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

## Master of Science in Geospatial Technologies

# Geostatistics Predictions with Deterministic Procedures

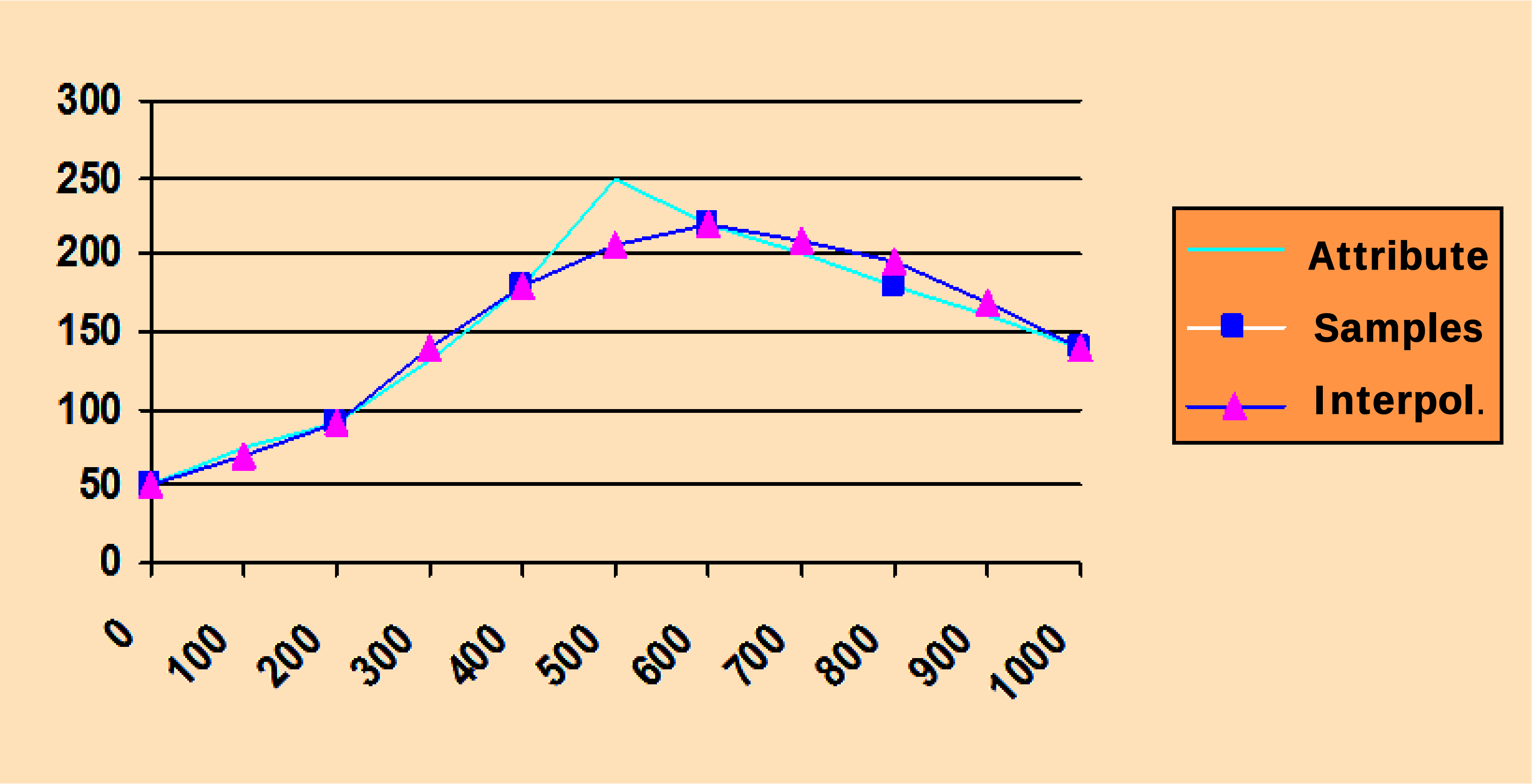
Carlos Alberto Felgueiras
cfelgueiras@isegi.unl.pt

# Predictions with Deterministic Procedures

## Contents

# Predictions with Deterministic Procedures

- **Introduction – Interpolation from Samples – General Idea**



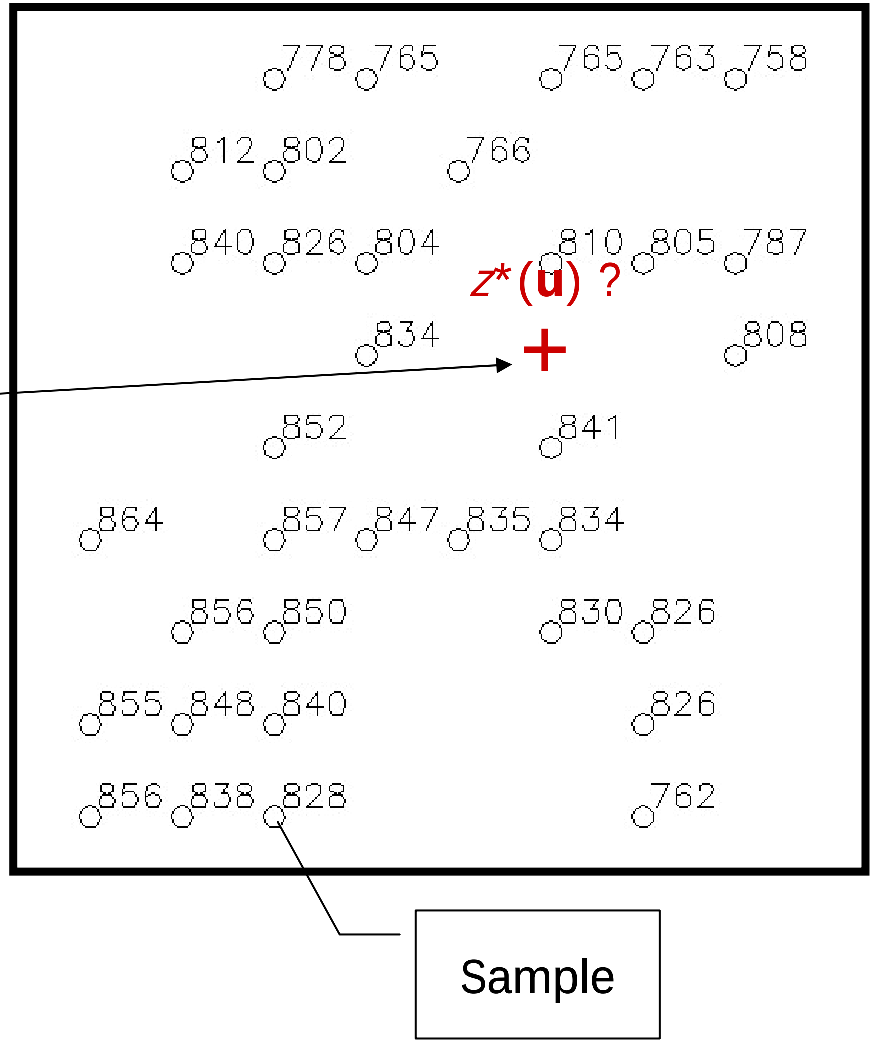
# Predictions with Deterministic Procedures

- **Introduction – General Concepts**

**Point Predictions, Estimations or Interpolations:**

- Using estimation procedures over sample information in order to calculate *z* values at unsampled locations.

