

Prostorne analize u *open source GIS* okruženju: R+SAGA+Google kartografski servisi

20-24 jun 2011. Građevinski fakultet, Univerzitet u Beogradu

Povezivanje geostatistike i GIS-a

Branislav Bajat

bajat@grf.bg.ac.rs

Geografski model podataka

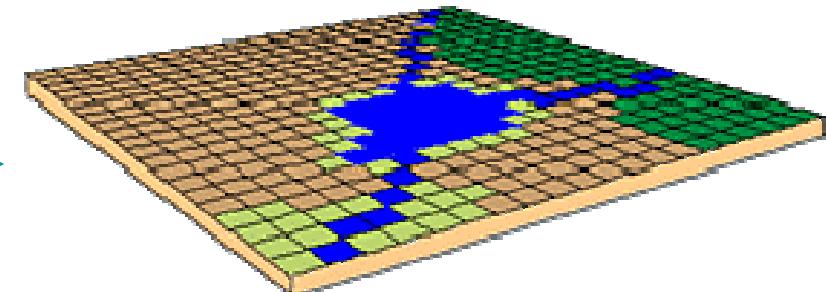
- *Geografski model podataka*- formalizovana shema prikazivanja podataka koji poseduju lokaciju i attribute.
- Metode prikazivanja geografskog prostora
 - Rasterski Model
 - Vektorski Model

Konceptualni model

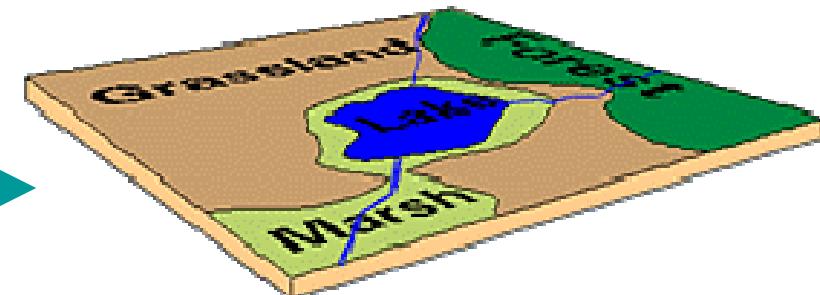
- Model baziran na poljima, gde zamišljamo da se ispitivani atribut menja u prostoru kao neka neprekidna matematička funkcija ili *polje*.
- Model baziran na entitetima gde smatramo da je prostor sastavljen od *entiteta* koji se opisuju svojim atributima ili svojstvima, i čije se lokacije kartiraju pomoću geometrijskog koordinatnog sistema.

Prikazivanje prostornih elemenata

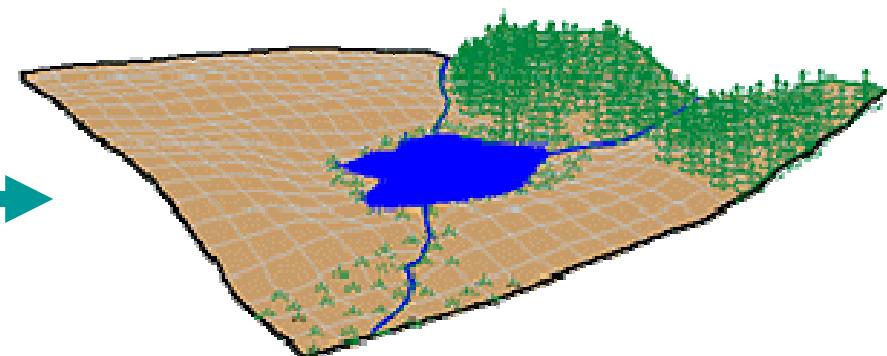
- RASTER



- VEKTOR

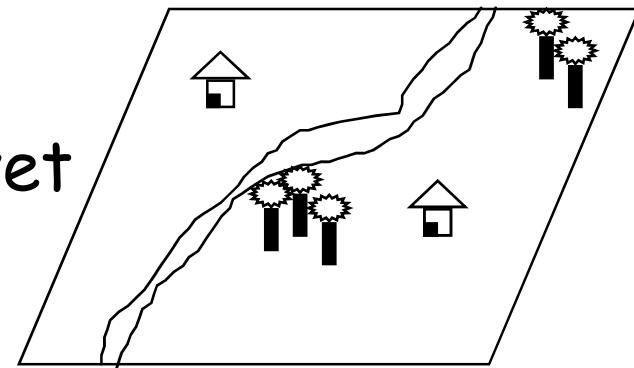


- Stvarni svet



Koncept Vektor i Raster

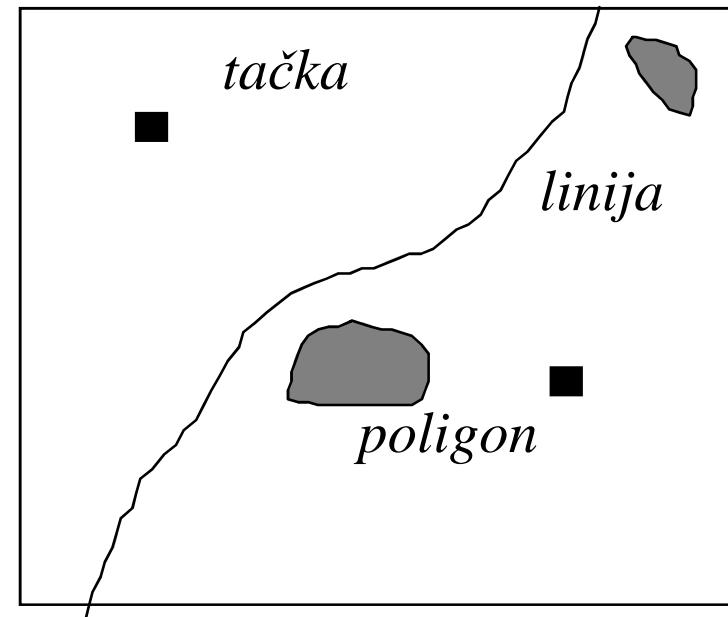
Realni svet



Rasterski prikaz

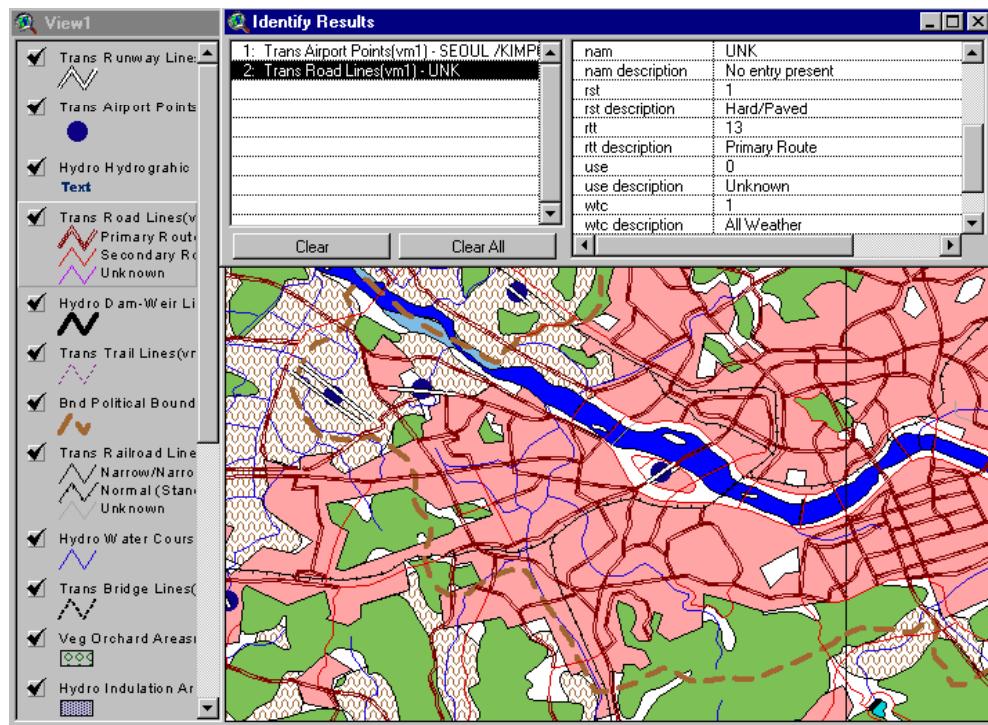
	0	1	2	3	4	5	6	7	8	9
0						R	T			
1						R			T	
2	H					R				
3						R				
4				R	R					
5			R							
6	R	T	T	T	H					
7	R	T	T	T						
8	R									
9	R									

Vektorski prikaz

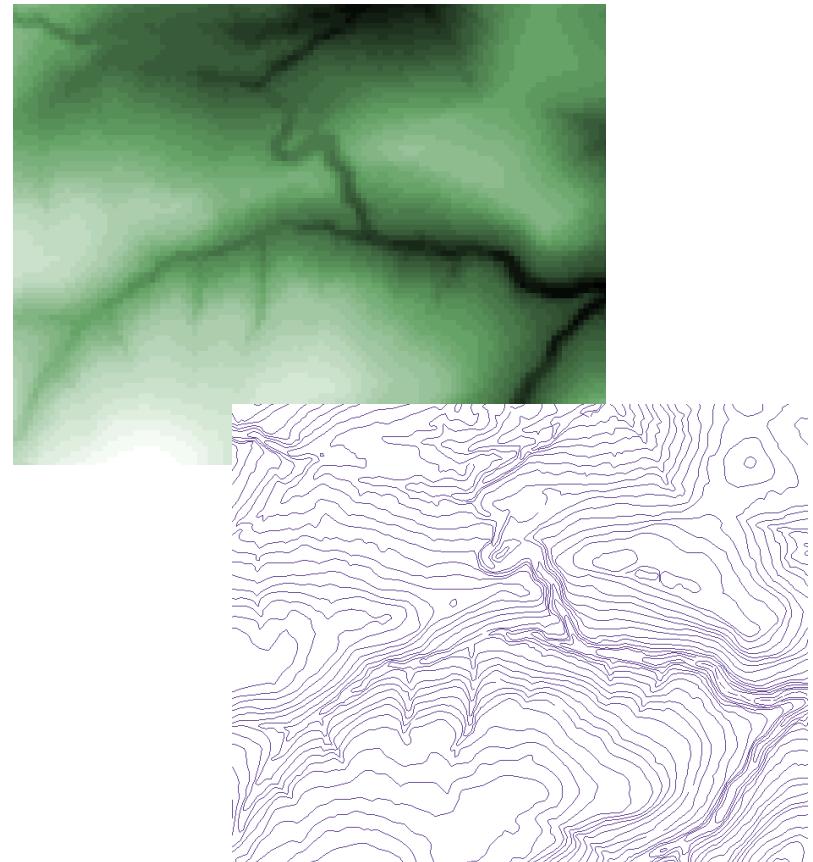


Konceptualni modeli prostornih fenomena u GIS-u

a) EntitetSKI model



b) Model polja



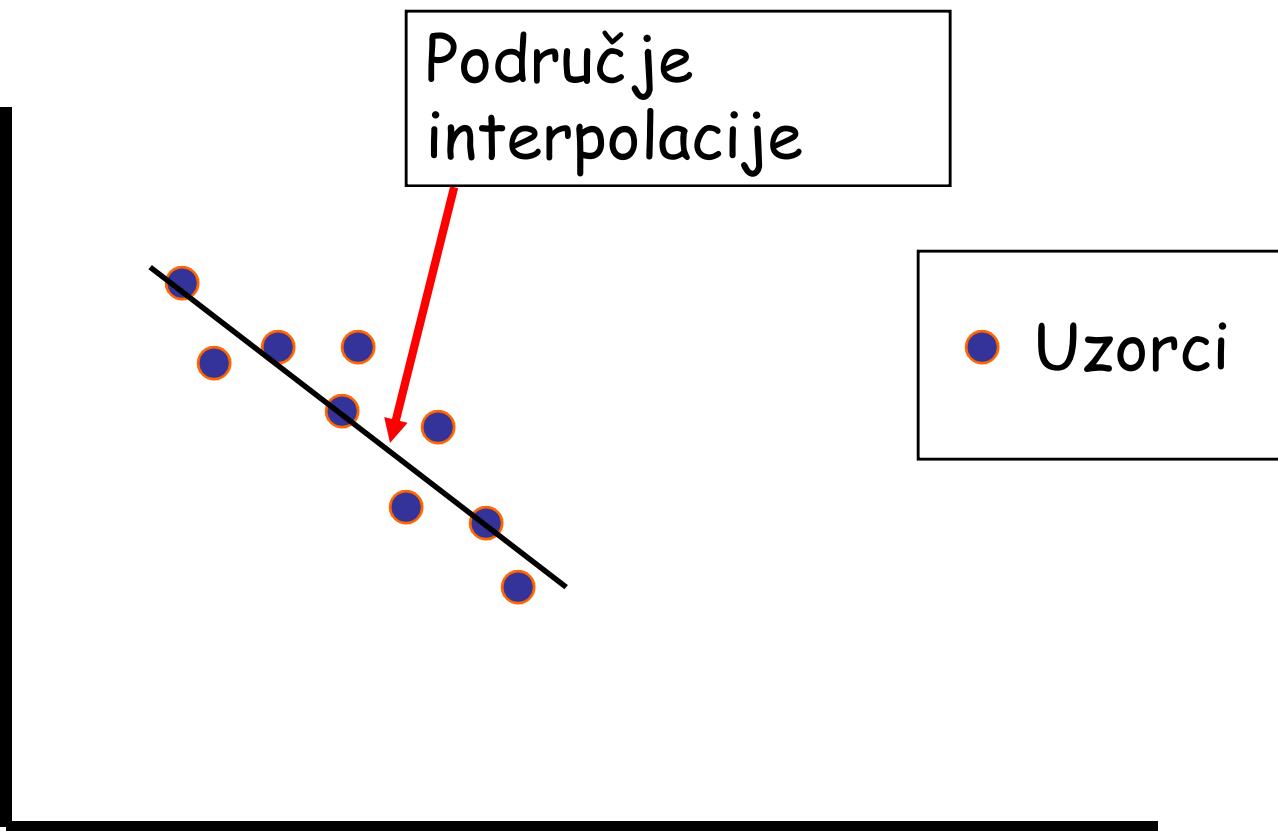
- Interpolacija

Procena vrednosti atributa unutar područja na kojem imamo merene uzorke.

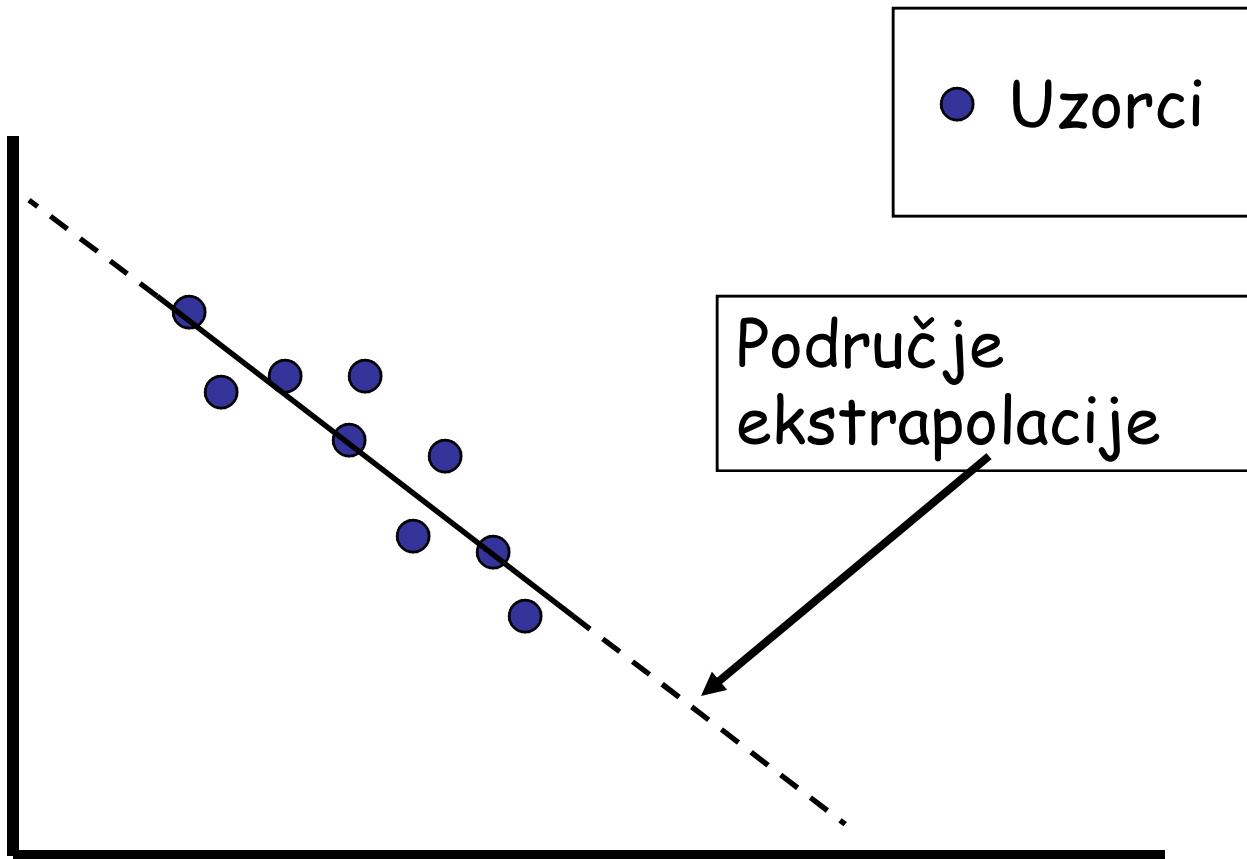
- Ekstrapolacija

Procena vrednosti atributa van područja na kojem imamo merene uzorke.

