Credibility Distributional Grade Geostatistics for Spatial Inequalities

Presentation: IBS Conference, Brazil
5th – 10th December 2010

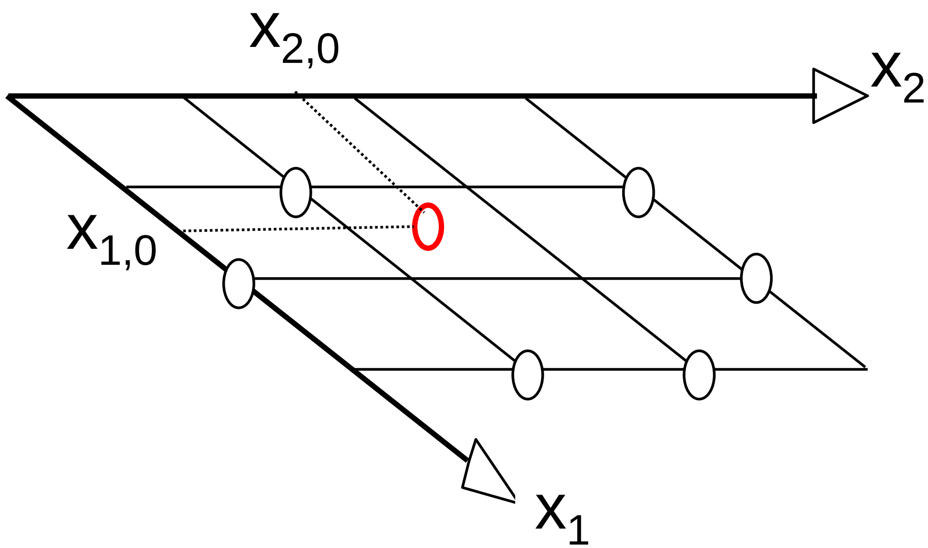
Acknowledge
Prof Renkuan Guo
Prof Christien Thiart
University of Cape Town, South Africa

Njeri Wabiri, Ph.D.
Human Sciences Research Council
South Africa
nwabiri@hsrc.ac.za or njeri.wabiri@yahoo.com

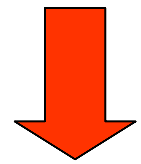
Introduction

- Context
- Fuzzy Spatial Data
- Credibility Measure Theory // Probability Measure Theory
- Credibility Geostatistics
- Conclusion

Spatial Estimation Problem ????



- What is the dependency between observed values?

