evropský
sociální
fond v ČR

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

Querying on Fuzzy Surfaces with Vague Queries

Jan Caha, Jiří Dvorský

StatGIS 2013

StatGis Team

Table of Contents

Introduction

- **information** used for modelling is often **uncertain** (ie. incomplete, vague, . . .)
- it is necessary to include the uncertainty of **inputs** and processes to asses uncertainty of **outputs**
- **queries** to the data can also be **vague** - ie. definition of "steep slopes"
- applications in **geography** and **geoinformatics** - for example fuzzy surfaces and their analyses

Fuzzy sets and Fuzzy numbers

- special cases of sets that does not have strictly defined criterion of membership
- transitional interval where objects have increasing or decreasing value of membership
- **fuzzy number** is a special case of fuzzy set that represents vague, imprecise or ill-known value
- conditions that fuzzy number has to satisfy - defined on $\mathbb{R}$, normal and convex fuzzy set, membership function must be at least piecewise continuous
- if the conditions are satisfied then fuzzy number can be used for calculation in so called **fuzzy arithmetic**

Ranking of fuzzy numbers

- several possible ways for ranking fuzzy numbers
- the most complex method of ranking fuzzy number is done in the framework of **Possibility theory** (Dubois and Prade, 1983)
- Possibility theory (Dubois and Prade, 1986) is an extension of fuzzy set theory and offers a new ways to handle uncertain data
- ranking system uses **two measures** - **possibility** and **necessity**
- 4 indexes - possibility and necessity of **at least equality** and possibility and necessity of **strict exceedance**

Ranking of fuzzy numbers - equations

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

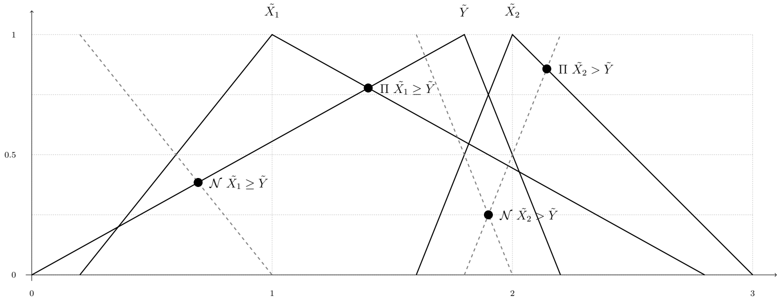
$$\mathcal{N}_{\tilde{X}}([\tilde{Y},\infty)) = \inf_{x} \max(1-\mu_{\tilde{X}}(x), \sup_{y \leq x} \mu_{\tilde{Y}}(y))$$

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

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

Ranking of fuzzy numbers - example

- the questions “Is $\tilde{X}_1$ equal or greater than $\tilde{Y}$?”
- and “Is $\tilde{X}_2$ strictly greater than $\tilde{Y}$?”



Vagueness in geography

- widely discussed topic in geography
- within it is domain there are many concepts that are **naturally vague**
- applications to modelling fuzzy geographical regions, surfaces with uncertainty and decision making based on vague datasets
- **reliability** and **certainty** of data and its analyses should be questioned
- problem of querying the data - almost no **flexibility**
- flexibility of queries is a necessary concept for more **complex decision support**

Soft spatial queries

- classic queries cannot introduce any measure of preference in the result, because they are based on **boolean logical expressions**
- spatial data are not well suited for such crisp queries, because these **query** might be far to **restrictive**
- reasons for creating **soft threshold** can be summarized as following:
  - the concept of Y is naturally fuzzy i.e. definition of "steep slopes"
  - there are more than one acceptable definitions of Y and there is no indications that any of them is more correct or precise than the others
  - Y is based on the expert opinion that is provided as an interval of values rather then precise value, or there is a need to merge definitions of Y from several such expert opinions

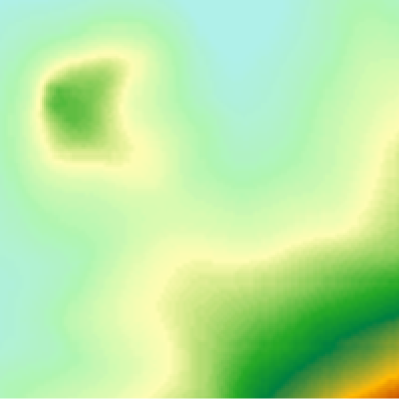
Fuzzy surfaces

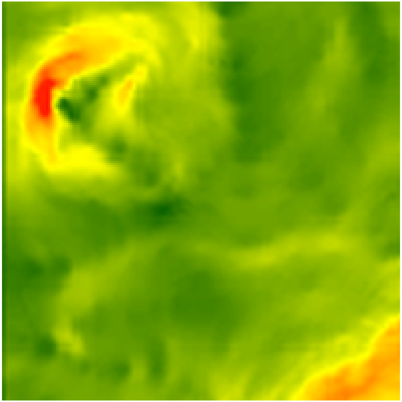
- in GIS the surface is represented by the mathematical function $Z = f(x, y)$
- **fuzzy surface** incorporates uncertainty of input data and/or processes of interpolation
- as a result there is a **fuzzy number** for each pair of coordinates $\tilde{Z} = f(x, y)$

