

A Spatial Fuzzy Influence Diagram for Modelling Spatial Objects Dependencies: a Case Study on Tree-related Electricity Outages

Zhe Zhang^a, Urška Demšar^b, Shaowen Wang^{cd} and Kirsi Virrantaus^a

^a *Research group on Geoinformatics, Department of Built Environment, Aalto University, Finland*

^b *School of Geography & Sustainable Development, University of St Andrews, U.K.*

^c *CyberGIS Center for Advanced Digital and Spatial Studies, National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, U.S.A*

^d *CyberInfrastructure and Geospatial Information Laboratory, Department of Geography and Geographical Information Sciences, University of Illinois at Urbana-Champaign, U.S.A*

Abstract

Spatial objects can be interconnected and mutually dependent in complex ways. In GIScience, spatial objects' topological relationships are not discussed together with their attributes' dependencies, and the vagueness of spatial objects is often ignored during the spatial modelling process. To address this, a spatial fuzzy influence diagram is introduced. Compared to the traditional statistical or fuzzy modelling approach, the influence diagram brings advantages in helping decision makers structure complex interdependency problems. A questionnaire was developed to evaluate the applicability of using an influence diagram in modelling spatial objects' dependencies. As a case study, a spatial fuzzy influence diagram is applied to tree-related electric outages. The result of the case study is represented as a vulnerability map of electrical networks. The map shows areas at risk due to tree-related electric outages. The results were first validated by using a visual comparison of the vulnerability map and electricity fault

data. In the second validation step, the percentage of fault-data, which has received values in different vulnerability categories was calculated. The results of the case study can be used to support the decision-making process of electrical network maintenance and planning.

Keywords: influence diagram, spatial fuzzy influence diagram, critical infrastructure, tree-related electric outages, fuzzy logic

1. Introduction

In geographical information science (GIScience) literature, spatial objects' attribute dependencies are not often discussed together with topological relationships. A typical example of this combined relationship is critical infrastructure dependency. For instance, if critical infrastructure such as electricity distribution and telecommunication networks are modelled as spatial objects, the dependency relationships between the two networks can be seen from two perspectives. The first one is attribute dependency: the failure of an electrical network will cause the failure of a telecommunications network because a telecommunications network needs electric power to work. The second one is topological dependency: an electricity distribution failure may cause failures in the telecommunications network in a specific geographical area; this is the case where spatial thinking should take place. Critical infrastructure dependencies have been well discussed in the literature (Chang *et al.*, 2007; Espada *et al.*, 2015; Utne *et al.*, 2011; McDaniels *et al.*, 2007; Rinaldi *et al.*, 2001). However, these studies mostly focus on attribute dependency without considering topological dependency and spatial context.

The motivation for this research work came from a research project called Alueellinen Varautuminen (ALVAR) (Rautvuori, 2014), funded by the National Emergency Supply Agency of Finland. This project involved several major critical

infrastructure companies of Finland with the goal of studying failure dependencies of critical infrastructure from an emergency management point of view. The project attempted to solve several problems. First, critical infrastructures are highly interconnected and mutually dependent in complex ways, which makes the interpretation of failure dependencies difficult. Second, expert knowledge is important in defining failure dependencies, but this knowledge is difficult to collect and represent in the modelling process. In the ALVAR project, experts from different critical infrastructure companies were asked to participate in several brainstorming sessions to collect ideas of how the failure of one critical infrastructure could affect others. However, there were difficulties in building common semantic models between the experts since they all had different backgrounds. Their opinions on critical infrastructure failure dependency were difficult to represent in the model in a unified way. In addition, all the experts thought spatial context should be involved in the failure dependency modelling process, but did not know how best to do so. Spatial context is important, as the failure of an electrical network will only affect other critical networks within a specific spatial distance. Moreover, involving spatial context can create a big data problem since the geographical data can become large and complex when the number of spatial objects and their dependencies increases. Therefore, there is a need to develop a general framework, which considers two types of dependencies (attribute and topology), data vagueness, and how expert knowledge could be better collected and represented in the modelling process.

In this article, a new methodology, a spatial fuzzy influence diagram (SFID) to model the dependencies of spatial objects, is proposed. Using a spatial influence diagram for decision-making is not new. For example, Zhu et al. (1996) modelled spatial object dependencies as a knowledge-based decision support system, where a

spatial influence diagram was used to support the decision-making process of land use planning. However, the vagueness of their data was not discussed in their work, and there was no decision node in the spatial influence diagram. Both are important issues for this research. Therefore, a spatial influence diagram, like Zhu's, can be extended by using a fuzzy approach. Compared to simple spatial fuzzy modelling, an influence diagram allows decision makers to structure complex dependency problems. An influence diagram is a conceptual modelling tool that graphically represents the causal relationships between decisions, uncertainties, and outcomes. The decision process will become more efficient when the decision structure is clear and all the decision elements are specified in advance.

Spatial data are often not precise and have inherent uncertainties (Longley et al., 2005). If the spatial object is poorly defined, the definition of a class or set within the universe is a matter of vagueness or ambiguity, which can be treated by using a fuzzy set theory (Fisher et al., 2006). In fuzzy set theory, values of an object are described as linguistic values and fuzzy set membership functions instead of crisp values, which allows the incorporation of uncertainty (i.e., vagueness) in the modelling process (Zadeh, 1996). This is an advantage for solving real-life problems, where linguistic information is frequently hard to quantify using classical mathematical techniques, while at the same time very important, as it represents subjective expert knowledge. For instance, fuzzy set theory has been used in this context for decision-making in the fields of transportation engineering (Teodorović, 1999), power system control (Kripakaran et al. 2014), and telecommunications (Sawhney et al., 2014). Additionally, while human reasoning plays an important role in every type of spatial data analysis and modelling, it is complex and therefore difficult for a computer to replicate. Fuzzy logic is one of the methods that facilitates computer emulation of some aspects of human reasoning, by

generating fuzzy membership functions and sets of if-then rules using expert knowledge. This has been demonstrated to be useful in various GIScience studies. For instance, Zhang et al. (2014) developed a fuzzy multiple-attribute decision-making model for vulnerability analysis on the basis of population information for disaster management. Expert knowledge was collected using a dynamic questionnaire, which was used to define the fuzzy membership function and fuzzy rules, and acted as one important parameter in model validation. Sicat et al. (2005) demonstrated a fuzzy modelling of farmers' knowledge for agricultural land suitability classification. Zhu et al. (2001) combined expert knowledge and fuzzy modelling with GIS for soil mapping.

