Palacký University, Olomouc

Faculty of Science

Department of Geoinformatics

Study programme: **P1304 Geography**

Specialization: **1302V011 Geoinformatics and Cartography**

UNCERTAINTY PROPAGATION IN FUZZY SURFACE ANALYSES

Dissertation thesis

Mgr. Jan CAHA

Supervisor: Doc. Mgr. Jiří DVORSKÝ, Ph.D.

Olomouc 2014

Author's Statement

I declare that this PhD thesis has been completed independently, under the supervision of Doc. Mgr. Jiří Dvorský, Ph.D. All the materials and resources are cited with regard to the scientific ethics, copyrights and the laws protecting intellectual property. This thesis or its parts were not submitted to obtain any other or the same academic title.

Olomouc, August 11, 2014 Mgr. Jan CAHA

Acknowledgement

Above all, I would like to thank to my family and my friends for the support not only during my studies but also generally in life.

I am also very grateful to my supervisor Jiří Dvorský for all the advice, consultations and continuous support during my work on this thesis. I even more appreciate his attitude towards me and my work.

A lot of my gratitude goes to Pete Fisher, not only for providing countless consultations, opinions, ideas and amazing advice related to my work but also for being so supportive towards my research. I am also very grateful for the chance to be Pete's colleague and friend. My internship in Leicester was by far the best academic experience of my life. I also thank to Jill Fisher for making my stays in Leicester as pleasant as possible.

My colleague and friend Alena Vondráková also deserves a lot of my thanks for helping me to pass all the administrative procedures, providing countless advice but above all for being a friend who understands me as very few people do.

I would also like to thank to several students that I had the chance to work with, so I would like to thank especially to Martin, Tomáš, Veronika and the others for never-ending discussions about my work, GIS and anything else. It was a great experience and I always enjoyed it. I am also very grateful to Peťa for the help with my English and for proofreading the whole text.

Table of Contents

List of symbols

$\tilde{A}$	fuzzy set or fuzzy number
$\mu_{\tilde{A}}$	membership function
$\mu_{\tilde{A}}(x)$	membership value of element x
$\tilde{A}_\alpha$	α-cut of $\tilde{A}$
$\tilde{A}_{0.5}$	0.5-cut of $\tilde{A}$ (specific α-cut)
$\mathrm{supp}(\tilde{A})$	support of $\tilde{A}$
$\mathrm{core}(\tilde{A})$	core of $\tilde{A}$
$\mathrm{hgt}(\tilde{A})$	height $\tilde{A}$
A	interval
$[\underline{a}, \overline{a}]$	interval described by the endpoints $\underline{a}, \overline{a}$
$\chi_{\tilde{A}_\alpha}$	characteristic function of the α-cut of a fuzzy set $\tilde{A}$
$\chi_{\tilde{A}_\alpha}(x)$	characteristic value of an element x to the α-cut of a fuzzy set $\tilde{A}$
Π	measure of possibility
$\mathcal{N}$	measure of necessity
(z^-, z^m, z^+)	alternative description of a triangular fuzzy number $\tilde{A} = (min, modal, max)M$

Introduction

Absolute certainty is a privilege of uneducated minds–and fanatics. It is, for scientific folk, an unattainable ideal.

– Cassius J. Keyser (Keyser, 1922)

During last 70 years or so, there has been a great development of mathematical theories, tools, and of course, their practical applications. These practical applications led to the emerge of computer science which later enabled the creation of many fields, including amongst others, also geoinformatics. The first experiments with geographical information systems (GIS) date back to the 1960s. It was mainly the work of Tomlinson (1974), who developed the first GIS in the world. Since then GIS has changed significantly, from rather simple tools designed for specific operations into a complex software handling all possible types of data and also operations and analyses with those data with impact on many fields where GIS is used mainly as a tool for solving practical problems (Goodchild, 2000).

The development of computers allows the users to perform the operations that were either impossible or took a significant amount of time just 20 year ago, in the matter of seconds. Fisher (2007b) reported that in his previous research (Fisher, 1991, 1992) the calculation of visibility on a small dataset (a grid of 200×200 cells) lasted up to 1 hour. Today, it is no problem to calculate the analysis on a grid four times bigger (400×400 cells) under one second. This example nicely illustrates the progress that was made during these years. The progress is mainly a result of better hardware but also partially of better software because undoubtedly today's algorithms are more optimized than the older ones. In spite of this development there are still issues embedded deeply within GIS that are widely unrecognized besides the scientific research.

One of these issues is uncertainty of spatial data and the problem of uncertainty propagation through the operations and analyses of such data (Devillers et al., 2010). Fisher (2007a) includes the issues of uncertainty, including fuzzy formalisms, amongst five most important research themes of representation in geosciences. The research regarding uncertainty and related topics has existed in geosciences since the late 1980s (Heuvelink et al., 1989). However, most of the research is connected to uncertainty in geodata and the topic of uncertainty propagation is limited to the usage of statistical methods (Zhang and Goodchild, 2002; Shi, 2010). That is a result of a long tradition of utilization of these methods in various fields of science (Hanss, 2005). Devillers et al. (2010) state that while spatial data have a number of very specific aspects, there is still a lot of good ideas about data uncertainty and quality to get from the research done in other fields. This idea mainly regards modelling of uncertain data, however in this thesis it is extended also on the methods of uncertainty propagation. Two main methods are used in geoinformatics – the analytical approach and the Monte Carlo method (Zhang and Goodchild, 2002). This thesis presents the usage of another method – fuzzy arithmetic (or fuzzy mathematics according to Fisher and Tate (2006)) that has been successfully used in other fields like engineering (Hanss, 2005), the weapon systems evaluation (Chen, 1996) and the land-cover accurancy assessment (Sarmento et al., 2013). As mentioned by Goodchild (2000), the early work in uncertainty in GIS was based heavily on the probability theory and the probabilistic methods (e.g. work by Heuvelink (1998)), but later it became clear that also another frameworks (including fuzzy sets) are needed to handle all aspects of uncertainty. The selection of the best framework, despite their incompatibilities, should depend purely on the problem at hand.

Fuzzy arithmetic and fuzzy numbers were introduced in the late 1970s (Nahmias, 1978; Dubois and Prade, 1978, 1979). The topic was frequently discussed in mathematical and computer science literature during the eighties and nineties (Dubois and Prade, 1983; Kaufmann and Gupta, 1985; Baekeland and Kerre, 1988; Buckley and Qu, 1990; Zimmermann, 1991) but it was not until 1995 when a detailed computer implementation was described (Anile et al., 1995). The missing practical implementation is one of the reasons why fuzzy arithmetic is the least developed subfield of the fuzzy set theory research (Hanss, 2005). Even though some implementations do exist (Gagolewski, 2014; Anile et al., 1995; Fonte et al., 2008b; Spinella, 2008), none of them is actually suitable to be used for complex programmes. For the other subfields of fuzzy set theory like fuzzy logic, some implementations that can be used even for complex programmes do exist (Cingolani and Alcalá-Fdez, 2012; Cingolani and Alcalá-Fdez, 2013). The existence of such implementation (usually in the form of a software library) helps with the expansion of mathematical theory into practical applications.

Curiously, fuzzy arithmetic has been employed only rarely despite the fact that the usage of fuzzy set theory has been widely researched in geoinformatics for various purposes (Fisher and Tate, 2006). Since there are methods for the creation of fuzzy surfaces (Diamond, 1989; Bardossy et al., 1990a), the next logical step would be to analyse these surfaces. A fuzzy surface can be analysed only with the usage of fuzzy arithmetic which so far has been done only in a few examples (Anile et al., 2003; Waelder, 2007; Caha et al., 2012). Goodchild (2000) is pointing out that fuzzy sets and the associated framework provide a comprehensive approach for handling uncertainty in modelling, reasoning and analysing uncertain data but he also mentions that there is more research to be done before the benefits of such research become meaningful to the majority of GIS users. This practically means that a lot of research involving a good amount of experiments needs to performed in order to figure the best possible approach to the issue of uncertainty in geosciences. Only after finishing this part of research is it sensible to start the explanation of outcomes to the majority of GIS community. However, so far only a relatively small portion of existing mathematical tools and approaches for handling uncertainty has been tested within the scope of geographic information science (GISc).

Reitsma (2013) states that GISc has been defined as a branch, or subfield, of the information science with a connection to many other disciplines, such as statistics (through spatial statistics), computer science, geography and other related fields. The same can be deduced from the information provided by Fisher (2007a). Raper (2009) describes that before 1999 there were only a few collaborations between GISc and classic information science because the information science did not account for the role of geographic information. In the early to mid-1990s GISc was considered a part of geography but the later development clearly proved that GISc actually covers more than just topics of classic geography (Raper, 2009). Obviously, cartography and geodesy can be perceived as original predecessors of GIS but unfortunately several cartographic dogmas and influences are actually slowing down the progress of GISc (Fisher, 1998b). In order to move forward, GISc needs to embrace its place within the information science with all the consequences coming with it. That especially includes the necessity to introduce new knowledge from other fields (Devillers et al., 2010) like mathematics and computer science into our own research and even everyday practice. In this thesis the main focus is to describe fuzzy arithmetic, to explain how it can be used for the uncertainty propagation in fuzzy surface analyses, to describe the obtained results and their further usage in the decision making process. Besides that, a brief comparison of the uncertainty propagation techniques and uncertainty modelling theories is provided.

Chapter 1

Aims of the thesis

The main aim of this thesis is to define the utilization of fuzzy arithmetic as a method for uncertainty propagation for analyses of fuzzy surfaces. This aim will be completed by the presentation of calculation of slope, aspect and visibility on fuzzy surfaces. In order to achieve this main goal several minor goals need to be reached:

- Description of uncertainty from the mathematical point of view. Listing of the mathematical theories that can be used for the uncertainty modelling and propagation with their brief comparison.
- Summary of theoretical foundations of the fuzzy set theory and fuzzy arithmetic necessary for their practical usage in the surface analyses.
- Analysis of the mathematical methods that can be and are used in geosciences for modelling uncertainty and the uncertainty propagation with a special focus on the topic of surfaces and their analyses.
- Presentation of methods, approaches and algorithms for the calculation of slope, aspect and visibility on fuzzy surfaces.
- Description of further utilization of the results of analyses of fuzzy surfaces in the decision making.
- Presentation of an illustrative case study to show the obtained results from the fuzzy surface analyses.

Together these minor steps will help to achieve the main goal.

As mentioned in the previous chapter, geographical information science is perceived as a part of information sciences that deals especially with spatial data. As a consequence of this fact, the thesis is an interdisciplinary text that draws mainly from the fields of mathematics, computer science, geoinformatics and also geography and other fields. Because of that, some parts of the thesis might be more technical than it is usual in the field of geoinformatics. However, these technical sections are necessary to achieve the main goal of the work.

The author intends to present fuzzy arithmetic as a method having a great potential usage in geoinformatics due to some of its properties distinguishing it from the currently used methods of uncertainty propagation. However, it is not the object of this thesis to compare the methods for the uncertainty propagation as competitive methods because they are not. The author would rather introduce a new method for the uncertainty propagation, with its set of advantages but also disadvantages, to the existing set of methods used in geosciences with a hope that fuzzy arithmetic will find its usage and will became a substantial part of the topic of uncertainty propagation.

Chapter 2

Uncertainty

As far as the propositions of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality. It seems to me that complete clarity as to this state of things became common property only through that trend in mathematics which is known by the name of "axiomatics".

– Albert Einstein (Einstein, 1954)

Uncertainty expresses the inability of precise description and/or expression due to variability, ambiguity, imprecision and/or vagueness of information (Klir and Yuan, 1995). Uncertainty can be also seen as ignorance or incomplete knowledge that does not allow precise statements (Dubois and Prade, 1986). It affects all data, information and of course all models of reality (Bandemer, 2006). Because all the data and models contain some kind and amount of uncertainty, there exists a need to incorporate uncertainty into these models and process it through the analysis in order to obtain complete results.

Uncertainty has always been present in the scientific world in one way or another. The Greek mathematician Archimedes devised the value of π as $\frac{223}{71} < \pi < \frac{22}{7}$ (Arndt and Haenel, 2006) around year 250 BC. This definition of π involves uncertainty because Archimedes was unable to calculate it with better precision. It can be also considered as the first example of an interval uncertainty in science (Lodwick et al., 2008). However, such definition of π is sufficient for many applications, it allows the calculation of circumference of a circle based on its diameter with a precision that is sufficient for some purposes. Obviously, more precise definitions of π were later introduced by various mathematicians, but it is a well known fact that the value of π cannot be expressed as a precise number. Even though 39 decimal digits are sufficiently precise even for cosmological calculations, the current number of the known digits of π is about 10^{14} (Arndt and Haenel, 2006).

Until the late 19$^{\text{th}}$ century uncertainty in science was seen as something that should be eliminated because scientific facts should be presented in precise numbers and statements (Klir, 2006). Around that time a new field of physics called statistical mechanics emerged, a domain where precise calculations were replaced by statistical methods and reasonable assumptions (Klir, 2006). It was the first occasion when the approach towards uncertainty in science was revised and started to be considered as useful and even essential in certain scientific inquiries (Klir, 2006).

In 1927 Heisenberg formulated his *uncertainty principle* which proved that in some situations uncertainty is an inevitable part of observation (Ayyub and Klir, 2006). This was yet another significant step in acknowledging that uncertainty is an important part of science.

Before World War II the only practical approach for handling uncertainty was probability, after the war the development of technology (mainly computers) and mathematical theories enabled the creation of new theories and approaches to uncertainty (Klir, 2006). These new theories and methods included for example the Monte Carlo method (Metropolis and Ulam, 1949), the interval arithmetic (Moore, 1966), the fuzzy sets (Zadeh, 1965), the Dempster-Schafer method (Dempster, 1967) and another theories.

In 1954 Albert Einstein said: *As far as the laws of mathematics refer to reality, they are not certain, as far as they are certain, they do not refer to reality* (Einstein, 1954). This statement points to the fact there are no completely perfect models and that every treatment of problems by mathematical instruments contains some amount of uncertainty. Tukey (1962) said that *An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate problem.* This quote summarizes the fact that very often a lot of effort is spent on obtaining very precise (numerically-wise) results while the fitness of the question (that was asked) is not discussed at all. Both these researchers were aware of the importance of uncertainty in science and wanted to stress this fact to the scientific community.

The work of Shannon lead to the universal acceptance that information is statistical in nature (Zadeh, 2006). In combination with the introduction of the Monte Carlo method this fact caused the spread of statistical methods through the different fields of science as the most common approach to model uncertainty. The Monte Carlo method allowed the propagation of uncertainty through complex models and it was quickly adopted as a possible solution for many applications. However many authors later provided examples of situations for which the statistical approach is not suitable (Dubois and Prade, 1986; Moore, 1966; Zadeh, 1965, 2005).

Zadeh (1973) postulated the *Principle of Incompatibility* as: stated informally, the essence of this principle is that as the complexity of a system increases, our ability to make precise and yet significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics. Or in other words it can be phrased as: The closer one looks at a real world problem, the fuzzier the solution appears (Zadeh, 1973). This principle implies that every problem involves uncertainty if it is studied in sufficient detail. The main idea presented by this principle could be also perceived as a statement that everything is uncertain but it is the amount of uncertainty and its influence on the outcomes that matters the most. In addition to the actual problem that is to be modelled, there is also a question regarding uncertainty that should be raised. This includes mainly the questions: *a*) Where does uncertainty originate in the model? *b*) What are the impacts of uncertainty on the results? *c*) How can uncertainty in the model be described? *d*) What level of certainty do the results need to have? Only after answering these questions and finding the answers the user should continue with their analysis.

2.1 Components of uncertainty

Uncertainty is a complex phenomenon that consists of several components (Celikyilmaz and Turksen, 2009). However, there is no agreement between the authors about the number of components or even about the components themselves. Celikyilmaz and Turksen (2009) name eight components, while Klir and Yuan (1995) mention three and Viertl (2011) only two (Fig. 2.1). The important thing is that all these authors agree that there are two main types of uncertainty. The first one originates in variability/randomness and the second in ambiguity/vagueness/imprecision (Klir and Yuan, 1995; Dubois and Prade, 1986). Helton and Oberkampf (2004) distinguish these two components as aleatory and epistemic uncertainty. Aleatory uncertainty is the one that originates in variability and randomness and epistemic uncertainty is a result of lack of knowledge (Helton et al., 2004). These two types define two main theories used to handle uncertainty. Statistics and its methods should be used to handle uncertainty as a result of variability and fuzzy theory (or in some cases the possibility theory) and other methods closely connected to it are suitable in the case of uncertainty being a result of incomplete knowledge (Zadeh, 1995, 2005; Viertl, 2011). However, this distinction of the uncertainty types is a relatively new approach. For a long time both aleatory and epistemic uncertainties were modelled by statistics (Helton and Oberkampf, 2004).

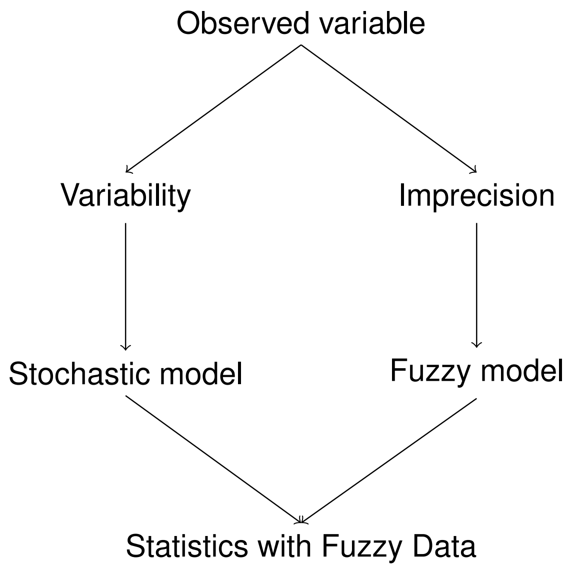


Figure 2.1: The components of uncertainty of the variable.
(according to: Viertl (2011))

2.2 The influence of uncertainty on a model

Every problem that should be treated scientifically requires a model representing a mathematical description of the problem and methods to derive some conclusions based on the model and the data. According to Bandemer (2006), there are three main components for a mathematical treatment of the problem: *a*) a mathematical model, *b*) a mathematical solving procedure, *c*) data. All of these are connected to the process of translating the real world problem into a mathematically processable model. Uncertainty in general is connected to all these components so that every step of a mathematical treatment of a problem should be considered in the context of uncertainty affecting this specific component. Different types of uncertainties are associated with the model, data and solving procedures.

For example the mathematical model of the real world suffers from two effects that may introduce uncertainty. These effects are simplification and idealization (Hanss, 2005). Idealization occurs in the process of translating the real world into a mathematical model because the model is never exhaustive, only the most important parts of reality are translated into it. Simplification occurs in models as the intentional uncertainty introduced by the model's creator. The reasons might vary from attainment of analytical solutions, reduction of calculation time to applicability of existing theories (Hanss, 2005).

Solving procedures used in the mathematical model are not often mentioned as a source of uncertainty but actually a lot of computing methods provide only approximations or estimations of the actual result (Bandemer, 2006). As a result, uncertainty can originate even in a purely computational part of the model.

Data are probably the most common source of uncertainty. Data as understood here do not include only the data passed as arguments to the model, for example the initial conditions, but also any constants in the model. Even the constants might be subjects to uncertainties, as they are sometimes based on the experts' knowledge.

Data may contain all types of uncertainties that will later affect the results of the model. There are many books related to the topic of modelling uncertainty in data, however most of them introduce only one specific method for modelling uncertainty (Celikyilmaz and Turksen, 2009; Dubois and Prade, 1986; Drosg, 2007; Kadane, 2011; Lindley, 2006; Moore et al., 2009) or they are very general (Halpern, 2003). There are also books related to modelling uncertainty directly from the measured data (Gupta, 2012). In geographical information science there are several books covering the topic of uncertainty modelling (Heuvelink, 1998; Shi, 2010; Zhang and Goodchild, 2002).

2.3 Mathematical theories for modelling uncertainty

There are several mathematical theories that are able to capture uncertainty. According to Oberguggenberger (2005), there are three main aspects that the uncertainty theory has to satisfy to be considered complete. The theory is obliged to have well established: *a*) definitions and axiomatics, *b*) numerics, *c*) semantics. Definitions and axiomatics tell us how uncertainty is described and what are the combinational rules for uncertainties. Numerics describes how uncertainty is propagated through the computational model and semantics describes the meaning of results and what do these results mean in the concept of reality.

There are several theories carrying all three aspects of uncertainty modelling and that have been used in practical applications. Oberguggenberger (2005) mentions these theories: *a*) deterministic values, *b*) interval arithmetic, *c*) probability and sets of probability measures, *d*) random sets (Dempster-Shafer theory), *e*) fuzzy set, fuzzy arithmetic and the possibility theory. Besides those, there exist several newer theories that are still developing and have not yet to prove their usability in practical applications. Those theories are: *a*) the Uncertainty theory (Liu, 2010), *b*) the Generalized theory of uncertainty (Zadeh, 2006). There are also other theories existing but currently they are not widely accepted and used in practical applications.

2.3.1 The uncertainty propagation

The propagation of uncertainty is a process that assesses uncertainty of the result of the model based on uncertainty of the data and uncertainty of the model itself (Crosetto and Tarantola, 2001). According to the definition of aspects of uncertainty by Oberguggenberger (2005), the process of uncertainty propagation is numerics of the modelling theory. The results of uncertainty propagation are the main reasons for modelling and calculation with uncertainty.

The outcome of the model Y is a function $f()$ of the input variables $X_1, \dots, X_n$ and model parameters $P_1, \dots, P_m$:

$$Y = f(X_1, \dots, X_n, P_1, \dots, P_m).$$

The processes of uncertainty propagation determines how uncertain is the output Y based on uncertainties of the input variables and model parameters.

The concrete calculation is done according to the theory used for the uncertainty modelling. A quite common technique used with the statistical description of uncertainty is the Monte Carlo method (Helton and Oberkampf, 2004), calculation with fuzzy numbers can be done by the usage of fuzzy arithmetic (Kaufmann and Gupta, 1985). The other theories of uncertainty have their own methods allowing uncertainty propagation, otherwise such theory could not be considered complete (Oberguggenberger, 2005). Depending on the used uncertainty method, the results of the uncertainty propagation might not be the same because different methods search for diverse results.

2.4 Semantics of uncertainty

Zadeh (2006) noted that as science moves towards the automated decision making a basic limitation of the probability theory becomes a problem. The same author also noted that there exists a sustaining idea that any information is statistical in nature (Zadeh, 2005). However, lately there have been many researches pointing out limitations and problems that could be brought by the purely statistical handling of uncertainty. One of these is called the *law of decreased credibility*, stating that the credibility of inference decreases with the strength of the assumptions maintained (Oberguggenberger, 2005).

It is also known that the probabilistic models often require more information than the user actually has. A specification of probability distributions for uncertain model parameters is a classic example because the user often does not have the necessary amount of information to specify the distribution (Helton and Oberkampf, 2004). In such situations it is preferable to conceptualize uncertainty in an alternative theory that might be better adapted to the uncertainty type and could avoid unwarranted assumptions (Helton and Oberkampf, 2004; Oberguggenberger, 2005). As will be shown later, these assumptions play an essential role in the topic of uncertainty propagation for the surface analysis.

Oberguggenberger (2005) points out that the interpretation of uncertainty theory is an essential ingredient for the translation from reality to the model and vice versa. Also the statements derived from the model are meaningful only in the context of the underlying semantics (Oberguggenberger, 2005). Different semantics have different meanings which causes the results obtained by various theories to be directly incomparable. If the results are to be compared it is necessary to compare not only the values but also their semantics and consequences.

Semantics of uncertainty provides the pointers towards the theories that should be used for modelling of uncertainty. In certain situations one method allows to capture uncertainty completely but some problems might be so complex that they require the combination of several theories together. An example might be the fuzzy statistics (Viertl, 2011) or statistics under the interval and fuzzy uncertainty (Nguyen et al., 2012).

2.5 Example

Previously it was mentioned that the selection of the uncertainty theory depends on the semantics of uncertainty and also the combination rules for uncertainty. In this chapter a simple practical example will be presented, with various types of uncertainty shown. The types of uncertainty that will be used in this example are: *a*) deterministic values, *b*) intervals, *c*) variability (statistics), *d*) vagueness/imprecision. Semantics of the input and results will be discussed and the obtained resulting values will be compared. For the sake of maintaining simplicity the details of this calculation will not be shown, however the calculations will be done according to the definitions specified in Moore et al. (2009) in case of interval arithmetic, Hanss (2005) for fuzzy arithmetic and Rubinstein and Kroese (2008) in a case of calculation with statistic uncertainty. The main aspect determining the selection of method for the uncertainty modelling will be semantics of uncertainty. An example in the same sense is provided by Lodwick et al. (2008) to compare the calculations of the Malthus law under probabilistic, interval and fuzzy uncertainty. Hanss (2005) performed an experiment of the calculating model:

$$z = f(x_1, x_2, x_3) = \sin(x_1) + x_2^2 - x_3,$$

with the uncertain values represented by the guassian fuzzy numbers and normally distributed random variables. The random values and fuzzy numbers in the experiment had the same range of values. The Monte Carlo simulation was done with 10 000 iterations and the outcomes proved that the method almost neglected any extreme possible results when compared to the results of fuzzy arithmetic. This result illustrates the fact that the marginal cases with low probability for each variable have even lower probability to be combined with another value having small probability. This leads to the omission of results that are perfectly possible but have very small probability.

The problem to solve is simple, the gravitational force that applies on an object with mass m in a gravitational field of Earth is to be computed. To calculate this force F_g, the value of a gravitational constant g needs to be known. The force is then calculated according to the formula:

$$F_g = mg.$$

This equation is a model of gravitational force. For our purposes the value of m is considered as a variable and g as a parameter of the model. In fact both of these can be considered as variables under some circumstances, but for the sake of explanation m will be considered as a variable and g as a parameter of the model.

In the first case the deterministic values are used, meaning that both values necessary for the calculation are considered precisely known. The value of g is $9.80655\mathrm{m\,s^{-2}}$ and the mass of the object is 5kg. The resulting force is 49.03275N. In this case no uncertainty is involved, both values are considered accurately known. But as mentioned previously, this approach is rather naive because all the data contain some uncertainty.

Let's assume that the mass of the object was measured several times (at least ten times) and various values of the mass were obtained. From these data the probability distribution can be constructed. The normal (gaussian) distribution with the mean 5kg and the standard deviation 0.2kg is the resulting distribution of the mass of the object. This variability of the value can be caused by imprecise measurement equipment, human error or some other reason. The value of g is still considered a precise value. We can perform a simple Monte Carlo experiment where 500 random values will be selected from the probability distribution of the mass. For these five hundred values the calculations will be performed and the results can be statistically evaluated. The results have mean 49.04666N and the standard deviation 0.9609622N. The minimal value is 46.25667N and the maximal 52.14151N. In this case uncertainty is present in the variability of the mass of the object. As a result, the value of the force F_g is also uncertain. In this case uncertainty was caused by the variability (repeated measurements) and thus statistics was a suitable method for its modelling. This type of uncertainty is often written in the format 49.04666 ± 0.9609622N, however this notation is inappropriate because it does not stress that the values are statistical. The notation semantically points towards the description by a fuzzy number.

At this point let's assume that the mass of the object is known only approximately. It is known that the mass is higher than 4.7kg and lower than 5.3kg. The value of g is again considered precise. In this case uncertainty is not caused by variability but by the lack of knowledge. Since there is no preference over the interval of the possible values, the suitable theory for modelling uncertainty is interval arithmetic. The mass m is represented as the interval $[4.7, 5.3]$ kg. The resulting force F_g will be represented as the interval $[46.09078, 51.97471]$N. The downside of the interval arithmetic is that it does not offer any preference measure of the result, it only provides bounds of the results.

The last case means that the value of the mass is known approximately, but with some preference measure. For example the mass was measured to have the value of 5kg but it is known that the precision of the measurement is ± 0.3 kg. Again, in this case uncertainty is caused by the imperfect knowledge, there is no variability involved. The assumption in this case is that the values close to what was actually measured are more likely to be correct. Consequently, the value of the mass can be modelled by a fuzzy number with the minimum value of 4.7kg, the modal value of 5kg and the maximal value of 5.3kg. The result will be also a fuzzy number with the minimal value of 46.09078N, the modal value of 49.04666N and the maximal value of 51.97471N. This result provides more information than the interval in the previous case because it shows that values close to the modal value of the fuzzy number are more likely to represent the proper solution than the limit values of the fuzzy number.

The basic difference between these examples lies in semantics of the results. The deterministic case does not acknowledge uncertainty at all. The statistic approach searches for the most probable result but cannot guarantee the provision of the bounds of the possible uncertainty. On the other hand, the interval analysis focuses on the bounds of the result. Fuzzy arithmetic provides the bounds in same way as interval arithmetic and also the most possible result. The presented example shows how different semantics of uncertainty specify which uncertainty theory is suitable for treating the problem and that all the theories can be used to handle the same problem.

2.6 Uncertainty in the geographical sciences

Uncertainty in various forms is naturally contained within geography as a subject, due to the fact that many geographical phenomena have vague or imprecise definitions (Fisher, 2000). Besides that, as mentioned previously, all the data and models contain uncertainty (Bandemer, 2006) and the geographical data and models are no exceptions to this rule (Shi, 2010; Zhang and Goodchild, 2002). The geographical data and analysis should recognize this uncertainty and emphasize the fact that the results are not absolutely precise but vague/imprecise (Fisher, 2000). However, in most of the cases this uncertainty of the data or uncertainty of the model are not acknowledged. Actually, both the data and the model contain some amount of uncertainty (Zhang and Goodchild, 2002), but the question emerges on the influence and significance of this uncertainty on the result. This influence can be evaluated by the uncertainty analysis and sensitivity analysis (Crosetto et al., 2000; Crosetto and Tarantola, 2001).

The topic of uncertainty modelling is well established in the geographical sciences (Shi, 2010; Zhang and Goodchild, 2002), the earliest attempts date into the late eighties and early nineties of the 20th century (Heuvelink et al., 1989; Heuvelink and Burrough, 1993). However, the practical implementation of the uncertainty propagation through the complex model is rather rare in geosciences (Fisher and Tate, 2006; Heuvelink, 2002) and almost exclusively limited to the scientific studies. Heuvelink (2002) mentions that in many cases of complicated models the uncertainty analysis does not pay off because it is more time-demanding than the analysis itself. The correct setting of the uncertainty propagation also requires a lot of knowledge that a common GIS user may not have, so there may be a need to somehow simplify the initial setting to make the analysis more comprehensible for the users (Heuvelink, 2002).

The topic of uncertainty modelling in geosciences is often related to the problem of data quality (Devillers et al., 2010). However, the theme is much broader because, as mentioned above, not only the data are the source of uncertainty but also the models and even the mathematical solving procedures can introduce some uncertainty into the outcome. When the topic of uncertainty propagation in geosciences is mentioned, two methods for performing it are usually named – the analytical approach and simulations (usually the Monte Carlo method) (Shi, 2010). Both these techniques are commonly used (Heuvelink, 1998; Oksanen and Sarjakoski, 2005a) but, as mentioned by Zhang and Goodchild (2002), the analytical approach is not practical for complex GIS applications. This fact leaves the Monte Carlo method as the main approach for the uncertainty propagation in geosciences (Heuvelink, 1998, 2002). As mentioned by Fisher and Tate (2006), even though the fuzzy sets and fuzzy logic have been widely researched in geosciences, the employment of fuzzy mathematics as a tool for the uncertainty propagation is very rare. However, several studies with the usage of fuzzy arithmetic were performed and all of them are related to the topic of surface derivatives (Fonte and Lodwick, 2005; Waelder, 2007; Caha et al., 2012). The reason why all these studies share the topic could be that the surface derivatives represent a very demonstrative example. Examples actually can be found, they are just very rare. Still, in every case the usage of fuzzy mathematics was found to be beneficial and useful. However, the calculations are more complex than in the case of the Monte Carlo method, which is probably the reason why Monte Carlo is preferred.

Devillers et al. (2010) mention that when studying the problem of spatial data quality, it might be useful to take a look at various other fields because the topic of data quality is not unique for geosciences. The knowledge from other fields is undeniably to be valuable because the basic concepts apply in all fields. The same thing can be stated for the topic of uncertainty propagation. The most important development is performed in the fields like mathematics, information sciences, informatics etc., and the geosciences should accept the knowledge from those fields and apply them to the spatial problems in order to figure out if these new methods and approaches are useful and have a potential for solving and overcoming problems (that geosciences currently have). Devillers et al. (2010) warn that if this new knowledge does not get into geosciences, it may happen that geosciences will become a microcosm that operates separately and for its own satisfaction. To avoid this problem, there is a need to include people, methods and approaches from other fields into geosciences to keep the field moving (Devillers et al., 2010).

Chapter 3

Fuzzy set theory

To what degree is something true or false?

– Lotfi A. Zadeh, attributed to Zadeh by Blair (1994)

As noted by Fisher (2000), vagueness is endemic in geographical thinking and geographical information. Because of that there is a need for tools and theories to model vagueness as a part of uncertainty. The fuzzy set theory provides instruments for describing vague or imprecise sets in a way that is relatively close to the style of human reasoning. Such sets that lack strict boundaries might be for example: *small numbers*, *cheap cars* or *suitable solutions*. All of these have in common that humans can understand them well but it is rather difficult to describe them in mathematical terms. But with the development of fuzzy set theory and fuzzy logic not only the description but also reasoning with such sets became possible (Zadeh, 1975). Further development introduced tools for modelling of ill-known numbers and algebraic operations with them (Dubois and Prade, 1978). Fuzzy arithmetic is an extension of interval arithmetic proposed by Moore (1966) that expands options by introducing a preference measure that is not included in interval arithmetic. Fisher and Tate (2006) are emphasizing that the topic of fuzzy set theory and its usage in geosciences has received a lot of attention but fuzzy mathematics was so far not used to analyse fuzzy surfaces. The following chapters summarize the necessary theoretical basis for the use of fuzzy arithmetic in terms of fuzzy surface analysis.

3.1 Fuzzy set

A fuzzy set is a generalization of a classic set that allows elements not only to belong or not to belong to a specific set, but it also allows a partial membership to the set (Zadeh, 1965). Let X be a set of objects that is called space (or universe), and elements of that space that are denoted as x. A classic set A is defined as a collection of objects $x \in X$ that do have certain property or properties. Each element either belongs or does not belong to such set. A fuzzy set is a generalization of a classic set and it is defined by a membership function $\mu_{\tilde{A}}$ that maps the elements of X on the values from the interval $[0, 1]$:

$$\mu_{\tilde{A}}(x) \to [0, 1]. \tag{3.1}$$

The membership value $\mu_{\tilde{A}}(x)$ then describes how much the element x belongs to the set $\tilde{A}$. The value of $\mu_{\tilde{A}}(x) = 1$ indicates a full membership to the set, $\mu_{\tilde{A}}(x) = 0$ denotes that the element does not belong to the set at all. All remaining values indicate a partial membership of the object to the set on a specific membership value. Obviously, the higher is the membership value, the higher is the membership to the set (Fig. 3.1). According to Klir and Yuan (1995), the usage of other intervals to describe the membership is possible but not very common.

