Credibility Kriging of Spatial Inequalities

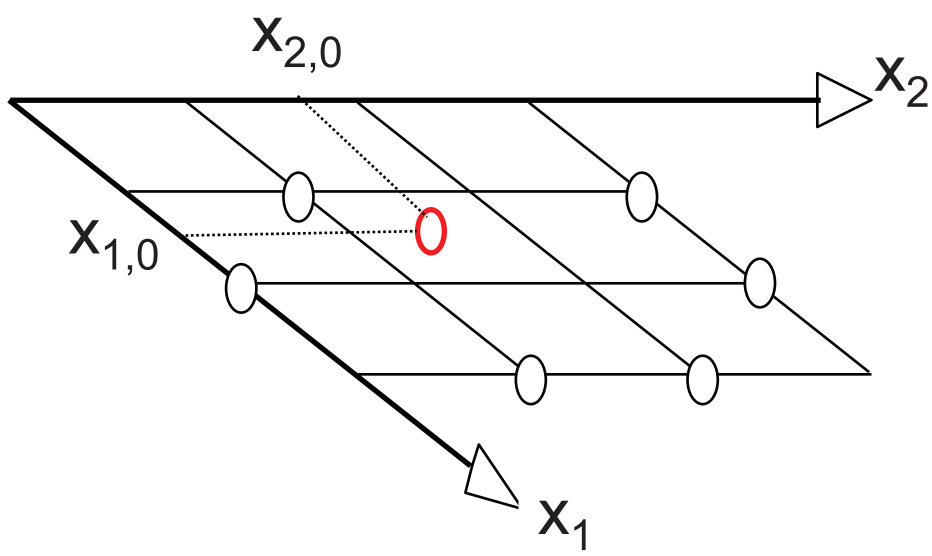
Njeri Wabiri, Ph.D.
Human Sciences Research Council
South Africa

Presentation At the TWOWS 4th General Assembly and International Conference, Beijing 25th-28th June 2010

Introduction

- Context
- Fuzzy Spatial Data
- Credibility measure theory
- Credibility Geostatistics
- Conclusion

Spatial Estimation Problem



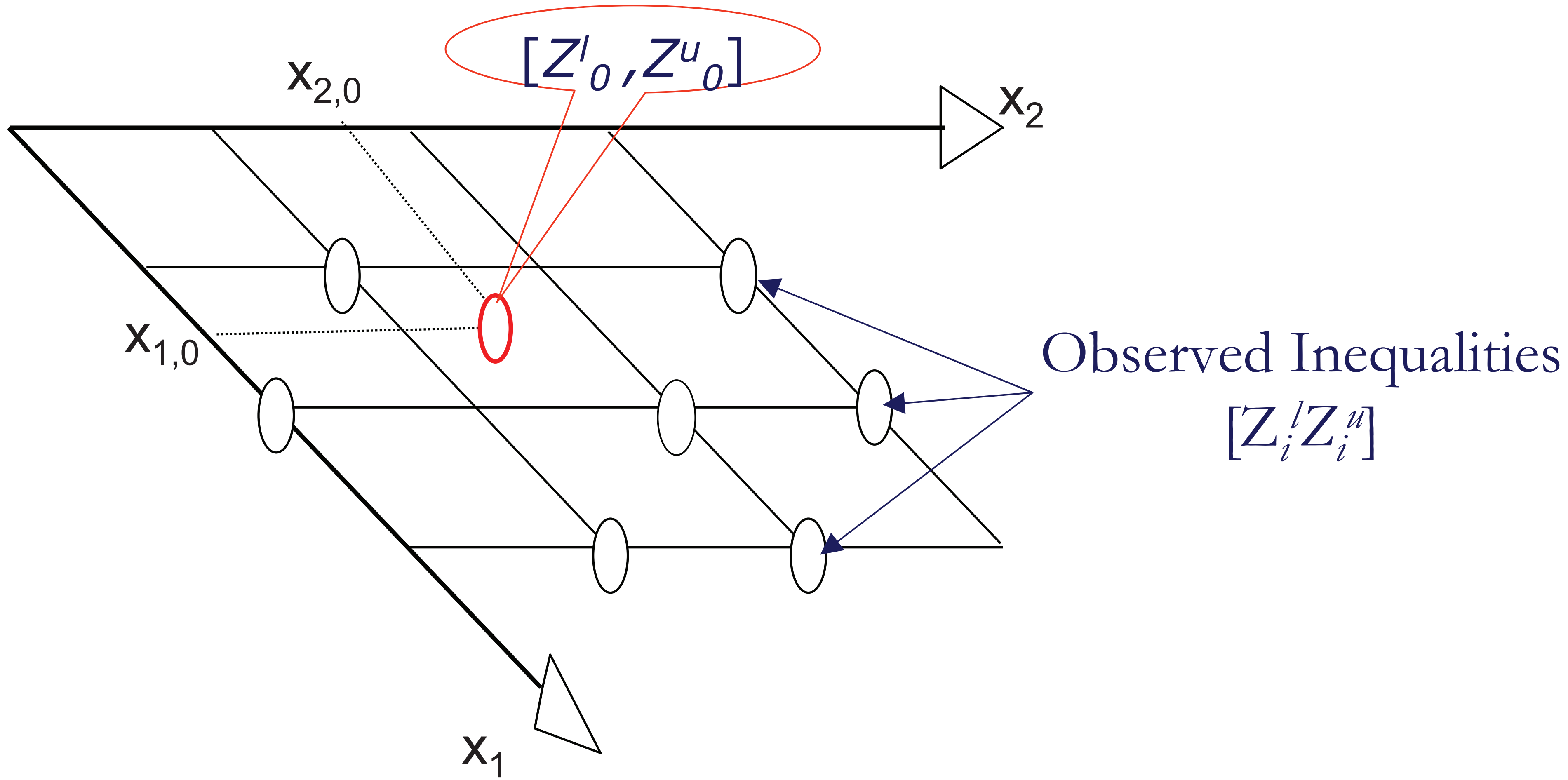
- What is the dependency between observed values?