Interpolacija

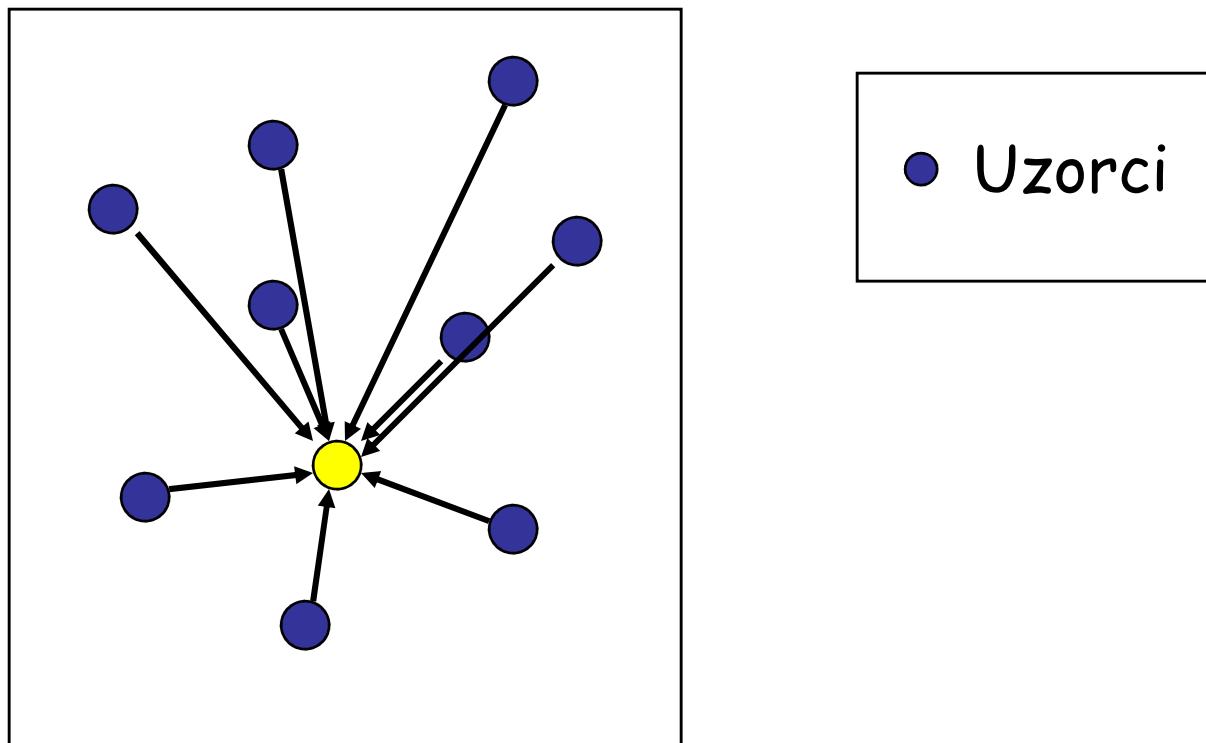


Ekstrapolacija



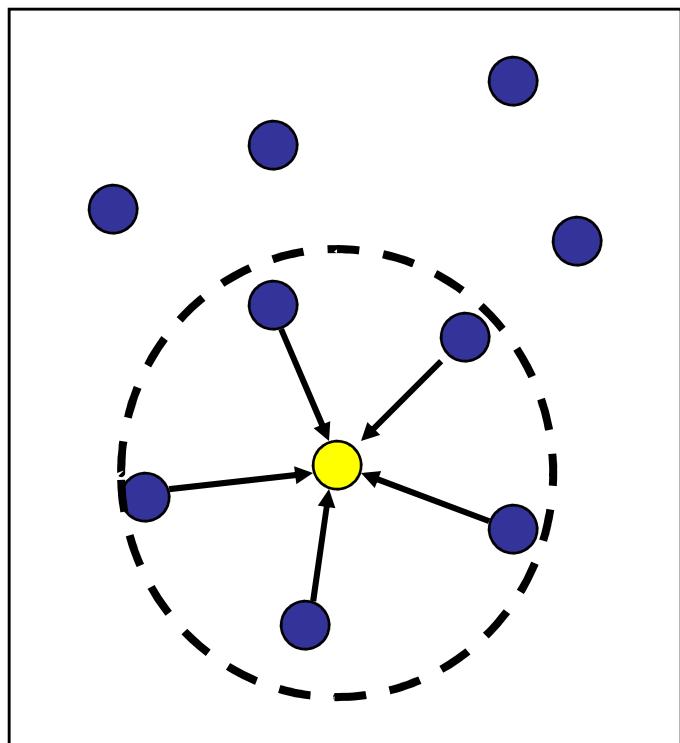
Globalna Interpolacija

Koristi **sve** poznate uzorke na području kako bi se procenila vrednost na traženoj lokaciji



Lokana Interpolacija

Koristi samo susedne uzorke za procenu vrednosti na traženoj lokaciji



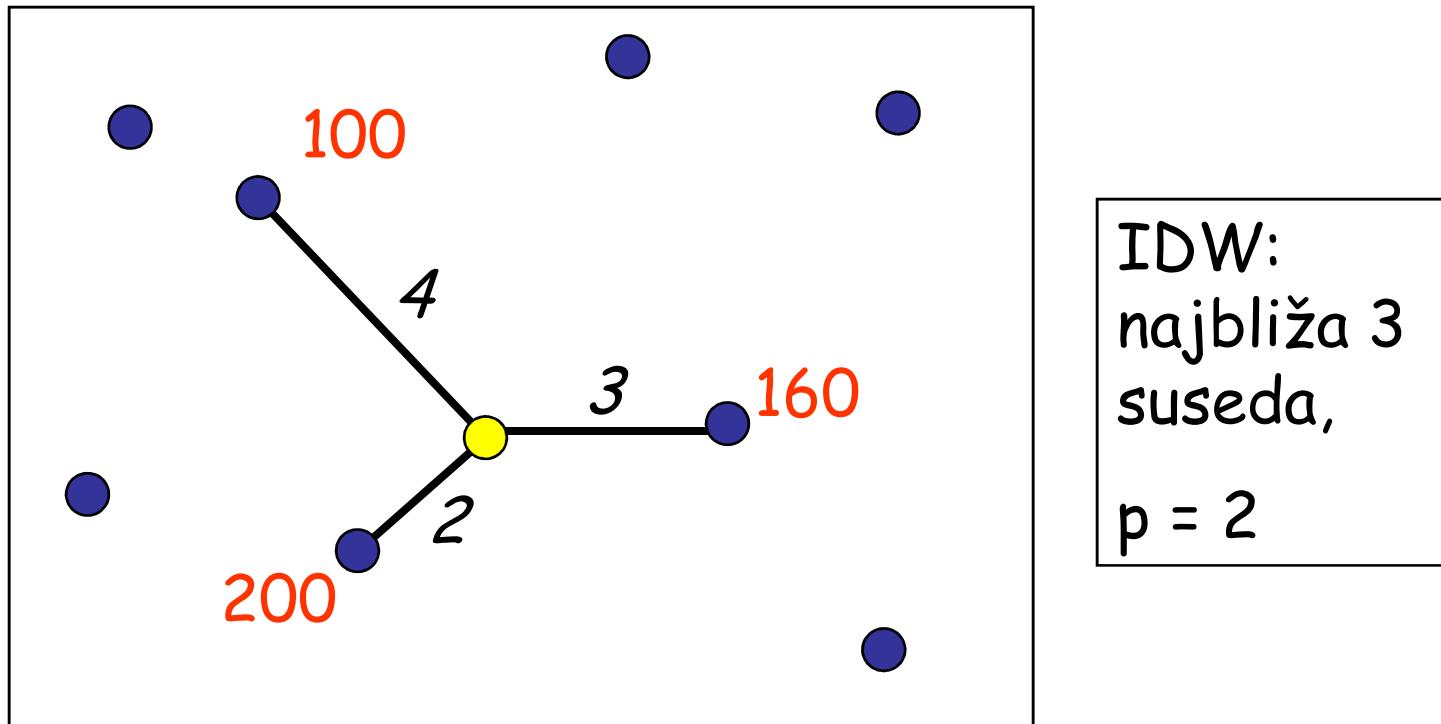
• Uzorci

Susedni uzorci se biraju tako što se zada n broj najbližih uzoraka ili se zada radijus pretraživanja

**Inverse Distance Weighted
(IDW)**

Metoda inverzne distance

Inverse Distance Weighted (primer)

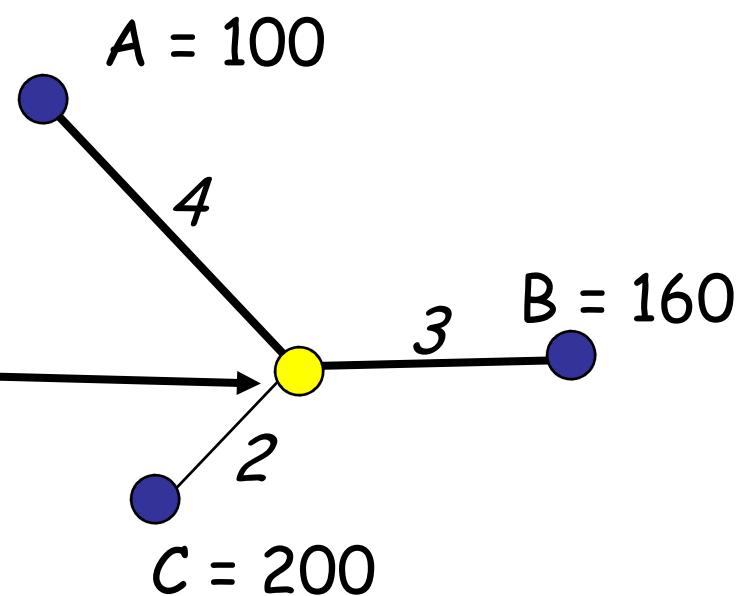


Inverse Distance Weighted (primer)

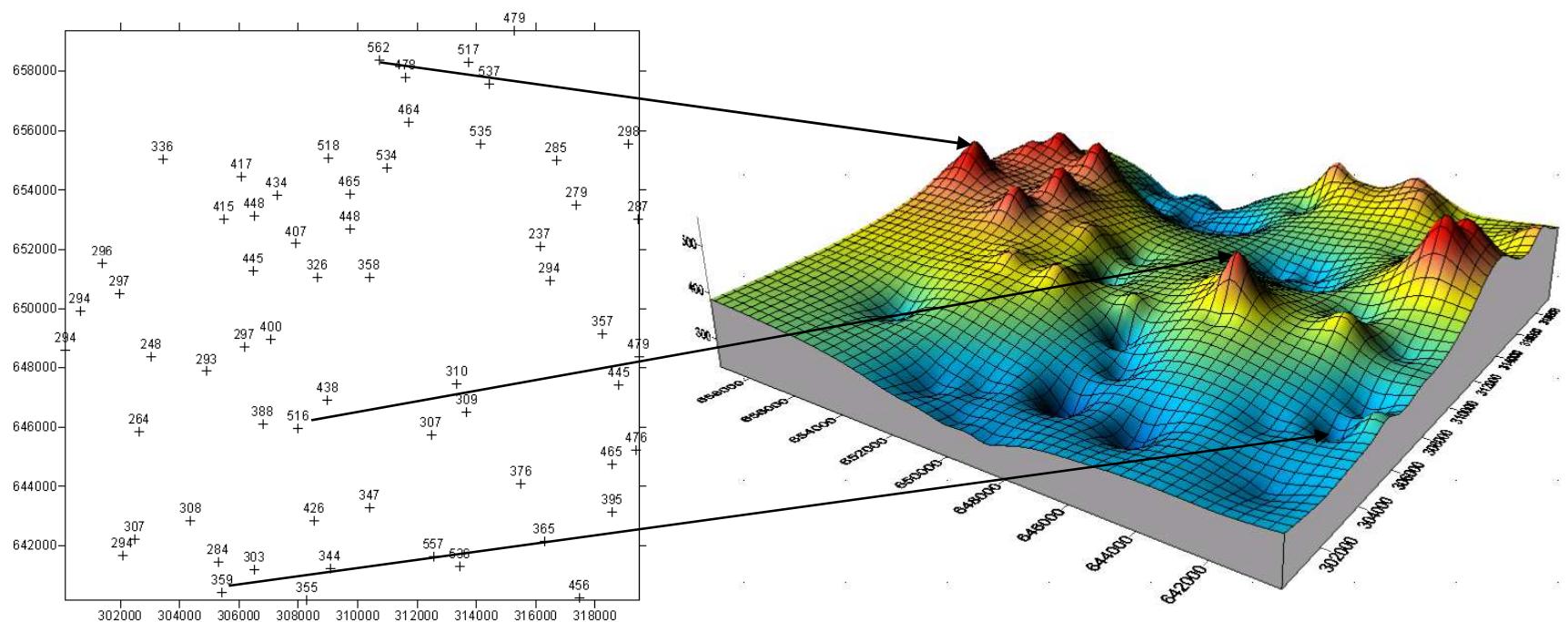
	Težine	Težine × vrednost
A	$1 / (4^2) = .0625$	$.0625 \times 100 = 6.25$
B	$1 / (3^2) = .1111$	$.1111 \times 160 = 17.76$
C	$1 / (2^2) = .2500$	$.2500 \times 200 = 50.00$
Suma = .4236		

