$, and N is number of all saved units in the database.

Determination of dU value is thus feasible in two ways regarding the desired information structure. When assessing *individual variables effects* on the individual functionality value, while the other variables keep constant values, it is necessary to differentiate U function as follows:

$$dU = \frac{dU_n}{du_{n,i}} \frac{du_{n,i}}{dx} \tag{5}$$

where x is one of the 20 mentioned variables.

In practice, however, such situations may arise that multiple factors may change at the sometime, e.g. the technical quality of database changes in all its parameters—the secondary data derivation methods will improve location and attribute accuracy and the data integrity will increase, and moreover the data are stored in a geodatabase accessible to all authorised users. In this database the data are maintained properly with respect to all topologic, thematic and time relations. In such a case it is suitable to define dU value as a total differential of all variables describing the modified factors.

Database functionality degree is comparable to the cost necessary for provisions—direct used material, wages, other expenses (HW, SW, amortisation, costs for co-operations, tax and social payments etc.), research and development cost, overhead cost and others. Functionality and cost imply *relative cost efficiency (RCE)* calculated as follows:

$$RCE = \frac{F}{\sum_{i=1}^{n} E_i} \tag{6}$$

where $i = 1,\ldots, N$.

Similarly to individual utility value, it is possible to consider the impact of particular variables of expenses E_i on final *RCE*. The goal is to find such solution as the functionality will be maximised and the expenses will be minimize.

The GDB benefit cost assessment including individual benefit cost is a task for a data manager or a geographer-analyst which is responsible to provide demanding project. The system enables him to consider which quality parameters are possible to improve in given time, with given technological conditions, with given sources, with given co-workers etc.

5 Pilot Study

In order to verify our methodology the task of Cross Country Movement (CCM) was chosen as an example. CCM can be solved as a common problem or with consideration of certain types of vehicles (the most frequent or the weakest in the unit, but in case of armed forces usually off road vehicles). The detailed theory of CCM is in (Rybanský, 2009).

The solution can offer to the commander not only one possibility, but the variants from which he can choose according to his intentions and the current situation at the given area.

5.1 Cross Country Movement

Let us recall the basis of the CCM theory. The main goal of CCM theory is to evaluate the impact of geographic conditions on of movement of vehicles in terrain. For the purpose of classification and qualification of geographic factors of CCM, it is necessary to determine:

- particular degrees of CCM
- typology of terrain practicability by kind of military (civilian) vehicles
- geographic factors and features with significant impact on CCM

As a result of the geographic factors impact evaluation we get three known degrees of CCM:

- GO - passable terrain
- SLOW GO - passable terrain with restrictions
- NO GO – impassable terrain

Geographic factors determining CCM and the selection of the access routes are follows:

- gradient of terrain relief and micro relief shapes
- vegetation cover
- soil conditions
- meteorological conditions
- water sheets, water courses

- settlements
- communications
- other natural and manmade objects

The impact of given geographic factor can be evaluated as a *coefficient of deceleration* 'C_i' from the scale of 0 to 1. The coefficient of deceleration shows the real (simulated) speed of vehicle v_j in the landscape in the confrontation with the maximum speed of given vehicle v_{max}. The impact of the whole n geographic factors can be expressed as the formula:

$$v_j = v_{\max} \prod_{i=1}^{n} C_i, \qquad n = 1, \ldots, N. \tag{7}$$

The main coefficients of deceleration are listed in the next table (see Table 3).

Table 3 Main coefficients of deceleration

Basic coefficient	Geographic signification and impact
C_1	Terrain relief
C_2	Vegetation
C_3	Soils and soil cover
C_4	Weather and climate
C_5	Hydrology
C_6	Build-up area
C_7	Road network

Each coefficient consists of several coefficients of 2nd grade. For example C_2 is express after simplification as:

$$C_2 = C_{21} C_{22}$$

where:

- C_{21} is deceleration coefficient by impact of trunks spacing,
- C_{22} is deceleration coefficient by impact of trunks diameter.

The values of deceleration coefficients are counted for given vehicle (its technical properties) from ascertained properties of geographic objects stored in the spatial geodatabase. Using formula (7) it is possible to create a cost map in which the value of each pixel is the final (modelled) speed. The cost map can be as a source for the fastest path, reliable path etc. calculation.