A fuzzy influence diagram models data vagueness using fuzzy reasoning instead of probabilities. It contains a set of nodes, which can be used to describe decision problems. If the values of the nodes are represented as probabilities, such as using the Bayesian theorem, the diagram is called a Bayesian network (Neapolitan, 2004). The Bayesian network has been often used in GIS applications (Aspinall, 1992; Li et al., 2010; Stassopoulou et al., 1998). However, it has a number of disadvantages, for example, the value of the node cannot be assigned numerically or expressed as a probability (Mateou et al., 2005). In such cases, fuzzy logic is more suitable than using a Bayesian framework. For example, Zhi and Jin (2010) used a fuzzy influence diagram to support safety risk assessment for highway tunnel construction.

In this article, a spatial influence diagram is combined with fuzzy modelling to create a new representation of the critical interdependencies: a spatial fuzzy influence diagram (SFID). This approach is applied to a case study on tree-related electric outage data from the ALVAR project. The purpose of this research is to introduce this method, which can then be utilised by experts to produce accurate scientific results using their own knowledge and study data. Furthermore, a questionnaire is developed in order to

investigate the applicability of using an influence diagram to model spatial objects' dependencies.

The rest of the article is organized as follows: section 2 gives an overview of the theoretical background relevant to the development of SFID. The general structure of SFID and the survey questionnaire are also presented in this section. Section 3 introduces the motivation and the background of a tree-related electric outage case study, as well as implementing SFID on the case study scenario. The results of the case study and validation of the results are illustrated in section 4. Sections 5 and 6 present the discussion and conclusion.

2. Methodology

In this section, the basic concepts of an influence diagram, fuzzy logic, and spatial component integration are introduced. This section also covers how to combine all these concepts to solve complicated decision-making problems using SFID.

2.1 Influence diagrams

An influence diagram graphically describes the structure of a decision problem, where relationships are shown with nodes and arcs (Howard and Matheson, 2005). A decision represents sets of mutually exclusive and exhaustive actions that a decision maker can take. Figure 1 illustrates an example of an influence diagram, which consists of three kinds of nodes and links (Tatman and Shachter, 1990). A chance node corresponds to a random variable that cannot be controlled by the decision-maker and is represented as oval shape in the diagram. Value nodes (octagon shaped) correspond to the utility functions, which represent a decision maker's preferences. A decision node corresponds to each decision that needs to be made, and is represented as a rectangle in the diagram.

Functional arcs end in value nodes and indicate that the utility function is a function of all the nodes that are connected to it. Conditional arcs end in chance nodes, and indicate that the uncertainty at their heads is probabilistically conditioned on all the nodes at their tails. Informational arcs ending in a decision node indicate that the decision at their heads is made with the outcome of all the nodes at their tails known beforehand.

Figure 1 Somewhere here.

2.2 Fuzzy logic

Fuzzy set theory was first introduced by Zadeh (1965) as a mathematical theory of vagueness. He defined a fuzzy set as: If X is the universe of discourse, and its elements are denoted by x , then the fuzzy set A in X is defined as a set of ordered pairs called a membership function (MF) of x in A . The MF maps each element of X to a value between 0 and 1. A fuzzy MF can be described in terms of linguistic variables such as *low*, *middle*, and *high*. For instance, a mathematical form of a triangular fuzzy MF curve is shown by equation 1 (Kaufmann and Gupta, 1985). The parameters a , b , c (with $a < b < c$) determine a x coordinate of the three corners of the triangular MF.

$$\text{triangle}(x; a, b, c) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ \frac{c-x}{c-b}, & b < x \leq c \\ 0, & x > c \end{cases} \quad (1)$$

A triangular MF is simple and commonly used, while other types of MFs such as Gaussian or Sigmoidal shapes are more complicated and computationally expensive (Zhao and Bose, 2002).

Figure 2 illustrates an example of a fuzzy MF and fuzzy rules. In this example, there are two input variables x and y , and each of them has three triangular MFs. The

degree of membership function μ is used to measure how likely an input value belongs to a corresponding MF. For instance, input value x more likely belongs to low MF than middle MF if the value of $\mu_{x(\text{low})}$ is greater than $\mu_{x(\text{middle})}$. Fuzzy rules are then used to obtain the relationships between the input and output values. These rules work as follows (Zimmermann, 2001): the *if part* (antecedent part) of the rule partitions the input space into a number of fuzzy regions (MFs), and the *then part* (consequent part) describes the behaviour of the system in these fuzzy regions. An example of a fuzzy rule is: ‘If input X is *high* and input Y is *high* then output Z (decision maker’s preference) is *low*.’ Fuzzy operators OR, AND and NOT in the fuzzy rule can be used to describe a fuzzy union, intersect, or complement operations of the input MFs. The mathematical expressions of the fuzzy operators are illustrated in Equations 2, 3 and 4 where $\mu_A(u)$ and $\mu_B(u)$ are MFs.

$$\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} \quad x \in X \text{ (AND operator)} \quad (2)$$

$$\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} \quad x \in X \text{ (OR operator)} \quad (3)$$

$$\mu_{\bar{A}} = 1 - \mu_A(x), \quad x \in X \text{ (NOT operator)} \quad (4)$$

Figure 2 Somewhere here.

2.3 Spatial fuzzy influence diagram

In the SFID, an uncertain spatial object is modelled as a chance node with attributes and geometry data, as shown in Figure 3. In this SFID, the two spatial objects’ chance nodes X and Y are connected to two types of fuzzy value nodes. A fuzzy value node translates spatial object data values into MFs, which represents the vagueness of the spatial object from the decision maker’s perspective. A spatial value node represents the topological

dependency of spatial objects and provides a space in the diagram where spatial information can be integrated into the links. A spatial node represents the spatial value of an uncertainty node and demonstrates how a spatial value can be represented in fuzzy value nodes. For instance, a spatial overlay function can be used to help a decision maker select which uncertainty node is closer (touches or intersects the other uncertainty node at a specific distance). Only the uncertainty nodes of spatial interest will be selected and represented in a fuzzy value node. A spatial node can also be used to demonstrate the spatial dependencies of different types of critical infrastructure and help to create MFs in the fuzzy value nodes. The *fuzzy output MF* value node represents the consequent part of the fuzzy rules as MFs according to the decision maker's preferences. The decision node (rectangular shaped) uses fuzzy rules to describe decisions for a specific decision problem. The antecedent part of a fuzzy rule comes from the MFs, which are created in the *fuzzy input MF* value node; the consequent part is derived from the *fuzzy output MF* value node. Fuzzy operators can be used to connect the antecedent part and consequent part of the rule.