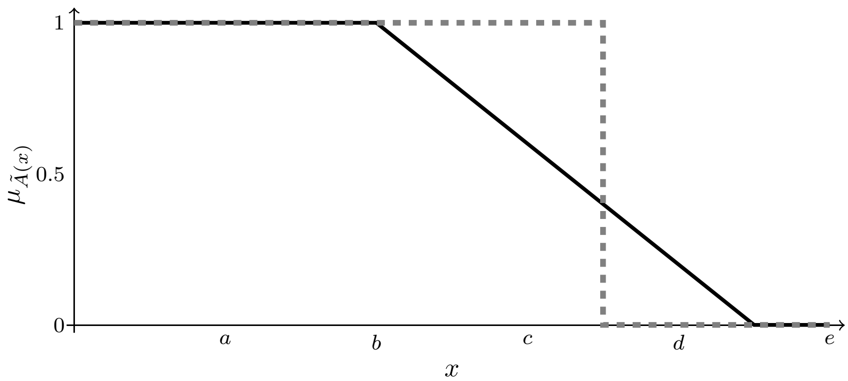


Figure 3.1: A crisp set (dotted gray line) and a fuzzy set (black line).

If $\tilde{A}$ is a finite set, the fuzzy set can be characterized by a set of pairs:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\}. \tag{3.2}$$

If the fuzzy set is not finite then it can be written as (Dubois and Prade, 1980):

$$\tilde{A} = \int_x \mu_{\tilde{A}}(x)/x. \tag{3.3}$$

The membership of a fuzzy set may have various forms (Klir and Yuan, 1995; Klimke, 2006), for example:

$$\mu_{\tilde{A}}(x) = \frac{1}{1 + (x - 5)^4}, \tag{3.4a}$$

$$\mu_{\tilde{B}}(x) = \begin{cases} 1 & \text{if } x \leq 10 \\ \dfrac{20 - x}{10} & \text{if } 10 < x < 20 \\ 0 & \text{if } x \geq 20 \end{cases} . \tag{3.4b}$$

3.2 α-cuts of a fuzzy set

An α-cut of a fuzzy set is a crisp (classic) set of elements from X that have the membership value higher or equal to α (Fig. 3.2). Some authors also refer to it as an α-level set (Klimke, 2006). Formally, an α-cut is written as (Zadeh, 1975):

$$\tilde{A}_\alpha = \{x \mid \mu_{\tilde{A}}(x) \geq \alpha\}, \tag{3.5}$$

where $\alpha \in [0, 1]$.

As visible from the definition, an α-cut is no longer a fuzzy set but a crisp set (Dubois and Prade, 1980) that can be written as an interval $\tilde{A}_\alpha = a \;= [\underline{a}, \overline{a}]$ with its characteristic function $\chi_{\tilde{A}_\alpha}$ (Lodwick et al., 2008). In some cases the interval could be degenerative, meaning that $\underline{a} = \overline{a}$ (Moore et al., 2009). The fact that an α-cut is a crisp set is very useful for further processing. There also exists a strong α-cut where the condition is more strict $\mu_{\tilde{A}}(x) > \alpha$ (Hanss, 2005). The natural property of an α-cut $\tilde{A}_{\alpha 1}$ is that it is a subset of $\tilde{A}_{\alpha 2}$ if $\alpha 1 > \alpha 2$ (Dubois and Prade, 1980). As a result of this property, α-cuts are referred to as the nested sets (Zimmermann, 1991).

There are two α-cuts that have a specific meaning. The core $(\text{core}(\tilde{A}))$ or the kernel of a fuzzy set is a set of x which have $\mu_{\tilde{A}}(x) = 1$ and the support $(\text{supp}(\tilde{A}))$ is a set of x where $\mu_{\tilde{A}}(x) > 0$. Both of them have special meaning as the former describe the elements with a full membership and the latter global extent of a fuzzy set.

As noted by Klimke (2006), for practical computational applications it is useful to also define the compact support of $\tilde{A}$ as:

$$\tilde{A}_0 = [\inf(\text{supp}(\tilde{A})), \sup(\text{supp}(\tilde{A}))], \tag{3.6}$$

which only makes sense if the support of $\tilde{A}$ is bounded. It can later be referred to as a 0-cut of $\tilde{A}$ denoted as $\tilde{A}_0$. Even though technically the zero α-cut of a fuzzy set is the universe X, it is very useful to perceive the 0-cut as a strong α-cut especially for practical implementation. The compact support will be of special use for the processing of fuzzy numbers.

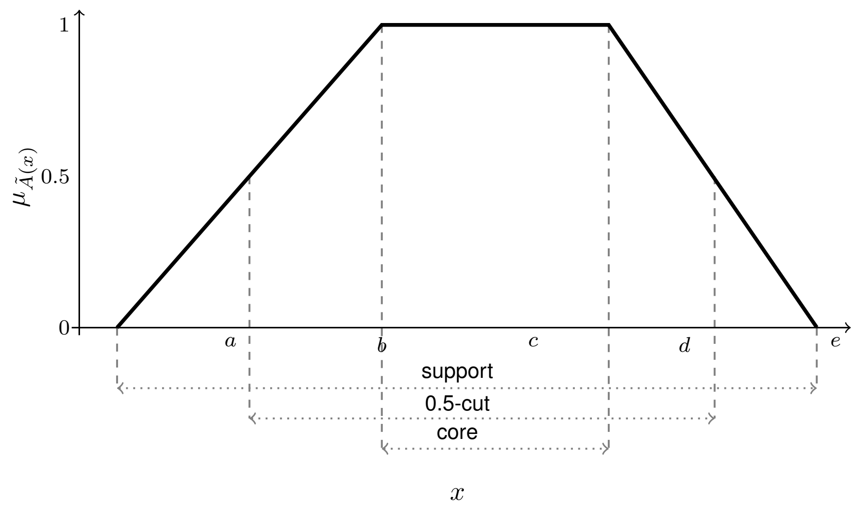


Figure 3.2: The α-cuts of a fuzzy set.

3.3 The decomposition theorem

The decomposition theorem states that every fuzzy set can be represented by an associated sequence of its α-cuts. Zadeh (1971) named this property of a fuzzy set as a resolution identity. Formally it is defined as (Hanss, 2005):

$$\mu_{\tilde{A}}(x) = \sup_{\alpha \in [0,1]} \chi_{\tilde{A}_\alpha}(x). \tag{3.7}$$

The decomposition theorem is very useful because it allows a transformation of description of a fuzzy set from the membership function $\mu_{\tilde{A}}$ to the α-cut intervals and vice versa. It is also often used in mathematical proofs because if a proposition can be proved to be true for every α-cut then it is also true for the whole fuzzy set. This is often more easy to prove than proving the proposition directly for a fuzzy set.

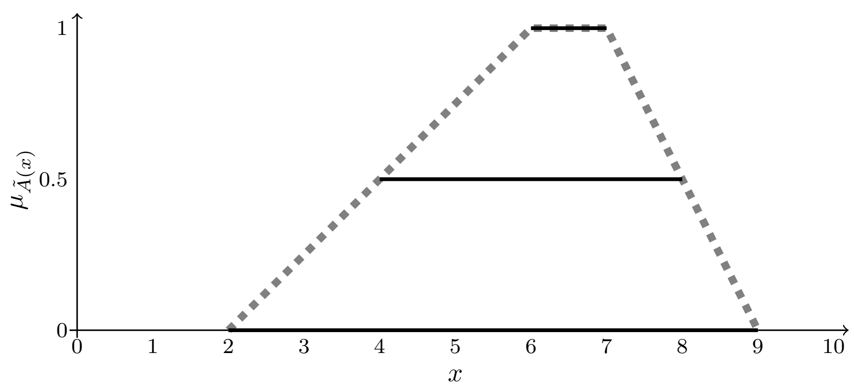


Figure 3.3: A fuzzy number described by a membership function (grey dashed line) and three α-cuts (black thick line) with the membership values 0, 0.5 and 1.

The decomposition theorem specifies how to separate a fuzzy set into a theoretically infinite number of α-cuts (Fig. 3.3). However, for practical applications there needs to be a finite number of α-cuts. The concrete values of α depend on the number m of the intervals to which the interval $[0,1]$ should be subdivided. m specifies the number of intervals with the length $\Delta\mu$ according to (Hanss, 2005):

$$\Delta\mu = \frac{1}{m}. \tag{3.8}$$

The discrete values of μ_j are given by the equation:

$$\mu_j = \frac{j}{m}, \qquad j = 0, \ldots, m. \tag{3.9}$$

The fuzzy set $\tilde{A}$ is then decomposed into a set of intervals:

$$\tilde{A} = \{A^0, \ldots, A^m\}, \tag{3.10}$$

where the superscript A^0 denotes the order of the α-cut. Each such decomposition consists of $m+1$ α-cuts, where A^0 is the support (or compact support) of the fuzzy set and A^m the core of the fuzzy set. The minimal number of m is 1 which, according to Eq. (3.9), results into two α-cuts with the values of $j=0$ and $j=1$. These two α-cuts are the least necessary minimum required to describe any fuzzy set. Obviously, for some fuzzy sets some of the α-cuts can be an empty set.

3.4 Properties of a fuzzy set

A fuzzy set has two main properties – height and convexity. Height refers to a maximal value of $\mu_{\tilde{A}}(x)$ that can be found in X.

$$\mathrm{hgt}(\tilde{A}) = \sup\{\mu_{\tilde{A}}(x) \mid x \in X\} \tag{3.11}$$

If the height $\mathrm{hgt}(\tilde{A}) = 1$, then the fuzzy set is called normal, otherwise it is called subnormal (Fig. 3.4).

For a fuzzy set, unlike for a crisp set, convexity refers to the properties of the membership function rather than to the support of a fuzzy set (Klimke, 2006). The original definition of a convex fuzzy set was provided by Zadeh (1965):

$$\mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geq \min((\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2))) \quad \forall\lambda \in [0,1], \forall x_1, x_2 \in X, \tag{3.12}$$

but Klimke (2006) provided a more comprehensible definition. A fuzzy set is convex if $\forall a, b, c \in X$, where $a \leq b \leq c$ and it fulfils:

$$\mu_{\tilde{A}}(a) \leq \mu_{\tilde{A}}(b) \text{ and } \mu_{\tilde{A}}(b) \geq \mu_{\tilde{A}}(c). \tag{3.13}$$

Otherwise it can be stated that every α-cut of $\tilde{A}$ must be convex in the sense of a classic set theory. If a fuzzy set is not convex it is called non-convex (Fig. 3.4).

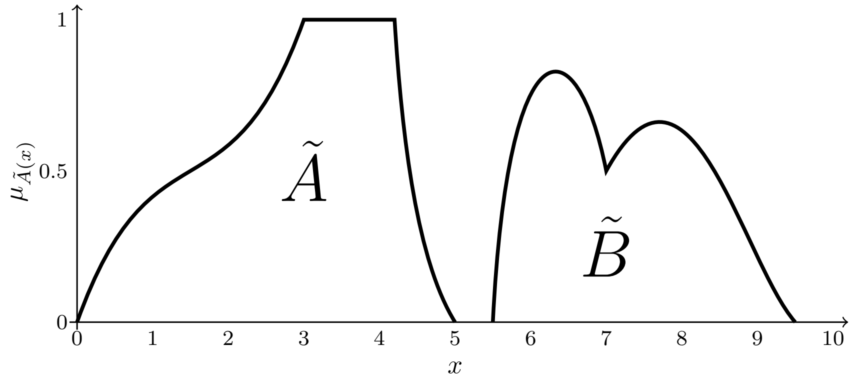


Figure 3.4: Types of fuzzy sets. $\tilde{A}$ - normal convex fuzzy set, $\tilde{B}$ - subnormal non-convex fuzzy set.

3.5 Operations with fuzzy sets

There are many operations with fuzzy sets that form analogies to crisp set operations. These are intersection, union and complement of a fuzzy set (Zadeh, 1965; Klir and Yuan, 1995; Zimmermann, 1991). Each of these operations has several definitions that are described by Dubois and Prade (1980) and also by Klir and Yuan (1995). In combination with the aggregation operators these operations form the basis for fuzzy logic (Zimmermann, 1991). Fuzzy logic is another branch of fuzzy science focused on the usage of fuzzy sets for approximate reasoning. However, such utilization of fuzzy sets is beyond the scope of this research.

Chapter 4

Fuzzy numbers

Basically, fuzzy logic is a precise logic of imprecision and approximate reasoning. More specifically, fuzzy logic may be viewed as an attempt at formalization/mechanization of two remarkable human capabilities. First, the capability to converse, reason and make rational decisions in an environment of imprecision, uncertainty, incompleteness of information, conflicting information, partiality of truth and partiality of possibility – in short, in an environment of imperfect information. And second, the capability to perform a wide variety of physical and mental tasks without any measurements and any computations.

– Lotfi A. Zadeh, (Zadeh, 2008)

Fuzzy numbers are special cases of fuzzy sets that model vague, imprecise or ill-known value (Hanss, 2005; Kaufmann and Gupta, 1985; Klimke, 2006). As noted by Viertl (2011), the measurements of continuous quantities are never precise numbers, they always contain some amount of uncertainty. Usually these measurements are idealized as precise numbers and sometimes statistics is used to describe uncertainty. However, even such statistical models are not correct as they are suitable to describe variability, but there cannot be variability in single measurements' uncertainty (Viertl, 2011). Consequently, there is a need for another formalism to capture this type of uncertainty. Unlike statistics the fuzzy set theory is well suited to describe this type of uncertainty. Semantics of fuzzy numbers is described in detail in chapter 7.3 on the page 60 and the following page.

As mentioned previously, the fuzzy numbers are special cases of fuzzy sets and as such they have to fulfil several requirements. There is only a partial agreement amongst different authors on which requirements really need to be fulfilled. The fuzzy set $\tilde{A}$ is a fuzzy number if it satisfies these conditions:

- $\tilde{A}$ is a normal fuzzy set, $\text{hgt}(\tilde{A}) = 1$.
- $\tilde{A}$ is convex (Eq. (3.12)).
- The membership function $\mu_{\tilde{A}}(x)$ is at least piecewise continuous.

Besides those conditions that can be found in several sources (Hanss, 2005; Klimke, 2006; Viertl, 2011) there are some other properties mentioned by some authors but not respected by others. These are:

- $\tilde{A}$ is defined on the universe of real numbers $\mathbb{R}$ (Klimke, 2006), however this is violated by introducing fuzzy numbers defined on the integers $\mathbb{Z}$ (Kaufmann and Gupta, 1985).
- There is only one x with $\mu_{\tilde{A}}(x) = 1$ (Hanss, 2005; Klimke, 2006), but other authors (Coroianu et al., 2013; Tran and Duckstein, 2002) introduce trapezoidal fuzzy numbers that do not respect this condition.
- $\text{supp}(\tilde{A})$ is a closed and bounded interval (Voxman, 1998; Viertl, 2011), this is not mentioned directly in other literature e.g. (Hanss, 2005; Klimke, 2006) but the existence of compact or finite support of a fuzzy set is mentioned.

For the purposes of this work a fuzzy number is a normal convex fuzzy set with at least a piecewise continuous membership function that is defined on $\mathbb{R}$ and has a closed and bounded support. Even the fuzzy sets that have more than one x that has $\mu_{\tilde{A}}(x) = 1$ are considered fuzzy numbers. The problem with such numbers lies merely in terminology as such fuzzy sets are sometimes referred to as fuzzy intervals (Klimke, 2006). The problem with this naming convention was also discussed by Zimmermann (1991) who clearly stated that even fuzzy intervals can be called fuzzy numbers.

If a fuzzy number has only one x with $\mu_{\tilde{A}}(x) = 1$, then this value is called a peak, a modal, a center or a mean value of a fuzzy number (Hanss, 2005). The last two expressions are preferable for the symmetric fuzzy numbers. If there is more than one x with $\mu_{\tilde{A}}(x) = 1$ than there exists the midpoint of $\tilde{A}_1$ that can be referred to as a modal value of a fuzzy number.

4.1 Properties of fuzzy numbers

A fuzzy number is called strictly positive if $\underline{a} > 0$ applies for its compact support $\tilde{A}_0 = [\underline{a}, \overline{a}]$ (Zimmermann, 1991). A fuzzy number is called strictly negative if $\overline{a} < 0$ is true. In special cases when $\overline{a} \leq 0$ and $\underline{a} \geq 0$ the fuzzy numbers are called negative and positive. If $0 \in \tilde{A}_0$ than a fuzzy number is called a fuzzy zero (Hanss, 2005).

A fuzzy number is called symmetric if its $\mu_{\tilde{A}}$ satisfies (Hanss, 2005):

$$\mu_{\tilde{A}}(\hat{a} - k) = \mu_{\tilde{A}}(\hat{a} + k) \qquad \forall k \in \mathbb{R}. \tag{4.1}$$

In this equation $\hat{a}$ denotes the midpoint of $\tilde{A}_1$ and k is an arbitrary real number. All other fuzzy numbers are asymmetric (Fig. 4.1).

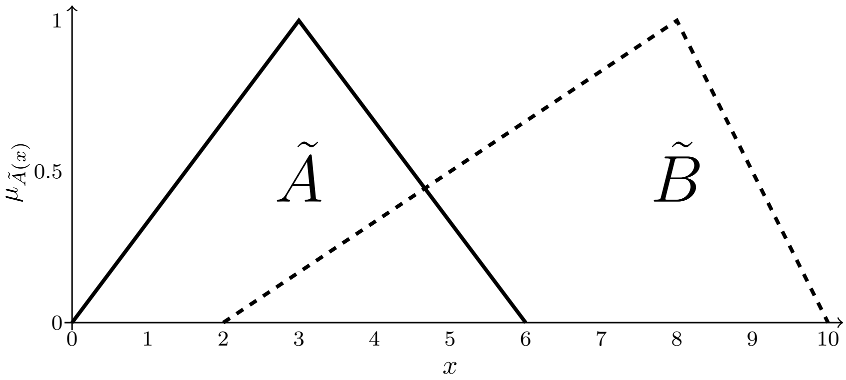


Figure 4.1: Types of fuzzy numbers. $\tilde{A}$ - symmetric, $\tilde{B}$ - asymmetric.

4.2 Types of fuzzy numbers

The set of fuzzy sets that can be qualified as fuzzy numbers is theoretically infinite, but there are several specific types of fuzzy numbers that are of a particular importance. These fuzzy numbers are usually important because their membership functions have specific properties that somehow facilitate their further use. Among those the following are the most important types - $L - R$ fuzzy numbers (Dubois and Prade, 1980), piecewise linear fuzzy numbers (Baekeland and Kerre, 1988), triangular, gaussian, quasi-gaussian, quadratic, exponential fuzzy numbers (Hanss, 2005), trapezoidal fuzzy numbers (Zimmermann, 1991) and also fuzzy singletons (Hanss, 2005). In the following chapters the most important types of fuzzy numbers will be described.

4.2.1 Triangular fuzzy numbers

It can also be referred to as a linear fuzzy number because of its very simple membership function. It is rather frequently used in applications mainly due to its simplicity (Hanss, 2005).

A triangular fuzzy number is either defined by a triplet $\tilde{A} = (a, \alpha_l, \alpha_r)$, where a denotes the peak value and α_l and α_r denote the left and right spread of the fuzzy number. The membership function is then (Hanss, 2005):

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & \text{if } x \leq a \ - \alpha_l \\ 1 + (x-a)/\alpha_l & \text{if } a \ - \alpha_l < x < a \\ 1 - (x-a)/\alpha_r & \text{if } a \ < x < a \ + \alpha_r \\ 0 & \text{if } x \geq a \ + \alpha_r \end{cases}. \tag{4.2}$$

Or the fuzzy number can be directly defined by triplet $\tilde{A} = (min, modal, max)$ where labeling is rather self explanatory. The membership function has the form:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & \text{if } x < min \\ \dfrac{x - min}{modal - min} & \text{if } min \leq x \leq modal \\ \dfrac{max - x}{max - modal} & \text{if } modal \leq x \leq max \\ 0 & \text{if } max < x \end{cases}. \tag{4.3}$$

4.2.2 Trapezoidal fuzzy numbers

As previously mentioned, trapezoidal fuzzy numbers are sometimes called fuzzy intervals to emphasize the difference between them and fuzzy numbers. In practical applications there is a negligible difference as trapezoidal fuzzy numbers pose no problem for calculations or any other integration (Zimmermann, 1991). A trapezoidal fuzzy number is defined by a quadruplet $\tilde{A} = (a, b, c, d)$ and the membership is defined as (Coroianu et al., 2013; Dutta et al., 2011):

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & \text{if } x < a \\ \dfrac{x-a}{b-a} & \text{if } a \ \leq x \leq b \\ 1 & \text{if } b \leq x \leq d \\ \dfrac{d-x}{d-c} & \text{if } c \leq x \leq d \\ 0 & \text{if } d < x \end{cases} . \tag{4.4}$$

As visible from the definition, a trapezoidal fuzzy number has a linear membership function and is very similar to a triangular fuzzy number. In fact a triangular fuzzy number can be viewed as a special case of a trapezoidal fuzzy number where $b = c$.

4.2.3 Fuzzy singletons

In the same way as a classic (crisp) set can be seen as a special case of a fuzzy set, crisp numbers can be considered as a special case of fuzzy numbers as they pose all necessary properties (Hanss, 2005). If a is a crisp number that represents a fuzzy singleton, the membership functions is:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & \text{if } x < a \\ 1 & \text{if } x = a \\ 0 & \text{if } a \ < x \end{cases} . \tag{4.5}$$

In other words a fuzzy singleton is a fuzzy number where $\tilde{A}_1 = \tilde{A}_0$ and both of them are degenerative intervals $[a, a]$.

4.2.4 Piecewise linear fuzzy numbers

A piecewise linear fuzzy number is a special type of fuzzy number defined by $2n + 4$ points (Baekeland and Kerre, 1988). In case when $n = 0$, a piecewise linear fuzzy number is equal to a trapezoidal fuzzy number. The addition of additional α-cuts is a way to specify the fuzzy number more precisely (Coroianu et al., 2013). Obviously, two points are necessary to define an α-cut, thus $2n$ points are needed. The value of n specifies the number of additional α-cuts. $\tilde{A}_0$ and $\tilde{A}_1$ are defined by four points that are absolutely necessary for the definition of a piecewise linear fuzzy number.

Coroianu et al. (2013) propose to define a piecewise linear fuzzy number as $\tilde{A} = (a_\alpha, v)$, where a_α is a set of additional α-cut values (without α-cut 0 and 1) of size n, and v is a set of values, sorted in ascending order, of size $2n + 4$.

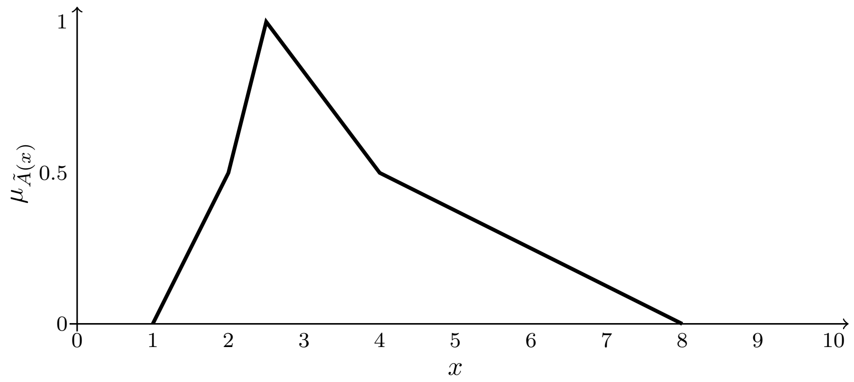


Figure 4.2: A piecewise linear fuzzy number $\tilde{B}$ defined as $\tilde{B} = ((0.5), (1, 2, 2.5, 2.5, 4, 8))$.

Consequently, the definition of a fuzzy number $\tilde{B}$ with $\tilde{B}_{0.5}$ added would look like: $\tilde{B} = ((0.5), (1, 2, 2.5, 2.5, 4, 8))$. The visualization of the fuzzy number $\tilde{B}$ is shown in Fig. 4.2.

The piecewise linear model of a fuzzy number is very useful for practical implementation because it does not depend on a specific membership function (the function between the points is linear) but with a sufficient number of n it can represent any shape of a fuzzy number. Figure 4.3 shows examples of four piecewise linear fuzzy numbers that approximate a gaussian fuzzy number. It is visible that with an increasing number of n the approximation is getting more precise.

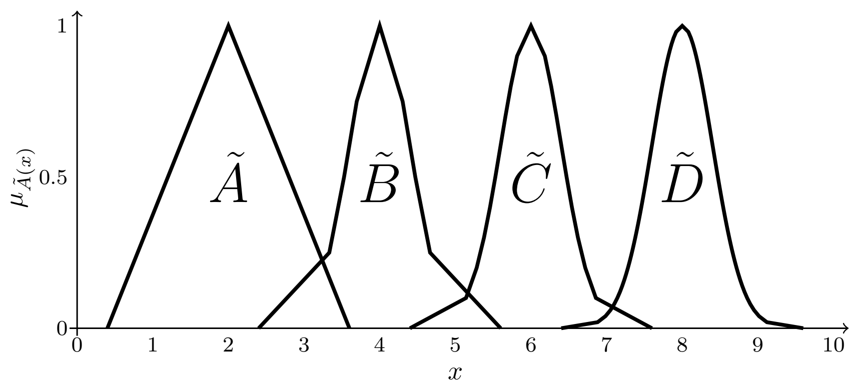


Figure 4.3: Piecewise linear fuzzy numbers approximating a gaussian fuzzy number with a different number of n. $\tilde{A} - n = 0$, $\tilde{B} - n = 3$, $\tilde{C} - n = 9$, $\tilde{D} - n = 49$.

Chapter 5

Fuzzy arithmetic

> *As human beings we must learn to accept that uncertainty is part of our life and will continue to be part of it in the future as well. As scientists we must also accept the fact that uncertainty is an ever-increasing part of our work day.*
>
> – Arnold Kaufmann and Madan M. Gupta (Kaufmann and Gupta, 1985)

Fuzzy arithmetic is an extension of classic arithmetic to operations with fuzzy numbers (Kaufmann and Gupta, 1985). The first practical examples of calculations with fuzzy numbers were presented by Dubois and Prade (1978, 1979) and Nahmias (1978). The original focus of fuzzy arithmetic was mainly the basic operations of addition, subtraction, multiplication and division (Dubois and Prade, 1978), later the term was used as more general – even for the calculation of functions of fuzzy numbers (Kaufmann and Gupta, 1985). Hanss (2005) distinguishes standard fuzzy arithmetic, which contains the basic operations, and advanced fuzzy arithmetic, which is used for more complex operations. For the purpose of this thesis the term fuzzy arithmetic will be used for all mathematical operations with fuzzy numbers.

Several approaches to the topic of fuzzy arithmetic were proposed over the years. For example fuzzy arithmetic based on $L-R$ fuzzy numbers (Dubois and Prade, 1980), fuzzy arithmetic for discretized fuzzy numbers (Hanss, 2005), fuzzy arithmetic based on the decomposed fuzzy numbers and interval arithmetic (Hanss, 2005; Kaufmann and Gupta, 1985; Klimke, 2006). The latest approach is identified as the best for practical implementation (Hanss, 2005; Kaufmann and Gupta, 1985; Klimke, 2006) so the main focus will be given to this approach. This approach based on the piecewise linear fuzzy numbers decomposed into α-cuts and the combination using interval arithmetic is used in all practical applications that were described in literature so far (Anile et al., 1995; Fonte et al., 2008b; Gagolewski, 2014; Spinella, 2008).

5.1 The extension principle

The extension principle as defined by Zadeh (1975) forms theoretical background for almost all operations with fuzzy sets. The principle allows an extension of crisp mathematical operations to their alternatives with fuzzy numbers. Several variants of the definition exist, even though the equation may vary, all the variants describe functionally the same thing. The described definition of the extension principle is based on the definitions provided by Hanss (2005) and Zimmermann (1991).

Let X be a cartesian product of universes $X = X_1, \ldots, X_d$ and $\tilde{A}_1, \ldots, \tilde{A}_d$ fuzzy sets on $X_1, \ldots, X_d$ respectively, with the membership functions $\mu_{\tilde{A}_1}(x_1), \ldots, \mu_{\tilde{A}_d}(x_d)$. The function f is mapping from X to the universe Y, $y = f(x_1, \ldots, x_d)$. Then the fuzzy set $\tilde{B}$ can be defined in Y by:

$$\tilde{B} = \{(y, \mu_{\tilde{B}}(y)) \mid y = f(x_1, \ldots, x_d), (x_1, \ldots, x_d) \in X\} \tag{5.1}$$

where

$$\pi(x_1, \ldots, x_d) = \min\{\mu_{\tilde{A}_1}(x_1), \ldots, \mu_{\tilde{A}_d}(x_d)\}, \tag{5.2}$$

and

$$\mu_{\tilde{B}}(y) = \begin{cases} \sup\limits_{y=f(x1,\ldots,x_d)} \pi(x_1, \ldots, x_d), & \text{if } \exists y = f(x1, \ldots, x_d) \\ 0 & \text{otherwise.} \end{cases} \tag{5.3}$$

This description of the extension principle is very general, for the simple arithmetic operations that involve only two fuzzy numbers $\tilde{A}_1$ and $\tilde{A}_2$ it can be simplified to (Hanss, 2005):

$$\tilde{B} = E(\tilde{A}_1, \tilde{A}_2) \tag{5.4}$$

where E symbolizes one of the elementary operations: $+, -, \times, /$. The resulting fuzzy number is obtained from:

$$\mu_{\tilde{B}}(z) = \sup_{z=E(x_1,x_2)} \min\{\mu_{\tilde{A}_1}(x_1), \mu_{\tilde{A}_2}(x_2)\} \qquad \forall x_1, x_2 \in \mathbb{R}. \tag{5.5}$$

This definition of the extension principle is nicely done from the mathematical point of view, however, as noted by Hanss (2005), Klimke (2006) and others, it is rather impractical for any implementation. The main issues regard solving the global optimization problems to find the upper and lower bounds of results. There are also a few other issues that are discussed in several sources (Hanss, 2005; Kaufmann and Gupta, 1985; Klimke, 2006). Due to those issues it is more practical to base the computational methods on the alternative approach that simplifies the calculation while respecting the extension principle.

5.2 Fuzzy arithmetic based on the decomposed fuzzy numbers

The most practical implementation of fuzzy arithmetic operations described in literature is based on the decomposition theorem (Eq. (3.7)) and interval arithmetic (Hanss, 2005; Kaufmann and Gupta, 1985; Viertl, 2011). This approach takes the advantage of the fact that a fuzzy number can be described by a finite set of α-cuts that can be treated, combined and used in calculations as intervals. Interval arithmetic was firstly introduced by Moore (1966), who referred to it as the interval analysis. Interval arithmetic provides a method for uncertainty representation and propagation. It captures uncertainty as the interval $A = [\underline{a}, \overline{a}]$ that signifies the bounds or the best/worst case assumption about the variable. The biggest drawback of the method is that there is no detailed information about uncertainty except for its bounds (Oberguggenberger, 2005). However, the method itself is relatively simple for implementation, which is one of the reasons why it became a rather popular method for the uncertainty propagation in computer science (Moore et al., 2009). The other reason is that unlike other methods for the uncertainty propagation, interval arithmetic always provides rigorous enclosures of solution (Moore et al., 2009) which is not true for other methods as e.g. the Monte Carlo method.

The relation between the extension principle and the usage of decomposed fuzzy numbers along with interval arithmetic is well described in literature (Hanss, 2005; Klimke, 2006) and especially in the pioneering work of Kaufmann and Gupta (1985). The approach is proven to provide the same results as the direct usage of the extension principle (Eqs. (5.1,5.2,5.3) and interval arithmetic simplifies the calculation significantly in terms of both the implementation as well as the computational complexity.

5.2.1 The procedure

The process of calculations with fuzzy numbers is done in several steps that can be described generally and that (with slight modifications) apply to any mathematical calculation with fuzzy numbers. These steps are (Hanss, 2005):

- decomposition of the input fuzzy numbers,
- application of interval arithmetic (or interval analysis),
- recomposition of the output fuzzy number.

The first step is based on the decomposition theorem (chapter 3.3). All input fuzzy numbers are decomposed into $m+1$ α-cuts. Then the expression is evaluated separately for each membership value μ_j. The evaluation is done according to interval arithmetic (Moore, 1966; Moore et al., 2009). The output fuzzy number is recomposed from the α-cuts that were calculated in the previous step.

5.3 Basic arithmetic operations

The fuzzy number $\tilde{Z}$ calculated as $\tilde{Z} = \tilde{X} \diamond \tilde{Y}$, where $\diamond$ is one of the operations $+, -, \times, /$, is obtained by the decomposition of fuzzy numbers into α-cuts and calculating the output for each α-cut according to (Kaufmann and Gupta, 1985):

$$\tilde{Z} = \tilde{X} \diamond \tilde{Y}, \tag{5.6}$$

$$\tilde{Z}_\alpha = \tilde{X}_\alpha \diamond \tilde{Y}_\alpha, \tag{5.7}$$

$$[\underline{z}, \overline{z}] = [\underline{x}, \overline{x}] \diamond [\underline{y}, \overline{y}] = [\min(G), \max(G)], \tag{5.8}$$

$$G = \{\underline{x} \diamond \underline{y}, \underline{x} \diamond \overline{y}, \overline{x} \diamond \underline{y}, \overline{x} \diamond \overline{y}\}. \tag{5.9}$$

If the operation $\diamond$ is division then it is assumed that $0 \notin [\underline{y}, \overline{y}]$, or in general $0 \notin \tilde{Y}$, otherwise the operation is not valid.

5.4 Functions of a fuzzy number

A function of a fuzzy number may have a form of:

$$\tilde{Y} = f(\tilde{X}), \tag{5.10}$$

or if it is a function of several fuzzy numbers then:

$$\tilde{Y} = f(\tilde{X}_1, \ldots, \tilde{X}_n). \tag{5.11}$$

Obviously, the calculation can be done according to the extension principle, but as mentioned previously, a direct usage of the extension principle is not reasonable. Instead, several optimization techniques can be used to simplify the calculation. Some of these methods are based on the properties of functions, the others are based on the subdivisions of the α-cuts and finding at least an approximation of the result (Moore et al., 2009).

Both approaches to the problem utilize the decomposition theorem and calculate the results by the α-cuts, from these α-cut outcomes the resulting fuzzy number is composed. There exist three main categories of functions: monotonic, nonmonotonic functions and special functions that have specific definitions. The details for functions of fuzzy numbers are provided in Kaufmann and Gupta (1985).

5.4.1 Monotonic functions

Non-decreasing or non-increasing functions of fuzzy numbers are very simple to calculate. Suppose that we have $\tilde{A}$, with $\tilde{A}_\alpha = [\underline{a}, \overline{a}]$, and the monotonic function $f()$. Then the α-cuts of $\tilde{B} = f(\tilde{A})$ can be calculated as (Kaufmann and Gupta, 1985):

$$\tilde{B}_\alpha = [\min(f(\underline{a}), f(\overline{a})), \max(f(\underline{a}), f(\overline{a}))]. \tag{5.12}$$

This step is repeated for every α-cut from the decomposed fuzzy number/numbers. The same applies even for the functions of more than one variable if the function is monotonic with respect to all variables (Moore et al., 2009).

5.4.2 Non-monotonic functions

Functions that are not monotonic pose a special challenge because their range of values over the interval $\tilde{A}_\alpha$ can be difficult to determine. However, solutions to this problem exist. The most complex one is the transformation method proposed by Hanss (2002) and also the utilization of sparse grids proposed by Klimke et al. (2004). Both approaches allow a calculation of non-monotonic functions with more than one fuzzy variable. However, since non-monotonic functions of fuzzy variables do not occur in problems that will be solved later in this thesis, the detailed description of these methods is not necessary. The description can be found in the literature mentioned above.

5.5 Special functions

Several functions of fuzzy numbers are of special interest for the analyses of fuzzy surfaces. Amongst these functions the integer powers of a fuzzy number and $\mathrm{atan2}(\tilde{Y}, \tilde{X})$ need to be explained for further use in this thesis.