- Defining mathematical models that allow performing attribute predictions at any location of a region of the earth surface.




**Spatial Interpolation Methods**

- *Deterministic x Stochastic Methods*
- *Global x Local Methods*

# Predictions with Deterministic Procedures

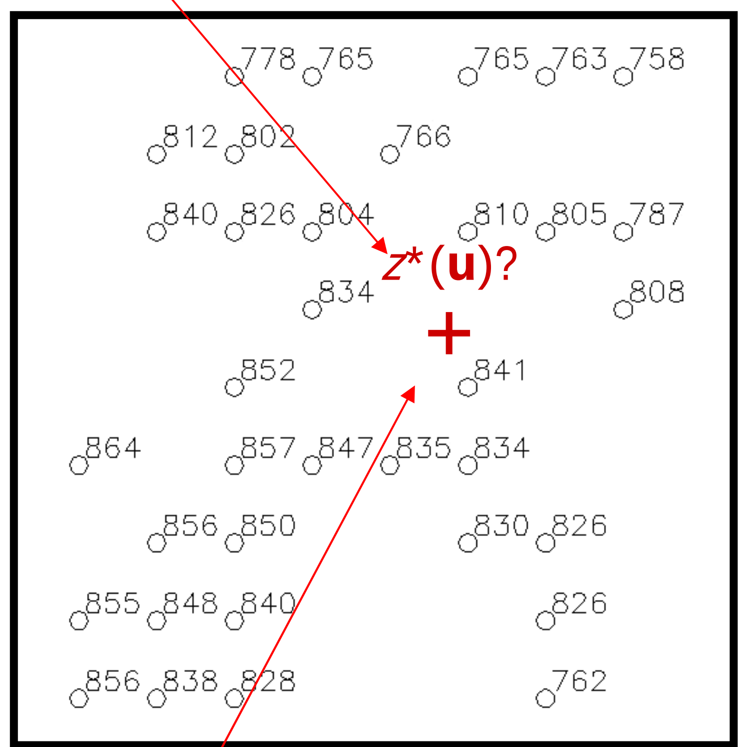
**Deterministic x Stochastic Methods**

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

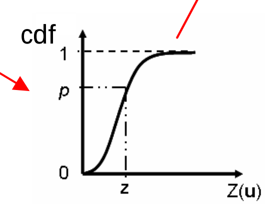
**o** sample locations

**+** estimation locations

- **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



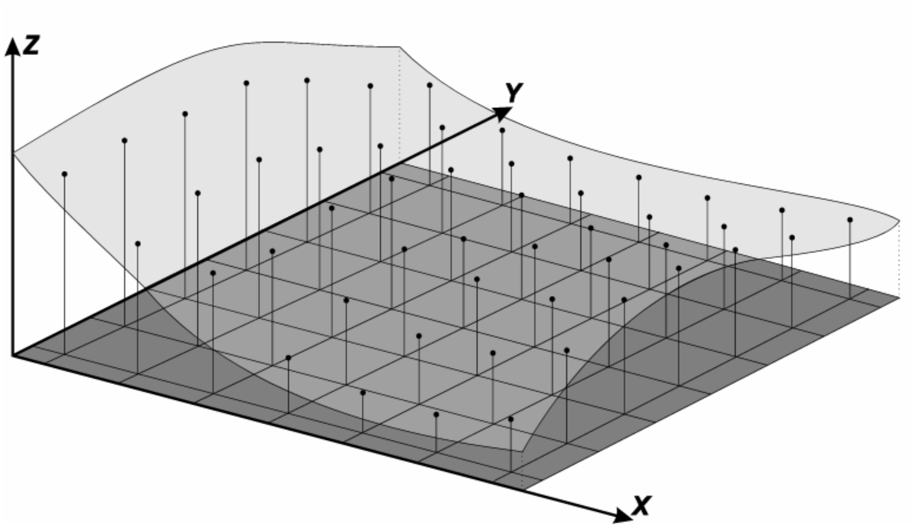
# Predictions with Deterministic Procedures

## Global Estimations

- Use of only one mathematical model, generally a polynomial, to represent the attribute variation for the hole area.
- Difficulties to represent erratic (complex) attribute variations. Show only the general tendencies, filtering details. Problems with polynomial oscillations.
- Used mostly in engineering projects for modeling smooth curves and surfaces.
- Examples: polynomials, splines (exact), bezier, ...

$$Z = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 \qquad n>0$$

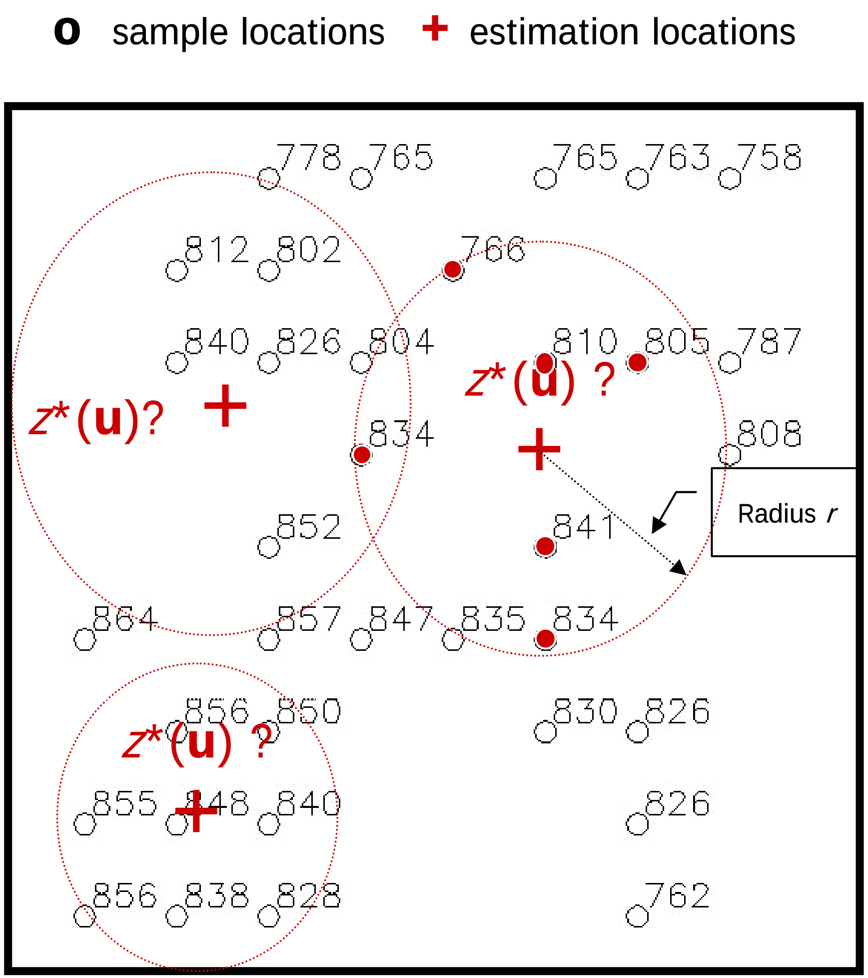
(use best fit to the samples criteria)



