MODELLING DATA QUALITY WITH POSSIBILITY-DISTRIBUTIONS

G. Navratil

Institute for Geoinformation and Cartography, Vienna University of Technology, Gusshausstr. 27-29, A-1040 Vienna, Austria – navratil@geoinfo.tuwien.ac.at

KEY WORDS: GIS, Data Quality, Metadata, Modelling, Fuzzy Logic, Possibility Theory, Interoperability

ABSTRACT:

Description of data quality relies heavily on numbers. When dealing with data sets, which have been collected during longer periods, however, variation in the data quality will become evident. Older data may have different quality than newer data and other aspects including height or accessibility may influence the quality. Precise numbers do not reflect the uncertainty, thus I propose the use of fuzzy numbers to specify data quality.

Fuzzy numbers are based on the specification of a distribution function. The distribution function may be a probability or a possibility function. Probability is more difficult to determine than possibility. Still, the possibility function may provide all information relevant for the user. Thus, providing the possibility function may be sufficient to improve the data quality description.

KURZFASSUNG:

Die Beschreibung von Datenqualität basiert vorwiegend auf Zahlenangaben. Beim Arbeiten mit Datensätzen, die über einen längeren Zeitraum erfasst wurden, erkennt man schnell Unterschiede in der Qualität verschiedener Daten desselben Datensatzes. Alte Daten wurden mit anderer Qualität erfasst als neuere Daten und auch Rahmenbedingungen wie Höhenlage oder Zugänglichkeit können die erzielte Qualität beeinflussen. Schafte Zahlenangaben können diese Schwankungen nicht vermitteln. Daher schlage ich die Verwendung unscharfer Zahlen zur Spezifikation der Datenqualität vor.

Unscharfe Zahlen basieren auf der Angabe von Verteilungsfunktionen. Das kann entweder eine Wahrscheinlichkeitsfunktion oder eine Möglichkeitsfunktion sein. Die Bestimmung der Möglichkeit eines Ergebnisses ist einfacher als die Bestimmung der Wahrscheinlichkeit. Trotzdem beschreibt die Möglichkeitsfunktion eventuell bereits alle Umstände, welche für den Nutzer eines Datensatzes relevant sind. Daher kann die Verwendung einer Möglichkeitsfunktion ausreichen, um die Beschreibung der Datenqualität zu verbessern.

1. INTRODUCTION

The amount of available data increased with the development of new technologies. Availability of data and capability of processing more data than before led to new applications like online route planning or visualizations in landscape planning and architecture. The outcome of the application depends on adequate selection of data sets. However, data quality varies with the source. Data quality descriptions like positional quality, temporal quality, or completeness have been defined to cope with that problem (Guptill and Morrison 1995).

Automation of processing requires automatic handling of data quality. Fitness for use is a method to describe, if a specific data sets is suitable for a specific task (Chrisman 1984). Automatic pre-selection of the data sets can simplify the selection process for the user. Automatic pre-selection requires automatic determination of the fitness for use. The necessity for the discussion of user needs has been pointed out by Byrom (2003). Approaches to define the user needs and use them to specify the fitness for use have been presented by Jahn (2004), Grum and Vasseur (2004), and Pontikakis and Frank (2004). A first step in that process is the suitable description of the data quality.

In many cases, data sets are collected over long periods. Graphs of street networks, for example, are determined once and updated periodically to reflect changes in the network. Since the quality of the determination may change due to improved technology, the quality of the data varies within the data set. A solution to this problem is using the worst case for providing a number for the quality. However, this solution may give a wrong impression if a small part of the data is of significantly worse quality than the rest of the data set. The Austrian cadastre, for example, provides data on parcel boundaries. The data is either from the coordinate-based cadastre and thus has a positional quality of 15 cm or it is from the boundary mark system with a positional accuracy of up to a hundred meters. However, most of the low quality boundaries are located in mountainous regions or forests where the determination of the boundary is less important than in urban areas. Thus specifying 100 m as positional quality of the Austrian cadastre would comply with above definition but would provide a wrong impression of the overall quality. Similar discussions are possible for other aspects of data quality and other data sets. Quality should thus not be described by a single value.

The paper uses the example of the Austrian cadastre to specify possibility distributions. Different aspects of quality are modelled with different distributions. Then user requirements are assumed and the paper presents a method to compare the data quality with the user requirements. This allows answering the question for fitness for use.