Example for C_2 coefficient evaluation

Let us to define a task 'Looking for the appropriate place for hidden command point and find out an appropriate route there'. If the place is found in a forest, its properties is necessary to evaluated concerning to given army vehicles.

The forest properties are saved in the DLM25 database where the trunks spacing is specified as TSC parameter and trunks diameter as SDS parameter measured in the high of 1.2 metres above the surface. The vehicle can pass forest up to TSC_{max} spacing and SDS_{max} thickness without some serious problems. The terrain vehicle can pass forest with reduced speed, if the trunks spacing is smaller but passable (TSC_{min}) and vehicle has to turn among trees, or trees thickness enables to break them (SDS_{min}). If the size of obstacle is bigger, the vehicle velocity is 0. Properties TSC_{min} and SDS_{min} are given by the technical description of given vehicles validated at the field tests, and comparative values are read from spatial geodatabase. In the mathematical formula the condition can be express:

$$C_{21} = \begin{cases} 1 & for \quad TSC > TSC_{max} \\ 0{,}5 & for \quad TSC = (TSC_{min}; TSC_{max}) \\ 0{,}25 & for \quad TSC < TSC_{min} \wedge SDS < SDS_{max} \\ 0 & for \quad TSC < TSC_{min} \wedge SDS > SDS_{max} \end{cases} \qquad (8)$$

$$C_{22} = \begin{cases} 1 & for \quad SDS < SDS_{min} \\ 0{,}5 & for \quad SDS = (SDS_{min}\,; SDS_{max}) \\ 0 & for \quad SDS > SDS_{max} \wedge TSC < TSC_{min} \end{cases} \qquad (9)$$

5.2 Geospatial Database Utility Value Evaluation

The master DGI database is usually utilised as a base for spatial data analyses. The national or international databases as DLM25, VMAP1or MGCP are very detailed, carefully maintained and used in many applications. But nobody can suppose that those databases contain all information he could need.

The task of CCM solution could require more information that is available in the master database. Geographer-analyst has to consider which information and in what quality can he obtain from master database. E.g. all forests in the area of interest are necessary to select for mentioned C_{2i} coefficients. Further he has to find out all their properties and their accuracy or count how many characteristics are missing. The system presented in the Table 2 serves as a manual. Next step is the individual utility value of given part of master database evaluation.

Attributes are usually defined as the characteristics or nature of objects. In geospatial sense, an attribute is regarded as a property inherent in a spatial entity (Shi, 2010). In our case an attribute is a characteristics or variable constituting the base for the computation of basic coefficients C_1 - C_7. These attributes differ in their nature according to the real world phenomena they represent.

Not all attributes are available within the used thematic spatial databases. So far the incompleteness of attributes has been omitted. Thus the real state-of-the-art has not been taken into account and the resulting CCM path has been considered as ‘certain’. One of the possibilities to make the resulting path closer to reality is to take the data attribute incompleteness into account and inform the decision maker (commander) about the uncertain parts of the path.

Two variants of the DLM25 database were utilised for the pilot project. The feature properties were defined according to the Feature Attribute Coding Catalogue (FACC) adapted as Catalogue of the Topographic Objects (CTO) (MTI,

2005) in the first variant updated in 2005. The missing values of object's attributes were marked as 0 in given domains. The 4th edition of CTO was transformed in accordance with the DGIWG Feature Data Dictionary (DGIWG-500, 2010) in 2010 and transformed edition (updated in 2010) was used in the second variants (MoD-GeoS, 2010). The missing properties were marked in several attribute categories as:

- -32767 for unknown variables,
- -32766 for unpopulated variables,
- -32765 for not applicable variables,
- -32764 for other variables, and
- -32768 for No or Null values.

The smaller personal database was created in the area of interest round Brno of the size approximately 400 km^2 and all objects and phenomena necessary for CCM evaluating was selected from DLM25 master databases of both variants. The individual utility value was counted for both variants, but with a small simplification. At the first step we didn't do any independent tests for position accuracy determination, further we didn't consider the software independency, and landscape importance. Then we suppose the whole database is complete, the position and thematical resolution corresponds to our task, and the data are properly protected.

On the base of statistical analyse 12.65% objects have any problems mainly incomplete attributes in the first variant of DML25 while 3.45% objects have any similar problems in the second one. The time difference is 5 years between both variants. Hence the individual utility value was calculated by the use of the formula (3) as 0.6887 for the 2005 variant and 0.8825 for the 2010 variant. The ideal quality level is 1.0068. Both variants were used for CCM of TATRA 815 evaluation.