Case study

- How can **soft queries** be used to query **fuzzy surfaces**?
- find out areas with higher than “medium slope” on a fuzzy surface
- ie. part of a complex decision of finding suitable areas for waste disposal site
- no universal definition of “medium slope”, definition according to the expert opinion as a triangular fuzzy numbers with support values $[5^\circ, 9^\circ]$ and kernel value 7°
- using **4 indexes** for ranking fuzzy numbers identify areas that **fulfil the criterion**

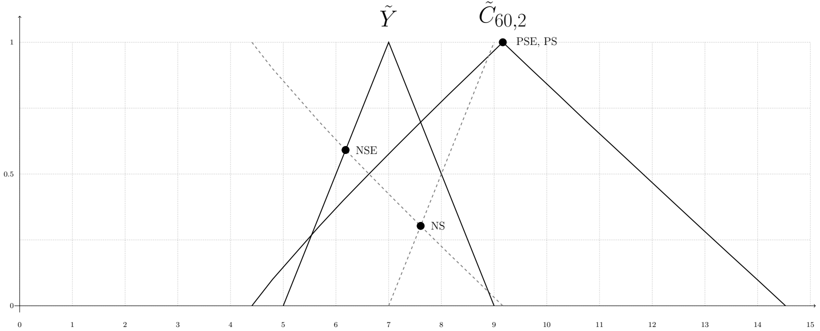
Case study - input data





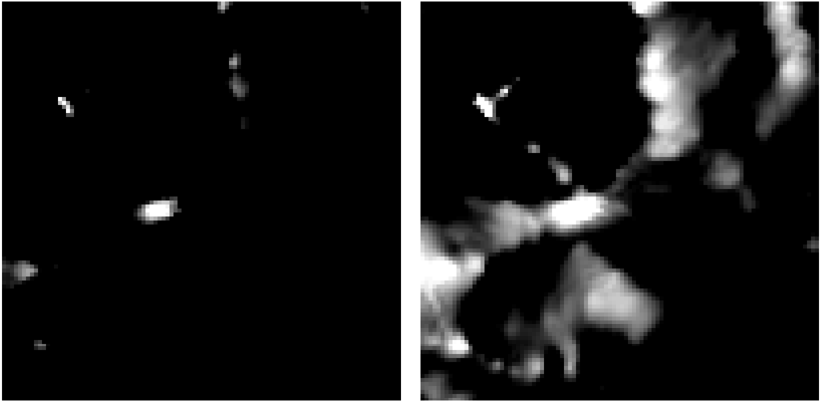
Visualization of kernel values of the fuzzy surface (left) and fuzzy slope (right). Terrain colours are from light blue (low values) to brown (high values), while slope is from green (low) to red (high).

Case study - processing the data



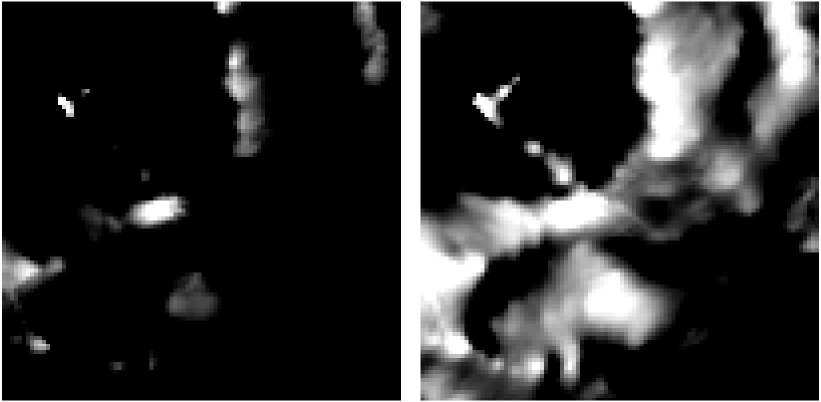
Four indices comparing one cell of the grid $\tilde{C}_{60,2}$ to the threshold $\tilde{Y}$.

Case study - results



Visualization of possibility (left) and necessity (right) of exceedance of the threshold. Black colour means value 1 and white 0.

Case study - results



Visualization of possibility (left) and necessity (right) of strict exceeder of the threshold. Black colour means value 1 and white 0.

Conclusions

- use of **fuzzy mathematics** for **uncertainty propagation** is not common for geographical problems
- the proposed method for querying fuzzy surfaces with **vague queries** should allow better handling of fuzzy data and show that it is improvement to te existing solutions
- approach offers **more information** for the **decision maker** when the uncertainty of data is present and the decision criterion does not have to be specified exactly
- four indices form natural ordering - the possibility of exceedance is the less restrictive, followed by the possibility of strict exceedance, then the necessity of exceedance and the necessity of strict exceedance is the most restrictive

Conclusions

- approach extends possibilities of **spatial decision support** by the use of **mathematical methods for handling uncertainty and vagueness** that is very common in geographical sciences
- further studies should be focused on **more complex decision making problems**

Thank you for your attention.