The SFID can be applied to different spatial decision-making applications by adding more nodes, modifying chance nodes' corresponding value node membership functions or re-computing fuzzy rules in the decision node.

Figure 3 somewhere here.

2.4 Evaluation survey

An influence diagram is a graphical representation of a decision problem which helps decision makers to understand and formulate decision problems. However, the efficiency of using an influence diagram is often difficult to quantify since it represents the decision makers' problem structuring and knowledge discovery process. This article

claims that SFID can be used to structure complex decision problems, which are related to spatial objects' dependencies. Therefore, it is interesting to show how useful this approach can be. For this reason, a questionnaire is developed to evaluate the applicability of using influence diagrams in modelling spatial objects' dependencies.

Twenty GIS experts participated in this survey. Most of the participants had worked with GIS before, but none of them had experience with influence diagrams. The questionnaire contains an introduction to influence diagrams, and example failures of four types of critical infrastructures (e.g. electricity, telecommunication, transportation network and water supply), and a scenario map. The participants were first encouraged to consider the failure dependencies of critical infrastructures without using an influence diagram. In the next step, a map that contained different types of critical infrastructures and houses was given to the participants. They were asked to formulate a decision problem, draw an influence diagram and record the time that they used in completing the task. Finally, they were asked to assign a grade from 1 (didn't help at all) to 5 (yes, helped me) to evaluate the applicability of using an influence diagram to model spatial objects' dependencies. The results of the survey are presented in section 4.

3. Case study: tree-related electric outages

The vulnerability of electric infrastructure is a crucial issue for modern society. One scenario that leads to electric outages is a tree falling or leaning onto power lines, which causes a so-called tree-related electric outage. This scenario is usually weather related because extreme weather conditions such as storms can cause trees to fall and damage electrical networks. This problem is more serious in countries with a high level of forest cover (such as Finland), where electrical networks run through or parallel to forested areas. In Finland, this has to some extent been addressed through a line clearance tree

trimming programme, which is carried out each year to improve the reliability of the electrical network. In Finnish urban areas, critical electric cables have also been moved underground to avoid tree-related electric outages. However, the electrical networks face many further challenges that can impact the reliability of the electric supply. For instance, it is hard to know the average number of hazardous trees and their locations. Re-locating cables underground, monitoring tree health and growth, and making judgements about appropriate inspections for the line clearance tree trimming programme are expensive and time consuming. Therefore, tree-related outages are not only a risk to human safety but also a financial and economical liability.

Tree-related outages have been extensively studied in literature. Guggenmoos (2003) studied the effects of tree mortality on power line security. Rees *et al.* (1994) designed a testing approach to identify factors in tree health, customer concerns, and tree outage relationships. Simpson and Bossuyt (1996) collected data on 22 factors that describe the conditions of tree-related outages. In addition, various types of forest and vegetation simulators were used to estimate the number of trees that are either dead or in a declining condition, which have a higher probability of falling (Croft, 2008; Janney, 2008). The relationship between weather conditions and electric outages have also been studied in the literature (Guggenmoos, 2011). If the electrical network and trees are modelled as spatial objects in a spatial data model, tree-related outages can be considered as the outcome of their dependency: electric outages are caused by fallen trees. Most of the literatures on tree-related outages only consider attribute dependency between trees and electrical networks, and ignore topology aspects.

In this article, an SFID is proposed to identify areas in the electrical network that are vulnerable to tree-related outages. In the end, the results are validated against real tree-related outages data to demonstrate the suitability of the proposed method.

3.1 Factors causing tree-related outages

Factors that cause tree-related outages can be classified into four categories: 1) factors related to nature, 2) weather conditions in the area, 3) the condition of the electrical network, and 4) the risk distance from the electrical network (the so-called clear width). Considering that tree-related outages are typically due to falling trees, the characteristics of the condition of a tree, such as height, age, and species are the main factors that need to be taken into consideration. The stability of individual trees also heavily depends on root size, with larger roots leading to higher stability (Coutts *et al.*, 1999).

Soil types and tree species have a major impact upon the structure of the roots. Pyatt *et al.* (2001) proposed soil classification using three dimensions: climate, soil moisture regime, and soil nutrients. Crow (2005) further developed soil classification into seven classes according to how well the soil can support root growth. In Finland, the Scots pine (*Pinus Sylvestris*), the European white birch (*Betula Pubescens*), and the Norway spruce (*Picea Abies*) are the most common tree species. Table 1 illustrates how deep these three species' roots can grow in different soil types (this table is reworked from the study by Crow, 2005): a higher rank number represents shorter (poorer) roots. The ranks of pine, birch and spruce are totalled for each soil type category and the results are sorted.

The risk of tree-related outages is also directly related to the number of trees within striking distance of the electrical network (the so-called clear width). Guggenmoos (2003) proposed a probability chart and shows how the risk of tree-related outages changes with the clear width.

Table 1 somewhere here.

3.2 Applying spatial fuzzy influence diagrams to the problem of tree-related electric outages

The aim of this case study is to use the SFID to identify the locations in the electrical network that have a high risk of tree-related outages during extreme weather conditions. The following data were used in the study: forest resource data, soil type data, and electrical network data along with the locations of tree-related outages. The forest resource dataset is in raster format, containing information on height, age, and location of trees. It was retrieved from the National Resources Institute Finland (LUKE). Soil type data (also in raster format) were retrieved from the Geological Survey of Finland (GTK). This dataset originally categorises soil into 14 classes. The electric distribution network dataset and the network outage dataset were retrieved from our company partner. The electric distribution network dataset contains power line data from a specified area. The network outage dataset contains 223 fault locations, which describes registered tree-related outages within the study area from 2013 to 2015.

