

KRIGING – geostatistické metody interpolace

Krigování je základní geostatistickou metodou určování lokálního odhadu. Metoda se často označuje akronymem **BLUE** (Best Linear Unbiased Estimator – tedy nejlepší lineární nezkreslený odhad). Toto označení má vystihnout výchozí podmínky krigování:

- odhadovaná hodnota je vypočtena jako lineární kombinace vstupních hodnot:

$$\hat{z}(x_0) = \sum_{i=1}^n \lambda_i \cdot z(x_i)$$

kde pro váhy platí

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

- nezkreslený (nestranný) odhad značí, že průměrná chyba tohoto odhadu je rovna nule

$$\sum (\hat{z}_i - z_i) = 0$$

- je minimalizován rozptyl odhadu

$$\sum (\hat{z}_i - z_i)^2 = \min.$$

Pokud prostorově závislá náhodná kolísání nejsou překryta nekorelovaným šumem, potom může být semivariogram využit k určení vah λ_i potřebných pro interpolaci. Procedura je podobná jako v případě metody vážených klouzavých průměrů s tím rozdílem, že právě váhy jsou odhadnuty geostatistickými metodami. Váhy λ_i jsou zvoleny tak, aby odhad $\hat{z}(x_0)$ byl nestranný a odhad rozptylu σ_e^2 byl menší, než jakákoliv jiná lineární kombinace pozorovaných hodnot (minimální).

Přitom pro minimální rozptyl hodnot $\hat{z}(x_0)$ platí výraz :

$$\hat{\sigma}_e^2 = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \phi$$

kde:

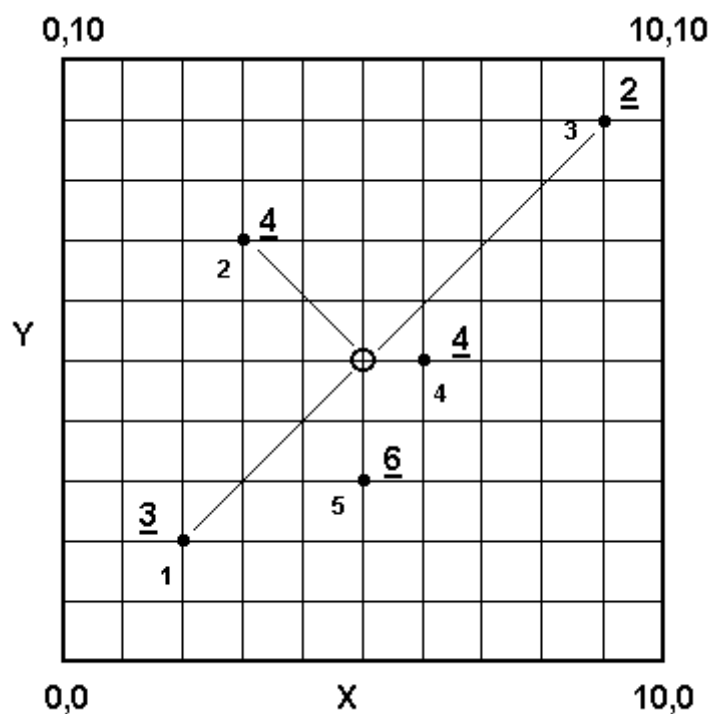
$$\sum_{i=1}^n \lambda_i \gamma(x_i, x_j) + \phi = \gamma(x_j, x_0) \text{ pro všechna } j.$$

Hodnota $\gamma(x_i, x_j)$ je semivariance proměnné z mezi body x_i a x_j . Hodnota $\gamma(x_i, x_0)$ je semivariance proměnné z mezi bodem x_i a bodem x_0 , pro který je hodnota proměnné z zjišťována. Obě hodnoty lze získat z vhodného teoretického modelu semivariogramu. Hodnota ϕ je tzv. Lagrangeův multiplikátor, který zajišťuje požadavek minimalizace odchylek a zároveň podmínku, že suma vah je rovna jedné.

Uvedená metoda se označuje jako **základní (ordinary) kriging** a je možné ji použít pro interpolaci v pravidelné mřížce hodnot, ke konstrukci map (např. izolinií).

PŘÍKLAD: Výpočet neznámé hodnoty v bodě metodou základního krigingu.

Na základě změřených hodnot veličiny Z v pěti bodech ($i=1, \dots, 5$) (viz. obrázek) máme za úkol odhadnout hodnotu Z bodě ($i=0$) o souřadnicích ($x=5, y=5$) metodou krigingu.



Obr. 1 Vstupní data pro lokální odhad metodou základního krigování (podtržená čísla značí hodnotu atributu v bodě)

Na základě předem provedené strukturální analýzy použijeme **sférický semivariogram**.

$$\gamma(h) = c_0 + c_1 * \left[\frac{3h}{2a} - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] \dots \dots \dots \text{pro } h \leq a$$

$$\gamma(h) = c_0 + c_1 \dots \dots \dots \text{pro } h > a$$

Parametry použitého teoretického semivariogramu jsou:

$$c_0 = 2,5$$

$$c_1 = 7,5$$

$$a = 10,0 \text{ (dosah)}$$

Data v pěti měřených bodech mají následující souřadnice

<u>i</u>	<u>x</u>	<u>y</u>	<u>z</u>
<u>1</u>	2	2	3
<u>2</u>	3	7	4
<u>3</u>	9	9	2
<u>4</u>	6	5	4
<u>5</u>	5	3	6