$$6.25 + 17.76 + 50.00 = 74.01$$

$$74.01 / .4236 = 175$$



Nedostaci IDW-a

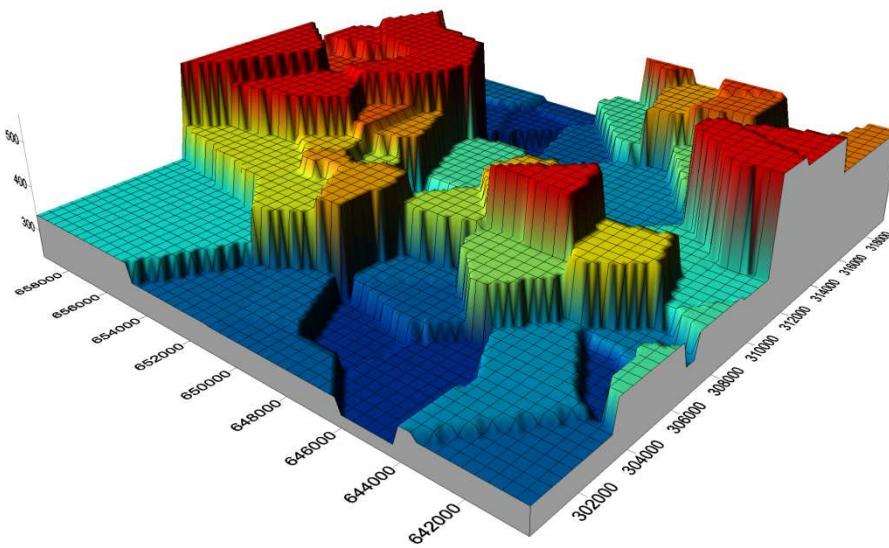


Uzorci

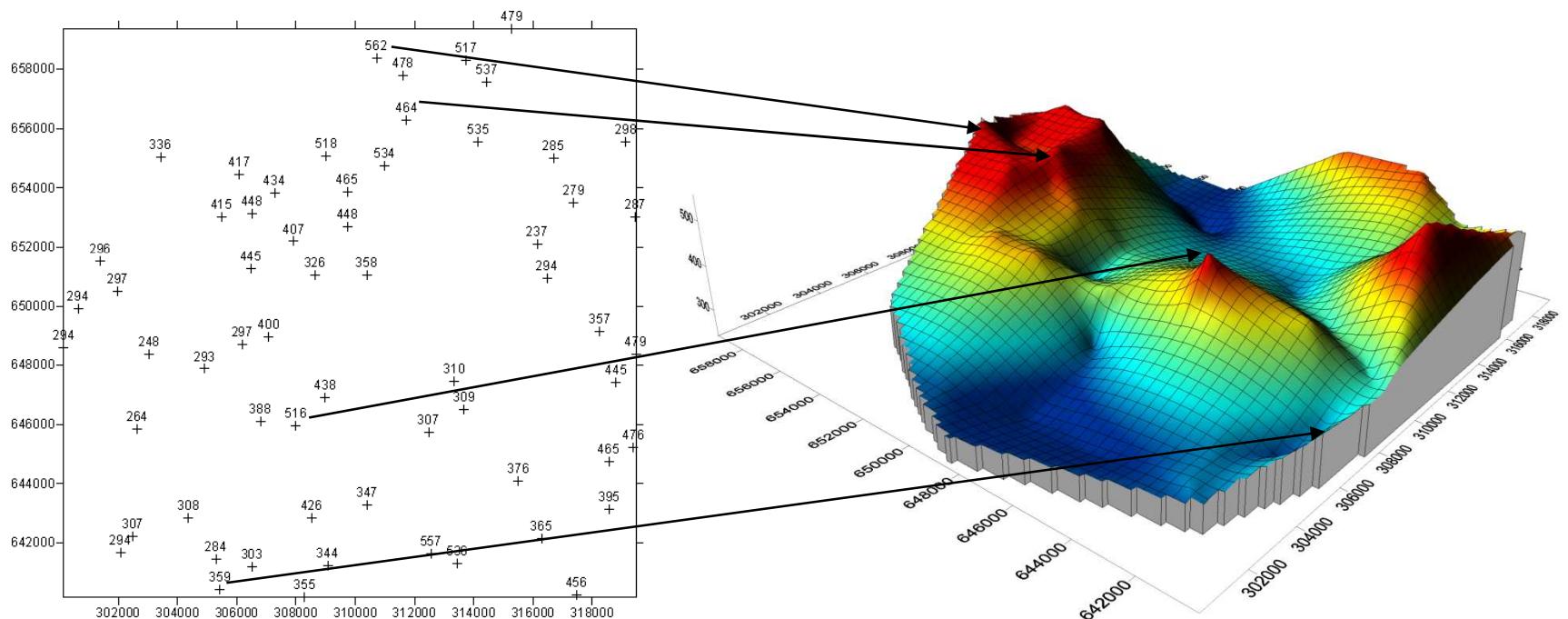
Interpolovaná površ

Pregled ostalih
metoda
interpolacije

Nearest neighbour/najbliži sused



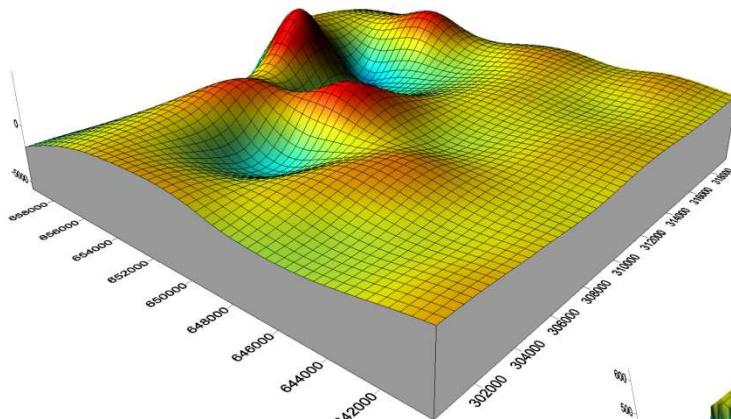
Natural neighbour/prirodni sused



Uzorci

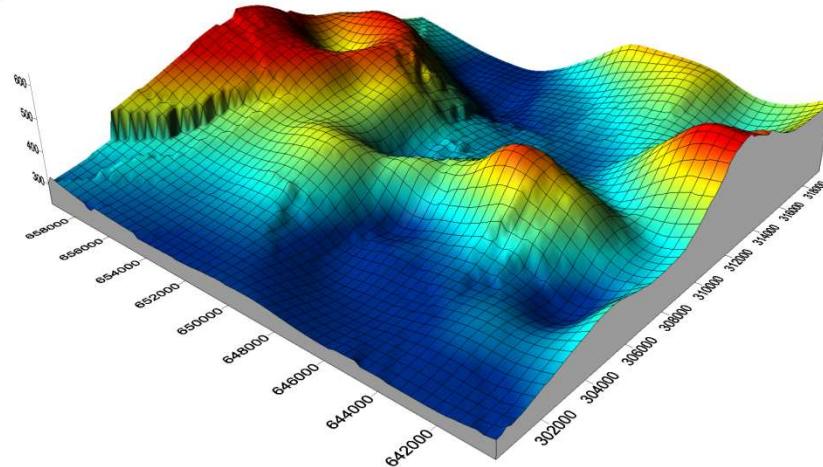
Plate spline/spline funkcija

Korišćene sve 62 tačke

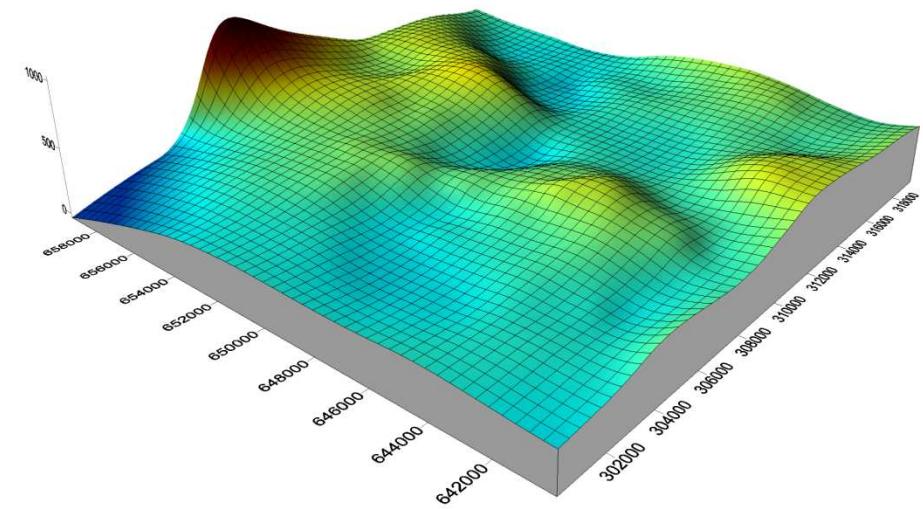


$$\phi(r) = (c^2 + r^2) \ln(c^2 + r^2)$$

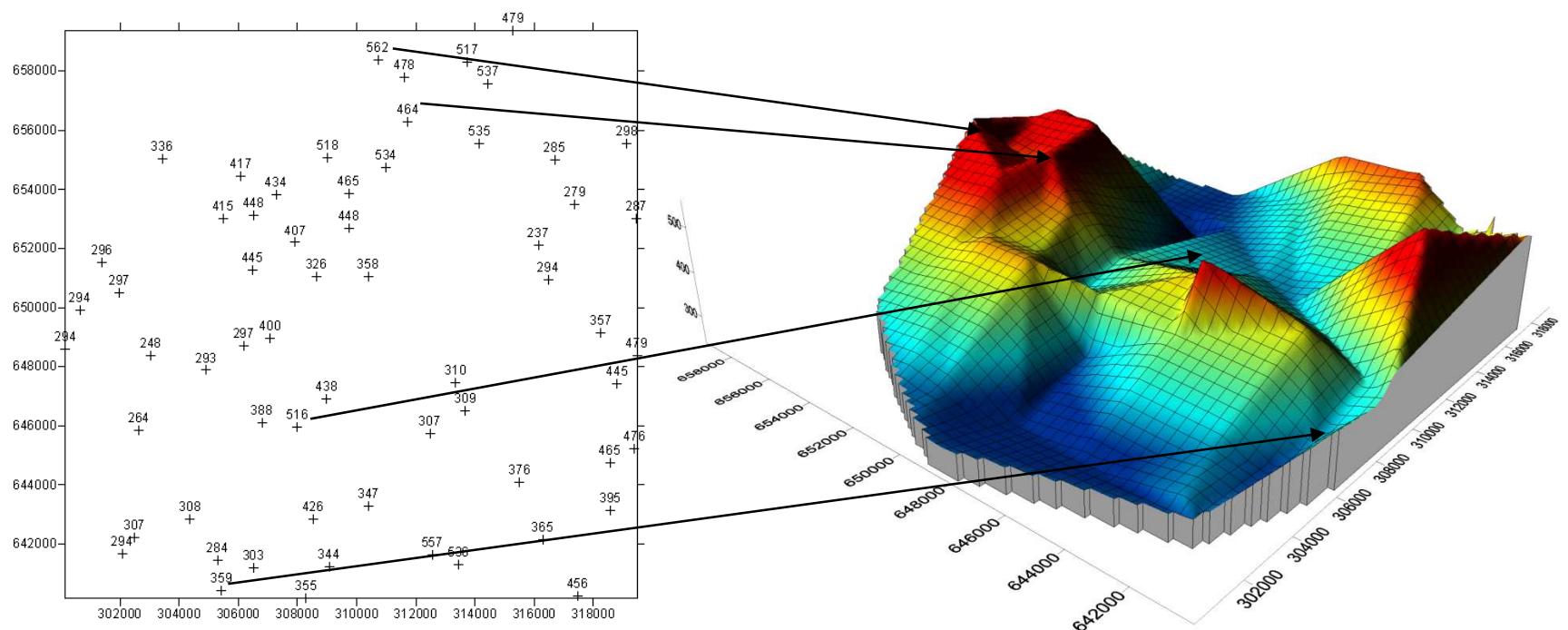
Korišćeno 12 susednih tačaka



Modified Shepard/ modifikovaná Šepardova

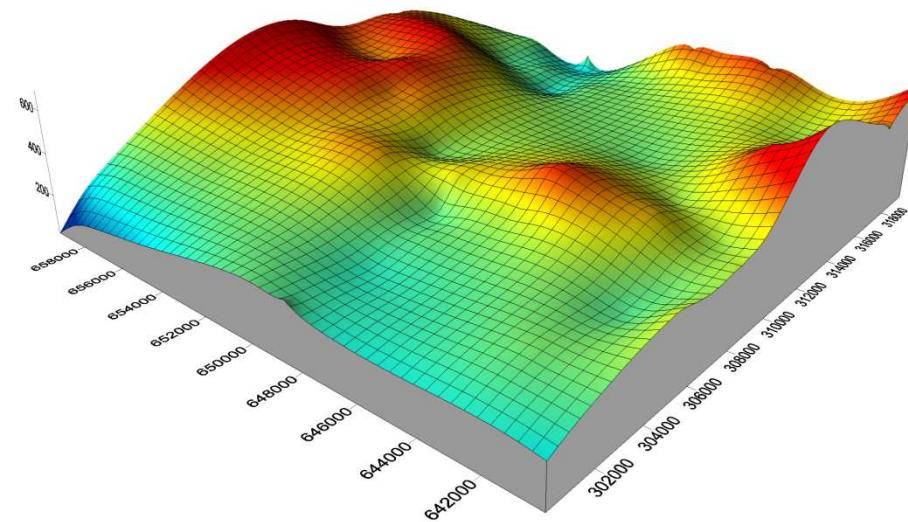


Triangulacija sa linearom interpolacijom

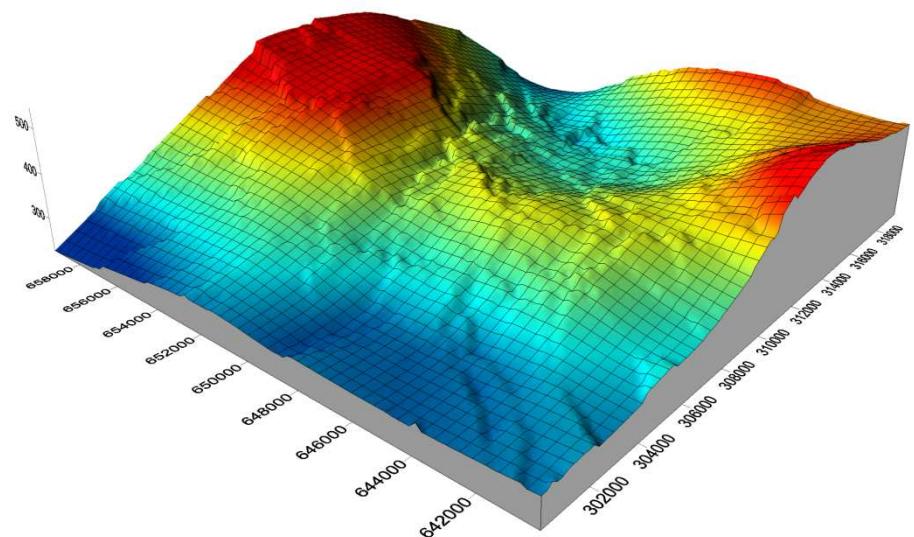


Uzorci

Minimum curvature/ minimalna zakrivljenost



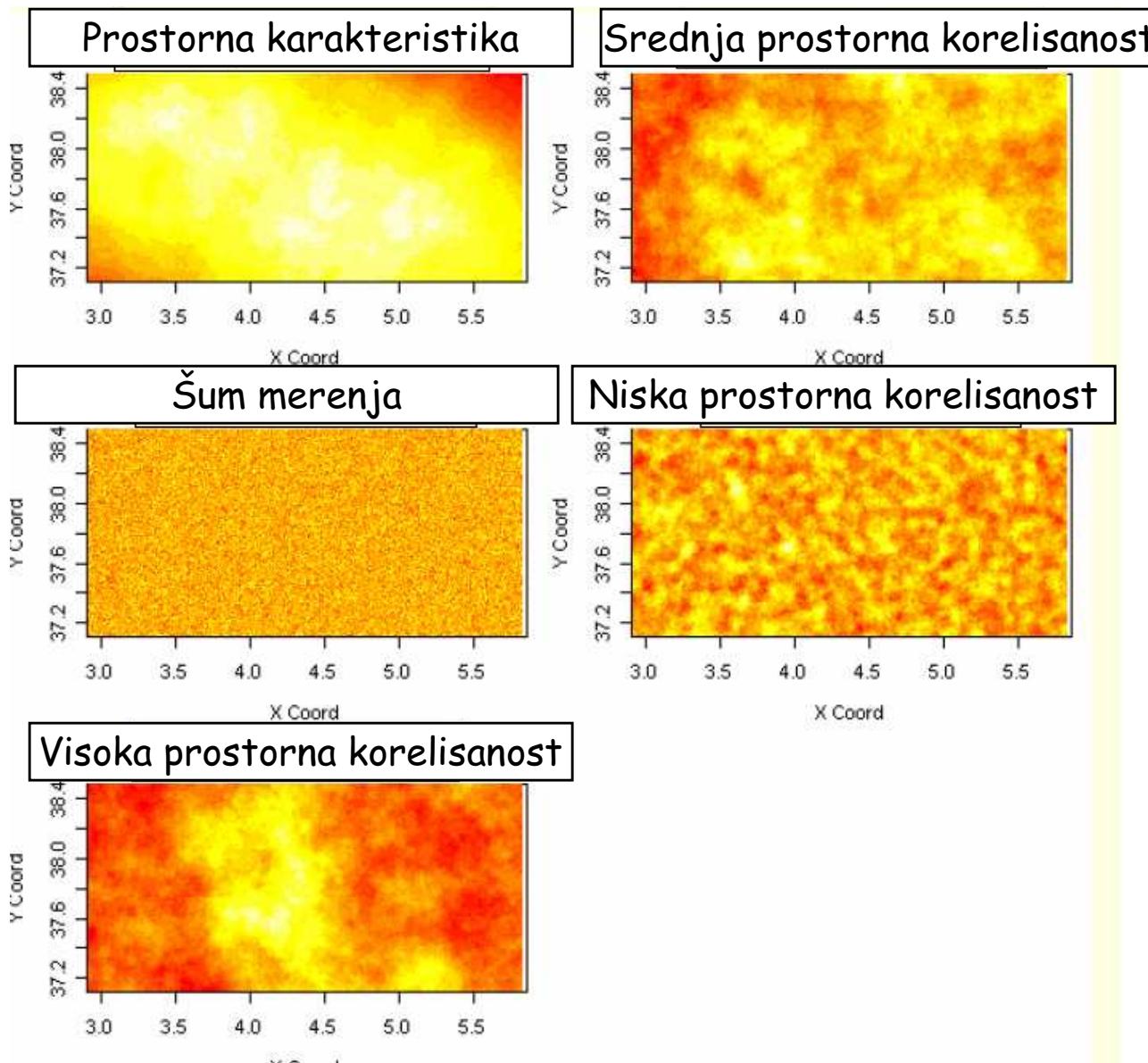
Local polynomial / lokalni polinomi



Geostatistika

- *Geostatistika*, nudi mogućnost analitičke obrade i opisa koncepta prostornog kontinuiteta, koji prisutan u svim geonaukama .
- *Prvi zakon u geografiji*, glasi da je sve u nekom odnosu sa ostalim, ali da su vrednosti koje su prostorno bliže sličnije od onih udaljenijih.
- Geostatistika za razliku od klasične statistike koja obrađuje slučajno promenljive veličine, nije skoncentrisana samo na vrednosti promenljive, već i na položaj promenljive u prostoru, odakle i potiče termin *regionalizovane promenljive*.

Primeri prostorno promenljivih veličina



- Kriging je jedna od najpoznatijih interpolacionih geostatističkih metoda.
- Naziv ove metode vezan je za ime D.G. Krige-a koji je prvi koristio prostornu korelaciju u valorizaciji rudarskih ležišta.
- Tehniku računanja ove metode postavio je francuski geomorfolog Georges Matheron, razvijajući teoriju regionalizovane promenljive.
- Matheron je ovu metodu razvio otrprilike u isto vreme kada je prof. Moritz razvio kolokaciju metodom najmanjih kvadrata, krajem šezdesetih godina dvadesetog veka.
- Obe metode su bazirane na stohastičkim procesima i praktično daju iste rezultate.

Kriging je interpolaciona metoda uz koju se često vezuje skraćenica "*BLUE*", koja označava *najbolju linearu nepomerenu ocenu* (best linear unbiased estimation).

- Najbolja jer nastoji da minimalizuje disperziju grešaka.
- On je linearna ocena pošto u njegovim formulama figuriše linearna kombinacija merenih podataka.
- Kriging metoda je "nepomerena" ocena jer je očekivana srednja vrednost ocene grešaka jednaka 0.

Semivariance

Poluvariijansa

Poluvarijansa

- Teorija regionalizovane promenljive koristi ***poluvarijansu*** da bi iskazala stepen uzajamne zavisnosti između između tačaka (uzoraka) na nekoj površi.
- Poluvarijansa je polovina vrednosti varijanse razlika vrednosti između svih uzoraka na konstantnom rastojanju.

Poluvarijansa je mera stepena prostorne zavisnosti između uzoraka

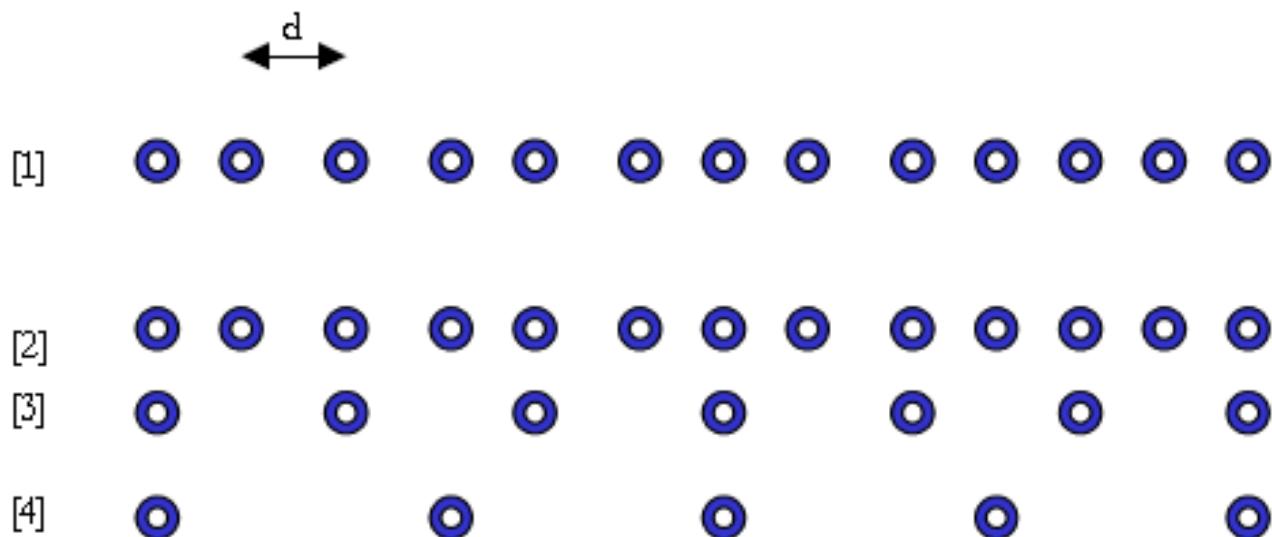
Poluvarijansa

- Vrednost poluvarijanse između tačaka zavisi od rastojanja između njih. Manje rastojanje rezultira manjom vrednošću poluvarijanse, a veće rastojanje većom vrednošću.

Računanje poluvarijanse (Između tačaka na pravilnom rastojanju)

- Razmotrimo uzorke pravilno udaljene (d), poluvarijansa se može odrediti za sva rastojanja h koja su umnožak koraka (d):

$$\gamma(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} (z_i - z_{i+h})^2$$



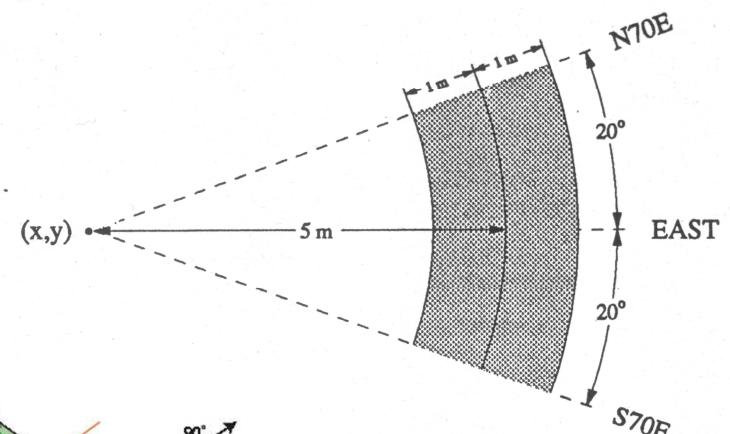
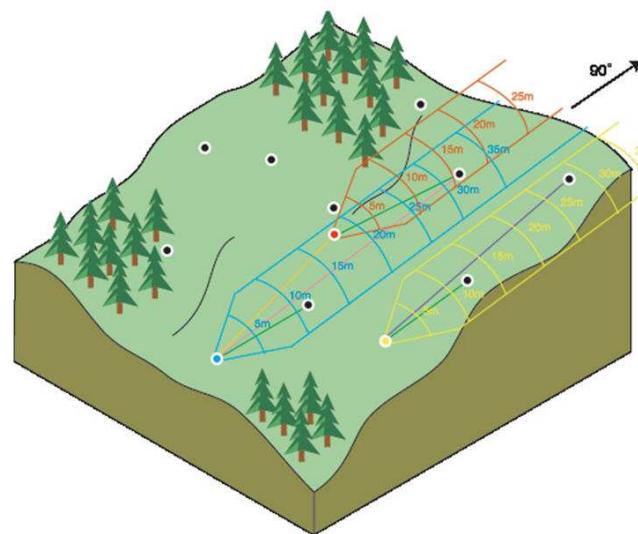
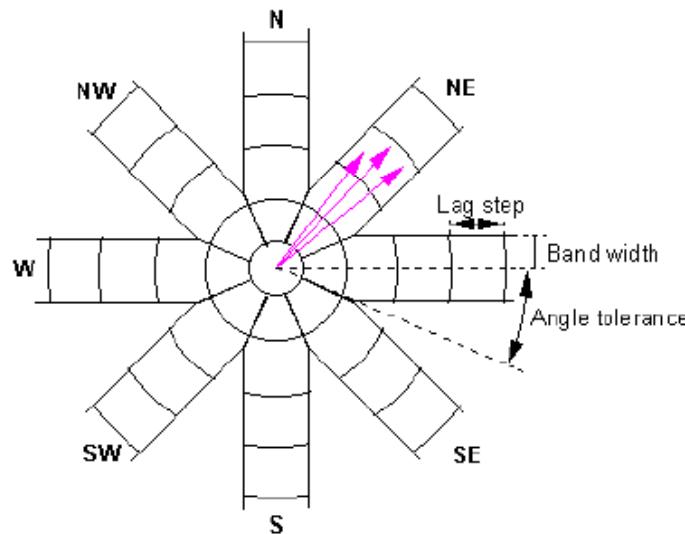
Poluvrijansa

$$\gamma(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} (z_i - z_{i+h})^2$$

- Z_i izmerena vrednost regionalizovane promenljive na lokaciji i ,
- Z_{i+h} je druga vrednost izmerena na lokaciji $i+h$ (h =umnožak rastojanja d)
- N_h broj parova unutar intervala h

Računanje poluvarejanse (Između tačaka na nepravilnom rastojanju)

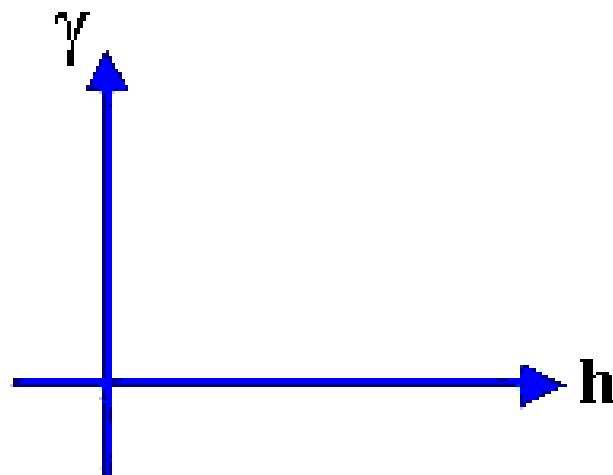
- U praksi je veom teško naići na tačke koje se nalaze na striktnom rastojanju [d].
- Zato se rastojanja grupišu u klase.



Variogram

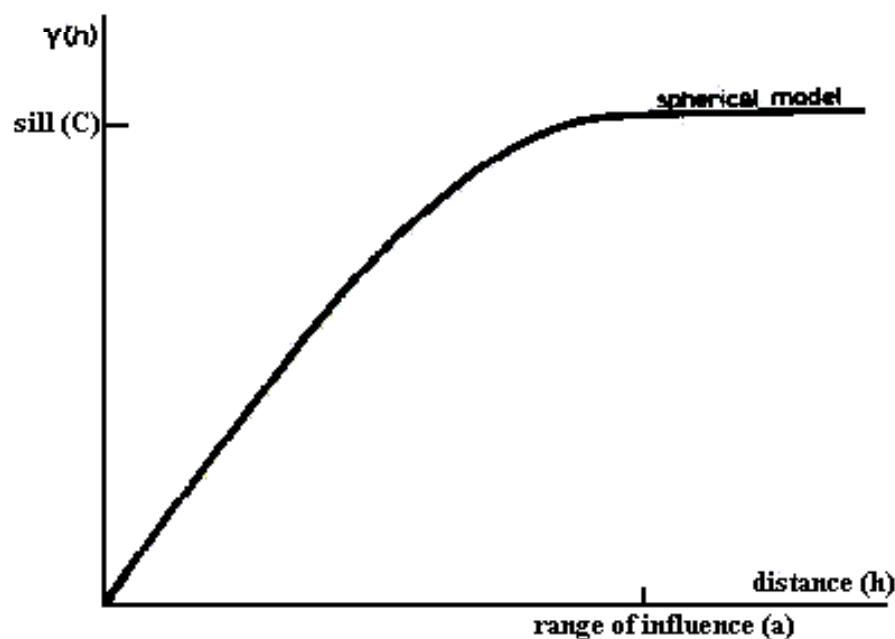
Variogram

- Predstavlja grafik poluvrijansi u funkciji od rastojanja između uzoraka.