Figure 4 illustrates the SFID developed to model tree-related electric outage problems. In the first step, four spatial object chance nodes were created: a soil type node, a tree age node, a tree height node, and an electrical network node. In the second step, a decision node, *trees and soil are within the clear width*, was created and it was conditioned on the outcome of the chance nodes. This decision node indicates that the trees and soils that are located within a clear width of the electrical network are of interest. Guggenmoos (2003) found that the clear width of an electrical network should be between 10 and 18 meters according to a statistical analysis. In the tree data set, the tallest tree is 23.7 meters tall; therefore, the clear width limit was set at 30 meters, which also takes into account the error in the data set. One of the geoprocessing methods, buffer analysis, was used to create a 30 meter buffer zone around the electrical

network and select trees and soils located inside the buffer to be studied. This buffer analysis tool was represented as spatial value node in the diagram.

In the third step, a decision node, *hazard tree*, was created. It represents a tree that has poor roots and is located in an area with a soil type that is not good for tree growth, and is therefore at a higher probability of falling (see Table 1). On the other hand, the tree is only hazardous if it is located within the clear width of the network, so the classification of a hazard tree is made depending on the outcome of the decision node *trees and soil are within the clear width of the electrical network*.

The decision node *hazard tree* is defined by reclassifying the selected soil type, the tree's height, and age according to a risk value (1 to 5) – these reclassifications are shown in Table 2 and Table 3 for tree and soil properties, respectively. Guggenmoos (2003) found that the tree's mortality rate is positively correlated with the tree's age and starts to increase rapidly when the tree is over 30 years old. The increased mortality rate can be related to hazard trees; therefore, the boundary of the first age category is set at 30 years (Table 2). Trees have a higher risk value if the height is higher because it is impossible for small trees to fall onto the network if their height is lower than the height of the network. The quality of tree height reclassification can be improved in the future when there is the electrical network height data available.

Finnish soil types can be categorized based on how well they can support the root growth according to Crow's (2005) soil classification table (Table 1). Finnish soil type data was later merged into 12 classes according to their similarity and presented in the first column of Table 3. The second column is soil classification according to Crow's (2005) work, and the number refers to the *ID* number in Table 1. The third column refers to its corresponding suitability for root growth, and the value comes from a *rank value* column in Table 1. A *New Rank* column was created according to the value

of the *Old Rank* column, so the higher *Old Rank* value received a higher *New Rank*. The value in the *New Rank* column will be used for soil reclassification and it represents soil vulnerability, which causes trees to fall.

Table 2 somewhere here.

Table 3 somewhere here.

In the last step, SFID is extended with the fuzzy logic to model the vagueness of the decision. The decision node, *risk area*, represents vulnerable locations of the electrical network, which contains hazard trees within the clear width of the network. In addition, tree-related outages occur more often under extreme weather conditions, therefore, the weather condition chance node is also considered.

There are many vagueness factors involved in defining risk areas. For instance, it is difficult to say exactly how tall or how old a tree constitutes a high risk of hazard, or which types of soil will cause a tree to grow a poor root network. Therefore, the classifications of high or low risk network areas on the basis of tree and soil data sets should not have clear boundaries, and a fuzzy approach is more reasonable to use. Three value nodes were created to translate reclassified soil types, tree age, and tree height risk values into MFs. According to the fuzzy set theory (see section 2.2), variables x_1 , x_2 , and x_3 represent the reclassified soil type, tree age, and tree height data values. For each input variable, three triangular MFs were constructed and are illustrated in Figure 4. According to Equation 1, a , b and c are equal to 1.5, 3, and 4.5 respectively for the middle MF of soil type data. The peak value of the middle MF was defined as 3, which is the middle value of the input data that varies from 1 to 5. By doing so, low and high MFs can be placed on the left and right side of the middle MF's peak value symmetrically. The peak value of the low MF is assigned to be 1, which refers to the smallest input value. A degree of MF μ represents how likely it is that the input value

belongs to a corresponding MF. For instance, an input value of 1 will belong to a low MF since it doesn't touch the other MFs' curves at all, and the μ is equal to the maximum μ value of 1. In another example, the input value 1.7 touches both the low and middle MFs' curves, but the μ value for low MF is higher than for middle MF, which indicates the input value more likely belongs to low MF.

Finally, a value node *fuzzy output MFs* was used to describe output MFs (consequent part) of the fuzzy rule, which represents the vulnerability value of the electrical network segments (see fuzzy rules in Figure 4). Here vulnerability refers to the degree to which a system (e.g. electrical network segments) responds adversely to the occurrence of hazardous events. Each electrical network segment has its unique ID, location information, and vulnerability value. The unique ID is used as a key to incorporate the computation results into its corresponding electrical network segment on the map to observe the vulnerability of the electrical network segments spatially. In the *risk areas* decision node, 27 fuzzy rules were constructed and used to represent decision makers' decisions (Figure 4). Those MFs are used as inputs (antecedent part) of the fuzzy rules. The fuzzy OR operator, which returns the maximum value (or union) of the fuzzy MFs, was used because the research focus is to identify which places have a high risk of tree-related outages. This refers to a place, evaluated based on reclassified soil type, tree age, or tree height data values, that has more hazard trees.

Figure 5 illustrates an example of a surface view of the fuzzy rules where the X and Y axes represent the tree's height and soil type's reclassified risk value, respectively, and the Z axis represents the area's vulnerability value. The fuzzy surface view indicates that the vulnerability value of the network increases when the input variables' reclassified risk value increases. In this research work, expert knowledge that was used to create fuzzy MFs and fuzzy rules was obtained through a literature review

of the domain problem area. On the other hand, all the MFs, the MFs' shapes, and fuzzy rules are not developed by the actual domain experts, such as electricity companies and forestry department personnel, but it is possible that real domain expert knowledge could be explored for this purpose in the future.

Figure 4 somewhere here.

Figure 5 somewhere here.

4. Results

4.1. Results of the evaluation survey

Results of the questionnaire are illustrated in Table 4. The average time used to complete an influence diagram, with an average of 5 chance nodes, 1 value node, and 2 decision nodes, was 13.44 minutes. The average grade for measuring the usefulness of using influence diagrams for modelling spatial objects' dependencies was 3.1, which is higher than the middle of the grade range (1 to 5). The highest and the lowest grades that the participants gave were 4 and 1. Eighteen participants graded the usefulness measure higher than 3, one participant gave the measure a 2, and one participant gave the measure a 1.

Table 4 somewhere here.

4.2. Results of the case study

The results of the case study are represented as a vulnerability map (Figure 6), where the vulnerability of the electrical network segments refers to the value of fuzzy output MFs of the fuzzy rules in the SFID model. That is, the vulnerability is given as the degree that network segments will suffer from tree-related outages and is evaluated on the basis of the number of hazard trees located within the clear width of the segments.

