

České vysoké učení technické v Praze

Fakulta stavební

Katedra mapování a kartografie



Interpolace bodových dat v GIS

Bakalářská práce

Linda Křikavová

květen 2009

Vedoucí bakalářské práce: Ing. Jiří Cajthaml, Ph.D.

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně. Veškerou použitou literaturu a podkladové materiály jsem uvedla v přiloženém seznamu.

V Praze dne 22. 5. 2009

Na tomto místě bych ráda poděkovala vedoucímu bakalářské práce Ing. Jiřímu Cajthamlovi, Ph.D. za navržení zajímavého tématu a za čas, který mi věnoval. Dále bych chtěla poděkovat rodičům za všestrannou podporu v průběhu dosavadního studia.

Interpolace bodových dat v GIS

Abstrakt

Práce se zabývá interpolací bodových dat a její realizací v Geografickém informačním systému (GIS). První část je teoretická, uvedeno je třídění interpolačních metod, jednotlivé metody jsou popsány a vysvětlen jejich matematický základ. V druhé, praktické části je ukázáno použití interpolačních metod v programu ArcGIS a jeho nadstavbách 3D Analyst, Spatial Analyst a Geostatistical Analyst. Výsledky jednotlivých metod jsou porovnány a zhodnoceny.

Klíčová slova

interpolace, extrapolace, odhad, autokorelace, *Thiessenovy polygony*, *metoda přirozeného souseda*, *IDW*, *spline*, *trend*, *kriging*, *topo to raster*, ArcGIS

Interpolation of point data in GIS

Abstract

The thesis deals with the interpolation of point data and its application in Geographic Information System (GIS). The first theoretical part includes the classification of interpolation methods. Each method is explained and defined on the mathematical base. The second practical part copes with the implementation of interpolation in the ArcGIS software and its extensions 3D Analyst, Spatial Analyst and Geostatistical Analyst. The results of interpolation methods are evaluated and compared to each other.

Key words

interpolation, extrapolation, estimation, autocorrelation, *Thiessen polygons*, *natural neighbour*, *IDW*, *spline*, *trend*, *kriging*, *topo to raster*, ArcGIS

Obsah

Úvod.....	6
1 Interpolace bodových dat	7
2 Interpolační metody	7
2.1 Klasifikace interpolačních metod.....	7
2.2 Deterministické interpolační metody	8
2.2.1 Thiessenovy polygony	8
2.2.2 Metoda přirozeného souseda (natural neighbour).....	9
2.2.3 Inverse distance weighting (IDW)	10
2.2.4 Spline	12
2.2.5 Trend	13
2.3 Stochastické interpolační metody	15
2.3.1 Kriging	15
3 Interpolace bodových dat v ArcGIS.....	22
3.1 Metoda přirozeného souseda (natural neighbour).....	23
3.2 Inverse distance weighting (IDW)	24
3.3 Spline	29
3.4 Trend	33
3.5 Kriging	37
3.6 Topo to raster	46
3.7 Hodnocení kvality interpolace	51
Závěr	53
Literatura	54

Úvod

Často se setkáváme se skutečností, že máme k dispozici údaje o daném jevu pouze prostřednictvím určitého množství bodových dat, ať už je to z důvodu náročného sběru dat, vysoké ceny nebo proto, že realizace vyčerpávajícího sběru dat je nemožná. Většinou nás ale zajímají i hodnoty v místech, kde měření provedeno nebylo. Potřebujeme znát celkové prostorové rozložení hodnot sledované proměnné, čehož docílíme interpolací známých bodových dat. Interpolací tedy získáme tzv. mapu sledovaného jevu, která může být vyjádřena i ve formě prostorového modelu. Příkladem je klimatický model vytvořený z údajů meteorologických stanic, model terénu z nadmořských výšek nebo model ložisek nerostných surovin vytvořený na základě vrtů.

V první části této práce bude nejdříve představena klasifikace interpolačních metod dle různých hledisek. Dále budou vysvětleny jednotlivé metody obecně, bude uveden jejich matematický základ. Druhá část bude věnována praktické ukázce použití metod v programu ArcGIS (verze 9.2). Jedná se o nadstavby ArcGIS 3D Analyst, Spatial Analyst a Geostatistical Analyst. U každé metody bude popsána volba parametrů, ukázána aplikace na dvou vzorkách dat a navrženo vhodné použití. Výsledky některých vybraných metod budou mezi sebou porovnány.

V obou částech bude zahrnuta *metoda přirozeného souseda*, *IDW*, *spline*, *trend* a *kriging*. *Thiessenovy polygony* budou zmíněny pouze v první teoretické části, nejedná se totiž o metodu, která by byla používána jako taková. Uvedena bude z důvodu, že je základem jiných metod. Naopak pouze v druhé praktické části bude ukázána metoda *topo to raster*, která je výsadním produktem ArcGIS a její přesné matematické pozadí není k dispozici.

V textu se budou často objevovat anglické termíny. Pokud jsou přeloženy, budou uvedeny v závorce. V případě, že český ekvivalent neexistuje a přeložení pojmu by bylo nepřesné nebo zavádějící, výraz je ponechán v angličtině. Zachování anglických výrazů je vhodné zvláště v druhé části práce, kde jsou zmiňovány pojmy užívané v programu ArcGIS, jehož pracovní prostředí je v angličtině.

1 Interpolace bodových dat

Termín interpolace bodových dat vyjadřuje proces výpočtu neznámých hodnot určitého jevu na základě známých bodových dat. Tato práce se zabývá případem, kdy body leží v dvourozměrném prostoru. Místo označení interpolace se často také používá pojem odhad nebo predikce neznámých hodnot. Aby bylo možné bodová data interpolovat, musí být splněna podmínka, že určitý sledovaný jev je spojitý nebo prostorově závislý (autokorelovaný).

K interpolaci se váže pojem extrapolace. Rozdíl mezi těmito postupy je v tom, že při interpolaci odhadujeme neznámé hodnoty v rámci tzv. prostorové domény známých údajů, zatímco při extrapolaci odhadujeme hodnoty mimo její rozsah. Extrapolace je obecně méně přesný postup, kterému je doporučeno se spíše vyhnout, důvodem je například možnost získání záporných odhadů striktně pozitivní proměnné.

2 Interpolační metody

2.1 Klasifikace interpolačních metod

Existují různá hlediska, podle kterých můžeme rozdělovat interpolační metody:

A. Deterministické a stochastické metody

Deterministické metody provádějí interpolaci přímo z měřených hodnot vstupních bodů. Není zde využita teorie pravděpodobnosti, pokaždé bude vypočten stejný výsledný odhad.

Stochastické metody zahrnují prvek náhodnosti. Výsledná prostorová predikce je chápána jako jedna z mnoha, která mohla být vytvořena. Metoda je založena na statistickém modelu, který předpokládá prostorovou závislost mezi vstupními body.

Toto rozdělení bude dále použito ve struktuře následujícího textu.

B. Globální a lokální metody

Globální metody zahrnují všechny měřené body do výpočtu odhadu. Jedna matematická funkce je použita k popisu zkoumaného jevu na celém zájmovém území. Změna jediného vstupního bodu ovlivní výsledný odhad v celé lokalitě.

Lokální metody aplikují algoritmus výpočtu opakovaně na malou část z celkového počtu bodů. Změna vstupní hodnoty ovlivní pouze výsledek lokální oblasti.

C. Exaktní a aproximující metody

Exaktní metody přesně respektují původní naměřené hodnoty, na kterých je interpolace založena. Výsledný odhad prochází všemi body, jejichž hodnota je známa. V opačném případě se jedná o aproximující metody.

D. Metody váženého průměru a metody základních funkcí

Metody založené na váženém průměru hodnot známých bodů lze obecně vyjádřit jako:

$$\hat{z}(s_0) = \sum_{i=1}^n \lambda_i z(s_i) ,$$

kde $\hat{z}(s_0)$ je hodnota odhadovaná v bodě s_0 umístěném v r -rozměrném prostoru. Ve dvourozměrném prostoru platí $s = f(x, y)$. $z(s_i)$ je měřená hodnota v bodě s_i , n je počet měření a λ_i je váha měřené hodnoty v bodě s_i . Předpokládáme, že suma vah je rovna jedné:

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

Tímto způsobem je možné přistupovat k na první pohled diametrálně odlišným metodám. Na tomto základě budou dále vysvětleny metody *Thiessenovy polygony*, *metoda přirozeného souseda*, *IDW* a *kriging*.

Druhou skupinou jsou metody používající lineární kombinaci základních matematických funkcí. Podle typu těchto funkcí rozeznáváme *spline*, *trend* a další.