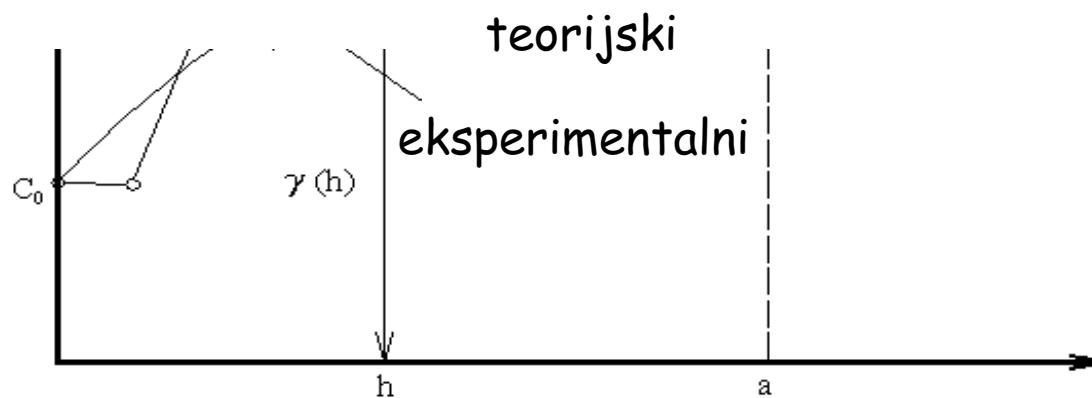


Variogram

- Poluvarijansa na rastojanju $d = 0$ trebala bi da bude jednaka nuli, pošto bi to teorijski bila razlika vrednosti promenljive u istoj tački.
- Kako su tačke udaljenije jedna od druge, tako bi i vrednost poluvarijanse trebalo da raste.



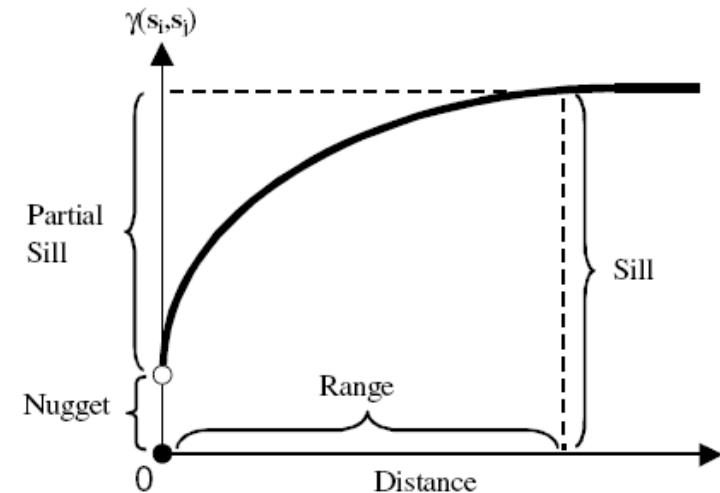
Modeliranje variograma



$$\gamma(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} (z_i - z_{i+h})^2$$

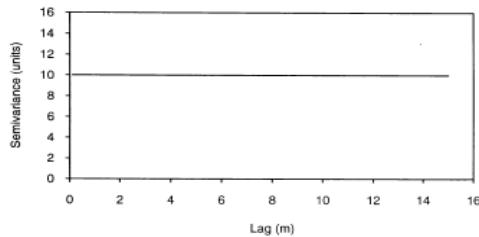
Poluvariijansa

- **Prag/Sill:** Maksimalna vrednost (asimptotski maksimum) koju poluvariijansa može dostići
- **Domet/Range:** Definiše unutar kog rastojanja su uzorci međusobno zavisni (korelaciono).
- **Grumen/Nugget:** varijansa greške merenja u kombinaciji sa vrednošću varijacija na rastojanjima koja su manja od rastojanja između merenih tačaka.

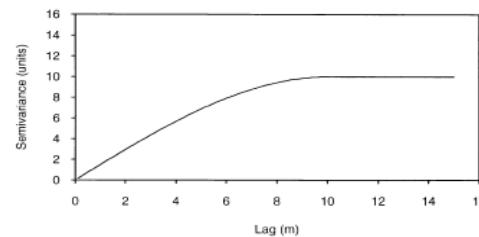


Modeli variogramma

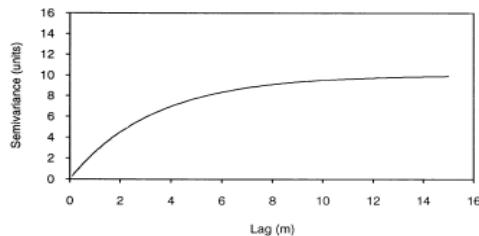
Nugget model



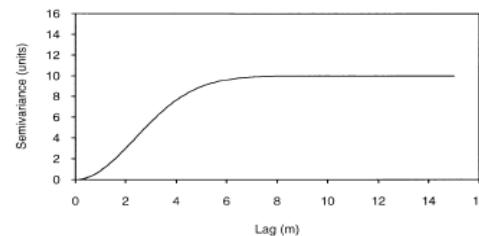
Spherical model



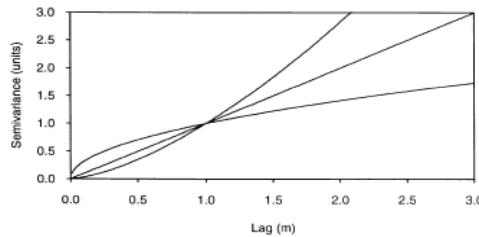
Exponential model



Gaussian model



Power model

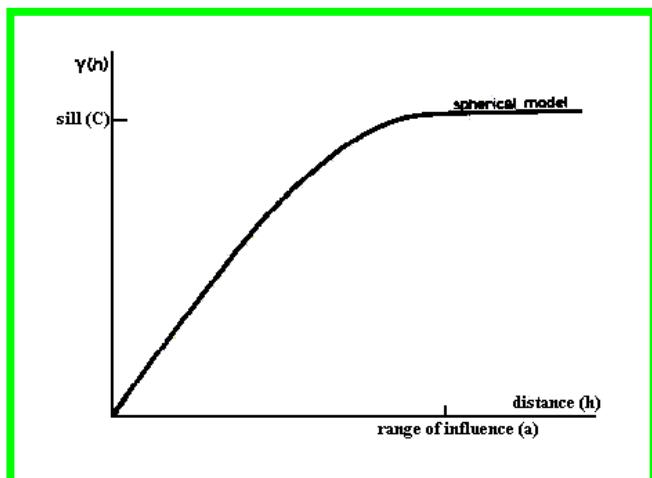


Modeli variogram

Model	Formula	Primedbe
Nugget effect	$\gamma(0) = C_0$	Variogram je konstantna veličina.
Linear	$\gamma(h) = C_1(h)$	Nema praga. Obično se kombinuje sa drugim funkcijama. Često se koristi kao funkcija nagiba.
Exponential Exp()	$\gamma(h) = C_1 \left(1 - e^{-kh}\right)$	k je konstanta, obično $k=1$ or $k=3$. Koristi se kod variograma sa velikim grumenskim efektom, koji polako rastu do vrednosti praga.
Spherical Sph()	$\gamma(h) = C_1 \left(\frac{3h}{2} - \frac{1}{2}h^3\right), h < 1$ $\gamma(h) = C_1, h \geq 1$	Koriste se u slučaju kada je grumenski efekat mali ali je bitan. Često se pojavljuje kao default model kod mnogih programskih paketa.

Variogrami (Modeli)

- Sferni (spherical) model.
- a domet variograma.
- C je prag.

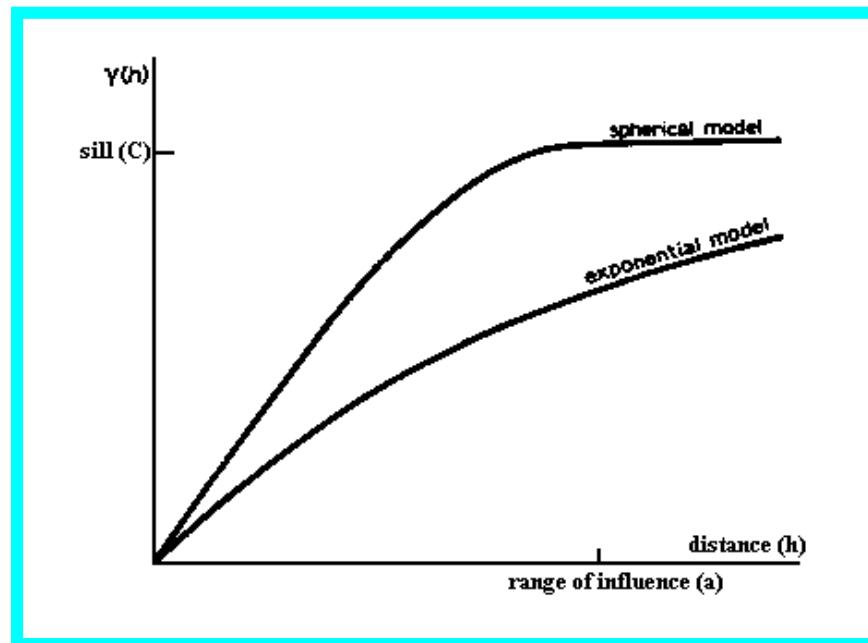


$$\gamma(h) = \begin{cases} C\left(\frac{3}{2}\frac{h}{a} - \frac{1}{2}\frac{h^3}{a^3}\right) & \text{gde } h \leq a \\ C & \text{gde } h \geq a \end{cases}$$

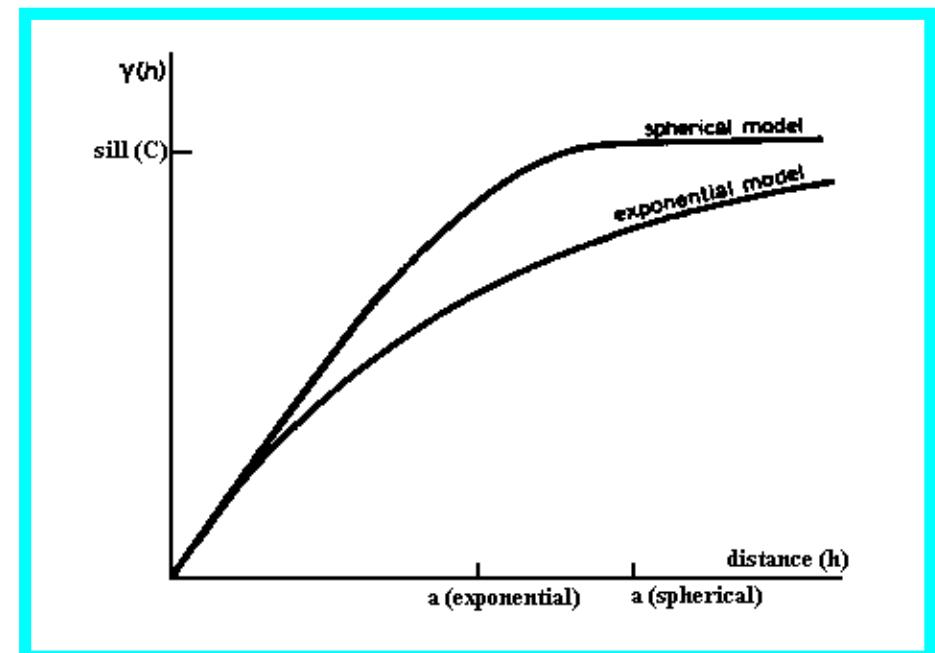
Variogrami (Modeli)

Eksponencijalni (Exponential) Model

$$\gamma(h) = C \left(1 - e^{-h/a} \right)$$



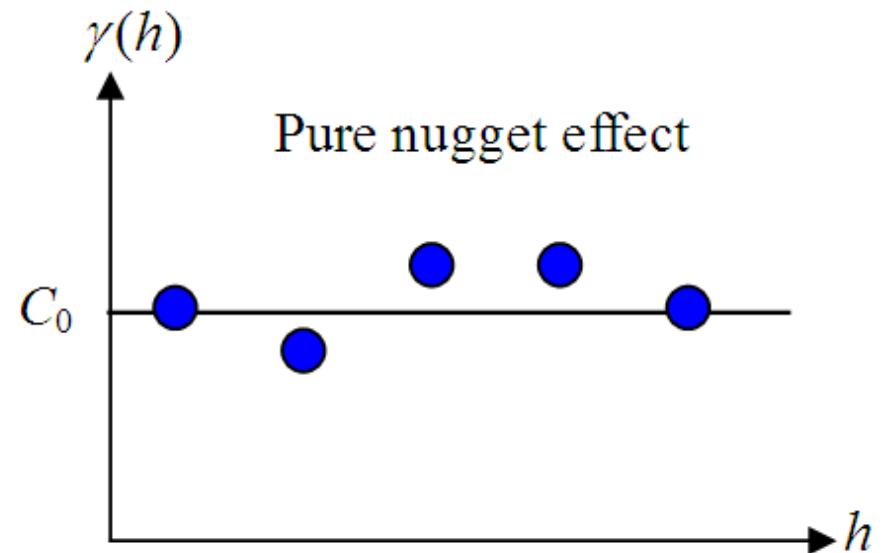
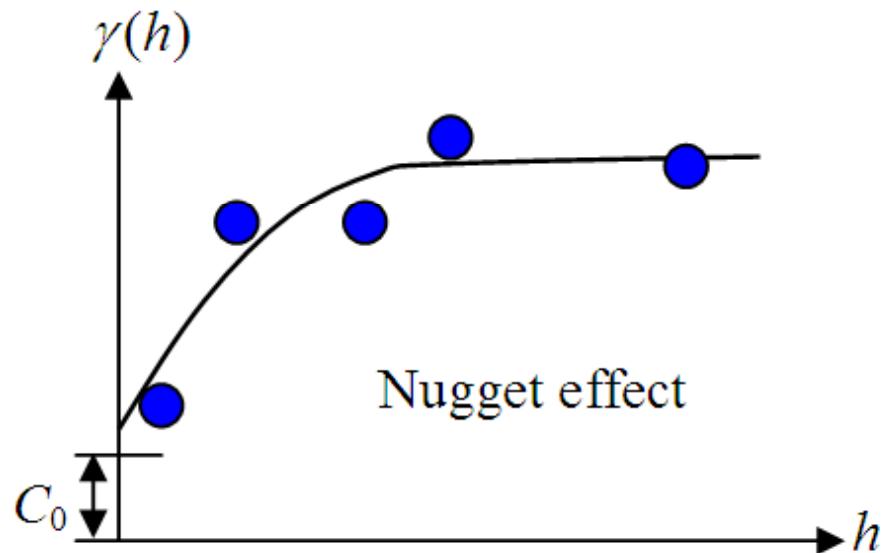
sferni i eksponencijalni sa istim pragom i dometom



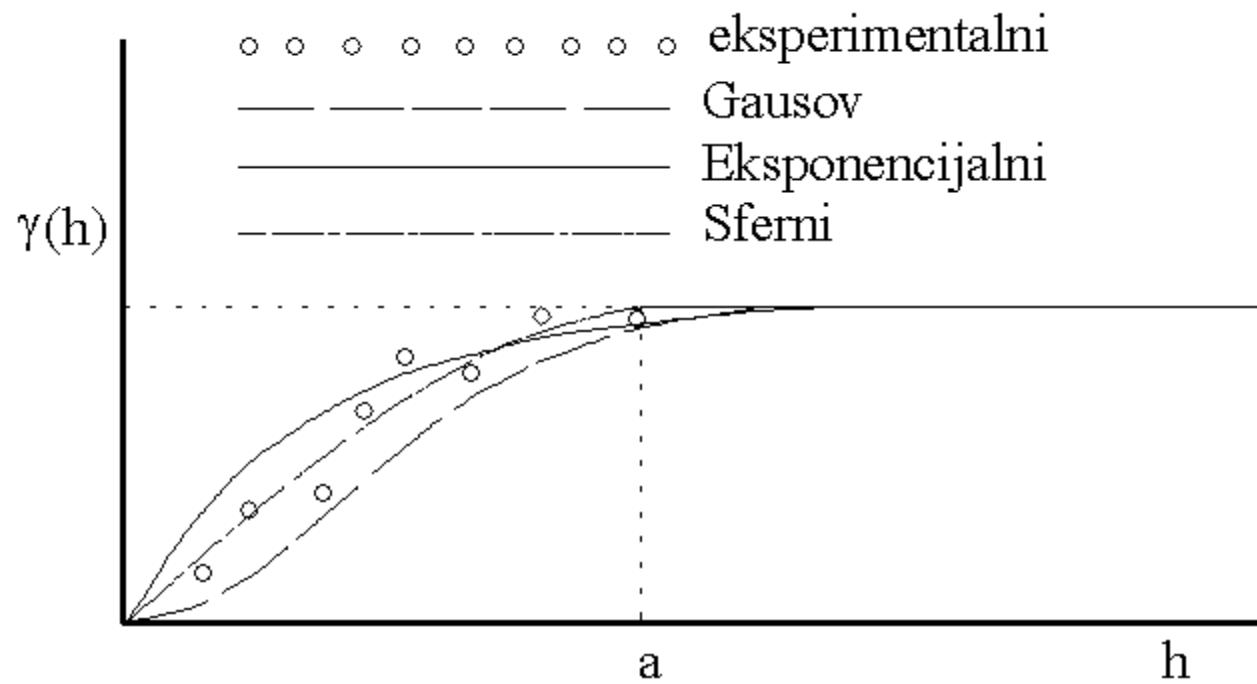
sferni i eksponencijalni sa istim pragom i istim početnim nagibom

Variogrami (Modeli)