Pokud budeme dále značit:

A – matice semivariancí mezi všemi dvojicemi bodů

b – vektor semivariancí mezi všemi body a bodem predikovaným

λ – vektor vah jednotlivých bodů

Φ – tzv. Lagrangeův člen

potom základní vztah pro odhad metodou krigování lze psát jako:

$$A \cdot \lambda = b$$

Pro vlastní řešení je nutné vypočítat váhy λ , které musí splňovat podmínku $\sum \lambda = 1$

Uvedený základní vztah lze vyjádřit jako soustavu rovnic:

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdot & \cdot & \cdot & \gamma_{1n} & 1 \\ \gamma_{21} & \gamma_{22} & \cdot & \cdot & \cdot & \gamma_{2n} & 1 \\ \cdot & \cdot & & & & & \cdot \\ \cdot & \cdot & & & & & \cdot \\ \cdot & \cdot & & & & & \cdot \\ \gamma_{n1} & \gamma_{n2} & \cdot & \cdot & \cdot & \gamma_{nn} & 1 \\ 1 & 1 & \cdot & \cdot & \cdot & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_n \\ \phi \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \cdot \\ \cdot \\ \cdot \\ \gamma_{n0} \\ 1 \end{bmatrix}$$

V tomto zápisu poslední řádek a poslední sloupec v první matici a hodnota Lagrangeova členu ϕ jsou použity pro zajištění podmínky sumy vah $\sum \lambda = 1$. Hodnota Lagrangeova multiplikátoru ϕ také slouží pro výpočet rozptylu odhadnuté hodnoty. Uvedená soustava rovnic nám poskytne hodnoty všech vah λ a hodnotu ϕ . V maticovém zápisu lze tedy psát:

$$A^{-1} \cdot b = \begin{bmatrix} \lambda \\ \phi \end{bmatrix}$$

Aby bylo možné vyčíslit hodnoty semivariací, je v prvním kroku zapotřebí vytvořit matici vzdáleností mezi datovými body:

i	1	2	3	4	5
1	0,000	5,099	9,899	5,000	3,162
2	5,099	0,000	6,325	3,606	4,472
3	9,899	6,325	0,000	5,000	7,211
4	5,000	3,606	5,000	0,000	2,236
5	3,162	4,472	7,211	2,236	0,000

Vektor vzdáleností mezi měřenými body a bodem predikovaným:

i	0
1	4,234
2	2,828
3	5,657
4	1,000
5	2,000

Těchto vzdáleností využijeme k výpočtu semivariací pro sférický model semivariogramu s výše uvedenými parametry c_0 , c_1 , a – tedy k sestavení matice A a vektoru b :

Matice A:

i	1	2	3	4	5	6
1	2,500	7,739	9,999	7,656	5,939	1,000
2	7,739	2,500	8,667	6,381	7,196	1,000
3	9,999	8,667	2,500	7,656	9,206	1,000
4	7,656	6,381	7,656	2,500	4,936	1,000
5	5,939	7,196	9,206	4,936	2,500	1,000
6	1,000	1,000	1,000	1,000	1,000	1,000

Ve výše uvedené matici má řádek navíc (i=6) zajistit podmínku, že váhy budou mít sumu rovnou jedné.

Vektor b :

i	0
1	7,151
2	5,597
3	8,815
4	3,621
5	4,720
6	1,000

Inverzní matce A^{-1} :

i	1	2	3	4	5	6
1	-,172	,050	,022	-,026	,126	,273
2	,050	-,167	,032	,077	,007	,207
3	,022	,032	-,111	,066	-,010	,357
4	-,026	,077	,066	-,307	,190	,030
5	,126	,007	-,010	,190	-,313	,134
6	,273	,207	,357	,003	,134	-6,873

Řešením výše uvedené soustavy rovnic lze pro jednotlivé vzdálenosti získat hodnoty vah λ :

i	λ	
1	0,0175	
2	0,2281	
3	-0,0891	vypočtené hodnoty vah
4	0,6437	
5	0,1998	
6	0,1182	vypočtená hodnota ϕ

Pro váhy $i=1,\dots,5$ platí, že jejich suma se rovná jedné, v posledním řádku je hodnota Lagrangeova členu ϕ .

Vzdálenosti měřených bodů od bodu predikovaného, již přísluší výše určené váhy:

i	0
1	4,234
2	2,828
3	5,657
4	1,000
5	2,000

Potom odhad hodnoty Z v bodě ($i=0$) o souřadnicích ($x=5, y=5$):

$$Z(x_{i=0}) = 0,0175 \cdot 3 + 0,2281 \cdot 4 - 0,0891 \cdot 2 + 0,6437 \cdot 4 + 0,1998 \cdot 6 =$$

$$Z(x_{i=0}) = \mathbf{4,560}$$

Rozptyl odhadu:

$$\sigma_e^2 = [0,0175 \cdot 7,151 + 0,2281 \cdot 5,597 - 0,0891 \cdot 8,815 + 0,6437 \cdot 3,621 + 0,1998 \cdot 4,720] + \phi =$$

$$\sigma_e^2 = 3,890 + 0,1182 =$$

$$\sigma_e^2 = \mathbf{4,008}$$

Typy krigování

Na rozdíl od deterministických metod interpolace nabízí metody krigingu vedle odhadů vlastní interpolované hodnoty také odhady pravděpodobnosti výskytu těchto hodnot a dále odhady chyb predikce. Pro charakterizování jednotlivých typů krigingu předpokládáme jednoduchý model:

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

kde $Z(x_i)$ je proměnná v bodě x_i , která se skládá z deterministické hodnoty trendu $\mu(x_i)$ a autokorelované náhodné proměnné $\varepsilon(x_i)$. Protože ve většině případů hodnotu trendu často pouze odhadujeme a neznáme ji přesně, je určitou chybou zatížena též náhodná složka. Pro tuto platí, že její průměrná chyba je rovna nule a autokorelace hodnot $\varepsilon(x_i)$ $\varepsilon(x_i+h)$ nezávisí na aktuální pozici, ale pouze na hodnotě vzdálenosti h .

