Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Master of Science in Geospatial Technologies

Geostatistics
Predictions with Geostatistics

Carlos Alberto Felgueiras
cfelgueiras@isegi.unl.pt

Predictions with Geostatistics

Contents

Predictions with Geostatistics

Deterministic x Stochastic Methods

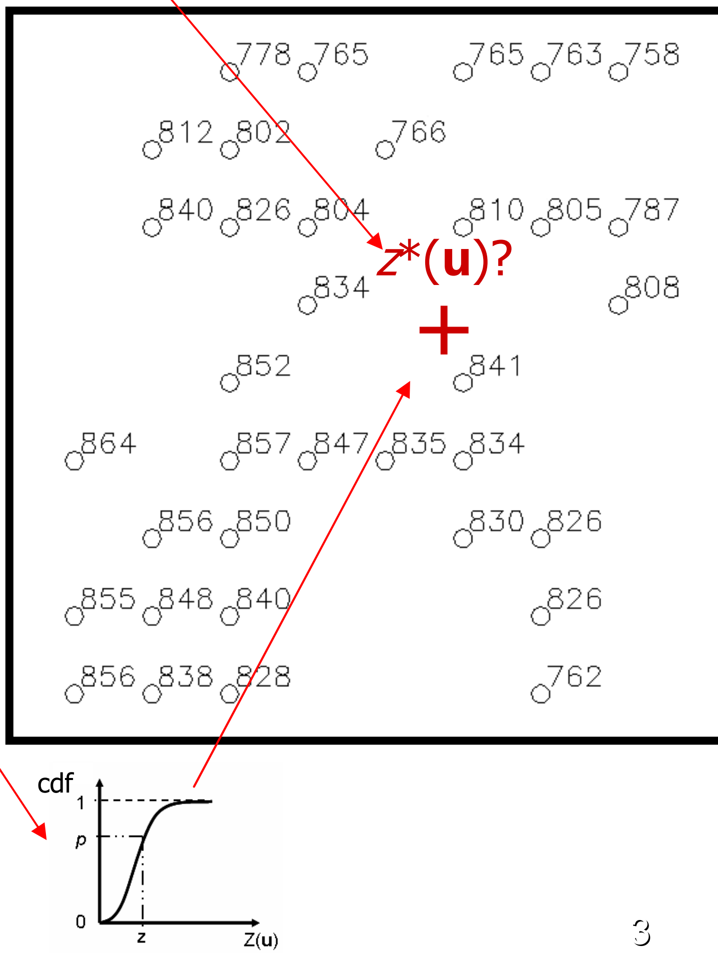
- **Deterministic**
  - The $z^*(\mathbf{u})$ value is estimated as a **Deterministic Variable**. An unique value is associated to its spatial location.
  - No uncertainties are associated to the estimations

- **Stochastic**
  - The $z^*(\mathbf{u})$ value is considered a **Random Variable** that has a *probability distribution function* associated to its possible values
  - Uncertainties can be associated to the estimations
  - Important: Samples are realizations of Random Variables

$z^*(\mathbf{u}) = K$

o sample locations

+ estimation locations



Predictions with Geostatistics

- **Random Variables and Random Fields**

  - A **Random Variable** (RV) is a variable that can take a variety of outcome values according to some probability (frequency) distribution. Deutsch and Journel, 1998.

  - A RV has a cumulative distribution function (cdf) which models the uncertainty about its *z* values. The cdf of a continuous RV $Z(\mathbf{u})$ is denoted:

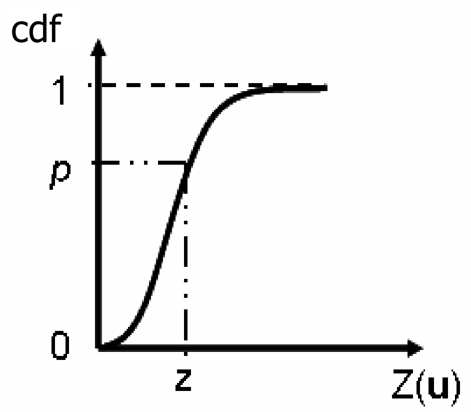
$$F(\mathbf{u}; z) = \text{Prob}\{Z(\mathbf{u}) \leq z\}$$

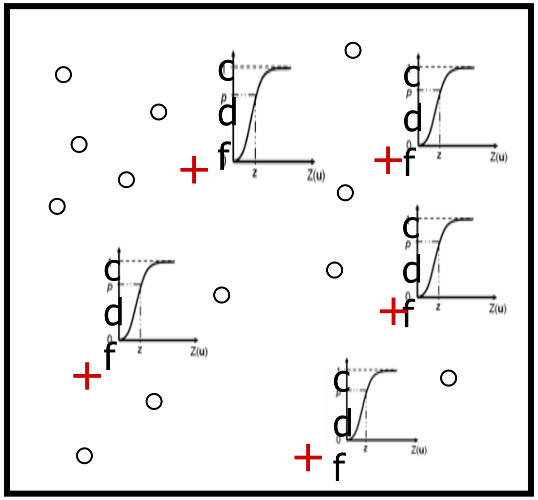
  - A **Random Function** (RF) is a set of RVs defined over some field of interest.