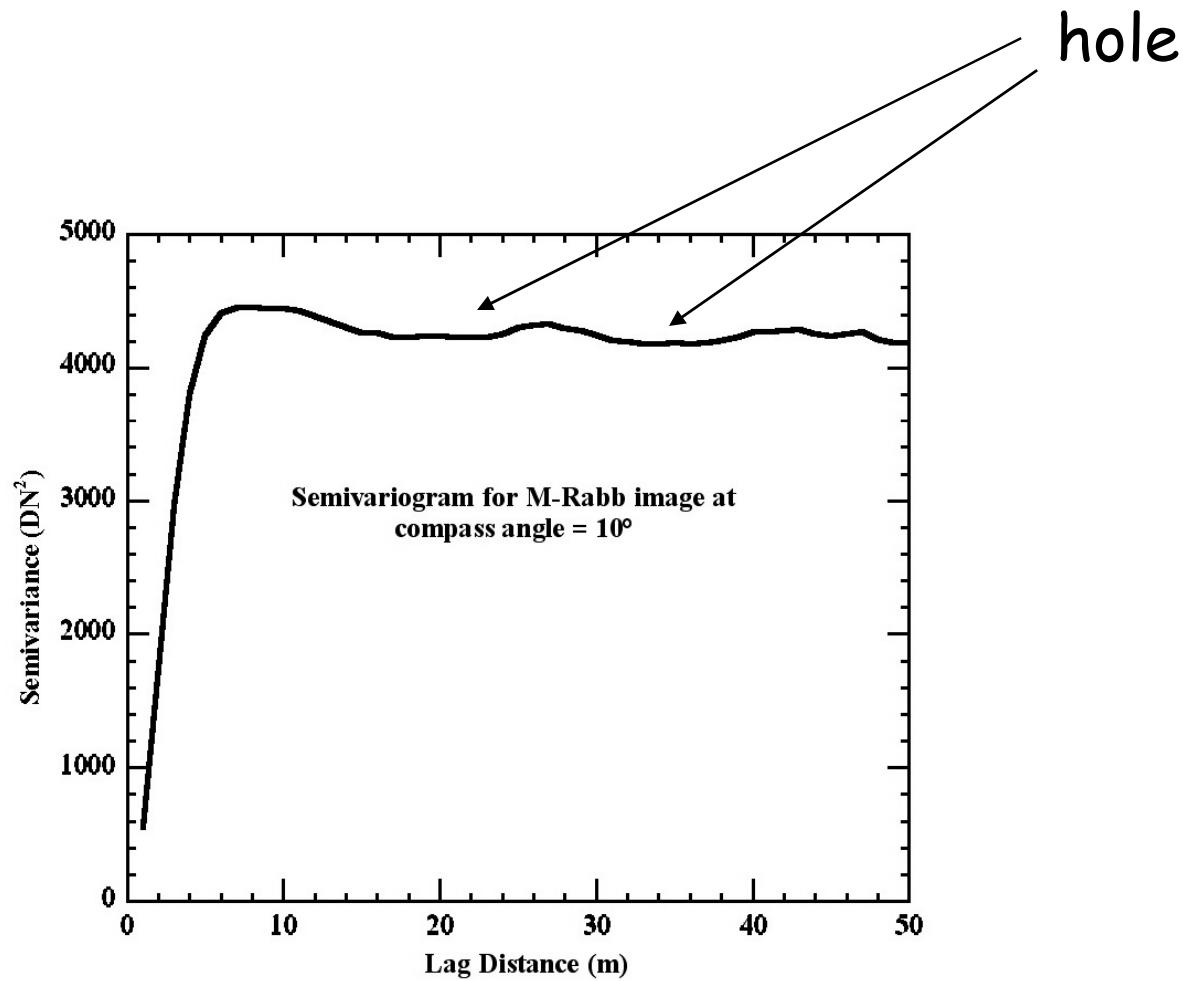
Grumenski (nugget) model



Poređenje variograma

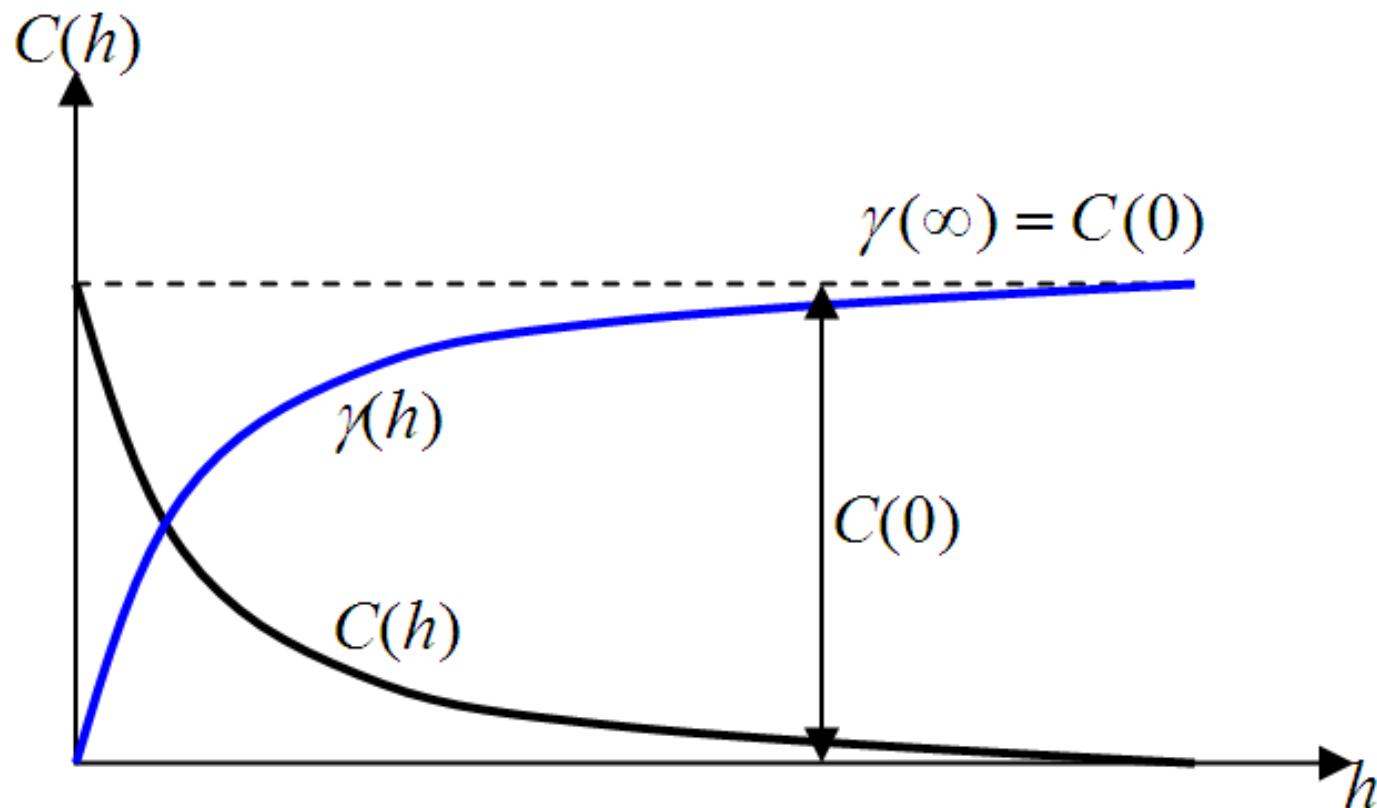


Efekat rupe (hole)



Odnos između kovarijanse i variograma u slučaju stacionarnih promenljivih

$$\gamma(h) = C(0) - C(h) = \sigma^2 - C(h)$$



Kriging interpolacija

Kriging Interpolacija

- Ovaj postupak uključuje i grešku merenja i nesigurnost prilikom procene vrednosti.
 - Kriging interpolacija u svakoj vrednosti koju procenjuje pridružuje i standardnu grešku ili varijansu.
-
- Bazira se određivanju optimalnih težinskih koeficijenata koji se pridružuju poznatim (uzorkovanim) vrednostima na osnovu poznavanja variograma. S obzirom da su variogrami u funkciji rastojanja, težine će zavistti od distribucije uzorkovanih tačaka.

Ordinary (obični) Kriging

Ordinary Kriging

- Ordinary kriging je najjednostavnija forma kriging-a.
- Kod Ordinary kriging-a, predpostavlja se da je regionalizovana promenljiva stacionarna (slaba stacionarnost)

- Uslovi koji zadovoljavaju hipotezu o prirođenosti (intrinsic hypothesis) su:

$$E[Z(x + h) - Z(x)] = m(h)$$

$$\text{Var}[Z(x + h) - Z(x)] = 2\gamma(h)$$

- U ovom slučaju uvodi se pretpostavka *oslabljene stacionarnosti drugog reda*.

$$\gamma(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} (z_i - z_{i+h})^2$$

Ordinary Kriging

- U našem slučaju $Z_e(p)$ procenjena vrednost Z u tački p , računa se kao ponderisana aritmetička sredina poznatih vrednosti:

$$z_e(p) = \sum w_i \cdot z(p_i)$$

- Greška ocene (*estimation error*) je razlika između procenjene i stvarne vrednosti u tački p $Z_a(p)$:

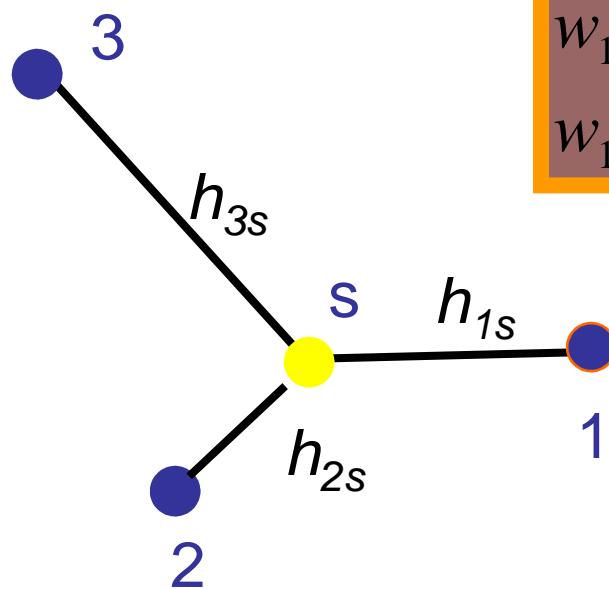
$$\varepsilon_p = z_e(p) - Z_a(p)$$

- Ukoliko ne postoji trend i suma težina je jednaka 1 (jedan), onda je ocena nepomerena (**unbiased**).
- Rasipanje ocena oko tačne vrednosti se naziva greška ili varijansa ocene (**estimation variance**).

$$\sigma_z^2 = \frac{\sum_{i=1}^n [z_e(s_i) - z_a(s_i)]_i^2}{n}$$

Ordinary Kriging

- Kriging nastoji da izabere optimalne težine kako bi dobili minimalnu grešku procene.
- Optimalne težine, koje daju nepomerenu ocenu i minimalnu grešku procene, dobijaju se rešavanjem sledećeg sistema jednačina.



$$w_1\gamma(h_{11}) + w_2\gamma(h_{12}) + w_3\gamma(h_{13}) = \gamma(h_{1s})$$

$$w_1\gamma(h_{21}) + w_2\gamma(h_{22}) + w_3\gamma(h_{23}) = \gamma(h_{2s})$$

$$w_1\gamma(h_{31}) + w_2\gamma(h_{32}) + w_3\gamma(h_{33}) = \gamma(h_{3s})$$

$$w_1 + w_2 + w_3 = 1$$

Ordinary Kriging

- Uvodi se pomoćna promenljiva poznata kao Lagranžov množitelj (**Lagrange multiplier**)

$$w_1\gamma(h_{11}) + w_2\gamma(h_{12}) + w_3\gamma(h_{13}) = \gamma(h_{1s})$$

$$w_1\gamma(h_{21}) + w_2\gamma(h_{22}) + w_3\gamma(h_{23}) = \gamma(h_{2s})$$

$$w_1\gamma(h_{31}) + w_2\gamma(h_{32}) + w_3\gamma(h_{33}) = \gamma(h_{3s})$$

$$w_1 + w_2 + w_3 = 1$$

$$\begin{bmatrix} \gamma(h_{11}) & \gamma(h_{12}) & \gamma(h_{13}) & 1 \\ \gamma(h_{21}) & \gamma(h_{22}) & \gamma(h_{23}) & 1 \\ \gamma(h_{31}) & \gamma(h_{32}) & \gamma(h_{33}) & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma(h_{1p}) \\ \gamma(h_{2p}) \\ \gamma(h_{3p}) \\ 1 \end{bmatrix}$$

Ordinary Kriging

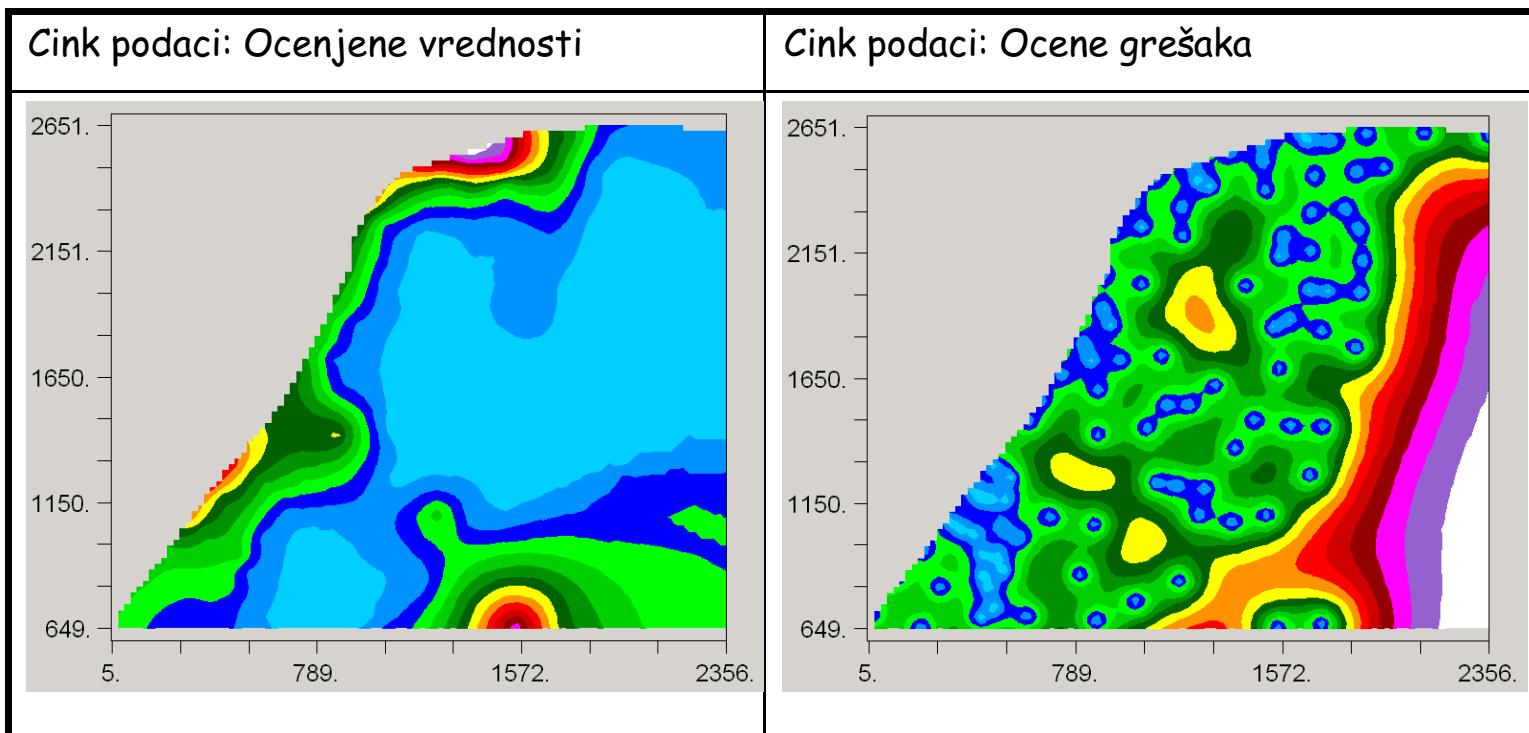
- Kada se odrede težine, ocenjena vrednost je:

$$z_e(p) = w_1 z_1 + w_2 z_2 + w_3 z_3$$

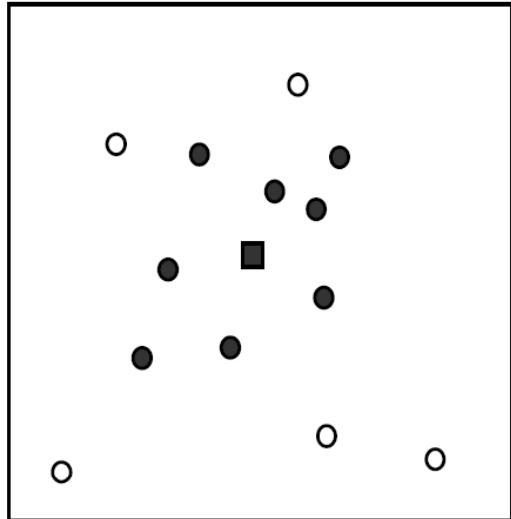
- Ocena greske se računa kao:

$$\sigma_z^2 = w_1 \gamma(h_{1p}) + w_2 \gamma(h_{2p}) + w_{13} \gamma(h_{3p}) + \lambda$$

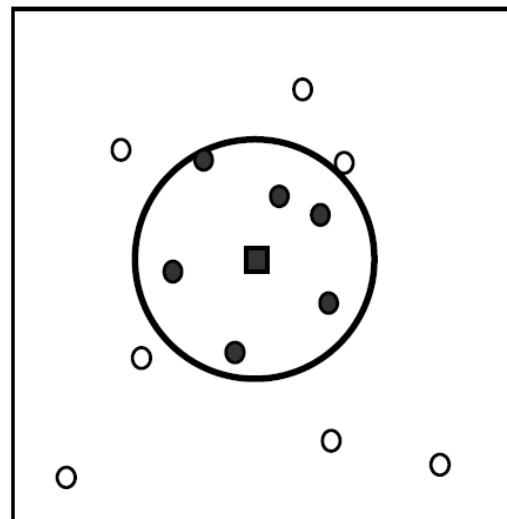
Ordinary Kriging



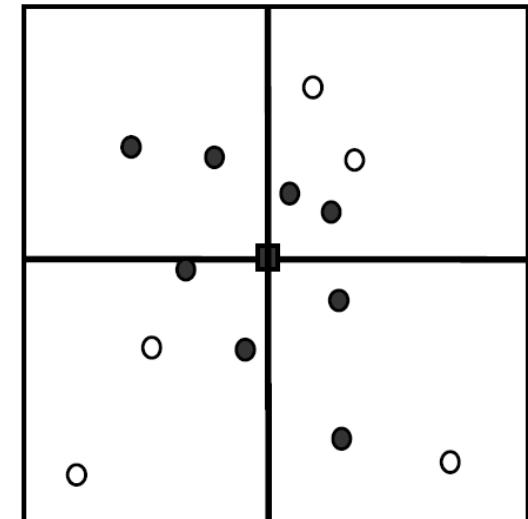
Izbor tačaka



(a)



(b)



(c)

- (a)biramo najbliže tačke,
- (b)koristimo zadati radijus
- (c)biramo podjednak broj tačaka u kvadrantu.