2. POSSIBILITY FUNCTIONS

Reasoning can be defined as testing the correspondence of a specified hypothesis with given statements. The statements can be data stored in a database and the hypothesis is a query on these data. The typical example is a database containing the heights of persons and the question, if a specific person is 'tall'. Four different situations can be determined (Dubois and Prade 1988):

- Both, the data in the database and the definition of 'tall' are crisp. The entry in the database for the person could be 1.7 m and 'tall' is defined as '>1.65 m'. This leads to traditional, two-valued logic.
- The data in the database is vague and the definition of 'tall' is crisp. Here the definition for 'tall' is the same as above but the entry in the database is expressed with a possibility distribution. The possibility distribution describes the degree of possibility that the height of the person corresponds with the value. This leads to possibility theory as published by Zadeh (1978; 1979) and expanded by Dubois and Prade (1988).
- The data in the database is crisp and the definition of 'tall' is vague. Here the entry in the database could be 1.7 m but the concept of 'tall' is uncertain. This leads to many-valued logic.
- Both, the data in the database and the definition of 'tall' are vague. This leads to fuzzy logic (Zadeh 1975).

Which of these types of logic shall we use for modelling data quality? The data describes the world. Since the world changes, the data must change, too. Thus the data acquisition is a continuous process. Data quality parameters shall describe the quality of this data. It will not be possible to use a crisp description. Crisp descriptions are possible for single elements within the data set. However, the quality will vary throughout the data set and this variation should be reflected by the data quality description. Thus description of the whole data set will be vague. There may be general statements like:

- The standard deviation of coordinates is 10 cm.
- The last update of the topographic map was done in 2001.

These expressions only set limits for the quality. The first proposition does not say that each coordinate in the data set has a standard deviation of 10 cm. The proposition may be meant as an average value or (which is more likely) as an upper boundary. The proposition could then be rephrased as "The value of standard deviation for an arbitrary point coordinate in the data set is unknown but it is less than 10 cm." The second statement is similar. It only defines that changes in the real world that happened after 2001 will not be included in the map. Thus we deal with uncertain data.

The questions we raise are crisp or can be made crisp. From a user's perspective we have two different questions:

- I need data set with a specific quality – is it available?
- There is a data set with specific quality – can I use it for the purpose at hand?

Both questions are crisp. In the first case there may be several parameters for the data quality. All of these parameters must be fulfilled. Thus a data set either fulfils the quality specification or it does not. This gives a crisp answer to the question. The second question is more complex. Like in the first case data quality issues must be considered but in addition a cost-benefit-analysis is necessary. According to Frank (1995) and Krek (2002) the value of a data set emerges from better decisions the user can make. The value can be compared to the costs of acquisition and processing of the data. The data set is applicable if the costs are lower than the benefits and there is no other possible outcome than using or not using the data set.

Such a discussion of processes requires a method to describe the outcome of processes. Possibility functions (Zadeh 1978) are such a method. Another method is probability functions. In general, the use of fuzzy methods is suitable for the results of precise observation processes and they can be used to make statistical analysis (Viertl 2006). Viertl uses probability functions, which assign probabilities to each possible outcome. This requires detailed knowledge since the possibilities must be determined first. Possibility distributions avoid that problem by not providing probabilities. They only specify the possibility of the result: The value 0 shows impossibility and 1 shows possibility. Values between 0 and 1 provide information on the plausibility of the outcome. Thus, a result with value 0.4 is possible but less plausible than a result with 0.8.

An introduction to the mathematical definition of possibility distributions based on propositions has been shown by Wilson (2002): The most common way to express propositions is using a set Θ of mutually exclusive and exhaustive possibilities. A possibility distribution π assigns a value of possibility to each element of the set. If there is an element with value 1 then the function is said to be normalized.

$$\pi: \Theta \rightarrow [0,1] \quad (1)$$

Values between 0 and 1 express orderings of propositions.

3. QUALITY OF CADASTRAL DATA

The Austrian cadastral data is an example for a large data set collected over an extended period. The data set includes parcel identifiers, parcel boundaries, and current land use. Details on the Austrian cadastral system can be found in different publications (Angst 1969; BEV 1991; Twaroch and Muggenhuber 1997). An important aspect is the definition of boundary. Whereas evidence in reality (like boundary marks, fences, or walls) defines the boundary in the traditional Austrian cadastre, the new system uses coordinates to specify the position of the boundary. This change allows the creation of data sets reflecting reality since the data provides the legal basis for the boundaries.

The determination of data quality must take the creation process into account. There are three different influences on data quality (Navratil and Frank 2005)

- technology,
- legality, and
- usability.