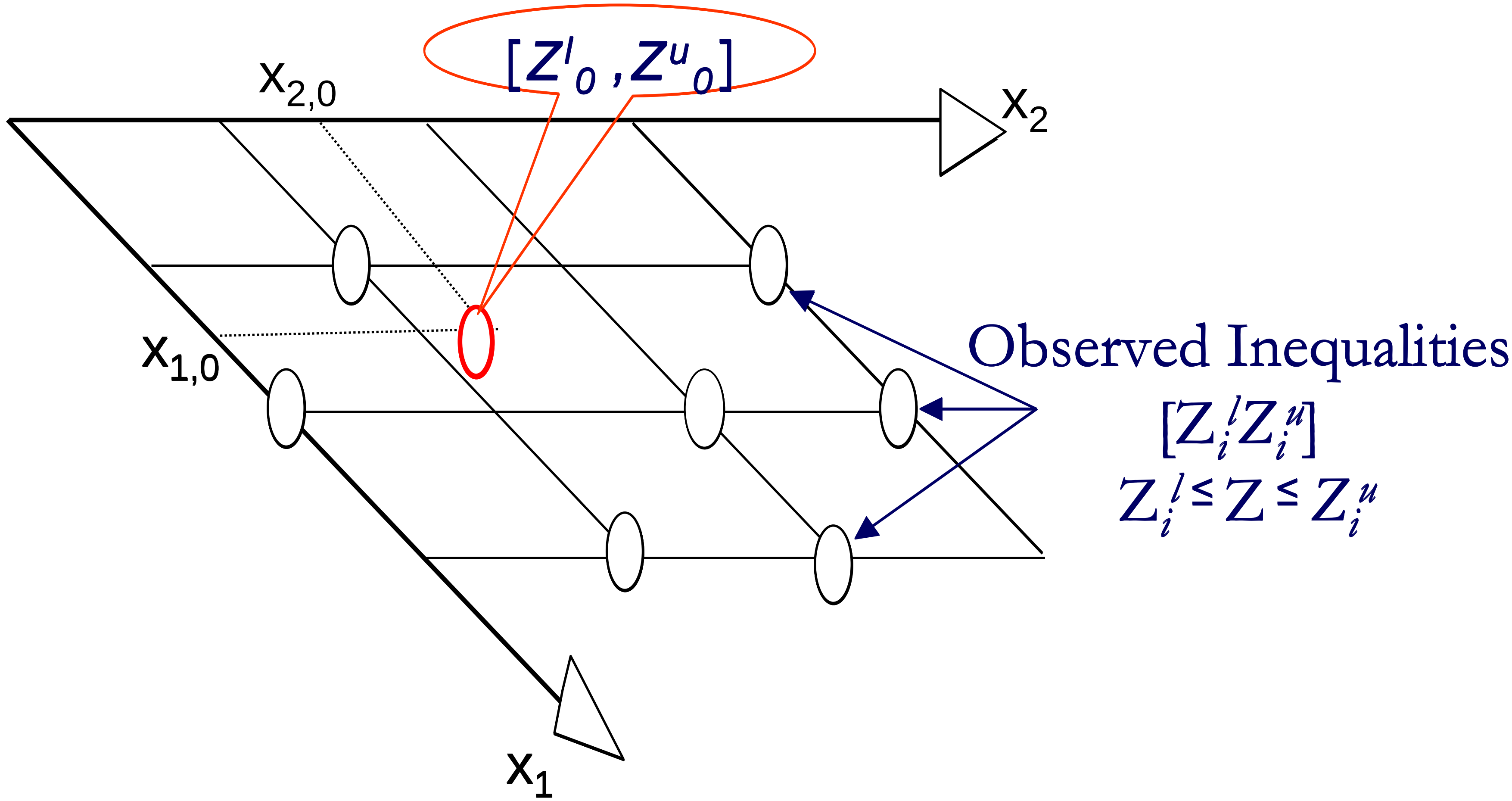


How to estimate at intermediate unknown spatial coordinates ?

Goal:

- Generate information with spatial continuity, using samples at selected locations
  - Pollution prediction maps, crucial for decision-making.
  - Early warnings on increase of gamma dose levels above certain thresholds.

Imprecise Spatial Point Data



- Inequalities are the Basic presentation of uncertainty
- Captures randomness and fuzziness in real life measurements
- Other presentation of imprecise spatial point data: pdf, fuzzy sets

Modeling With Spatial Inequalities

- Use précis number (mean, median, quartiles)-loose information
- Interval Analysis –lacks gradation (Diamond, 1988)

Better Approach

- Interval or Inequality can be view as a fuzzy set
- Apply the basic concept of fuzzy mathematics (Kaufmann, 1975; Zadeh, 1965, 1978).
- Set-based operations and set-based outputs- Mathematics complex

???? simpler approach

Basics

- Probability theory: Random variable
  Probability measure (self-duality property)
  Probability distribution function
  Real-valued operations /Outputs

- Fuzzy methods theory: (Random) fuzzy events, fuzzy random events
  Possibility measure (no self-duality property)
  Membership function
  Set-based operations/ set-based outputs

Zadeh (1965, 1978), Kaufmann (1975)

Fuzzy Geostatistics

- Fuzzy set operations,
- Apply Membership function (subjective);
- Set-based predictions; Difficult with GIS