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.

Uncertain/Imprecise Spatial Data



- Spatial inequalities are the Basic presentation of uncertainty
- Captures randomness and fuzziness in real life measurements

Modeling With Spatial Inequalities

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

Better Approach

- Apply the basic concept of fuzzy mathematics (Kaufmann, 1975; Zadeh, 1965, 1978).

Note--

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

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

  Zadeh (1965, 1978), Kaufmann (1975)

- Fuzzy geostatistical: Fuzzy set operations, difficult with GIS

  Bandemer and Gebhardt, 2000; Bardossy et al.,1988,1990b;
  Burrough and McDonnell, 1998; Kacewiez, 1994

Our Approach- Two steps

Random interval set	→	Fuzzy variable set	→	Scalar Fuzzy random variable
Interval Arithmetics Diamond, 1988		Membership function Possibility Measure		Credibility Distribution Credibility Measure (Self Dual)
		Fuzzy set Mathematics		Credibility Measure theory

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 $\mathrm{Cr}_k(\cdot) : 2^{\Theta_k} \rightarrow [0, 1]$ 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 Cr: $2^{\Theta} \rightarrow [0, 1]$ satisfying Axioms 1-4 is called a $(\vee, \wedge)$-Credibility measure, and the triplet, $(\Theta, 2^{\Theta}, \mathrm{Cr})$ is referred to as the $(\vee, \wedge)$-credibility measure space

Credibility Distribution

A fuzzy variable, ξ, is a mapping from the credibility space $(\Theta, 2^{\Theta}, \mathrm{Cr})$ to a set of real numbers. Parallel to a random variable, a fuzzy variable is fully specified by its credibility distribution function.

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

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

This represents the accumulated credibility grade that the fuzzy variable ξ takes a value less than or equal to $z \in \Re$.