# Predictions with Deterministic Procedures

## Local Estimators

- Use local neighborhood to determine the closer local samples to be taken into account in the interpolation.
- Polynomials are applied locally, using only the closer samples.
- Problems with continuity must be considered.
- Criteria for neighborhood definition
  - Number of closer samples? How many?
  - Distance from the location to be interpolated? How far?
  - What is better?



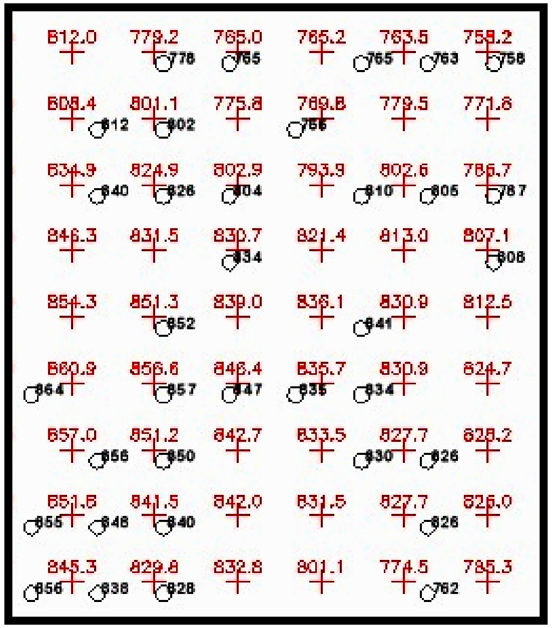

Closer samples for local estimation

# Predictions with Deterministic Procedures

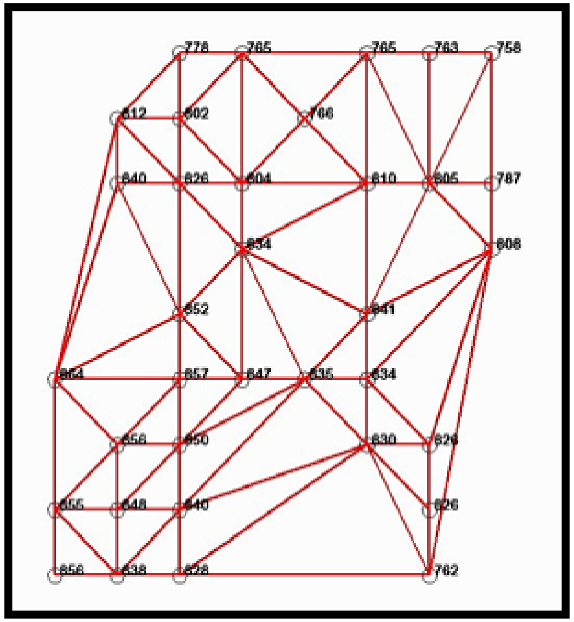
- **Digital Terrain Models (DTM or DEM for elevations)**

**GRID x TIN :** Usually predictions are used in GIS to create rectangular and triangular mesh structures – Digital Terrain Representations

**Rectangular Regular Networks**
(Grid or Image)



**Triangular Irregular Networks**
(TIN)



**o** sample locations **+** estimation locations

# Predictions with Deterministic Procedures

- **Rectangular Grids with Local Mean Interpolators**
- General equation for estimating each grid point at location **u** $(x_{\mathbf{u}}, y_{\mathbf{u}})$

$$z^*_{\mathbf{u}} = \frac{\sum_{\alpha=1}^{n} w_{\alpha\mathbf{u}} z_{\alpha}}{\sum_{\alpha=1}^{n} w_{\alpha\mathbf{u}}}$$

**u** is the spatial location of z*

*n* is the number of samples inside r

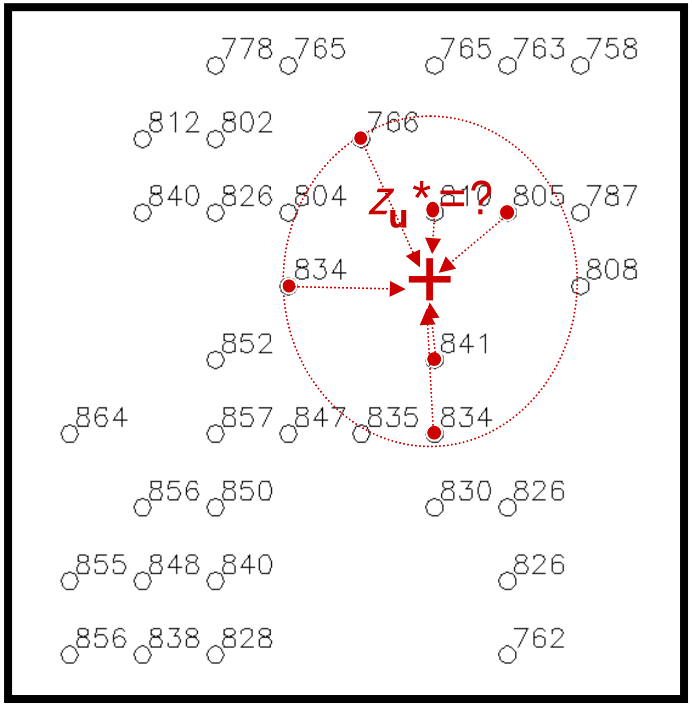
$w_{\alpha\mathbf{u}}$ is the weight of sample $\alpha$ at location **u**