Technological possibilities limit the achievable quality. In general there is a maximum quality as well as a minimum quality. The usual method to create a terrain model, for example, is either terrestrial 3D-measurements or evaluation of aerial images. In both cases the quality of the terrain model depends on the expenditure. Lower flight height will improve the quality as well as condensing the network of terrestrial measured points. Different equipment used for measurement will result in different precision of the height points, too. This cannot be continued infinitely since maximum quality is determined by the roughness of the world's surface (e.g., a field after ploughing).

the world's surface (e.g., a field after ploughing). Reduction of quality is also limited. Having less than a single height point in a terrain model is impossible.

Laws also have an impact on the quality of available data sets (Navratil 2004). Laws directly influence quality and extent of data sets required for tasks specified in laws. Laws may also prohibit the use of data with higher quality than a specified one due to data protection laws or security reasons. In contrast to the technological limitations legal influences are not 'hard'. It is possible to specify the maximum technical quality of a length measurement and this quality cannot be exceeded. This is not completely true for laws. The Austrian law stipulates a minimum precision of 15 cm for boundary points. Since it is extremely difficult to prove the quality of coordinates in a subdivision document, a point with a precision of only 20 cm may also be accepted. Thus legal rules on data quality can be often seen as guidelines to develop technical solutions and thus the results of the process meet these rules.

Usability of data also affects data quality. Data used more often will usually have higher quality since it produces more revenue and thus more money is available for collecting the data. Maps provide a good example: Nautical maps have higher quality where needed. Users do not need detailed nautical maps in the middle of the Atlantic Ocean but they do need them in coastal areas where the danger of hitting the ground is high. Thus more money is spent on mapping coastal areas than on mapping the ocean.

So how can we describe data quality? ISO 19113 "Quality Principles" (ISO 19113 2002) defines the framework for a quality model. The data quality elements are

- completeness,
- positional accuracy,
- temporal accuracy,
- logical consistency, and
- thematic accuracy.

Each of these elements provides a set of possibilities for the possibility functions. Since the functions are similar for all elements, I will use the first three elements only in this paper.

Completeness for the boundary data in the cadastre can be defined as the percentage of boundary lines, which are either missing or classified in a wrong way. Positional accuracy connects to the elements defining the boundary lines. The Austrian cadastre uses boundary points to define the boundary. Thus the positional accuracy of the boundary points can be used, stipulate the positional accuracy of the data set. Temporal accuracy of the cadastre is the time between a change in reality and it's reflection in the cadastral data set.

Frank, Grum et al. show, that the quality elements completeness, positional accuracy, and temporal accuracy are not independent (Frank, Grum et al. 2004). Positional accuracy and completeness dilute over time. The underlying assumption is that data is obtained once and there is no update process. This is not true for the Austrian cadastre. Changes in the real world do not affect the position of the boundaries since the boundaries are based on the position of boundary points. Coordinates define the position of the boundary points and these coordinates are not influenced by changes in the real world. They can only be changed by cadastral processes (Navratil and Frank 2003). Thus the positional accuracy of the boundaries is independent of the time. However, this is only true for the coordinate-based system. The traditional tax cadastre still running in parallel to the coordinate-based one does have a dilution of precision over time because the boundaries are defined by evidence (boundary stones, fences, walls, etc) in the real world, which can change. The same is valid for completeness. In the tax cadastre land can be acquired by adverse possession. Adverse possession creates a new boundary, which is not part of the cadastral data. This, too, is not possible in a coordinate cadastre.

4. MODELLING DATA QUALITY WITH POSSIBILITY DISTRIBUTIONS

4.1 Technological Influence

From a technological point of view the Austrian cadastral data can be up to 100% complete. The cadastral data set can be kept complete if the original cadastral data set is complete and each change is included in the data set. Completeness of the original data set cannot be proven although there are tests, which can be used to check it. Cadastral maps of the country, for example, must not have areas without defined ownership. Ownership on land is defined by the land register and the land register uses the cadastre as spatial reference. Thus missing boundaries in the cadastre would result in inconsistent or incomplete data sets in the land register because a piece of land would have two different ownership situations or no defined owner.
Reduction in completeness would be technically possible by removing boundary lines. This is a simple process and thus each level of completeness from 0% (all lines removed) to 100% is possible. The possibility distribution for the completeness is defined for the different levels of percentage thus ranging from 0% to 100% (see Figure 1).