5.5.1 Integer powers of a fuzzy number

Calculation of the integer power depends on whether the power is an even or odd number and also on whether the fuzzy number is positive, negative or a fuzzy zero (Kaufmann and Gupta, 1985; Dutta et al., 2011). The calculation itself can be performed α-cut wise. The fuzzy number $\tilde{B}$ with the α-cuts $\tilde{B}_\alpha = [\underline{b}, \overline{b}]$ that is defined as $\tilde{B} = \tilde{A}^n$, can be calculated by these equations.

For a positive fuzzy number $\tilde{A}$ the α-cuts of $\tilde{B}$ are defined regardless of the evenness or oddness of the power as:

$$\tilde{B}_\alpha = [\underline{a}^n, \overline{a}^n]. \tag{5.13}$$

For the fuzzy zeroes ($0 \in \tilde{A}_0$) the α-cuts that contain zero are calculated for the odd powers as:

$$\tilde{B}_\alpha = [\underline{a}^n, \overline{a}^n], \tag{5.14}$$

and for the even powers as:

$$\tilde{B}_\alpha = [0, \max(\underline{a}^n, \overline{a}^n)]. \tag{5.15}$$

The α-cuts that do not contain zero can be calculated according to the positivity or negativity of the α-cut.

For negative fuzzy numbers the calculation depends on whether the power is even or not. If n is even, then:

$$\tilde{B}_\alpha = [\overline{a}^n, \underline{a}^n], \tag{5.16}$$

if n is odd:

$$\tilde{B}_\alpha = [\underline{a}^n, \overline{a}^n]. \tag{5.17}$$

These definitions are in agreement with the equations provided by Moore et al. (2009) for the calculation of integer power of the intervals and with definitions for the calculation of integer power of fuzzy numbers provided by Kaufmann and Gupta (1985).

5.5.2 Function atan2

The function atan2(y, x) is a version of function arctan that takes two arguments and that are transformed from the cartesian coordinates into the polar coordinates (Gaile and Burt, 1980; Mardia and Jupp, 1999). The polar coordinates are expressed as a counter clock-wise angle defined by a point with coordinates $0, 0$, the x axis and coordinates x, y of the point. The range of the function is $[-\pi, \pi]$, which is sometimes mapped to $[0, 2\pi]$ by adding 2π to the negative results. The definition of the functions is:

$$\operatorname{atan2}(y, x) = \begin{cases} \arctan \frac{y}{x} & \text{if } x > 0 \\ \arctan \frac{y}{x} + \pi & \text{if } y \geq 0 \text{ and } x < 0 \\ \arctan \frac{y}{x} - \pi & \text{if } y < 0 \text{ and } x < 0 \\ +\frac{\pi}{2} & \text{if } y > 0 \text{ and } x = 0 \\ -\frac{\pi}{2} & \text{if } y < 0 \text{ and } x = 0 \\ \text{undefined} & \text{if } y = 0 \text{ and } x = 0 \end{cases}. \tag{5.18}$$

From the definition and Fig. 5.1 it is visible that the function is not defined for $y = 0$ and $x = 0$ and it is discontinuous around the line $x < 0$ and $y = 0$. Those two facts pose the main problem for the calculation of atan2 with fuzzy numbers. Despite that, the function can be viewed as monotonic with respect to both variables (Fig. 5.1) so the results can be calculated quite easily, without the need of advance techniques for the calculation of functions of fuzzy numbers.

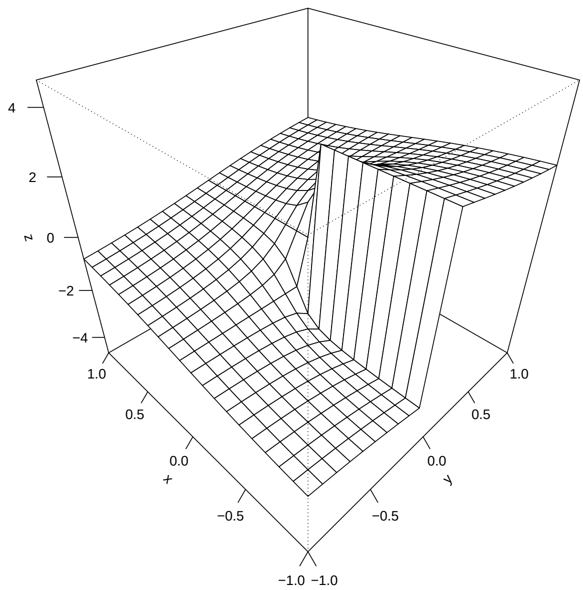
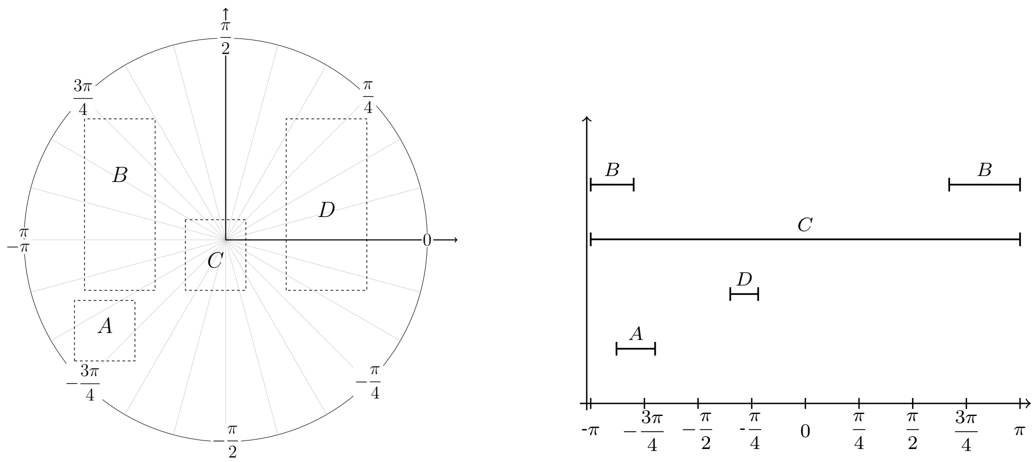


Figure 5.1: Visualization of the values of the atan2(x, y) function.

Fig. 5.2a shows the problem with discontinuity of the atan2 function on four examples that represent the α-cuts of fuzzy arguments of the function. Examples A and D can be calculated easily according to the Eq. (5.12) by the approach shown by Klimke (2006). The resulting intervals are approximately $[-0.88\pi, -0.7\pi]$ for the example A and $[-0.22\pi, 0.35\pi]$ for D. These results pose no problem as they form valid intervals that can be used as the α-cut of the resulting fuzzy number. The same cannot be said for the examples C and B. For C the outcome is clearly the whole range of the function $[-\pi, \pi]$ because both arguments contain 0. For the example B the result is not one interval but actually two of them: $[-\pi, -0.8\pi]$ and $[0.67\pi, \pi]$ (Fig. 5.2b). This is a direct result of discontinuity of the atan2 function when $y = 0$ and $x < 0$.



(a) Examples of four α-cut combinations. (b) Resulting intervals of the four examples.

Figure 5.2: Visualizations of four examples of the atan2 function calculation.

In order to allow the calculation with fuzzy numbers there is a need for a modified version of the function. As noted by Gaile and Burt (1980), there exists a zero direction problem in the directional statistics. The problem that is encountered when atan2 is calculated for fuzzy arguments is quite similar. It can be avoided but the obtained results will require a slightly more work in order to be interpreted correctly.

The modified version of the atan2 function is defined:

$$\text{atan2modified}(y,x) = \begin{cases} \arctan\frac{y}{x} - 2\pi & \text{if } y \leq 0 \text{ and } x > 0 \\ \arctan\frac{y}{x} & \text{if } y < 0 \text{ and } x > 0 \\ \arctan\frac{y}{x} - \pi & \text{if } y \geq 0 \text{ and } x < 0 \\ \arctan\frac{y}{x} - \pi & \text{if } y < 0 \text{ and } x < 0 \\ +\frac{\pi}{2} - \pi & \text{if } y > 0 \text{ and } x = 0 \\ -\frac{\pi}{2} - \pi & \text{if } y < 0 \text{ and } x = 0 \\ \text{undefined} & \text{if } y = 0 \text{ and } x = 0 \end{cases} . \tag{5.19}$$

If the calculation of atan2$(\tilde{Y},\tilde{X})$ is to be performed, firstly it needs to be determined whether the problem with discontinuity of the function will occur. This will happen if there exists an α-cut such that $0 \in \tilde{Y}_\alpha$ and $0 > \overline{x}$ from $\tilde{X}_\alpha$. If this condition is true, then the modified (rotated) version of atan2 needs to be used (Eq. (5.19)). The rotated variant of the function has a modified range $[-1.5\pi, 0.5\pi]$ instead of the original range $[-\pi,\pi]$ and it is discontinuous if $x = 0$ and $y < 0$ (see Fig. 5.3). The problem with the undefined value of both functions is solved by setting the result interval to a full range of values if $0 \in \tilde{X}_\alpha$ and also $0 \in \tilde{Y}_\alpha$. In either case both functions atan2 and atan2modified are continuous with respect to both variables and as such can be propagated by the usage of a simple approach according to Eq. (5.12).

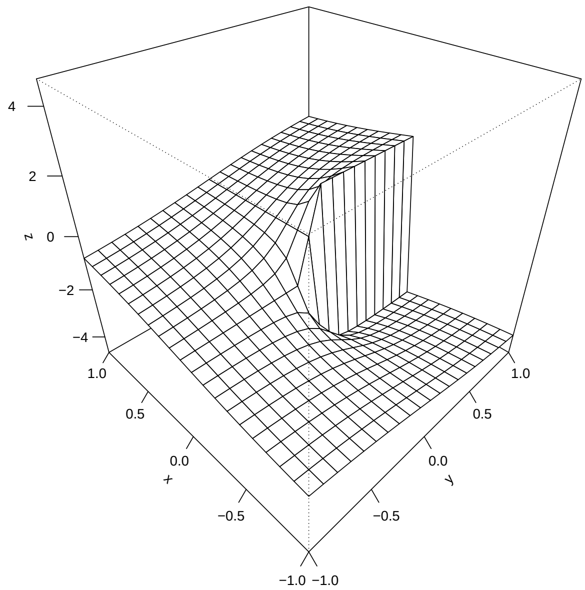


Figure 5.3: Visualization of the values of the modified (rotated) atan2(x,y) function named atan2modified(x,y).

The algorithm for the calculations of $\text{atan2}(\tilde{Y}, \tilde{X})$ with fuzzy parameters is provided in Algorithm 2, for better readability the part of the code that determines if the use of the modified version of the function is needed is earmarked into separate Algorithm 1.

Algorithm 1 Function determining the necessity of using a rotated variant of the atan2 function for two fuzzy numbers

---

Require: Fuzzy numbers $\tilde{Y}, \tilde{X}$ and the integer $m > 1$ that defines the number of α-cuts
Ensure: a boolean value stating whether the rotated variant of atan2 is required for the calculation

procedure ROTATED($\tilde{Y}, \tilde{X}, m$)
    $rotated \leftarrow false$
    for all $i \in \{\frac{0}{m}, \frac{1}{m}, \ldots, \frac{m}{m}\}$ **do**
        if $(0 \in \tilde{Y}_i) \wedge (0 > \overline{\tilde{X}_i})$ **then**
            $rotated \leftarrow true$
        end if
    end for
    return *rotated*
end procedure

Algorithm 2 Function atan2 for Fuzzy Variables

Require: Fuzzy numbers $\tilde{Y}, \tilde{X}$ and the integer $m > 1$ that defines the number of α-cuts
Ensure: Fuzzy number $\tilde{Z}$ containing m α-cuts
procedure ATAN2($\tilde{Y}, \tilde{X}, m$)
▷ the following function is described in Algorithm 1
$rotated \leftarrow$ ROTATED($\tilde{Y}, \tilde{X}, m$)
for all $i \in \{\frac{0}{m}, \frac{1}{m}, \ldots, \frac{m}{m}\}$ **do**
if $rotated$ **then**
if $(0 \in \tilde{Y}_i) \wedge (0 \in \tilde{X}_i)$ **then**
▷ the range of the modified function atan2modified
$Z \leftarrow [-1.5\pi, 0.5\pi]$
else
▷ the following calculations according to Eq. (5.19)
$v1 \leftarrow$ atan2modified($\underline{Y_i}, \underline{X_i}$)
$v2 \leftarrow$ atan2modified($\underline{Y_i}, \overline{X_i}$)
$v3 \leftarrow$ atan2modified($\overline{Y_i}, \underline{X_i}$)
$v4 \leftarrow$ atan2modified($\overline{Y_i}, \overline{X_i}$)

$min \leftarrow \min(v1, v2, v3, v4)$
$max \leftarrow \max(v1, v2, v3, v4)$
$Z \leftarrow [min, max]$
end if
else
if $(0 \in \tilde{Y}_i) \wedge (0 \in \tilde{X}_i)$ **then**
▷ the range of the function atan2
$Z \leftarrow [-\pi, \pi]$
else
▷ the following calculations according to Eq. (5.18)
$v1 \leftarrow$ atan2($\underline{Y_i}, \underline{X_i}$)
$v2 \leftarrow$ atan2($\underline{Y_i}, \overline{X_i}$)
$v3 \leftarrow$ atan2($\overline{Y_i}, \underline{X_i}$)
$v4 \leftarrow$ atan2($\overline{Y_i}, \overline{X_i}$)

$min \leftarrow \min(v1, v2, v3, v4)$
$max \leftarrow \max(v1, v2, v3, v4)$
$Z \leftarrow [min, max]$
end if
end if
$\tilde{Z}_i \leftarrow Z$
end for
return $\tilde{Z}_i$
end procedure

5.6 Limitations and disadvantages of fuzzy arithmetic

The biggest problem of fuzzy and also interval arithmetic is called the overestimation effect (Hanss, 2005). It is also referred to as the dependency problem (Moore et al., 2009). This effect occurs if one fuzzy variable occurs multiple times in the expression that is to be evaluated. This leads to situations where $\tilde{X}^2 \neq \tilde{X} \times \tilde{X}$ for some fuzzy numbers because in later case the fuzzy numbers are considered independent variables while actually they are not. The same case is shown by Hanss (2002) where the example is presented proving that $\frac{\tilde{P}_1+\tilde{P}_2}{\tilde{P}_1} \neq \tilde{P}_2$. This is a consequence of the fact that the distributive law does not hold for either intervals or fuzzy numbers (Moore et al., 2009; Hanss, 2005). For the same reason $\tilde{X} - \tilde{X} \neq 0$ except for a few limit cases.

These issues of fuzzy arithmetic have been addressed by several authors (Hanss, 2005; Kaufmann and Gupta, 1985) and some solutions were provided (Hanss, 2002; Klimke et al., 2004). The main issue is that these methods significantly complicate the calculation from the implementation perspective. However, the transformation method presented by Hanss (2002) can overcome all the issues mentioned above.

Fortunately, none of the issues mentioned in this chapter are relevant to the analyses of a fuzzy surface that will be presented later in the thesis. Nevertheless, it is important to mention these issues and disadvantages because some of them affect specific surface analyses. The analyses present in this thesis are not the case, none of these issues is relevant to those analyses. But the user should still be aware of these drawbacks and limitations in order to avoid potential problems.

Chapter 6

Further operations with fuzzy numbers

There are many operations that can be performed with fuzzy numbers, but for the purpose of this thesis identification of a minimum, a maximum and ranking of fuzzy numbers are of interest.

6.1 Minimum and maximum of fuzzy numbers

The minimum or the maximum of two crisp numbers are relatively easy to obtain, however it is not true for two fuzzy numbers. The outcome of these operations is again a fuzzy number that needs to capture the information contained in both fuzzy numbers that enter the operation. The minimum and maximum of fuzzy numbers $\tilde{A}$ (with α-cut $\tilde{A}_\alpha = [\underline{a}, \overline{a}]$) and $\tilde{B}$ ($\tilde{B}_\alpha = [\underline{b}, \overline{b}]$) are defined by Kaufmann and Gupta (1985) for each α-cut as:

$$\min(\tilde{A}_\alpha, \tilde{B}_\alpha) = [\min(\underline{a}, \underline{b}), \min(\overline{a}, \overline{b})] \tag{6.1}$$

and

$$\max(\tilde{A}_\alpha, \tilde{B}_\alpha) = [\max(\underline{a}, \underline{b}), \max(\overline{a}, \overline{b})]. \tag{6.2}$$

Visualization of the minimum and maximum of fuzzy numbers is shown in Fig. 6.1.

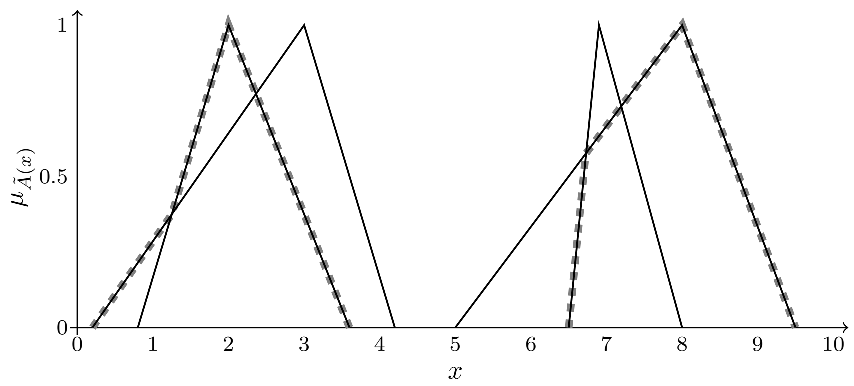


Figure 6.1: Left side – minimum of the fuzzy numbers (dashed line).
Right side – maximum of fuzzy numbers (dashed line).

6.2 Ranking fuzzy numbers

The issue of ranking fuzzy numbers is rather a complex one. There exist many approaches to the problem, however most of them consider only one specific point of view and the others are counterintuitive in certain situations (Dubois and Prade, 1983). Several approaches are based on the extension of the classic ranking operations by the usage of the extension principle (Zadeh, 1971). However, most of these methods fail to recognize the problem of indistinguishability of fuzzy numbers. The main problem is that most of the approaches do not form a total-ordering structure (Dubois and Prade, 1986). The review of these methods is provided by Bortolan and Degani (1985).

The problem of indistinguishability is becoming obvious when two fuzzy numbers that should be ranked overlap significantly. The decision if one is bigger than the other or vice versa is then complicated and it cannot be simply done by one index (Fig. 6.2). Based on this fact, Dubois and Prade (1983) proposed ranking of fuzzy numbers in the setting of the possibility theory (Dubois and Prade, 1986) that uses two indices to evaluate the comparison operators. The comparison is pairwise for two fuzzy numbers, however a more complex variant that allows ranking of n fuzzy numbers was proposed as well (Dubois and Prade, 1986).

Ranking is done in the framework of the possibility theory that uses two measures – possibility and necessity. The measure of possibility is an optimistic one, it evaluates if there is at least some chance of a predicament being true. On the other hand, necessity evaluates to what extent there exist strong indicators that a predicament is true, or, in other words, whether the predicament is necessarily true. As such, necessity is a pessimistic measurement that evaluates how much of the information is actually against the predicament. Both measurements take the values on the interval

$[0, 1]$, where 1 means a complete fulfilment, 0 means no fulfilment and the values between mean a partial fulfilment.

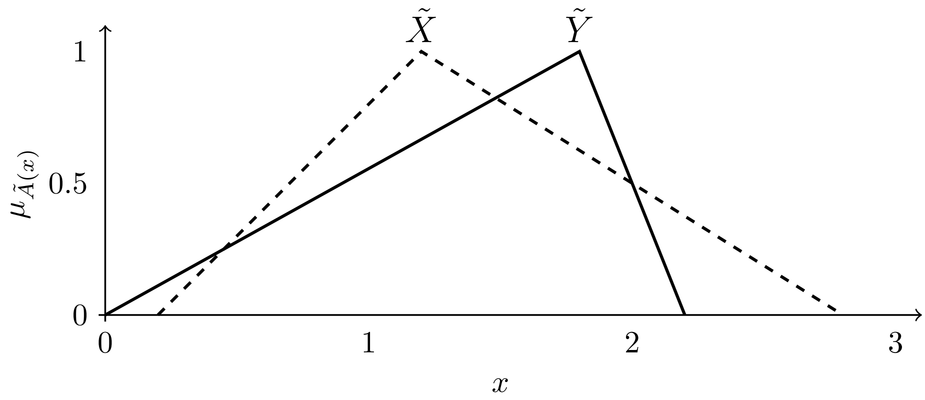


Figure 6.2: Comparison of $\tilde{X}$ and $\tilde{Y}$.

To evaluate to what extent $\tilde{X}$ is greater than $\tilde{Y}$ four indices are needed (Dubois and Prade, 1983). These are $\Pi_{\tilde{X}}([\tilde{Y}, \infty))$ and $\mathcal{N}_{\tilde{X}}([\tilde{Y}, \infty))$ to assess possibility and necessity that $\tilde{X}$ is greater or at least equal to $\tilde{Y}$. And $\Pi_{\tilde{X}}(]\tilde{Y}, \infty))$, $\mathcal{N}_{\tilde{X}}(]\tilde{Y}, \infty))$ for evaluating a strict exceedance of $\tilde{Y}$ by $\tilde{X}$. The expressions for calculating those indices are the following (Dubois and Prade, 1983):

$$\Pi_{\tilde{X}}([\tilde{Y}, \infty)) = \sup_x \min(\mu_{\tilde{X}}(x), \sup_{y \leq x} \mu_{\tilde{Y}}(y)), \tag{6.3}$$

$$\mathcal{N}_{\tilde{X}}([\tilde{Y}, \infty)) = \inf_x \max(1 - \mu_{\tilde{X}}(x), \sup_{y \leq x} \mu_{\tilde{Y}}(y)), \tag{6.4}$$

$$\Pi_{\tilde{X}}(]\tilde{Y}, \infty)) = \sup_x \min(\mu_{\tilde{X}}(x), \inf_{y \geq x} 1 - \mu_{\tilde{Y}}(y) j), \tag{6.5}$$

$$\mathcal{N}_{\tilde{X}}(]\tilde{Y}, \infty)) = \inf_x \max(1 - \mu_{\tilde{X}}(x), \inf_{y \geq x} 1 - \mu_{\tilde{Y}}(y)). \tag{6.6}$$

These results answer the questions “Is $\tilde{X}$ greater than $\tilde{Y}$?” and “Is $\tilde{X}$ strictly greater than $\tilde{Y}$?” in terms of both possibility and necessity (Fig. 6.3). The details on the implementation, proofs and the process of answering the inverse problem are provided by Dubois and Prade (1983) and Dubois and Prade (1986).

In Fig. 6.3 it is visible that the possibility and necessity values are calculated as the intersections between specific sections of membership functions and/or its inverse function. For example the point $\Pi\, \tilde{X}_1 \geq \tilde{Y}$ in the figure is an intersection between the part of $\mu_{\tilde{X}_1}$ that comes after the peak of $\tilde{X}_1$, and the part of $\mu_{\tilde{Y}}$ that is before the peak of $\tilde{Y}$. As such, the possibility values are determined by the definitions of these two membership functions. The same applies for the three other indices. Together these indices form a complete set of comparison indices, meaning that they characterize all the respective configurations of two fuzzy numbers (Dubois and Prade, 1983).

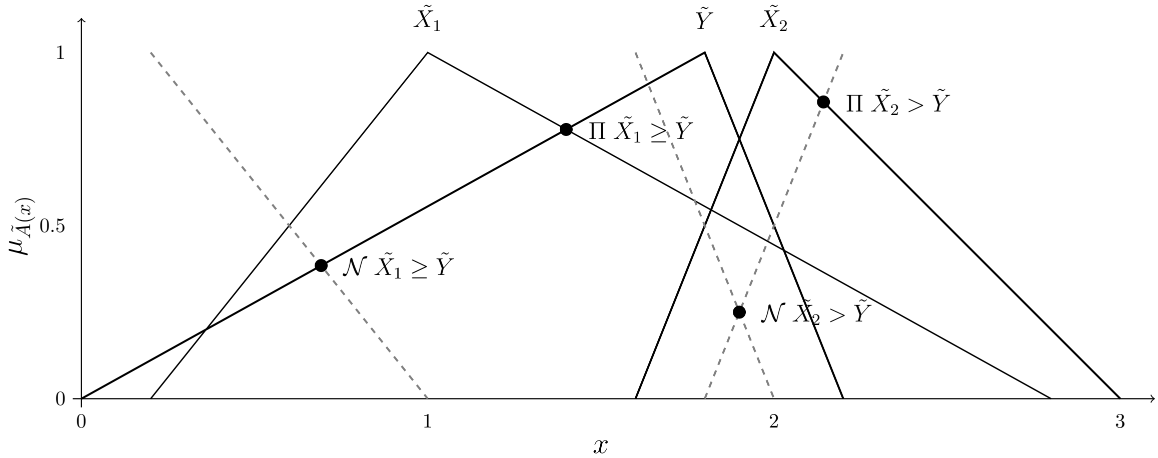


Figure 6.3: Possibility and necessity of $\tilde{X}_1 \geq \tilde{Y}$ and $\tilde{X}_2 > \tilde{Y}$.

Fuzzy numbers can be also compared to the crisp values (Dubois and Prade, 1983). This can either be done by perceiving a crisp number as a special case of a fuzzy number, or more commonly by defining a special approach for such comparison (Dubois and Prade, 1983). Such comparison is described also by Fisher and Caha (2014).

Chapter 7

Surface Uncertainty

Any variable that can be considered a continuous field through the geographical space is in GIS usually represented as a surface, by a field-based model (Longley et al., 2005). A field-based model approximates a continuous variable by a regular raster that most of the times consists of rectangular cells (pixels). In GIS terminology it is often referred to as a grid and it is described by a number of rows M, a number of cells N in each row and a cell size S. A field-based model is widely used in all types of spatial analyses (Fisher, 1997; Longley et al., 2005). There are, however, several issues related to this model that the user needs to be aware of. Firstly, there are two interpretations of the meaning of a cell of the grid – a point (a centre of the grid) or an area (an interior of the cell) (Fisher, 1997). Secondly, several different ways can be used to determine the value recorded for a single cell. The value stored in a cell may represent: *a*) the majority of area of the cell, *b*) a value at some systematic location with the cell (e.g. central point), *c*) a representative value from the cell, based on a specific rule (minimum, maximum, modal value etc.) (Fisher, 1997). For models of surface the value of a cell usually represents the value at the centre because surface models are created by the process of interpolation that predicts a surface value at a point (Wilson and Gallant, 2000). Another issue is the selection of the correct size of a cell, that is discussed by Hengl (2006).

Fisher (1997) summarizes that the issues mentioned previously cannot be solved easily and that the raster (or field-based) model of the data is irreplaceable within the scope geosciences. Thus, it is necessary to inform the users about the potential issues that the model may have. As a consequence of these issues, the information stored within the raster model should be considered uncertain (Fisher, 1997) because the values that are stored may not be as precise representation of the reality as the user would expect.

The field-based model is the most widely used data model for surface data and also the majority of research on error and uncertainty of surfaces was performed on this model (Fisher and Tate, 2006). Such representation may be used to represent the variables either from physical geography (elevation, atmospheric pressure or amount of precipitation) or human geography (population density, average wage or value of property). While the physical surfaces are quite understandable it might not be the same for the surfaces from the field of human geography. These are more abstract, however even those fit the definition that their value denoted as z is a function of the space $z = f(x, y)$, where z is a value of the observed variable and x and y are the coordinates of the point with this value.

The definition of a surface in the geographical sense says that a surface is a statistical representation of a continuous variable by a large number of selected points with known x, y and z coordinates in an arbitrary coordinate field (Miller and Laflamme, 1958). As the definition suggests, only the selected points are used for surface creation. The main reasons for this are time and costs demands of data acquirement (Longley et al., 2005). There are numerous methods of interpolation that can be used to construct a surface from the sampling points, each with its own set of advantages and disadvantages.

The surface models are often used as error–free and completely certain even though they definitely are not. The main issue of the surface models is always their quality and precision (Zhang and Goodchild, 2002). It is affected by quality and suitability of the sample points and by the interpolation method that is used for the creation of surface. The main issues that introduce uncertainty into the model of surface are: *a*) uncertainty of the input points (the values of x, y and z), *b*) density of the sample points through the area of interest, *c*) knowledge of additional processes affecting the surface, *d*) epistemic uncertainty that the user can introduce into the interpolation process. These problems are discussed by El-Sheimy et al. (2005); Hengl (2009); Loquin and Dubois (2010b); Wilson and Gallant (2000).

The main reasons why uncertainty of surfaces is of interest are the facts that uncertainty of surfaces propagates into products derived from these surfaces. The derived products are later used in many practical applications (Fisher and Tate, 2006). Due to these facts, uncertainty of geographical surfaces and their analyses are studied quite commonly (Carlisle, 2005; Fisher, 1992; Fisher and Tate, 2006; Hebeler and Purves, 2009; Heuvelink, 1998; Loquin and Dubois, 2010b; Oksanen and Sarjakoski, 2005a, 2006). Basically, all the mentioned studies utilize the statistical approach to uncertainty of a surface and thus use the Monte Carlo method for the uncertainty propagation. The reasons why this method is often used are mainly that it is rather easily implemented for any operation (Hanss, 2005) and that there is rather long tradition in processing the surface uncertainty with the usage of statistic methods (Heuvelink, 2002). However, for further development of the topic of uncertainty propagation it might be necessary to introduce new mathematical models of uncertainty into the topic of surface analyses.

7.1 Fuzzy surface models

The process of modelling surface from a finite set of samples is a common problem in geosciences. As mentioned previously, surfaces are often treated as certain and error-free models (Zhang and Goodchild, 2002), even though there is a wide set of reasons why they are not. Perhaps the biggest issue arises from incomplete knowledge of the surface under study (Santos et al., 2002). A user cannot be sure that the sample of surface values contains samples that are representative enough to construct a precise surface. There is also the issue of a measurement precision of the individual sample point, some authors are pointing out that every measurement is fuzzy, at least to some extent because there exist no absolutely precise measurements (Viertl, 2011; Lodwick and Santos, 2003; Waelder, 2007). Another uncertainty can be introduced to a surface by the selection of interpolation technique (Santos et al., 2002). Not only there is a range of methods that can be used for interpolation (IDW, spline interpolators, kriging etc.) but some of these methods have specific parameters (e.q. tension of spline, variogram in kriging) that may contain epistemic uncertainty. This means that the values of these parameters are selected by the user and their selection is partially arbitrary (Loquin and Dubois, 2010b). In fact these parameters are better described as a set of possible values than a single value which may not be correct. Lodwick et al. (2008) argue that much (if not the most) of uncertainty of surfaces in geosciences is interval, fuzzy or possibilistic in its nature. Fisher (2005) mentions that the fuzzy set theory should be used in case that the definition of class or individual object is vague. The individual object, in the case of a surface (represented by a grid) cell, respectively its value is definitely vague because it can be based on uncertain data or influenced by epistemic uncertainty in the interpolation method (Loquin and Dubois, 2010b) and even the grid model itself is simplification and idealization of a real surface.

Based on these facts, the model of surface that would account for its inherent uncertainty (Lodwick and Santos, 2003) is needed. Such model was firstly proposed by Diamond (1989) and Bardossy et al. (1990a). A fuzzy surface as described by Diamond (1989) was a result of interpolation with imprecise (fuzzy) data, while the model of Bardossy et al. (1990a) was based on precise data but an imprecise variogram in the kriging interpolation process. These two studies were the first to introduce fuzzy numbers into the spatial data modelling and spatial prediction but the applications of fuzzy approaches for predictions and modelling were used in mathematics before (Tanaka et al., 1982; Heshmaty and Kandel, 1985). Later, more techniques and approaches for the construction of fuzzy surfaces emerged, including bayesian fuzzy kriging (Bandemer and Gebhardt, 2000), improved kriging with imprecise variograms (Loquin and Dubois, 2010b, 2011), inverse-distance weighting method (Waelder, 2007) and also spline interpolators (Anile and Spinella, 2004; Lodwick and Santos, 2003; Santos et al., 2002). All these methods can be used to create valid fuzzy surfaces.

The definition of a fuzzy surface is only slightly different from the definition of an ordinary surface. A fuzzy surface is described by a set of points with known x, y coordinates and a fuzzy number $\tilde{Z}$ that represents the possible values of z at this location. Since the methods for creation of a fuzzy surface are either based on interpolation with imprecise input data, imprecise parameters of interpolation and rarely other techniques the outcome naturally contains a specific type of uncertainty. For every location there are three values of $\tilde{Z}$ that are of special interest to the user. These are z^- denoting the lowest value ($\underline{z}$ from the interval of the α-cut $\tilde{Z}_0$), z^+ denoting the highest value ($\overline{z}$ from $\tilde{Z}_0$), and z^m standing for the modal value of $\tilde{Z}$ (either the peak value or the midpoint in case that there are more than one z with $\mu_{\tilde{Z}}(z) = 1$). Obviously, since the fuzzy surface is defined by fuzzy numbers, the α-cut $\tilde{Z}_\alpha = [\underline{z}, \overline{z}]$ can be extracted for any location. In the same way, any α-cut can be extracted through the whole surface, forming an interval surface (Anile and Spinella, 2004).

The data model for the storage of fuzzy surfaces does not exist within any geographical information software. A fuzzy surface is usually stored as three separate surfaces z^-, z^m and z^+ and either only one of them is visualized as a classic surface (Waelder, 2007), or all three surfaces are visualized in a 3D view (Lodwick and Santos, 2003; Santos et al., 2002; Loquin and Dubois, 2010b). Neither of these visualizations is good because visualization of only one of the surfaces shows only a part of the information that the user should obtain and showing all three of them makes it almost impossible to correctly interpret the data. The issue of visualization of fuzzy data has been discussed since the first introduction of fuzzy set theory into geosciences (Leung et al., 1993; Goodchild et al., 1994) but as the recent examples show, the issue is still unresolved (Caha et al., 2012; Loquin and Dubois, 2010a; Waelder, 2007). Only Anile and Spinella (2004) specify that a fuzzy surface should be stored in what is called a raster FDEM, which is described as a rectangular array of fuzzy numbers of size $N \times M$. However, so far almost no research has been done on how a fuzzy surface should be visualized in order to provide the user with complete information about the possible values and their uncertainty (Vondráková and Caha, 2014). The visualization able to provide a lot of information is a profile of a fuzzy surface (Fig. 7.1) but it still provides only a small portion of the actual information contained within the whole surface. Santos et al. (2002) proposed the approach for visualization that renders not a continuous surface but only the fuzzy numbers in 3D (Fig. 7.2). However, such visualization works only for a rather small surface. Fuzzy surfaces containing more points are not comprehensible with the usage of this visualization.

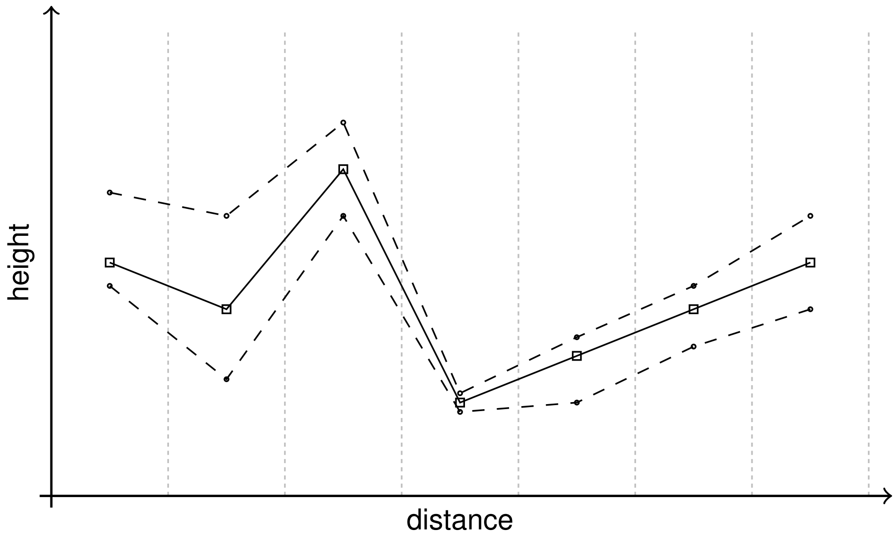


Figure 7.1: The profile of a fuzzy surface. The full line represents z^m, dashed lines represent z^- and z^+. Vertical gray lines show the division of surface into the cells of the grid.