2.2 Deterministické interpolační metody

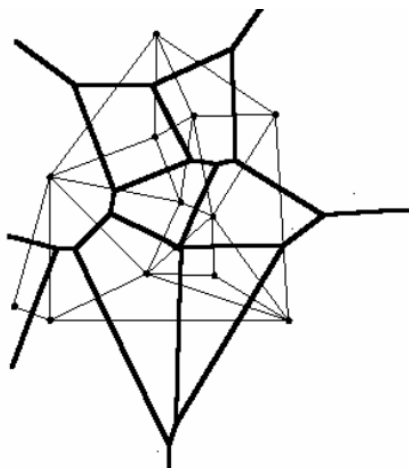
2.2.1 Thiessenovy polygony

Thiessenovy polygony, též nazývané *Voronoiovy polygony*, jsou definovány jednoduchým geometrickým pravidlem. Každý polygon obsahuje právě jeden měřený bod. Ostatní body uvnitř polygonu jsou blíže tomuto měřenému bodu než kterémukoliv jinému měřenému bodu.

Postup tvorby *Thiessenových polygonů* je relativně jednoduchý. V prvním kroku jsou určeny spojnice sousedních měřených bodů. Celá zájmová plocha je následně rozdělena liniemi kolmými na tyto spojnice, tím jsou definované hranice polygonů (obr. 1). Kolmé linie jsou umístěny tak, aby v každém polygonu ležela část území bližší k měřenému bodu než ke kterémukoliv jinému. To znamená v polovině spojnice

sousedních bodů. Odhad hodnoty každého bodu polygonu P_i je rovný hodnotě zdrojového bodu, který v něm leží.

$$\lambda_i \begin{cases} 1 & s_i \in P_i \\ 0 & s_i \notin P_i \end{cases} \quad \hat{z}(s_0) = z(s_i)$$

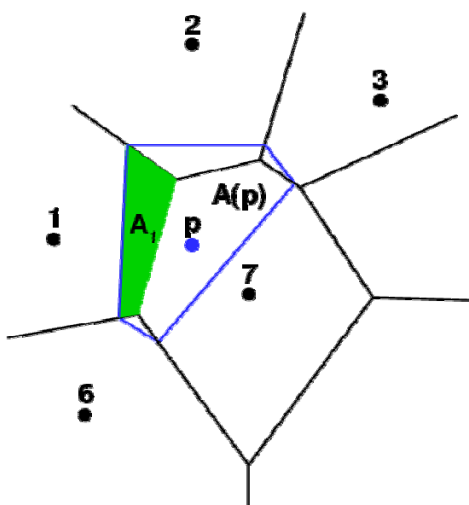


Obr. 1 *Thiessenovy polygony*

Nedostatek této metody je evidentní, každý odhad je založený jen na jedné hodnotě. Navzdory tomu, že zkoumaný jev je spojitý, výsledná struktura bude vždy diskrétní. Dalším problémem je, že polygony okrajových bodů mají teoreticky nekonečnou plochu, z tohoto důvodu musí být ořezány hranicemi zkoumaného území.

Využití *Thiessenových polygonů* jako samostatné metody je minimální, přesto jsou základem některých jiných interpolačních metod, jako TIN nebo přirozený soused.

2.2.2 Metoda přirozeného souseda (natural neighbour)



Obr. 2 *Metoda přirozeného souseda*

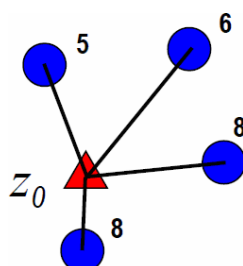
Metoda přirozeného souseda využívá pro určení vah *Thiessenovy polygony*. Vložení interpolovaného bodu do sítě *Thiessenových polygonů* způsobí její přebudování v okolí tohoto bodu (obr. 2). Nyní máme dvě sítě polygonů - původní a novou, která vznikla přidáním určovaného bodu. Polygon nového bodu překrývá určité části původních polygonů známých bodů. Tyto body, kterým se říká přirození sousedé, budou zahrnuty do interpolace bodu nového. Váhy přirozených sousedů jsou plochy, které jsou odděleny z původních polygonů jednotlivých sousedů.

$$A(p) = \sum_{i=1}^n A_i, \quad \lambda_i = \frac{A_i}{\sum_{i=1}^n A_i},$$

kde A_i je část plochy původního polygonu přirozeného souseda i , $A(p)$ je plocha polygonu nového bodu p .

Metoda je velice efektivní, pokud jsou měřené hodnoty umístěny pravidelně. Výsledná struktura modelovaného jevu je spojitá a vyhlazená, bez extrapolovaných hodnot.

2.2.3 Inverse distance weighting (IDW)

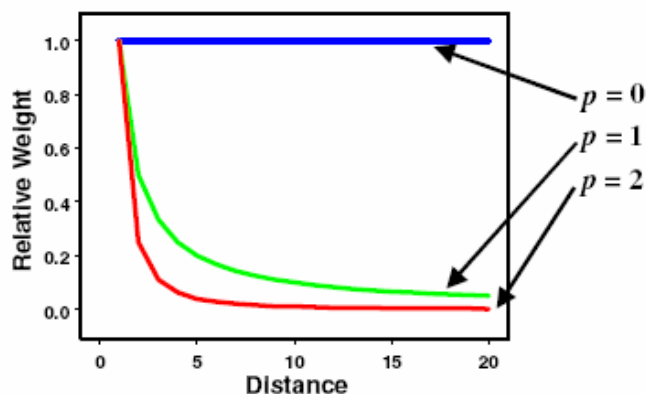


Obr. 3 IDW

IDW uplatňuje základní geostatistický princip: jevy, které jsou v prostoru blíže k sobě, se více podobají než jevy, které jsou prostorově vzdálenější. Určení vah vstupních bodů je založeno na tomto principu. Základní koncept je použití inverzní vzdálenosti od určovaného bodu. Tato vzdálenost je dále upravovaná exponentem p . Z toho vyplývá, že váhy se řídí jediným prostorovým modelem - inverzní vzdáleností modifikovanou vhodnou mocninou:

$$\lambda_i = \frac{\frac{1}{d_{i,0}^p}}{\sum_{i=1}^n \frac{1}{d_{i,0}^p}},$$

kde $d_{i,0}$ je vzdálenost mezi známým a určovaným bodem, p je parametr, který se volí $p = 1, 2$ nebo 3 . Čím nižší je hodnota tohoto parametru, tím podobnější váhy jsou přiřazené měřeným bodům a výsledný odhad se blíží aritmetickému průměru. V případě, že p by se rovnalo nule, všechny váhy budou stejné bez ohledu na rozmístění měřených bodů a odhad bude rovný aritmetickému průměru. Čím vyšší je hodnota parametru, tím rozdílnější jsou váhy – vzdálenější body mají nižší váhu, bližší body vyšší (obr. 4). Nejčastěji používaný exponent je 2.

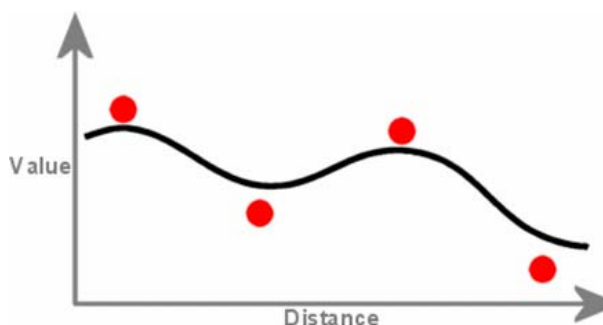


Obr. 4 Volba exponentu p

Nedostatkem této metody je tvorba koncentrických izolinií (bull eyes) okolo vstupních bodů. Důvodem je silný vliv těchto bodů ve svém okolí, především při volbě vyšší hodnoty exponentu. Pro zmírnění tohoto jevu se zavádí vyhlazovací parametr σ . Při použití parametru σ se tato metoda mění z exaktní na aproximující (obr. 5).

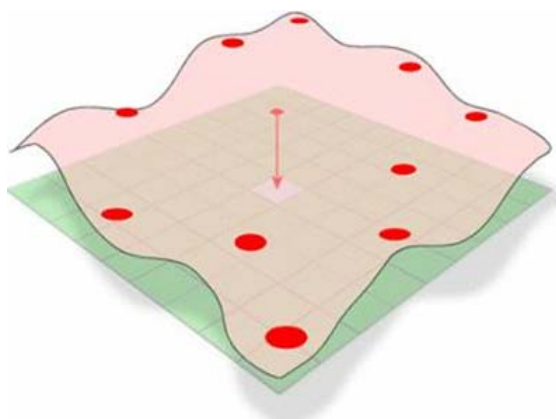
$$\lambda_i = \frac{1}{\sum_{i=1}^n \frac{1}{(d_{i,0} + \sigma)^p}}$$

Metoda *IDW* nedokáže vypočítat hodnoty vyšší nebo nižší než jsou hodnoty vstupních dat. Pokud tedy nemáme k dispozici body měřené v místech s extrémními hodnotami, dochází v těchto místech ke zkreslení.