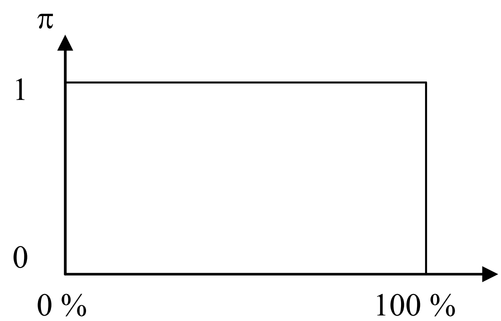


Figure 1. Possibility distribution for technological influence on completeness

Positional accuracy for cadastral boundaries depends on the accuracy of boundary points. Modern technical solutions for point determination use GPS and high-precision measurement equipment. This results in a standard deviation of 1 – 5 cm. This accuracy can be reached if the whole data set is re-measured to eliminate all influences of outdated measurement methods like triangulation networks as performed in Austria in the 19th century. Reduction of quality is easily possible, e.g., by using cheaper equipment. The lower limits are reached if the standard deviation affects the topology described by the data set. In case of the Austrian cadastre effects will start with an accuracy of approximately 1m and the data set will become unusable with an accuracy of approximately 10 m.

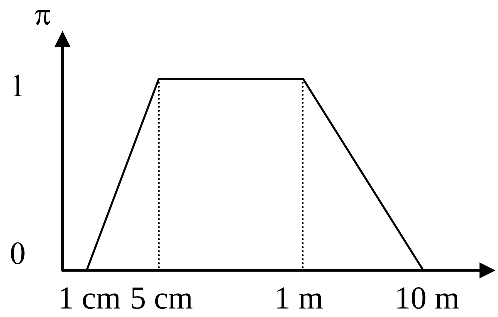


Figure 2. Possibility distribution for technological influence on positional accuracy

Temporal accuracy depends on processing time for changes. Changes in the real world require documentation and must then be inserted in the cadastral data. This takes time since it must be guaranteed that there are no errors in the data. In addition publication must be secure. Thus, the database visible outside the cadastral organization is a copy of the original database. Updates are done in regular intervals to keep the update process simple. The Austrian database is updated daily. This further restricts the temporal accuracy. We can assume that the process from the definition of the boundary to the published change requires a week. It is possible, however, that the process takes longer. The maximum age of the data is the age of the cadastre itself, which is less than 200 years. The possibility function thus looks like in Figure 3.

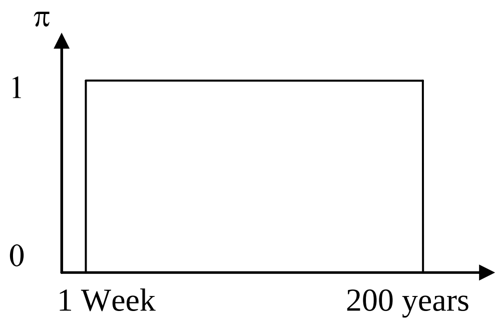


Figure 3. Possibility distribution for technological influence on temporal accuracy

4.2 Legal Influence

Boundaries in the cadastre reflect changes in the ownership situation as described. As mentioned above Austria has two different cadastral systems: the coordinate cadastre and the tax cadastre. The tax cadastre allows for adverse possession. Thus it is legally possible that the data set is not completely describing the boundaries. The percentage of these boundaries cannot be estimated but the assumption that the number is not high seems plausible since many boundaries are fixed by walls or fences. However, 100% completeness is not plausible since a single case of adverse possession would result in a deviation from 100%.

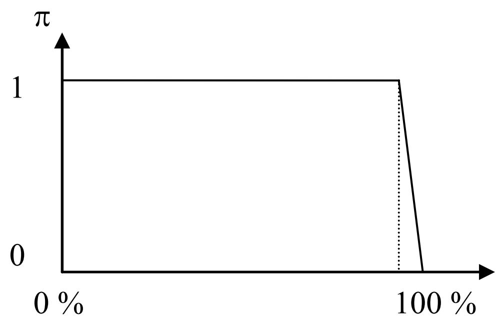
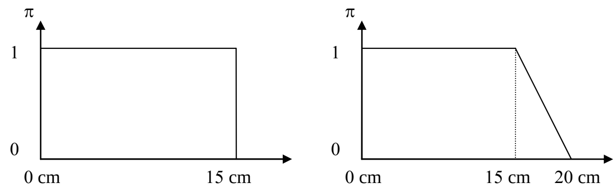


Figure 4. Possibility distribution for legal influence on completeness

The decree for surveying (Austrian Ministry for Economics 1994) stipulates the positional accuracy for boundary points. The limit is 15cm and thus theoretically the possibility distribution for the positional accuracy looks like in Figure 5a. This rule is strict as law does not know standard deviation as a basis for decision making (Twaroch 2005). However, as stated in section 3 it is difficult to control the actual accuracy of a boundary point. We thus have to assume that there may be points with lower accuracy. We can model this by gradually reducing the value of the possibility distribution as shown in Figure 5b.