  - A RF has a set of all its *K*-variate cdfs for any number *k* and any choice of the *K* locations $\mathbf{u}_k$, $k = 1,\ldots K$

  - The multivariate cdf is used to model joint uncertainty about the K values $z(\mathbf{u}_1),\ldots,z(\mathbf{u}_K)$. The multivariate cdf of a continuous RF $Z(\mathbf{u})$ is denoted:

$$F(\mathbf{u}_1,\ldots,u_K; z_1,\ldots,z_K) = \text{Prob}\{Z(\mathbf{u}_1) \leq z_1,\ldots,Z(\mathbf{u}_K) \leq z_K\}$$





o sampled location (realizations)

+ unsampled locations

Predictions with Geostatistics

- **Regionalized variables**

- What is a **regionalized variable**?
  - It is a variable that is distributed in a region of the earth space.
  - It is used to represent spatial phenomena, which means, phenomena occurring in an earth region. Examples
    - grade of clay, or sand, in the soil;
    - grade of mineral in a rock, or soil.
    - variations of elevation, temperature, pressure, etc…
    - index of human development, illness taxes, etc…

- **Geostatistics provide tools to perform analysis and modeling, for estimations and simulations, on regionalized variables.**

Predictions with Geostatistics

- **Regionalized variables**

The spatial variation of a regionalized variable, at a spatial location **u**, can be expressed as an addition of three components Burrough (1987):

$$Z(\mathbf{u}) = m(\mathbf{u}) + \varepsilon'(\mathbf{u}) + \varepsilon''$$

where:

$m(\mathbf{u})$ is the variable tendency (global or mean local information)

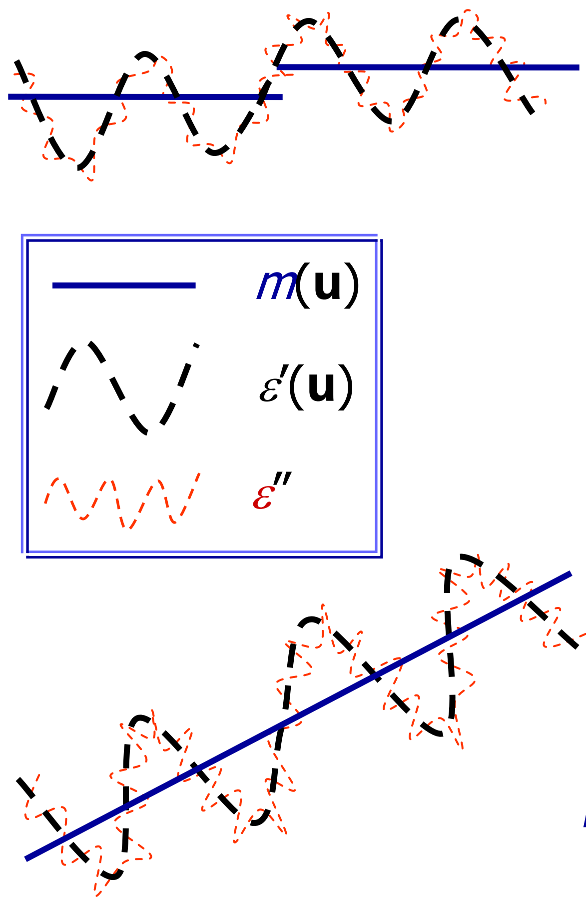
$\varepsilon'(\mathbf{u})$: is the random term that varies locally and has a spatial dependency with $m(\mathbf{u})$;

ε'' : is a uncorrelated random noise having a Gassian distribution with *mean* equal 0 and standard deviation σ.

Important: The geostatistics approaches works mainly on the modeling of the random term ε'(u):

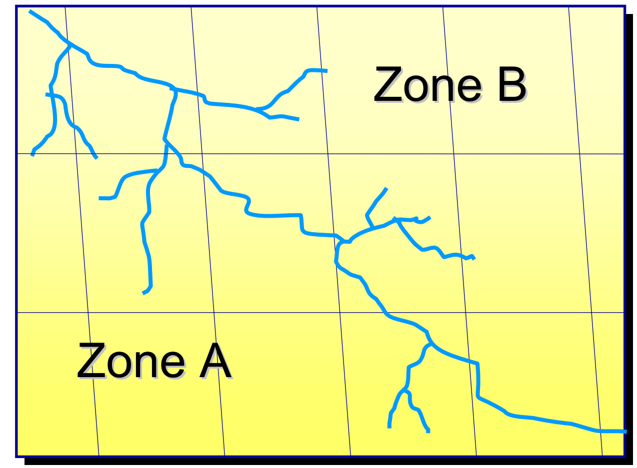
Predictions with Geostatistics

- **Regionalized variables - Tendencies**



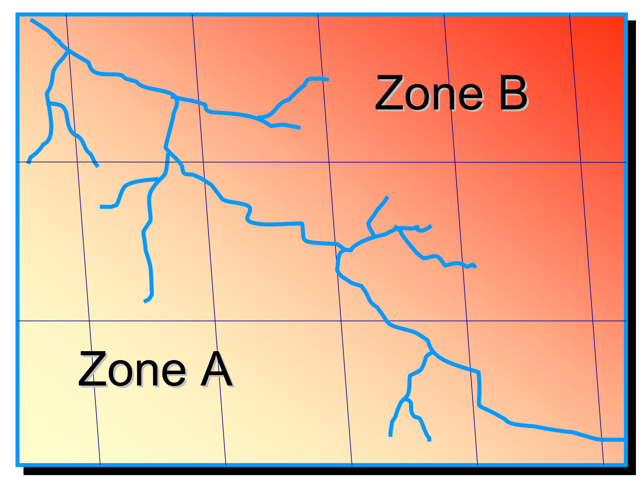
Without Tendency

m(**u**) is ~ constante



With Tendency

m(**u**) is a deterministic function



Fonte: Modified from Burrough (1987).