**Interpolators:**

- **Nearest Neighbor** when $n = 1$
- **Simple Means** when $w_{\alpha\mathbf{u}} = 1$

o sample locations

\+ estimation locations




Closer samples for estimation a $z^*$ value at a grid location

# Predictions with Deterministic Procedures

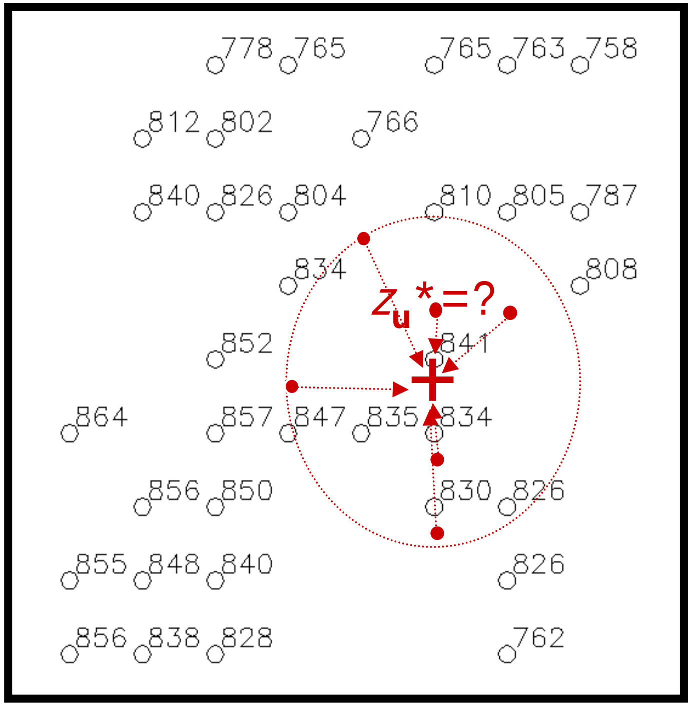
- **Rectangular Grids with Local Mean Interpolators**

o sample locations

+ estimation locations

- General equation for estimating each grid point at location **u**

$$z_{\mathbf{u}}^{*} = \frac{\sum_{\alpha=1}^{n} w_{\alpha\mathbf{u}} z_{\alpha}}{\sum_{\alpha=1}^{n} w_{\alpha\mathbf{u}}}$$




$w_{\alpha\mathbf{u}}$ is the weight of sample $\alpha$ at location **u**

**Interpolators:**

- **IDW – Inverse Distance Weighted** when
- **IDW Optimized (Quadrants)**

$$w_{\alpha\mathbf{u}} = 1 \Big/ d_{\alpha\mathbf{u}}^{k}$$

$d_{\alpha u}$ is the Euclidean distance between sample $\alpha$ and location **u**

*k is the power of the distance* $d_{\alpha\mathbf{u}}$

$$d_{\alpha\mathbf{u}} = \sqrt{(x_{\alpha} - x_{\mathbf{u}})^2 + (y_{\alpha} - y_{\mathbf{u}})^2}$$

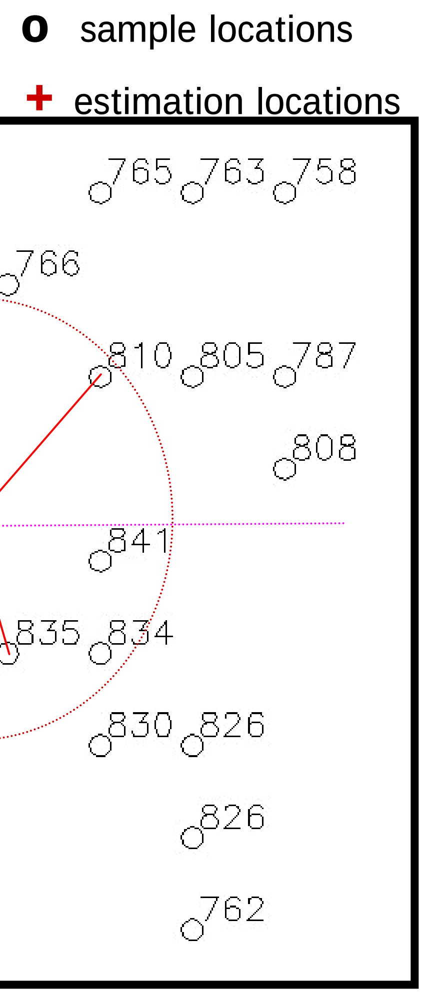
# Predictions with Deterministic Procedures

- **Rectangular Grids with Local Mean Interpolators**

## IDW Optimized (Quadrants)

- The space is split in 4 quadrants relative to the interpolation location.
- Only 1 (or 2, ..., or *n*) closer sample(s) of each quadrant is considered for the interpolation.
- This optimization avoid using clustered samples
- Some authors suggest split the space in octants (or more ants).

**Important:** The *z* variation inside each rectangle of a rectangular grid can be modelled by bilinear, bicubic or other patch function.

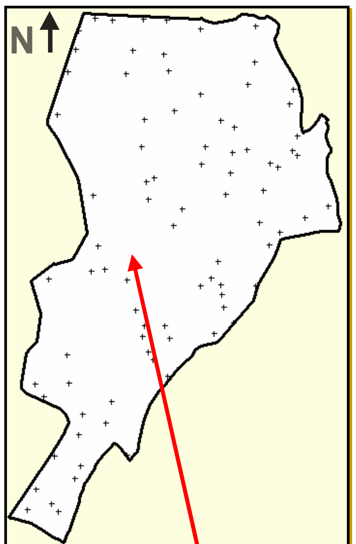



IDW considering quadrant locations

# Predictions with Deterministic Procedures

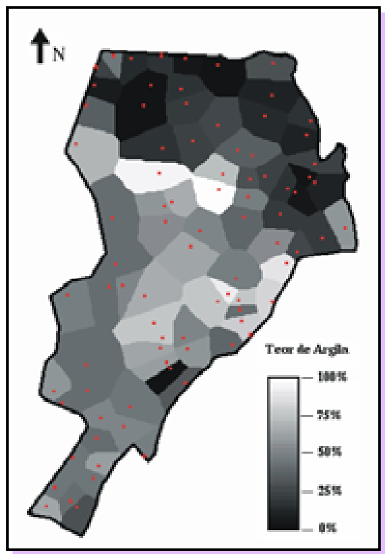
- **Rectangular Grids with Local Mean Interpolators**