(a) theoretical distribution (b) practical distribution

Figure 5. Possibility distribution for legal influence on positional accuracy

An assessment of the temporal accuracy of cadastral data must consider the cadastral processes. Changes of boundaries emerge mainly from subdivisions and require surveys of the parcel boundaries. The resulting map must be registered in the cadastral office to be inserted in the cadastral map. The subdivision map is checked on correctness and the proof of correctness is documented. The owner then must register the subdivision at the land register and only after this step the new boundary is shown consistently in all cadastral data sets. The legally important action for the boundary definition, however, is the agreement of the neighbors on the boundary definition, which is expressed during the survey by a signature. The period between the date of signature and the visibility of the new line in the cadastral data set has been assessed as a week in section 4.1. However, the owner of the land must also submit the subdivision map to the land register and this must be done within a year. Thus the time span between survey and registration may be up to 1 year.

The situation deteriorates with land consolidation. The purpose of land consolidation is to improve the efficiency of agriculture by increasing the parcel areas. Due to differences in soil quality the amount of land owned by a person may change during land consolidation because the basis is not the area itself, but the productivity of the land. This may lead to dissatisfied land owners, who may go to court to get their wishes fulfilled. Such trials can take a long time and if there are several of those trials, the time between survey (and definition of the boundaries in the field) and update of the data set may be decades. During this period the situation in the world (the new boundaries) and the cadastral data (the old boundaries) will be different. However, the number of land consolidations per year is little and most of them do not involve courts. Still, there is the possibility for such a case, which shows in the possibility distribution.

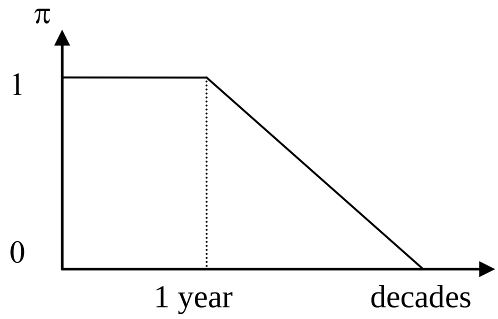


Figure 6. Possibility distribution for legal influence on temporal accuracy

4.3 User Influence

Cadastral data are used by two different groups of users (Navratil, Twaroch et al. 2005):

- Users of the boundary itself: Owners of land need data on their and the neighboring parcels with high accuracy. Data on these parcels must be complete. Data on the remaining land is of no greater relevance to them.
- Users of the positional reference in general: The cadastre is the only large-scale map available for the whole area of Austria and thus it is often used to provide spatial reference. These users are not interested in a high quality of the boundary description. However, they need the data sets to be complete in the area of interest.

Since these two groups have different requirements we have to treat them differently. The differences will clearly show in the possibility functions. In contrast to the technological and legal influence the possibility functions now show if it is possible to use the data set for the specific task.

4.3.1 User of the boundary

Owners of land need data on their parcel completely. This knowledge, for example, is necessary to plan the construction of a building accordingly. Complete knowledge on the boundaries of neighboring parcels may also be necessary, e.g., for subdivisions. Thus the data set is useful as long as all necessary data is included.

Completeness of less than 100% does not eliminate the possibility that all necessary information is contained in the data set. However, reduced completeness increases the plausibility of the fact that some data is missing. Thus we have to assume that there is a level of completeness, below which the data set is unusable for the owner because at least one piece of data will be missing. Starting with this level the possibility increases to 1, which is reached with the complete data set. Figure 7 shows the resulting distribution.