Hodnota trendu může nabývat konstantní hodnoty ($\mu(x_i) = \mu$). V závislosti na tom, zda hodnota μ představuje konstantu či zda data obsahují trendovou složku a v závislosti na tom, za hodnota μ představuje zámou hodnotu či zda ji odhadujeme definujeme různé metody (modely) krigování. Mezi základní metody krigování, kterými se provádí odhad na základě přímo naměřených hodnot patří především:

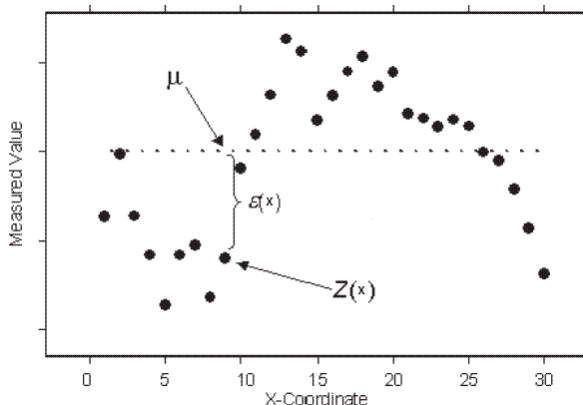
- základní (ordinary) krigování s bodovým odhadem
- základní (ordinary) krigování s blokovým odhadem
- jednoduché (simple) krigování
- univerzální krigování
- pravděpodobnostní krigování
- co-kriging
- lognormální krigování.

Základní krigování (ordinary kriging)

Obecný model základního krigování:

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

kde μ je neznámá hodnota trendu.



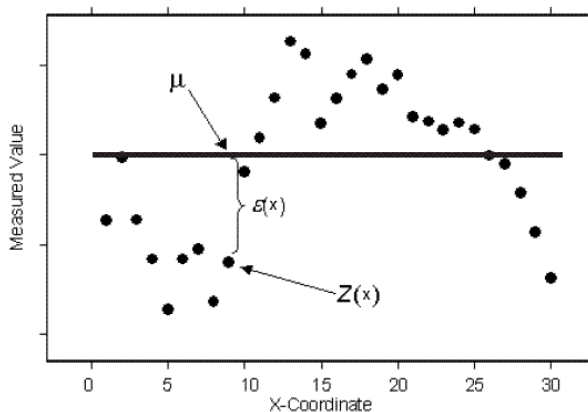
Obr. 2 Princip základního krigování

Jednoduché krigování (Simple kriging)

Nejjednodušší variantou krigování je tzv. **jednoduché krigování** (simple kriging). K výpočtu je potřebná znalost průměrné hodnoty veličiny v poli (μ) a dále předpoklad stacionarity. Obecný model jednoduchého krigování má opět tvar:

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

kde μ je známá konstanta.



Obr. 3 Princip základního krigování

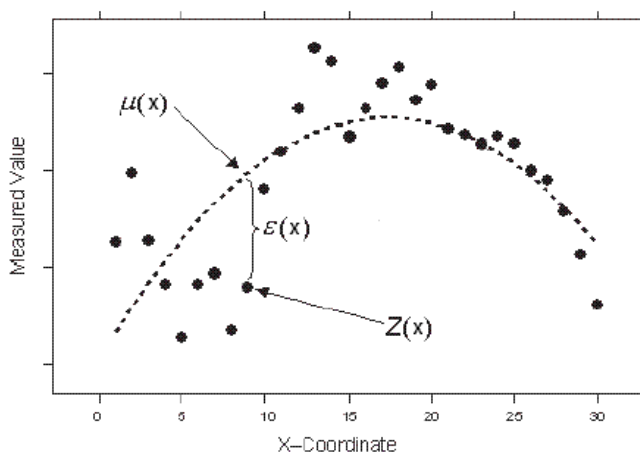
V tomto modelu, protože známe hodnotu μ , potom v bodech měření známe přesně také $\varepsilon(x)$. V případě základního krigování jsou obě hodnoty odhadovány. Základní korigování tedy nabízí přesnější odhad autokorelace, předpoklad znalosti μ je však často nereálný. Často se však používá některá z běžných trendových funkcí, kterou se nejprve vyjádří rezidua. Trend reziduí se potom považuje za nulový a aplikuje se základní krigování.