On this map, the colour red represents extremely high vulnerability values and light green represents low vulnerability values. Note that the location of the case study is removed due to the data privacy agreement with our company partner in the ALVAR project, therefore the results can be only presented as a schematic map of the network.

To validate the model, the vulnerability map was overlaid with the timestamped tree-related outage data (fault data), and each fault location was annotated with its corresponding weather information. There are 223 fault data points. In Figure 6, numbers on the map refer to the fault locations with their corresponding wind speed (km/h). A visual comparison of the vulnerability map with these faults provides a possibility to observe the match between the predicted vulnerability value of the network and the fault location. Theoretically, the electrical network segments that have higher vulnerability values should refer to the locations of fault data, and therefore receive a yellow or a red colour on the map. According to the results shown in Figure 6, most of the electric fault locations are in the main branches and in the southern part of the electrical network. In some places, such as the southwestern and southeastern branches of the electrical network, there are good matches between fault values and extremely high vulnerability values.

To further validate the method, the percentages of the fault locations that received values for different vulnerability categories were calculated (Table 5). For this, the vulnerability raster map was first converted into points, where each point represented the centre location of the original raster cell. Then, a spatial join between the electrical network fault data and the point vulnerability map was created so that each fault data point would have all the attribute information of the nearest vulnerability map point. Finally, the fault locations that received extremely high vulnerability values were counted by selecting the fault points that have a vulnerability value greater than 0.5889

(see Figure 6). According to Table 5, 19% of fault points that were theoretically expected to receive at least a high vulnerability value received a low value instead. On the other hand, over half of the fault data received at least a high vulnerability value.

Table 5 somewhere here.

Figure 6 somewhere here.

5. Discussion

In this article, the general structure of an SFID is proposed, which can be used to model spatial objects' dependencies. The emphasis of an SFID is on the knowledge-based expert system design, which is a new way of structuring both software and information systems that closely match the human reasoning process.

Most participants (18 out of 20), after generating their own influence diagrams to solve a specific decision problem, favourably graded the usefulness of the influence diagram (grade 3 or higher). This indicates the potential of the influence diagram to facilitate the easier formulation of a decision problem. Two participants, however, evaluated the diagram unfavourably (grade 2 or less). They claimed that their low grades were due to the difficulty of understanding the influence diagram theory since none of them had used an influence diagram before. However, their performance in drawing an influence diagram contradicts this: both of them used less than 6 minutes to complete the task where the average time used to generate a diagram was 13.44 minutes. This suggests that the outlying results from these two participants are due to individual preferences and opinions rather than actual ability to generate an influence diagram.

It is also interesting to see the complexity of the participant-generated influence diagrams, as shown by the number of the nodes in the average participant's diagram.

According to the results, the average number of chance nodes, value nodes, and decision nodes are 5, 1, and 2, respectively, which indicates that most of the participants could formulate a reasonably complex decision structure even without being previously familiar with this representation. This, together with the participants' favourable perception of the influence diagram (an average grade over 3), indicates that an influence diagram could be a potentially useful way for structuring spatial decision problems. However, given the limited number of participants in our study, larger evaluation studies are necessary, with experts not only in GIScience, but also in various domains where SFID could be useful, such as the partners in the ALVAR project.

Moreover, some implications of the case study results related to identifying the vulnerability of the electrical network to tree-related outages are discussed. According to these results, the main branches and southern part of the electrical network received most of the electric faults. There are 15 fault points with a low wind speed value but high segment vulnerability value in the northeastern branch of the electrical network. This indicates there are many hazard trees in this area and since trees can fall onto a power line with even low wind speed, more investigation should take place in the future.

The network outage dataset contains only 223 fault locations and many vulnerable locations of the electrical network do not have fault data available to validate, which causes inaccuracy in the validation results. Therefore, more fault data should be included in the future to make the validation process more complete and accurate. In addition, a new literature study of how Finnish soil types support tree root growth should be developed in order to improve the soil type risk reclassification, as well as take into consideration more tree species. SFID is an expert based system, so fuzzy MFs and fuzzy rules should be developed by electricity management and forestry

experts in the future. In this article, weather data is mainly used to visually explore how fault data is related to corresponding weather conditions. In the future, dynamic weather data can also be used as input parameters to SFID in order to make the results more realistic.

From a methodology point of view, the SFID represents a set of geoprocessing tools used to model topological dependencies of spatial objects. When the number of spatial objects increases, traditional software tools used to implement research work will not be able to handle large raster datasets. Work is ongoing to evaluate cutting edge cyberGIS approaches (Wang *et al.* 2013; Wang *et al.* 2015) to enable SFID for both computation-intensive and data-intensive problem solving. It is also useful to consider sensitivity analysis as part of SFID to see how results change according to the variations in expert knowledge.

6. Conclusion

In this article, a new methodology, SFID, was proposed to model spatial object dependencies by taking into consideration both the topological relationship and the attribute dependency of the object. Compared to traditional influence diagrams and fuzzy influence diagrams, the spatial context of the objects can be incorporated into the SFID, which provides the possibility to visualize the results on a map and answer the question: *where* are the areas that are at risk? For instance, buffer analysis was used to create a 30-meter clear width buffer zone around the electrical network since trees are hazardous to the electrical network only if they are within the buffer. In addition, the structure of an SFID is easily adapted for different applications. Decision nodes were added to the diagram to enable decisions to be made as the outcome of the diagram. Fuzzy logic can systematically emulate human reasoning, and allows for experts'

knowledge to be incorporated into the modelling process. Moreover, fuzzy logic was also used to model the vagueness of spatial objects. Compared to spatial fuzzy modelling, SFID uses the structure of an influence diagram to help decision makers formulate decision problems, which are related to spatial objects' dependencies. The usefulness of using an influence diagram in this study is evaluated by using a questionnaire and most of the participants thought influence diagrams could be a useful way to model spatial objects' dependencies.

SFID was tested with a tree-related electric outages case study, and the results showed that SFID is a suitable method to model tree-related electric outages. As a result, vulnerable locations of an electrical network were visualised on a map in order to support electrical network maintenance and planning. For instance, decision makers can prioritise areas for a tree-trimming programme or optimise repair resources in the case of outages caused by extreme weather conditions.