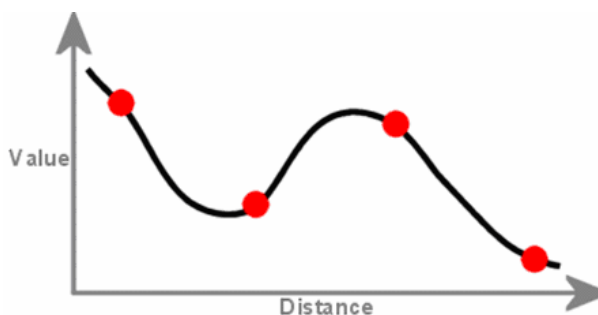
Obr. 5 *IDW*

2.2.4 Spline



Obr. 6 Spline

Metoda *spline* odhaduje neznámé hodnoty použitím matematických funkcí. Metoda je založena na dvou podmínkách – prokládaná matematická funkce prochází měřenými body a má minimální křivost. Splinová funkce imituje tenký flexibilní povrch nucený procházet přesně měřenými body (je exaktní) a zároveň tento povrch musí mít minimální zakřivení (musí být co nejhladší).



Obr. 7 Spline

Každá část povrchu je reprezentována samostatnou polynomicou funkcí odvozenou z lokálních hodnot, přitom musí být zajištěna spojitost sousedních polynomických funkcí na jejich styku (v případě kubických polynomů musí být totožné první dvě derivace). Výsledkem je hladký povrch respektující vstupní body.

Funkce *spline* se zapisuje následujícím vzorcem:

$$S(s) = T(s) + \sum_{i=1}^n \lambda_i R(r_i) ,$$

kde $T(s)$ je polynomická funkce, λ_i jsou neznámé koeficienty měřených bodů, $R(r_i)$ je radiální základní funkce (radial basis function) závislá na vzdálenosti r_i , r_i je vzdálenost mezi bodem $s = f(x, y)$ a i -tým měřeným bodem, n je počet měřených bodů. Funkce

$T(s)$ a $R(r_i)$ jsou voleny různě dle typu metody *spline*. Místo názvu *spline* se často používá označení radial basis function (RBF). Existuje mnoho druhů RBF, například:

- *thin plate spline (TPS)*
- *regularized spline*
- *tension spline*

Podmínka minimální křivosti je matematicky vyjádřena minimalizací sumy kvadrátů členů druhé derivace povrchu.

Thin plate spline (TPS) je základní formou *spline*. V jeho první derivaci povrchu (sklon) se mohou objevit prudké změny v okolí měřených bodů. Z tohoto důvodu není TPS vhodný k odhadu druhé derivace povrchu (křivost).

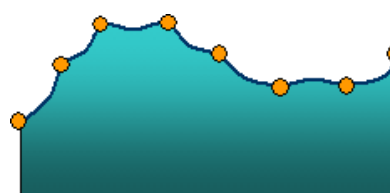
Typ *regularized spline* obměňuje podmínku minimalizace křivosti přidáním členů třetí derivace. Parametr váhy vyjadřuje váhu přiřazenou třetí derivaci při minimalizaci. Vyšší váha znamená menší křivost neboli hladší povrch.

Typ *tension spline* obměňuje podmínku minimalizace křivosti přidáním členů první derivace. Parametr váhy vyjadřuje váhu přiřazenou první derivaci při minimalizaci. Pokud je váha rovna nule, jedná se o TPS. Čím vyšší hodnota váhy, tím je povrch hrubší.

Metoda *spline* umí vypočítat i vyšší nebo nižší hodnoty než jsou extrémní hodnoty vstupních bodů, což výše popsaná metoda *IDW* nedokáže. Porovnání těchto dvou metod je ukázáno na příčném řezu měřenými body (obr. 8 a obr. 9). Interpolovaná místa u *spline* podléhají podmínce co nejhladšího povrchu, u *IDW* závisí na váženém průměru z okolních bodů.



Obr. 8 *Spline*



Obr. 9 *IDW*

2.2.5 Trend

Metoda *trend* používá pro odhad neznámých hodnot následující model:

$$z(s) = \mu(s) + \varepsilon(s) ,$$

kde $z(s)$ je hodnota funkce sledovaného jevu v bodě s , $\mu(s)$ je střední hodnota (trend) v bodě s , $\varepsilon(s)$ je náhodná chyba v tomto bodě. Cílem metody je proložit vstupní body polynomickeou funkcí zpravidla prvního až třetího řádu ve tvaru:

$$\mu(s) = \sum_{u+v \leq p=0}^P a_p x^u y^v = \sum_{p=0}^P a_p f^p(s) .$$

Z tohoto vyjádření vyplývá, že očekávanou hodnotu v bodě s je možné vyjádřit jako lineární kombinaci P známých funkcí $f^p(s)$ a neznámých koeficientů a_p , kde p je stupeň polynomické funkce. Vybraná polynomická funkce je nastavena na dané body metodou nejmenších čtverců. Je potřeba dosáhnout toho, aby čtverec rozdílu mezi danou hodnotou a odhadovanou hodnotou v témže bodě byl v úhrnu minimální:

$$\sum_{i=1}^n \left(z(s_i) - \sum_{p=0}^P a_p f^p(s_i) \right)^2 = \min .$$

V případě lineárního *trendu* v dvojrozměrném prostoru dostáváme následujícími vztahy:

$$z = a_0 + a_1 x + a_2 y ,$$

$$z = S \cdot a ,$$

$$\text{kde } z = \begin{pmatrix} z(s_1) \\ z(s_2) \\ \vdots \\ z(s_n) \end{pmatrix}, \quad S = \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{pmatrix}, \quad a = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} .$$

Matice S obsahuje souřadnice vstupních bodů (případně jejich mocniny) v n řádcích podle n vstupních bodů a P sloupcích podle počtu neznámých koeficientů. Měřené hodnoty v těchto bodech vyjadřuje vektor z . a je vektor neznámých koeficientů, který získáme takto:

$$a = (S^T S)^{-1} S^T z .$$

Výsledný odhad je pak:

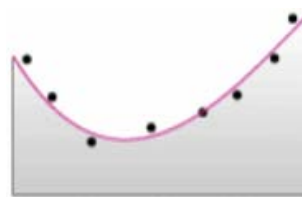
$$\hat{z}(s_0) = s_0 \cdot a ,$$

$$\text{kde } s_0 = (1 \quad x_0 \quad y_0) .$$

Při vybírání stupně polynomické funkce se lze řídit obecným pravidlem: jakýkoliv příčný řez polynomické funkce řádu p může mít nejvýše $p - 1$ střídajících se maxim a minim. Polynom prvního řádu popisuje nakloněnou rovinu (obr. 10), polynom druhého řádu popisuje například jednoduchý pahorek nebo údolí (obr. 11).



Obr. 10 Lineární *trend*



Obr. 11 Kvadratický *trend*

Tato metoda může být použita pro globální odhad trendu zkoumaného jevu použitím jedné funkce pro celé zájmové území. Hodnoty koeficientů trendové funkce však mohou být určeny i pro určité omezené okolí odhadu. Tomto případě cílem není určení globálního trendu, ale odhad konkrétních neznámých hodnot. Globální trend je odhadovaný v případě, kdy je okolí odhadu rozšířené na celé zájmové území. Nedostatkem tohoto postupu je deformace výsledné struktury v okrajových částech území, se snahou co nejlepšího odhadu v centrální části.

Metoda *trend* je aproximující, v důsledku toho může poskytovat v extrapolačních oblastech neslučitelné výsledky. V případě, že se v okrajových částech nachází nízké hodnoty měřených bodů, metoda může poskytovat záporné odhady, což nemusí vyhovovat podstatě zkoumaného jevu.

2.3 Stochastické interpolační metody

Druhá skupina interpolačních metod zahrnuje prvek náhodnosti. Stochastické metody jsou založeny na statistickém modelu, který zahrnuje autokorelaci – statistický vztah mezi měřenými body. Tyto metody vedou nejen k vytvoření prostorové predikce, ale umožňují rovněž určit i přesnost této předpovědi.