Univerzální krigování (Universal kriging)

Obečný model univerzálního krigování má opět tvar:

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

kde $\mu(x)$ je deterministická funkce - např. polynom druhého stupně jako na přiloženém obrázku. Pokud odečteme hodnotu polynomu od originálních dat, dostaneme chybovou složku $\varepsilon(x)$, která má charakter náhodné proměnné s průměrem rovným nule. Autokorelace je modelována právě z této náhodné proměnné. Jak je patrné z obrázku, univerzální kriging je v tomto případě analogií regresní závislosti. Na rozdíl od regrese, kde složku $\varepsilon(x)$ považujeme za nekorelovanou, v případě krigování ji modelujeme jako složku autokorelovanou. Stejně jako v případě základního krigingu, vhodná dekompozice na obě výše uvedené složky ze samotných dat nelze provést (*proto je třeba model ???*). Univerzální kriging používá k popisu autokorelované náhodné složky semivariogramu nebo kovariance.



Obr. 4 Princip univerzálního krigování

Indikátorové krigování

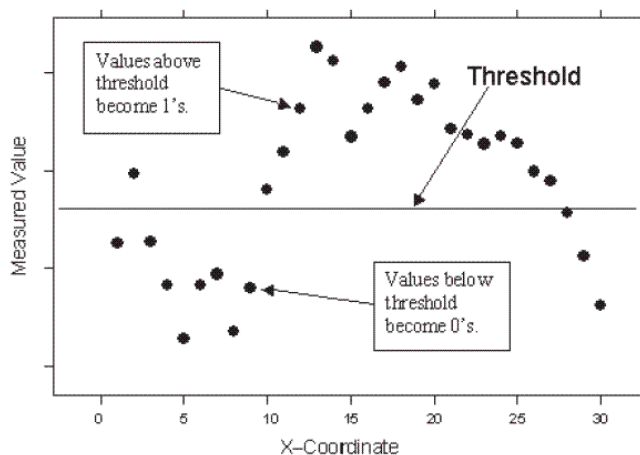
V řadě úloh nás nezajímá, zda je v daném místě nebo na dané ploše nejpravděpodobnější průměrná koncentrace 0,016 nebo 0,015, ale odhad pravděpodobnosti, s jakou je překročena limitní hodnota např. 0,012. Tyto úlohy řeší tzv. indikátorové krigování, které náleží do skupiny neparametrických geostatistických metod. Dále sem patří také soft kriging a pravděpodobnostní krigování. Pro tyto metody krigingu se zavádí pojem prahové hodnoty, kterou lze spojitou proměnnou převést na binární (viz. obr.)

Indikátorové krigování předpokládá model ve tvaru:

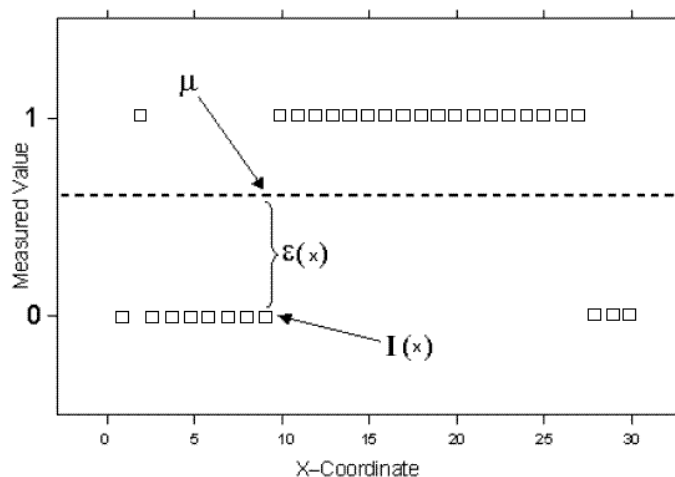
$$I(x_i) = \mu + \varepsilon(x_i)$$

kde μ neznámá konstanta, $\varepsilon(x)$ autokorelovaná náhodná proměnná a $I(x)$ je buďto přímo zjištěná binární proměnná a nebo binární proměnná, kterou obdržíme ze spojitých dat prahováním. Jinak se model neliší od základního krigování. Protože indikátorová proměnná nabývá hodnot 0 nebo 1, potom interpolované hodnoty budou

nabývat hodnot od 0 do 1 a lze je interpretovat jako pravděpodobnost, že proměnná nabývá hodnoty 1. Pokud bylo pro vytvoření indikátorové proměnné použito postupu prahování, potom lze vytvořenou mapu interpretovat jako pravděpodobnost překročení prahové hodnoty.



Obr. 5 Princip prahování hodnot proměnné



Obr. 6 Princip indikátorového krigování

Uvedený princip lze obecně rozšířit a použitím dvou více hodnot vytvořit např. dvě indikátorové proměnné (viz. dále - co-kriging).

Pravděpodobnostní krigování (probability kriging)

Pravděpodobnostní krigování předpokládá model ve tvaru:

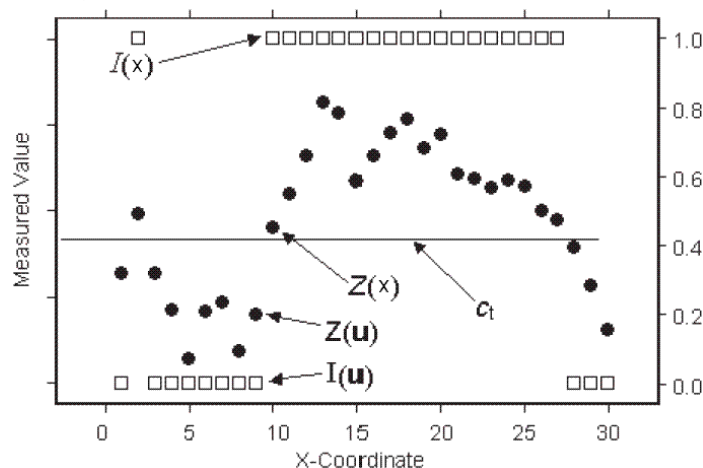
$$I(x) = I(Z(x) > c_1) = \mu_1 + \varepsilon_1(x)$$