Acknowledgements

We would like to acknowledge Tero Kauppinen and Christian Fjäder who from the National Emergency Supply Agency of Finland for funding this research work. We would like to express thanks to Jari Ahlstedt who from Caruna for providing us help in the case study. We would like to thank our colleagues, Hannes Seppänen and Pekka Luokkala, for their good cooperation during the project. We would like to acknowledge the colleagues from CyberInfrastructure and Geospatial Information Laboratory(CIGI), University of Illinois, Urbana-Champaign, PhD. students from Department of Nuclear, Plasma, and Radiological Engineering, University of Illinois, Urbana-Champaign, visiting professors from Dalian university of technology (P. R China), Chinese Academy of Science, Xihua University (P.R. China), for participating in answering the

evaluation survey. Finally, we would like to thank Becky Vandewalle who did the language editing for the article.

References

- Aspinall, R., 1992. An inductive modelling procedure based on Bayes' theorem for analysis of pattern in spatial data. *International Journal of Geographical Information Science*, 6(2), 105–121.
- Coutts, M.P., *et al.*, 1999. The development of symmetry, rigidity and anchorage in the structural root system of conifers. *Plant and Soil*, 217, 1–15.
- Croft, J., 2008. East Cascades (EC) Variant overview, forest vegetation simulator [online]. Department of agriculture, forest service, United States. Available from: http://www.fs.fed.us/fmnc/ftp/fvs/docs/overviews/FVSec_Overview.pdf [Accessed 15 March 2016]
- Crow, P., 2005. The influence of soils and species on tree root depth [online]. Forestry Commission. Available from: [http://www.forestry.gov.uk/pdf/FCIN078.pdf/\\$FILE/FCIN078.pdf](http://www.forestry.gov.uk/pdf/FCIN078.pdf/$FILE/FCIN078.pdf) [Accessed 15 March 2016]
- Chang, S., *et al.*, 2007. Infrastructure failure interdependencies in extreme events: power outage consequences in the 1998 Ice Storm. *Natural Hazards*, 41, 337–358.
- Espada, R., *et al.*, 2015. Vulnerability assessment and interdependency analysis of critical infrastructures for climate adaptation and flood mitigation. *International Journal of Disaster Resilience in the Built Environment*, 6 (3), 313–346.

- Fisher P., Comber A., and Wadsworth R., 2006. Approaches to uncertainty in spatial data. *Fundamentals of spatial data quality*, R. Devillers and R. Jeansoulin eds., ISTE: London, 43-59.
- Guggenmoos, S., 2003. Effects of tree mortality on power line security. *Journal of arboriculture*, 29(4).
- Guggenmoos, S., 2011. Tree-related electric outages due to wind loading. *Arboriculture & Urban Forestry*. 37(4), 147-151.
- Howard, R. A., and Matheson, J. E., 2005. Influence diagrams. *Decision Analysis*, 2(3), 127-143.
- Janney, G., 2008. *The Southern (SN) Variant overview, forest vegetation simulator* [online]. Department of agriculture, forest service, United States. Available from: http://www.fs.fed.us/fmsc/ftp/fvs/docs/overviews/FVSSn_Overview.pdf [Accessed 15 March 2016]
- Kaufmann, A., and Gupta, M. M., 1985. *Introduction to fuzzy arithmetic: theory and applications*. New York: Van Nostrand Reinhold.
- Kripakaran, P., Vigneshan, M., Asvanth, S., Sivachandiran, B., 2014. *Performance of Power System Stabilizer using Fuzzy Logic Controller*. 9(3), 42-49.
- LUKE (The Natural Resources Institute Finland), 2012. *State of Finland 2012: Finnish forests in European context demonstrated with selected indicator* [online]. Available from: <http://www.metla.fi/metinfo/sustainability/finnish.htm> [Accessed 15 March 2016]
- Longley, P., et al., (2005). *Geographic information systems and science*. John Wiley & Sons.

- Li, L., et al., (2010). Using spatial analysis and Bayesian network to model the vulnerability and make insurance pricing of catastrophic risk. *International Journal of Geographical Information Science*, 24(12), 1759-1784.
- McDaniels, T., et al., 2007. Empirical framework for characterizing infrastructure failure interdependencies. *Journal of Infrastructure Systems*, 13(3), 175-184.
- Mateou, N. H., et al., 2005. Fuzzy Influence Diagrams: an alternative approach to decision making under uncertainty. *Proceedings on the Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference*.
- Neapolitan, R.E., 2004. *Learning Bayesian Networks*, Chicago: Pearson Prentice Hall.
- Pyatt, G., et al., 2001. *An ecological classification for forestry in Great Britain*. Edinburgh: Forestry Commission.
- Rees, W.T., et al., 1994. Priority trimming to improve reliability. *In ISA conference presentation*, Halifax, Nova Scotia.
- Rinaldi, S., et al., 2001. Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies. *IEEE Control Systems Magazine*, 21 (6), 11-25.
- Rautvuori M., 2014. ALVAR-toimikunnat rasvaavat alueellisen varautumisen rattaita. National Emergency Supply Agency. Available from web https://www.varmuudenvuoksi.fi/aihe/toimintaymparisto/113/alvar-toimikunnat_rasvaavat_alueellisen_varautumisen_rattaita (Cited 21.3.2017)
- Simpson, P., and Bossuyt, R. V., 1996. Tree-Caused Electric Outages. *Journal of Arboric*. 22(3).

- Stassopoulou, A., *et al.*, 1998. Application of a Bayesian network in a GIS based decision making system. *International Journal of Geographical Information Science*, 12(1), 23-46.
- Sicat, R. S., Emmanuel J. M. C., and Uday B., N., 2005. Fuzzy modeling of farmers' knowledge for land suitability classification. *Agricultural systems* 83(1), 49-75.
- Sawhney, A., Bhatia, R., and Mahajan, P., 2014. Congestion control in wireless communication network using fuzzy logic and machine learning techniques. *International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering*, 3(11).
- Tatman, J. A., and Shachter. R. D., 1990. Dynamic programming and influence diagrams. *IEEE transactions on systems, man, and cybernetics*, 20(2).
- Teodorović D., 1999. Fuzzy logic systems for transportation engineering: the state of the art. *Transportation Research*, part A (33), 337-363.
- Utne, I. B., *et al.*, 2011. A method for risk modelling of interdependencies in critical infrastructures. *Reliability Engineering and System Safety*, 96, 671-678.
- Wang, S., *et al.*, 2015. Open CyberGIS Software for Geospatial Research and Education in the Big Data Era. *SoftwareX*, 5:1-5
- Wang, S., *et al.*, 2013. CyberGIS Software: A Synthetic Review and Integration Roadmap. *International Journal of Geographical Information Science*, 27(11): 2122-2145
- Zadeh, L.A., 1965. Fuzzy sets, *Information and control*, 8(3), 338-353
- Zhao, J., and Bose, B. K., 2002. Evaluation of membership functions for fuzzy logic controlled induction motor drive. In IECON 02, IEEE 2002 28th Annual Conference of the Industrial Electronics Society, 1, 229-234.