2.3.1 Kriging

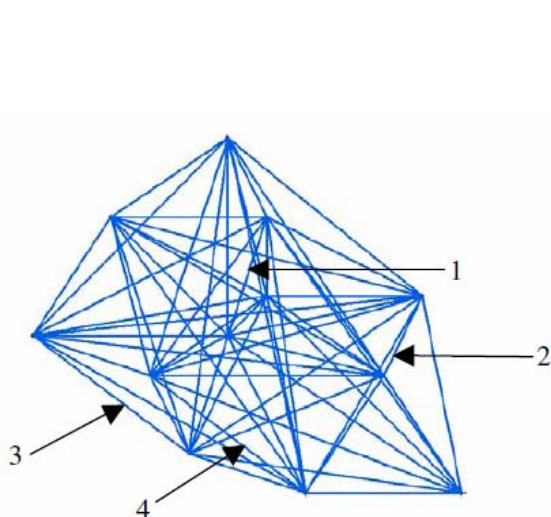
Metoda *kriging* je podobná deterministické metodě *IDW*. V případě *IDW* váhy závisí výlučně na vzdálenosti od předpovědního místa. U metody *kriging* váhy nezávisí pouze na vzdálenosti mezi měřenými body a předpovědním místem, ale také na prostorovém uspořádání měřených bodů okolo místa předpovídané hodnoty. Abychom mohli použít prostorové uspořádání měřených bodů pro výpočet vah, musí být určena prostorová autokorelace (prostorová závislost).

Prostorovou autokorelaci jevu vzhledem ke vzdálenosti a směru působení vyjadřuje semivariogram. Výpočet empirického semivariogramu se zapisuje ve tvaru:

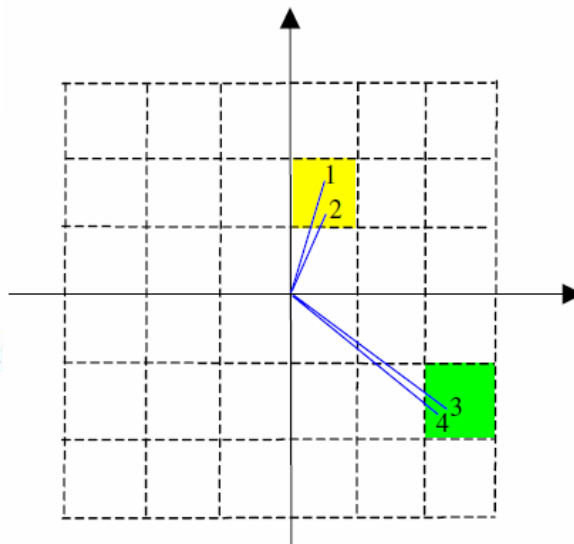
$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(s_i) - z(s_i + h)]^2,$$

kde $\gamma(h)$ je odhadovaná semivariance pro vzdálenost h , $n(h)$ je počet všech párů měřených bodů oddělených vzdáleností h , $z(.)$ je měřená hodnota v bodě $(.)$.

Prvním krokem výpočtu empirického semivariogramu je určení vzdáleností všech párů měřených bodů (obr. 12).



Obr. 12 Vzdálenosti všech párů

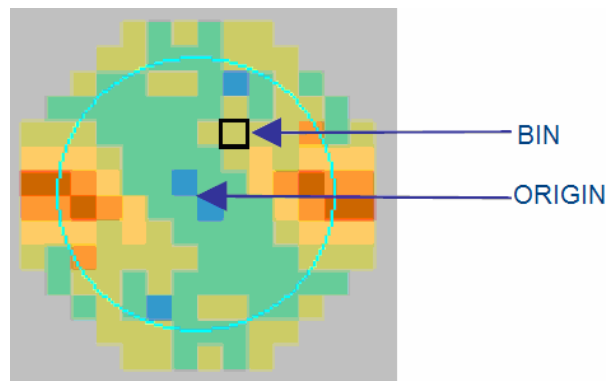


Obr. 13 Směry a vzdálenosti párů

Je zřejmé, že s rostoucím počtem měřených bodů prudce stoupá počet vzdáleností a výpočet se stává nezvladatelným. Řešením tohoto problému je seskupování hodnot (binning). Vzdálenosti mezi měřenými body jsou transformovány do skupin (bins) na základě jejich vzdálenosti a směru (obr. 13). Seskupování hodnot se provádí v souřadnicovém systému, ze kterého data pochází. Tím vzniká síť, jejíž buňky tvoří základ pro výpočet reprezentantů vzdáleností jednotlivých skupin. Následující výpočty jsou pak prováděny s těmito vzdálenostmi. Výpočet empirické semivariance pro reprezentující vzdálenost každé skupiny (bin) se poté provede dle už výše zmíněného vzorce:

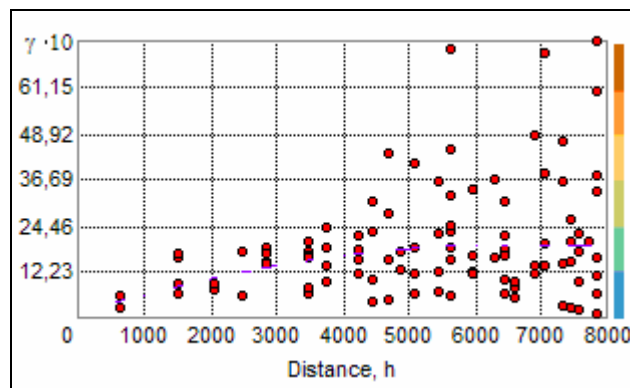
$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(s_i) - z(s_i + h)]^2 .$$

Vzorec vyjadřuje umocněný rozdíl měřených hodnot všech párů ve skupině, který je zprůměrnován. Dále je výraz vydělen dvěma, protože vzdálenost párů je počítána obousměrně (z místa s_i na místo $s_i + h$ a obráceně). Tímto je tedy dána jedna hodnota empirického semivariogramu pro každou skupinu (bin). Tyto hodnoty mohou být barevně kódovány (obr. 14).



Obr. 14 Barevné kódování empirického semivariogramu

Barvy vyjadřující hodnotu semivariance jsou souměrné podle počátku (origin). Studené barvy (modrá a zelená) vyjadřují nižší hodnoty, teplé barvy (červená a oranžová) symbolizují vyšší hodnoty semivariance.



Obr. 15 Empirický semivariogram

V grafu (obr. 15) jsou hodnoty empirického semivariogramu pro každou skupinu (bin) zobrazeny červenou tečkou. Na ose y je vynesena semivariance, na ose x vzdálenost od počátku ke středu buňky (bin).

Jak již bylo řečeno, semivariogram vyjadřuje prostorovou autokorelaci. Z grafu je vidět, že prostorová autokorelace splňuje základní geostatistický princip: jevy, které jsou v prostoru blíže k sobě, se více podobají než jevy, které jsou prostorově vzdálenější. Páry měřených bodů, které mají mezi sebou menší vzdálenost (na empirickém semivariogramu zobrazeny jako body vlevo na ose x), nabývají podobnějších hodnot (na empirickém semivariogramu body na ose y dole).

V dalším kroku dochází k vytvoření modelu pomocí empirického semivariogramu. Tento proces je spojnicí mezi prostorovým popisem určitého jevu a jeho prostorovou predikcí. Cílem modelu je předpověď hodnot v neměřených místech.

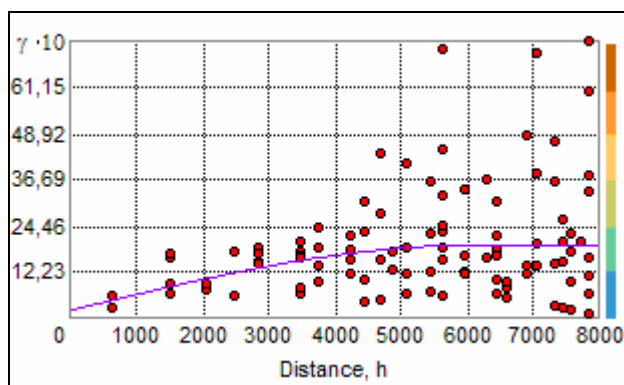
Empirický semivariogram poskytuje informace o hodnotách v místech měření, potřeba je ale určit hodnoty ve všech směrech a vzdálenostech.