Bandemer and Gebhardt, 2000; Bardossy et al.,1988; Burrough and McDonnell, 1998; Kacewiez, 1994

Our Approach- Two steps

Random interval set →	Fuzzy (variable) set →	Scalar Fuzzy variable
Closed Random Interval Theory	Membership function Possibility Measure (Lacks self duality)	Credibility Distribution Credibility Measure (Self Duality)
Interval Arithmetics	Fuzzy set Mathematics	Real-valued Mathematics

Credibility Measure Theory

Let Θ denote a nonempty set, with corresponding power set 2^Θ.
We refer to the elements $B \in 2^\Theta$ as events. In addition, let $\mathrm{Cr}(B)$ denote a number assigned to event B such that $0 \leq \mathrm{Cr}(B) \leq 1$. The number $\mathrm{Cr}(B)$ indicates the credibility that the event B occurs. $\mathrm{Cr}(B)$ satisfies the following axioms Liu (2006):

Axiom 1. $\mathrm{Cr}(\Theta) = 1$

Axiom 2. $\mathrm{Cr}(\cdot)$ is non-decreasing, i.e. $\mathrm{Cr}(B) \leq \mathrm{Cr}(C)$ for $B \subseteq C$, $C \in 2^\Theta$.

Axiom 3. $\mathrm{Cr}(\cdot)$ is self-dual, i.e. $\mathrm{Cr}(B) + \mathrm{Cr}(B^c) = 1$ for $B \in 2^\Theta$.

Axiom 4. $\mathrm{Cr}\{\cup_i B_i\} \wedge 0.5 = \sup\left[\mathrm{Cr}\{B_i\}\right]$ for any $\{B_i\}$ with $\mathrm{Cr}(B_i) \leq 0.5$

Axiom 5. Assume that a given set of functions $\mathrm{Cr}_k(\cdot) : 2^{\Theta_k} \to [0, 1]$ satisfy Axioms 1-4, and $\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_q$, then for each $(\theta_1, \theta_2, \ldots, \theta_q) \in \Theta$

$$\mathrm{Cr}\left(\theta_1, \theta_2, \ldots, \theta_q\right) = \mathrm{Cr}_1\{\theta_1\} \wedge \mathrm{Cr}_2\{\theta_2\} \wedge \ldots \wedge \mathrm{Cr}_q\{\theta_q\}$$

Credibility Measure Space

Any set function $\text{Cr}: 2^{\Theta} \to [0,1]$ satisfying Axioms 1-4 is called a $(\wedge,\vee)-$Credibility measure, and the triplet $(\Theta, 2^{\Theta}, \text{Cr})$ is referred to as the $(\wedge,\vee)$-Credibility Measure Space

A fuzzy variable, ξ, is a mapping from the credibility space $(\Theta, 2^{\Theta}, \text{Cr})$ to a set of real numbers

Credibility Distribution of Fuzzy variable ξ

The credibility distribution $\Phi: \Re \to [0,1]$ of the fuzzy variable ξ on $(\Theta, 2^{\Theta}, \text{Cr})$ is:

$$\Phi_{\xi}(z) = \text{Cr}\{\theta \in \Theta : \xi(\theta) \le z\}$$

Represents cumulated credibility grade of the fuzzy variable, ξ, taking values less or equal to $z \in \Re$

Parallel to a random variable, a fuzzy variable is fully described by its credibility distribution function

….continued

The induced membership function of the fuzzy variable ξ on $\left(\Theta, 2^{\Theta}, \mathrm{Cr}\right)$ is

$$\mu_{\xi}(z) = \left(2\mathrm{Cr}\{\xi = z\}\right) \Lambda 1, \; z \in \Re$$

Fuzzy set/credibility measure theory

- Membership function: not correct starting point set-theoretical foundation of the fuzzy mathematics
- Proposed possibility measure, assumed counterpart of probability measure, lacks self-duality