5.3 CCM of TATRA 815 Analyses

The common army vehicle TATRA 815 ARMAX was chosen for a particular vehicle evaluation.

Table 4 The technical characteristics of TATRA 815 ARMAX (Tatra, 2010)

Parametr	Value
Length (m)	7,87
With (m)	2,5
High (m)	3,01
Maximum climbing capability at 26000 kg of cargo	36°
Maximum climbing capability at 41000 kg of cargo	18°
Maximum climbing capability up to rigid step (m)	0,5
Maximum width of trench (m)	0,9
Maximum road speed (kph)	85
Maximum depth of wade without water streaming (m)	1,2

For further simplification only data with full information (all attribute properties had to be filled) were considered *in both variants*. If some information was missing in any geographic object system didn't consider it and one reliable path was analysed. Authors knew that e.g. narrow streams are probably passable for TATRA lorry, but their passing was allowed only over bridges if no information about depth and banks characteristics were in the database.

The process was in progress according to next schema (Fig. 6):

- CCM evaluation - only reliable information are considered,
- final cost map calculation according to equation (7),
- minimum cost path calculation from one initial point to three destinations placed in the forests as some hidden position.

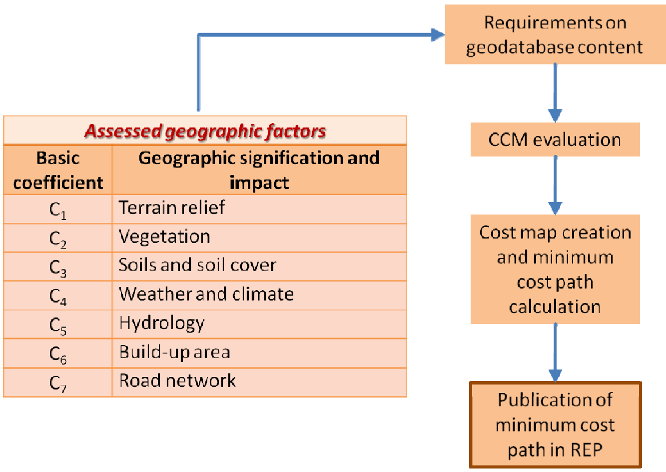

Fig. 6 Spatial analyse without database quality evaluation

Only one solution is offered to commander. This solution seems to be appropriate, but geographer-analyst generally doesn't know details about situation on area of responsibility and commander's intentions. Problems should appear when the tactical (or other) situation doesn't make the published path possible to use. Commander usually requires a new solution in such a case to miss prohibited area. The quality characteristics of temporally database are than to be considered by geographer to be sure, where are the weakest points of a new analysis. The weak points of analysis have to be sent to commander together with own analysis and it is up to him what will be the final decision. Two tasks for commander appear in CCM example:

1. Use less reliable path and consider that vehicles could stay in front of some obstacle
2. Wait and order to GEO team to improve spatial database (e.g. required properties) as soon as possible and then use new reliable path

The second case was simulated by the second variant of database in which the quality parameters were improved.

ArcGIS 9.3 was used for all calculations and analyses. The main analyses were described using ModelBuilder to have simpler possibility to change the input parameters. In the next figures there are the main results – cost maps. The cost of each pixel is symbolized in the gray scale where darker tone signifies higher cost, higher speed in this case.

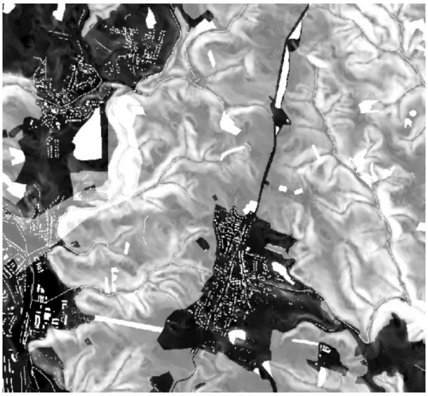

Fig. 7 The cut cost map created from DLM25 2005 version

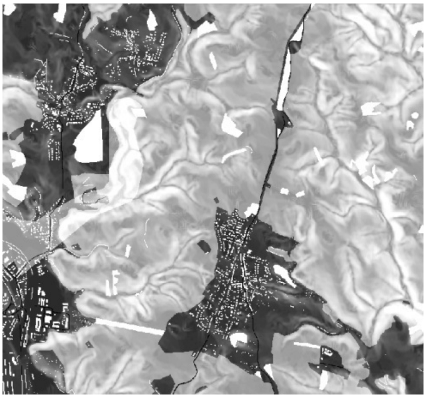

Fig. 8 The cut cost map created from DLM25 2010 version

The minimum cost paths were evaluated using both cost maps and the same process created in ModelBuilder were applied. The results are in the next figures.