Study Area



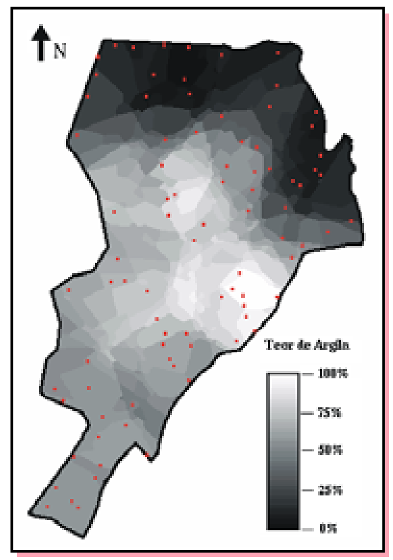

Sample Set

Deterministic Procedures

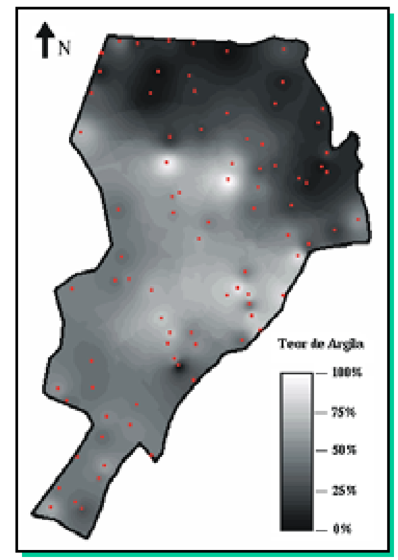



Nearest Neighbors

Dirichlet Map




Simple Means




Inverse Distance Weighted

# Predictions with Deterministic Procedures

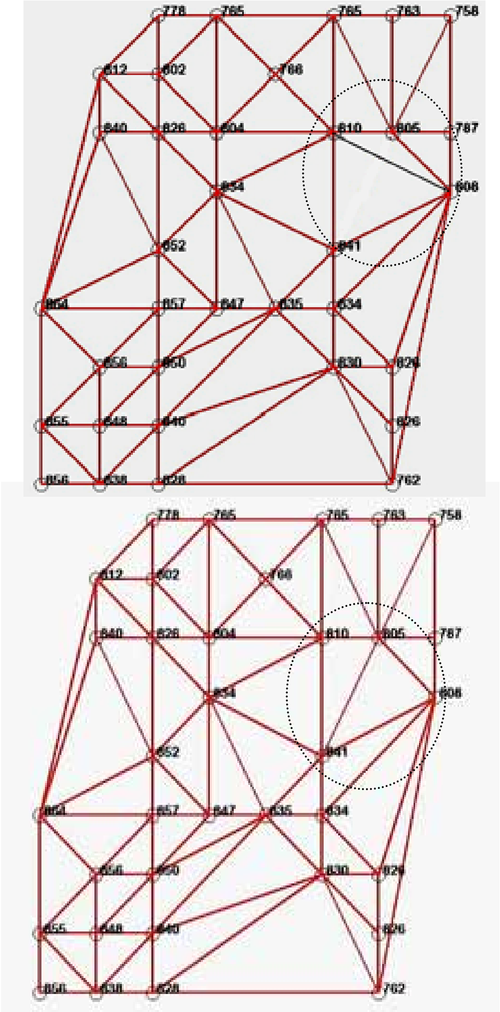
- **Triangular Irregular Networks - TIN**

  - The samples are connected in order to form a triangular partition in the region of the samples.

  - The mesh is created using the samples location as the vertices of the triangles. So, there is no interpolations in the process of the TIN construction

  - TIN is adaptative: Bigger triangles represent homogeneous regions and smaller ones represent more erratic areas.

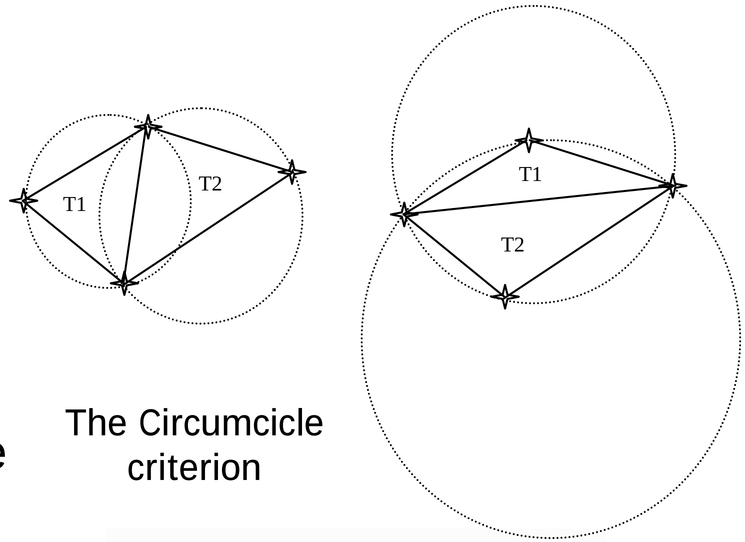
  - Problem: Given a set of n samples it is possible to construct many different triangulations. How to compare different triangulations ? What is the better triangulation for a given sample set?



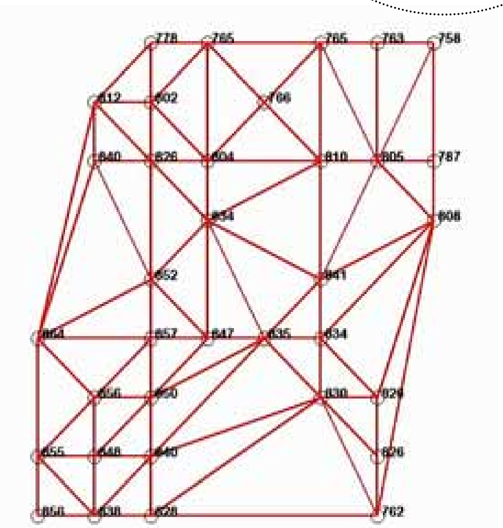
Two different TIN for a same sample set

# Predictions with Deterministic Procedures

- **TIN – The Delaunay Triangulation**
- The **Delaunay triangulation** is the most used TIN in DTM modeling.
- **Circumcircle criterion:** " Given the points pa, pb and pc $\in$ the Sample Set P where a$\neq$b$\neq$c, a T is a Delaunay triangulation only and if only $\forall$ *t* $\in$ T, with vertices at pa, pb and pc , the circumcircle formed by the vertices of *t* does not contain any other point pd pd $\in$ P / d $\neq$ a $\neq$ b $\neq$ c ".
- Uses the **circumcircle criteria** to define the triangles of an **unique** triangulation.
- Avoid creation of very thin triangles. Create triangles more closer to equilateral triangles
- **Important:** The z variation inside each triangle can be modeled by linear, cubic, quintic or other function.