$$Z(x) = \mu_2 + \varepsilon_2(x)$$

kde μ_1 a μ_2 jsou neznámé konstanty. $I(x)$ je binární proměnná vytvořená indikátorovým prahováním ($I(Z(x) > c_1)$). V tomto případě dostáváme dvě náhodné chyby $\varepsilon_1(x)$ a $\varepsilon_2(x)$. Cíle pravděpodobnostního krigování jsou stejné jako u krigování indikátorového, jsou však dosaženy využitím konceptu co-krigingu.

Na obrázku 7 má datový bod $Z(u=9)$ hodnotu indikátorové proměnné $I(u)=0$ a bod $Z(x=10)$ hodnotu $I(x)=1$. Pokud bychom chtěli predikovat hodnotu v polovině vzdálenosti mezi oběma body – na x-ové souřadnici 9,5, potom použitím modelu indikátorového krigování bychom obdrželi hodnotu 0,5. Z obrázku je však patrné, že datový bod $Z(x)$ je nepatrně nad hodnotou prahu, naopak bod $Z(u)$ je výrazně pod

prahovou hodnotou. Je tedy reálné předpokládat, že predikovaná proměnná v bodě 9,5 bude méně než 0,5.



Obr. 7 Princip pravděpodobnostního krigování

Pravděpodobnostní krigování se tedy snaží využít vedle indikátorové proměnné ještě další extra informace v původních datech. Nevýhodou pravděpodobnostního krigování je nutnost provádět odhady jako autokorelací pro jednotlivé proměnné, tak křížových korelací mezi nimi. Dalšími odhady neznámých autokorelací se vnáší do výsledného modelu větší míra nejistoty.

Nelineární kriging (log-normal)

Pokud nemají vstupní data normální rozdělení, je nutné je před vlastní interpolací transformovat. Nejběžnější je transformace lognormální. Originální data jsou transformována na přirozený logaritmus o základu 10. Tedy modelování variogramu a interpolace probíhá s proměnou $y(u)$:

$$y(u) = \ln z(u)$$

Predikované hodnoty je poté nutno transformovat nazpět, což může působit problémy (viz. Borrough et. al. 1992) a jako alternativa se nabízí indikátorový kriging. Pro některá FG data, která vykazují rozdělení s kladnou asymetrií, je však lognormální transformace výhodná (např. obsah chemických látek v půdě).

Kriging s využitím externí informace

K interpolaci kromě hodnot vlastní interpolované proměnné lze využít například:

1. vhodnou stratifikaci dat (**stratifikovaný kriging**)
2. hodnoty jiné proměnné, která koreluje s původní a kterou lze snadno měřit ve větším počtu bodů (např. výškové poměry) – **co kriging**
3. fyzikální či empiricky sestavený model, který podmiňuje rozložení hodnot studované proměnné

Stratifikovaný kriging spočívá v rozdělení oblasti na subregiony. Předpokládá dostatečný počet bodů pro výpočet hodnot variogramu. Může dávat vhodnější odhady, je však nutné řešit oblasti na styku subregionů. Např. obsah znečišťujících látek podle oblastí zaplavovaných podél vodního toku s různou frekvencí.

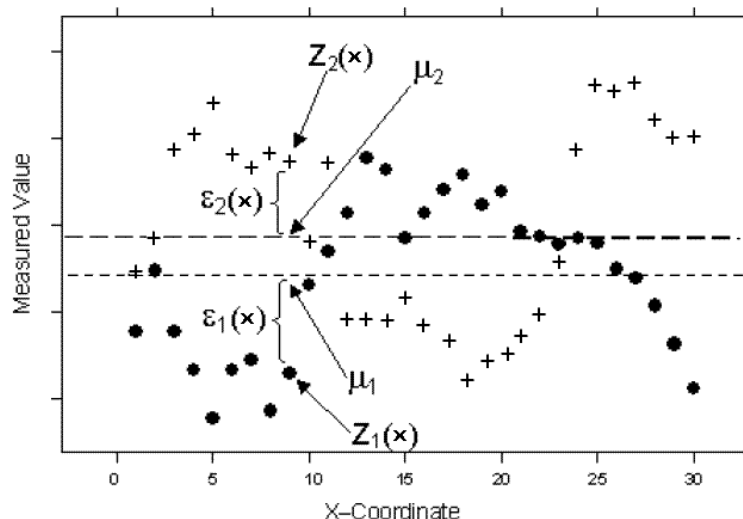
Co-kriging

Máme dvě proměnné z_1 a z_2 , které vykazují prostorovou korelaci. Pak lze využít hodnot proměnné z_2 k interpolaci hodnot proměnné z_1 . Tento koncept je vhodný zvláště v případech, kdy je proměnná z_2 snáze získatelná a rozšířitelná i na více než dvě proměnné. Přitom pro přesnější odhady se používá jak autokorelace jednotlivých proměnných, tak vzájemné (cross) korelace všech použitých proměnných. Základní co-kriging využívá následujících modelů:

$$Z_1(x) = \mu_1 + \varepsilon_1(x)$$

$$Z_2(x) = \mu_2 + \varepsilon_2(x)$$

kde μ_1 a μ_2 jsou neznámé konstanty. Dále dostáváme dvě náhodné chyby $\varepsilon_1(x)$ a $\varepsilon_2(x)$. Základní co-kriging odhaduje hodnotu proměnné $Z_1(x_0)$ stejně jako základní krigování, ovšem navíc využívá kovariance s hodnotu $Z_2(x)$.