- Zhu, X., *et al.*, 1996. ILUDSS: A knowledge-based spatial decision support system for strategic land-use planning. *In Computers and Electronics in Agriculture*, 15, 279-301.
- Zhu, A. X., Hudson, B., Burt, J., Lubich, K., & Simonson, D. (2001). Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Science Society of America Journal*, 65(5), 1463-1472.
- Zhi, Z., and Jin, F., X., 2010. Application study on fuzzy influence diagram in highway tunnel construction safety risk assessment. In proceeding of 2010 International Conference on Management Science & Engineering.
- Zedeh L. A. 1996. The roles of fuzzy logic and soft computing in the conception, design and deployment of intelligent systems. *BT technology journal*, 14(4), 32-36
- Zimmermann, H.J., 2001. *Fuzzy set theory and its application*. London: Kluwer Academic.
- Zhang, Z., Demsă U., Rantala J., and Virrantaus K., 2014. A fuzzy multiple-attribute decision making modelling for vulnerability analysis on the basis of population information for disaster management. *International Journal of Geographical Information Science*. Vol. 28 (9), 1922-1939

Tables

Table 1. Soil type classification and its corresponding suitability for tree root growth (Crow, 2005). Root size for pine, birch, and spruce tree according to soil classification. 1: root size <4 m, 2: root size < 3 m, 3: root size < 2 m, 4: root size < 1.5 m, 5: root size <0.5 m, 6: not good for growth

ID	Soil Type	Pine	Birch	Spruce	Total Score	Final Rank
1	Loose, deep well drained soils	1	3	3	7	1
2	Shallow soils over rock	5	5	6	16	7
3	Intermediate loamy soils	3	3	3	9	2
4	Impervious subsoils	3	5	5	13	4
5	Soils with moisture retaining upper horizons	6	4	5	15	6
6	Soils with wet lower horizons	6	4	4	14	5
7	Organic rich soils (drained)	4	4	4	12	3
8	Organic rich soils (waterlogged)	6	6	5	17	8

Table 2. Tree height and tree age data with reclassification.

Tree Height (meters)	Tree Age (years)	Reclassified Value (risk value)
0-5	1-30	1
5.1-10	31-80	2
10.1-15	81-130	3
15.1-20	131-180	4
20.1-23.7	181-239	5

Table 3. Soil type data with reclassification.

Soil Type	Soil Classification (Table 2)	Old Rank (Table 2)	New Rank (risk value)
Fragmented rock, bedrock physical trauma	2	7	5
Gravel and sandy till	1	1	1
Silt moraine	5	6	4
Hill moraine	1	1	1
Moraine esker or moraine covered gravel and sand deposit	1	1	1
Esker outside gravel and sand deposit	1	1	1
Litho gravel and sand deposition	1	1	1
Stream deposition	6	5	3
Homogeneous clay and silt deposition	5	6	4
Clay and silt composition	5	6	4
Peat composition	7	3	2
Bedrock outcrops	2	7	5

Table 4. Questionnaire results.

	Number of Chance Nodes	Number of Value Nodes	Number of Decision Nodes	Time	Grade	GIS working experience (years)
Mean	5	1	2	13.44	3.1	4.175
Range	9	3	7	40	3	13
Median	5	1	1	11	3	2.5
Mode	4	1	1	10	3	0

Table 5 Validation results of the case study.

Predicted Vulnerability	Number of Fault Points	Percentage of Fault Data
Low	42	19%
Middle	61	27%
High	80	36%
Extremely High	40	18%

Figures

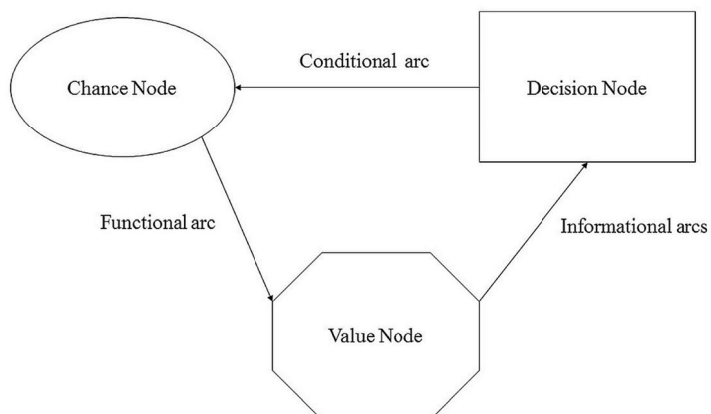


Figure 1. General structure of an influence diagram.

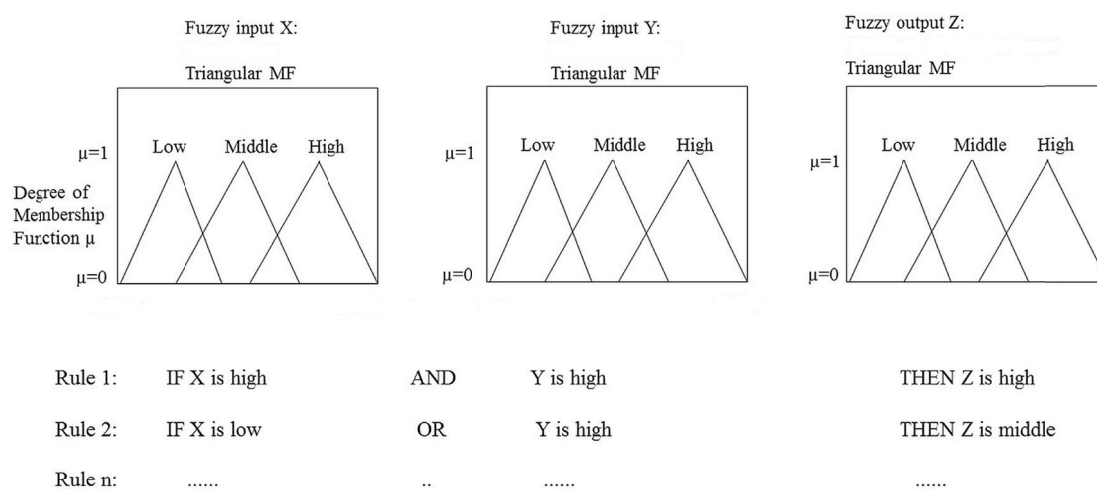


Figure 2. Examples of fuzzy membership functions and fuzzy rules.