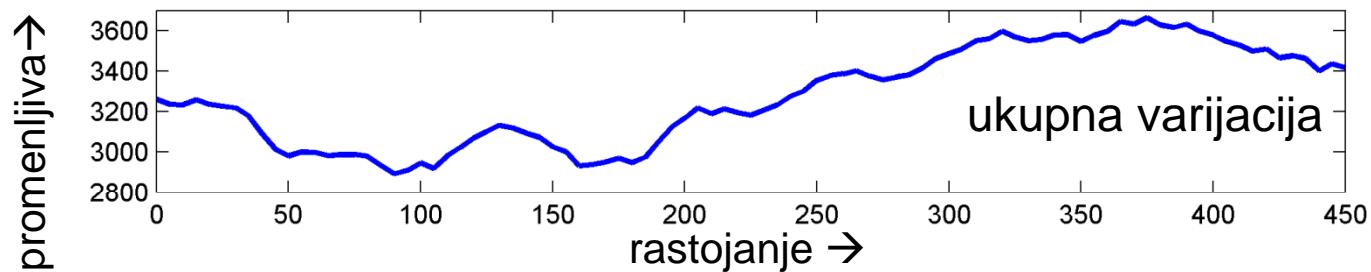
Kriging princip

Vrednosti Z na lokaciji s modeluju se kao suma:

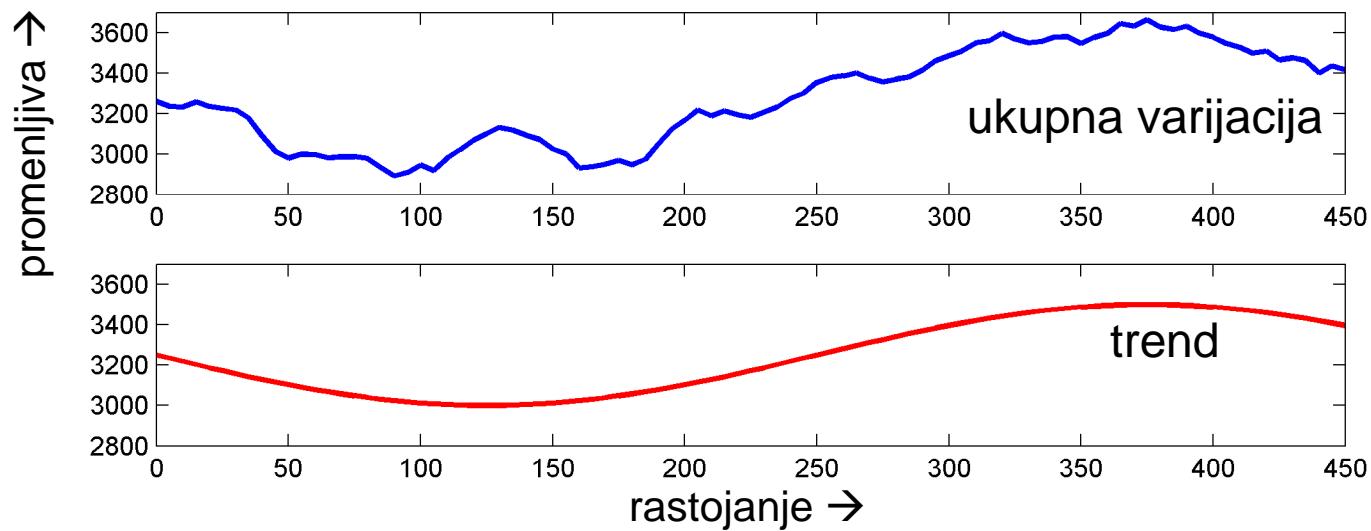
- trenda (m)
- prostorne autokorelaciјe (ε')
- Greške sa normalnom raspodelom (Gaussian error (ε'')))