Membership function

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

$$\mu_{\xi}(z) = (2Cr\{\xi = z\}) \wedge 1, z \in \Re$$

Random interval set and Fuzzy variable set

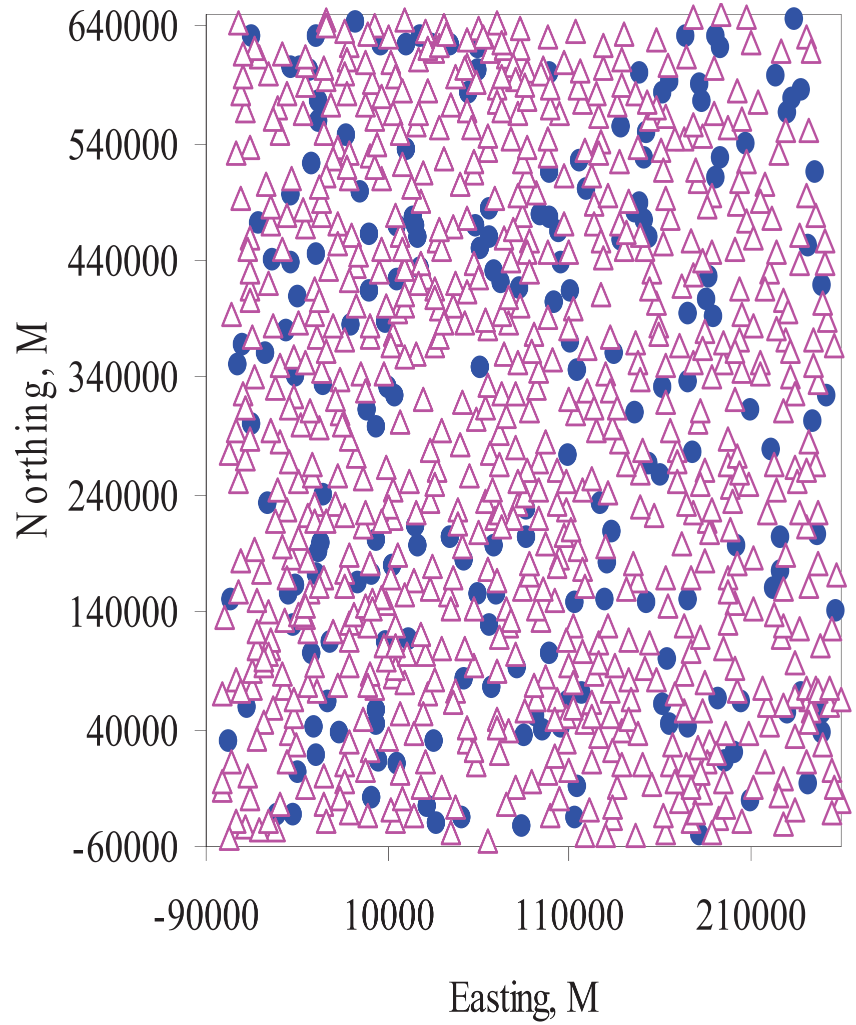
let $\tilde{\varphi}(Z)$ denote the fuzzy variable generated by the random interval Z, then its membership function is expressed by:

$$\mu_{\tilde{\varphi}(Z)}(z) = \int_{-\infty}^{z} \int_{z}^{\infty} f(z^l, z^u) dz^u dz^l \ , \forall z \in \tilde{\varphi}(Z)$$

An Example:

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

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



Fuzzy variable set with data-assimilated membership function:

$$\mu_{\tilde{\varphi}([Z^l,Z^u])}(z) = \int_{-\infty}^{z}\int_{z}^{\infty} p_k(z^l, z^u)dz^u dz^l$$

Data-assimilated credibility distribution:

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

Credibility 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

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

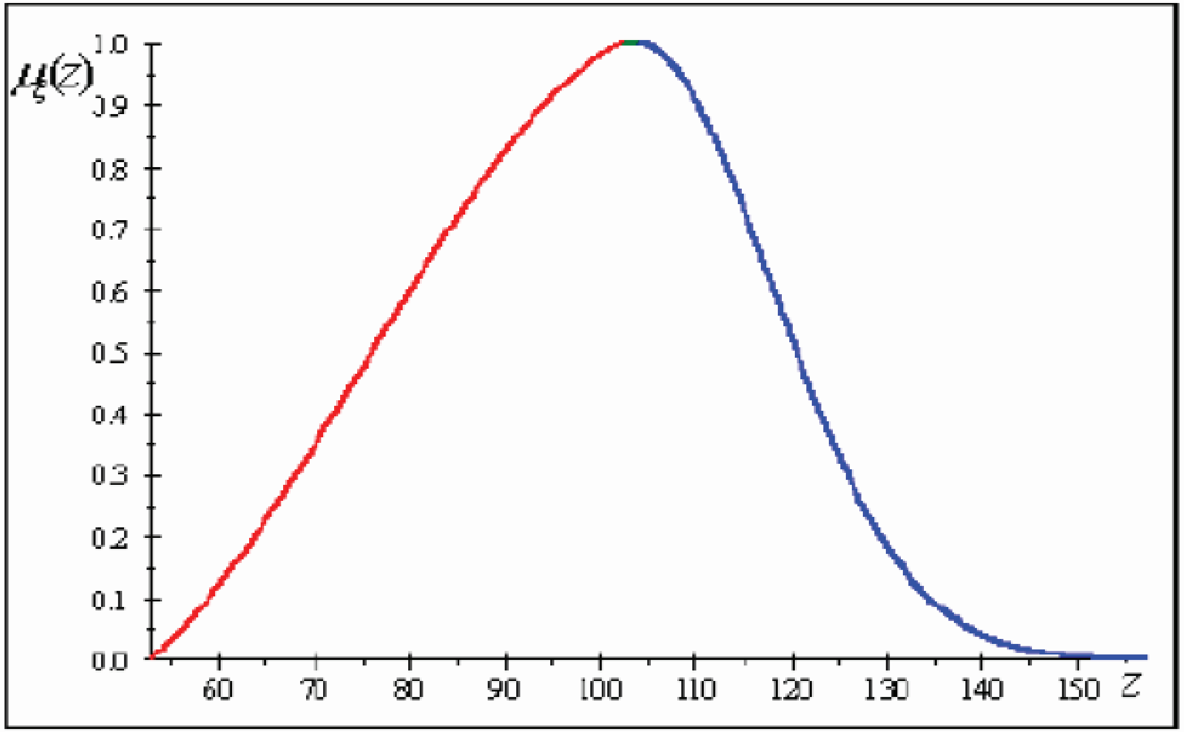
Credibility grade kriging system

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

Credibility grade predictor

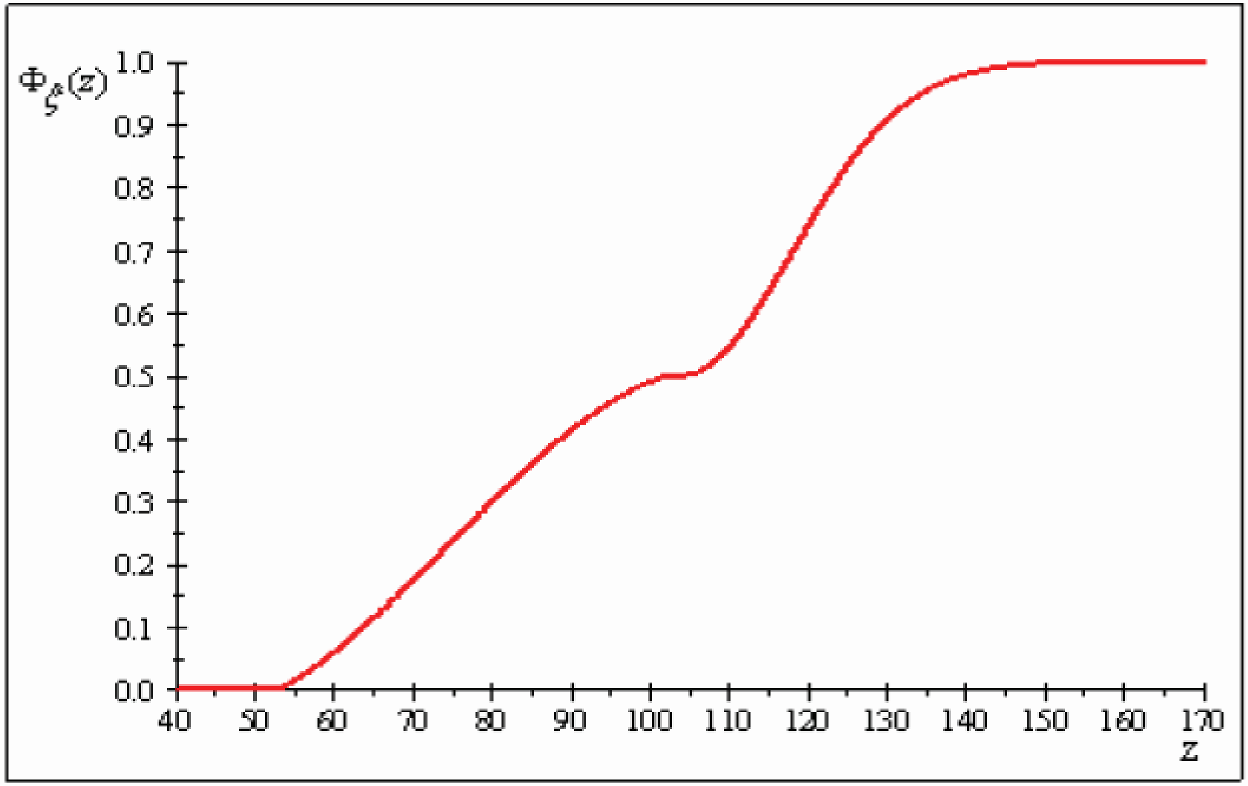
$$\hat{\Phi}\left(z(\boldsymbol{x}_0)\right) = \sum_{i=1}^{n(h)} \lambda_i \Phi\left(z(\boldsymbol{x}_i)\right), \quad \sum_{i}^{n(h)} \lambda_i = 1$$