Predictions with Geostatistics

- **Stationary hypothesis**
  - The **RF** is said to be **stationary** within field A if its multivariate cdf is invariant under any translation the K coordinate vectors **u** (Journel, 1988)
- **Stationary hypotheses of mean and covariance (intrinsic stationary)**

1. All the RVs has the same mean value *m* (stationary of the first moment) (or the increments $[Z(\mathbf{u})-Z(\mathbf{u}+\mathbf{h})]$ has the same expected value equal 0)

$$E[Z(\mathbf{u})] = E[Z(\mathbf{u}+\mathbf{h})] = m \quad \text{or} \quad E[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]=0 \,, \quad \forall\ \mathbf{u},\mathbf{h} \in A.$$

2. The variogram (also the covariogram) varies only in function of **h**:

$$Var[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})] = E[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]^2 = 2\gamma(\mathbf{h})$$

where: $2\gamma(\mathbf{h})$ is called the variogram function

3) Under the stationary hypothesis $\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$

Predictions with Geostatistics

- **Stationary Hypothesis**
  - Relation between the Covariogram and the Variogram on intrinsic stationary hypothesis – Formulations

$2\gamma(\mathbf{h}) = E[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2$ and $C(\mathbf{h}) = E[Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h})] - E[Z(\mathbf{u})]E[Z(\mathbf{u} + \mathbf{h})]$

For stationary random functions $E[Z(\mathbf{u})] = E[Z(\mathbf{u}+\mathbf{h})] = m$ so:

$$C(\mathbf{h}) = E[Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h}) - E[Z(\mathbf{u})]^2] = E[Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h})] - m^2]$$

Developing the square terms of the variogram:

$$2\gamma(\mathbf{h}) = E[Z(\mathbf{u})]^2 + E[Z(\mathbf{u} + \mathbf{h})]^2 - 2E[Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h})]$$

$$\gamma(\mathbf{h}) = E[Z(\mathbf{u})]^2 - E[Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h})]$$

Subtracting m^2 of each one of the terms

$$\gamma(\mathbf{h}) = E[Z(\mathbf{u})]^2 - m^2 - \{E[Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h})] - m^2\}$$

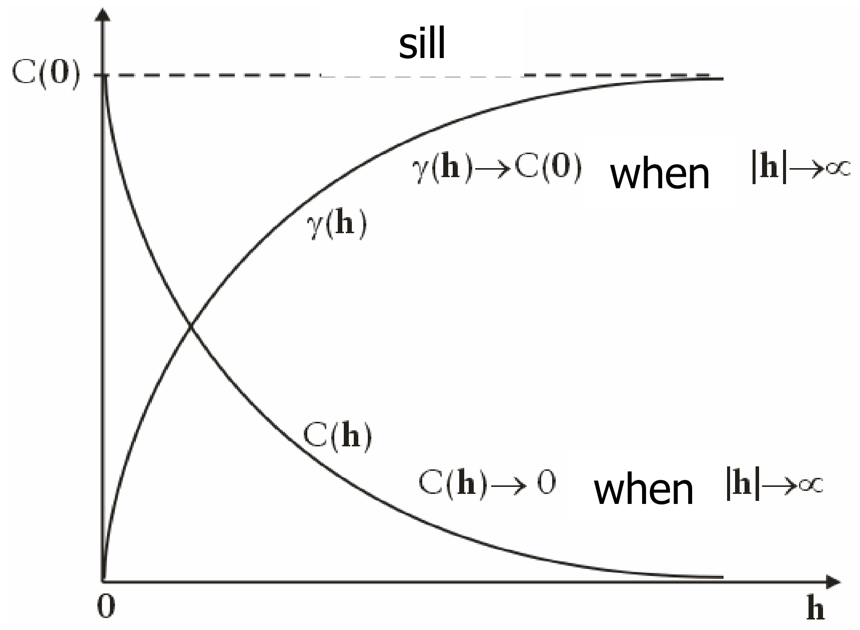
The relation between the variogram and the covariogram is reached

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$$

Predictions with Geostatistics

- **Stationary Hypothesis**
  - Relation between the Covariogram and the Variogram on intrinsic stationary hypothesis

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$$



Important: Under stationary hypothesis the covariance $C(\mathbf{h})$ and the variogram $2\gamma(\mathbf{h})$ are equivalent tools for spatial dependency characterization.

Predictions with Geostatistics

- **Stationary Hypothesis**
  - **Important**: *Stationary is a model hypothesis*
  - Geostatistics use the stationary hypothesis on these procedures to Estimate and Simulate values for Regionalized Variables
  - **Problems**
    - The Regionalized Variables to be modeled have an stationary behavior? What if they not?
    - How to use Geostatistics with attributes which have nonstationary behavior?
    - How to check stationary in a sample set of a Regionalized Variable?
      - Calculate local means with a moving window statistics
      - The variogram must exist (with a low value of nugget effect)
      - The sill C(0) must be approximately equal to the variance of the samples
      - The range value can not be too large compared to the region of study

Predictions with Geostatistics

- **Kriging prediction – Introduction** (Goovaerts, 1997)
  - **Kriging** is a generic name adopted by geostatisticians for a family of generalized *least-squares regression* algorithms.
  - All kriging estimators are but variants of the *basic linear regression estimator* $Z^*(\mathbf{u})$ defined as:

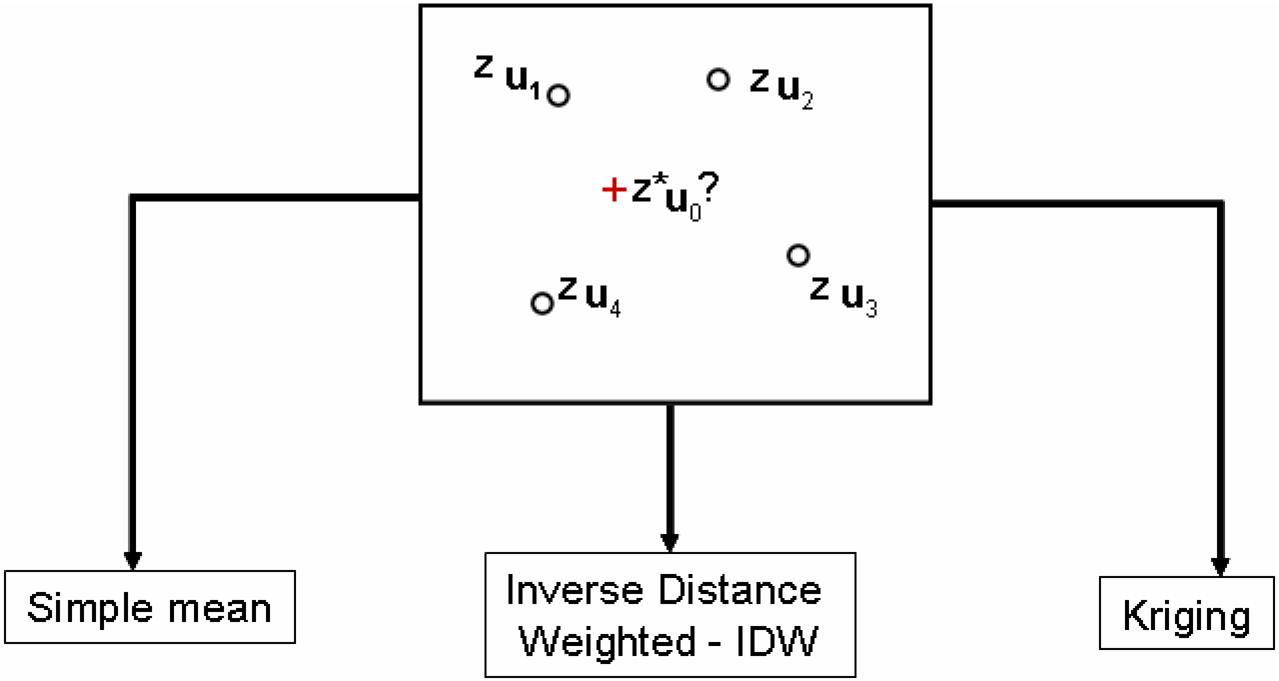
$$Z^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u})\left[Z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)\right]$$

where

- $\lambda_\alpha(\mathbf{u})$ is the weight assign to datum $z(\mathbf{u}_\alpha)$ interpreted as a realization of the RV $Z(\mathbf{u}_\alpha)$
- $m(\mathbf{u})$ and $m(\mathbf{u}_\alpha)$ are the expected values of RVs $Z(\mathbf{u})$ and $Z(\mathbf{u}_\alpha)$
- $n(\mathbf{u})$ is the number of data closest to the location **u** being estimated

Predictions with Geostatistics

- **Kriging prediction – Comparison with deterministic approaches**



Simple mean:

$$\hat{Z}_{\mathbf{u}_0} = \sum_{i=1}^{n} \lambda_i \, . \, Z_{\mathbf{u}_i}$$

$\lambda_i = 1/n$

Inverse Distance Weighted - IDW:

$$\hat{Z}_{\mathbf{u}_0} = \sum_{i=1}^{n} \lambda_i \, . \, Z_{\mathbf{u}_i}$$

$\lambda_i = 1/d^2$

Kriging:

$$\hat{Z}_{\mathbf{u}_0} = \sum_{i=1}^{n} \lambda_i \, . \, Z_{\mathbf{u}_i}$$

$\lambda_i = ?$

Predictions with Geostatistics

- **Kriging prediction – Introduction** (Goovaerts, 1997)
  - All kinds of kriging has the same objective of ***minimizing the estimation (or error) variance*** $\sigma^2_E(\mathbf{u})$ under the constraint of ***unbiasedness of the estimator***, that is:

$$\sigma_E^2(\mathbf{u}) = Var\{\mathrm{E}(\mathbf{u})\} = Var\{Z^*(\mathbf{u}) - Z(\mathbf{u})\}$$

is minimized under the constraint that

$$\mathrm{E}\{Z^*(\mathbf{u}) - Z(\mathbf{u})\} = 0$$

  - The RF $Z(\mathbf{u})$ is usually decomposed into a residual component $R(\mathbf{u})$ and a trend component $m(\mathbf{u})$: $Z(\mathbf{u}) = R(\mathbf{u}) + m(\mathbf{u})$
  - $R(\mathbf{u})$ is modeled as a ***stationary RF*** with zero mean and covariance $C_R(\mathbf{h})$:

$$\mathrm{E}\{R(\mathbf{u})\} = 0$$

$$\mathrm{Cov}\{R(\mathbf{u}), R(\mathbf{u}+\mathbf{h})\} = \mathrm{E}\{R(\mathbf{u}) \cdot R(\mathbf{u}+\mathbf{h})\} = C_R(\mathbf{h})$$

  - So, the expected value of Z at location **u** is $m(\mathbf{u})$ or $\mathrm{E}\{Z(\mathbf{u})\} = m(\mathbf{u})$

Predictions with Geostatistics

- **Kriging Variants –** (Goovaerts, 1997)
  - Three kriging variants can be used according the model considered for $m(\mathbf{u})$

1. **Simple Kriging** (SK) – $m(\mathbf{u})$ is known and is constant in the considered region *A*.