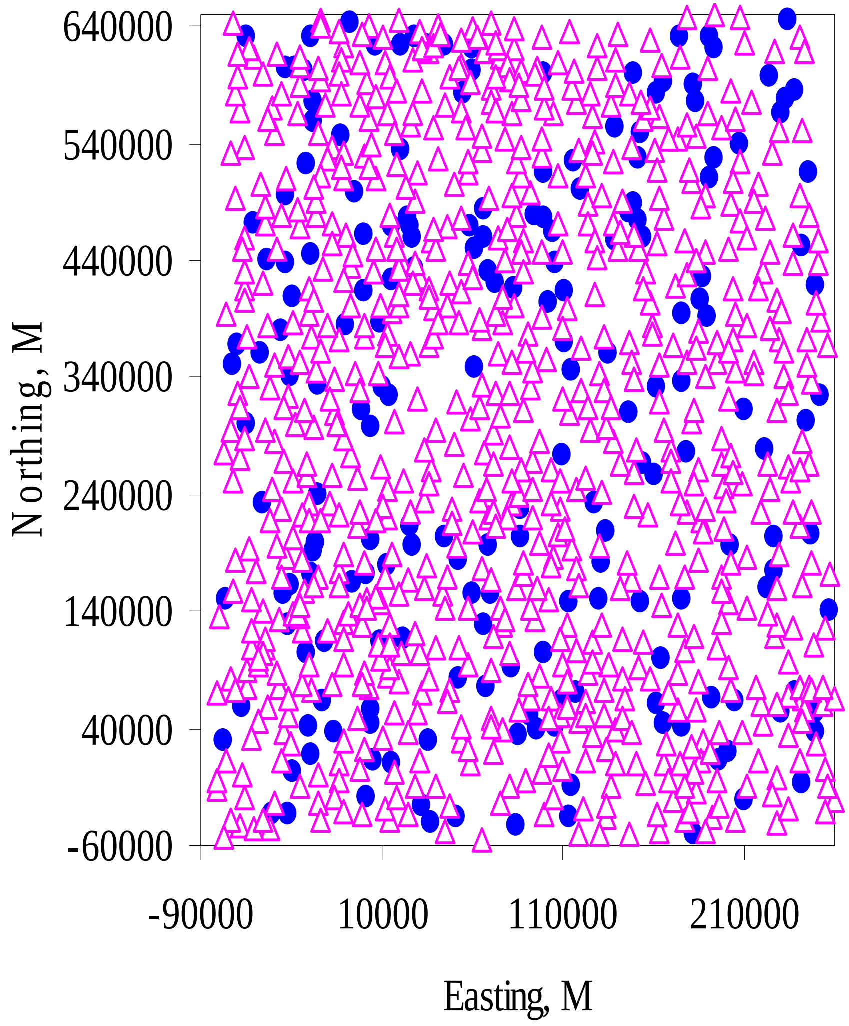
- Axioms of credibility measure provides a set-theoretical foundation for fuzzy variables
- Fuzzy variable should be characterized by its credibility distribution first.

- As a tradition, we provide fuzzy variable membership function; an *induced* function, conventional and convenient mathematical language for describing the fuzzy phenomenon.

An Example:

- Mean gamma dose rates of natural ambient radioactivity in Germany (inequalities)
- Total of 1008 monitoring stations
- Spatial inequalities at 200 monitoring stations. (Blue)

$$\left\{ \left(z_i^l, z_i^u \right), i = 1, 2, \ldots, n \right\}$$



Induced Maximum entropy fuzzy Variable

$$\mu_{\xi([z^l, z^u])}(z) = \int_{-\infty}^{z} \int_{z}^{\infty} p(z^l, z^u) dz^u dz^l$$

Maximum entropy data-assimilated credibility distribution function

$$\Phi_{\xi}(z) = \frac{1}{2}\left(\mu_{\xi([z^l, z^u])}(z) + 1 - \sup_{y \neq z}\left[\mu_{\xi([z^l, z^u])}(y)\right]\right)$$