The propagation of uncertainty into the derivatives of fuzzy surfaces is done by the means of fuzzy arithmetic (Lodwick et al., 2008) and some other related methods (e.g. comparison of fuzzy numbers). As described in chapters 4 and 5, the calculations with fuzzy numbers are more complicated than with classic numbers. The calculation cannot be done directly, it has to be modified to allow processing of fuzzy numbers which may not be straightforward (Fig. 7.3). This could be a reason why fuzzy arithmetic has received little attention when compared to the other branches of the fuzzy set theory. Besides that, the potential of fuzzy arithmetic for the solution of real world problems is often underestimated (Hanss, 2005). The topic of fuzzy mathematics is used very rarely for the analyses of fuzzy surfaces (Fisher and Tate,

2006) and the only book that introduces the topic of fuzzy arithmetic to geosciences is quite new (Mount et al., 2009).

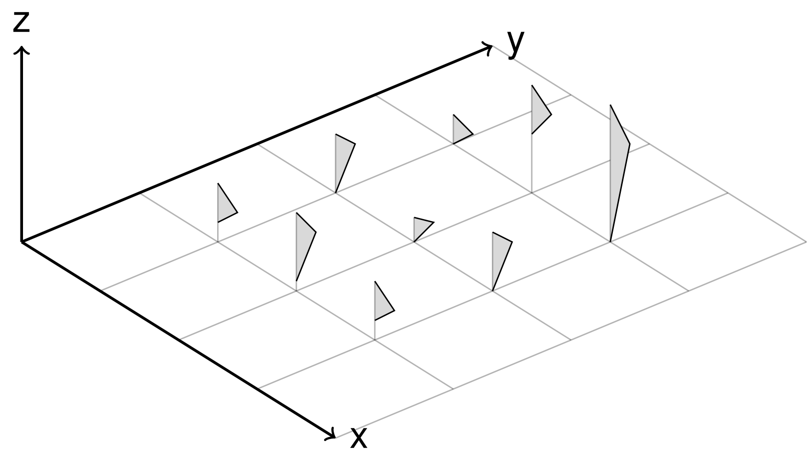


Figure 7.2: Visualization of a small fuzzy surface (3×3 cells). (according to: Santos et al. (2002))

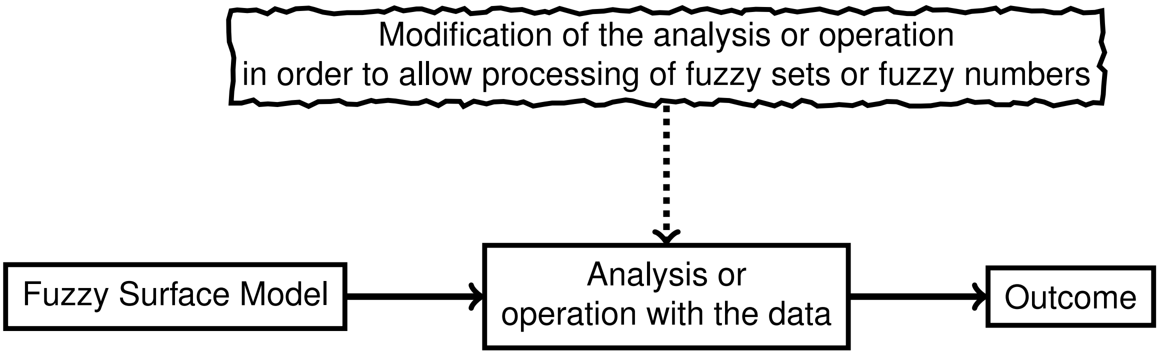


Figure 7.3: Uncertainty propagation based on fuzzy arithmetic requires the modification and adjustment of the analysis or operation.

7.2 Statistical models of a surface error

Statistical description of a surface error has a very long tradition not only in geosciences (Heuvelink et al., 1989). The basic assumption behind the statistical description of surface uncertainty, often referred to as an error, is that there is a difference between the real surface and the model of a surface. This error can have several types: *a*) an error with bias, *b*) a systematic error, *c*) a spatially autocorrelated error, *d*) a random error and of course various combinations of these types (Fisher and Tate, 2006).

The statistical treating of a surface error and its influence on products derived from these surfaces is done in two steps. Firstly, the model of the surface and the model of uncertainty are built. Then a random realization of the surface error is drawn from the uncertainty model (Oksanen and Sarjakoski, 2006). The random realization of uncertainty is added to the surface, creating a random realization of the surface (this surface contains uncertainty). On this surface with uncertainty an analysis or data operation is performed. The whole process is repeated n times and the results are statistically evaluated (Fig. 7.4).

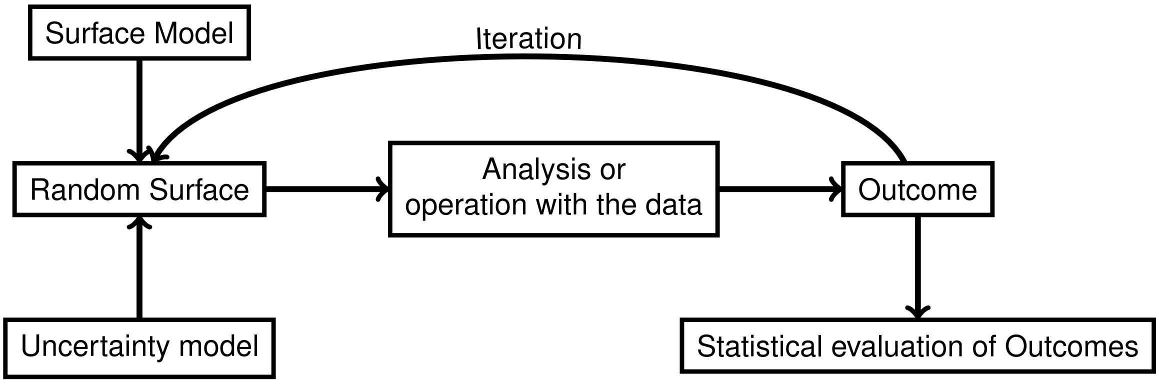


Figure 7.4: Uncertainty propagation based on the statistic description of uncertainty performed by Monte Carlo.

The basic description of the surface error are its mean and root mean square error (RMSE) (Fisher, 1998a; Hunter and Goodchild, 1997). These characteristics are obtained by the comparison of surface with measurements that have higher precision, sometimes the LIDAR data or geodetic measurements. RMSE is very often the only description of the surface accuracy (with mean of the error assumed to be zero) but by no means is it a good description of uncertainty of the surface (Fisher, 1998a; Hunter and Goodchild, 1997). Based on these two characteristics the only model of uncertainty that can be established is spatially uncorrelated and normally distributed error of the surface with the mean being 0 and the standard deviation equal to RSME (Fisher, 1998a; Heuvelink, 2002). Such model of the surface error was considered the worst case scenario (Heuvelink, 1998; Oksanen, 2006) but the derivatives of the surface did not have the highest variability for the uncorrelated error in a study presented by Oksanen and Sarjakoski (2005a).

Because the spatially uncorrelated model of surface uncertainty was considered the worst case scenario, several authors introduced spatial autocorrelation into their models of uncertainty (Fisher, 1998a; Hunter and Goodchild, 1997; Heuvelink, 1998; Oksanen and Sarjakoski, 2006). There are various methods how a spatially autocorrelated error of a surface can be created. Oksanen and Sarjakoski (2005a) mention that geostatistics recognizes several methods: *a*) simulated annealing (pixel swapping), *b*) a spatially autoregressive model, *c*) spatial moving averages, *d*) a sequential gaussian simulation. The problem with the creation of spatially autocorrelated errors for the surface is that the parameters of autocorrelation are usually not known and their selection depends on the user (Fisher, 1998a; Heuvelink, 1998). Oksanen and Sarjakoski (2005a) created 32 models (two variogram models with four values of two variables (sill and range) – $2 \times 4 \times 4 = 32$) of uncertainty of the surface. It is obvious that the authors wanted to model a wide range of possible uncertainties that could be associated with the model of the surface. The selection of the uncertainty model is always problematic because datasets that would allow the estimation of uncertainty, like the dataset used by Holmes et al. (2000), are rarely available. Sometimes these datasets are available only for a part of the study area and the information about uncertainty is then used even on the areas where it was not actually studied (Hebeler, 2008). However, this assumption that such information can be generalized for the whole area of study may not be correct (Fisher, 1998a).

In the study performed by Oksanen and Sarjakoski (2005a), the authors did 1 000 iterations for each model of uncertainty, meaning altogether 32 000 calculations. The authors wanted to avoid the common problem of the Monte Carlo method which is the convergence of results (Oksanen, 2006). If the number of simulations is low (tens of iterations) e.g. in the studies by Holmes et al. (2000) or Fisher (1991), the risk of obtaining unreliable results is very high (Heuvelink, 1998). Oksanen (2006) noted that the probable drainage basin appear to converge after 500 iterations. In a similar experiment Nackaerts et al. (1999) concluded that the probabilistic visibility converges after roughly 30 - 60 iterations, even though some fluctuations can be observed after more than 90 iterations. These examples show that the usage of Monte Carlo can be significantly time and calculation demanding. Even after the calculations, the evaluation of the results might not be an easy task for the user (Heuvelink, 2002).

7.3 Comparison of methods for the uncertainty propagation

Until now four methods of the uncertainty propagation that can be used in the surface analyses have been mentioned: *a*) the analytical approach, *b*) the probabilistic simulation (the Monte Carlo method), *c*) interval arithmetic, *d*) fuzzy arithmetic. Interval arithmetic can be considered as a special case of fuzzy arithmetic that does not provide any measure of preference amongst the results (Hanss, 2005; Oberguggenberger, 2005). The preference is rather important for the decision making based on the outcomes of uncertainty propagation. As a consequence of this, interval arithmetic will not be further considered in this thesis. The analytical approach will be also omitted because many applications are too complex and thus making the analytical approach impractical (Zhang and Goodchild, 2002). Two possible methods for the uncertainty propagation are remaining – the Monte Carlo method and fuzzy arithmetic. Here the differences will be summarized and some important remarks will be mentioned. The methods will be compared in terms of semantics, numerics (Oberguggenberger, 2005) and also practical aspects of calculation (Hanss, 2005) with the emphasis on surface analysis.

Oberguggenberger (2005) mentions that the interpretation of probability is a subject of scientific disputes for a long time. There exist three most important semantics of probability:

- classical probability (the fraction of favorable cases among the possible cases),
- frequentist probability (random occurrence of an event in the sequence of independent trials),
- subjective probability (measure of personal confidence).

In a surface analysis that utilizes the Monte Carlo approach the semantics of uncertainty is frequentist, e.g. the drainage basin delineation presented by Oksanen and Sarjakoski (2005b) or the probable visible areas shown by Fisher (1991, 1994). Certain aspects of uncertainty can be perceived as subjective probability because, as mentioned by Oksanen and Sarjakoski (2005a), some parameters of the model of uncertainty were selected as realistic estimates of the true values. The problem arising here is semantics of the result. What does the probability of visibility equal to a specific value e.g. 55 % mean? In reality there is only one surface so it cannot be said that in 55 cases out of 100 tries the point will be visible from the viewpoint. As a consequence, semantics of the results is at least problematic. The problem is less obvious for the numerical values like slope because for such variables the probability distribution is obtained. However, the semantics issue remains even in such results.

Two other important issues of the probabilistic treatment of uncertainty are mentioned by Oberguggenberger (2005). From the philosophical point of view it is unclear if probability of the studied object/process or the probability of the experiment which was designed to measure it is studied. Another issue is connected with the *law of decreasing credibility*. This law states that the credibility of inference decreases with the strength of assumptions maintained (Manski, 2003). As mentioned above, the models of surface uncertainty often rely on the assumption about rather specific spatial autocorrelation, distribution of errors and other assumptions which, according to this law, decrease the credibility of the experiment.

Semantics of fuzzy numbers is connected to the notion of possibility (Oberguggenberger, 2005). The membership value 1 denotes a completely possible value n of an uncertain number $\tilde{N}$. The smaller is the membership value, the smaller is the possibility that n can belong to $\tilde{N}$ or that $\tilde{N}$ can be described by n. Another possible description of the membership value is the degree of potential surprise (Oberguggenberger, 2005). The membership value 1 means the realization of $\tilde{N}$ that is not surprising at all. As the membership value decreases, the surprise is increasing. When the membership value is 0, n is considered an impossible realization of $\tilde{N}$. Hanss (2005) explains that the membership value 1 represents the deterministic value (no uncertainty) while the 0-cut represents the worst case scenario of deviations (maximal uncertainty) from this deterministic value. Kaufmann and Gupta (1985) describe semantics using the terms of intervals of confidence and the level of presumption. It could be argued that the membership values encode the risk assessment for the uncertain values (Oberguggenberger, 2005). There is a wide agreement that fuzzy numbers should be used to model the uncertain values that do not originate in variability (Kaufmann and Gupta, 1985) or if uncertainty of these values is a result of simplification and idealization of the model (Hanss, 2005). Surfaces as used in geosciences are always simplified and also idealized (El-Sheimy et al., 2005), which points towards the usage of fuzzy set theory to model their uncertainty.

There is an interesting connection between fuzzy numbers and statistical models of surface uncertainty. Fisher (1998a) points out that without a detailed study of uncertainty which is often impossible, the selection of parameters for the model of uncertainty is purely subjective. Oksanen and Sarjakoski (2005a) avoided the problem by creating 32 possible models. The authors selected four possible values of two parameters of the model and these were applied on two uncertainty models (models of spatial autocorrelation of the errors). All the values were considered realistic estimates of the parameters. However, such epistemic uncertainty about the models' parameters should be modelled using a fuzzy variogram (Loquin and Dubois, 2010b, 2011; Bardossy et al., 1990a) because in this case uncertainty is clearly caused by the lack of knowledge about the fixed but poorly known parameter (Helton et al., 2004). Applying the process proposed by Loquin and Dubois (2010b), the error could be modelled as a fuzzy surface which could contain all the possible combinations of parameters proposed by the authors.

From the numerical point of view, the biggest difference is that the statistical methods focus on the probable results and statements about probability of the outcomes while fuzzy arithmetic calculates the possible range of outcomes of the model. The outcome in form of a possible range of results is more valuable for further decision making (Hanss, 2005). In the decision making the worst and also the best case along with the most likely solution is an important piece of information for the decision maker. The statements about probability are less important because, especially for the critical risk applications, it is not important to what extent is some event probable. It is more important that the result is possible (Hanss, 2005). The analyses of surfaces are the same case as the critical risk or engineering applications, the decision maker will more likely be interested in the ranges of values than their probability distributions. The example of how this affects the calculation was provided by Hanss (2005), Lodwick et al. (2008) and also in chapter 2.5 of this thesis.

Chapter 8

Surface derivatives

Derivatives represent a useful characteristics as they provide a mathematical description of surface appearance. In geosciences, tools for their calculation are based on the approximation of a real surface by a finite number of elements (Wilson and Gallant, 2000). In the case of grid (rasters) structure these elements are represented by cells (pixels) (Waelder, 2007). This means that the derivative of a specific cell is calculated based on the values of neighbouring cells. There are two first derivatives of the surface: slope and aspect, several second derivatives describing various versions of curvature (Wilson and Gallant, 2000), a complete list of primary and secondary surface parameters and their significance is provided by Wilson (2012). All of those are commonly used in the geographical and environmental analyses, for example in fields like hydrology, geomorphology, geology, oceanography, ecology and others (Skidmore, 1989).

According to Wilson and Gallant (2000), two conditions have to be met to allow the calculation of derivatives of the surface. The cells of the grid have to be aligned to the geographical axes and the distance between the centres of the cell should be the same for the whole grid. If both these conditions are met, the calculation is rather straightforward. Otherwise, it is necessary to resample the grid according to those conditions. Another solution would be the modification of equations which is performed rather rarely due to the complexity of this process (Wilson and Gallant, 2000).

8.1 Methods of partial derivatives calculation

The basis of derivation determination is to calculate the partial derivatives of surface in two directions: North-South (denoted as z_y with respect to the alignment with this axis) and East-West (denoted as z_x). There are several methods for the calculation of those gradients, their comparison was performed by Jones (1998), Zhou and Liu (2004) and also by Skidmore (1989). The conclusion was that the 4-Cell method provides the most precise results, closely followed by Horn's method. The third best algorithm was a modified version of Horn's method and as the fourth the method of Sharpnack and Akin (Jones, 1998) was evaluated, these conclusions are not in a complete agreement with the conclusions made by Skidmore (1989). The study by Zhou and Liu (2004) was focused on the other elements of calculation than a comparison of various algorithms to establish their ranking. The algorithms were tested with respect to the data quality and resolution of the grid. However, findings from all these papers (Jones, 1998; Skidmore, 1989; Zhou and Liu, 2004) suggest that the 4-Cell method, Horn's method, Sharpnack and Akin's are all good estimators of the first derivatives of a surface. Based on these results, these three algorithms for the gradient calculation are considered in the thesis, they are also the most commonly implemented in GIS.

Skidmore (2007) noted that since his original research in 1989 (Skidmore, 1989), nearly nothing changed in the topic of slope and aspect calculation within GIS. Some improvements were made but the evolution is rather slow (Skidmore, 2007). He noted that there was a number of papers continuing to compare accuracy and efficiency of algorithms, confirming the main conclusions of the original paper. Both Skidmore (2007) and Zhou and Liu (2004) stress the fact that uncertainty (precision) of a surface is crucial for a correct estimation of the first derivatives of a surface, which makes the studies such as the one performed by Oksanen and Sarjakoski (2005a) important.

z_7	z_8	z_1
z_6	z_9	z_2
z_5	z_4	z_3

Figure 8.1: Node numbering convention in the neighbourhood of a central cell z_9 (edited from: (Wilson and Gallant, 2000))

In all upcoming formulæ the cells are labelled according to Fig. 8.1, the variable d denotes the size of the cell of the grid. The arrangement and numbering of the cells vary through literature and the formulæ for the calculation of derivations vary accordingly (Wilson and Gallant, 2000).

8.1.1 The 4-Cell Method

The 4-Cell method calculates the values of gradients only from the cells having a direct neighbourhood with the central cell. The method was firstly described by Fleming and Hoffer (1979). The equations for calculation are:

$$z_x = \frac{z_2 - z_6}{2d}, \tag{8.1}$$

$$z_y = \frac{z_8 - z_4}{2d}. \tag{8.2}$$

8.1.2 Horn's Method

Horn's Method considers even the cells in the neighbourhood having only one point common with the central cell. Cells having common edge have a higher weight assigned in the calculation. The method was presented by Horn (1981) and the equations are:

$$z_x = \frac{(z_1 + 2z_2 + z_3) - (z_7 + 2z_6 + z_5)}{8d}, \tag{8.3}$$

$$z_y = \frac{(z_7 + 2z_8 + z_1) - (z_5 + 2z_4 + z_3)}{8d}. \tag{8.4}$$

8.1.3 Sharpnack and Akin's method

Sharpnack and Akin's method is very similar to the Horn's method with the change that all cells have the same weight. The method was proposed by Sharpnack and Akin (1969) and the equations have the following form:

$$z_x = \frac{(z_1 + z_2 + z_3) - (z_7 + z_6 + z_5)}{6d}, \tag{8.5}$$

$$z_y = \frac{(z_7 + z_8 + z_1) - (z_5 + z_4 + z_3)}{6d}. \tag{8.6}$$

8.2 The calculation of slope and aspect

The three methods for the calculation of partial derivatives z_y and z_x that were mentioned in the previous section offer three possible ways to calculate the first derivatives. These partial gradients are further used to calculate the slope S and the aspect A. For the slope calculation in percentage (as a change of height within a distance unit) the following equation is used:

$$S = 100\sqrt{z_x^2 + z_y^2}. \tag{8.7}$$

If the result is to be provided in degrees, a slight modification is necessary, this slope is labelled as geographical:

$$S_g = \frac{180}{\pi}\arctan\left(\sqrt{z_x^2 + z_y^2}\right). \tag{8.8}$$

The calculation of aspect is a bit more complicated and requires the usage of the atan2 function:

$$A = \mathrm{atan2}(z_y, -z_x). \tag{8.9}$$

The mathematical aspect A is different from the geographical aspect A_g, A has the range $[-\pi, \pi]$ radians, the value of 0 for East and the values increase in a counter-clockwise direction. A_g has the range $[0, 2\pi]$ in radians or $[0°, 360°]$ in degrees, the value of 0 for North and the values increase in a clockwise direction (Wilson and Gallant, 2000). So, there is a need to adjust the values by this formula:

$$A_g = \begin{cases} 450° - \frac{180}{\pi}a & \text{if } \frac{180}{\pi}A > 90° \\ 90° - \frac{180}{\pi}a & \text{otherwise.} \end{cases} \tag{8.10}$$

Based on those equations the calculation of approximation of slope and aspect can be calculated from the surface represented by the grid.

8.3 The first derivatives of fuzzy surfaces

In any analysis calculated on a fuzzy surface uncertainty of a surface is propagated through the analysis into a result. Such result then shows uncertainty connected with the input data represented as fuzzy numbers. So far there are two examples of the calculation of a fuzzy slope in the literature provided by Fonte and Lodwick (2005) and Waelder (2007). Unfortunately, in both cases the fuzzy slope is not a main focus of the research so it is discussed only very briefly. Fonte and Lodwick (2005) use the fuzzy slope to identify areas having a slope potentially higher than 25% but the calculation serves as one of several examples in the article, so it is discussed very briefly. Waelder (2007) provided methods for the calculation of partial derivatives using a finite elements method but the presented method is focused on a fuzzy surface constructed using purely triangular fuzzy numbers. The equations are adjusted to work on such surface but it does not handle the calculation of a fuzzy slope in general, because triangular fuzzy numbers are only one type of a theoretically infinite set of fuzzy number types. These case–specific adjustments of equations are common for presenting methods utilizing fuzzy arithmetic (Hanss, 2005). There has been no attempt (of which the author is aware) to calculate the aspect of a fuzzy surface.

The basis for the determination of both slope and aspect is the calculation of gradients z_x and z_y. Considering the fact that all inputs are uncertain and represented by fuzzy numbers, the results will also contain uncertainty and they will also be represented by fuzzy numbers. The calculation of gradients itself is based on basic arithmetic operators that have fuzzy equivalents according to Eqs. (3.7), (3.9) and (5.6,5.7,5.8,5.9). This applies for all three methods of the gradient calculation (Eqs. (8.4),(8.2),(8.6)).

8.3.1 Slope

Calculating the slope of a fuzzy surface according to Eqs. (8.7, 8.8) does not need any special approaches. The square of the fuzzy number can be calculated according to the equations from chapter 5.5.1 and the square root is a monotonous function and can be calculated according to Eqs. (3.7) and (5.12). If the slope is to be provided in degrees, Eq. (8.8) is used. As previously mentioned, there is no problem with the usage of crisp numbers with fuzzy numbers while calculating. The arctan function is again a monotonous one and as such calculated according to Eq. (5.12). Obtaining the value of slope as a fuzzy number is therefore a relatively simple matter.

An example of a slope between two points with fuzzy height is shown in Fig. 8.2. The figure illustrates the three most important results of the fuzzy slope calculation – the modal slope being equal to the slope that would be obtained if the calculation is done with crisp numbers, and also two limits – the minimal and maximal slope. The minimal slope in this case is 0 because there is an overlap where the minimal value of P_1 is lower than the maximal value of P_2. Visualization of a three dimensional example is not possible because such visualization is not understandable.

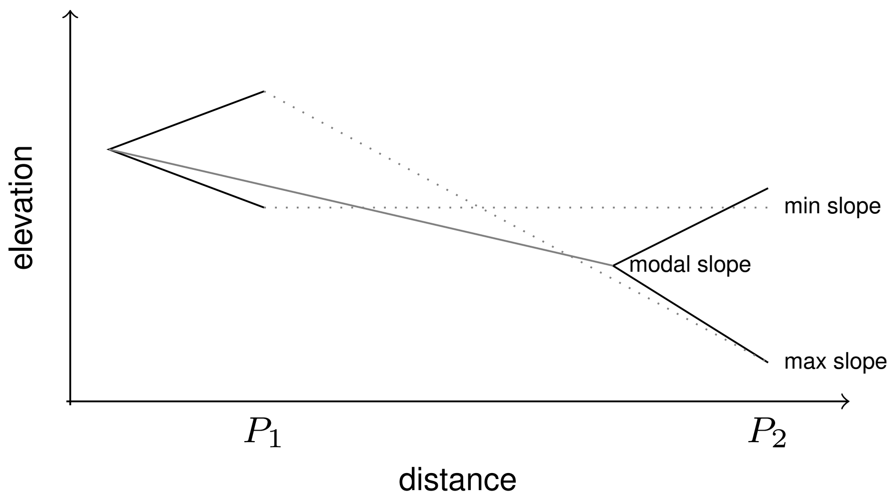


Figure 8.2: A fuzzy slope between the points P_1 and P_2. Minimal, maximal and modal slope between points with fuzzy height are displayed.

8.3.2 Aspect

The aspect calculation is more complicated then the slope calculation. Equation (8.9) contains the function atan2 (Gaile and Burt, 1980) that has to be calculated for two fuzzy arguments (atan2$(\tilde{Y}, \tilde{X})$). This in not a trivial operation and the function has to be modified to allow the calculation. The process of calculation of atan2$(\tilde{Y}, \tilde{X})$ is described in chapter 5.5.2 on the page 42. The resulting mathematical orientation needs to be recalculated into the geographical orientation according to Eq. (8.10). The value used for comparison is the maximal value of kernel of a fuzzy number – the value of $\overline{a}$ from the α-cut $\tilde{A}_1$ instead of purely the value of A in Eq. (8.10). After calculating A_g according to Eq. (8.10), the resulting values of A_g then do not fit the original range of aspect values $[0°, 360°]$, which is a result of the propagation of fuzzy numbers through the calculation. Actually, the resulting angles are from the range $[-90°, 630°]$, however, these values should be interpreted according to Fig. 8.3, which means that the negative values v have the same aspect as $360° + v$ and the positive values v higher than $360°$ are equal to $v - 360°$. The fuzzy orientation is more complicated for the interpretation but it is necessary to calculate them in such form to allow the correct propagation of fuzzy numbers through the calculation. The similar issue (called the zero direction problem) is found in directional statistics (Gaile and Burt, 1980; Mardia and Jupp, 1999) and it is common for all angular data. For the visualization and interpretation it is necessary to ensure that all those values will be interpreted correctly.

In Fig. 8.2 the aspect might be negative, indicating that P_1 is higher than P_2, but also positive because there is a small overlap that allows also a solution where P_2 is higher than P_1. In case like this, there exist arguments for both solutions because the information with uncertainty provides contradictory results. However, the evaluation of possibility of the result can be done, suggesting that the first outcome is more possible. This example nicely illustrates that the results of analysis with uncertainty are sometimes not clear and they require more detailed assessment in order to provide useful information for the user.

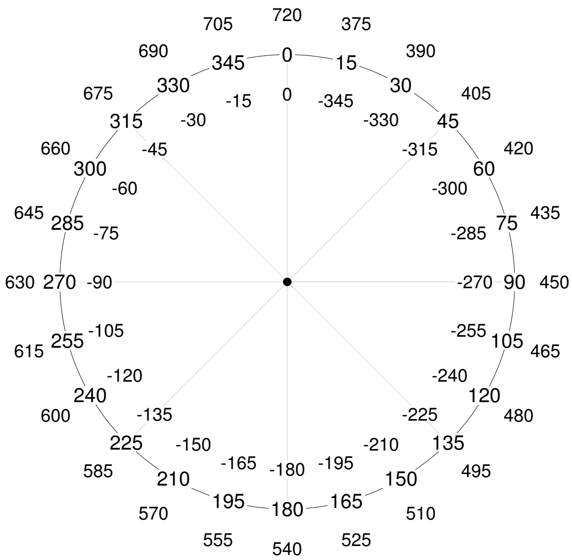


Figure 8.3: Normal, double positive and negative values of angles.

8.4 Example

In this section an example of the calculation of aspect and slope using Horn's method (Eq. (8.4)) will be shown. The method was chosen because that is the one most commonly implemented in GIS. For the sake of readability, the calculation will only be presented for the α-cuts 0 and 1, even though theoretically any number of the α-cuts can be chosen. Each alpha cut of the fuzzy number $\tilde{A}$ will be written as a set of α-cuts with each α-cut defined as $(\alpha : \underline{\tilde{X}_\alpha}; \overline{\tilde{X}_\alpha})$. The distance between the cells is $d = 10$ meters. The surface used in this example is visualized in Fig. (8.4). The input fuzzy numbers of the neighbouring cells are triangular fuzzy numbers and have the following definition:

$$z_1 = (0.0 : 382.81; 384.01)(1.0 : 383.41; 383.41)$$
$$z_2 = (0.0 : 384.34; 385.5)(1.0 : 384.92; 384.92)$$
$$z_3 = (0.0 : 385.83; 386.93)(1.0 : 386.38; 386.38)$$
$$z_4 = (0.0 : 385.63; 386.63)(1.0 : 386.13; 386.13)$$
$$z_5 = (0.0 : 385.46; 386.22)(1.0 : 385.84; 385.84)$$
$$z_6 = (0.0 : 384.13; 384.87)(1.0 : 384.5; 384.5)$$
$$z_7 = (0.0 : 382.63; 383.53)(1.0 : 383.08; 383.08)$$
$$z_8 = (0.0 : 382.74; 383.78)(1.0 : 383.26; 383.26)$$

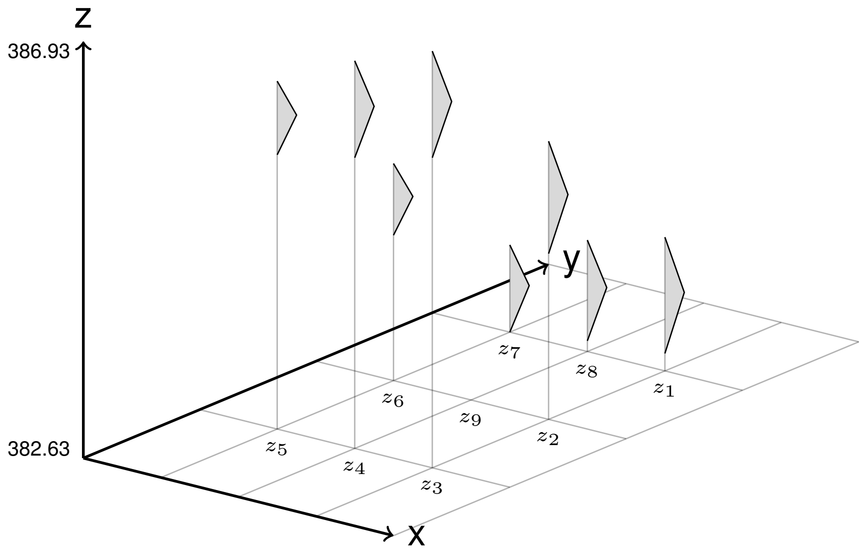


Figure 8.4: Visualization of a small fuzzy surface used in the example.

The first step is to calculate the values of z_x and z_y, to do that we firstly extract the necessary α-cuts from the fuzzy numbers according to Eq. (3.9) and then calculate the values for each α-cut according to Eq. (8.4), applying Eqs. (5.6,5.7,5.8,5.9) for each operation. The resulting fuzzy numbers have the following values:

$$z_x = (0.0 : -0.027; 0.07)(1.0 : 0.021; 0.021)$$
$$z_y = (0.0 : -0.19; -0.09)(1.0 : -0.14; -0.14)$$

With the knowledge of gradients, the calculation of slope is a simple matter (Eq. (8.7)), the equations from chapter 5.5.1 will be used to calculate the power of fuzzy numbers, the addition of fuzzy numbers is done according to

Eqs. (5.6,5.7,5.8,5.9) and the square root can be calculated as a monotonous function (Eq. (5.12)).

$$z_x^2 = (0.0 : 0.0; 0.005)(1.0 : 0.00; 0.00)$$
$$z_y^2 = (0.0 : 0.009; 0.037)(1.0 : 0.02; 0.02)$$
$$S = (0.0 : 0.093; 0.206)(1.0 : 0.145; 0.145)$$

The value of slope S can be further transformed into degrees by calculating $\arctan(S)$ (Eq. (5.12)) and then multiplying the result by $\frac{180}{\pi}$ or into percent by multiplying it by 100. The slope in percent is:

$$S = (0.0 : 9.3; 20.6)(1.0 : 14.5; 14.5).$$

To calculate the aspect Eq. (8.9) will be used. The calculation of mathematical aspect of a fuzzy surface is done according to the procedures mentioned in chapter 8.3.2.

$$A = (0.0 : 73.76; 126.88)(1.0 : 98.48; 98.48)$$

The mathematical aspect needs to be turned into the geographical aspect according to Eq. (8.10). For the comparison the value of $\overline{a}$ from the α-cut $\tilde{A}_1$ (in this example the value is 98.48) is used and the geographical aspect is obtained.

$$A_g = (0.0 : 323.12; 376.24)(1.0 : 351.52; 351.52)$$

As can be seen, this is the case when the range of a resulting fuzzy aspect is higher than 360°. In practice this means that the possible aspect for the cell ranges from 323.12° to 360° and from 0° to 16.24°, however, such result is not a valid fuzzy number because such fuzzy set is not convex and thus it cannot be a fuzzy number. To avoid this fact, an alternative range of values $[-90°, 630°]$ needs to be introduced for the calculation of a fuzzy aspect as described in chapter 8.3.2.

Through the whole presented example for all variables the kernel value of each fuzzy number is the same value as it would be in the case of calculation with crisp numbers. This fact shows that the propagation of uncertainty was done correctly because if the triangular fuzzy numbers, where the kernel value corresponds to what originally was a crisp number, are used, the kernel value of the result should be equal to the result of the crisp calculation (Hanss, 2005).

8.4.1 Comparison with Monte Carlo

As a comparison, the same calculations were performed with the usage of the Monte Carlo simulations using 100, 500, 1 000, 10 000 and 1 000 000 iterations. A similar experiment to determine the precision and time demands was also performed on an even simpler example (Caha and Dvorský, 2013b). The triangular probability distributions were used as they are specified by three values (Evans et al., 2000), which makes them very similar to the triangular fuzzy numbers. The results of the simulations are summarized in Tab. 8.1. It is obvious that as the number of simulations rises, the ranges get closer to the range identified by fuzzy arithmetic. However, it is very improbable for Monte Carlo to identify a complete range of results, even with a very high number of simulations.

Number of simulations	Slope range (%)	Aspect ranges (°)	
100	$[12.96, 16.72]$	$[341.58, 360]$	$[0, 2.45]$
500	$[11.95, 17.56]$	$[342.24, 360]$	$[0, 1.13]$
1 000	$[12.03, 17.05]$	$[342.36, 360]$	$[0, 4.40]$
10 000	$[11.31, 17.98]$	$[338.78, 360]$	$[0, 3.24]$
1 000 000	$[10.34, 18.45]$	$[336.57, 360]$	$[0, 6.56]$

Table 8.1: Intervals provided by the Monte Carlo simulations. Slope is in percent and aspect in degrees.