$$m(\mathbf{u}) = m,\ known\ \forall \mathbf{u} \in A$$

2. **Ordinary Kriging** (OK) – m(**u'**) is unknown but constant for sub regions *W* of *A*

$$m(\mathbf{u}') = \text{constant but unknown}\ \ \forall \mathbf{u}' \in W(\mathbf{u})$$

3. **Kriging with a trend model** (KT) or **Universal Kriging** – m(**u'**) is unknown but varies smoothly within each local neighborhood $W(\mathbf{u})$ of the region *A*. The trend component is modeled as a linear combination of functions $f_k(\mathbf{u})$ of the coordinates:

$$m(\mathbf{u}') = \sum_{k=0}^{K} a_k(\mathbf{u}') f_k(\mathbf{u}')$$

$$\text{with } a_k(\mathbf{u}') \approx a_k \text{ constant but unknown}\ \ \forall \mathbf{u}' \in W(\mathbf{u})$$

Predictions with Geostatistics

- **Kriging Predictions – The Simple Kriging** (Goovaerts, 1997)

The trend component $m(\mathbf{u}) = m$ is the stationary mean of the attribute in region *A*.

$$Z^*(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u}) \cdot [Z(\mathbf{u}_\alpha) - m] + m$$

$$= \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u}) \cdot Z(\mathbf{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u})\right] \cdot m$$

As for all krigings the $n(\mathbf{u})$ weights are determined such as to ***minimizing the estimation (or error) variance*** $\sigma^2_E(\mathbf{u})$ under the constraint of ***unbiasedness of the estimator***.

Applying these two conditions the weights are evaluated from the solution of the following linear system (see mathematical deductions in section 5.2 of the Goovaerts book):

$$\sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta C(\mathbf{u}_\alpha - \mathbf{u}_\beta) + \mu = C(\mathbf{u}_\alpha - \mathbf{u}_u), \qquad \alpha = 1, \ldots, n(\mathbf{u})$$

The minimum error variances, the SK variance, is given by:

$$\sigma^2_{SK}(\mathbf{u}) = C(0) - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{SK}(\mathbf{u}) \cdot C(\mathbf{u}_\alpha - \mathbf{u})$$

Important: $C(\mathbf{h})$ comes from the stationary relation

$C(\mathbf{h}) = C(\mathbf{0}) - \gamma(\mathbf{h})$

Predictions with Geostatistics

- **Kriging Predictions – The Ordinary Kriging** (Goovaerts, 1997)

The trend component $m(\mathbf{u})$ is unknown but is constant in sub-regions W of the region A.

$$Z^*(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u}) \cdot Z(\mathbf{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u})\right] \cdot m(\mathbf{u})$$

$$Z^*_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) \cdot Z(\mathbf{u}_\alpha) \quad \text{with} \quad \sum_{\alpha=1}^{n(u)} \lambda_\alpha^{OK}(\mathbf{u}) = 1$$

As for all krigings, the $n(\mathbf{u})$ weights are determined such as to ***minimizing the estimation (or error) variance*** $\sigma^2_E(\mathbf{u})$ under the constraint of ***unbiasedness of the estimator***.

Applying these two conditions the weights are evaluated from the solution of the following linear system (see mathematical deductions in section 5.3 of the Goovaerts book):

$$\begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{OK} C(\mathbf{u}_\alpha - \mathbf{u}_\beta) + \mu_{OK}(\mathbf{u}) = C(\mathbf{u}_\alpha - \mathbf{u}_u), & \alpha = 1,\ldots,n(\mathbf{u}) \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta^{OK}(\mathbf{u}) = 1 \end{cases}$$

Important: $C(\mathbf{h})$ comes from the stationary relation

$C(\mathbf{h}) = C(0) - \gamma(\mathbf{h})$

Predictions with Geostatistics

- **Kriging prediction – Ordinary Kriging equation system in matrix notation**

K.λ **= k** or λ **= K**$^{-1}$**.k**

$$\begin{bmatrix} \lambda_1 \\ \circ \\ \circ \\ \circ \\ \lambda_n \\ 1 \end{bmatrix} = \begin{bmatrix} C_{11} & \circ\circ\circ & C_{1n} & 1 \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ C_{n1} & \circ\circ\circ & C_{nn} & 1 \\ 1 & \circ\circ\circ & 1 & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} C_{10} \\ \circ \\ \circ \\ \circ \\ C_{n0} \\ 1 \end{bmatrix}$$

$$[\lambda] \qquad [K]^{-1} \qquad [k]$$

K is the matrix of Covariances between samples

k is the vetor of Covariances between samples and the estimation location $\mathbf{u}_0$

λ is the vector of Weights to be computed for each sample at $\mathbf{u}_\alpha$

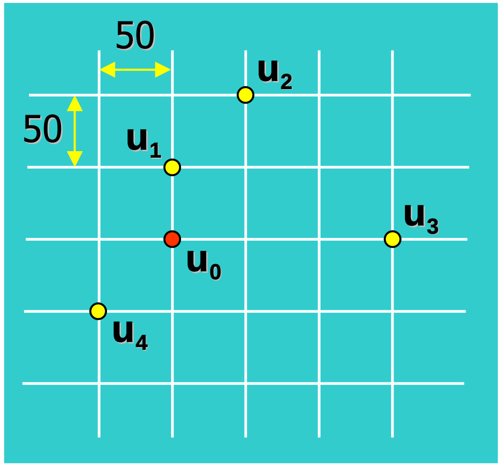
Important: $C(\mathbf{h})$ is calculated using the stationary relation

$$C(\mathbf{h}) = C(\mathbf{0}) - \gamma(\mathbf{h})$$

Predictions with Geostatistics

Kriging prediction - EXAMPLE

Given the below configuration how the Kriging estimates the value of the Z variable at location $\mathbf{u}_0$, using the samples $z(\mathbf{u}_1)$, $z(\mathbf{u}_2)$, $z(\mathbf{u}_3)$ e $z(\mathbf{u}_4)$. Consider that the semivariogram was fitted by a spherical model with the following parameters: a (range)= 200, C_1 (contribution)= 20, and C_0 (nugget)= 2.