Results: Membership function



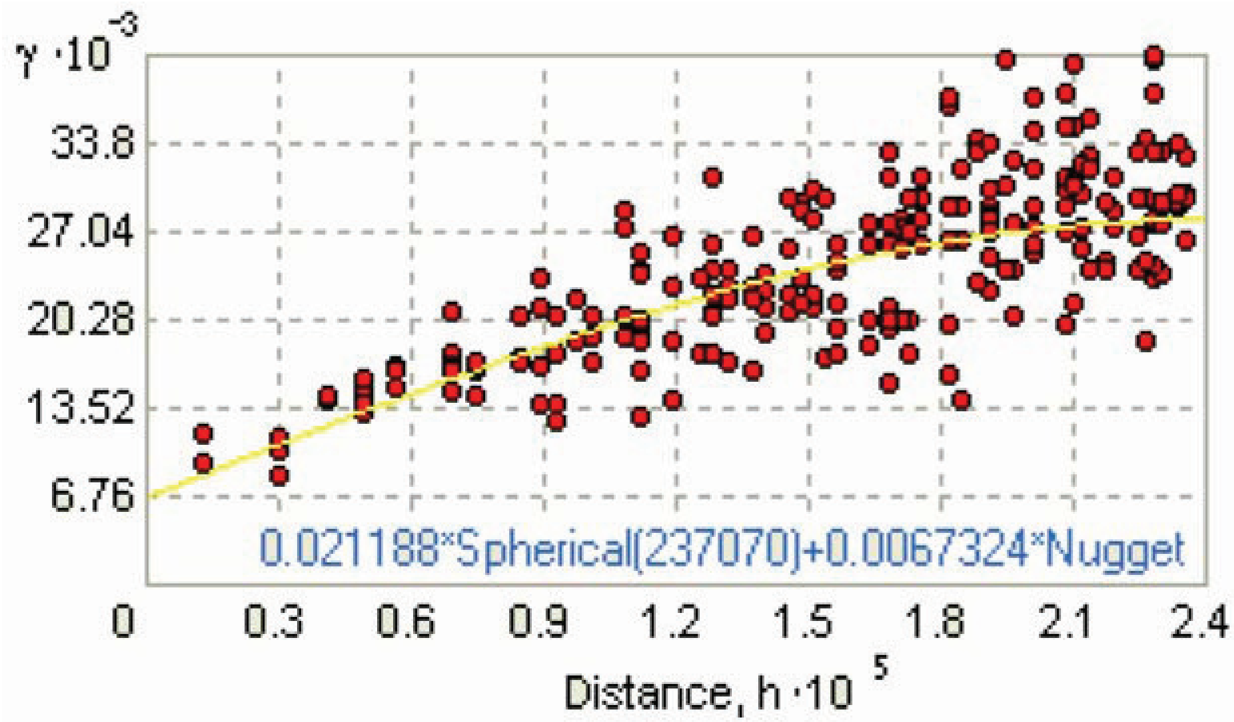
$$\mu_\xi(z) = \begin{cases} 1.9517 - 0.1097z + 0.0018z^2 - 0.000008z^3 & \text{if } 52.795 \leq z < 102.186 \\ 1 & \text{if } 102.186 \leq z \leq 103.780 \\ 1.0003\exp\left(-\frac{(z-103.7329)^2}{408.89}\right) & \text{if } 103.780 < z \leq 157 \\ 0 & \text{otherwise} \end{cases}$$