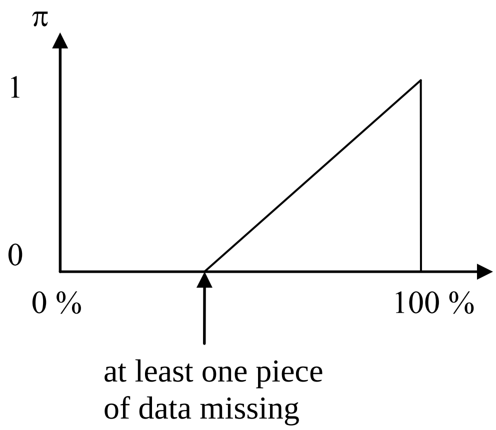


Figure 7. Completeness for users of boundary

Positional accuracy is important for land owners. Land owners want to use their land and one of the common types of use is the creation of a building. In Austria buildings must comply with legal rules specifying for example the maximum building height, the construction style, or the distance from the parcel boundary. The last point requires high positional accuracy to fit the strategies of courts. Thus, although an accuracy of 20 cm may be sufficient for some tasks of land owners, most tasks require a positional accuracy of at maximum 10 cm.

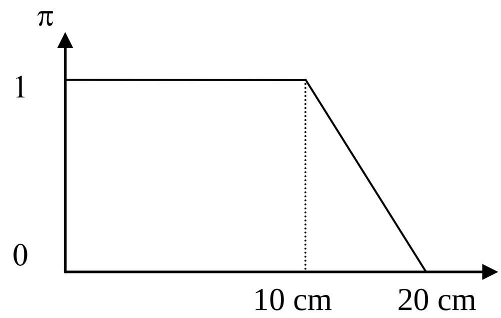


Figure 8. Positional accuracy for users of boundary

The demands of users for temporal accuracy are low because as owners of the land they are involved in all processes concerning their land or neighboring parcels. Thus they can validate, if the cadastral data set is up-to-date. The possibility function will thus be independent of the temporal quality and only depend on the correspondence with the knowledge of the owner.

The only exception emerges if the ownership changes. The buyer must be able to confirm the propositions of the current owner and thus the cadastral data must contain all changes. Thus in this case the temporal accuracy should be 1 day and checking the plausibility of the propositions becomes more difficult with deteriorating temporal quality.

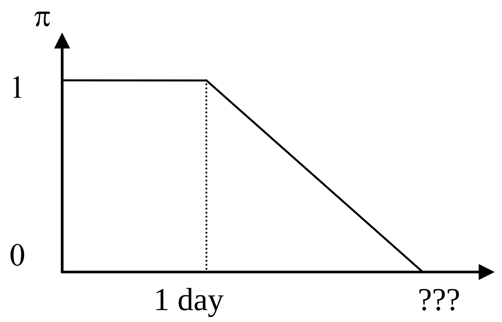


Figure 9. Temporal accuracy for users of boundary

4.3.2 Users of the spatial reference

If cadastral data is used for spatial reference, completeness is more important. Spatial reference is usually provided as background graphics. The maps produced will have a scale of 1:10.000 or smaller. Thus the level of detail provided by the map is low. Too many lines will confuse the viewer and thus the map maker will have to select lines for display. However, the removal will not be done arbitrarily. Lines will be removed mainly in areas with a high density of lines. Percentage of completeness must be high to allow the user to differentiate between essential and unnecessary lines. Decreasing completeness will also decrease usability. Figure 10 shows the resulting possibility distribution.

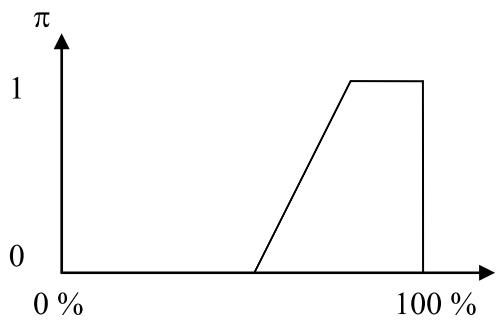


Figure 10. Completeness for spatial reference

Spatial reference has limited demands for positional accuracy. Assuming the above scale of 1:10.000 and accuracy on the map of 1/10mm then the accuracy of the points should be 1m. Higher mapping accuracy leads to higher accuracy demands but accuracy better than 0,5m is not needed for positional reference. The lower limit of accuracy depends on the type of visualization. Accuracy of less than 10m in built-up areas may result in less plausible data sets because is will not be possible to determine on which side of a street a point is and the number of display errors will be high due to small parcels.