$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \alpha \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & 1 \\ C_{21} & C_{22} & C_{23} & C_{24} & 1 \\ C_{31} & C_{32} & C_{33} & C_{34} & 1 \\ C_{41} & C_{42} & C_{43} & C_{44} & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} C_{01} \\ C_{02} \\ C_{03} \\ C_{04} \\ 1 \end{bmatrix}$$

Theoretical Model

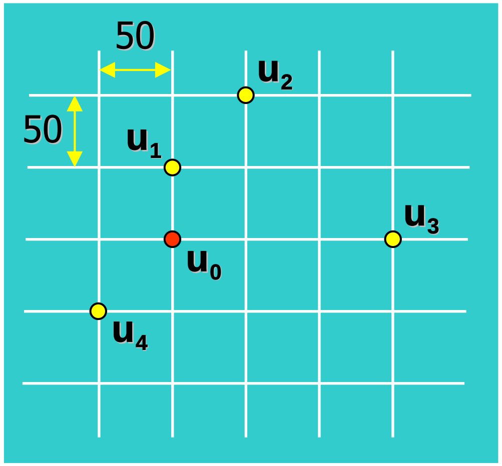
The matrix elements are evaluated as: $C_{ij} = C_0 + C_1 - \gamma(\mathbf{h})$

$$C_{12} = C_{21} = C_{04} = C_0 + C_1 - \gamma(50\sqrt{2})$$

$$= (2+20) - \left[2 + 20\left(1{,}5\frac{50\sqrt{2}}{200} - 0{,}5\frac{(50\sqrt{2})^3}{(200)^3} \right) \right] = 9{,}84$$

Predictions with Geostatistics

Kriging prediction - **EXAMPLE**



$C_{13} = C_{31} = (C_0 + C_1) - \gamma [\sqrt{(150)^2 + (50)^2}] = 1,23$

$C_{14} = C_{41} = C_{02} = (C_0 + C_1) - \gamma [\sqrt{(100)^2 + (50)^2}] = 4,98$

$C_{23} = C_{32} = (C_0 + C_1) - \gamma [\sqrt{(100)^2 + (100)^2}] = 2,33$

$C_{24} = C_{42} = (C_0 + C_1) - \gamma [\sqrt{(100)^2 + (150)^2}] = 0,29$

$C_{34} = C_{43} = (C_0 + C_1) - \gamma [\sqrt{(200)^2 + (50)^2}] = 0$

$C_{01} = (C_0 + C_1) - \gamma (50) = 12,66$

$C_{03} = (C_0 + C_1) - \gamma (150) = 1,72$

$C_{11} = C_{22} = C_{33} = C_{44} = (C_0 + C_1) - \gamma (\mathbf{0}) = 22$

Predictions with Geostatistics

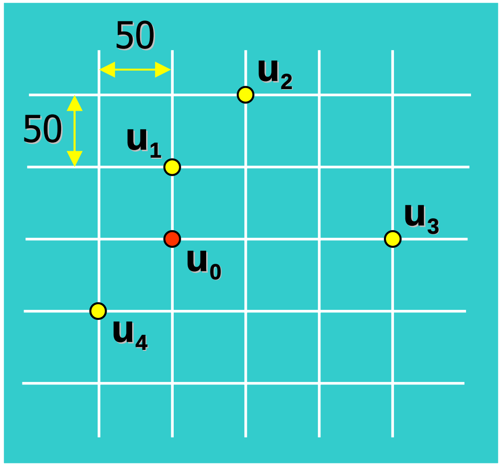
Kriging prediction - EXAMPLE

Filling out the C_{ij} values in the matrices the following weights will be found:

$$\lambda_1 = 0{,}518 \quad \lambda_2 = 0{,}022 \quad \lambda_3 = 0{,}089 \quad \lambda_4 = 0{,}371$$

Finally the estimated value at $\mathbf{u}_0$ is given by:

$$\hat{Z}(\mathbf{u}_0) = 0{,}518\ z(\mathbf{u}_1) + 0{,}022\ z(\mathbf{u}_2) + 0{,}089\ z(\mathbf{u}_3) + 0{,}371\ z(\mathbf{u}_4)$$



Observation: although the samples Z_2 and Z_3 have small influence in the final estimate value of Z_0, their influences are not linear with respect to their distances to Z_0. The sample Z_3 is far than Z_2, but its contribution is more than 4x the contribution of sample 3. This occurs due the fact that Z_0 has a direct influence of Z_3, while Z_2 is to closer to the Z_1. The introduction of the covariances between samples, besides the covariance between the samples and Z_0, in the weight evaluation avoid the association of undue weights to "clustered" samples. This is a great advantage of the Kriging compared to the others that consider only the distances between each sample and the point to be estimated (IDW for example).