Credibility spatial random function

For a given fuzzy variable ξ *with credibility distribution* Φ_ξ*, if* $\xi = z_i$ *at location* $\boldsymbol{x}_i = (x_i, y_i)$*, then* $\Phi_\xi(z_i)$ *is called the credibility grade for fuzzy variable* ξ *at location* (x_i, y_i)*. The collection of spatially distributed credibility grades, denoted as* $\{\Phi_\xi(z_i),\ \boldsymbol{x}_i \in D \subset \mathbb{R}^2, i = 1, 2, \cdots, n\}$*, is called sampled credibility grades over re-gion* D*. The credibility grades range from 0 to 1 and forms an alternative generalization to 0/1 indicator codes as used in indicator kriging*

Credibility grade geostatistics

Sample Credibility grade Semivariogram

$$\hat{\gamma}_{\Phi}(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \left[\left(\Phi\left(z(\boldsymbol{x}_i + h)\right) - \Phi\left(z(\boldsymbol{x}_i)\right) \right)^2 \right]$$

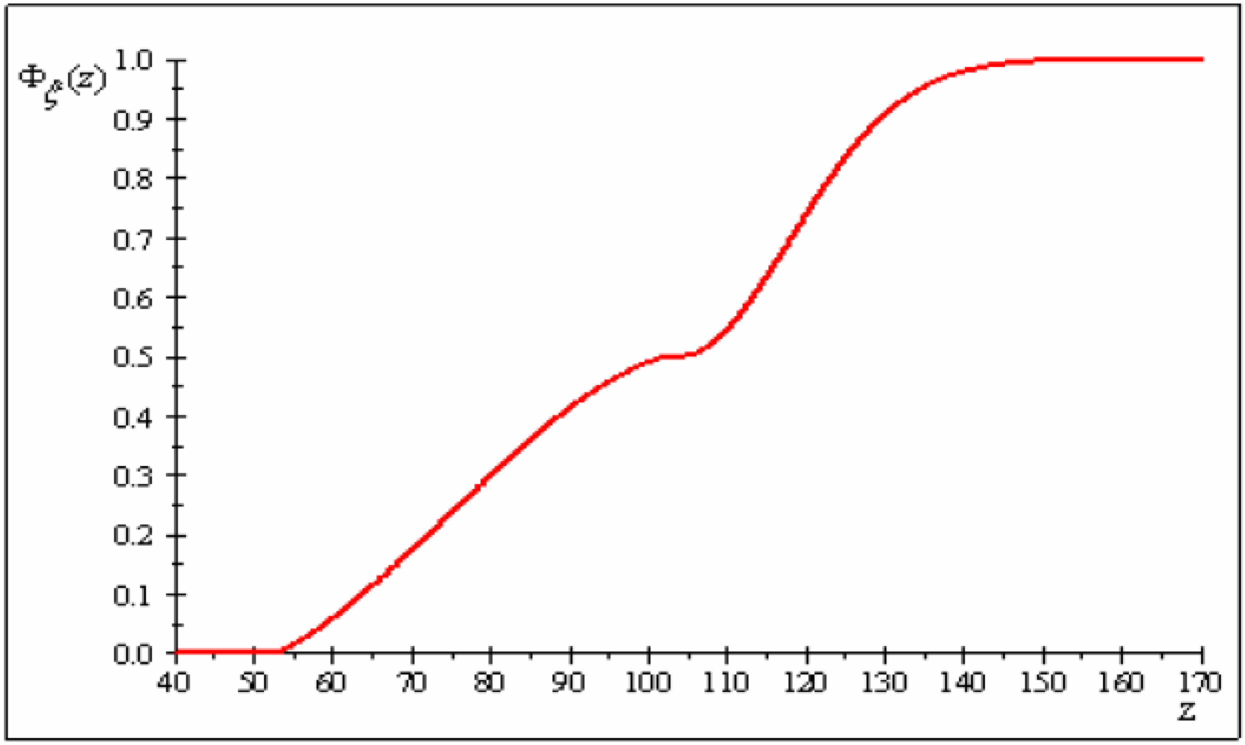
Credibility grade kriging system

$$\begin{aligned} \sum_{j=1}^{n(h)} \lambda_j \gamma_{\Phi}(\boldsymbol{x}_i - \boldsymbol{x}_j) + \psi &= \gamma_{\Phi}(\boldsymbol{x}_0 - \boldsymbol{x}_i) \quad i = 1, \ldots, n(h) \\ \sum_{j}^{n(h)} \lambda_j &= 1 \end{aligned}$$

Credibility grade predictor

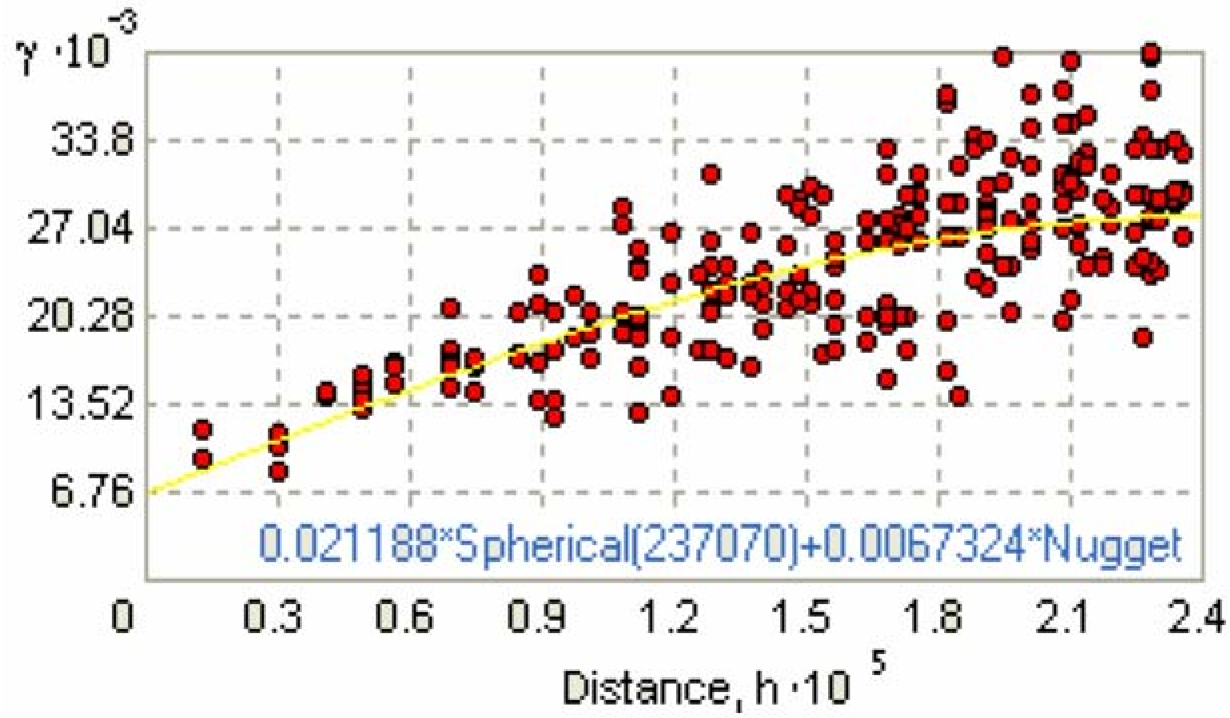
$$\hat{\Phi}\left(z(\boldsymbol{x}_0)\right) = \sum_{i=1}^{n(h)} \lambda_i \Phi\left(z(\boldsymbol{x}_i)\right), \quad \sum_{i}^{n(h)} \lambda_i = 1$$