Credibility distribution



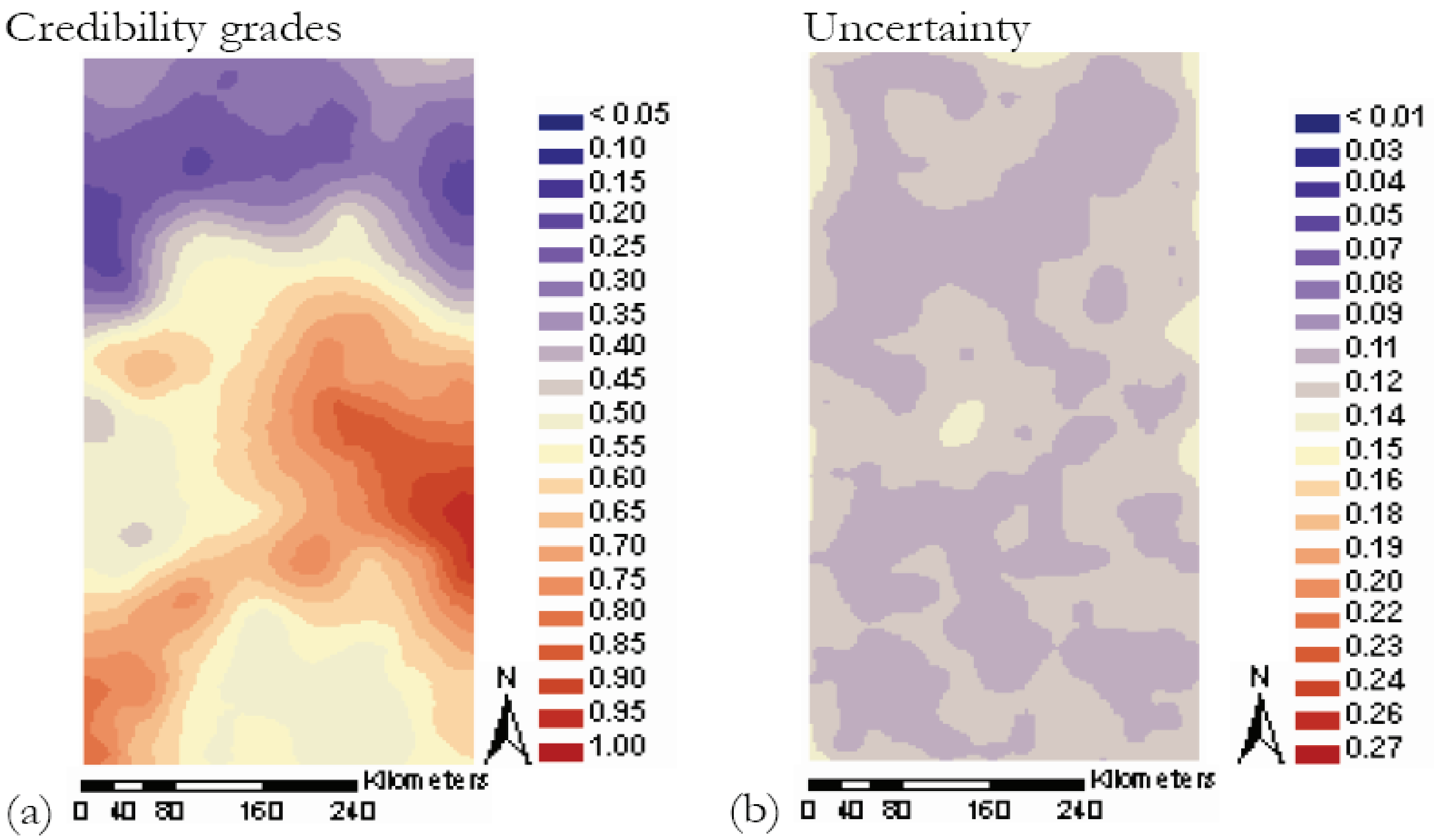
$$\Phi_\xi(z) = \begin{cases} 0 & \text{if } z < 52.795 \\ \frac{1}{2}\left(1.9517 - 0.1097z + 0.0018z^2 - 0.000008z^3\right) & \text{if } z \in [52.795, 102.186) \\ \frac{1}{2} & \text{if } z \in [102.186, 103.780] \\ 1 - 0.500\,16\exp\left(-\frac{(z-103.73288)^2}{408.89}\right) & \text{if } z \in (103.780, 157.000] \\ 1 & \text{if } z > 157.000 \end{cases}$$