Predictions with Geostatistics

- **Kriging prediction**

Summary

Following Journel, 1988: $\mathbf{K}.\boldsymbol{\lambda} = \mathbf{k} \Rightarrow \boldsymbol{\lambda} = \mathbf{K}^{-1}\mathbf{k}$

$$\begin{vmatrix} \lambda_1 \\ \lambda_2 \\ : \\ \lambda_n \\ \alpha \end{vmatrix} = \begin{vmatrix} C_{11} & C_{12} \ldots\ldots C_{1n} & 1 \\ C_{21} & C_{22} \ldots\ldots C_{2n} & 1 \\ : & : \quad\quad : & : \\ C_{n1} & C_{n2} \ldots\ldots C_{nn} & 1 \\ 1 & 1 \ldots\ldots 1 & 0 \end{vmatrix}^{-1} \begin{vmatrix} C_{10} \\ C_{20} \\ : \\ C_{n0} \\ 1 \end{vmatrix}$$

The elements of the matrices are evaluate by the relation: (*Journel, 1988*):

$$C_{ij} = C(\mathbf{0}) - \gamma(\mathbf{h}) = C_0 + C_1 - \gamma(\mathbf{h})$$

Replacing the C_{ij} values in the matrices one find the weights λ_1, λ_2, ..., e λ_n.

The Kriging Estimator is given by: (*Journel, 1988*): $Z^*_{\mathbf{x}_0} = \sum_{i=1}^{n} \lambda_i Z(\mathbf{x}_i)$

The Kriging Variance (*Journel, 1988*): $\sigma^2_{ko} = (C_0 + C_1) - \lambda^T k$

Predictions with Geostatistics

- **Cokriging prediction**

-It is a predictor that works with 2 or more variables.

-Primary (main) variable values $Z_1^*(\mathbf{u})$ are estimated using a sample set of Z_1 and sample set of secondary variables $Z_2,\ldots,Z_v$ .that are correlated to the primary.

$$Z_1^*(\mathbf{u}) - m_1(\mathbf{u}) = \sum_{\alpha_1=1}^{n_1(\mathbf{u})} \lambda_{\alpha 1}(\mathbf{u})\left[Z_1(\mathbf{u}_{\alpha 1}) - m_1(\mathbf{u}_{\alpha 1})\right] + \sum_{i=2}^{N_v} \sum_{\alpha_i=1}^{n_i(\mathbf{u})} \lambda_{\alpha i}(\mathbf{u})\left[Z_i(\mathbf{u}_{\alpha i}) - m_i(\mathbf{u}_{\alpha i})\right]$$

- The secondary variables must have a high degree of correlation with the primary.

- The cokriging estimators make use of direct and crossvariograms.

- As for the kriging estimator, one can work with simple, ordinary and with a trend cokriging.

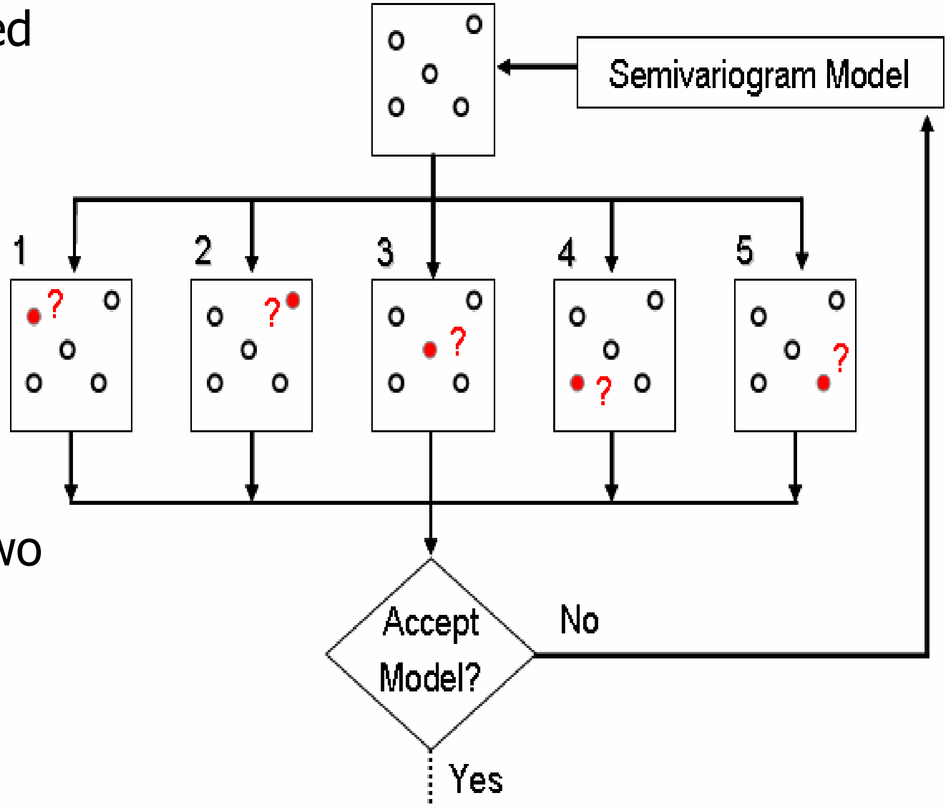
Predictions with Geostatistics

- **Cross Validation and Validation**

The modeled variogram can be considered valid or not by some exploratory analysis on the results of a (cross)validation procedure.

Cross Validation - Each sample αi is taken off the sample set and its $z^*(\mathbf{u}_{\alpha i})$ value is estimated using the current semivariogram model.