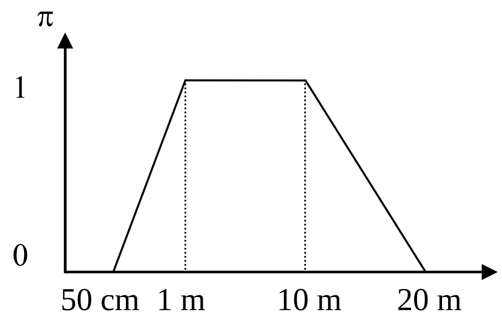


Figure 11. Positional accuracy for users of boundary

Temporal accuracy is not important for spatial reference. A data set will be useful for spatial reference as long as the parcels can be matched with reality. The number of significant changes in an area is usually rather small. The most significant changes emerge from land consolidation, which is rare. In addition, land consolidation is only performed for agricultural areas. Still the plausibility of a generally unchanged data set will reduce with the age of the data set. However, it can be assumed that a cadastral data set can be used for providing spatial reference for 5 years without significant problems. Then the number of changes may start changing the boundaries so that the possibility of using the data set reduces. After 10 years we may have to obtain a new data set.

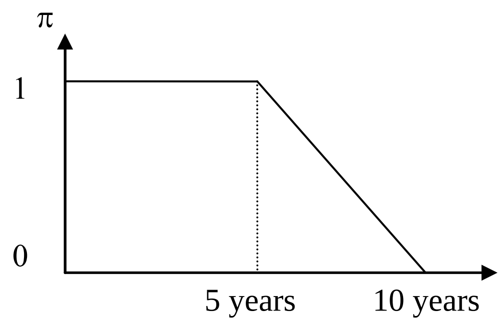


Figure 12. Temporal accuracy for users of boundary

5. COMBINATION OF POSSIBILITY DISTRIBUTIONS

A reasonable condition for data quality is that several conditions must be met. This is a logical ‘and’-relation. The minim-function combines possibility distributions according to the logical ‘and’. The ‘or’-relation would lead to the maximum function for the combination of possibility distributions (viertl and Hareter 2006). Figure 13 shows the combination of the possibility distributions for positional accuracy. The full line represents the influence of technology, the line with short dashes the legal influence and the line with long dashes the needs of land owners. The grey area marks the overlap of the possibilities. Figure 14 shows the same combination for users of the spatial reference. This combination has no solution.

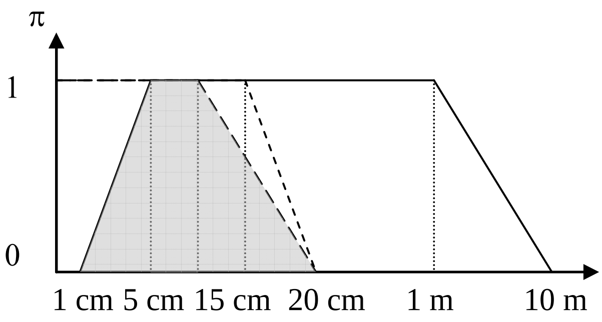


Figure 13. Combination of possibility distributions for positional accuracy for users of boundary

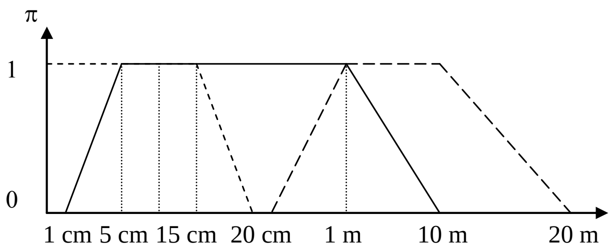


Figure 14. Combination of possibility distributions for positional accuracy for users of spatial reference

The example shows that the technical solutions and legal rules for cadastral systems do meet the demands of owners of land. Other types of use have different demands and thus the possibility function is different. Users, who need spatial reference only, require a different technical solution and different legal rules. This matches the empirical results reported by Navratil, Twaroch et al. (2005).

6. CONCLUSIONS

As we have seen it is possible to model the influences on data quality with possibility distributions. It was possible to specify all necessary possibility distributions. The combination of influences produced a result that can be verified by practical experience. The method thus can be used to assess the correspondence of the influences on data quality.

Left for future investigation is the application for data set selection. The paper showed how to model possibility distributions for influences on data quality. It also showed a simple method to combine these influences. A general method will be needed to create the possibility distribution for more general examples. These distributions might require more sophisticated methods of combination.

7. REFERENCES

Angst, J. (1969). "Das neue Vermessungsgesetz." Österreichische Juristenzeitung **337**.

Austrian Ministry for Economics (1994). Verordnung des Bundesministers für wirtschaftliche Angelegenheiten über Vermessung und Pläne (Vermessungsverordnung 1994 - VermV). **BGBl.Nr. 562/1994**.