Existuje celá řada funkcí, které je možné použít pro modelování empirického semivariogramu, například:

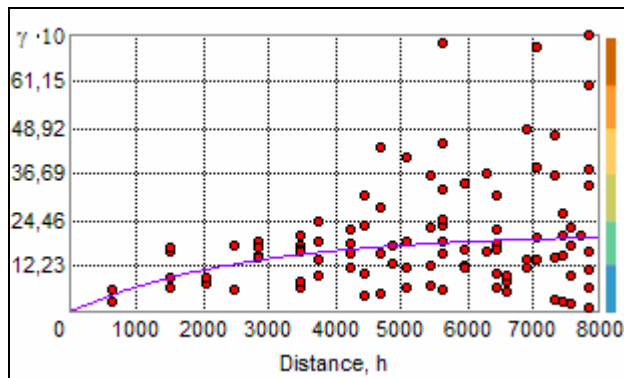
- kruhové
- sférické
- exponenciální
- Gaussovy
- Besselovy

Zvolený typ funkce ovlivňuje predikci neznámých hodnot. Obecně platí, čím lépe funkce vystihuje empirický průběh, tím přesnější je prostorová predikce. Volba modelu závisí na prostorové autokorelaci a na znalosti vlastností zkoumaného jevu. Nejčastěji se používá model sférický nebo exponenciální.

Ve sférickém a exponenciálním modelu se s rostoucí vzdáleností zmenšuje vliv autokorelace. Ve sférickém modelu (obr. 16) vliv autokorelace od určité vzdálenosti prakticky zaniká (hodnota semivariance se ustálí). V exponenciálním modelu (obr. 17) působení autokorelace zaniká až v nekonečnu (autokorelace klesá exponenciálně se vzrůstající vzdáleností).

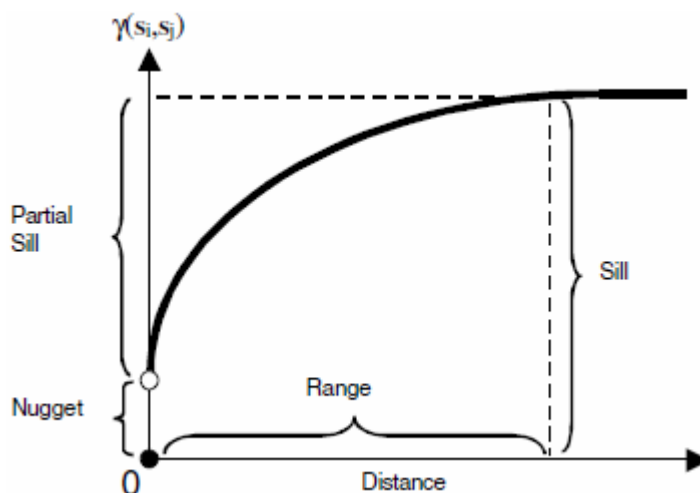


Obr. 16 Sférický model semivariogramu



Obr. 17 Exponenciální model semivariogramu

Vzdálenost, ve které přestává autokorelace působit se nazývá range (obr. 18). Páry měřených bodů oddělené vzdáleností menší než hodnota range jsou prostorově autokorelovány, zatímco místa vzdálenější nejsou. Hodnota autokorelační funkce ve vzdálenosti range se označuje sill. Počáteční hodnota vlivu je označena jako nugget. Rozdíl sill a nugget představuje hodnotu partial sill. Když je hodnota nugget nenulová, znamená to, že opakovaná měření na stejném bodě dávají rozdílné výsledky.



Obr. 18 Model semivariogram

V této fázi je známa prostorová autokorelace, tudíž můžeme přejít k výpočtu predikce. Existují dva typy komponent, které ovlivňují kvalitu prostorové predikce – střední hodnota (trend) a náhodná chyba. V případě konstantní střední hodnoty se jedná o jednoduchou metodu *kriging* (*simple kriging*), kterou lze zapsat ve tvaru:

$$z(s) = \mu + \varepsilon(s),$$

kde $z(s)$ je hodnota funkce sledovaného jevu v bodě s , μ je konstantní střední hodnota a $\varepsilon(s)$ je náhodná chyba v bodě s . Odhad hodnoty funkce v předpovědním místě lze zapsat jako:

$$\hat{z}(s_0) = \sum_{i=1}^n \lambda_i z(s_i),$$

kde $z(s_i)$ je měřená hodnota v bodě i , λ_i je neznámá váha měřené hodnoty v bodě i , n je počet měření. Metoda výpočtu vah vychází z předpokladu, že suma vah je rovna jedné:

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

Proto je nutné dosáhnout toho, aby rozdíl mezi skutečnou hodnotou $z(s_0)$ a předpovídanou hodnotou $\hat{z}(s_0)$ byl minimální:

$$\left(z(s_0) - \sum_{i=1}^n \lambda_i z(s_i) \right)^2 = \min .$$

Řešení této minimalizace lze zapsat takto:

$$K_{ij} \cdot \lambda_i = K_{i0} .$$

A odtud vztah pro výpočet vah měřených bodů:

$$\lambda_i = K_{ij}^{-1} \cdot K_{i0} ,$$

$$\text{kde } \lambda_i = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ m \end{pmatrix}, \quad K_{ij} = \begin{pmatrix} k_{11} & k_{12} & k_{13} & \dots & k_{1n} & 1 \\ k_{21} & k_{22} & k_{23} & \dots & k_{2n} & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ k_{n1} & k_{n2} & k_{n3} & \dots & k_{nn} & 1 \\ 1 & 1 & 1 & \dots & 1 & 0 \end{pmatrix}, \quad K_{i0} = \begin{pmatrix} k_{10} \\ k_{20} \\ \vdots \\ k_{n0} \\ 1 \end{pmatrix} .$$

K_{ij} je matice modelových hodnot semivariogramu mezi všemi páry měřených bodů, λ_i je vektor vah měřených bodů, m je Lagrangeův multiplikátor a K_{i0} je vektor modelových hodnot semivariogramu mezi měřeným bodem a předpovědním místem. Vektor K_{i0} je různý pro každý předpovídaný bod.

V univerzální metodě *kriging* (*universal kriging*) se předpokládá model ve tvaru:

$$z(s) = \mu(s) + \varepsilon(s) ,$$

kde trend $\mu(s)$ je deterministická funkce, zpravidla polynom prvního až třetího řádu.

Soustavu rovnic pak stačí rozšířit o členy popisující trend:

$$\lambda_i = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ m \\ a_0 \\ a_1 \\ a_2 \end{pmatrix}, \quad K_{ij} = \begin{pmatrix} k_{11} & k_{12} & k_{13} & \dots & k_{1n} & 1 & 1 & x_1 & y_1 \\ k_{21} & k_{22} & k_{23} & \dots & k_{2n} & 1 & 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ k_{n1} & k_{n2} & k_{n3} & \dots & k_{nn} & 1 & 1 & x_n & y_n \\ 1 & 1 & 1 & \dots & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & \dots & 1 & 0 & 0 & 0 & 0 \\ x_1 & x_2 & x_3 & \dots & x_n & 0 & 0 & 0 & 0 \\ y_1 & y_2 & y_3 & \dots & y_n & 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_{i0} = \begin{pmatrix} k_{10} \\ k_{20} \\ \vdots \\ k_{n0} \\ 1 \\ 1 \\ x_0 \\ y_0 \end{pmatrix},$$

kde x_i, y_i jsou souřadnice bodu i ; a_1, a_2 jsou koeficienty polynomu prvního řádu.

Po výpočtu vah měřených bodů může být určena predikce míst s neznámou hodnotou:

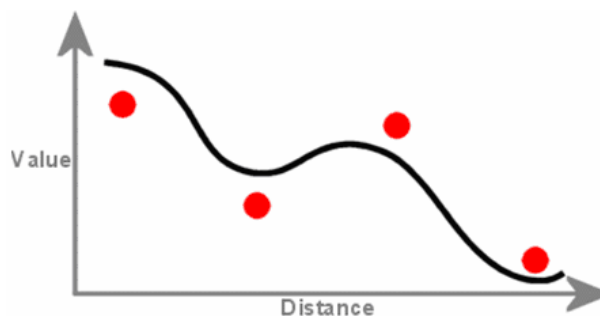
$$\hat{z}(s_0) = \sum_{i=1}^n \lambda_i z(s_i) .$$

Celkové řešení prostorového modelu se skládá ze dvou kroků:

- prostorová autokorelace (výpočet empirického semivariogramu a následné vytvoření modelu semivariogramu)
- prostorová predikce (metodou *kriging* jsou určeny váhy, výpočet prostorového modelu je založen na těchto vahách)

Kvůli těmto dvěma rozdílným krokům jsou měřené hodnoty použity dvakrát, nejprve k odhadu autokorelace, poté k určení predikce.

Metoda *kriging* je výpočtově jedna z nejsložitějších. Pokud je zvolen model semivariogramu, který má nenulovou hodnotu nugget, exaktní metoda *kriging* se mění na aproximující (obr. 19). Výsledný model přesahuje minimální a maximální hodnoty měřených bodů.