$$Z(s) = m(s) + \varepsilon'(s) + \varepsilon''(s)$$

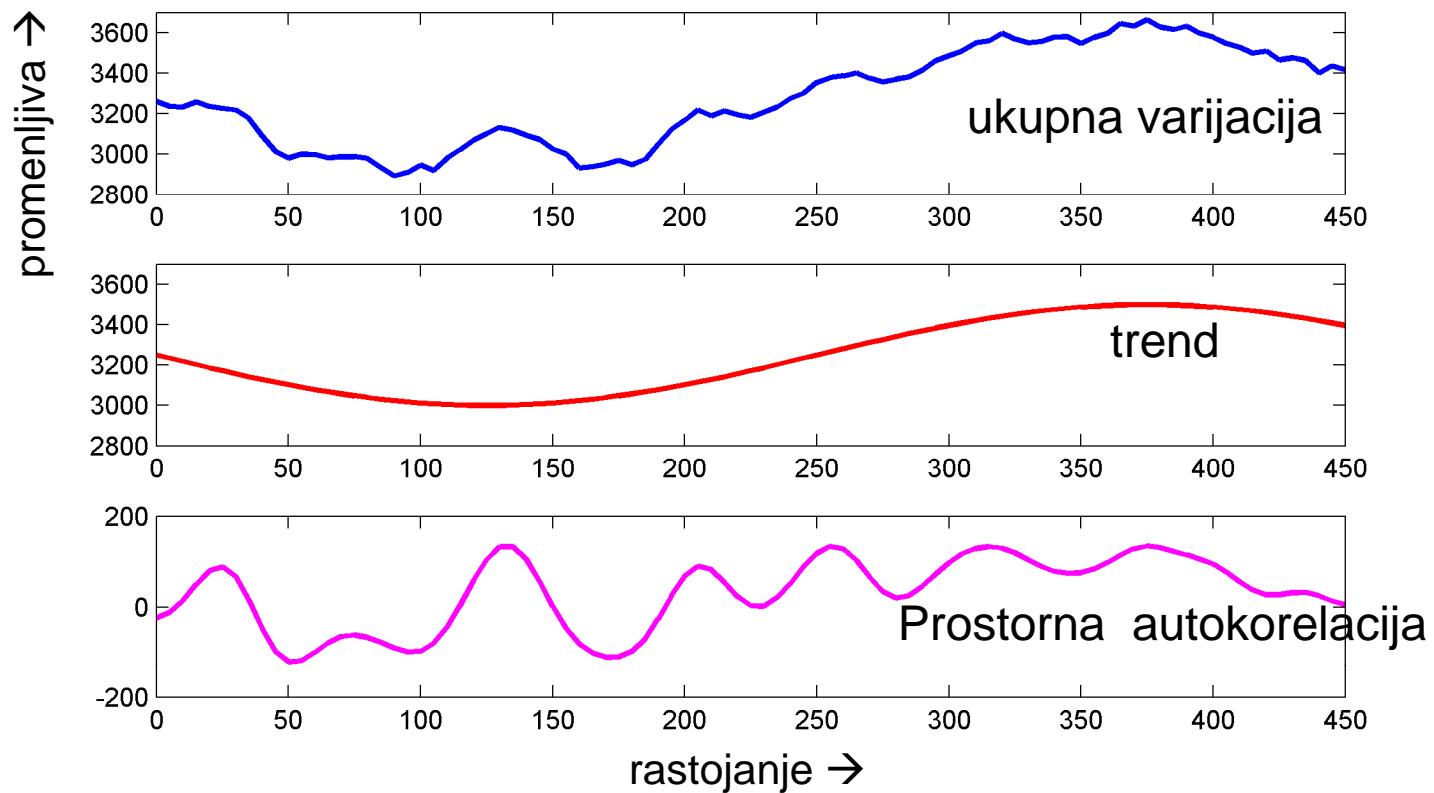
Kriging princip



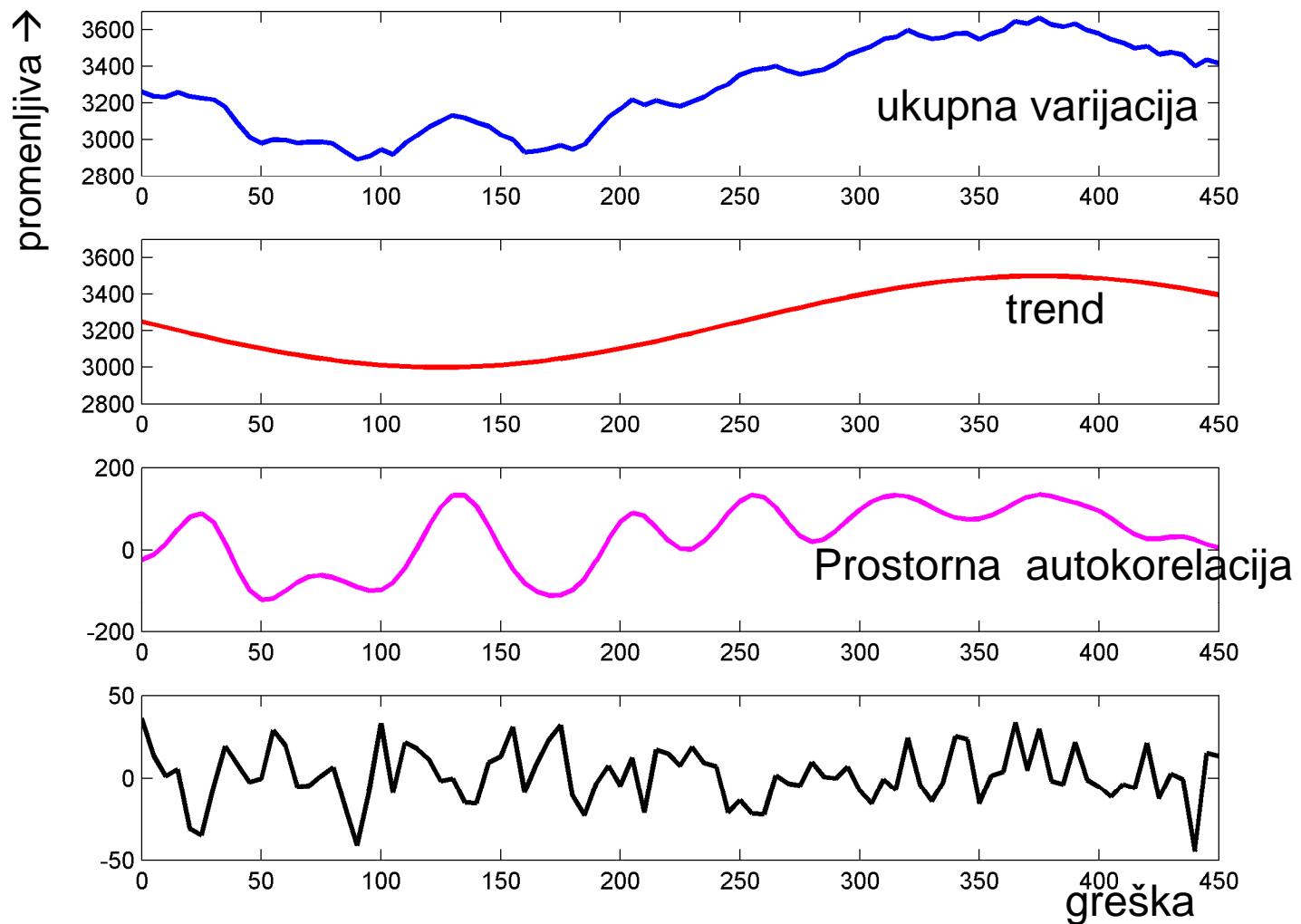
Kriging princip



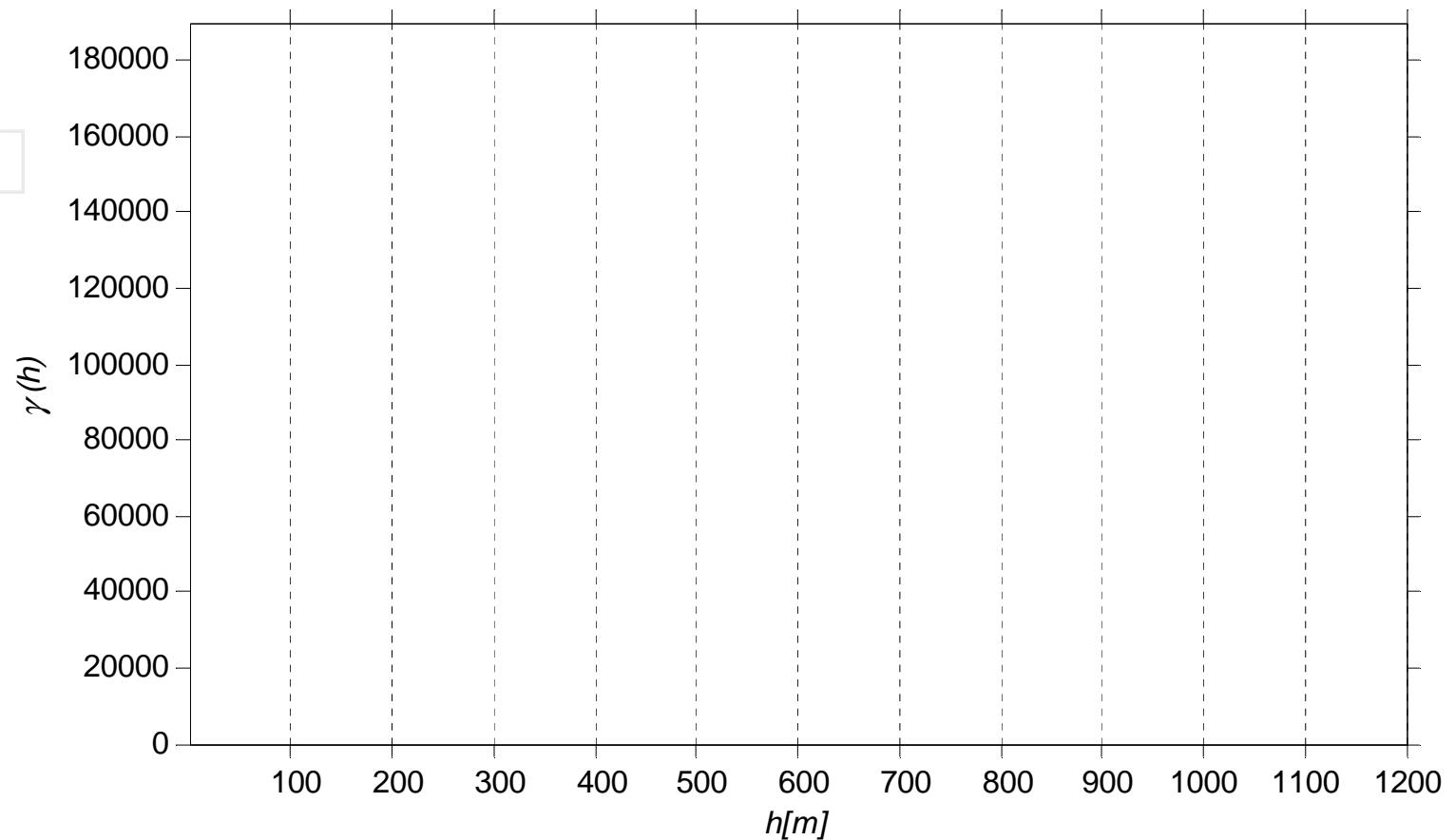
Kriging princip



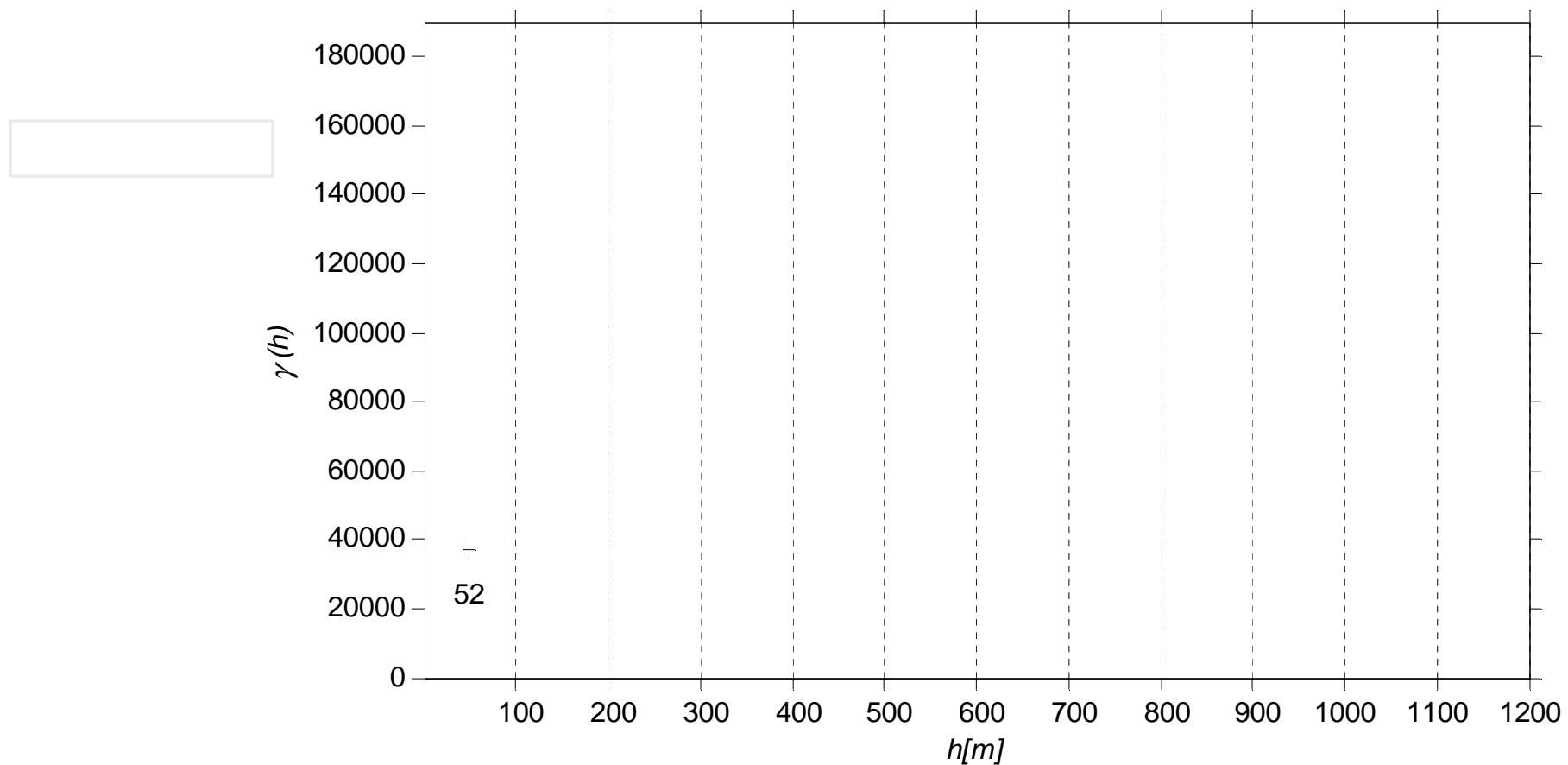
Kriging princip



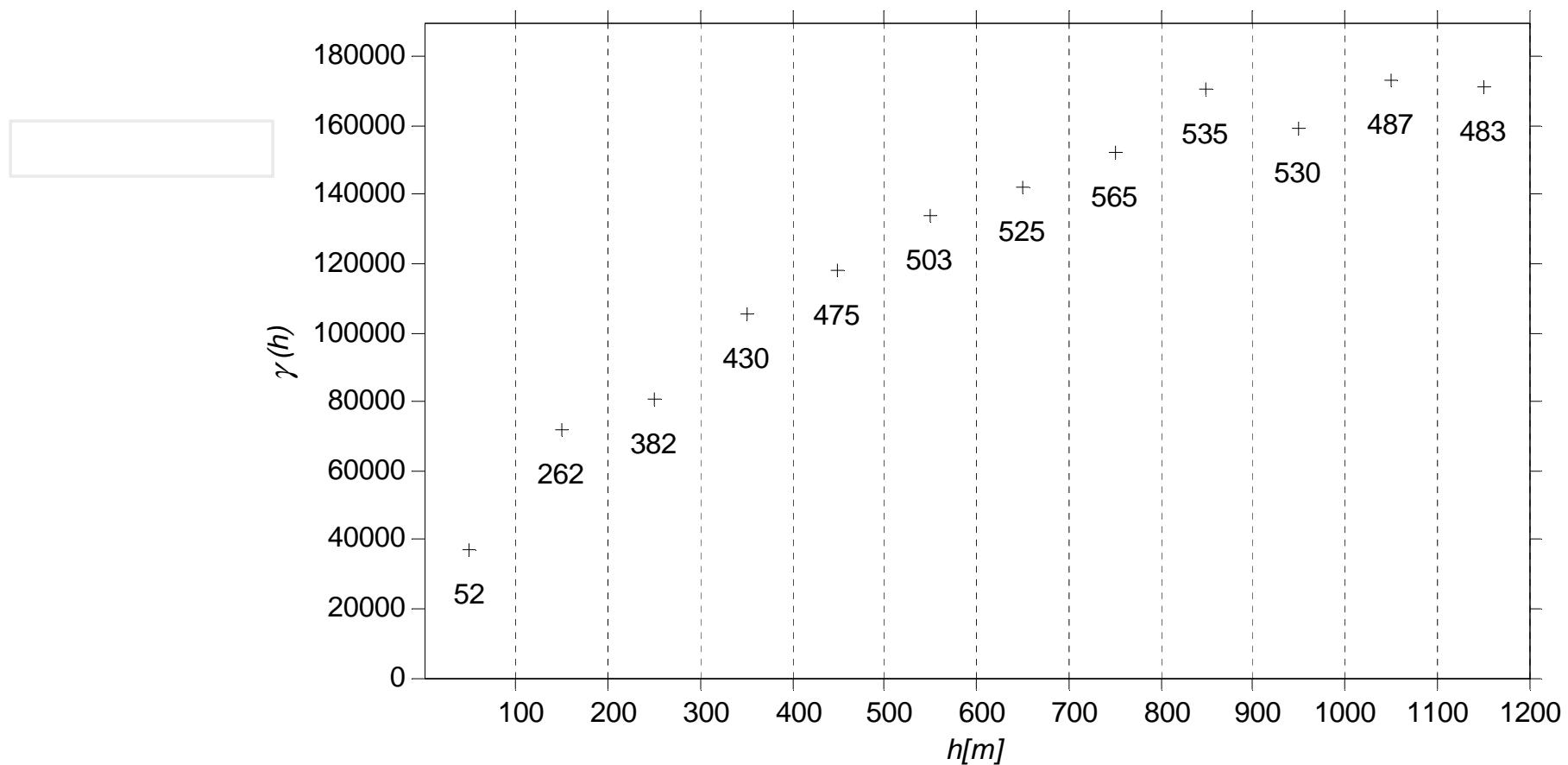
Eksperimentalni variogram



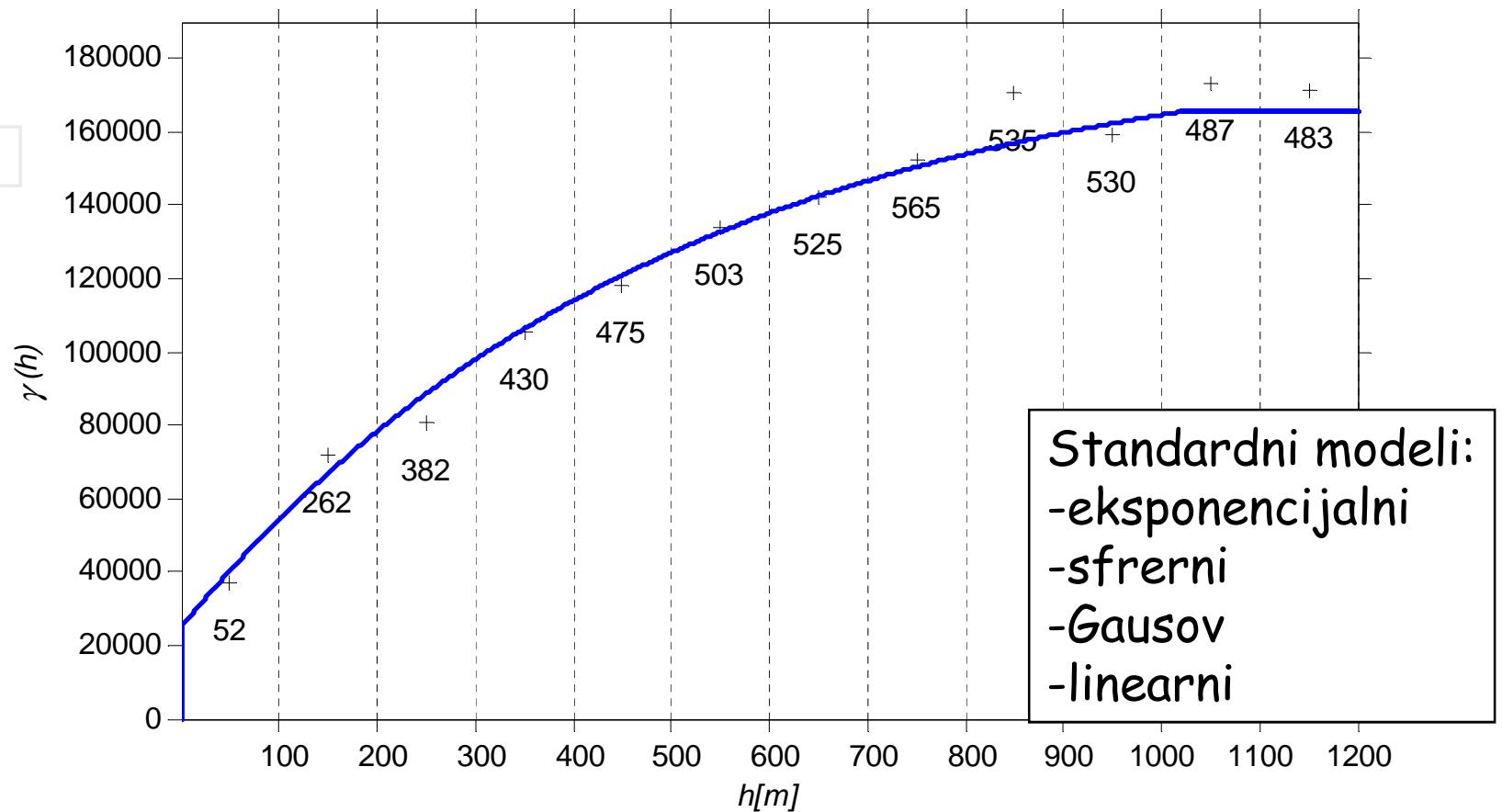
Eksperimentalni variogram



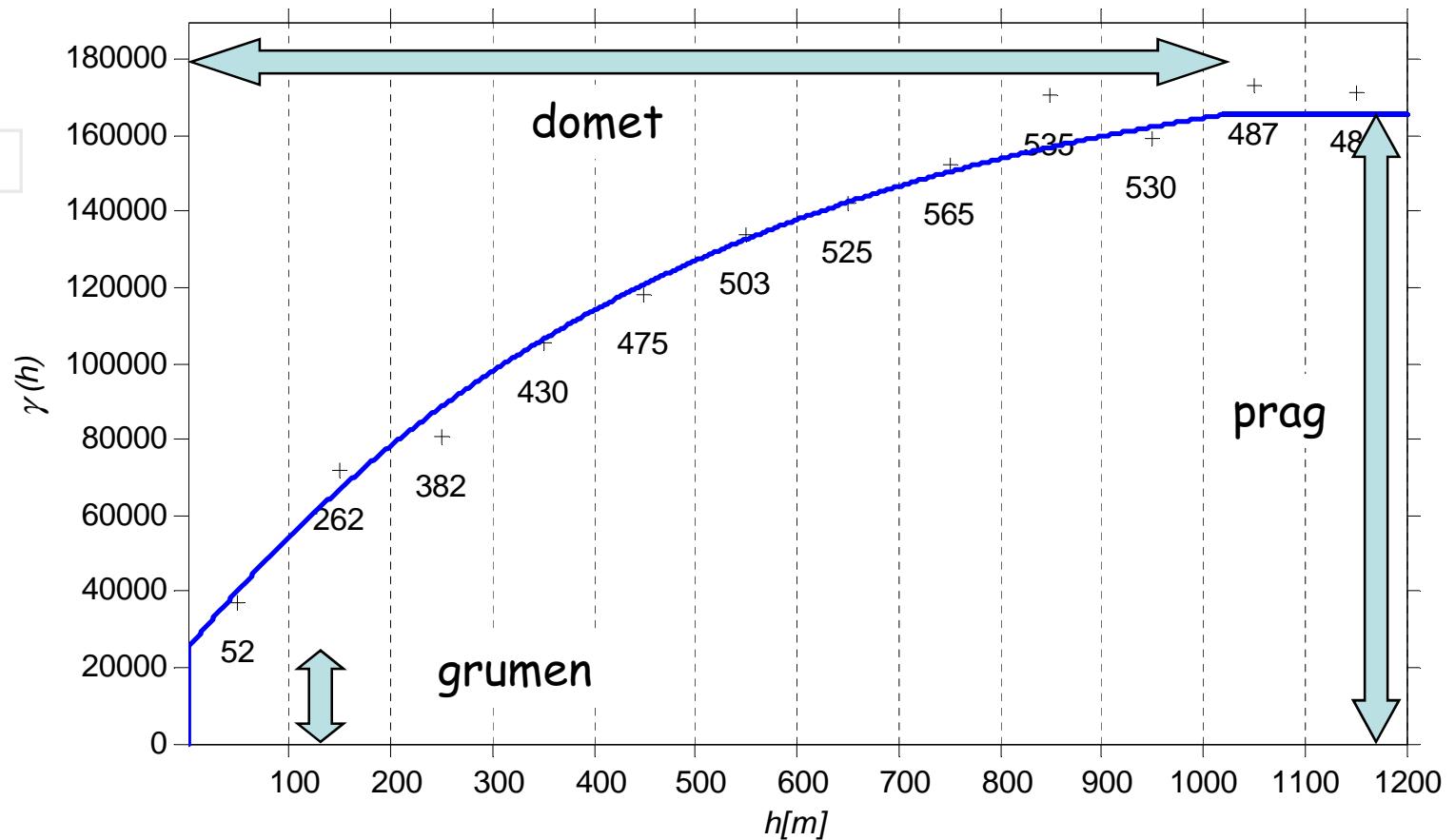
Eksperimentalni variogram



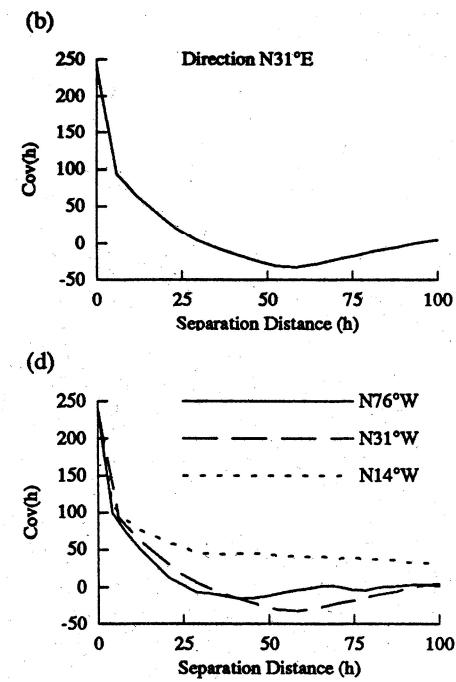
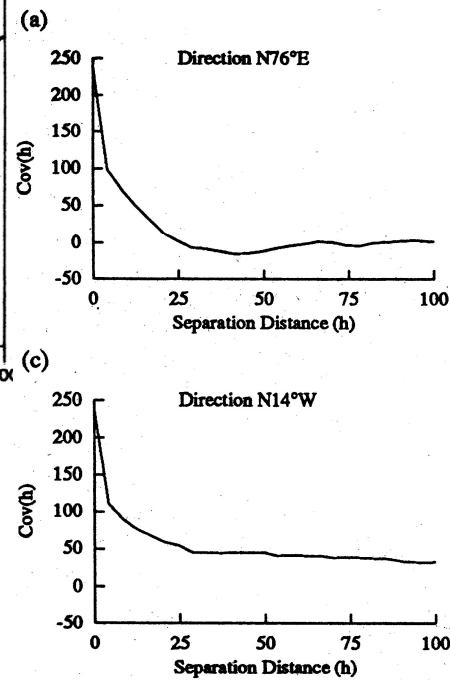
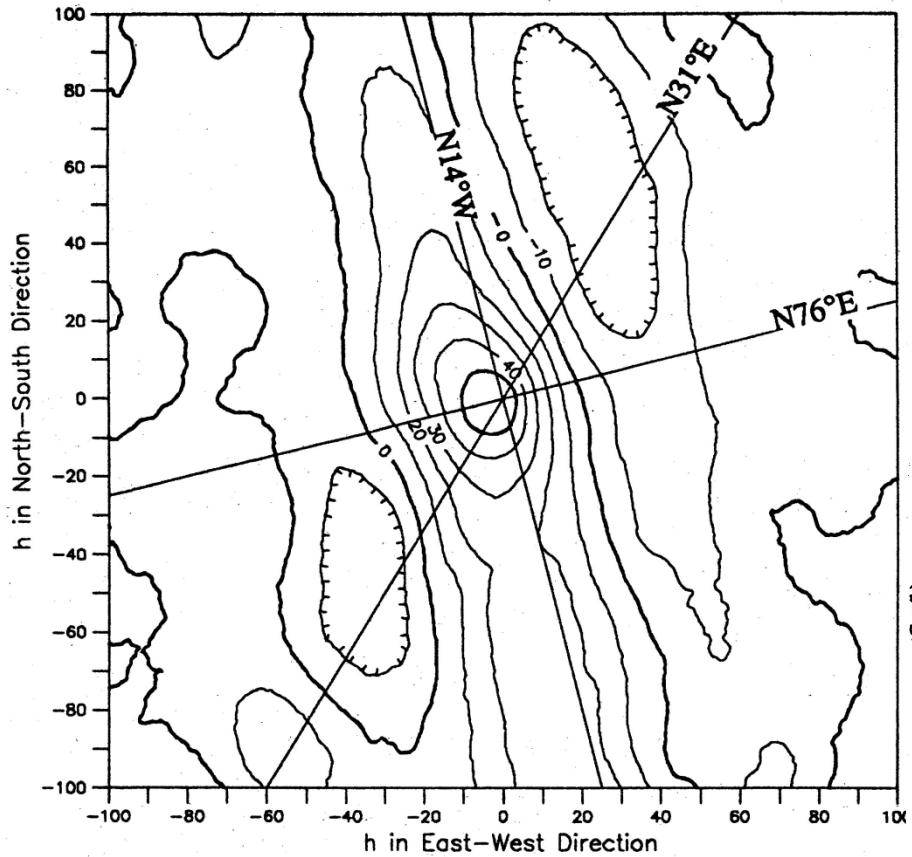
Eksperimentalni variogram



Eksperimentalni variogram

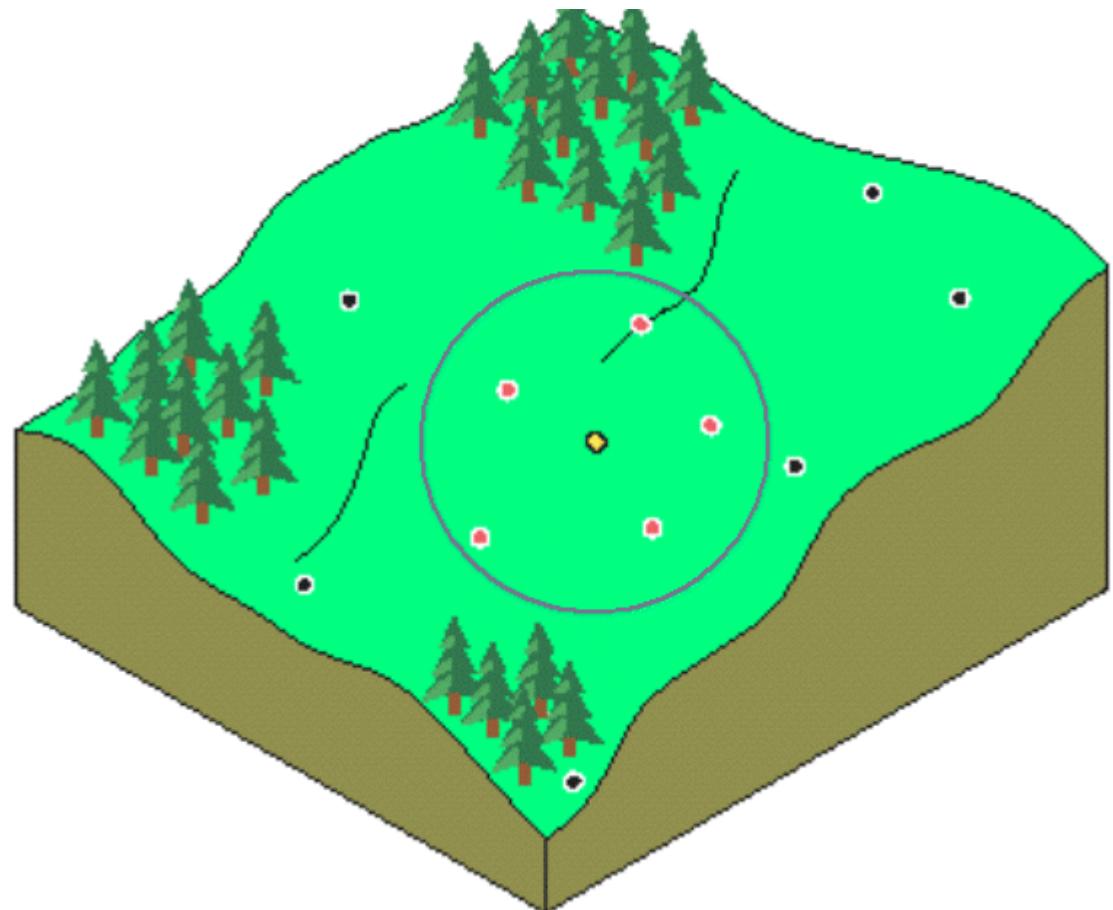


Anizotropija



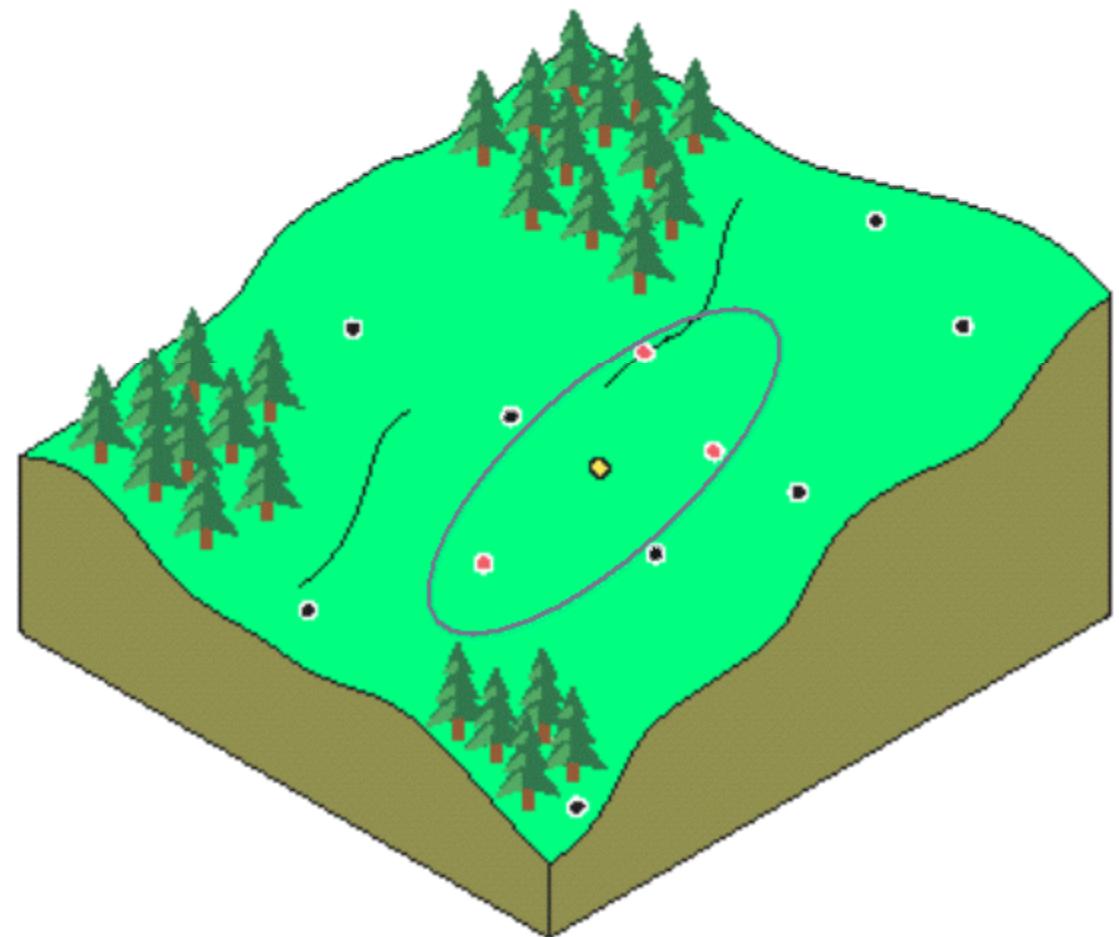
Slučaj izotropije

- Ukoliko ne postoji poseban pravac u kojem je uticaj okolnih uzoraka od većeg značaja tada se prozor za pretraživanje zadaje u obliku kružnice sa zadatim radiusom.



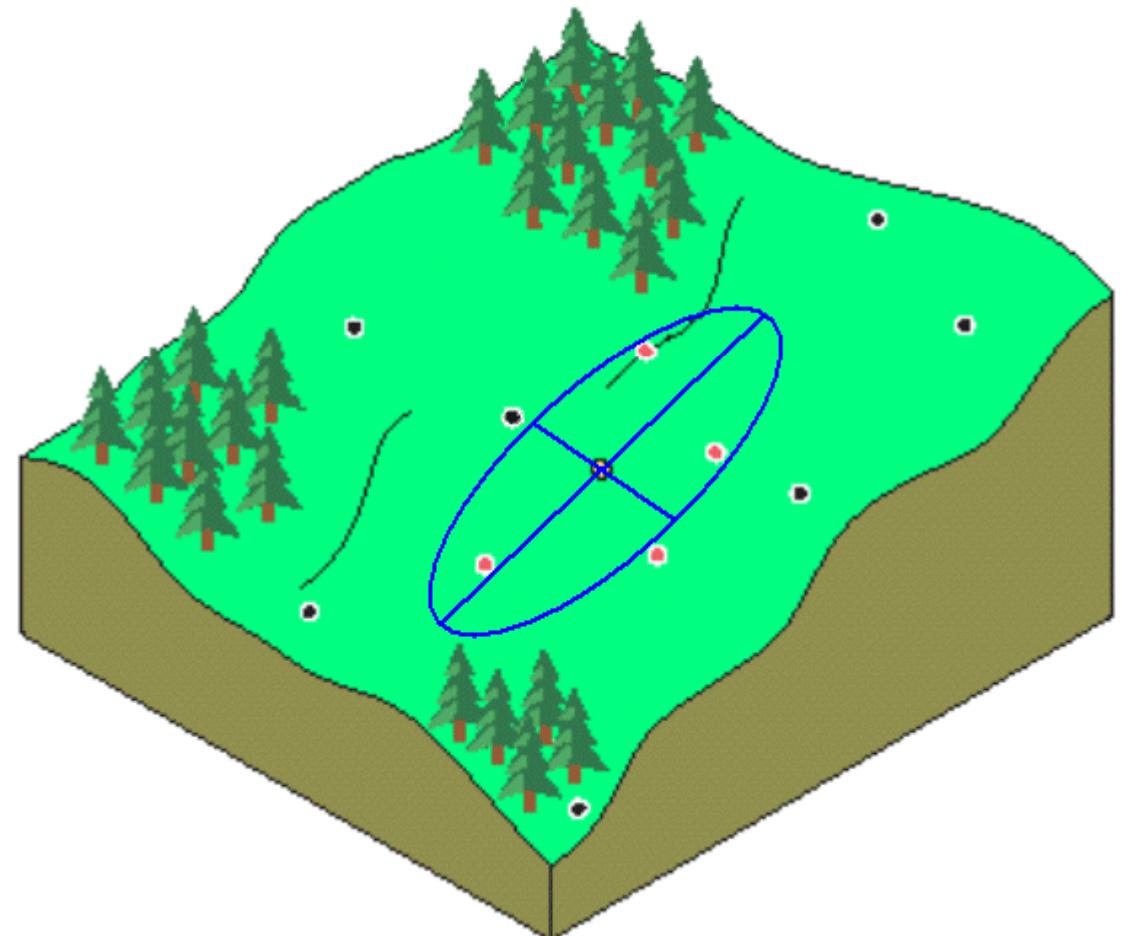
Definisanje smera uticaja

- U slučaju anizotropije, kada je uticaj podataka u određenom pravcu značajniji u odnosu na druge (npr. pravac površinskog oticanja voda), tada se umesto prozora u obliku kružnice može definisati i elipsa.