Credibility distribution



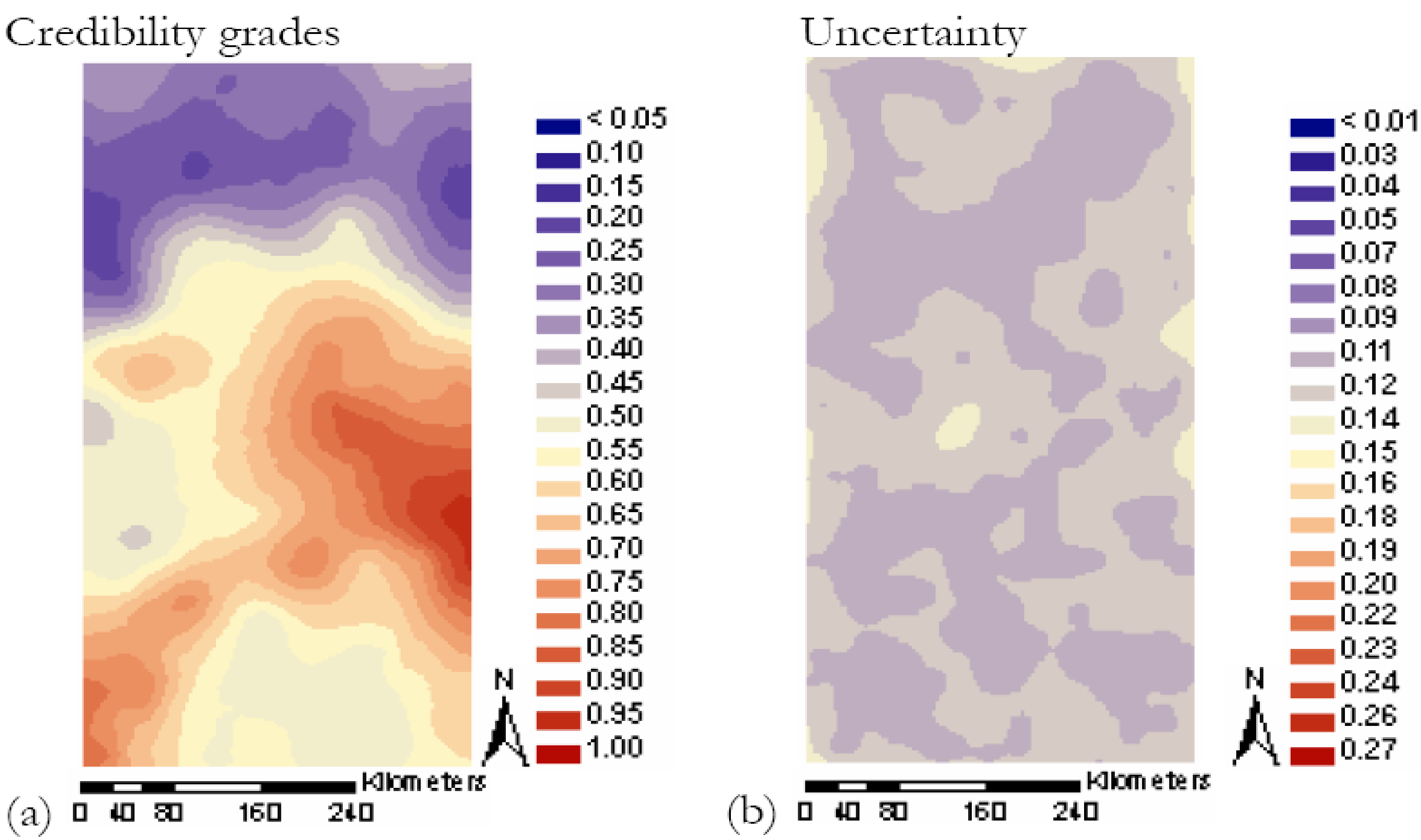
$$\Phi_\xi(z) = \begin{cases} 0 & \text{if } z < 52.795 \\ \frac{1}{2}\left(1.9517 - 0.1097z + 0.0018z^2 - 0.000008z^3\right) & \text{if } z \in [52.795, 102.186) \\ \frac{1}{2} & \text{if } z \in [102.186, 103.780] \\ 1 - 0.50016\exp\left(-\frac{(z-103.73288)^2}{408.89}\right) & \text{if } z \in (103.780, 157.000] \\ 1 & \text{if } z > 157.000 \end{cases}$$