The Circumcicle criterion



A Delaunay Triangulation

# Predictions with Deterministic Procedures

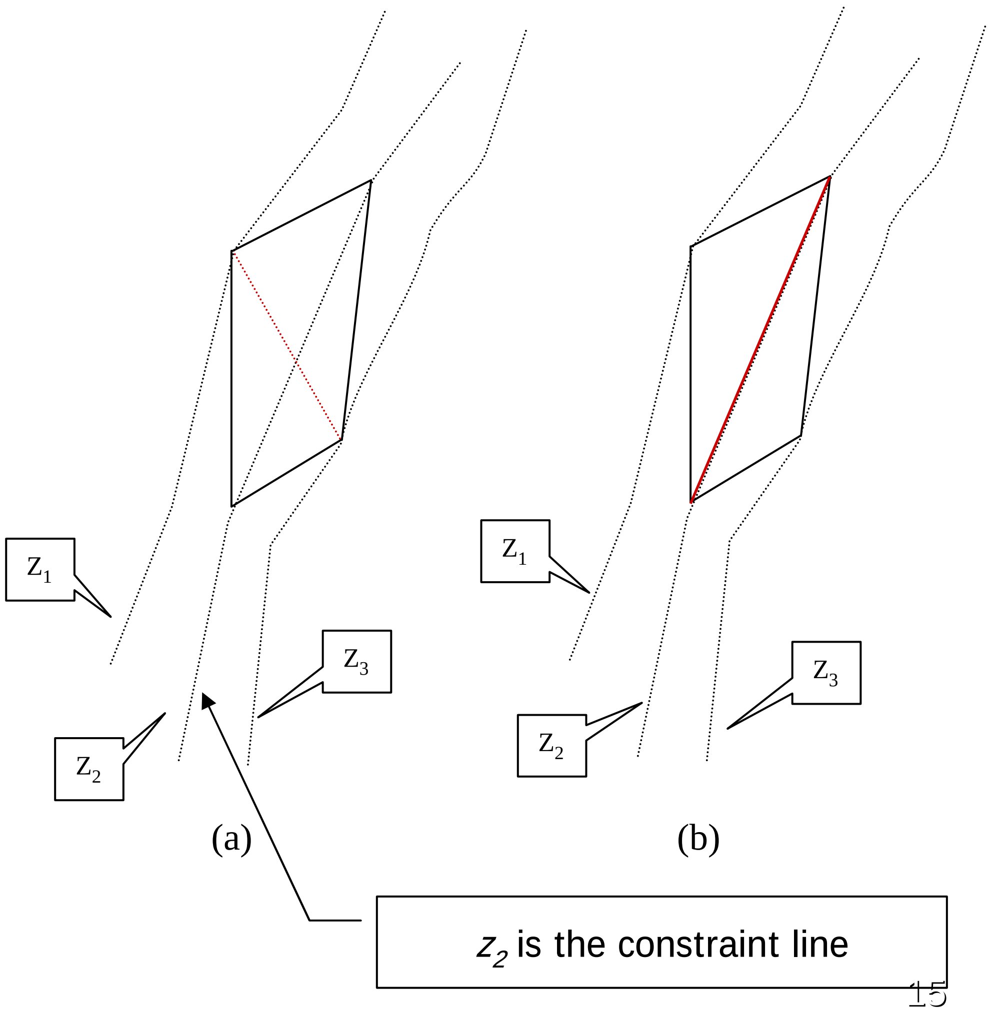
- **TIN – The Delaunay Triangulation with constrainties**

**TIN with constrainties**

Constraints lines are lines that represent morphological characteristics of the surface and can not be “broken” by triangles sides

It is possible to include constraint information (break lines) to be considered in the TIN construction process.

This is important for getting better representations for attributes presenting morphological structures (ex. elevations)



# Predictions with Deterministic Procedures

- **Grid from TIN**

  Given a TIN it is possible to calculate a rectangular grid for the same region.

  Algorithm:

  For each **u** grid location{

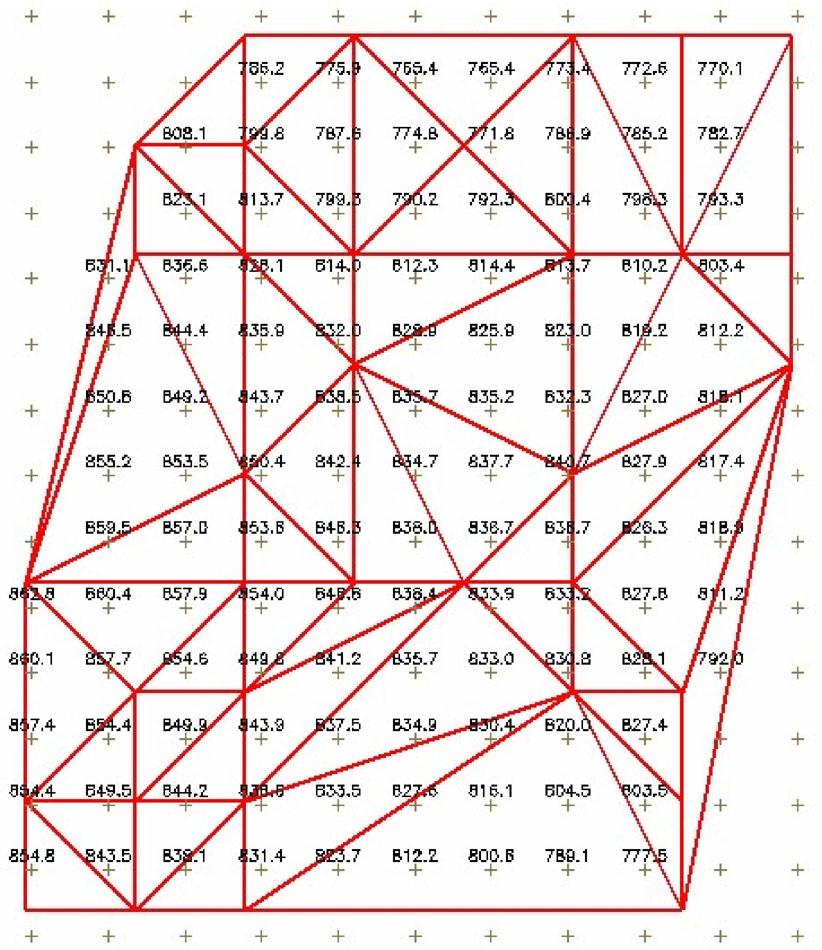
  Find the triangle *t* that contains **u**

  Calculate the z grid value at **u** using a linear (or other) function fitting *t*

  }

  Obs. If the **u** location is out of any triangle the location is associated to a dummy (not valid) value.

  It is also possible to construct a TIN from a rectangular GRID. How?