Izbor tačaka po sektoru

- U okviru prozora za pretraživanje definišu se sektori (kvadranti ili oktanti) u okviru kojih se zadaje minimalni i maksimalni broj tačaka čije vrednosti se koriste pri lokalnoj interpolaciji.



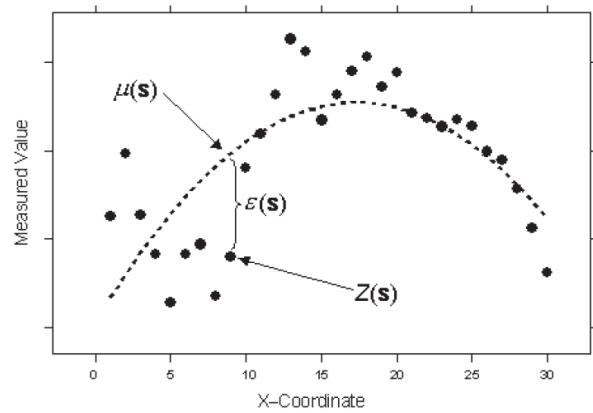
Kriging metode

- Universal Kriging - Kriging sa trendom
- Indicator Kriging - sa definisanim pragom vrednosti
- Stratified Kriging - regionalizovani pristup
- Co-Kriging - korišćenje pridruženih podataka

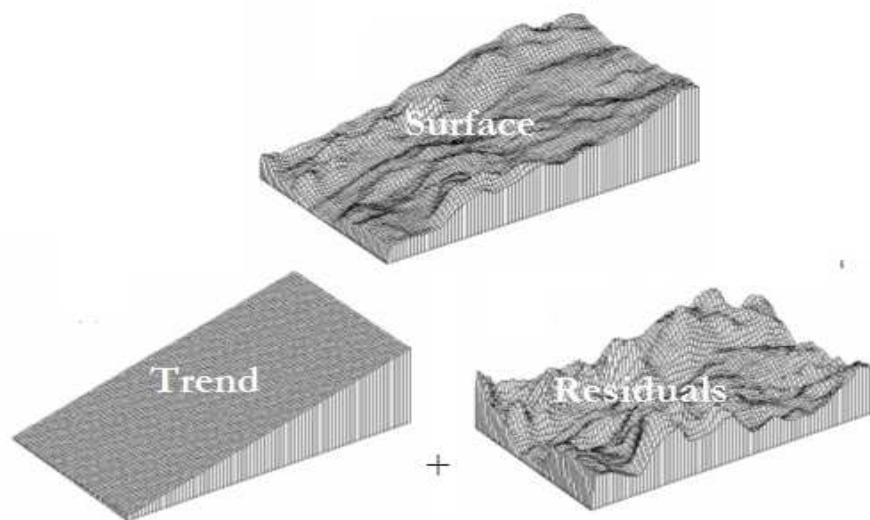
Kriging metode

- *Ordinary Kriging* najopštiji model i najčešće korišćeni. Predpostavlja da je srednja vrednost promenljive konstantna veličina (ne postoji trend među merenim podacima).
- *Simple Kriging* predpostavka poznavanja srednje vrednosti promenljive (konstantna veličina).
- *Universal Kriging* predpostavlja prisustvo trenda među merenim podacima, npr. preovlađujući vetrovi gde se trend može modelovati determinističkim polinomskim funkcijama. U ovom slučaju se trend oduzima od merenih uzoraka (dobijaju se reziduali) i tada se vrši ocena prostorne korelacijske.

Universal Kriging



$$Z(s) = \mu(s) + \varepsilon(s)$$



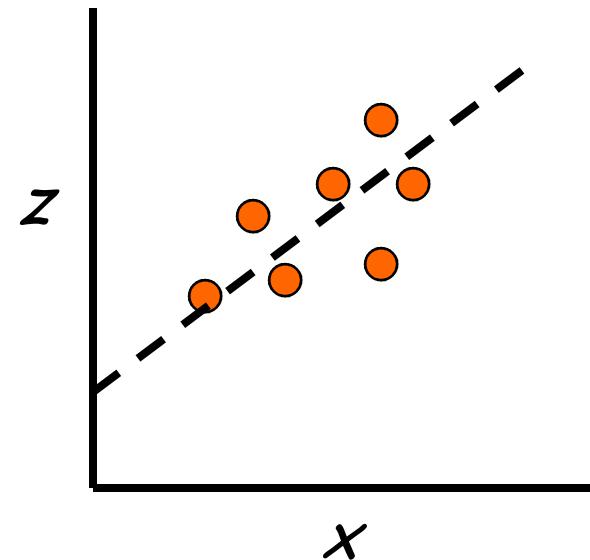
Trend se ocenjuje u funkciji prostornih koordinata, UK je dobro rešenje kada je evidentna prisutna nestacionarnost među merenim podacima, tj srednja vrednost podataka varira u prostoru.

Trend površi

Trend površi 1. reda

U jednodimenzionalnom slučaju: **Z** varira kao
linerana funkcija oko **X**

$$z = b_0 + b_1 x + e$$

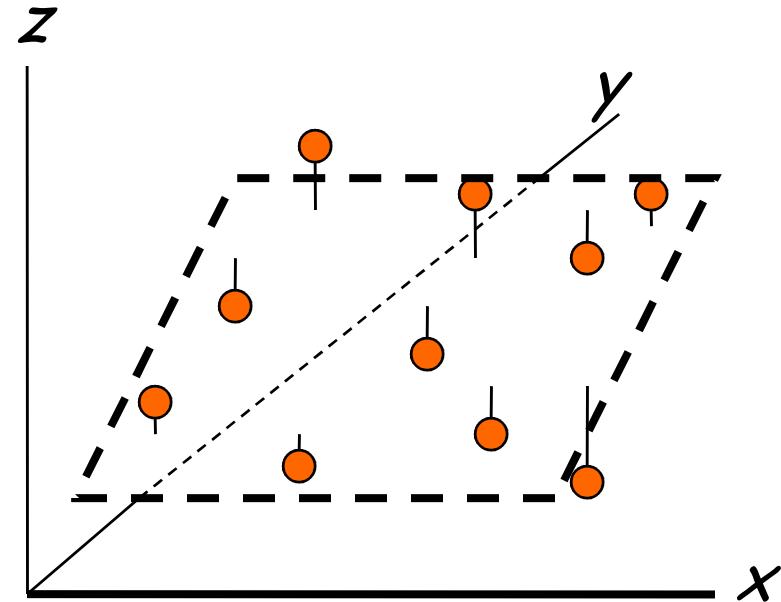


Trend površi

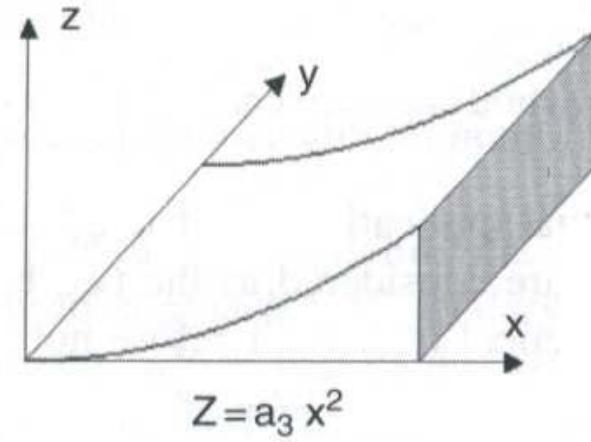
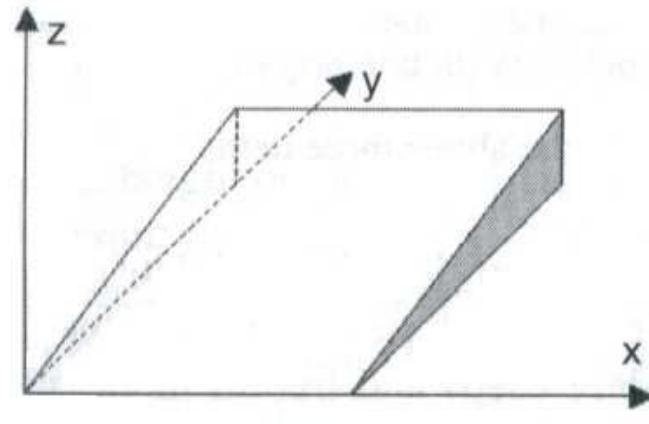
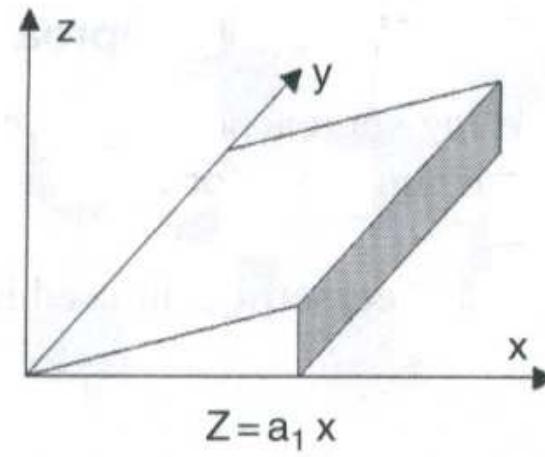
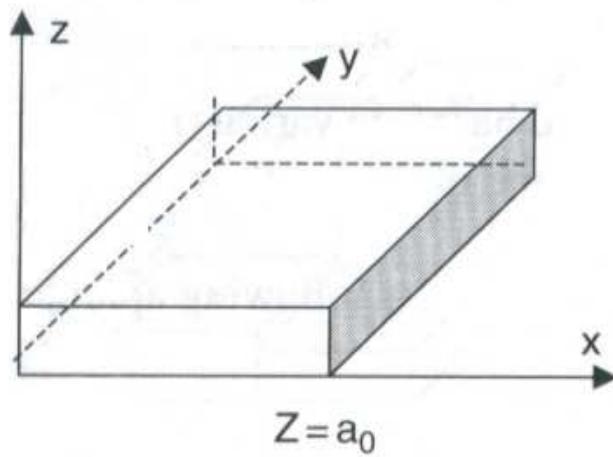
Trend površi 1. reda

U dvo dimenzionalnom slučaju: **z** varira kao linerana funkcija oko **x** i **y**

$$z = b_0 + b_1x + b_2y + e$$



Trend površi



Procena pomoću dodatnih podataka

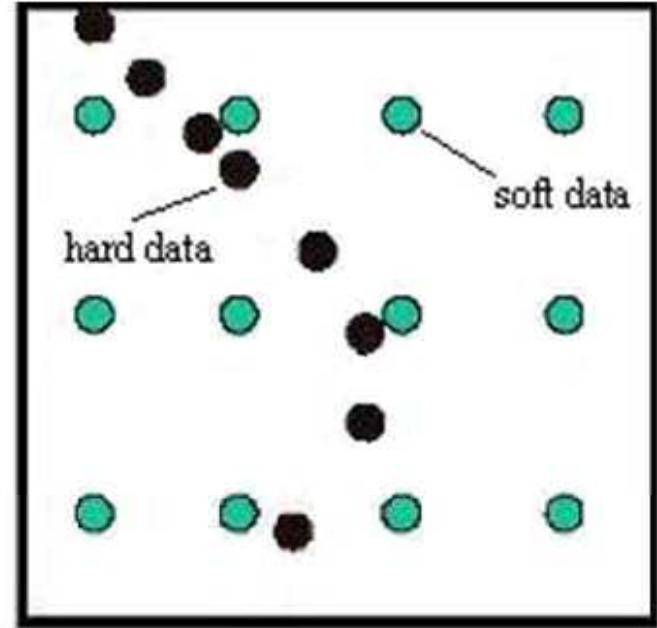
- Co-kriging
- Regresioni kriging ili kriging sa eksternim trendom (with external drift)

Cokriging

- Želimo da procenimo ciljanu promenljivu koja je obično ili teška za uzorkovanje i laboratorijsku obradu, pa imamo samo mali broj uzoraka.
- Međutim, imamo sekundarnu promenljivu (obično laku za uzorkovanje) koja je u korelisana sa ciljanom promenljivom.

Cokriging

$$\hat{u}_0 = \sum_{i=1}^n a_i \cdot u_i + \sum_{j=1}^m b_j \cdot v_j$$



u_1, \dots, u_n - Primarna promenljiva (*hard data*)

v_1, \dots, v_m - sekundarna promenljiva (*soft data*)

Cokriging

- Cokriging sistem jednačina:

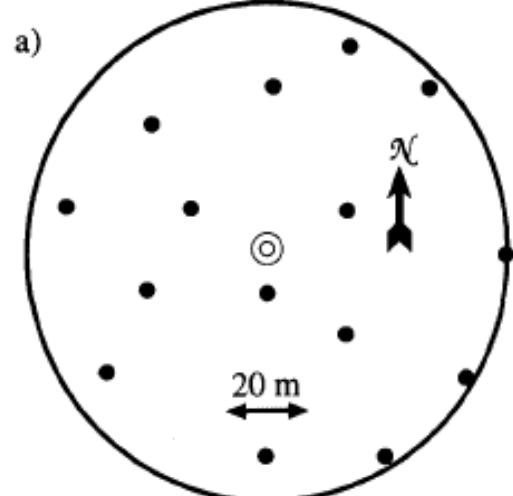
$$\sum_{i=1}^n a_i \text{Cov}\{U_i U_j\} + \sum_{i=1}^m b_i \text{Cov}\{V_i U_j\} + \mu_1 = \text{Cov}\{U_0 U_j\} \text{ za } j = 1, \dots, n$$

$$\sum_{i=1}^n a_i \text{Cov}\{U_i V_j\} + \sum_{i=1}^m b_i \text{Cov}\{V_i V_j\} + \mu_2 = \text{Cov}\{U_0 V_j\} \text{ za } j = 1, \dots, m$$

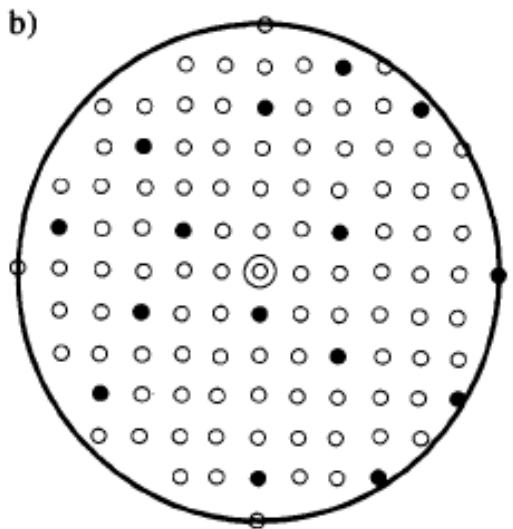
$$\sum_{i=1}^n a_i = 1$$

$$\sum_{i=1}^m b_i = 0$$

Cokriging

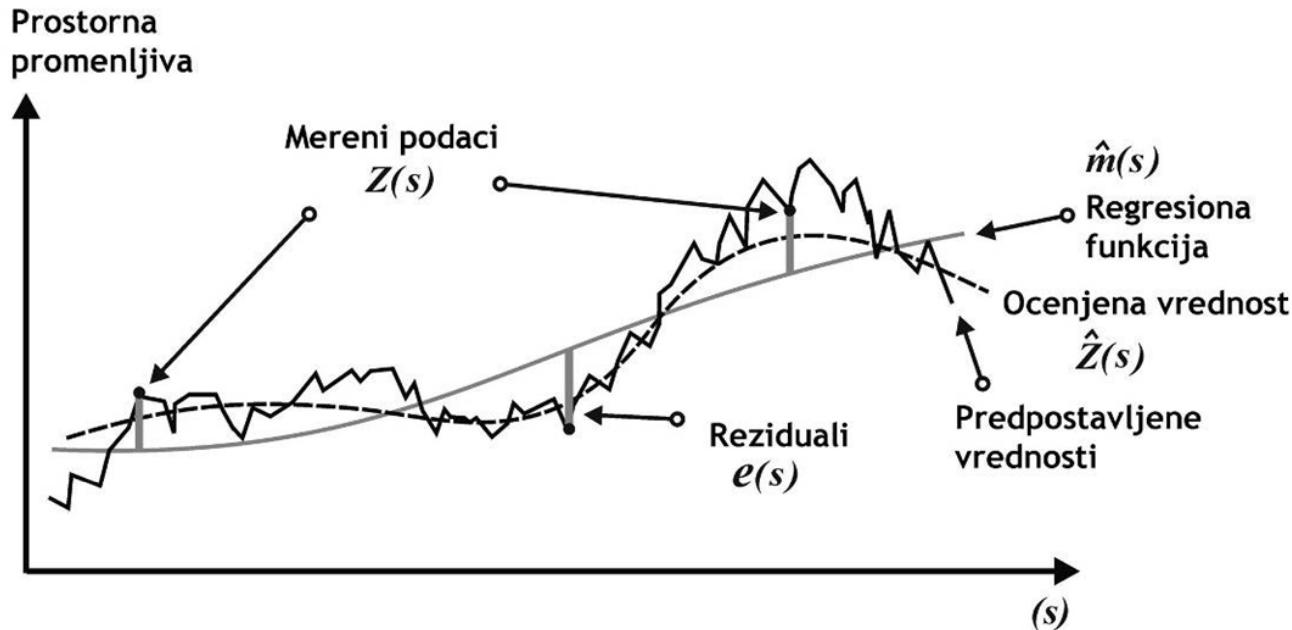


a) Isotopic



b) Heterotropic

Regresioni kriging



$$\hat{z}(s_0) = \hat{m}(s_0) + \hat{e}(s_0)$$

$$\hat{z}(s_0) = \sum_{k=0}^p \hat{\beta}_k \cdot q_k(s_0) + \sum_{i=1}^n w_i(s_0) \cdot e(s_i);$$

$$q_0(s_0) = 1$$

$\hat{m}(s_0)$ modelovani deterministički deo

$\hat{e}(s_0)$ -interpolovani reziduali

$\hat{\beta}_k$ -ocenjeni težinski koeficijenti trenda

$q_k(s_0)$ -je k-ta pomoćna promenljiva ili prediktor na lokaciji s_0

Pomoćne (eksterne) promenljive

- Digitalni modeli terena
- Primarni topografski parametri (nagib, ekspozicija, zakriviljenost ...).
- Sekundarni topografski parametri (topographic wetness index, stream-power indexes, radiation index, natural drainage networks...).
- Podaci iz satelitskih snimaka (NDVI, ili drugi vegetacioni indeksi...).
- Digitalizovane tematske karte (biljni pokrivač, namena zemljišta...).

Literatura

- Isaaks and Srivastava "An Introduction to Applied Geostatistics", Oxford University Press 1989
- Hengl, T. 2009. A Practical Guide to Geostatistical Mapping, 2nd Edt. University of Amsterdam, www.lulu.com, 291 p.