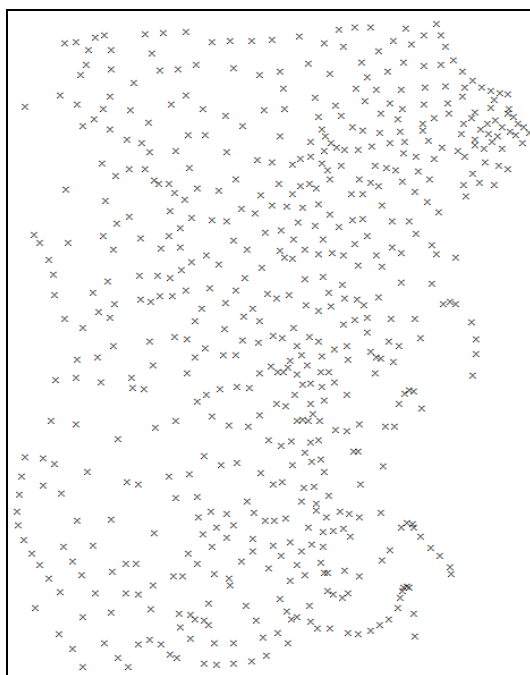


Obr. 19 *Kriging*

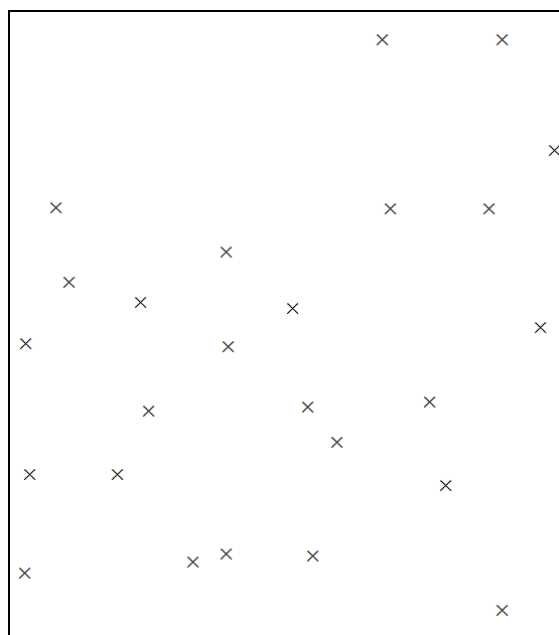
3 Interpolace bodových dat v ArcGIS

Praktické použití interpolace bodových dat bude ukázáno v programu ArcGIS (verze 9.2) firmy ESRI. Interpolace je možné provádět v nadstavbách programu ArcGIS: 3D Analyst, Spatial Analyst a Geostatistical Analyst. Nabídka a možnosti nastavení interpolačních metod v nadstavbě 3D Analyst a Spatial Analyst jsou zcela shodné. Tyto nadstavby nabízejí použití *metody přirozeného souseda*, *IDW*, *spline*, *trend*, *kriging* a *topo to raster*. Metoda *topo to raster* je speciální metoda programu ArcGIS. Nadstavba Geostatistical Analyst umožňuje použití *IDW*, *spline*, *trend* a *kriging*. V tomto případě je zde možnost nastavení více parametrů než ve zbývajících dvou nadstavbách.

Data pro praktickou ukázkou interpolačních metod byla zvolena následovně. Prvním vzorkem jsou nadmořské výšky bodů (obr. 20) na poměrně malém území (cca 0,3 km²). Tento soubor tedy reprezentuje data s vysokou hustotou, bodů je téměř 600. Maximální převýšení mezi body tohoto souboru je asi 70 m, nadmořské výšky bodů jsou poměrně rozdílné. Druhým vzorkem jsou měřené teploty vzduchu na bodech pokrývajících plochu 70 km² (obr. 21). Tento soubor představuje rozptýlená data, obsahuje pouze 25 bodů. Interval měřených teplot vzduchu je 4,5°C. Vybrány tedy byly alespoň tyto dvě datové sady s různými vlastnostmi. Důvodem je, že jedna metoda dosahuje rozdílných výsledků v závislosti na různém typu měřených dat. Jedna metoda vykazuje lepší výsledky pro určitý typ vstupních dat než metoda jiná atd.



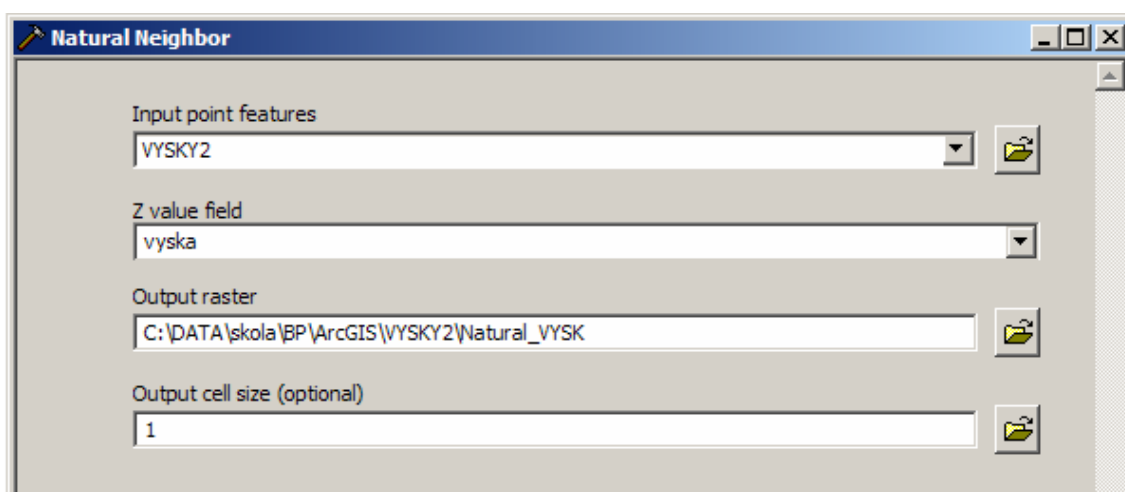
Obr. 20 Nadmořské výšky



Obr. 21 Teploty vzduchu

3.1 Metoda přirozeného souseda (natural neighbour)

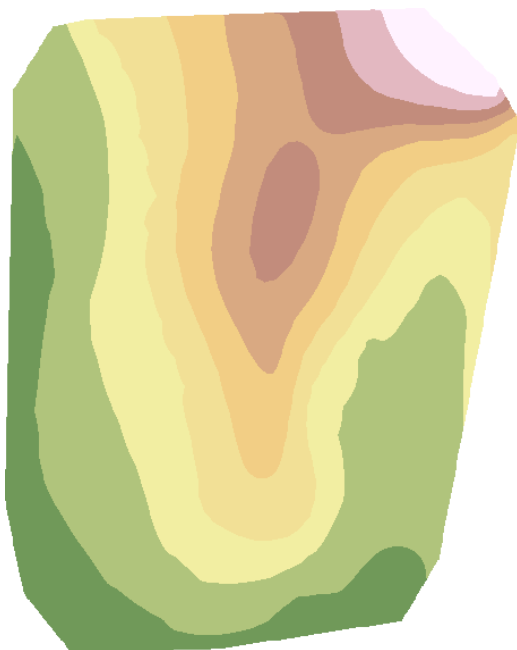
Metoda přirozeného souseda nemá žádné parametry, kterými by se dal ovlivnit její výsledek. Počet sousedů je dán rozmístěním bodů v ploše. Před výpočtem se dá zvolit pouze velikost pixelu výsledného rastru (cell size). Ta závisí na požadovaném rozlišení rastru, které se odvíjí od velikosti interpolovaného území. Standardní nastavení (default) je stanoveno na hodnotu vypočtenou z kratší strany obdélníku, který ohraničuje plochu vstupních bodů ve vstupním souřadnicovém systému. Délka této kratší strany je vydělena ještě číslem 250.