Credibility grade variogram



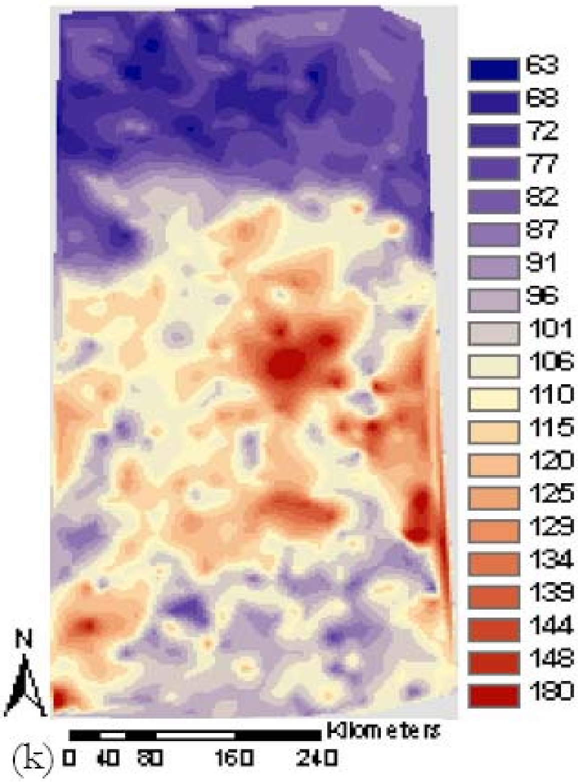
Increases slowly from the origin; an indication of smooth imprecise random process

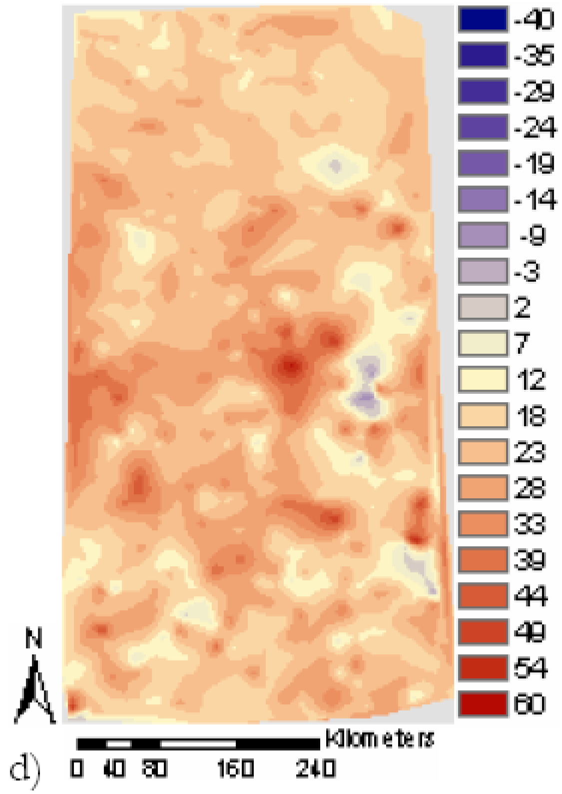
Pollution map and Error Map



ME= 0.0003; RMSE=0.1145

Ordinary kriging of central values





ME= -0.13; RMSE=11.97

QUESTIONS