The results show that Monte Carlo failed to identify results having very small probability of occurrence, but these are feasible solutions to the problem. This outcome is in an agreement with the example provided by Hanss (2005). These solutions can be perceived as the best/worst possible solutions and possibly they can be very important for the decision making. The complete range of outcomes should be $[9.3, 20.5]$ for the slope and $[323.12, 376.24]$ for the aspect, which should be divided into two intervals $[323.12, 360]$ and $[0, 16.24]$ for the interpretation. Monte Carlo did not reach these widths of intervals but it is visible from Tab. 8.1 that as the number of simulations increases, the estimates are actually converging towards the results provided by fuzzy arithmetic. However, the number of simulations to obtain the true range is likely to be very high since the extension of the intervals is not significant even for the significant increase in the number of simulations e.g. the change between the fourth and the fifth row of Tab. 8.1.

Chapter 9

The visibility analysis

The visibility analysis (also called the viewshed operation by Fisher (1992)) in GIS is used to identify the areas of a surface that are visible from a viewing point (Fisher, 1992, 1994). The outcome of the analysis is presented as a boolean product dividing the surface into visible and invisible areas. Sometimes the variant of an algorithm producing the visibility angle for the visible areas and no data values for the invisible areas is used (Neteler and Mitasova, 2008). However, for the purpose of this thesis only the variant producing the results in form of visible and invisible areas is more important.

The visibility analysis has a widespread usage in various fields e.g. landscape and urban planning (Fisher, 1995, 1996; Hernández et al., 2004; Ohsawa and Kobayashi, 2005), archaeology (Fisher et al., 1997; Ogburn, 2006), location of various transmitters and receivers (O'Sullivan and Turner, 2001) and management of environmental resources (O'Sullivan and Turner, 2001). The analysis of visibility is quite sensitive on the quality (precision) of the surface, so the introduction of a relatively small error can lead to very divergent results (Fisher, 1992). Because the analysis is rather sensitive to uncertainty of the surface, the topic has been a subject of extensive research (Felleman and Griffin, 1990; Fisher, 1991, 1992, 1993, 1994; Sorensen and Lanter, 1993). Visibility in GIS can be calculated over a triangulated irregular network (TIN) (Nagy, 1994), but usually the grid structure is used (Fisher, 1991, 1994). Fisher (1993) noted that the implementation of a viewshed analysis contains several important design decisions. These design decisions affect the way how the terrain aproximation (grid) is translated into straight segments allowing the assessment of visibility. The author concluded in that there is rather a small chance that two implementations will actually provide the same results because it is nearly impossible that any two programmers would implement the algorithm exactly in the same way (Fisher, 2007b). That makes it very difficult to actually compare the viewshed algorithms from different software (Fisher, 1993).

9.1 The algorithm

The processes for determining the visibility of a surface from a viewpoint can be divided into two steps (Fisher, 1993). The first step is to infer elevations located on a line-of-sight (LOS) from the grid. This can be done in many ways, all of them are summarized in the article by Fisher (1993). The second step is to determine which points on the LOS are visible and which are not visible. This part of the calculations is rather simple, the algorithm is straightforward, but the result is highly dependent on the elevations that were inferred in the first step. The calculation of visibility is rather time demanding since it needs to be calculated from the viewpoint to every cell in the grid.

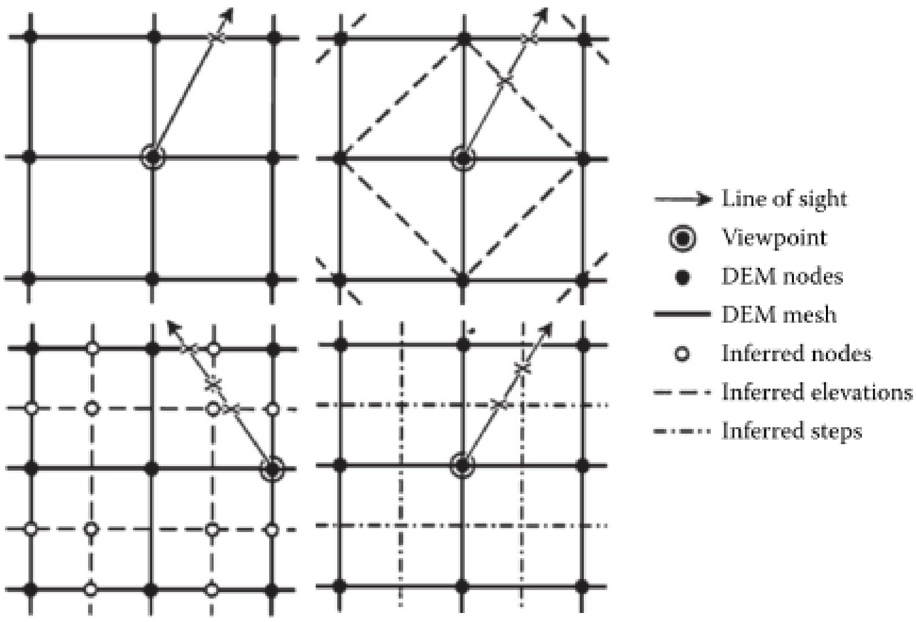


Figure 9.1: Different algorithms for inferring heights on a grid. (adapted from Fisher (1993))

As mentioned previously, the grid as a surface model is rather simplified and idealized. In its most simplistic meaning (the whole cell has a constant value) the model is rather inappropriate for the visibility calculation (Fisher, 1993). The overview of this simplistic approach was presented by Sorensen and Lanter (1993), who provided a complete description of the visibility calculation on a grid where each cell has a constant value, and all issues associated with it. To overcome some of these issues, several ways how to infer more height values on the grid are used for the visibility calculation. According to Fisher (1993), the main variants are: *a*) linear interpolation between the grid neighbours (top left), *b*) triangulation of the grid (top right), *c*) grid constraint of the mesh (bottom left), *d*) the stepped model (bottom right), all the variants are shown in Figure 9.1.

Besides that, the approximation of a viewpoint and a target cell can influence the result of the analysis (Fisher, 1993). Both the viewpoint and the target cell can be approximated by 4 points (in each corner of the cell), normally both cells are considered points. So there exist four possible variants how the approximation may look like: *a*) point-to-point, *b*) cell-to-point, *c*) point-to-cell, *d*) and cell-to-cell (Fisher, 1993). However, for practical application only the point-to-point variant is used, otherwise the calculation can become too complex. Theoretically the cell-to-cell variant requires 16 (4 points approximating the viewpoint cell × 4 points on the target cell) LOS calculations to proclaim the target cell is invisible, yet it is enough to find the one that results as visible to identify the target cell as visible (Fisher, 1993). Even though, this does not affect the majority of algorithms (as they use the point-to-point approximations), it is a topic that should be mentioned, because it is a possible reason why algorithms might differ. Fisher (1993) reported nearly 50 % increase of a viewable area after the change from point-to-point to cell-to-cell cell approximation.

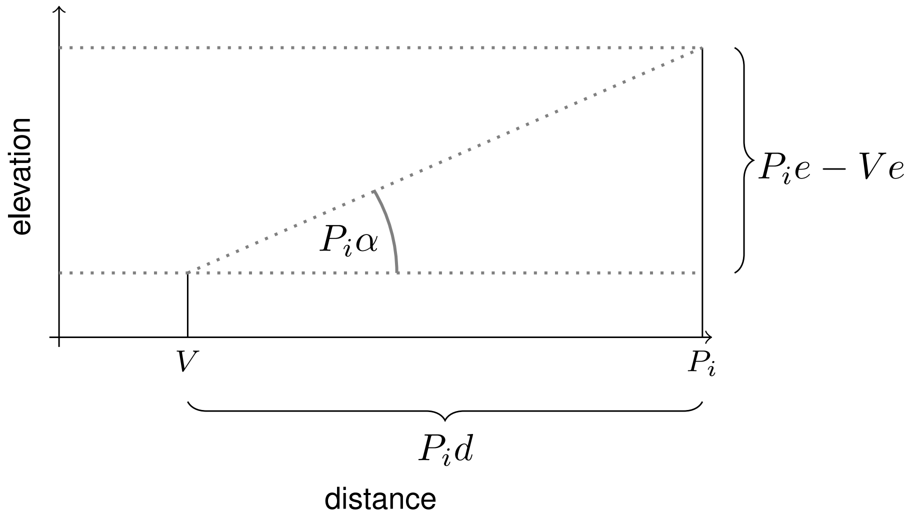


Figure 9.2: Calculation of a viewing angle from the viewpoint V to the point P_i on LOS.

After obtaining LOS, with all relevant points on it, it is possible to determine which points on LOS are visible and which are not. LOS consists of the viewpoint V and a set of points $P = P_1, \ldots, P_n$. The viewpoint has the elevation Ve, which consists of the elevation at the viewpoint plus some offset, denoting height of the observer over the surface. Each point P_i on LOS has a distance from the viewpoint $P_i d$ allowing to order the points from the closest to the farthest, and also the elevation $P_i e$. The distance and elevation are necessary to calculate the viewing angle (in radians):

$$P_i\alpha = \arctan \frac{P_i e - Ve}{P_i d}, \tag{9.1}$$

which denotes the vertical angle from the viewpoint to the point P_i (Fig. 9.2). The angle is positive for the points higher than the viewpoint and negative for the points lower than the viewpoint. The point P_x is visible if the angle $P_x\alpha$ is higher than all $P_p\alpha$ where $p < x$ (Fig. 9.3). The process is algorithmically described in Algorithm 3.

Algorithm 3 A function for the determination of visibility of the last point on LOS

Require: a set of points P on LOS in an ascending order according to the distance from the viewpoint ($P_i d$) and viewpoint V
Ensure: $True$ if the point is visible and $False$ if it is not

```
procedure DETERMINEVISIBILITY(P, V)
    visibile ← False
    l ← length(P)
    maxAngle ← (P_1 e − Ve)/P_1 d
    if l == 1 then
        visible ← True
    else
        for all i ∈ (2, . . . , l − 1) do
            angle ← (P_i e − Ve)/P_i d
            if maxAngle < angle then
                maxAngle ← angle
            end if
        end for
        angle ← (P_l e − Ve)/P_l d
        if maxAngle < angle then
            visible ← True
        end if
    end if
    return visible
end procedure
```

An example of LOS shown in Fig. 9.3 clearly shows that the points P_1, P_2 and P_8 are visible and all the others are invisible. The angle between the viewpoint and the point P_2 hides all the following points on LOS except the point P_8.

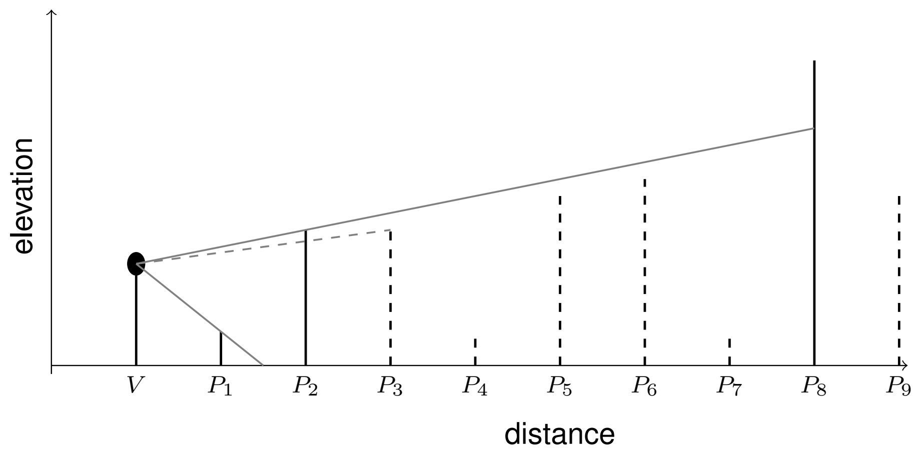


Figure 9.3: Determination of the visible points on LOS, the visible points have vertical lines as full lines, invisible as dashed lines. Three viewing angles are shown as examples.

9.1.1 A Variant of the algorithm

An interesting variant of the visibility algorithm can be provided by returning the difference between the angle of the last point on LOS and the maximal angle. That is replacing the comparison between the values in Algorithm 3 by their subtraction. Instead of depending on the result of $maxAngle < angle$, $visibility$ would be equal to $angle - maxAngle$. This outcome is potentially very interesting for assessing the appropriateness of visibility. The sign of this result indicates the visibility (positive values) or invisibility (negative values) and the magnitude of the value expresses how well the endpoint will be visible. Small positive values indicate the areas that are just barely above the highest point on LOS so they will not be as clearly visible as the points having the difference higher. In an analogous way, the points having a small negative angle may become visible because of e.g. high vegetation, buildings etc. This is not so much possible for large negative values.

The illustration of this variant of the algorithm is in Fig. 9.4. The point P_5 shows the variant of a visible point with a very small difference to the highest angle on LOS (P_1) consequently, P_5 will be barely visible. On the other hand, P_6 has a significantly high difference to the maximal angle (P_5) so it will be clearly visible. Point P_2 illustrates an example that has a small negative value of the difference and P_3 has a high negative value.

This variant of the visibility algorithm is closely related to the concept implemented in GRASS GIS (Neteler and Mitasova, 2008), providing the value of the vertical angle for the visible areas but only for the endpoint,not the difference of the angle to the highest angle on LOS. This variant is in a few ways inspired by the fuzzy visibility as presented by Fisher (1994). However, to the best of the author's knowledge this variant of the visibility algorithm has not been described in any literature yet.

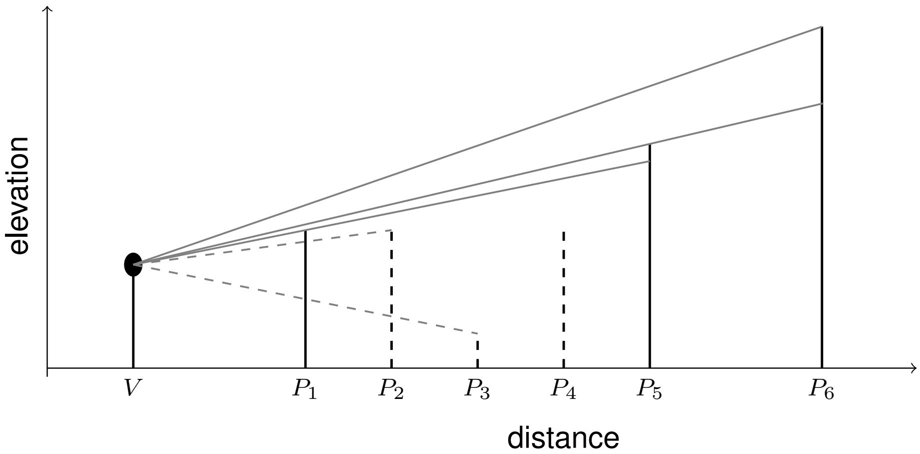


Figure 9.4: An example illustrating the difference between the clear visibility (point P_6) and worse visibility (point P_5) due to the difference between the visibility angle of the point and the highest angle on LOS.

9.2 The visibility on surfaces with uncertainty

The effects of surface uncertainty on the visibility analysis have been studied since the introduction of uncertainty propagation to GIS (Felleman and Griffin, 1990; Fisher, 1991). Since the first articles by Fisher (1991, 1992), there has been a lot of confusion about the correct naming of results. The results were originally denoted as fuzzy viewsheds in both cases, but the processes used for their calculation were definitely based on statistical methods of the uncertainty propagation. The first article (Fisher, 1991) even names explicitly the Monte Carlo method. This error was later recognized and corrected (Fisher, 1994, 2007b) by clearly stating that those results are actually the probable viewsheds. However, the problem remained and occasionally a wrong usage of this term can be found in literature (e.g. Loots et al. (1999)).

The definition of a probable viewshed was coined by Fisher (1994) after the realization that what was originally named as a fuzzy viewshed cannot be a fuzzy viewshed, even though it has some properties of a fuzzy viewshed (Fisher, 2007b). The term probable viewshed relates to the usage of probabilistic methods and statistical evaluation of the result. The probable viewshed is calculated in the same manner as the slope and other derivatives by the Monte Carlo method as in Fig. 7.4. After the calculation the outcomes for each cell can be summarized by dividing the number of results when the cell was visible by the number of iterations. The semantics of the result is frequentist, showing the ratio of favourable outcomes (the cell is visible) from all experiments. So the resulting value 1 means that the cell was visible in all results, 0 means it was never visible. Since the calculations in the older articles (Fisher, 1991, 1992) were done with extremely low number of iterations (20) and the issue of convergence was never mentioned, the results might be questionable. The convergence issue of the probabilistic visibilities was studied by Nackaerts et al. (1999). The conclusion of this research was that the major fluctuations reduced rapidly after 30 iterations and the results converged with sufficient accuracy after approximately 60 iterations. Even though in some cases fluctuations can be seen even after 100 iterations (Nackaerts et al., 1999). The confidence intervals for the probabilistic visibility 0.5 are equal to ± 0.2 after approximately 20 iterations, ± 0.1 after 90 iterations and after 200 iterations are roughly equal to ± 0.08 (Nackaerts et al., 1999). This shows that even the usage of a relatively high number of iterations may not guarantee completely precise estimates.

The fuzzy viewshed in its more correct naming was introduced by Fisher (1994). The distance decay function is used to incorporate the fact that the visual clarity drops with increasing distance in the otherwise binary viewshed. It means that objects farther from the viewpoint are less likely to be well visible even if they are on the visible part of a surface. This approach was later extended by Ogburn (2006) by the introduction of Higuchi viewsheds (Higuchi, 1983) into the process. However, since this approach does not consider uncertainty of the surface, it is not important for the purpose of this thesis, except to the correct characterization of the terms. Fisher (1994) noted that it is indeed possible to combine the probable and fuzzy viewsheds to produce probable fuzzy viewsheds that would contain uncertainty of the surface as well as the distance decay function.

So far the only approach trying to calculate the visibility on a fuzzy surface was proposed by Anile et al. (2003). However, if the algorithm is studied more closely it quickly becomes obvious that the algorithm is partially optimizing towards the production of larger visible areas. The algorithm tries to confirm that LOS, which would allow the endpoint to be visible from the viewpoint, does exists. Such LOS can be often found and thus the resulting viewshed will usually be larger then a classic binary viewshed, but the result is usually far too optimistic. The algorithm acts in such way as if uncertainty works in favour of visibility. It does not account for the contradictory information that is actually contained in the fuzzy surface.

9.3 The possibilistic visibility on fuzzy surfaces

The term possibilistic visibility or possibilistic viewshed is selected with respect to the previous research that coined the term fuzzy visibility (Fisher, 1994). To avoid collision with his usage of the term fuzzy visibility, it was necessary to introduce another term to describe the result. The name is also fitting because of the usage of methods from the possibility theory in the centre of the algorithm and the result actually comes in the form of possibly and necessary visible areas.

The part of the algorithm that is affected by the fuzziness of a surface is the determination of visibility on LOS. Even though the part where the heights are inferred is also affected it is not as much important for the algorithm. When LOS points and the viewpoint are located on a fuzzy surface their height P_ie and Ve are actually the fuzzy numbers $\tilde{P_i}e$ and $\tilde{V}e$. As a logical consequence of this fact the vertical angle $P_i\alpha$ between the point P_i and the viewpoint V must also be the fuzzy number $\tilde{P_i}\alpha$. The means of fuzzy arithmetic are applied to Eq. (9.1). Firstly, the fraction is calculated according to Eqs. (3.7,3.9) and (5.6,5.7,5.8,5.9). The function arctan is a monotonous one and as such calculated according to Eq. 5.12. In Fig. 9.5 the limit values for the 0-cut of the angle $\tilde{P_i}\alpha$ are shown. From the image it is obvious that the angle between the points with fuzzy elevation will also be fuzzy.

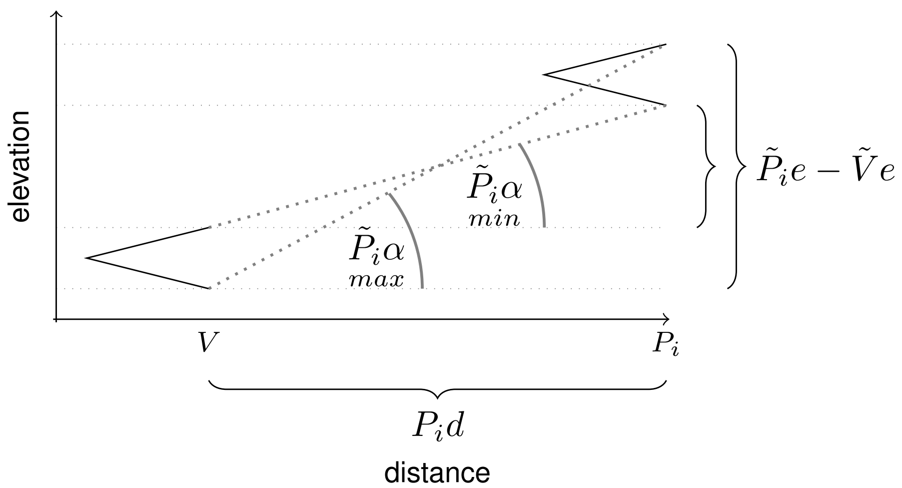


Figure 9.5: The angle between the viewpoint V and the point P_i on a fuzzy surface as a fuzzy number.

In the classic algorithm for visibility (Algorithm 3) the highest angle between the viewpoint and the target point has to be found. This is rather easy for the crisp values, however for the fuzzy numbers selection of the higher out of two values might not be so straight forward (Kaufmann and Gupta, 1985) because the maximum of fuzzy numbers might actually contain parts of both compared numbers (chapter 6.1). Using

Eq. (6.2) the maximum of fuzzy numbers is obtained. This operation is repeated for all P_i on LOS except for the endpoint. The obtained maximal value of the fuzzy angle is compared to the fuzzy angle of the last point on LOS. Because the ranking of fuzzy numbers is not as easy as comparing the crisp values, the indices described in chapter 6.2 are used. All four indices are of interest here, but especially the possibility of exceedance (Eq. (6.3)) and the strict necessity of exceedance (Eq. (6.6)). The possibility of exceedance evaluates the situation when all uncertainty in the surface works in favour of visibility. The result is as much of the visible area as possible. On the other hand, the strict necessity of exceedance shows the situation when all uncertainty in the surface works against the visibility. The resulting visible area is thus smaller. Since all the indices return a value from the range $[0, 1]$, visibility is naturally graded, showing the areas having a small possibility of being visible as well as completely possible areas. Also the areas having small necessity values which might be visible and values of high necessity, which show the areas that are definitely visible, are shown.

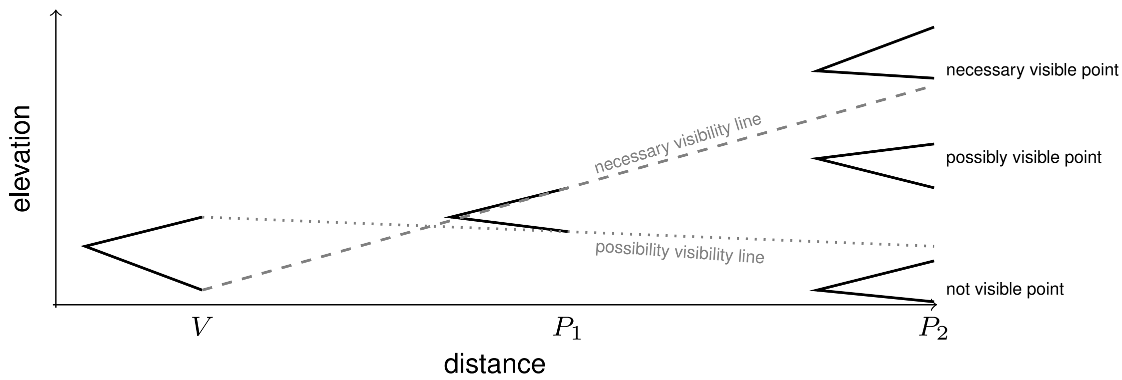


Figure 9.6: The possible and necessary visibility line from the viewpoint V through the point P_1 to three variants of P_2.

In Fig. 9.6 the example of fuzzy LOS is shown. The possible visibility line has the lowest value of the angle $\tilde{P}_i\alpha$ while the necessary visible line shows the highest value of the angle $\tilde{P}_i\alpha$. Amongst those, there exists and infinite number of other visibility lines, associated to different α-cuts of the fuzzy numbers. The point P_2 shows three possible relations the point can have to the lines of sight. The lowest example of P_2 is not visible under any circumstances as it lies under the possible visibility line. The highest example is always absolutely visible because it is above the necessary visible line. The middle example of P_2 will have the value of possibility higher than 0 showing that there are chances that it is visible, while the necessity will be equal to 0 which means that the chances or real visibility are not strong.

A more complex example showing how the propagation of a maximal fuzzy number affects LOS is shown in Fig. 9.7. It is obvious how a possible line of sight changes through the points P_1, P_2 and P_3, while the necessary LOS does not change through the example because the maximal angle is defined by the viewpoint and the point P_1. The point P_4 in this example will have the possibility 1 of being visible, but the strict necessity will be lower than 1 indicating that it is not completely sure that the point will be visible.

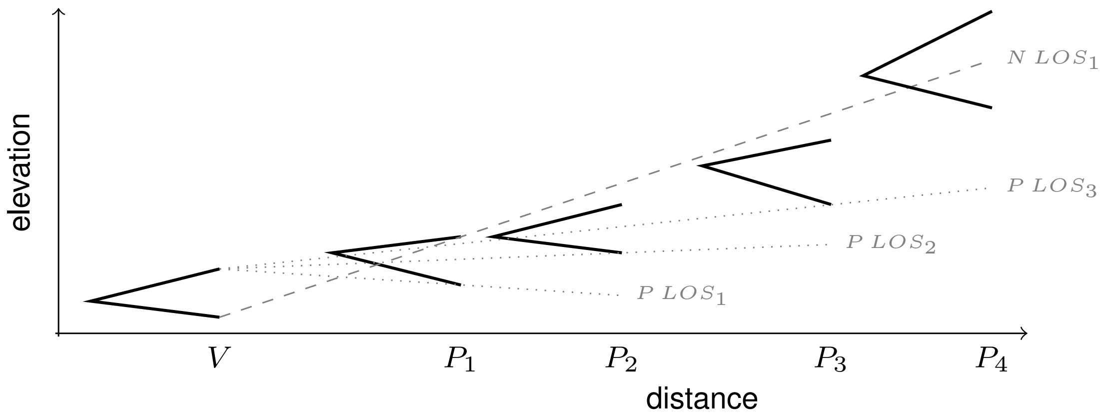


Figure 9.7: The necessary line of sight $N\ LOS_1$ and the possible lines of sight $P\ LOS_1$, $P\ LOS_2$, $P\ LOS_3$. The example shows how the propagation of a maximal angle affects the possible line of sight.

9.3.1 The algorithm

As mentioned previously, there are two important parts of the algorithm for the calculation of visibility on a grid. Firstly, points on LOS need to be determined for LOS under consideration and then visibility can be calculated.

For the practical implementation a method named as the grid constrain of the mesh (Fisher, 1993) was selected. The algorithm is described in Fig. 9.8. The method thickens the grid by adding inferred points at the corners of the cells, these inferred points are calculated as mean of the four neighbouring cells. The calculation of these corner points is described by Domingo-Santos et al. (2011). The example is shown in Fig. 9.8 for the point IP_4. The points on LOS are placed when LOS intersects with the cell border or passes directly through either the centre of the cell or the inferred point. If the LOS point is placed on the intersection of LOS with the cell border then its height is calculated as the weighted mean of two closest inferred points. E.g. the height of P_2 is calculated based on the heights of IP_2 and IP_1. The weighting factor is the distance between the LOS point and the inferred point, so in this example the point IP_2 will have higher influence because it is closer to P_2 than IP_1. The detailed description of the calculation can be found in Sorensen and Lanter (1993); Fisher (1993). The result of this procedure is a set of points P_i on LOS between the viewpoint and the target cell. Since the algorithm is designed for a fuzzy surface, it is obvious that the elevation of each cell is a fuzzy number, so also the elevations of inferred points and the line of sight points are fuzzy numbers that were calculated by means of fuzzy arithmetic.

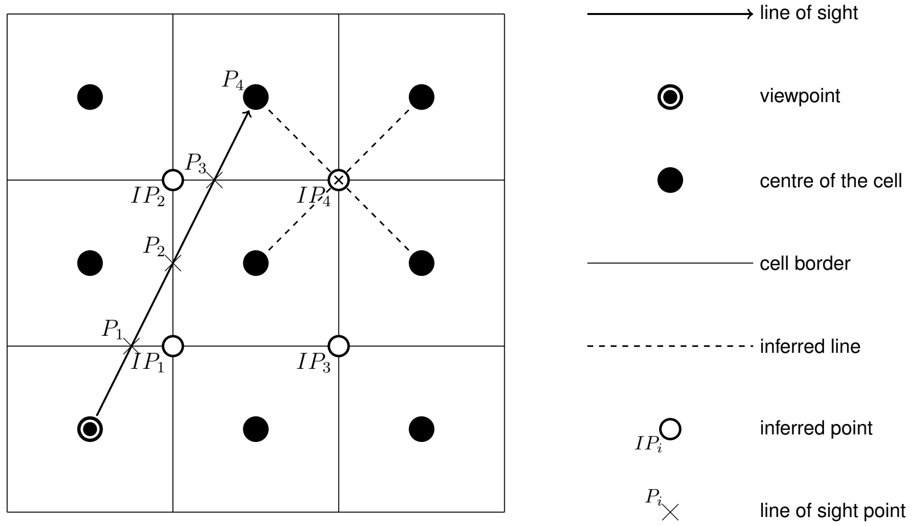


Figure 9.8: The inference of heights on LOS. Based on: Fisher (1993) and Sorensen and Lanter (1993).

After LOS is determined and constructed, the determination of visibility between the viewpoint and the target cell can be done. The process is done according to Algorithm 4. For each point on LOS a vertical angle needs to be calculated. The maximal angle is then calculated by the calculation of a maximal fuzzy number from the points between the viewpoint and the target point (excluding the target point itself). This ensures the propagation of maximal values through the line of sight. Then the comparison of the maximal angle and the angle for the target cell can be done by ranking those two fuzzy numbers (according to Eqs. (6.3,6.4,6.5,6.6)).

For each target point four values of visibility are obtained. These values provide the user with a lot more information than the classic boolean visibility and even more than the probable visibility. Such outcome allows a rather complex assessment of the situation as well as complex reasoning about it.

Algorithm 4 A function for determination the visibility of the last point on fuzzy LOS

Require: a set of points P on the LOS in ascending order according to the distance from the viewpoint ($P_i d$) with the fuzzy height $\tilde{P_i} e$ and the viewpoint V with the fuzzy height $\tilde{V} e$

Ensure: PSE - exceedance possibility, NSE - exceedance necessity, PS - strict exceedance possibility, NS - strict exceedance necessity; all of them with the values from the range $[0, 1]$

procedure DETERMINEVISIBILITY(P, V)

    $PSE, NSE, PS, NS \leftarrow 0$
    $l \leftarrow length(P)$
    $max\tilde{Angle} \leftarrow (\tilde{P}_1 e - \tilde{V} e)/P_1 d$
        ▷ The first point is always absolutely visible
    if $l == 1$ **then**
        $PSE, NSE, PS, NS \leftarrow True$
    else
        for all $i \in (2, \ldots, l-1)$ **do**
            $\tilde{P}_i \alpha \leftarrow (\tilde{P}_i e - \tilde{V} e)/P_i d$
                ▷ Produces the maximal angle based on the angle of the current
                ▷ point and the maximal angle on LOS so far
            $max\tilde{Angle} \leftarrow \max(max\tilde{Angle}, \tilde{P}_i \alpha)$
        end for
            ▷ The angle of the last point
        $\tilde{P}_l \alpha \leftarrow (\tilde{P}_l \alpha - \tilde{V} e)/P_l d$
            ▷ Calculated according to Eqs. (6.3,6.4,6.5,6.6)
        $PSE \leftarrow \Pi_{\tilde{P}_l \alpha}([max\tilde{Angle}, \infty))$
        $NSE \leftarrow \mathcal{N}_{\tilde{P}_l \alpha}([max\tilde{Angle}, \infty))$
        $PS \leftarrow \Pi_{\tilde{P}_l \alpha}(]max\tilde{Angle}, \infty))$
        $NS \leftarrow \mathcal{N}_{\tilde{P}_l \alpha}(]max\tilde{Angle}, \infty))$
    end if
    return PSE, NSE, PS, NS
end procedure

9.3.2 The variant of the possibilistic algorithm

In the same way as presented in chapter 9.1.1, the possibilistic algorithm can be altered to provide a fuzzy angle. Such fuzzy angle ($\tilde{P}_i \alpha$) is calculated as the difference between the angle of the last point on LOS and the maximal angle on LOS. The subtraction is calculated by means of fuzzy arithmetic. The result can be classified into three main categories: positive fuzzy numbers indicating definitely visible areas, negative fuzzy numbers showing areas that are certainly invisible and fuzzy zeros defining areas with uncertain visibility. A more detailed assessment of the result can be done by comparing the resulting angle to a specific threshold, most likely 0, by the process described by Dubois and Prade (1983). The same process was used to query the crisp data with soft queries by Caha et al. (2014c). This variant of the algorithm is described here because of its potential for further research.

9.4 Example

In the same manner as in chapter 8.4 where an example of the slope and aspect calculation was shown, in this chapter a simple calculation of LOS will be presented. The fuzzy numbers will again be simplified into solely triangular representations with two α-cuts for the sake of readability. The example is completely artificial, only to demonstrate the process of calculation. The calculation is done according to the Algorithm 4.

Let's consider LOS with the viewpoint V and four points $P_i, i \in 1, \ldots, 4$, the example is in fact shown in Fig. 9.7. The points have the elevation:

$$Ve = (0.0 : 5; 20)(1.0 : 10; 10)$$
$$\tilde{P}_1 e = (0.0 : 15; 30)(1.0 : 20; 20)$$
$$\tilde{P}_2 e = (0.0 : 25; 40)(1.0 : 30; 30)$$
$$\tilde{P}_3 e = (0.0 : 40; 60)(1.0 : 52; 52)$$
$$\tilde{P}_4 e = (0.0 : 70; 100)(1.0 : 80; 80)$$

The distances from the viewpoint are:

$$P_1 d = 75$$
$$P_2 d = 125$$
$$P_3 d = 190$$
$$P_4 d = 240$$

All values can be measured in any distance units, because only the angles will be compared, but let's assume that they are in meters. Based on these values the angles can be calculated:

$$\tilde{P}_1\alpha = (0.0 : -3.814; 18.435)(1.0 : 11.31; 11.31)$$
$$\tilde{P}_2\alpha = (0.0 : 2.29; 15.642)(1.0 : 9.09; 9.09)$$
$$\tilde{P}_3\alpha = (0.0 : 6.009; 16.144)(1.0 : 12.465; 12.465)$$
$$\tilde{P}_4\alpha = (0.0 : 2.29; 15.642)(1.0 : 16.26; 16.26)$$

From the values of $\tilde{P}_1\alpha$, $\tilde{P}_2\alpha$ and $\tilde{P}_3\alpha$ the maximal angle is obtained:

$$max\tilde{A}ngle = (0.0 : 6.009; 18.435)(1.0 : 12.465; 12.465)$$

As visible from the description, the values of $\tilde{P}_1\alpha$ and $\tilde{P}_3\alpha$ contributed to the maximal angle. Now the value $\tilde{P}_4\alpha$ is compared to $max\tilde{A}ngle$ to find out what are the possibility and necessity values of exceedance and strict exceedance of $max\tilde{A}ngle$ by $\tilde{P}_4\alpha$. The results are as follows:

$$PSE = 1.0$$
$$NSE = 0.936$$
$$PS = 0.808$$
$$NS = 0.363$$

The outcomes show the chance of point P_4 being visible from the viewpoint V are relatively high because the value of possibility of exceedance is equal to 1 and the values of necessity of exceedance and the strict possibility of exceedance are both significantly high. However, the value of strict necessity of exceedance is only 0.363 which points to the fact that it is not necessary. Yet the chances that the point P_4 will be visible from the viewpoint V are quite high, even if uncertainty that may occur is taken into account.