Grid from TIN construction

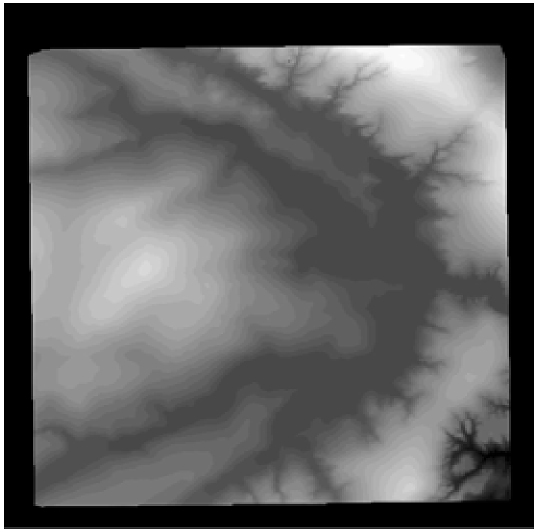
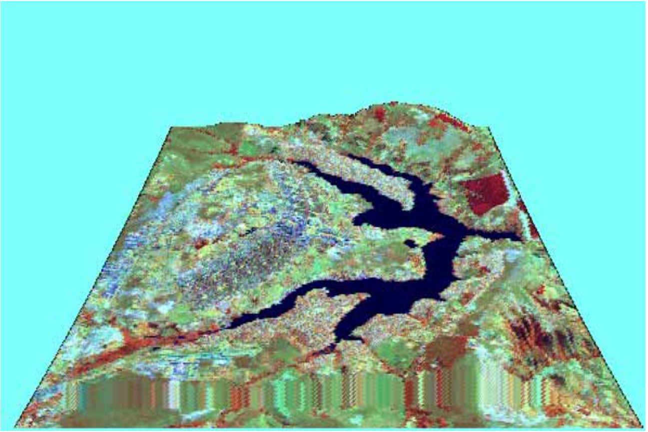
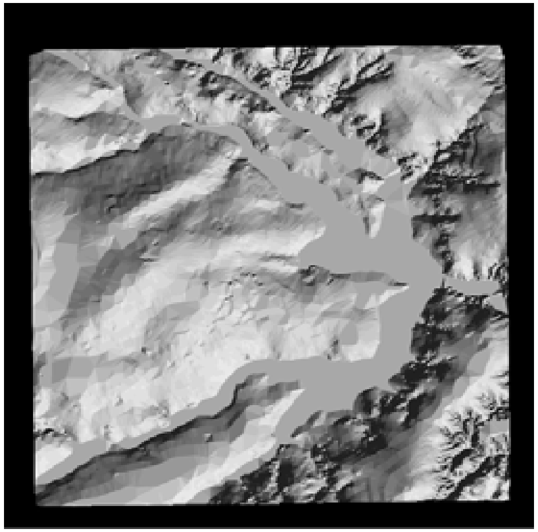
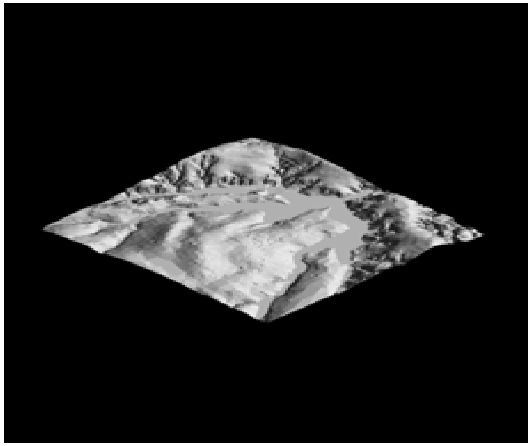
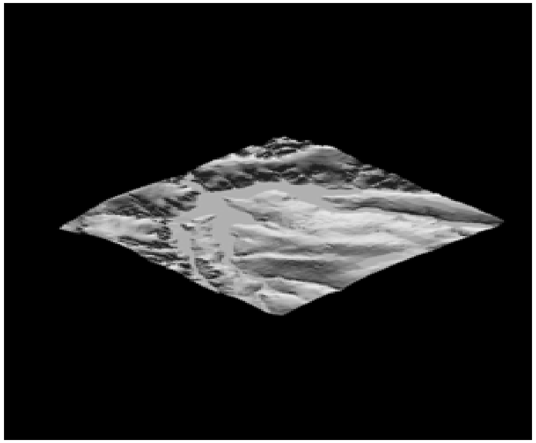
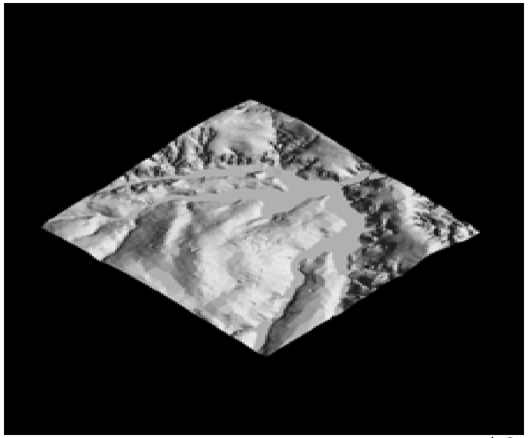
# Predictions with Deterministic Procedures

- **Grid x TIN comparison (table?)**

| | ***TIN*** | ***Rectangular Grid*** |
|---|---|---|
| Advantages | 1. Better to represent attributes with complex spatial variations<br>2. Allows incorporation of constraints (topographic top and bottom lines)<br>3. Better for quantitative analysis (measures, slope evaluation, volume calculations, etc.…) | 1. Easy structures and algorithms to be manipulated in computers<br>2. Better for qualitative analysis (visual)<br>3. Suitable for visualization with planar projections |
| Problems | 1. More complex structures and algorithms to be manipulated<br>2. Not suitable for visualizations with planar projections | 1. Representation of atributes with complex variations<br>2. Not suitable for quantitative analysis (loss of details) |

# Predictions with Deterministic Procedures

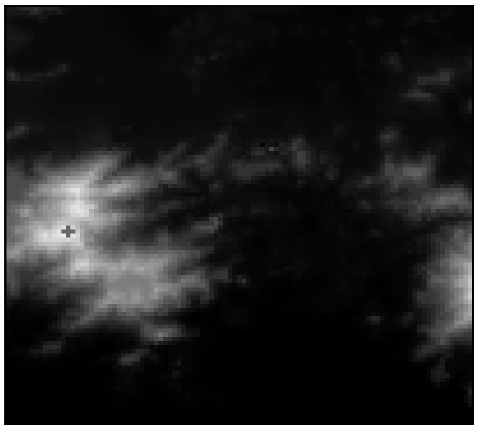
- **Examples Rectangular Regular applications (visualizations)**