Credibility grade variogram



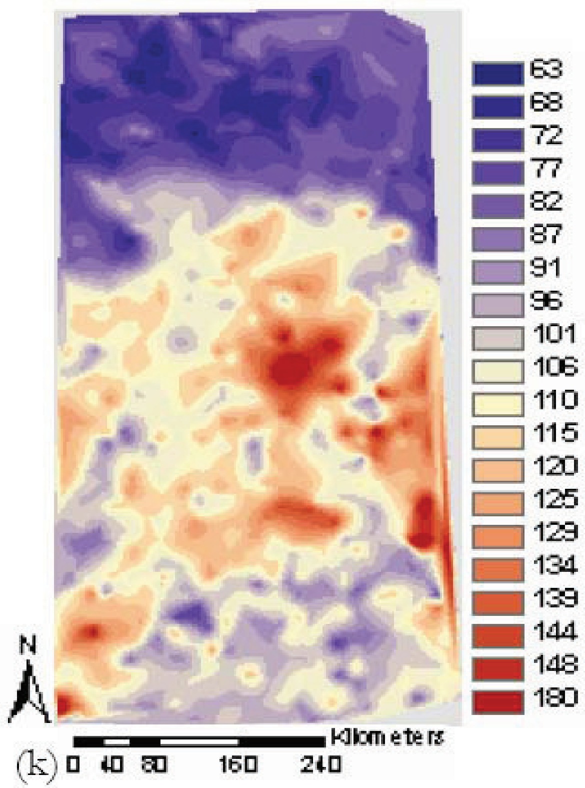
- Increases slowly from the origin; an indication of smooth random process

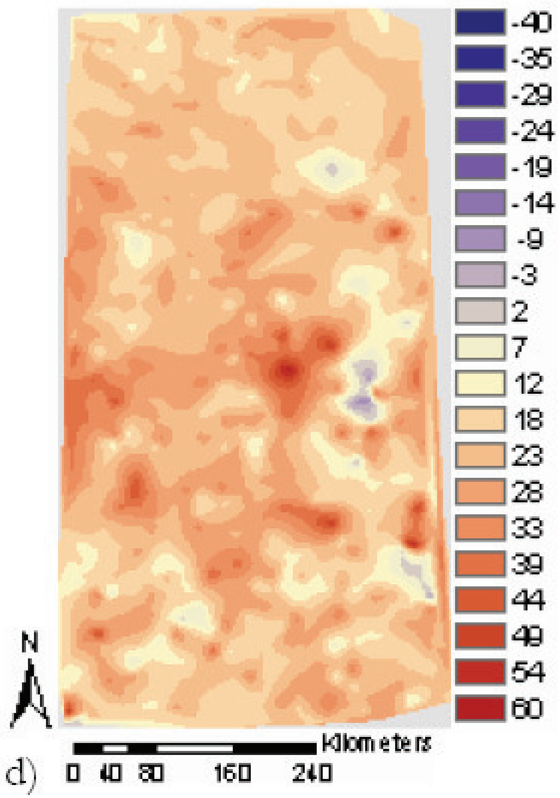
Predicted pollution map



ME= 0.0003 and RMSE=0.1145

Ordinary kriging of central values





ME= -0.13

RMSE=11.97

QUESTIONS