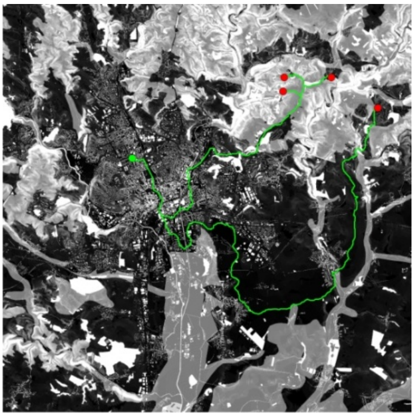

Fig. 9 The minimum cost paths in CM of 2005 version. The initial point is green, the destinations are red

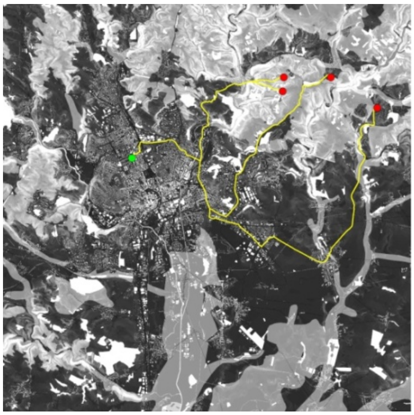

Fig. 10 The minimum cost paths in CM of 2010 version. The initial point is green, the destinations are red

The comparing of both results presented over the topographic situation is shown in the next picture (Fig. 11).

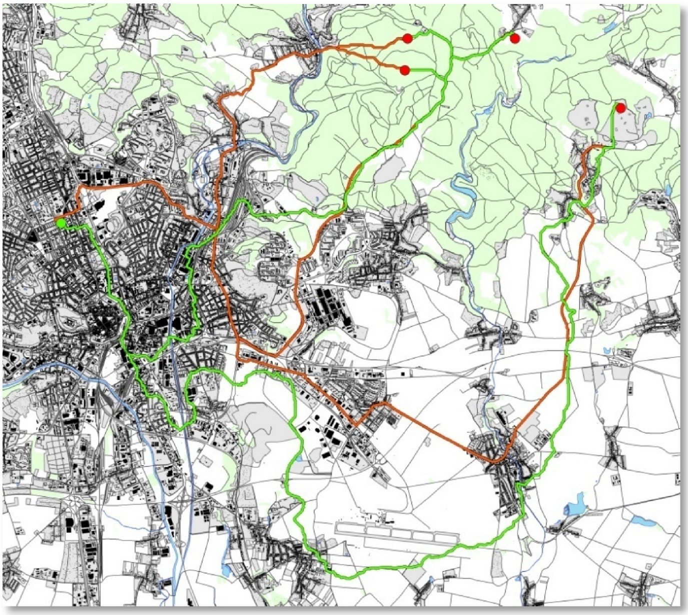

Fig. 11 Comparing of two variant of minimum cost paths. Red ones answer to 2005 version and the green ones to 2010 version

5.4 Discussion

The obtained results proved that the results of spatial analyses are highly dependent on overall quality of digital spatial data. Relatively small changes in the database can cause significant differences in received results.

Authors purposely considered only original stages of selected objects from master databases in both variants where no uncertainty is evident. But the real situation in the landscape is quite different. Some thematic properties are uncertain and they can change in shorter or longer time (trees growth, soil moisture, river with and depth etc.), some are uncertain because of no possibility to measure them in the required precision (e.g. soil types borders). There exist several methods how to handle the attribute uncertainty and incompleteness (Shi, 2010).

Slightly other situation is when the missing properties of data are considered. Some properties are missing on purpose because their application has no sense or is impossible because of some restrictions. Some properties are missing because of lack of time to add all declared object properties and from the geographer point of view the object importance seems not to be very high. Moreover, the geographer can assume that the user could know the area of interest and that he is able to add missing properties (with respect to his experience). E.g. Small streams, round Brno, have in the summer its width up to 2 meters and depth up to 0.5 meters, so they are easily passable. On the other hand, banks of small streams can be high, muddy etc. and so impassable. To complete all declared properties is costly and time demanding, so sometimes the question is whether to add the missing data and spend money for it or not to add the data and rely on the user‘s knowledge about the environment.