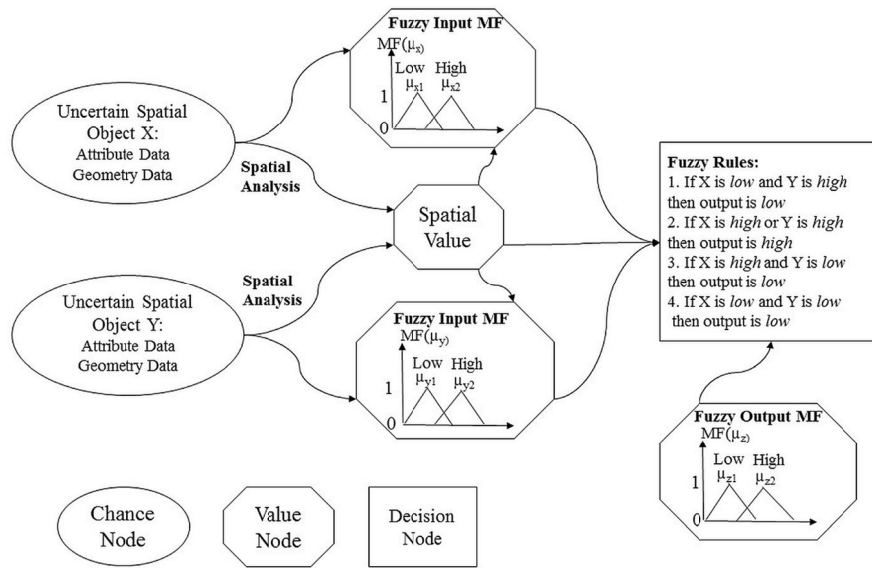


Figure 3. General structure of a spatial fuzzy influence diagram.

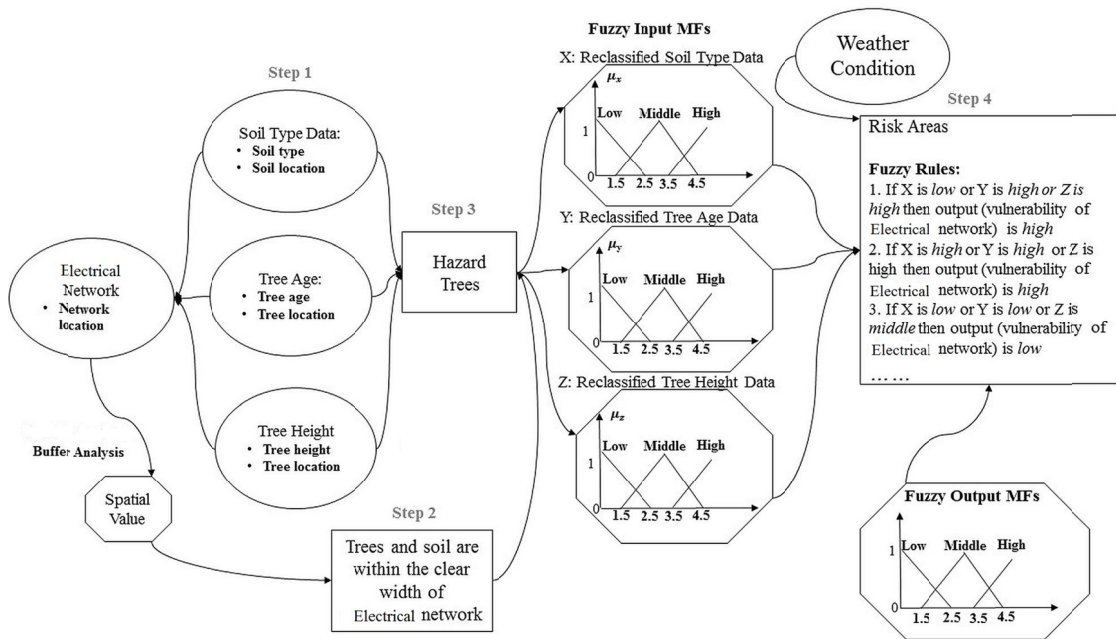


Figure 4. An example of a spatial fuzzy influence diagram modelling tree-related electric outages.

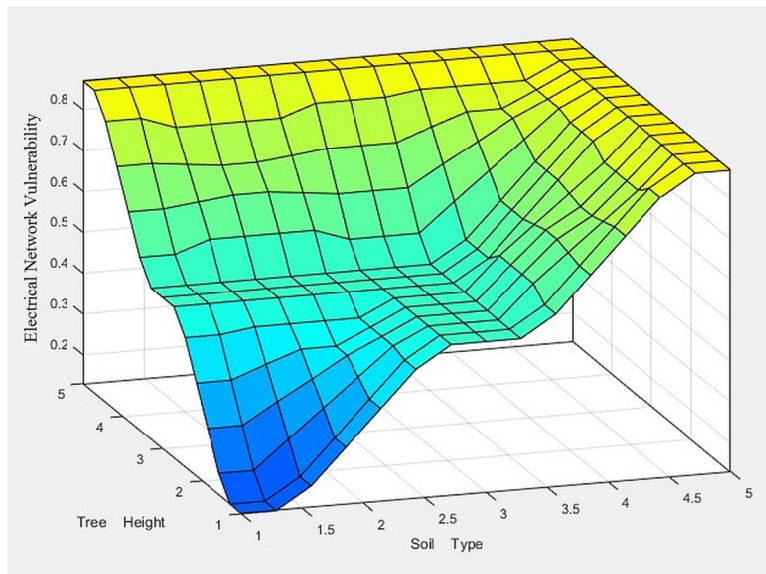


Figure 5. Surface view of fuzzy rules. Input attributes (x and y axes) are tree height and soil type reclassified risk value, and the output attribute (z axis) is the vulnerability of the electrical network.