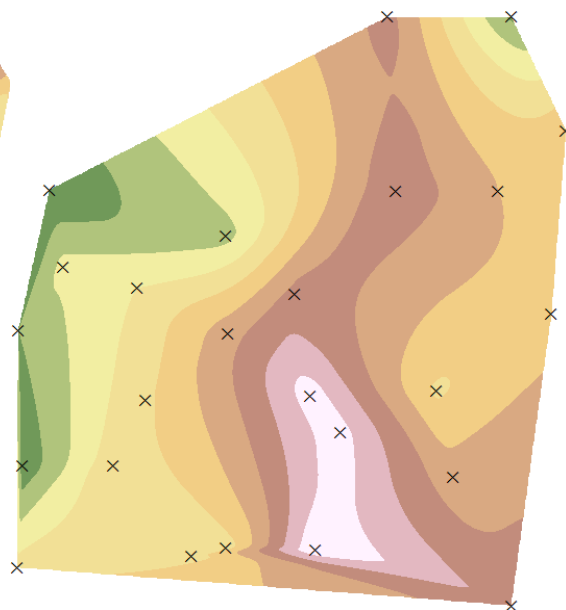
Obr. 22 Metoda přirozeného souseda v ArcGIS Spatial Analyst

Výsledný odhad této metody nezahrnuje extrapolované oblasti, jako je tomu u ostatních metod. Aby *Thiessenovy polygony* na obvodu zájmového území netvořily nekonečnou plochu, okolo měřených bodů se vytvoří konvexní n-úhelník ohraničující všechny body. Interpolace poté proběhne pouze v rámci této plochy. Výstupem ostatních metod je jinak vždy rastr obdélníkového tvaru obsahující i extrapolované oblasti.

Při použití *metody přirozeného souseda* na oba druhy vstupních dat jsou výsledkem hladké odhady (obr. 23 a obr. 24).



Obr. 23 *Metoda přirozeného souseda*



Obr. 24 *Metoda přirozeného souseda*

Tato metoda se používá pro vstupní data různé povahy (hustá, řídká, rovnoměrná, rozptýlená). Je obecnou metodou pro různé druhy modelovaných jevů. Výsledná struktura odhadu je spojitá a vyhlazená.

3.2 Inverse distance weighting (IDW)

Výsledek interpolace metodou *IDW* může být ovlivněn nastavením následujících parametrů:

Mocnina (power)

Mocnina představuje hodnotu parametru p . Čím vyšší mocnina, tím vyšší váha bude přiřazena bližším bodům a výsledná struktura bude mít více detailů, což znamená, že bude méně hladká. Nastavení nižší mocniny přikládá větší význam vzdálenějším bodům a výsledný odhad bude hladší. Hodnota mocniny se volí od 0,5 do 3. Nejčastěji je používána mocnina 2, která je standardním nastavením (default).

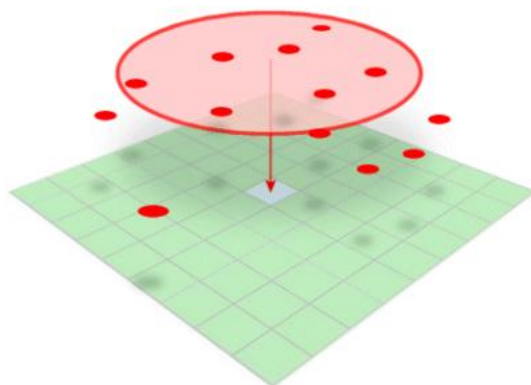
Poloměr (search radius)

Nastavení poloměru omezuje počet vstupních bodů, ze kterých je počítán odhad neznámého bodu. Poloměr může být pevný (fixed) nebo proměnný (variable).

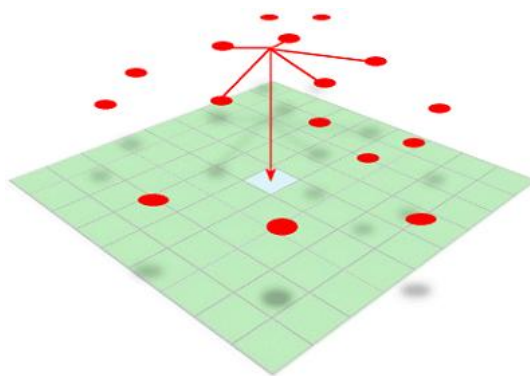
Při nastavení pevného poloměru (obr. 25) je požadováno zadání vzdálenosti a minimálního počtu vstupních bodů. Vzdálenost určuje poloměr kruhu kolem určovaného bodu, všechny měřené body v tomto kruhu budou použity pro výpočet odhadovaného bodu. V případě, že je zde méně bodů než zadaný minimální počet,

poloměr kruhu se bude zvětšovat, dokud nedosáhne stanoveného minimálního počtu měřených bodů. Zadání minimálního počtu vstupních bodů však není povinné. Je důležité si uvědomit, že pevný poloměr bude použit pro každé místo odhadu, takže když měřené body nebudou rozmístěny rovnoměrně, pro každý odhad bude pravděpodobně použit různý počet vstupních bodů. Pevný poloměr má standardní nastavení vzdálenosti (default) rovno pětinásobku velikosti pixelu výstupního rastru.

U proměnného poloměru (obr. 26) je počet vstupních bodů určen, což se projeví různou velikostí poloměru pro každý interpolovaný bod. Poloměr kruhu tedy bude záviset na hustotě měřených bodů kolem určovaného bodu. Dále může být definována maximální vzdálenost, kterou poloměr nesmí přesáhnout. V případě, že hodnota poloměru dosáhne maximální vzdálenosti před docílením zadaného počtu vstupních bodů, odhad v tomto místě bude proveden na základě měřených bodů v maximální vzdálenosti. Proměnný poloměr je standardním nastavením (default) s počtem vstupních bodů 12.



Obr. 25 Pevný poloměr

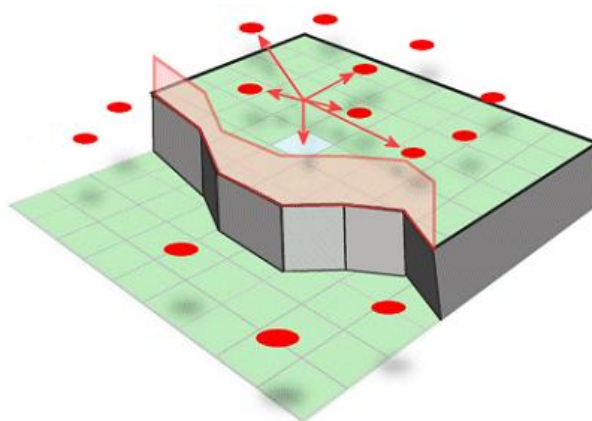


Obr. 26 Proměnný poloměr

Pro výběr poloměru platí obecné pravidlo. Pokud je k dispozici velký počet vstupních bodů pravidelně rozmístěných, potom je vhodné použít fixní poloměr. V opačném případě, když je měřených bodů málo a jsou náhodně rozmístěny, je doporučena volba variabilního poloměru.

Bariéra (input barrier polyline feature)

Bariéra je liniová vrstva fungující jako zlom, který omezuje prostor, ve kterém jsou hledány vstupní body. Linie může vyjadřovat například hranu či útes v terénu, nebo kterékoliv jiné přerušení daného jevu. Do výpočtu interpolovaného bodu budou zahrnuty pouze body na té straně bariéry, kde se tento určovaný bod nachází (obr. 27).



Obr. 27 Bariéra

Input point features
VYSKY2

Z value field
vyska

Output raster
C:\DATA\skola\BP\ArcGIS\VYSKY2\Idw_VYSKY

Output cell size (optional)
1

Power (optional)
2

Search radius (optional)
Variable
Fixed
Variable

Number of points:
12

Maximum distance:

Input barrier polyline features (optional)

Obr. 28 Nastavení parametrů *IDW* v ArcGIS Spatial Analyst

Nadstavba Geostatistical Analyst poskytuje oproti Spatial Analyst nastavení směrového vlivu na výpočet vah, sektoru vstupních bodů a vyhlazení.