Using the cost map it is possible to create not only reliable path, but also risky path if some ‘unimportant’ obstacles are sign as passable. If the commander obtains the risky path, the additional information about obstacles which were not included into consideration is necessary to be added to him. Then it is up to commander to consider the level of risk decision taken.

6 Conclusion

The pilot project has demonstrated a strong relationship between quality data and the results of spatial analysis. Likewise, it pointed to the problem of defining quality. It is not possible to assess only the technical properties of the spatial database, but it is necessary to consider the quality of the entire complex. For example, when changing standard used, which at first glance may not be large, its influence on the resulting analysis can be substantial. Specifically, during the solution of the CCM task, we faced a problem of classification of thematic properties. In an earlier version, blank data were marked only to the value 0, while in the new version, the blank data disintegrate into several groups identified by the values of -32765, -32766, etc. according to specifications (DGIWG-500, 2010).

In the pilot project, we have dealt only marginally with uncertainty in setting the boundaries of geographic objects and phenomena and with the uncertainty of their thematic properties. The authors are aware of the fact that this uncertainty, given by the natural conditions, can significantly affect the results of spatial analysis. For example, the width of the water flow is usually stored in a database as the value of the width of the flow in the normal state level. If it spill over due to the heavy rains, this width may be much larger, or vice versa during prolonged drought it may diminish. As a result, the conditions of negotiability of this flow changes significantly. The problem of implementation of the principles of uncertainty and their mathematization will be the task of the solution of our future project.

The actual pilot project showed that the proposed way to address the relationship between the quality of spatial data and spatial analysis is possible and should continue.

Acknowledgements. Research results presented above were kindly supported by the project „The evaluation of integrated digital geospatial data reliability“ funded by the Czech Science Foundation (Project code 205/09/1198) and also by the research task „Dynamic geovisualization in crises management “supported by the Czech Ministry of Education (grant MSM0021622418).

References

DGIWG-500. Implementation Guide to the DGIWG Feature Data Dictionary (DFDD) (2.2.2 - July 19, 2010 ed.). DGIWG (2010)

Jacobsson, A., Giversen, J.: Eurogeographics (2007), http://www.eurogeographics.org/documents/Guidelines_ISO19100_Quality.pdf (retrieved 2009)

Konečný, M., Kubíček, P., Staněk, K.: Mezinárodní standardizační aktivity v oblasti prostorově orientovaných informačních systémů. In: Sborník příspěvků Konference Geomatika, vol. 98, pp. 35–41. CAGI, Praha (1998)

MoD-GeoS, Catalogue of the Topographic Objects DLM25 (7.3 ed.). Ministry of Defence of the Czech Republic, Geographic Service, Dobruska (2010)

MTI. Catalogue of the Topographic Objects DLM25 (4 ed.). Military Topographic Institute, Dobruska (2005)

Rybanský, M.: Cross-Country Movement, The Imapct And Evaluation Of Geographic Factors, 1st edn. Akademické nakladatelství CERM, s.r.o. Brno, Brno, Czech Republic (2009)

Shi, W.: Principles of Moddeling Uncertainties on Spatial Data and Spatial Analysis. CRC Prass, Taylor and Francis Group, Boca Raton (2010)

Talhofer, V., Hofmann, A.: Possibilities of evaluation of digital geographic data quality and reliability. In: 24ht International Cartographic Conference, The World's Geo-Spatial Solutions, ICA/ACI, Santiago de Chilie (2009)

Talhofer, V., Hoskova, S., Hofmann, A., Kratochvil, V.: The system of the evaluation of integrated digital spatial data realibility. In: 6th Conference on Mathematics and Physics at Technical Universities, pp. 281–288. University of Defence, Brno (2009)

Tate, J.: Open standards the Gateway to Interoperability? GEOINT Summit 2010. Jacobs Flaming Conferences, Vienna (2010)

Tatra, A.: Tatra is the solution, TATRA (2010), http://partners.tatra.cz/exter_pr/vp/new/typovy_listprospekt.asp?kod=341&jazyk=CZ (retrieved May 17, 2010)