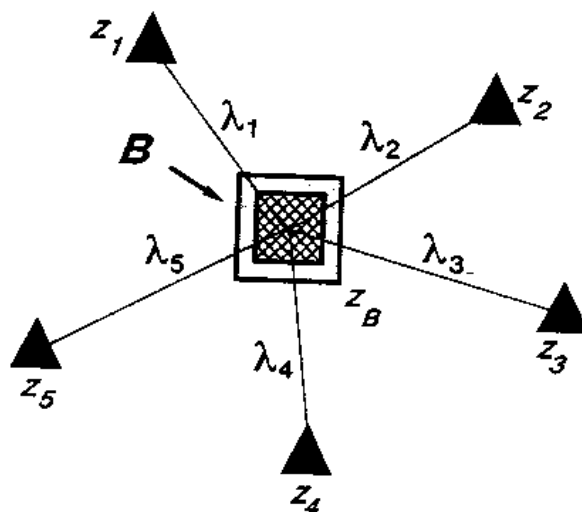


Obr. 8 Princip co-kringu

Z obrázku je patrné, že data Z_1 a Z_2 se jeví jako nekorelovaná. Dále pokud Z_1 je pod průměrem μ_1 , potom Z_2 je často nad průměrem μ_2 a naopak. Tedy Z_1 a Z_2 vykazují negativní cross korelaci. Vedle základního co-kringu jsou dalšími variantami např. jednoduchý, univerzální, indikátorový či pravděpodobnostní co-kriging.

Blokový odhad při základním krigování (Block kriging)

Lokální (bodový) odhad metodou krigingu lze určitým způsobem vztáhnout k ploše či objemu v prostoru interpolovaných dat. Mnoho přírodních jevů vykazuje značnou variabilitu a výsledkem bodového odhadu může být mapa obsahující značný počet ostrých vrcholů a depresí. Tento efekt lze potlačit tak, že modifikujeme výše uvedené rovnice a odhadneme průměrnou hodnotu $z(B)$ proměnné z pro jistou plochu či objem B (viz. obr). Tato modifikace je vhodná, pokud výsledkem interpolací být struktura pravidelných buněk (grid).



Obr. 9 Princip blokového krigování

Průměrná hodnota z pro blok B

$$z(B) = \int_B \frac{z(x)dx}{plocha_B}$$

bude odhadnuta z výrazu:

$$\hat{z}(B) = \sum_{i=1}^n \lambda_i \cdot z(x_i)$$

Kde stejně jako u bodového odhadu je suma všech vah λ_i rovna jedné.

Minimální rozptyl nyní bude:

$$\hat{\sigma}^2(B) = \sum_{i=1}^n \lambda_i \bar{\gamma}(x_i, B) + \phi - \bar{\gamma}(B, B)$$

a získáme ho, když

$$\sum_{i=1}^n \lambda_i \gamma(x_i, x_j) + \phi = \bar{\gamma}(x_j, B) \text{ pro všechna } j.$$

Rozptyly odhadů pro blokový kriging jsou daleko menší než pro bodový kriging. Výsledný interpolovaný povrch je obecně více shlazený a neobsahuje takové množství lokálních extrémů. Blokové korigování je aproximující metodou.

Hodnocení a verifikace modelů

Krigování jako interpolační metoda umožňuje pro každý interpolovaný bod odhadnout potenciální velikost chyby odhadu. Vedle map predikovaných hodnot tak lze především konstruovat mapy hodnot σ_e^2 (rozptyl krigingu), které vypovídají o spolehlivosti interpolovaných hodnot. Tyto hodnoty se obvykle prezentují v podobě map druhé mocniny σ_e^2 - tzv. **směrodatné chyby** (odchylky) **krigingu (Standard error map)**, protože tyto mají stejné jednotky jako predikované hodnoty. V některých případech se stanovuje také tzv. přesnost (relativní chyba) odhadu:

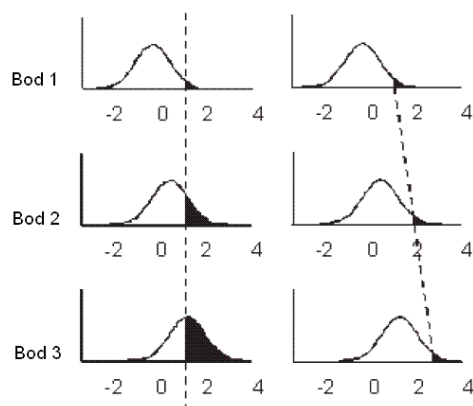
$$\delta_k = 100 \frac{\sigma}{\hat{z}}$$

Vydeme-li z výše uvedeného příkladu, kdy rozptyl odhadu je $\sigma_e^2 = 4,008$. Potom směrodatná chyba krigingu bude $\sigma_e = 2,002$. Budeme-li předpokládat, že chyby predikce mají normální rozdělení, potom 95% interval spolehlivosti predikovaných hodnot lze určit z následujícího vztahu:

$$Z(x_0) \pm 1,96 \cdot \sqrt{\sigma_e^2}$$

kde $Z(x_0)$ je odhad hodnoty proměnné z v bodě x_0 a σ_e^2 je rozptyl odhadu. V našem případě tedy při opakovaném použití stejného modelu padne 95 % odhadovaných hodnot do intervalu $(4,560 \pm 1,96 \cdot 2,002)$ tj. $(0,64; 8,48)$

Konstrukce dalších dvou typů map, které nabízí např. ArcGIS a kterými lze zhodnotit kvalitu interpolace vychází následujícího obrázku.