9.4.1 Comparison with Monte Carlo

As a comparison the example shown in the previous section was calculated by the Monte Carlo method. The triangular probability distributions were used as they are specified by three values (Evans et al., 2000) and thus closely related to triangular fuzzy numbers. The experiment is done in the same manner as presented by Fisher (1992). The results are shown in Tab. 9.1.

From the results of the fuzzy visibility calculation we obtain the results of the crisp calculation if only the 1-cut of fuzzy numbers is taken into account, so the maximal angle from the points P_1, P_2 and P_3 is 12.465. If it is compared to the angle 16.26 of the point P_4, it is obvious that the point is visible. The same result can be estimated from Fig. 9.7, if only the peaks of fuzzy numbers are considered. The results of possibilistic visibility suggest that it is very likely that the point will be visible, but the value of the strict necessity of exceedance suggests that one cannot be definitely sure about visibility. The result nicely shows the contradictory facts in the uncertain information. The probabilistic results shown in Tab. 9.1 fail to do that, even with a high number of iterations the results show the point as almost definitely visible (probability over 99.8 %). This is caused by the fact that the combination of input values that render the point invisible is highly improbable, however it is absolutely feasible. This example illustrates the difference in approach to the uncertainty propagation problem by different methods.

Number of simulations	Probability of visibility (%)
20	100
100	100
1 000	99.8
10 000	99.85
1 000 000	99.8343

Table 9.1: The probability of visibility of the point P_4 using various number of the Monte Carlo iterations.

Chapter 10

Decision making with the results of fuzzy surface analyses

> *Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.*
>
> – John W. Tukey (Tukey, 1962)

Decision making based on the alternatives with uncertainty is rather important because uncertainty can significantly affect the selection of alternatives. Figure 10.1 shows the situation with two alternatives, the area of alternative describes its suitability for the user. The smaller the area is, the more suitable it is as the solution of the decision making process. If uncertainty is not considered, it is clear that Alternative 1 is better than Alternative 2. However, if uncertainty is taken into account, the problem is more complicated. Alternative 2 with uncertainty provides a better solution than Alternative 1 with uncertainty. The actual solution is dependent on the criteria for the selection that the decision maker will use. Will the solution with smaller risk but worse results (Alternative 2) be used? Or is the decision maker more inclined to use the better solution even if there is bigger risk involved (Alternative 1)? The actual outcome of such situation depends on the decision maker, their preference and the problem at hand. The example illustrates the need for the decision making with uncertain alternatives.

The usage of results based on the analysis of a fuzzy surface is not as straightforward as in the case of a classic crisp analysis. There are two types of results that such analysis can produce. The first case is a situation where the outcome of analysis is a fuzzy number as e.g. the calculations of slope and aspect on a fuzzy surface. Such analysis is a direct result of using fuzzy arithmetic. The second case is a situation where the result of the analysis is a classification into categories, e.g. the visibility analysis – the visible and invisible areas of the surface. The categories can be produced by the classification of the result, by some kind of comparison (e.g. visibility) or another technique. Regardless of the way the result were derived, their primary usage is intended to support the decision making. This is indeed the main aim of all geographical analyses.

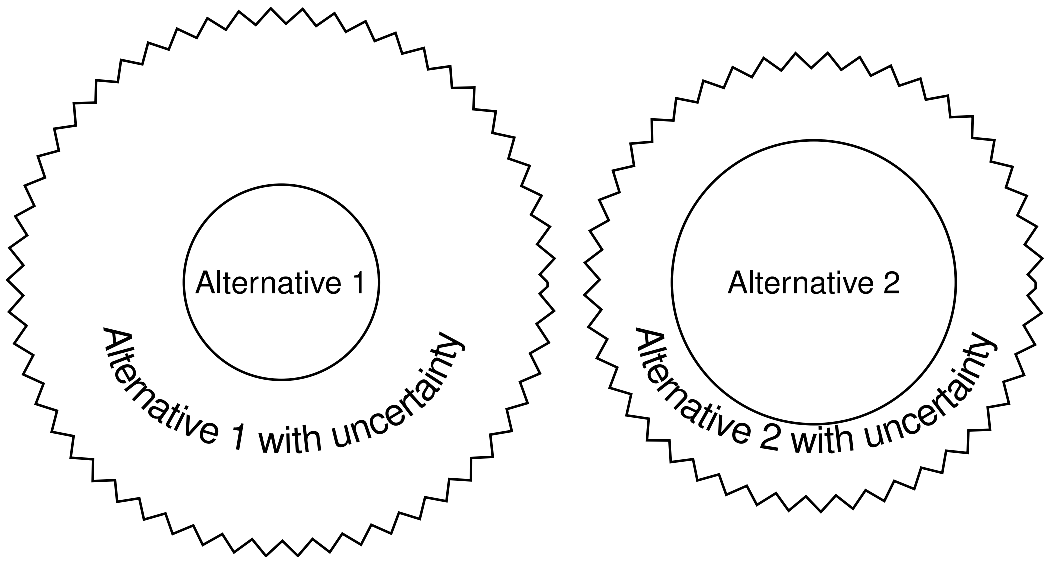


Figure 10.1: Uncertainty in decision making. The area of the alternative indicates its suitability – smaller is better.

Since the fuzzy surface and its derivatives contain uncertainty, the decision support done with these data cannot be exact (Bellman and Zadeh, 1970). To include this contradiction of data in the decision making process more than one measure is necessary. This is achieved by utilization of the possibility theory.

There has been an extensive research regarding the usage of fuzzy sets (de Bruin, 2000; Fonte and Lodwick, 2005; Herrera et al., 2006; Jiang and Eastman, 2000; Witlox and Derudder, 2005), the possibility theory (Caha and Dvorský, 2013a; Caha et al., 2014b,c) and other techniques (Li et al., 2007; Sozer et al., 2008) in geographical queries and spatial decision making. A big part of this research is focused on the introduction of vague, soft or linguistic terms instead of strict values (de Bruin, 2000). The smaller part of the research deals with the application of these vague queries on datasets containing uncertainty (Caha and Dvorský, 2013a; de Bruin, 2000; Fonte and Lodwick, 2005). While the topic is relatively new in geosciences it has been a subject of research for a longer period of time in mathematics and computer science (Bosc and Prade, 1997; Dubois et al., 1996; Dubois and Prade, 1986).

There are several reasons why a vague query might be used:

- the object that should be found in the data is naturally vague, e.g. "steep slope",
- more than one definition of the object exists with no indications about their correctness,
- the object definition is based on one or more expert's opinions that should be merged and/or are provided as vague definitions.

These findings are based on several studies (Bosc et al., 2005; Fonte and Lodwick, 2005; Witlox and Derudder, 2005). Because of these facts soft queries are needed to enrich the possibilities of classic boolean queries.

So far there has been very few attempts to acquire the useful information from a fuzzy surface or some derivatives based on a fuzzy surface. Fonte and Lodwick (2005) and Fonte et al. (2008a) showed how a fuzzy number can be compared to the quality threshold to obtain the fuzzy classification of the result into suitable and unsuitable categories. Caha and Dvorský (2013a) presented a way for querying a fuzzy surface with a vague query that utilized the possibility theory. A subsequent research (Caha et al., 2014b,c) showed that such queries, even when applied to a crisp surface, provide significantly more information to the decision maker and thus allow better decision making.

The classic GIS query can have this form: "Select pixels within this surface where the variable X is higher (or lower) then the threshold T". In case of a fuzzy surface the variable X is actually a fuzzy number $\tilde{X}$ and the threshold T can also be a fuzzy number $\tilde{T}$, representing the vague threshold of the term like "steep slope". Such threshold can be represented by a triangular fuzzy number (Caha et al., 2014b). Then the comparisons are done according to Eqs. (6.3,6.4,6.5,6.6) (Caha and Dvorský, 2013a). The comparison is done for each pixel of the fuzzy grid and produces four results.

The four indices form a set of nested solutions – the possibility of exceedance $<$ the necessity of exceedance, the possibility of strict exceedance $<$ the necessity of strict exceedance. The possibility of exceedance is the most loose, providing the user with estimation if there is at least some chance of $\tilde{X}$ being higher or equal to $\tilde{T}$. The necessity of strict exceedance evaluates how big the chances that $\tilde{X}$ si necessary bigger than $\tilde{T}$ are. The two remaining indices pair up with these to provide two pairs of measures, allowing to assess the possibility and necessity of at least equality of strict exceedance of $\tilde{T}$ by $\tilde{X}$. The details of this process and a practical example are provided by Caha and Dvorský (2013a).

Querying data from a fuzzy surface can be seen as vaguely linked to the results of indicator kriging (Hengl, 2009; Shad et al., 2009). Instead of providing the probability of a value at a location being higher or lower than the threshold, the possibilistic query (Caha et al., 2014c) provides the measures of possibility and necessity. In the same way a result of Monte Carlo can be evaluated based on the cumulative density function (CDF) (Rubinstein and Kroese, 2008). The results of fuzzy arithmetic are queried with the usage of possibility theory. Those two processes are likewise but not completely the same. However, drawing this parallel between them is helpful for understanding the concept.

Chapter 11

Case Study

The purpose of the case study in this thesis is merely to illustrate the utilization of fuzzy arithmetic in analyses of a fuzzy surface and to present the obtained results as well as their further usage in subsequent decision making processes. Since the case study serves mainly as an illustrative example, an artificial dataset will be used. Artificial data are also referred to as the synthetic data (Barse et al., 2003). Such approach is not uncommon in computer science, especially for testing of algorithms, and some examples can be found even in geosciences, like the research presented by Zhou and Liu (2004). The main reasons for the usage of artificial datasets as summarized by Barse et al. (2003) are: *a*) the need to demonstrate the specific characteristics that might be hard to find in real datasets, *b*) data might be hard or almost impossible to obtain (at least a sufficient amount of data), *c*) artificial data provide high degree of freedom for testing, *d*) the processes for the simulation of synthetic data are reproducible. These reasons indicate that the usage of synthetic data might be suitable for studies not focusing on the data but rather on algorithms and procedures for handling data.

In this thesis the main focus is on the description of algorithms for fuzzy surface analyses. It is useful but not absolutely necessary to demonstrate these algorithms on a practical example. Due to the fact that the example could help the reader with understanding the presented algorithms, a case study showing the calculation of a fuzzy slope, fuzzy aspect and the possibilistic visibility on a fuzzy surface will be presented. The case study is based on an artificially generated dataset and a process described by Loquin and Dubois (2010a) to create a fuzzy surface. All presented algorithms would work on any fuzzy surface that was created by any process mentioned in chapter 7.1. The artificial dataset is used because the fuzzy data (as described by Diamond (1989)) are not very common in geosciences, the usage of real crisp data would be possible for the approach described by Loquin and Dubois (2010b) and also shown by Caha et al. (2014a). Even though it would be possible to use real data as an example, utilization of synthetic data is not only easier, but it will also allow a demonstration of some characteristics that might be hard to show otherwise.

All data presented in this case study are available in an Appendix 1 of this thesis.

11.1 The elevation points

The input for surface modelling in geosciences is usually a set of points with known x, y and z coordinate. To generate the artificial data representing the points with known values x, y, z, the package *geoR*, specifically its function *grf*, was used (Ribeiro Jr and Diggle, 2001). An irregulated grid of 400 points with a gaussian correlation function having the sill 200, range 400 and nugget 0 was simulated. The mean value of z was set to 150 and the ranges of x and y were $[0, 4000]$. The units of all coordinates are not explicitly specified because for further usage they do not matter, but for the sake of clarity we will assume that they are defined in meters. The outcome of a gaussian random field simulation is a very smooth surface so to make it more rough, as real terrains are, a random value, drawn from the normal distribution with mean 0 and the standard deviation 4, was added to z value of each point.

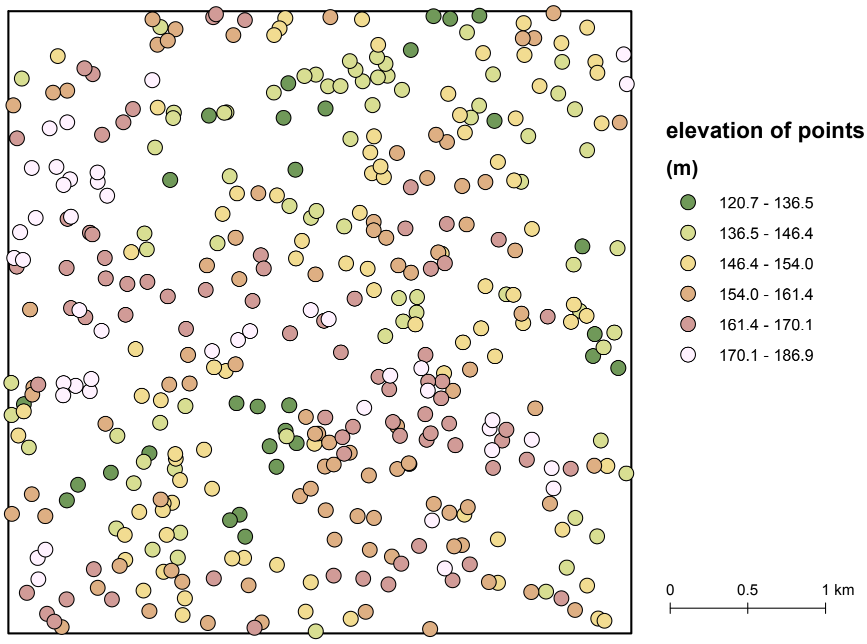


Figure 11.1: The points used in the case study example.

The result is a set of spatial points that generally follows the function mentioned previously with a small random component included. The visualization of this dataset is shown in Fig. 11.1. The dataset itself is attached in a form of csv file as an appendix *points.txt* and the R script used to generate the dataset as well as *case_study.R*.

11.2 The fuzzy surface creation

The previously created dataset obviously contains crisp data. To create a fuzzy surface from such dataset the assumption of existence of epistemic uncertainty is made. Epistemic uncertainty as described by Helton et al. (2004) and Helton and Oberkampf (2004) is a lack of knowledge about the fixed but poorly known parameters of the model. In the context of surface modelling, epistemic uncertainty affects the selection of parameters for the interpolation method. An approach for the creation of a fuzzy surface based on this fact was firstly described by Bardossy et al. (1990a,b) and later modified by Loquin and Dubois (2010a, 2011). In all these studies the authors mention that the selection of parameters for semivariogram (or variogram) used in the kriging interpolation method is dependent on the user and partially subjective. However, the selection of different parameters could lead to different results. To overcome this problem, the authors suggest to specify the parameters (sill, range and nugget) of semivariogram as fuzzy numbers, specifying their lowest, modal and highest possible values. Bardossy et al. (1990a) propose to specify three semivariograms that are used as limits for the calculation. Bardossy et al. (1990a) also mention that this approach is specially useful if the fit of an experimental variogram is more difficult. Such cases are rather common because real data rarely have “nice” experimental variograms allowing an easy fit of a theoretical variogram.

The process of fuzzy surface creation as proposed by Bardossy et al. (1990a) and Loquin and Dubois (2010a) is computationally significantly complex. To simplify the process, Loquin and Dubois (2010a) proposed an optimization scheme. The authors noted that extreme values of a fuzzy surface are usually produced by the combination of extreme values of the parameters (sill, range and nugget). As a result, it is not necessary to actually calculate kriging with fuzzy numbers (which would be very computationally and time demanding). It is enough to perform 8 (2^3 – two extreme values for each parameter) calculations of kriging and to select the highest and lowest value at each prediction point. The modal value of a fuzzy surface is obtained as a result of kriging with modal semivariogram. The resulting fuzzy surface (composed of triangular fuzzy numbers) might not be absolutely correct, deviations from a surface that would be calculated with true fuzzy variogram may occur (Loquin and Dubois, 2010a). It is suggested to check the results of this optimization procedure against the probabilistic metaheuristic method – the simulated annealing (Loquin and Dubois, 2010a). The optimization method was tested in order to determine its usefulness and it was determined that the results contain only a relatively small number of conflicts (Caha et al., 2014a).

The dataset shown in Fig. 11.1 is used to interpolate the fuzzy surface. Three semivariograms are defined for the dataset (Fig. 11.2) with the parameter values summarized in Tab. 11.1. The modal variogram was selected according to the *autofitVariogram* function of the *automap* package (Hiemstra et al., 2008) and the minimal and maximal semivariograms were selected based on the experts' opinion. Both semivariograms should be a limit (lower and upper) to the possible realizations of semivariogram. The fuzzy numbers forming a fuzzy surface have a triangular shape defined by these three values.

The resulting fuzzy surface was tested according to the methodology shown in Caha et al. (2014a) against 1000 simulated surfaces produced by the simulated annealing. In the worst case the ratio of the values that did not fit within the limits produced by the optimization scheme was 0.0075. It is a very good result from which a suggestion that a fuzzy surface produced by the optimization method is suitable for practical usage can be concluded.

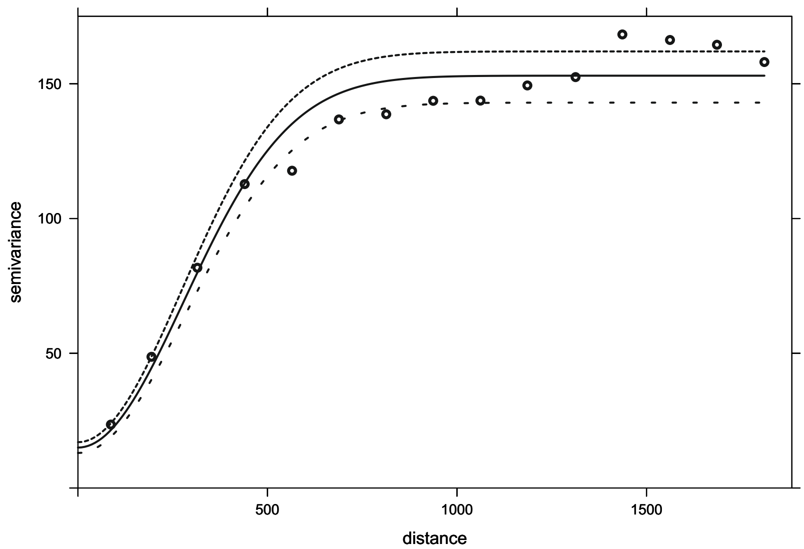


Figure 11.2: Three semivariograms used to create a fuzzy surface. The modal (full line), maximal (densely dashed) and minimal (loosely dashed) semivariogram.

A fuzzy semivariogram, as shown in Fig. 11.2, contains all semivariograms that could be formed from the combinations of parameters from Tab. 11.1. It helps the user to construct the variogram even in cases when the user is not absolutely sure about the values of parameters. The practical example as shown in this chapter might not be the best example in terms of the necessity to use the fuzzy interpolation but it still serves the purpose rather well.

Parameter	Minimal value	Modal value	Maximal value
sill	130	138	145
range	390	395	400
nugget	13	15	17

Table 11.1: The values of semivariogram parameters.

The resulting fuzzy surface is visualized as a modal value in Fig. 11.3 and the negative and positive differences from the modal value (Fig. 11.4). From the visualizations it can be seen that the surface is very smooth, which is a result of the artificial creation. The input data, even with a random component, still follow the functional definition very closely. However, this fact does not influence the analyses of a surface. The deviations from the modal value have a maximal absolute value of 2.35 which, for the grid with the cell size of 10 meters, is a reasonable value. Such magnitude will show the differences in the minimal, modal and maximal values of the outcomes sufficiently while it remains very reasonable. From the visualization in Fig. 11.4 it is visible that majority of area has relatively small uncertainty and areas with higher uncertainty (positive or negative) are clustered together. These clusters of higher uncertainty values are the places where different semivariograms captured different trends in the data, this effect is well described by Loquin and Dubois (2010b).

The calculations necessary for the creation of this fuzzy surface can be found in the appendix *case_study.R* and the three surfaces in form of an ASCII grid as another appendix – folder *2_fuzzy_surface*.

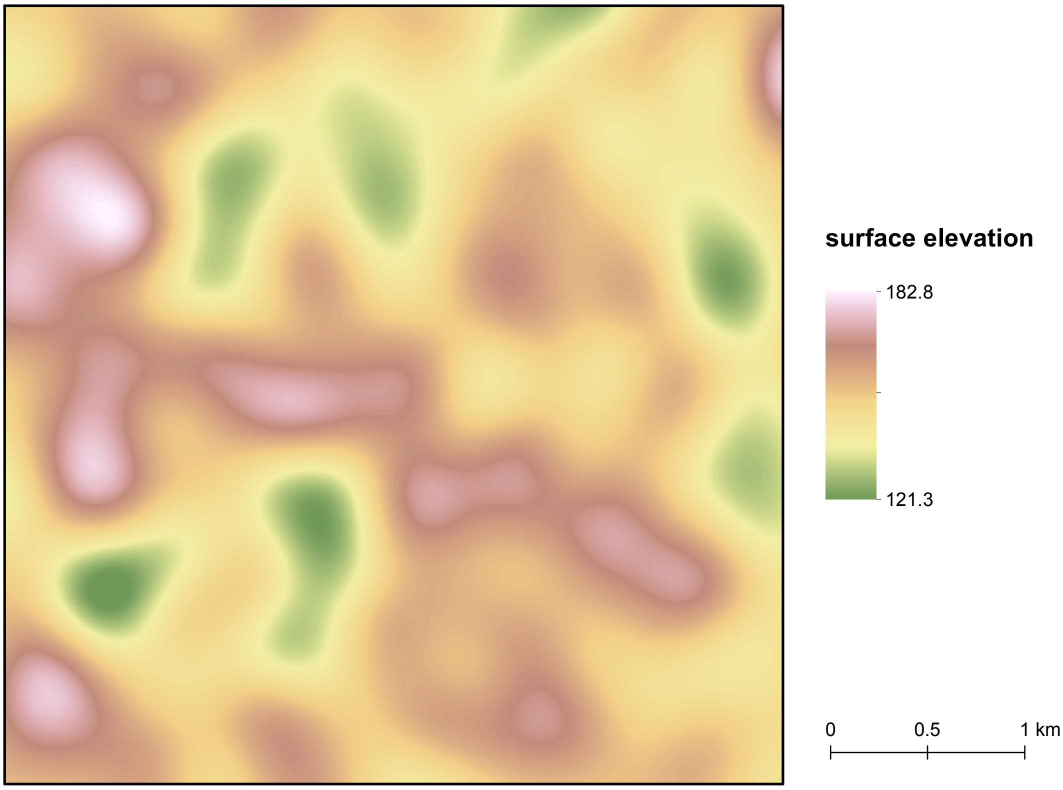


Figure 11.3: The modal value of a fuzzy surface.

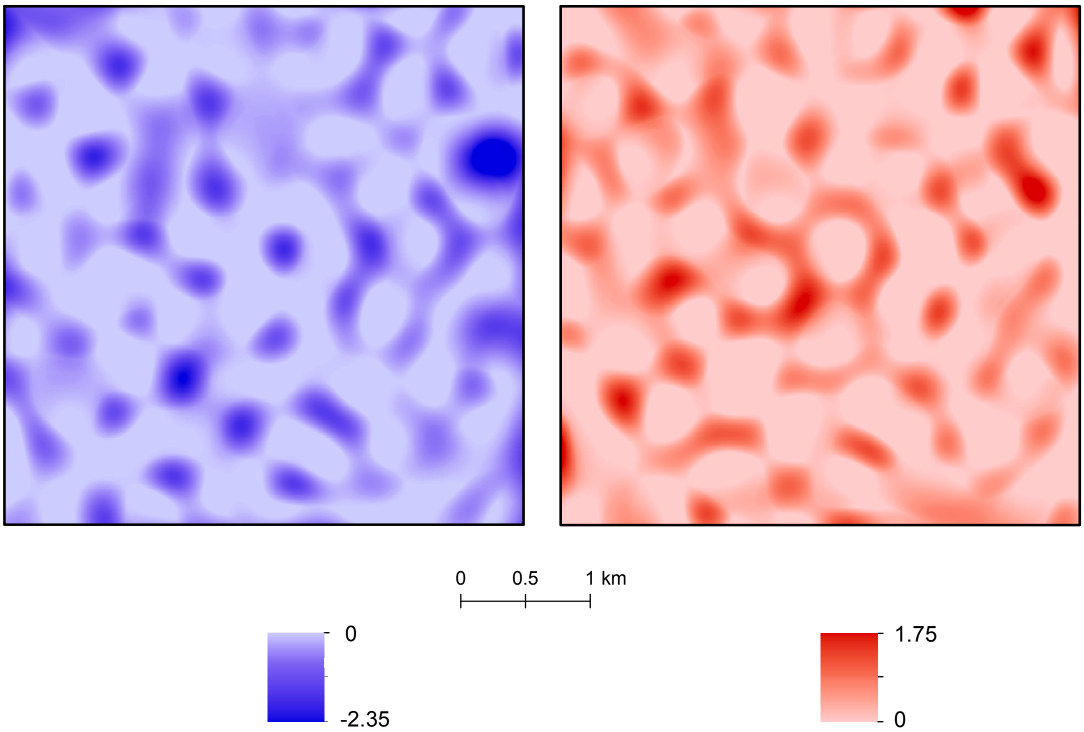


Figure 11.4: The differences of a fuzzy surface from the modal value. The difference between the minimal and modal value (left) and the maximal and modal value (right).

11.3 The first derivatives of a fuzzy surface

The calculation of first derivatives of a fuzzy surface is described in chapter 8.3. The fuzzy surface that was created in the previous chapter will be analysed and an example of how it can be used in further decision making will be shown. In case of both derivatives Horn's method (Eq. 8.4) was used.

The modal values of the resulting fuzzy slope and fuzzy aspect should be equal to the crisp calculation of slope and aspect values on a crisp surface (in this case the modal value of a fuzzy surface). The modal values of a fuzzy surface were compared to the ArcGIS implementation of slope calculation. Small deviations were found that were most likely caused by different handling of rounding of the numbers. But it is impossible to identify the problem exactly because the implementation used in ArcGIS is nowhere described in detail. The issue of non-public algorithms that do not allow any comparison was raised before by Fisher (1993, 2007b). In this study the differences between the modal value of a fuzzy slope were in the interval $[-0.113085, 0.30312]$ but the mean value was 0.00036 and the standard deviation 0.01127. Generally, these differences are not significant. In case of aspect, the modal value of a fuzzy aspect was in almost all cases (over 99 %) smaller than the value calculated by the ArcGIS implementation by the value 1.1571. Again, this is not an important difference but it raises an interesting question about the implementations used in ArcGIS. As noted by Fisher (1993), the comparisons of algorithms are not a common field for research in geosciences but the occurrence of differences should at least raise a question if the implementations are actually done precisely according to the definitions.

11.3.1 The fuzzy slope

The fuzzy slope was calculated using Horn's method (chapter 8.1.2) for calculating partial derivatives. The modal values of the fuzzy surface are shown in Fig. 11.5.

Figure 11.6 shows the minimal and maximal values of the fuzzy slope. When compared with each other, the influence of uncertainty in the calculation is rather obvious. It is visible that the lower limit of the slope is 0 % for a significant area of the surface. The maximal values obtained from the fuzzy slope show that with accounted uncertainty the maximal slope on the surface might be slightly higher than 22 % which is a significant rise from the maximal slope amongst the modal values which is 15 %.

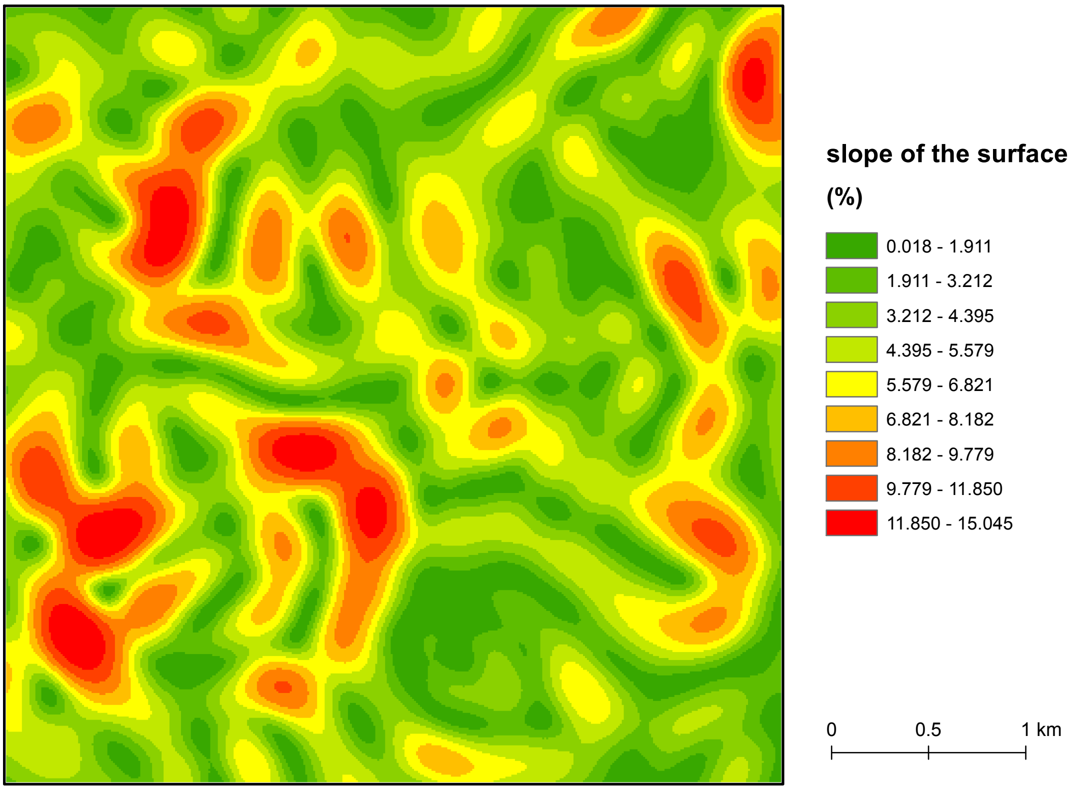


Figure 11.5: The modal value of the fuzzy slope.

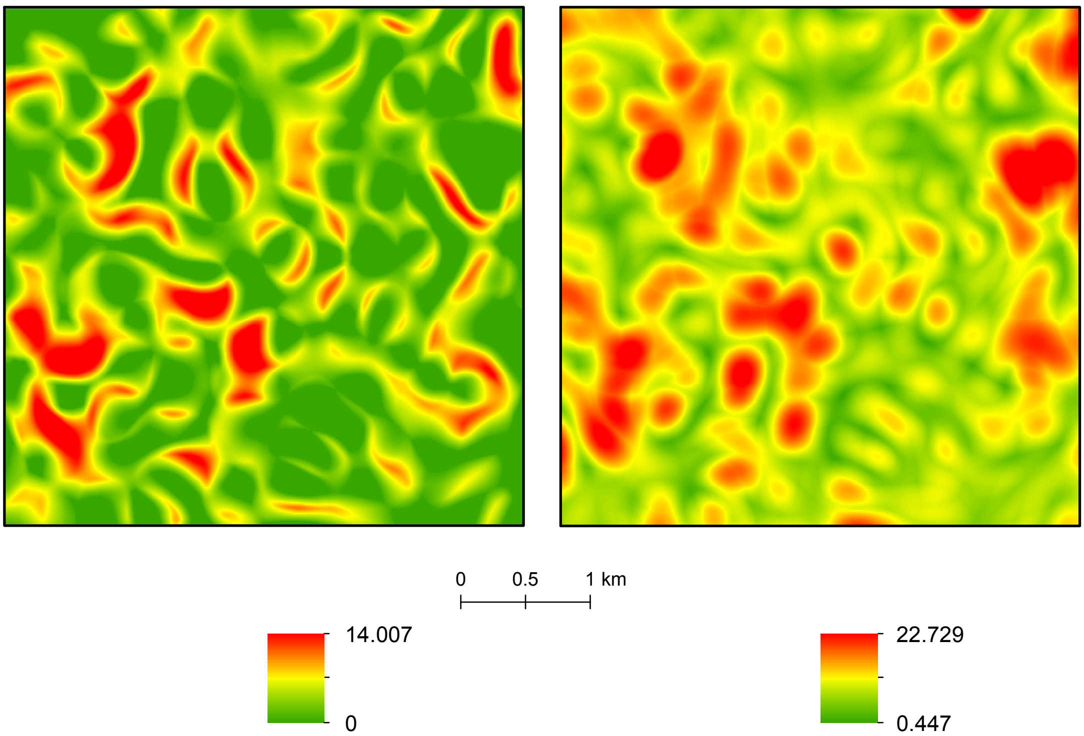


Figure 11.6: The minimal (left) and maximal (right) value of the fuzzy slope (in percent).

11.3.2 The fuzzy aspect

The process of calculation of aspect of the fuzzy surface is described in chapter 8.3.2. Again, Horn's method was used to calculate the partial derivatives of the surface. The modal value of fuzzy aspect is shown in Fig. 11.7, visualized with the usage of a standard ArcGIS palette for aspect values.

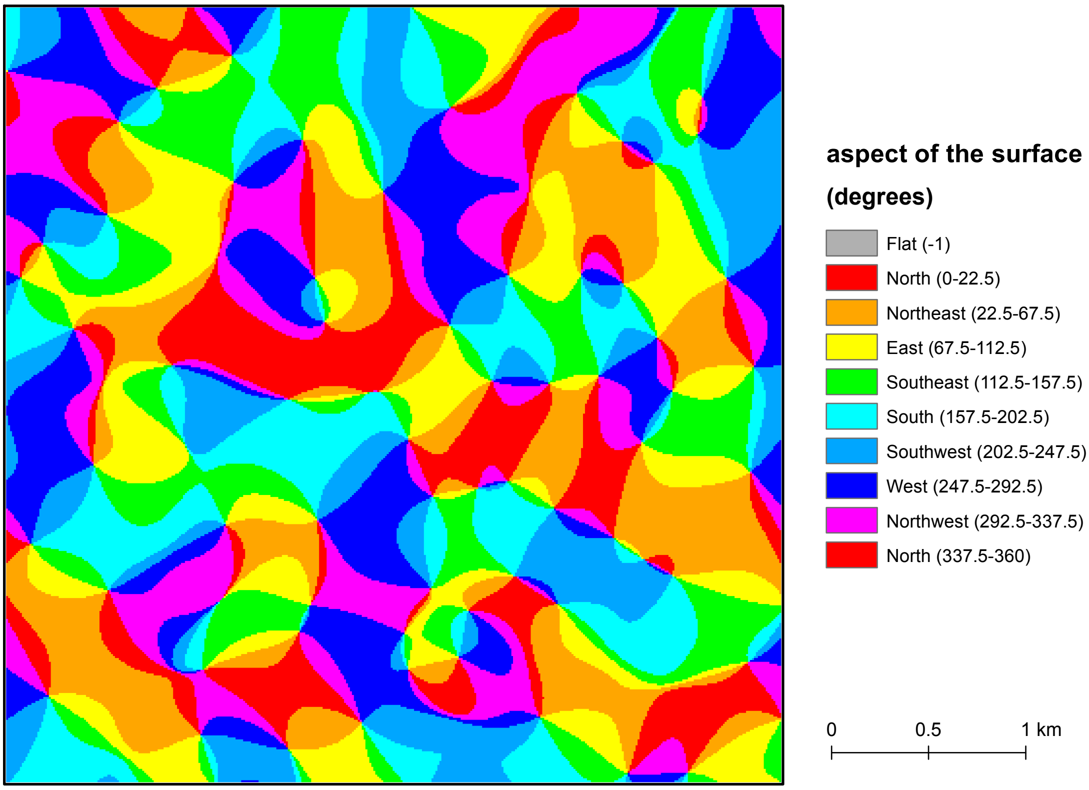


Figure 11.7: The modal value of the fuzzy aspect.