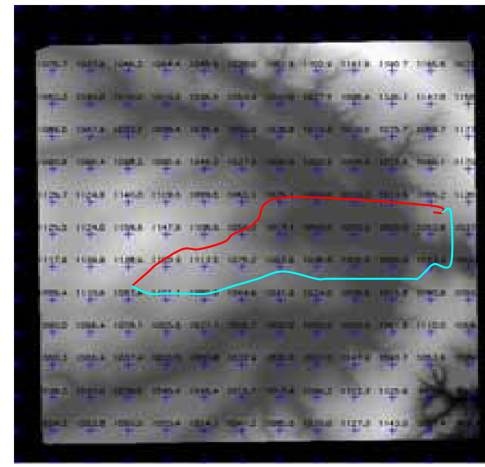
# Predictions with Deterministic Procedures

- **Examples of TIN applications (evaluations)**

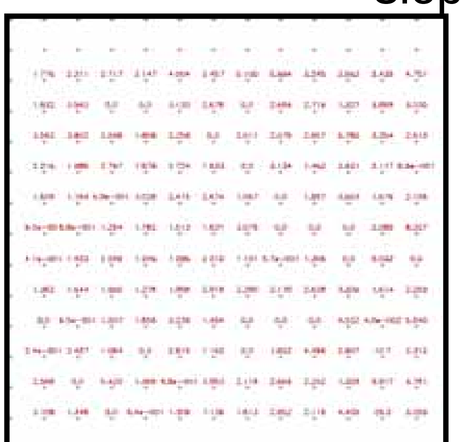
Visibility Analysis

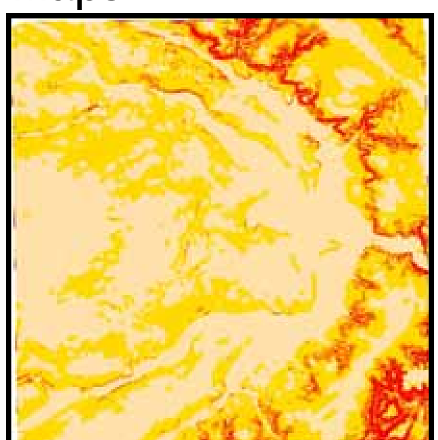


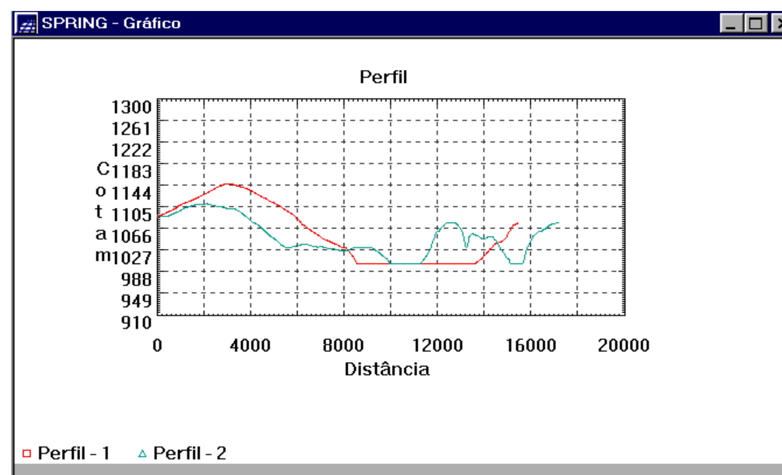
Profiles Determination



Slope Maps







# Predictions with Deterministic Procedures

- **Problems with deterministic procedures**

If the sample set is well distributed and too dense all the interpolators (deterministic or stochastic) perform satisfactory.

Deterministic procedures do not perform well when

the set has few samples

the set present clusters of samples

the spatial continuity of the attribute is not isotropic

For Rectangular Grid Modeling

Problems: Few samples, definition of radius of influence, number of neighbors and exponential parameter for IDW estimators.

For TIN Modeling

Problem: How triangles must be considered for estimation inside a triangle?

# Summary and Conclusions

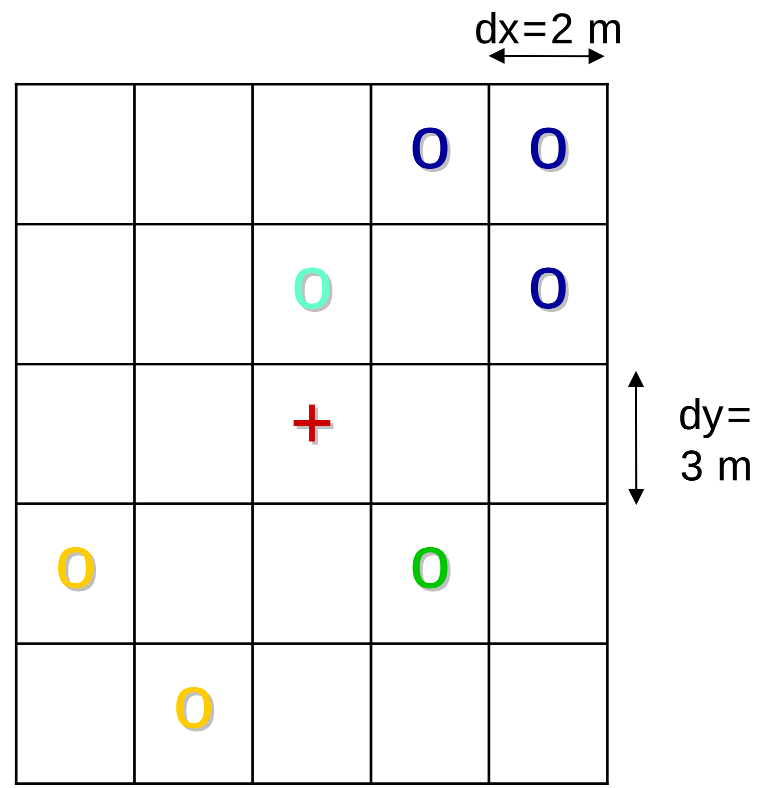
## Summary and Conclusions

- Deterministic estimators can be used to model spatial data.
- Current GISs allow users work with these tools mainly to Digital Terrain Modeling (DTM) tasks.
- Deterministic estimators perform better when the sample set is dense.
- When the sample set has few elements the deterministic models usually present some undesirable artifacts in their representations
- The user always should take care of (be worried with) the parameters used in deterministic estimations (black box).
- Usually deterministic modeling are not followed with uncertainty informations about the performed estimations

# Predictions with Deterministic Procedures

## Exercises

1. Run the Lab2 already available in the geostatistics course area of ISEGI online.
2. Given the following sample configuration estimate z values for the **u** location using the deterministic procedures (local means and TIN) you have learn in this presentation.
3. Report the results and advantages, or disadvantages, of each one.
4. In your opinion which one of the estimates procedures performs better to estimate the z value at the location **u**.
5. Send the reporter to the e-mail of the geostatistics professor before 25/10/2007.




o – 500  o – 800  + **u** location

o – 900  **o** - 1000

# Predictions with Deterministic Procedures

*END*

*of Presentation*