Obr. 10 Princip konstrukce Probability map a Quantile map (vysvětlivky viz. text)

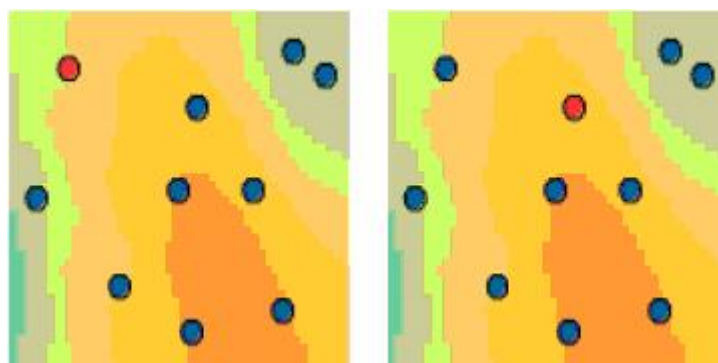
Předpokládáme, že krigováním predikované hodnoty mají ve třech různých bodech normální rozdělení a nacházejí se ve středu každé křivky rozdělení. Chceme-li určit pravděpodobnost, že predikovaná hodnota bude větší než prahová hodnota - např. 1, potom na obrázku vlevo tuto pravděpodobnost představuje na jednotlivých křivkách část plochy vpravo od prahové hodnoty (černé plochy). Při konstantní prahové hodnotě se její pravděpodobnost výskytu pro jednotlivé body mění – tedy lze z ní vytvořit mapu pravděpodobností (**probability map**).

Na obrázku vpravo je schematicky znázorněno, jakým způsobem určit kvantil s např. 5 procentní pravděpodobností výskytu. Tuto pravděpodobnost v tomto případě opět značí černá plocha vpravo od prahové hodnoty a hodnotu kvantilu odečteme na ose x. Při konstantní pravděpodobnosti se budou měnit hodnoty kvantilů a lze je opět prezentovat ve formě kvantilové mapy (**quantile map**).

Validace a křížová validace predikovaných hodnot metodou krigingu

Hodnocení přesnosti interpolace lze provádět také pomocí dále popsaných grafických nástrojů

Křížová validace modelu - k vytvoření spojitého povrchu jsou použita všechna vstupní data v měřených bodech. Poté jsou jednotlivé body měření (červené) po jednom postupně vynechány ze vstupní množiny dat a ze zbývajících (modrých) je vypočtena hodnota v místě vynechaného bodu.



Obr. 11 Princip křížové validace modelu

Statistické zhodnocení

Procesem křížové validace obdržíme veličiny, které mají následující význam:

$\hat{Z}(x_i)$ je predikovaná hodnota pro daný bod x_i , kterou obdržíme v procesu křížové validace

$\hat{\sigma}(x_i)$ je směrodatná chyba predikce, tedy druhá odmocnina z výrazu pro rozptyl krigování:

$$\hat{\sigma}_e^2 = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \phi$$

Pozorované a vypočtené hodnoty jsou následně porovnány dále uvedenými měrami:

MPE – mean prediction error - průměr rozdílů měřených a předikovaných hodnot - hodnoty chyb odhadů by měly být nestranné – tedy jejich průměr by se měl rovnat nule.

$$MPE = \frac{\sum_{i=1}^n (\hat{Z}(x_i) - z(x_i))}{n}$$

RMSPE (root mean square prediction error) – druhá odmocnina průměrného čtverce vzdálenosti vypočtených hodnot (červené body) od teoretických (zelená přímková v grafech). Tato hodnota slouží k porovnání několika různých modelů. Čím menší je RMSPE, tím vhodnější je model (tím bližší jsou vypočtené hodnoty hodnotám měřeným).

$$RMSPE = \sqrt{\frac{\sum_{i=1}^n (\hat{Z}(x_i) - z(x_i))^2}{n}}$$

ASE (average standard error) - průměrná směrodatná chyba

$$ASE = \sqrt{\frac{\sum_{i=1}^n \hat{\sigma}(x_i)}{n}}$$

Výše uvedené nástroje umožňují posoudit vhodnost modelu a také porovnat více modelů vzájemně mezi sebou.