As mentioned in chapters 8.3.2 and 5.5.2, the calculation of aspect is more complex due to the usage of the atan2 function. It was mentioned previously that the range of values for atan2 with fuzzy arguments is $[-90°, 630°]$. However, when interpreting, the value x smaller than zero has the same interpretation as the value $x + 360$ and the value x higher than 360 has the same interpretation as $x - 360$. The interpreted results with the assigned classes are shown in Fig. 11.8 and the true values are shown in Fig. 11.9 with the values outside the range $[0°, 360°]$ highlighted with the hatch. The example shown in Fig. 11.8 illustrates how uncertainty of the surface affects the aspect calculation.

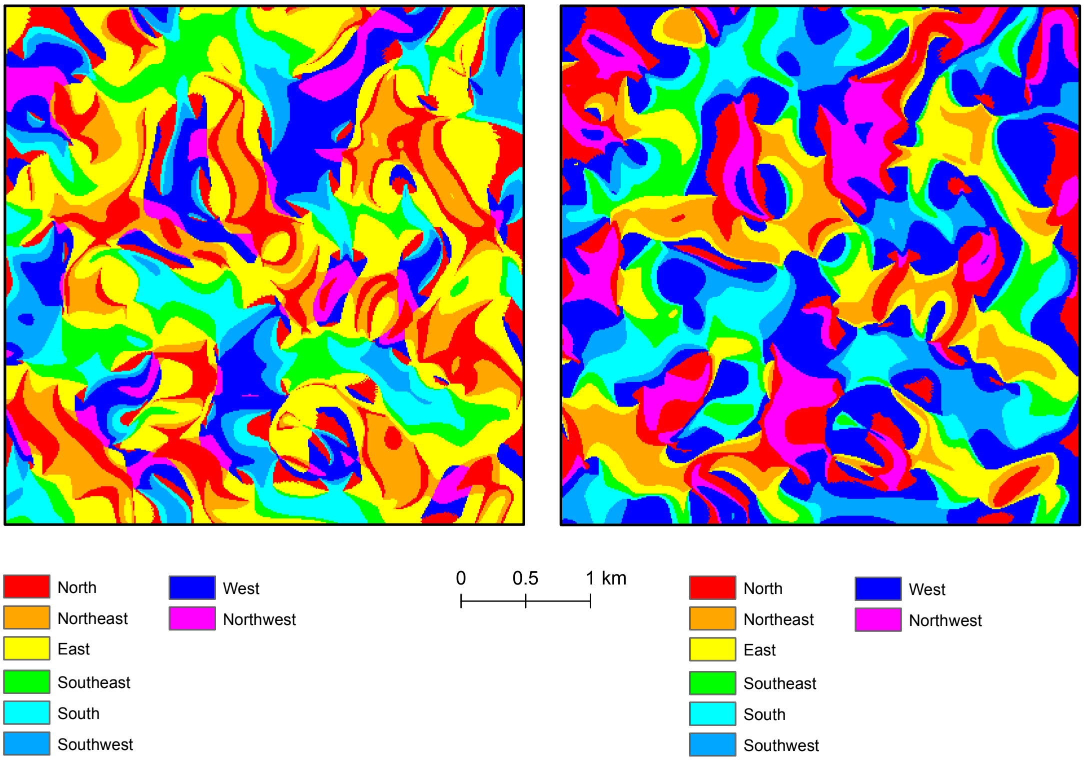


Figure 11.8: The minimal (left) and maximal (right) value of the fuzzy aspect. Shown as an orientation.

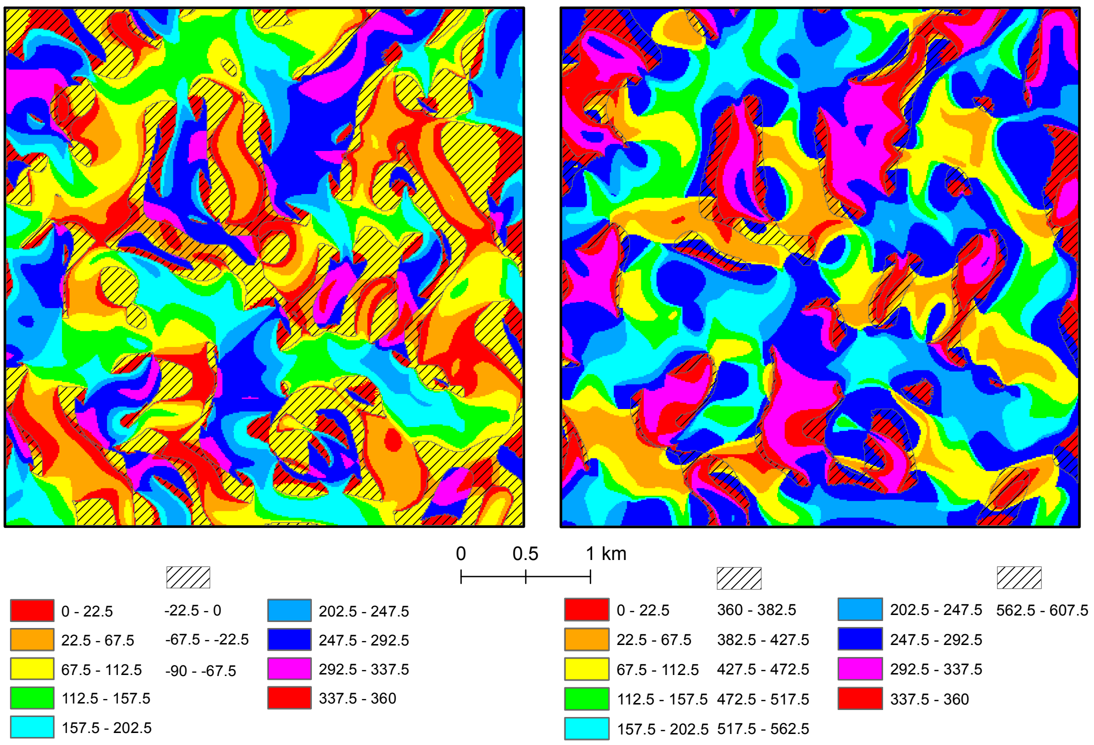


Figure 11.9: The minimal (left) and maximal (right) value of the fuzzy aspect. With the values outside the range of the crisp aspect highlighted with the hatch.

11.3.3 Querying the results of a fuzzy analysis

The issue of further usage of results of the fuzzy surface analysis in the decision support is even less developed than the issues of analyses of a fuzzy surface. Several studies showed how fuzzy surfaces and their derivatives can be queried in order to obtain useful data for the decision support (Caha and Dvorský, 2013a; Fisher and Caha, 2014). Besides that, a significant research has been performed regarding the usage of fuzzy numbers to create the possibilistic queries (Caha and Dvorský, 2013a; Caha et al., 2014c,b) but this topic is not directly relevant to this thesis. The theory regarding the decision making with the results of a fuzzy analysis is summarized in chapter 10.

The research presented by Fisher and Caha (2014) showed how a fuzzy surface can be compared to the crisp threshold. The outcome of such comparison are the measures of possibility, showing the areas where the uncertain values might exceed the threshold, and necessity, showing the areas where the uncertain values necessarily exceed the threshold. The example is shown in Fig. 11.10. Obviously, the fuzzy number $\tilde{B}$ exceeds the threshold T more than $\tilde{A}$ but even $\tilde{A}$ partly exceeds the threshold. To describe all these differences and complex situations that might occur, there is a need for more than just one measure of exceedance.

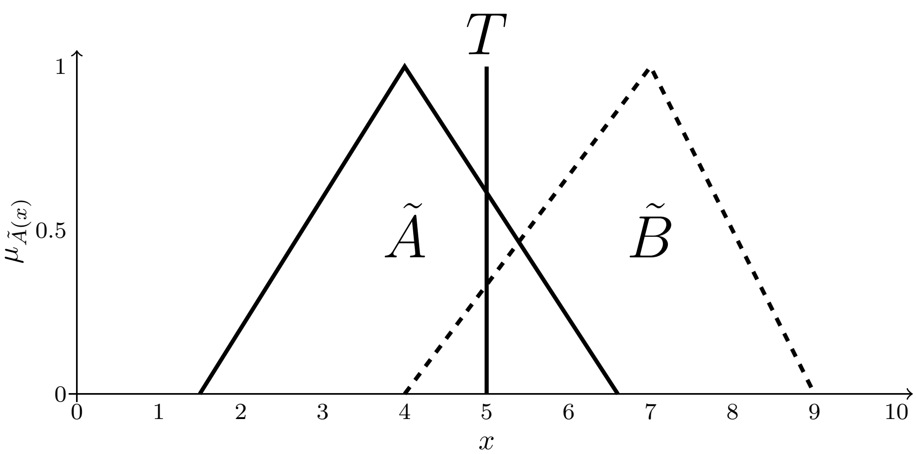


Figure 11.10: An example of the fuzzy number $\tilde{A}$ possibly exceeding the threshold T (possibility 0.62, necessity 0) and the fuzzy number $\tilde{B}$ necessary exceeding the threshold T (possibility 1, necessity 0.67).

To extend the approach to the spatial data represented by a grid, the comparison needs to be done for each cell of the fuzzy surface or the result of a fuzzy analysis. As a result, the measures of possibility and necessity are obtained (Fig. 11.11). It is obvious from the definitions of the possibility theory (Dubois and Prade, 1986) that the necessity outcome will be a subset of the possibility outcome because what is necessary must be possible. The possibility in this case identifies the areas where the fuzzy numbers exceed the threshold by even a small part while necessity identifies fuzzy numbers that exceed the threshold with their peak value (Fisher and Caha, 2014). The necessity value 1 is obtained only if all values of a fuzzy number are higher than the threshold.

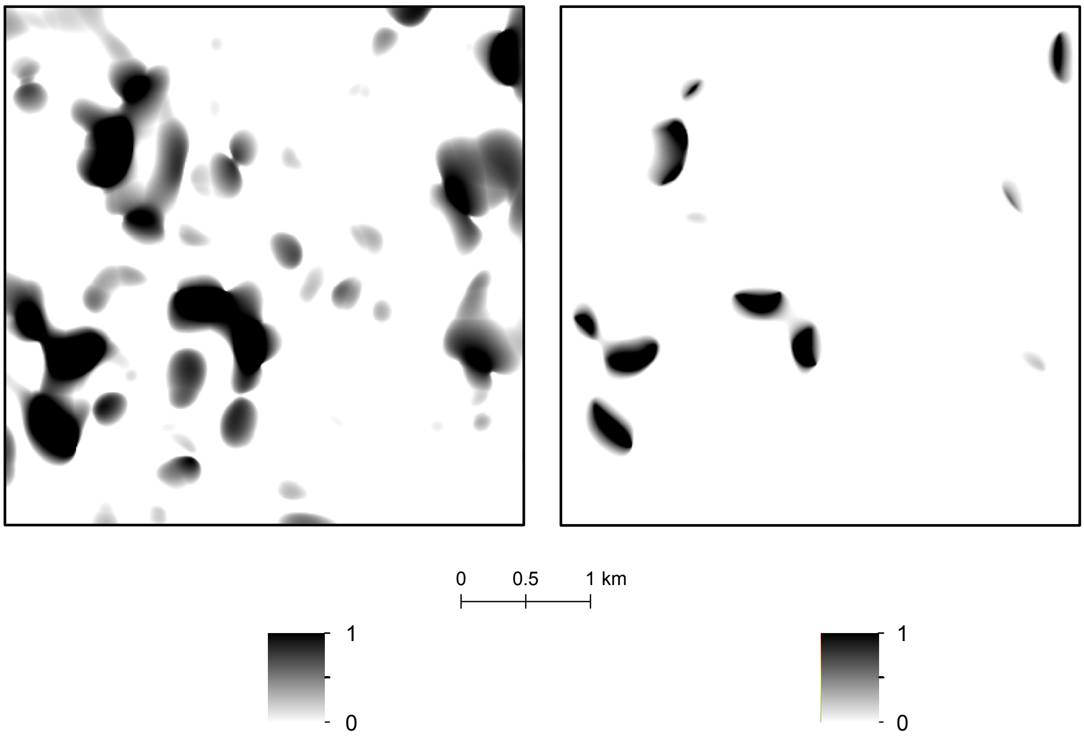


Figure 11.11: A comparison of the fuzzy slope calculated in the previous chapter to the crisp threshold 10 % to identify the areas with possibly (left) and necessary (right) higher slope.

The example presented here shows the classification of the fuzzy slope into the category of slope higher than 10 % (Fig. 11.11). As visible from the image, the area possibly exceeding this value is significantly larger than the area where the values are necessarily higher. The range of values $[0, 1]$ for each measure also provides more information than a classic crisp classification.

The outcomes of fuzzy analyses can also be queried by the vague queries as shown by Caha and Dvorský (2013a) and also used in more complex decision making problems than just a simple selection of an area that fulfils one criterion (Caha et al., 2014b). However, such utilization of the results is beyond the scope of this thesis.

11.4 The visibility analysis

This chapter shows an example of the possibilistic visibility calculated on a fuzzy surface. To establish a baseline for comparisons, the boolean visibility was calculated on a crisp surface (the modal value of a fuzzy surface). The observer has an offset of 1.8 meters from the surface. The offset can also be a fuzzy number, for example triangular defined as $[1.4, 1.8, 2.1]$ that would model an unknown height of the observer. In this case study the viewpoint was placed in a ridge in a relatively high part of the surface. As a result, the visible areas are stretched in the northeast and southwest direction from the viewpoint (Fig. 11.12). The crisp case visibility calculated in ArcGIS with the viewshed function is visualized in Fig. 11.12.

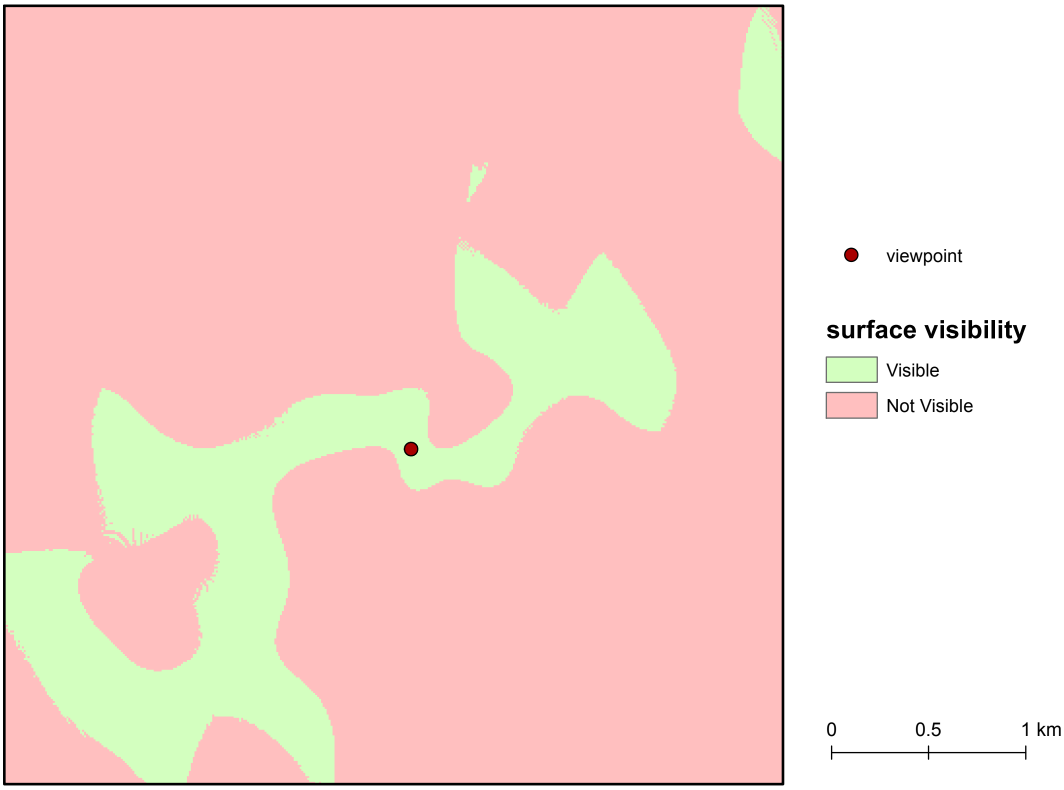


Figure 11.12: The visibility from the viewpoint (1.8 meter above the surface) as calculated by the viewshed operation in ArcGIS. The crisp surface is the modal value of the fuzzy surface.

The theory and algorithm for the calculation of the possibilistic visibility is summarized in chapter 9.3. The possibilistic visibility was calculated on a fuzzy surface with the same viewpoint as in the crisp case, as described in chapter 9.3.1 there are four indices that can be obtained as a result from the possibilistic visibility. The results are shown in Fig. 11.13 and 11.14. In all cases the value of 1 denotes visibility and 0 denotes invisibility of an area for the specific index.

From the results several interesting observations can be done. The four results form a set of nested solutions that can be lined up from the most optimistic (the possibility of visibility), through the strict possibility of visibility, the necessity of visibility and finally the most pessimistic index which is the strict necessity of visibility.

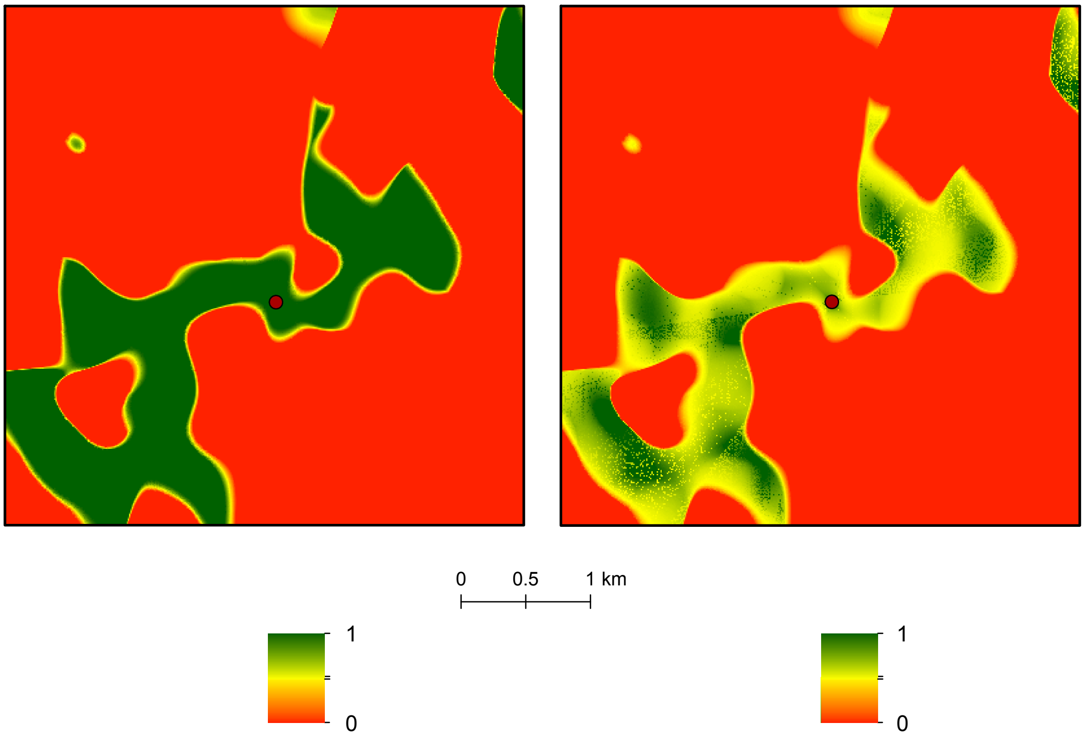


Figure 11.13: The possibilistic visibility from the viewpoint (1.8 meter above the surface). The possibility (left) and necessity (right) of visibility.

The solution shown in Fig. 11.13 on the left side shows the most optimistic realization of visibility when the whole uncertainty works in favour of the observer. As a result, the visible area is bigger than in the case of crisp visibility and the edges of this area are much smoother. The opposite situation is shown on the right side in Fig. 11.14. The strict necessity of visibility shows the situation when the whole uncertainty works against the observer, resulting in a very small visible area (the value of 1). A slightly bigger part of the surface is visible with smaller membership values but these cannot be considered as definitely visible. From these four indices the strict possibility of visibility is the one that is the most close to the crisp visibility in terms of a visible area. Obviously, the values differ because the crisp calculation is strictly boolean while possibility has gradual values of visibility.

The presented case study of the possibilistic visibility on the fuzzy surface shows how this approach can provide more information for the user about the possible

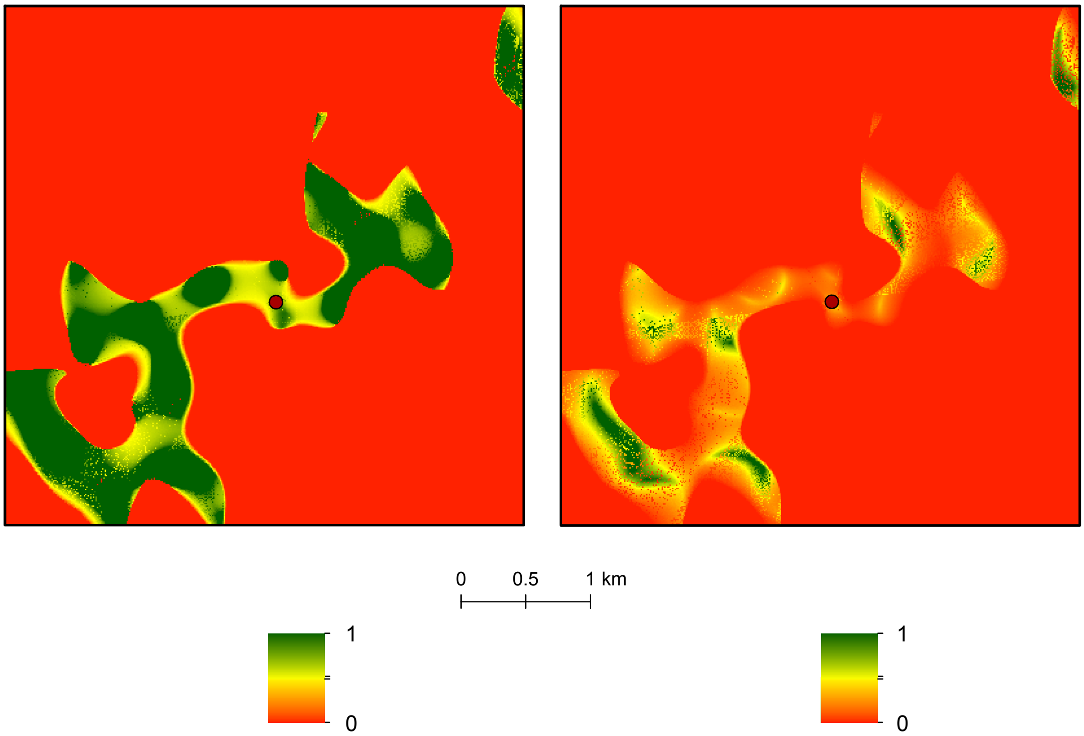


Figure 11.14: The possibilistic visibility from the viewpoint (1.8 meter above the surface). The strict possibility (left) and strict necessity (right) of visibility.

visibility. The approach accounts for uncertainty of the surface and provides the user four indices describing the chances of visibility much better than a classic boolean expression. Two of these indices, possibility and strict necessity of visibility, describe the extreme situations. In reality it is unlikely that the whole uncertainty would be either in favour or against the observer. The true solution will be somewhere between these extreme solutions. However, the ability to obtain these extreme solutions is very helpful for the decision making process.

The presented solution has a potential to impact the way how analyses are perceived in geosciences. Usually, the user reasons about the result as being precise, but providing them with the upper and lower limits of the solution could help them realize that the analysis is not as precise as they might think. This does not apply only for the case of visibility analysis but for many other analyses as well as e.g. the catchment delineation.

Chapter 12

Discussion

Anyone who has never made a mistake has never tried anything new.

– Albert Einstein

a large part of this thesis covers the theoretical foundations of fuzzy arithmetic and the topics closely associated to it that are later used to calculate the analyses of fuzzy surfaces. In his review of the book *Fuzzy Surfaces in GIS and Geographical Analysis: Theory, Analytical Methods, Algorithms, and Applications, edited by Weldon Lodwick, (Boca Raton, FL: CRC Press, 2008)* (Mount et al., 2009) Fisher noted that most of the readers from the GISc community are unlikely to be familiar with the concept of fuzzy arithmetic. Thus, a large introduction (Lodwick et al., 2008) to the topic is necessary. He also noted that the book is more dense with formulæ than is usual for books in the field of geoinformatics. The mentioned book was the reason why the author of this thesis got interested in fuzzy arithmetic and fuzzy surfaces, so it is not surprising that this thesis is also dense with formulæ and algorithms.

The most important part of the thesis are the methods of calculation described in chapters 8 and 9 and presented in practical usage in the case study in chapter 11. As described in these chapters, the approach to the usage of fuzzy arithmetic varies significantly when compared to the Monte Carlo method that is the most commonly used. As a consequence, fuzzy arithmetic has a potential to become an important method for the uncertainty propagation not only in surface analyses but in other applications as well. It was shown that network algorithms are affected by uncertainty as well (Caha and Dvorský, 2014). Many other geographical analyses would also benefit from the utilization of fuzzy arithmetic as a method for the uncertainty propagation.

Whenever a method using fuzzy sets is used, the question why statistics was not used to handle the problem is very likely to be raised. Zadeh (2005) describes this issue and tries to challenge its logical consequences. Several examples were provided not only in this thesis but also in work by Hanss (2005) and Lodwick et al. (2008) showing that the usage of the Monte Carlo method and fuzzy arithmetic does not provide the same results. These two methods (and also others like interval arithmetic) are not competitive but rather complementary, with each one focused on a different part of uncertainty. This fact is nicely illustrated in Fig. 2.1 (page 16) and explained in the work of Viertl (2011). However, in many fields of science the uncertainty propagation is very closely associated with the Monte Carlo method or in some cases with the analytical approach. This is significantly noticeable in the GISc literature (Heuvelink, 1998; Oksanen and Sarjakoski, 2005a; Shi, 2010; Zhang and Goodchild, 2002). That is not a problem as long as the field is open to new approaches and methods. Unfortunately, sometimes the approach to new methods and solutions is considerably negative (Zadeh, 2005). However, these new methods should be tested, compared to the existing approaches and their potential value for the field should be properly evaluated before they are rejected.

The introduction of fuzzy arithmetic into the analyses of spatial data poses interesting challenges and may offer some attractive solutions to the issues that so far have not been solved. These issues do not concern only the analyses of fuzzy surfaces but also almost all possible analyses because the fuzziness can be observed in almost any geographical dataset.

12.1 The implementation of fuzzy arithmetic

To achieve the aims and goals of this thesis it was necessary to implement some operations of fuzzy arithmetic. The implementation covered the representation of fuzzy numbers, basic arithmetic operations $(+, -, \times, /)$, several functions of fuzzy numbers (all of them monotonic – arctan, atan2, powers and square root of fuzzy numbers) and also the comparison indices (chapter 6.2). The implementation was done in Java programming language. It is based on a theory provided by Kaufmann and Gupta (1985), Hanss (2005), Klimke (2006), Lodwick et al. (2008), Dubois and Prade (1986) and heavily influenced by the code examples provided by Anile et al. (1995), Fonte et al. (2008b), Spinella (2008), Dubois and Prade (1986) and Gagolewski (2014). The origin of the library was also partially influenced also by Java fuzzy logic library jFuzzyLogic (Cingolani and Alcalá-Fdez, 2012; Cingolani and Alcalá-Fdez, 2013). This implementation was used in all author's articles. The description of the library is not included in the thesis because it is beyond the scope of the thesis.

The development of computers allows us to perform operations that would not be possible fifteen or twenty years ago due to time or memory demands. However, current computers (even laptops) have enough memory and computational power to perform calculations with fuzzy arithmetic even on a grid of the size of 400×400 cells. The calculation of slope and aspect is relatively fast, lasting less than a few seconds. The calculation of visibility is another matter that can last on such grid up to two hours but this is probably caused by the implementations of algorithms that were not optimized in any way. With optimization, the amount of time needed for the calculation of the possibilistic visibility could most likely be lowered or the algorithm could be parallelized. The visibility algorithm is a classic example of algorithms that can benefit from parallelization. This could also provide several interesting research topics mainly for the field of computer science.

12.2 Future work

This thesis presents the first examples of fuzzy surface analyses with fuzzy arithmetic. The results of this work showed that fuzzy arithmetic does have a possible usage in geoinformatics and that the results can provide rather interesting data for further studies. The consequent work could focus in several directions. The most obvious one is to focus on another analyses and describe their fuzzy equivalents. The interesting examples could be the second derivatives (curvatures) of surface, the catchment delineation or the optimal path identification on fuzzy surfaces. Besides that, there is a lot of space for the calculation of real world practical examples of a fuzzy slope, aspect and the possibilistic visibility. These outcomes should be thoroughly compared to the outcomes of other methods of the uncertainty propagation to explain the differences amongst these methods and to highlight the benefits of the usage of fuzzy arithmetic. Also the topic of visualization of fuzzy surfaces is largely unexplored (Vondráková and Caha, 2014). The outcomes of fuzzy analyses provide sets of possible results that are different from any other geographical datasets. These datasets should be properly visualized in order to provide the user or the decision maker with as much information as possible.

From the previous paragraph it is evident that there are many possible further utilizations of fuzzy arithmetic in geosciences. As a relatively new method it provides many research questions and other opportunities for a consequent research.

Chapter 13

Conclusions

The thesis is focused on the presentation of fuzzy arithmetic as a method for the uncertainty propagation in fuzzy surface analyses. To achieve this main goal several minor goals needed to be fulfilled.

The issue of modelling uncertainty from the mathematical point of view was summarized in chapter 2 in order to put a perspective on why there are several theories for modelling uncertainty, what are the differences amongst these theories and why the selection of theory matters. A simple example illustrating semantics and numerics differences amongst three most common theories of uncertainty was presented in chapter 2.5. This example should assist with understanding the differences in the meaning of uncertainty and also the differences of the obtained results.

Chapters 3, 4, 5 and 6 summarize the necessary theoretical foundations of the fuzzy set theory, fuzzy arithmetic and additional mathematical procedures for the usage in fuzzy surface analyses. These chapters provide a description of methods that are applied later. Figure 7.3 on page 57 shows that an analysis or operation with data needs to be modified in order to allow processing of fuzzy sets or fuzzy numbers. These modifications are based on the theoretical foundations provided in the above mentioned chapters.

The current state of the topic of surface uncertainty modelling in geosciences is summarized in chapter 7. This chapter provides an overview of methods and approaches (based on both statistics and the fuzzy set theory) that are used for modelling uncertainty of geographical surfaces. A special focus is put on the development of fuzzy surfaces and the rationale behind this approach as this topic is not sufficiently known as the statistical handling of surface uncertainty. Chapter 7.3 provides a comparison of the Monte Carlo method and fuzzy arithmetic with respect to the uncertainty propagation in surface analyses.

The most important parts of the thesis are chapters 8 and 9 as they explain how fuzzy equivalents of the first derivatives of surface and visibility on a fuzzy surface can be calculated. Some examples of the fuzzy slope calculations were presented in literature prior to this thesis (Fonte and Lodwick, 2005; Waelder, 2007) but in both cases they are relatively simple examples serving primarily as illustrative examples. In this thesis the focus was put on explaining the process of calculation in detail and it also provides methods for calculating the aspect of a fuzzy surface. The calculation of the fuzzy aspect was a challenge due to the necessity to calculate the atan2 functions with fuzzy arguments. The process is described in chapters 5.5.2 and 8.3.2. The calculation of the possibilistic visibility on a fuzzy surface as described in chapter 9 is largely a follow up to the research done in the early 1990s by Peter Fisher (Fisher, 1991, 1992, 1993, 1994). However, the approach presented in this thesis varies significantly because the mentioned research utilized the probabilistic approach to the surface uncertainty while in this thesis the approach completely based on fuzzy arithmetic and the possibility theory is used. A brief preview of how the results of fuzzy analyses can be used for further decision making is provided in chapter 10. This part of the research is based on author's publications dealing with querying the fuzzy data and the usage of soft queries in geosciences (Caha and Dvorský, 2013a; Caha et al., 2014c,b; Fisher and Caha, 2014).

A case study showing the calculation of the fuzzy slope, aspect and the possibilistic visibility is presented in chapter 11. The examples show the results obtained from the fuzzy surface analyses and imply how these results can be further used.

The main aim of the thesis as well as all the minor aims were reached and successfully completed. The applicability of fuzzy arithmetic for surface analyses was presented and the utilization of the result described.

Bibliography

ANILE, A. M., SPINELLA, S. Modeling Uncertain Sparse Data with Fuzzy B-splines. *Reliable Computing*, 10, 5, s. 335–355, 2004.

ANILE, A. M., DEODATO, S., PRIVITERA, G. Implementing Fuzzy Arithmetic. *Fuzzy Sets and Systems*, 72, 2, s. 239–250, 1995.

ANILE, A. M., FURNO, P., GALLO, G., MASSOLO, A. A fuzzy approach to visibility maps creation over digital terrains. *Fuzzy Sets and Systems*, 135, 1, s. 63–80, 2003.

ARNDT, J., HAENEL, C. *Pi Unleashed*. Springer-Verlag, 2006.

AYYUB, B. M., KLIR, G. J. *Uncertainty Modeling and Analysis in Engineering and the Sciences*. Boca Raton : Chapman & Hall/CRC, 2006.

BAEKELAND, R., KERRE, E. E. Piecewise linear fuzzy quantities : A way to implement fuzzy information into expert systems and fuzzy databases. In: BOUCHON, B., SAITTA, L., YAGER, R. R. (Eds.), *Uncertainty and Intelligent Systems SE - 16*, 313 / *Lecture Notes in Computer Science*, s. 119–126. Springer Berlin Heidelberg, 1988.

BANDEMER, H. *Mathematics of uncertainty : ideas, methods, application problems*. Č. 189. Berlin ; New York : Springer, 2006. ISBN 3540284575.

BANDEMER, H., GEBHARDT, A. Bayesian fuzzy kriging. *Fuzzy Sets and Systems*, 112, 3, s. 405–418, 2000.

BARDOSSY, A., BOGARDI, I., KELLY, W. E. Kriging with imprecise (fuzzy) variograms. I: Theory. *Mathematical Geology*, 22, 1, s. 63–79, January 1990a.

BARDOSSY, A., BOGARDI, I., KELLY, W. E. Kriging with imprecise (fuzzy) variograms. II: Application. *Mathematical Geology*, 22, 1, s. 81–94, 1990b.

BARSE, E., KVARNSTROM, H., JOHNSON, E. Synthesizing test data for fraud detection systems. In: *19th Annual Computer Security Applications Conference, 2003. Proceedings.*, s. 384–394. IEEE, 2003.

BELLMAN, R. E., ZADEH, L. A. Decision Making in a Fuzzy Environment. *Management Science*, 17, 4, s. 141–164, 1970.

BLAIR, B. Interview: Lotfi Zadeh and fuzzy logic theory. *Azerbaijan international*, 2, 4, s. 46–54, 1994.

BORTOLAN, G., DEGANI, R. A review of some methods for ranking fuzzy numbers. *Fuzzy Sets and Systems*, 15, s. 1–19, 1985.

BOSC, P., PRADE, H. An introduction to the fuzzy set and possibility theory-based treatment of soft queries and uncertain of imprecise databases. In: SMETS, P., MOTRO, A. (Eds.), *Uncertainty Management in Information Systems*, s. 285–324. Dordrecht : Kluwer, 1997.