BEV (1991). Database of Real Estates. Vienna, Austria**:** 27.

Byrom, G. M. (2003). Data Quality and Spatial Cognition: the perspective of a National Mapping Agency. International Symposium on Spatial Data Quality (ISSDQ), Hong Kong, The Hong Kong Polytechnic University.

Chrisman, N. R. (1984). "The Role of Quality Information in the Long-Term Functioning of a Geographical Information System." Cartographica **21**: 79-87.

Dubois, D. and H. Prade (1988). An Introduction to Possiblistic and Fuzzy Logics. Non-Standard Logics for Automated Reasoning. P. Smets, E. H. Hamdani, D. Dubois and H. Prade. London, Academic Press Limited**:** 287-326.

Dubois, D. and H. Prade (1988). Possibility Theory: An Approach to Computerized Processing of Uncertainty. New York, NY, Plenum Press.

Frank, A. U. (1995). Strategies for the Introduction of GIS. Basic Concepts of GIS (ISPRS), Budapest.

Frank, A. U., E. Grum, et al. (2004). Procedure to Select the Best Dataset for a Task. Third International Conference on Geographic Information Science (GIScience), Maryland.

Grum, E. and B. Vasseure (2004). How to Select the Best Dataset for a Task? International Symposium on Spatial Data Quality (ISSDQ), Bruck a. d. Leitha, Austria, Department of Geoinformation and Cartography.

Guptill, S. C. and J. L. Morrison, Eds. (1995). Elements of Spatial Data Quality, Elsevier Science, on behalf of the International Cartographic Association.

ISO 19113 (2002). Geographic Information - Quality Principles.

Jahn, M. (2004). User needs in a Maslow Schemata. International Symposium on Spatial Data Quality (ISSDQ), Bruck a. d. Leitha, Austria, Department of Geoinformation and Cartography.

Krek, A. (2002). An Agent-Based Model for Quantifying the Economic Value of Geographic Information. Vienna, Austria, Department of Geoinformation/TU Vienna.

Navratil, G. (2004). How Laws affect Data Quality. International Symposium on Spatial Data Quality (ISSDQ), Bruck a.d. Leitha, Austria, Department of Geoinformation and Cartography.

Navratil, G. and A. U. Frank (2003). Modeling Processes Defined by Laws. 6th AGILE Conference on Geographic Information Science, Lyon, France, Presses Polytechniques et Universitaires Romandes.

Navratil, G. and A. U. Frank (2005). Influences Affecting Data Quality. International Symposium on Spatial Data Quality (ISSDQ), Peking, China.

Navratil, G., F. Twaroch, et al. (2005). Complexity vs. Security in the Austrian Land Register. CORP 2005 & Geomultimedia05, Vienna, Austria, Selbstverlag des Institutes für EDV-gestützte Methoden in Architektur und Raumplanung.

Pontikakis, E. and A. U. Frank (2004). Basic Spatial Data according to User's Needs-Aspects of Data Quality. International Symposium on Spatial Data Quality (ISSDQ), Bruck a.d. Leitha, Austria, Department of Geoinformation and Cartography.

Twaroch, C. (2005). Richter kennen keine Toleranz. Intern. Geodätische Woche, Obergurgl, Wichmann.

Twaroch, C. and G. Muggenhuber (1997). Evolution of Land Registration and Cadastre. Joint European Conference on Geographic Information.

Viertl, R. (2006). Fuzzy Models for Precision Measurements. Proceedings 5th MATHMOD, Vienna, ARGESIM / ASIM.

viertl, R. and D. Hareter (2006). Beschreibung und Analyse unscharfer Information. Vienna, Springer.

Wilson, N. (2002). A Survey of Numerical Uncertainty Formalisms, with Reference to GIS Applications. Annex 21.1 to REV!GIS Year 2 Task 1.1 deliverable.

Zadeh, L. A. (1975). "Fuzzy Logic and approximate Reasoning (In Memory of Grigore Moisil)." Synthese **30**: 407-428.

Zadeh, L. A. (1978). "Fuzzy Sets as a Basis for a Theory of Possibility." Fuzzy Sets and Systems **1**: 3-28.

Zadeh, L. A. (1979). A Theory of Approximate Reasoning. Machine Intelligence, Vol. 9. J. E. Hayes, D. Michie and L. I. Mikulich. New York, Elsevier**:** 149-194.