MSPE (mean standardized prediction error) - průměrná standardizovaná chyba predikce

$$MSPE = \frac{\sum_{i=1}^n (\hat{Z}(x_i) - z(x_i)) / \hat{\sigma}(x_i)}{n}$$

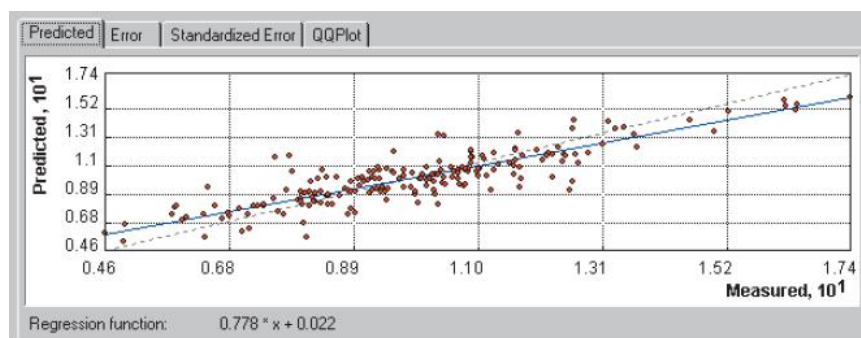
RMSSPE (root mean square standardized prediction error)

$$RMSSPE = \sqrt{\frac{\sum_{i=1}^n [(\hat{Z}(x_i) - z(x_i)) / \hat{\sigma}(x_i)]^2}{n}}$$

Validace modelu - vstupní soubor měřených hodnot rozdělí na dvě části – data trénovací a testovací. Trénovací množina dat se použije pro odhad trendu a autokorelačního modelu. Pokud sestavený model vyhovuje trénovacím datům, je ověřen na datech testovacích.

Pro oba zmíněné způsoby ověření vhodnosti modelu se využívá sady grafických nástrojů. Nejběžnějším je graf korelačního pole měřených a predikovaných hodnot. Obecnou vlastností krigingu jako interpolační metody je podhodnocení vysokých

hodnot a naopak nadhodnocení hodnot nízkých. Tato vlastnost se projeví menší hodnotou směrnice přímky proložené korelačním polem.



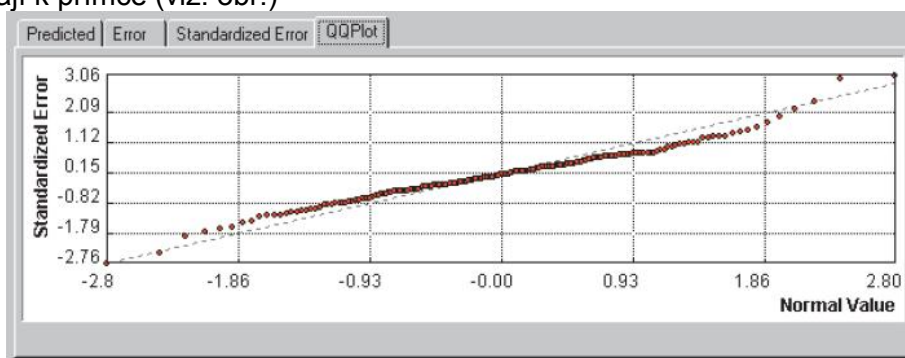
Obr. 12 Korelační pole měřených a predikovaných hodnot

Chybový graf (Error plot) – stejný jako předchozí, jsou však vynášeny hodnoty rozdílů mezi měřenými a predikovanými hodnotami

Standardizovaný chybový graf (Standardized Error) – hodnoty rozdílů mezi měřenými a predikovanými hodnotami jsou děleny odhadnutou směrodatnou chybou krigování.

V případě nulové autokorelace budou všechny predikované hodnoty stejné – budou odpovídat průměru a proložená přímka bude mít horizontální průběh. V případě prostorové autokorelace a vhodného modelu krigingu bude proložená přímka totožná s diagonálou a navíc body korelačního pole budou vykazovat malé odchylky od diagonálního směru.

Q-Q graf – znázorňuje graf kvantilů rozdílů mezi měřenými a predikovanými hodnotami dělenými odhadnutou směrodatnou chybou krigování a odpovídajících kvantilů normovaného normálního rozdělení. V případě, že odchylky měřených a odhadnutých hodnot mají normální rozdělení, potom se body v korelačním poli přibližují k přímce (viz. obr.)



Obr. 13 Příklad Q-Q grafu

Interpretace statistických charakteristik k hodnocení vhodnosti modelu:

- **Požadavek nestrannosti odhadu – unbiased** - průměrná chyba odhadu a standardizovaná průměrná chyba odhadu by se měly blížit k nule:
MPE \rightarrow 0
MSPE \rightarrow 0
- **Požadavek minimálních chyb** – aby predikované hodnoty byly co nejblíže hodnotám měřeným. Čím menší bude hodnota RMSPE, tím lepší model – tedy tuto podmínku lze použít k porovnání vhodnosti více modelů.
RMSPE \rightarrow min.
- **Požadavek vhodné variability predikovaných dat** – variabilita predikovaných hodnot je určována z hodnot měřených. Je tedy důležité, aby i variabilita interpolací vypočtených hodnot byla vhodná:

- ASE \approx RMSPE – vhodný model (vhodná variabilita predikovaných hodnot)
- ASE $>$ RMSPE – máš model nadhodnocuje variabilitu odhadnutých hodnot
- ASE $<$ RMSPE – máš model podhodnocuje variabilitu odhadnutých hodnot

V případě značného podílu šumové složky (např. v důsledku chyb v měření) či v případě značně komplexního povrchu nedává kriging lepší výsledky než jiné interpolátory. Na rozdíl o jiných metod kriging nabízí objektivní, a priori metodu odhadu vhodného okolí pro vlastní interpolaci. Řeší tedy otázku počtu bodů v okolí daného bodu, otázku velikosti a tvaru tohoto okolí. V případě existence bariér (náhlých skoků v hodnotách interpolovaného povrchu nedává kriging dobré výsledky a je nutné jej rozdělit na elementární části neobsahující bariéry.