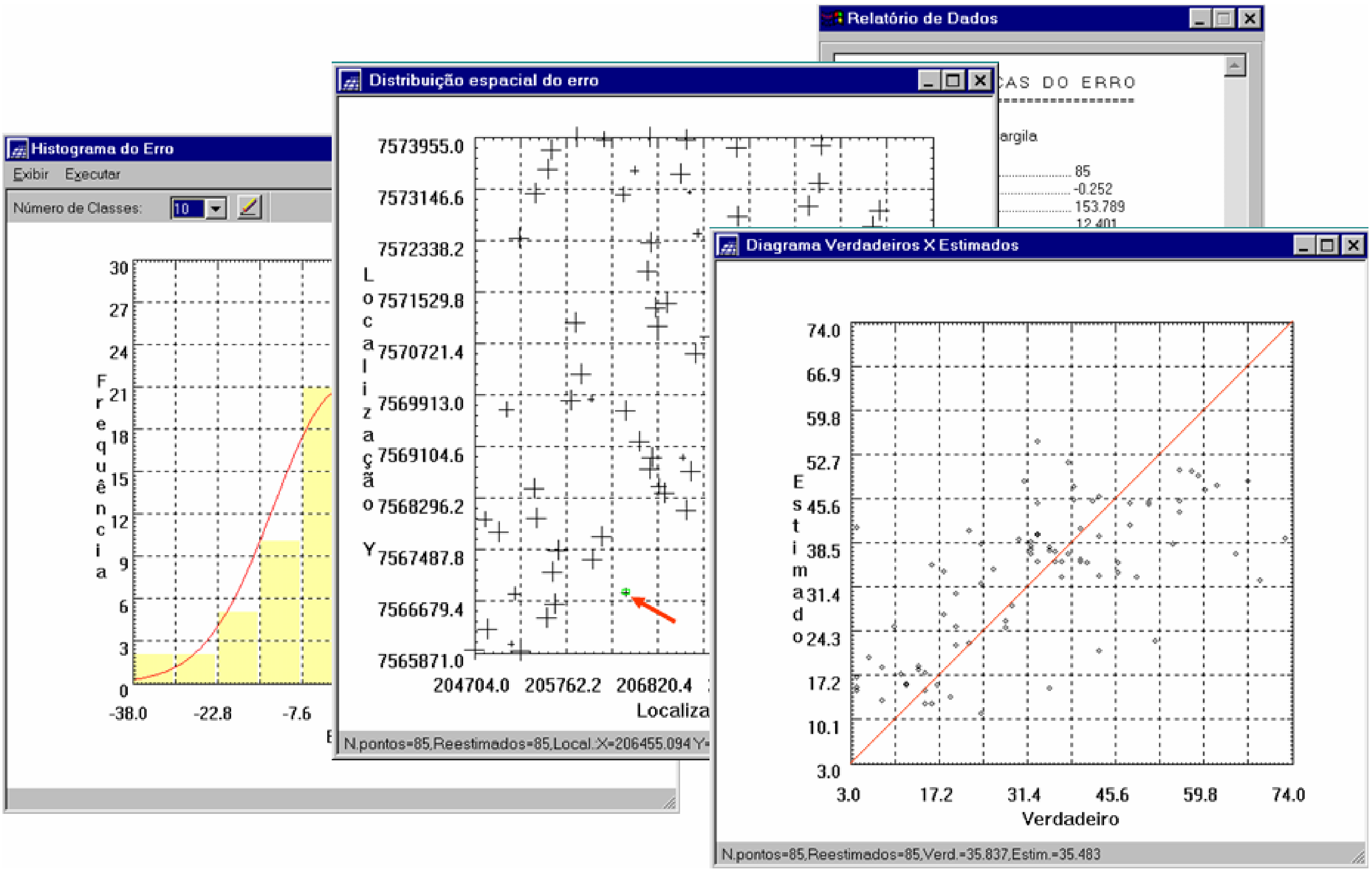
The differences ("errors") between the two sample sets can be compared by some exploratory spatial data analysis such as: summary statistics, histograms, spatial distribution diagram, scatter plots of the observed x the estimated z values, etc…



Validation – The above mentioned analysis can be done also if a different, and representative, sample set of Z is available for the same region.

Predictions with Geostatistics

- **ESDA examples for (Cross)Validation procedures**



Predictions with Geostatistics

- **Problems with geostochastic procedures**

The main drawback of using geostatistic approaches is the need of work on variogram generations and fittings. This work is interactive and require from the user knowledge of the main concepts related to basics of the geostatistics in order to obtain reliable variograms.

The kriging approach is an estimator based on weighted mean evaluations and is uses the hypothesis of minimizing the error variance. Because of these the kriging estimates create smooth models that can filter some details of the original surfaces.

Predictions with Geostatistics

- **Advantages on using geostochastic procedures**

  - Spatial continuity is modeled by the variogram
  - Range define automatically the region of influence and number of neighbors
  - Cluster problems are avoided
  - It can work with anisotropic phenomena
  - Allows prediction of the Kriging variance

Summary and Conclusions

Summary and Conclusions

- Geostatistic estimators can be used to model spatial data.
- Geostatistics estimators make use of variograms that model the variation (or continuity) of the attribute in space.
- Geostatistics advantages are more highlighted when the sample set is not dense
- Current GISs allow users work with these tools mainly in Spatial Analysis Modules.

Predictions with Geostatistics

Exercises

1. Run the Lab3 that is available in the geostatistics course area of ISEGI online.
2. Find out a sample set of points of an spatial attribute of your interest (in the internet or with a friend, for examples). Important: The sample set should have more than 50 samples and less than 500 samples.
3. Repeat the exercises you have realized in the Lab2 and Lab3 in order to model your attribute with geostatistical procedures. Use any software you want.
4. Report your work and results in an "article" that will be presented (in 10 minutes) to the other students of the geostatistics course (to be scheduled).
5. Send the report to the e-mail of the geostatistics professor before november 8, 2007.

Predictions with Deterministic Procedures

END

of Presentation