BOSC, P., KRAFT, D., PETRY, F. Fuzzy sets in database and information systems: Status and opportunities. *Fuzzy Sets and Systems*, 156, 3, s. 418–426, 2005.

BUCKLEY, J. J., QU, Y. On using α-cuts to evaluate fuzzy functions. *Fuzzy Sets and Systems*, 38, s. 309–312, 1990.

CAHA, J., DVORSKÝ, J. Optimal Path Problem with Possibilistic Weights. In: INSPEKTOR, T., IVAN, I., RŮŽIČKA, J. (Eds.), *Symposium GIS Ostrava 2014: Geoinformatics for Intelligent Transportation*, Ostrava, 2014. VSB - Technical University of Ostrava.

CAHA, J., DVORSKÝ, J. Querying on Fuzzy Surfaces with Vague Queries. In: PAN, J.-S., POLYCARPOU, M. M., WOŹNIAK, M., CARVALHO, A. C., QUINTIÁN, H., CORCHADO, E. (Eds.), *Hybrid Artificial Intelligent Systems SE - 55*, s. 548–557. Berlin : Springer Berlin Heidelberg, 2013a.

CAHA, J., DVORSKÝ, J. Comparison of Fuzzy Arithmetic and Stochastic Simulation for Uncertainty Propagation in Slope Analysis. In: *Proceedings of the 12th International Conference on GeoComputation*, Wuhan, China, 2013b. Wuhan University. Available from: `http://www.geocomputation.org/2013/papers/37.pdf`.

CAHA, J., TUČEK, P., VONDRÁKOVÁ, A., PACLÍKOVÁ, L. Slope Analysis of Fuzzy Surfaces. *Transactions in GIS*, 16, 5, s. 649–661, October 2012.

CAHA, J., MAREK, L., PASZTO, V. Spatial Prediction Using Uncertain Variogram. In: BAJAT, B., KILIBARDA, M. (Eds.), *Proceedings of DailyMeteo.org/2014 Conference*, s. 20–24, Belgrade, 2014a. Dosije studio doo. Available from: `http://dailymeteo.org/sites/default/files/DailyMeteo2014WEB.pdf`.

CAHA, J., NEVTÍPILOVÁ, V., DVORSKÝ, J. Constraint and Preference Modelling for Spatial Decision Making with Use of Possibility Theory. In: POLYCARPOU, M., CARVALHO, A. C., PAN, J.-S., WOŹNIAK, M., QUINTIÁN, H., CORCHADO, E. (Eds.), *Hybrid Artificial Intelligence Systems SE - 13*, 8480 / *Lecture Notes in Computer Science*, s. 145–155. Springer International Publishing, 2014b.

CAHA, J., VONDRÁKOVÁ, A., DVORSKÝ, J. Comparison of Crisp, Fuzzy and Possibilistic Threshold in Spatial Queries. In: ABRAHAM, A., KRÖMER, P., SNÁŠEL, V. (Eds.), *Innovations in Bio-inspired Computing and Applications SE - 22*, s. 239–248. Berlin : Springer International Publishing, 2014c.

CARLISLE, B. H. Modelling the Spatial Distribution of DEM Error. *Transactions in GIS*, 9, 4, s. 521–540, October 2005.

CELIKYILMAZ, A., TURKSEN, I. B. *Modeling uncertainty with fuzzy logic : with recent theory and applications*. Č. 240. Berlin : Springer, 2009. ISBN 3540899235 9783540899235.

CHEN, S.-M. Evaluating weapon systems using fuzzy arithmetic operations. *Fuzzy Sets and Systems*, 77, 3, s. 265–276, 1996.

CINGOLANI, P., ALCALá-FDEZ, J. JFuzzyLogic: A robust and flexible Fuzzy-Logic inference system language implementation. In: *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*, s. 1–8, 2012.

CINGOLANI, P., ALCALÁ-FDEZ, J. jFuzzyLogic: a Java Library to Design Fuzzy Logic Controllers According to the Standard for Fuzzy Control Programming. *International Journal of Computational Intelligence Systems*, 6, 1, s. 61–75, 2013.

COROIANU, L., GAGOLEWSKI, M., GRZEGORZEWSKI, P. Nearest piecewise linear approximation of fuzzy numbers. *Fuzzy Sets and Systems*, 233, s. 26–51, 2013.

CROSETTO, M., TARANTOLA, S., SALTELLI, A. Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture Ecosystems & Environment*, 81, 1, s. 71–79, 2000.

CROSETTO, M., TARANTOLA, S. Uncertainty and sensitivity analysis : tools for GIS-based model implementation. *International Journal of Geographical Information Science*, 15, 5, s. 415–437, 2001.

DE BRUIN, S. Querying probabilistic land cover data using fuzzy set theory. *International Journal of Geographical Information Science*, 14, 4, s. 359–372, 2000.

DEMPSTER, A. P. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38, 2, s. 325–339, 1967.

DEVILLERS, R., STEIN, A., BéDARD, Y., CHRISMAN, N., FISHER, P., SHI, W. Thirty Years of Research on Spatial Data Quality: Achievements, Failures, and Opportunities. *Transactions in GIS*, 14, 4, s. 387–400, 2010.

DIAMOND, P. Fuzzy Kriging. *Fuzzy Sets and Systems*, 33, 3, s. 315–332, 1989.

DOMINGO-SANTOS, J. M., VILLARÁN, R. F., RAPP-ARRARÁS, I., PROVENS, E. C.-P. The visual exposure in forest and rural landscapes: An algorithm and a GIS tool. *Landscape and Urban Planning*, 101, 1, s. 52–58, 2011.

DROSG, M. *Dealing with uncertainties : a guide to error analysis*. Berlin : Springer, 2007.

DUBOIS, D., PRADE, H. Operations on Fuzzy Numbers. *International Journal of Systems Science*, 9, 6, s. 613–626, 1978.

DUBOIS, D., PRADE, H. Fuzzy real algebra: Some results. *Fuzzy Sets and Systems1*, 2, s. 327–348, 1979.

DUBOIS, D., PRADE, H. *Fuzzy sets and systems : theory and applications*. Č. Nf. New York : Academic Press, 1980.

DUBOIS, D., PRADE, H. Ranking Fuzzy Numbers in the Setting of Possibility Theory. *Information Sciences*, 30, 3, s. 183–224, 1983.

DUBOIS, D., PRADE, H. *Possibility Theory: An approach to Computerized Processing of Uncertainty*. New York : Plenum Press, 1986. ISBN 0-306-42520-3.

DUBOIS, D., FARGIER, H., PRADE, H. Possibility theory in constraint satisfaction problems : Handling priority , preference and uncertainty. *Applied Intelligence*, 6, s. 287–309, 1996.

DUTTA, P., BORUAH, H., ALI, T. Fuzzy Arithmetic with and without using α-cut method : A Comparative Study. *International Journal of Latest Trends in Computing*, 2, 1, s. 99–107, 2011.

EINSTEIN, A. *Ideas and Opinions*. New York : Broadway Books, 1954.

EL-SHEIMY, N., VALEO, C., HABIB, A. *Digital terrain modeling : acquisition, manipulation, and applications*. Boston : Artech House, 2005.

EVANS, M., HASTINGS, N., PEACOCK, B. Triangular Distribution. In: *Statistical Distributions*, s. 187–188. New York : Wiley-Interscience, 3rd ed. edition, 2000.

FELLEMAN, J., GRIFFIN, C. The role of error in gis-based viewshed determination: A problem analysis. Technical report, IEPP Report EIPP-90-2, Syracuse, New York, 1990.

FISHER, P. The pixel : A snare and a delusion. *International Journal of Remote Sensing*, 18, 3, s. 679–685, 1997.

FISHER, P. Sorites paradox and vague geographies. *Fuzzy Sets and Systems*, 113, 1, s. 7–18, 2000.

FISHER, P. F. 1st Experiments in Viewshed Uncertainty - Simulating Fuzzy Viewsheds. *Photogrammetric Engineering and Remote Sensing*, 58, 3, s. 345–352, 1992.

FISHER, P. Improved Modeling of Elevation Error with Geostatistics. *GeoInfomatica*, 2, 3, s. 215–233, 1998a.

FISHER, P. Valediction. *International Journal of Geographical Information Science*, 21, 10, s. 1165–1170, 2007a.

FISHER, P., CAHA, J. *(in print)* on use of fuzzy surfaces to detect possible elevation change. In: *GIScience*, 2014.

FISHER, P., FARRELLY, C., MADDOCKS, A., RUGGLES, C. Spatial Analysis of Visible Areas from the Bronze Age Cairns of Mull. *Journal of Archaeological Science*, 24, 7, s. 581–592, 1997.

FISHER, P. F. Simulation of the Uncertainty of a Viewshed. In: *51st Annual Convention of the American Congress on Surveying and Mapping / 56th Annual Convention of the American Soc for Photogrammery and Remote Sensing*, s. 205—-218, Bethesda, 1991. Amer Congress Surveying & Mapping.

FISHER, P. F. Algorithm and Implementation Uncertainty in Viewshed Analysis. *International Journal of Geographical Information Systems*, 7, 4, s. 331–347, 1993.

FISHER, P. F. Probable and fuzzy models of the viewshed operation. In: WORBOYS, M. (Ed.), *Innovations in GIS 1*, s. 161–175. Taylor & Francis, 1994.

FISHER, P. F. An exploration of probable viewsheds in landscape planning. *Environment and Planning B: Planning and Design*, 22, 5, s. 527–546, 1995.

FISHER, P. F. Is GIS Hidebound by the Legacy of Cartography? *Cartographic Journal, The*, 35, 1, s. 5–9, 1998b.

FISHER, P. F. Models of uncertainty in spatial data. In: LONGLEY, P., GOODCHILD, M. F., MAGUIRE, D. J., RHIND, D. W. (Eds.), *Geographical Information Systems: Principles, Techniques, Management, and Applications*, chapter 13, s. 191–205. New York, NY : Wiley, 2nd edition, 2005.

FISHER, P. F. Algorithm and Implementation Uncertainty: Any Advances? In: FISHER, P. (Ed.), *Classics from IJGIS: Twenty years of the International Journal of Geographical Information Science and Systems*, s. 225–228. Boca Raton : CRC Press, 2007b.

FISHER, P. F., TATE, N. J. Causes and consequences of error in digital elevation models. *Progress in Physical Geography*, 30, 4, s. 467–489, 2006.

FISHER, P. Extending the Applicability of Viewsheds in Landscape Planning. *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 62, 11, s. 1297 – 1302, 1996.

FLEMING, M., HOFFER, R. M. *Machine processing of landsat MSS data and DMA topographic data for forest cover type mapping, Laboratory for Applications of Remote Sensing.* Purdue University, 1979.

FONTE, C. C., LODWICK, W. A. Modelling the Fuzzy Spatial Extent of Geographical Entities. In: PETRY, F., ROBINSON, V. B., COBB, M. A. (Eds.), *Fuzzy modeling with spatial information for geographic problems*, s. 120–142. Berlin : Springer, 2005. ISBN 3540237135.

FONTE, C., SANTOS, J., GONÇALVES, G., SPINELLA, S., ANILE, M. Applications. In: LODWICK, W. (Ed.), *Fuzzy surfaces in GIS and geographical analysis : theory, analytical methods, algorithms, and applications*, s. 111–128. CRC Press, 2008a.

FONTE, C., SANTOS, J., SPINELLA, S. Algorithms—Pseudo Code. In: LODWICK, W. (Ed.), *Fuzzy surfaces in GIS and geographical analysis : theory, analytical methods, algorithms, and applications*, s. 129–136. Boca Raton : CRC Press, 2008b.

GAGOLEWSKI, M. *FuzzyNumbers Package: Tools to deal with fuzzy numbers in R*, 2014. Available from: `http://FuzzyNumbers.rexamine.com/`.

GAILE, G. L., BURT, J. E. *Directional Statistics.* Geo Abstracts, University of East Anglia, 1980. ISBN 0860940322.

GOODCHILD, M. F., CHIH-CHANG, L., LEUNG, Y. Visualizing fuzzy maps. In: HEARNSHAW, H., UNWIN, D. (Eds.), *Visualization in GIS*, s. 158–167. New York : Wiley, 1994.

GOODCHILD, M. F. Introduction : special issue on ‘ Uncertainty in geographic information systems ’. *Fuzzy Sets and Systems*, 113, 1, s. 3–5, 2000.

GUPTA, S. V. *Measurement Uncertainties: Physical Parameters and Calibration of Instruments.* Dordrecht : Springer Heidelberg, 2012.

HALPERN, J. Y. *Reasoning about uncertainty.* Cambridge, Mass. : MIT Press, 2003.

HANSS, M. The transformation method for the simulation and analysis of systems with uncertain parameters. *Fuzzy Sets and Systems*, 130, 3, s. 277–289, 2002.

HANSS, M. *Applied Fuzzy Arithmetic : An Introduction with Engineering Applications.* Berlin ; New York : Springer-Verlag, 2005.

HEBELER, F. *Modelling Topographic Uncertainty : Impacts on Large Scale Environmental Modelling.* PhD thesis, Universitat Zurich, 2008.

HEBELER, F., PURVES, R. S. The influence of elevation uncertainty on derivation of topographic indices. *Geomorphology*, 111, 1-2, s. 4–16, 2009.

HELTON, J. C., OBERKAMPF, W. L. Alternative representations of epistemic uncertainty. *Reliability Engineering & System Safety*, 85, 1-3, s. 1–10, 2004.

HELTON, J. C., JOHNSON, J. D., OBERKAMPF, W. L. An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering & System Safety*, 85, 1-3, s. 39–71, 2004.

HENGL, T. *A Practical Guide to Geostatistical Mapping.* Amsterdam : University of Amsterdam, 2nd edition, 2009. ISBN 978-90-9024981-0.

HENGL, T. Finding the right pixel size. *Computers & Geosciences*, 32, 9, s. 1283–1298, November 2006.

HERNÁNDEZ, J., GARCÍA, L., AYUGA, F. Assessment of the visual impact made on the landscape by new buildings: a methodology for site selection. *Landscape and Urban Planning*, 68, 1, s. 15–28, 2004.

HERRERA, F., HERRERA-VIEDMA, E., MARTÍNEZ, L., WANG, P., BORDOGNA, G., CHIESA, S., GENELETTI, D. Linguistic modelling of imperfect spatial information as a basis for simplifying spatial analysis. *Information Sciences*, 176, 4, s. 366–389, 2006.

HESHMATY, B., KANDEL, A. Fuzzy Linear-Regression and Its Applications to Forecasting in Uncertain Environment. *Fuzzy Sets and Systems*, 15, 2, s. 159–191, 1985.

HEUVELINK, G. B. M., BURROUGH, P. A. Error propagation in cartographic modelling using boolean logic and continuous classification. *International Journal of Geographical Information System*, 7, s. 231–246, 1993.

HEUVELINK, G. B. M., BURROUGH, P. A., STEIN, A. Propagation of errors in spatial modelling with gis. *International Journal of GIS*, 3, 4, s. 303–322, 1989.

HEUVELINK, G. B. M. *Error Propagation in Environmental Modelling with GIS.* London : Taylor & Francis Ltd, 1998.

HEUVELINK, G. B. M. Analysing Uncertainty Propagation in GIS:Why is it not that Simple? In: FOODY, G. M., ATKINSON, P. M. (Eds.), *Uncertainty in remote sensing and GIS*, s. 155—-165. Chichester : Wiley, 2002.

HIEMSTRA, P., PEBESMA, E., TWENHÖFEL, C., HEUVELINK, G. Real-time automatic interpolation of ambient gamma dose rates from the dutch radioactivity monitoring network. *Computers & Geosciences*, 35, 8, s. 1711–1721, 2008. DOI: http://dx.doi.org/10.1016/j.cageo.2008.10.011.

HIGUCHI, T. *The Visual and Spatial Structure of Landscapes.* Cambridge, Massachusetts : MIT Press, 1983.

HOLMES, K. W., CHADWICK, O. A., KYRIAKIDIS, P. C. Error in a USGS 30-meter digital elevation model and its impact on terrain modeling. *Journal of Hydrology*, 233, s. 154–173, 2000.

HORN, B. Hill shading and the reflectance map. *Proceedings of the IEEE*, 69, 1, s. 14–47, 1981.

HUNTER, G. J., GOODCHILD, M. F. Modeling the Uncertainty of Slope and Aspect Estimates Derived from Spatial Databases. *Geographical Analysis*, 29, 1, s. 35–49, 1997.

JIANG, H., EASTMAN, J. R. Application of fuzzy measures in multi-criteria evaluation in GIS. *International Journal of Geographical Information Science*, 14, 2, s. 173–184, 2000.

JONES, K. H. A comparison of algorithms used to compute hill slope as a property of the dem. *Computers & Geosciences*, 24, s. 315–323, 1998.

KADANE, J. B. *Principles of uncertainty.* Boca Raton, FL : CRC Press, 2011.

KAUFMANN, A., GUPTA, M. M. *Introduction to Fuzzy Arithmetic.* New York : Van Nostrand Reinhold Company, 1985.

KEYSER, C. J. *Mathematical Philosophy.* New York : E. P. Dutton & Company, 1922.

KLIMKE, A., WILLNER, K., WOHLMUTH, B. Uncertainty Modeling Using Fuzzy Arithmetic Based on Sparse Grids: Applications to Dynamic Systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12, 06, s. 745–759, 2004.

KLIMKE, W. A. *Uncertainty Modeling using Fuzzy Arithmetic and Sparse Grids.* Dissertation, Universität Stuttgart, 2006.

KLIR, G. J. *Uncertainty and information : foundations of generalized information theory.* Hoboken, N.J. : Wiley-Interscience, 2006.

KLIR, G. J., YUAN, B. *Fuzzy Sets and Fuzzy Logic: Theory and Applications.* Upper Saddle River : Prentice Hall, 1995.

LEUNG, Y., GOODCHILD, M. F., LIN, C. C. Visualization of fuzzy scenes and probability fields. In: NEWTON, H. (Ed.), *Computing Science and Statistics, Volume 24: Graphics and Visualization (Proceedings of the 24th Symposium on the Interface)*, 24, s. 416–422, Fairfax Station, 1993. Interface Foundation of N America.

LI, R., BHANU, B., RAVISHANKAR, C., KURTH, M., NI, J. Uncertain spatial data handling: Modeling, indexing and query. *Computers & Geosciences*, 33, 1, s. 42–61, 2007.

LINDLEY, D. V. *Understanding uncertainty.* Hoboken, N.J. : Wiley-Interscience, 2006.

LIU, B. *Uncertainty theory a branch of mathematics for modeling human uncertainty.* Berlin, Heidelberg : Springer,, 2010.

LODWICK, W. A., SANTOS, J. Constructing consistent fuzzy surfaces from fuzzy data. *Fuzzy Sets and Systems*, 135, 2, s. 259–277, 2003.

LODWICK, W., ANILE, A. M., SPINELLA, S. Introduction. In: LODWICK, W. (Ed.), *Fuzzy surfaces in GIS and geographical analysis : theory, analytical methods, algorithms, and applications*, s. 1–46. Boca Raton : CRC Press, 2008.

LONGLEY, P., GOODCHILD, M. F., MAGUIRE, D., RHIND, D. *Geographical information systems and science.* Chichester : Wiley, 2nd edition, 2005.

LOOTS, L., NACKAERTS, K., WAELKENS, M. Fuzzy Viewshed Analysis of the Hellenistic City Defence System at Sagalassos , Turkey. In: DINGWALL, L., EXON, S., GAFFNEY, V., LAFLIN, S., LEUSEN, M. (Eds.), *Proceedings of the 25th Anniversary Conference of Computer Applications and Quantitative Methods in Archaeology*, Birmingham, 1999.

LOQUIN, K., DUBOIS, D. Kriging with Ill-Known Variogram and Data. In: DESHPANDE, A., HUNTER, A. (Eds.), *Scalable Uncertainty Management SE - 5leží*, 6379 / *Lecture Notes in Computer Science*, s. 219–235. Springer Berlin / Heidelberg, 2010a.

LOQUIN, K., DUBOIS, D. Kriging and Epistemic Uncertainty: A Critical Discussion. In: JEANSOULIN, R., PAPINI, O., PRADE, H., SCHOCKAERT, S. (Eds.), *Methods for Handling Imperfect Spatial Information SE - 11*, 256 / *Studies in Fuzziness and Soft Computing*, s. 269–305. Springer Berlin Heidelberg, 2010b. ISBN 978-3-642-14754-8.

LOQUIN, K., DUBOIS, D. A fuzzy interval analysis approach to kriging with ill-known variogram and data. *Soft Computing*, 16, 5, s. 769–784, 2011.

MANSKI, C. R. *Partial Identification of Probability Distributions.* New York : Springer- Verlag, 2003.

MARDIA, K. V., JUPP, P. E. *Directional Statistics.* Wiley, 1999. ISBN 0471953334.

METROPOLIS, N., ULAM, S. The Monte Carlo Method. *Journal of the American Statistical Association*, 44, 247, s. 335–341, 1949.

MILLER, C. L., LAFLAMME, R. A. *The digital terrain model - theory & application.* Č. no 117. Cambridge, Mass. : M.I.T. Photogrammetry Laboratory, 1958.

MOORE, R. E. *Interval Analysis.* Englewood Cliffs, NJ : Prentice Hall, 1966.

MOORE, R. E., KEARFOTT, R. B., CLOUD, M. J. *Introduction to interval analysis.* Philadelphia, PA : Society for Industrial and Applied Mathematics, 2009.

MOUNT, N., FISHER, P., BRUNSDON, C., FARMER, E. Book Reviews. *International Journal of Geographical Information Science*, 23, 4, s. 545–549, 2009.

NACKAERTS, K., GOVERS, G., ORSHOVEN, J. V. Accuracy assessment of probabilistic visibilities. *International Journal of Geographical Information Science*, 13, 7, s. 709–721, October 1999.

NAGY, G. Terrain Visibility. *Computers & Graphics*, 18, 6, s. 763–773, 1994.

NAHMIAS, S. Fuzzy variables. *Fuzzy Sets and Systems*, 1, 2, s. 97–110, 1978.

NETELER, M., MITASOVA, H. *OPEN SOURCE GIS: A GRASS GIS Approach.* New York : Springer, third edition, 2008.

NGUYEN, H. T., KREINOVICH, V., WU, B., XIANG, G. *Computing statistics under interval and fuzzy uncertainty applications to computer science and engineering.* Berlin ; New York : Springer, 2012.

OBERGUGGENBERGER, M. The mathematics of uncertainty: models, methods and interpretations. In: FELLIN, W., LESSMANN, H., OBERGUGGENBERGER, M., VIEIDER, R. (Eds.), *Analyzing Uncertainty in Civil Engineering*, s. 252. Berlin : Springer, 2005. ISBN 978-3540222460.

OGBURN, D. E. Assessing the level of visibility of cultural objects in past landscapes. *Journal of Archaeological Science*, 33, 3, s. 405–413, 2006.

OHSAWA, Y., KOBAYASHI, T. An Analytical Model to Assess the Visibility of Landmarks. *Geographical Analysis*, 37, 3, s. 336–349, 2005.

OKSANEN, J. *Digital Elevation Model Error in Terrain Analysis.* PhD thesis, University of Helsinki, 2006.

OKSANEN, J., SARJAKOSKI, T. Error propagation of DEM-based surface derivatives. *Computers & Geosciences*, 31, 8, s. 1015–1027, 2005a.

OKSANEN, J., SARJAKOSKI, T. Error propagation analysis of DEM-based drainage basin delineation. *International Journal of Remote Sensing*, 26, 14, s. 3085–3102, 2005b.

OKSANEN, J., SARJAKOSKI, T. Uncovering the statistical and spatial characteristics of fine toposcale DEM error. *International Journal of Geographical Information Science*, 20, 4, s. 345–369, 2006.

O'SULLIVAN, D., TURNER, A. Visibility graphs and landscape visibility analysis. *International Journal of Geographical Information Science*, 15, 3, s. 221–237, 2001.

RAPER, J. Geographic information science. *Annual Review of Information Science and Technology*, 43, 1, s. 1–117, 2009.

REITSMA, F. Revisiting the 'Is GIScience a science?' debate (or quite possibly scientific gerrymandering). *International Journal of Geographical Information Science*, 27, 2, s. 211–221, 2013.

RIBEIRO JR, P. J., DIGGLE, P. J. geoR: a package for geostatistical analysis. *R-NEWS*, 1, 2, s. 14—-18, 2001. Available from: `http://cran.r-project.org/doc/Rnews/`.

RUBINSTEIN, R. Y., KROESE, D. P. *Simulation and the Monte Carlo Method.* Hoboken, N.J. : John Wiley & Sons, Inc., 2nd edition, 2008.

SANTOS, J., LODWICK, W. A., NEUMAIER, A. A New Approach to Incorporate Uncertainty in Terrain Modeling. In: EGENHOFER, M., MARK, D. (Eds.), *Geographic Information Science*, 2478 / *Lecture Notes in Computer Science*, s. 291–299. Springer Berlin Heidelberg, 2002.

SARMENTO, P., FONTE, C. C., CAETANO, M., STEHMAN, S. V. Incorporating the uncertainty of linguistic-scale reference data to assess accuracy of land-cover maps using fuzzy intervals. *International Journal of Remote Sensing*, 34, 11, s. 4008–4024, 2013.

SHAD, R., MESGARI, M. S., ABKAR, A., SHAD, A. Predicting air pollution using fuzzy genetic linear membership kriging in GIS. *Computers Environment and Urban Systems*, 33, 6, s. 472–481, 2009.

SHARPNACK, D. A., AKIN, G. An algorithm for computing slope and aspect from elevations. *Photogrammetric Engineering*, 35, s. 247–248, 1969.

SHI, W. *Principles of modeling uncertainties in spatial data and analyses.* Boca Raton : CRC Press/Taylor & Francis, 2010. ISBN 9781420059274.

SKIDMORE, A. K. A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. *International journal of geographical information systems*, 3, 4, s. 323–334, 1989.

SKIDMORE, A. K. Evolution of Methods for Estimating Slope Gradient and Aspect from Digital Elevation Models. In: FISHER, P. (Ed.), *Classics from IJGIS: Twenty years of the International Journal of Geographical Information Science and Systems*, s. 111–117. Boca Raton : CRC Press, 2007.

SORENSEN, P. A., LANTER, D. P. Partial for Determining Two Algorithms Data Structure Visibility and Reducing Analysis Error in Viewshed lnduced. *Photogrammetric Engineering and Remote Sensing*, 59, s. 1129–1132, 1993.

SOZER, A., YAZICI, A., OGUZTUZUN, H., TAS, O. Modeling and querying fuzzy spatiotemporal databases. *Information Sciences*, 178, 19, s. 3665–3682, 2008.

SPINELLA, S. Appendices—Fuzzy Arithmetic and Fuzzy Query C++ Source. In: LODWICK, W. (Ed.), *Fuzzy surfaces in GIS and geographical analysis : theory, analytical methods, algorithms, and applications*, s. 137–158. Boca Raton : CRC Press, 2008.

TANAKA, H., UEJIMA, S., ASAI, K. Linear Regression Analysis with Fuzzy Model. *IEEE Transactions on Systems, Man and Cybernetics*, 12, 6, s. 903–907, 1982.

TOMLINSON, R. F. *Application of Electronic Computing Methods and Techniques to the Storage, Compilation and Assessment of Mapped Data.* PhD thesis, University College London, 1974.

TRAN, L., DUCKSTEIN, L. Comparison of fuzzy numbers using a fuzzy distance measure. *Fuzzy Sets and Systems*, 130, 3, s. 331–341, 2002.

TUKEY, J. W. The Future of Data Analysis. *The Annals of Mathematical Statistics*, 33, 1, s. 1–67, March 1962.

VIERTL, R. *Statistical methods for fuzzy data.* Chichester, West Sussex : Wiley, 2011.

VONDRÁKOVÁ, A., CAHA, J. Visualization of fuzzy surfaces. In: *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM*, s. 1069–1076, Sofia, 2014. STEF92 Technology Ltd.

VOXMAN, W. Some remarks on distances between fuzzy numbers. *Fuzzy Sets and Systems*, 100, 1, s. 353–365, 1998.

WAELDER, O. An application of the fuzzy theory in surface interpolation and surface deformation analysis. *Fuzzy Sets and Systems*, 158, 14, s. 1535–1545, 2007.

WILSON, J. P. Digital terrain modeling. *Geomorphology*, 137, 1, s. 107–121, 2012.

WILSON, J. P., GALLANT, J. C. *Terrain analysis : principles and applications.* New York : Wiley, 2000.

WITLOX, F., DERUDDER, B. Spatial Decision-Making Using Fuzzy Decision Tables: Theory, Application and Limitations. In: PETRY, F., ROBINSON, V. B., COBB, M. A. (Eds.), *Fuzzy modeling with spatial information for geographic problems*, s. 252–274. Berlin : Springer, 2005.

ZADEH, L. A. Fuzzy Sets. *Information and Control*, 8, 3, s. 338–353, 1965.

ZADEH, L. A. Similarity relations and fuzzy orderings. *Information Sciences*, 3, 2, s. 177–200, 1971.

ZADEH, L. A. Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. *EEE Transactions on Systems, Man and Cybernetics*, 3, 1, s. 16, 1973.

ZADEH, L. A. Concept of a Linguistic Variable and Its Application to Approximate Reasoning-I. *Information Sciences*, 8, 3, s. 199–249, 1975.

ZADEH, L. A. Discussion: Probability Theory and Fuzzy Logic Are Complementary Rather Than Competitive. *Technometrics*, 37, 3, s. 271–275, 1995.

ZADEH, L. A. Toward a generalized theory of uncertainty (GTU) - an outline. *Information Sciences*, 172, 1-2, s. 1–40, 2005.

ZADEH, L. A. Generalized theory of uncertainty (GTU) - principal concepts and ideas. *Computational Statistics & Data Analysis*, 51, 1, s. 15–46, 2006.

ZADEH, L. A. Is there a need for fuzzy logic? *Information Sciences*, 178, 13, s. 2751–2779, 2008.

ZHANG, J., GOODCHILD, M. F. *Uncertainty in geographical information.* London : Taylor & Francis, 2002.

ZHOU, Q. M., LIU, X. J. Analysis of errors of derived slope and aspect related to DEM data properties. *Computers & Geosciences*, 30, 4, s. 369–378, 2004.

ZIMMERMANN, H. J. *Fuzzy set theory - and its applications.* Boston : Kluwer Academic Publishers, 2nd rev. edition, 1991.

Shrnutí

Téma propagace nejistoty se v geoinformatické literatuře objevuje přibližně od konce 80. let 20. století. Přestože se zhruba ve stejné době objevují první aplikace pro modelování nejistoty pomocí fuzzy množin, je velká většina výzkumu v této oblasti spojená s využitím statistických metod, zejména pak metody Monte Carlo. Ostatní existující metody nejsou příliš často využívány a jen výjimečně byly tyto metody porovnány, aby byly zjištěny odchylky mezi poskytovanými výsledky. Tato práce se zaměřuje na možnost využití fuzzy aritmetiky jako metody pro propagaci nejistoty v analýzách fuzzy povrchů. Fuzzy povrchy jsou speciálním případem povrchů, které v sobě přirozeně obsahují jistou míru nejistoty. Jejich vhodnost a využitelnost pro modelování nejistoty povrchů byla již několikrát ověřena a existuje celá řada postupů pro jejich tvorbu. Nicméně analýzy těchto povrchů nebyly ve větší míře nikdy studovány.

Hlavním cílem disertační práce bylo popsání využitelnosti fuzzy aritmetiky pro analýzy fuzzy povrchů. K dosažení hlavního cíle bylo třeba splnění několika dílčích cílů. Tyto cíle zahrnují popis nejistoty z hlediska matematiky, vytvoření seznamu matematických teorií, které mohou být použity pro modelování nejistoty a její propagaci. Dále bylo nezbytné provést sumarizaci teoretických základů fuzzy množin a fuzzy aritmetiky v rozsahu postačujícím pro jejich využití v analýzách povrchů. Další cílem bylo shrnutí metod a postupů využívaných v geoinformatice pro modelování a propagaci nejistoty, se zaměřením zejména na analýzy povrchů. Nejdůležitější částí disertační práce je pak vytvoření a prezentace metod a postupů pro výpočet sklonu a orientace fuzzy povrchu a také viditelnosti nad tímto povrchem. V rámci uvedení získaných dat do širšího kontextu je stručně představeno možné využití získaných výsledků pro podporu rozhodování. Vytvořené metody pro výpočty fuzzy analýz jsou demonstrovány na případové studii.

Jednotlivé představované analýzy byly na jednoduchých ukázkových výpočtech porovnány s metodou Monte Carlo. Cílem těchto porovnání bylo demonstrovat, že každá z těchto metod propagace nejistoty se zaměřuje na odlišnou komponentu nejistoty a poskytují tudíž rozdílné výsledky, a to nejen číselně, ale i z hlediska sémantiky. Ukázkové výpočty všech tří analýz na rozsáhlejším území jsou demonstrovány v případové studii. V rámci případové studie je zmíněn i potenciál využití získaných výsledků pro podporu rozhodování.

Cíle práce byly beze zbytku naplněny, disertační práce představuje možnosti využití fuzzy aritmetiky jako metody pro propagaci nejistoty v analýzách fuzzy povrchů na příkladech výpočtu sklonu, orientace a viditelnosti. Výsledky získané na případové studii naznačují potenciál této metody v praktických aplikacích. Pro podporu rozhodování je zajímavá zejména vlastnost výsledků, které mají horní a dolní limitu. Tuto vlastnost nelze při využití metody Monte Carlo zaručit. Pro praktické aplikace je přitom tato vlastnost poměrně klíčová, neboť umožňuje vytváření limitních scénářů. Disertační práce také naznačuje další možné výzkumné směry, které se otevírají využitím fuzzy aritmetiky. Jedná se zejména o prezentaci dalších analýz fuzzy povrchů, využívání dat s nejistotou pro kvalitnější podporu rozhodování či vhodnost vizualizace získaných výsledků.

APPENDICES
TO THE DISSERTATION THESIS

Appendix 1 - CD with the text of dissertation thesis and datasets from the case study

```
CD Structure
├── thesis
│   └── dissertation_thesis.pdf ........................ Text of the thesis.
└── case_study ........................................ Data of the case study.
    ├── data_generation_scripts ......... R scripts used to generate the data.
    │   ├── _functionDefitions.R
    │   ├── case_study.R
    │   └── case_study.RData
    └── data ............. Data visualized in the case study (mostly ASCII grids).
        ├── 1_points
        │   └── points.txt .......... Points used to interpolate the fuzzy surface.
        ├── 2_fuzzy_surface
        │   ├── fuzzySurface_max.txt
        │   ├── fuzzySurface_min.txt
        │   └── fuzzySurface_modal.txt
        ├── 3_slope
        │   ├── fuzzySurface_slope_max.txt
        │   ├── fuzzySurface_slope_min.txt
        │   └── fuzzySurface_slope_modal.txt
        ├── 4_aspect
        │   ├── fuzzySurface_aspect_max.txt
        │   ├── fuzzySurface_aspect_min.txt
        │   └── fuzzySurface_aspect_modal.txt
        ├── 5_slope_exceedance .......... Query outcomes shown in chap. 11.3.3.
        │   ├── slope_exceedanceNecessity.txt
        │   └── slope_exceedancePossibility.txt
        └── 6_visibility
            ├── fuzzy_visibility_v1_nec.txt
            ├── fuzzy_visibility_v1_nec_strict.txt
            ├── fuzzy_visibility_v1_poss.txt
            ├── fuzzy_visibility_v1_poss_strict.txt
            └── viewpoint1.shp ........ Shapefile of viewpoint used in case study.
```