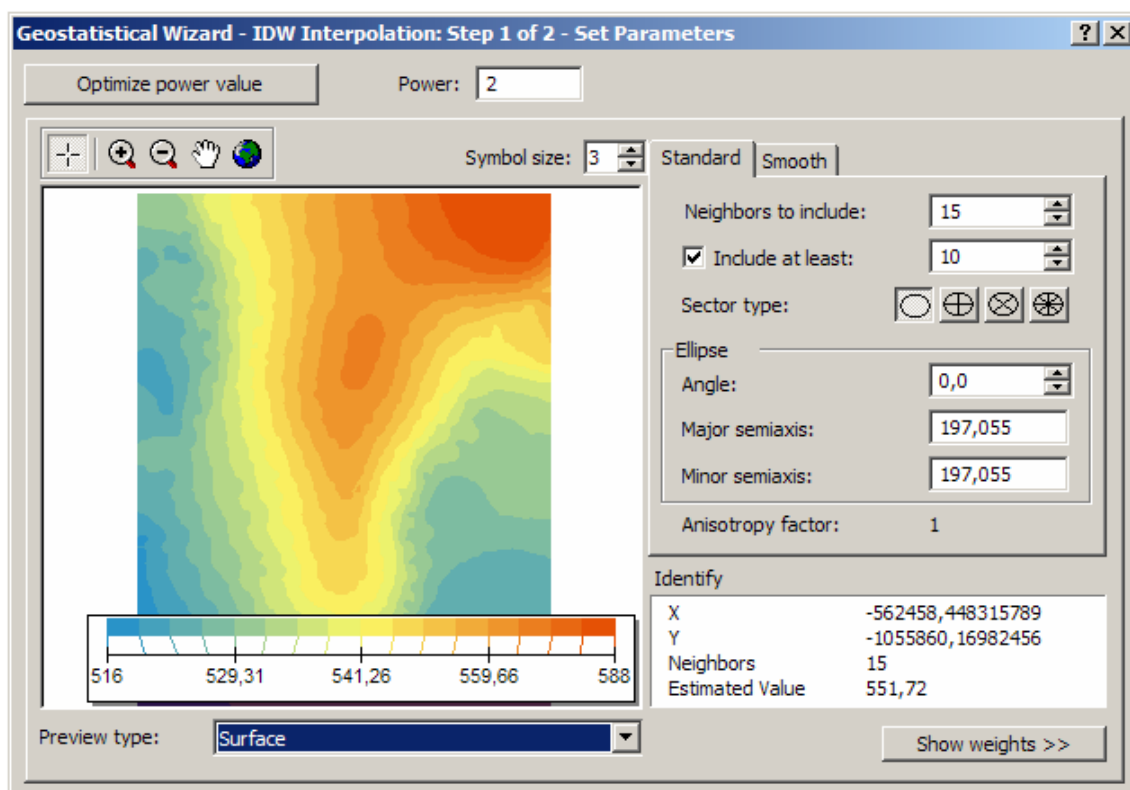
Geostatistical Analyst umožňuje dva typy nastavení, buď standardní nebo vyhlazovací (smooth). U obou typů nastavení je možné zahrnout směrové působení na výpočet vah. Pokud neexistuje žádný směrový vliv, potom uvažujeme body rovnoměrně ve všech směrech a zvolíme tvar okolí určovaného bodu jako kruh s určitým

poloměrem. V případě, že existuje nějaké směrové působení například převládající vítr, změníme tvar okolí na elipsu s hlavní poloosou rovnoběžnou se směrem větru. Okolí určovaného bodu tedy může být upravováno natočením hlavní poloosy a velikostí obou poloos.

U standardního nastavení se kromě maximálního a minimálního počtu vstupních bodů může volit i sektor, ze kterého budou tyto body vybírány. Když rozdělíme okolí určovaného bodu na sektory, minimální a maximální omezení počtu bodů bude aplikováno na každý sektor. Různé typy sektorů (jeden, čtyři nebo osm) jsou ukázány na obr. 29 (sector type).

Při použití nastavení s vyhlazením se míra vyhlazení volí v intervalu od 0 do 1. Standardním nastavením (default) je hodnota 0,5.

Výhodou prostředí ArcGIS Geostatistical Analyst je okno s náhledem výsledné interpolace, díky kterému lze pozorovat, jak se rozdílná volba jednotlivých parametrů projeví na výsledném odhadu.



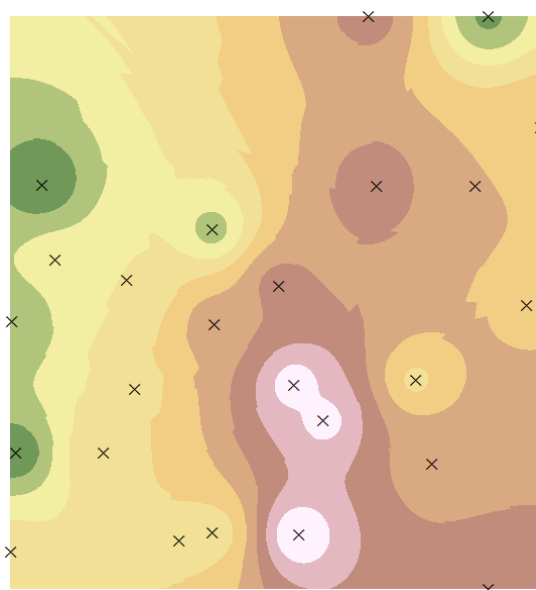
Obr. 29 Nastavení parametrů *IDW* v ArcGIS Geostatistical Analyst

V případě použití metody *IDW* na rozptýlená data teplot (obr. 30) se zřetelně projevil bull eyes efekt v důsledku silného vlivu vstupních bodů na své okolí. Na druhém, hustém souboru vstupních dat se tento vliv projevil pouze vlnkami v jeho výsledné struktuře – nejméně při volbě nízké hodnoty mocniny p (obr. 31). Při

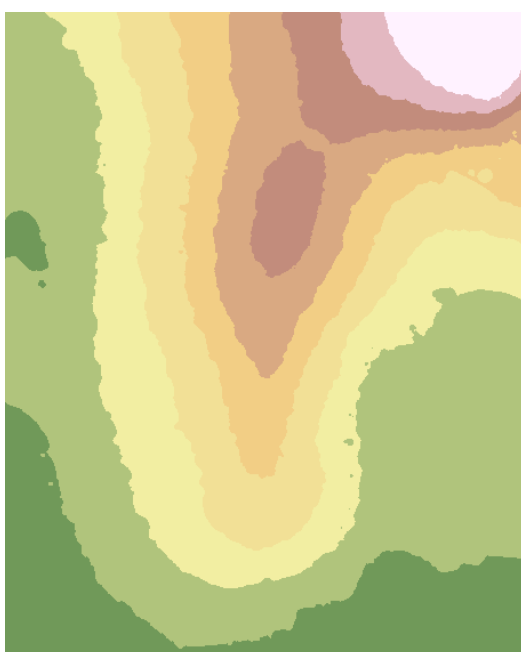
zvyšování mocniny se tento vliv stává výraznějším (obr. 32). Jak je vidět z obrázků, lepší výsledky tato metoda vykazuje na husté síti bodů. Nejlepšího výsledku by se dalo dosáhnout v případě rozmístění bodů v zájmové oblasti stejnoměrně.

Při volbě okolí, které bude mít na odhad vliv, byla zvolena varianta proměnného poloměru pro oba soubory. U obou z nich totiž nejsou body pravidelně rozprostřené v ploše.

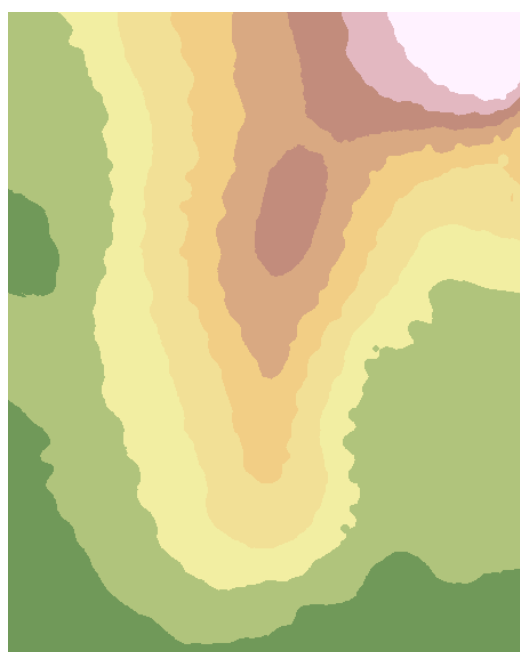
Interpolované hodnoty obou souborů nepřekročily interval vstupních hodnot. To je adekvátní tomu, že *IDW* je metodou váženého průměru.



Obr. 30 *IDW*, parametr $p = 2$



Obr. 31 *IDW*, parametr $p = 1$



Obr. 32 *IDW*, parametr $p = 2$

Metodu *IDW* je vhodné použít, pokud jsou vstupní body husté a rovnoměrně rozmístěné. Nehodí se pro data, která vytváří shluky. Při nevhodné struktuře dat dochází k tvorbě rušivých bull eyes. Výhodou je možnost ovlivnění počtu bodů a tvaru okolí, který bude mít vliv na výsledný odhad. Nevýhodou je, že metoda neumí vypočítat hodnoty nad maximum a pod minimum měřených dat. Proto není žádoucí ji použít, pokud nejsou měřeny body v místech s extrémními hodnotami.

Metoda *IDW* je jednoduchá na výpočet, hodně se používá, například pro interpolaci nadmořských výšek nebo meteorologických jevů, jako jsou srážky nebo teploty.

3.3 Spline

Výsledek interpolace *spline* lze ovlivnit nastavením následujících parametrů:

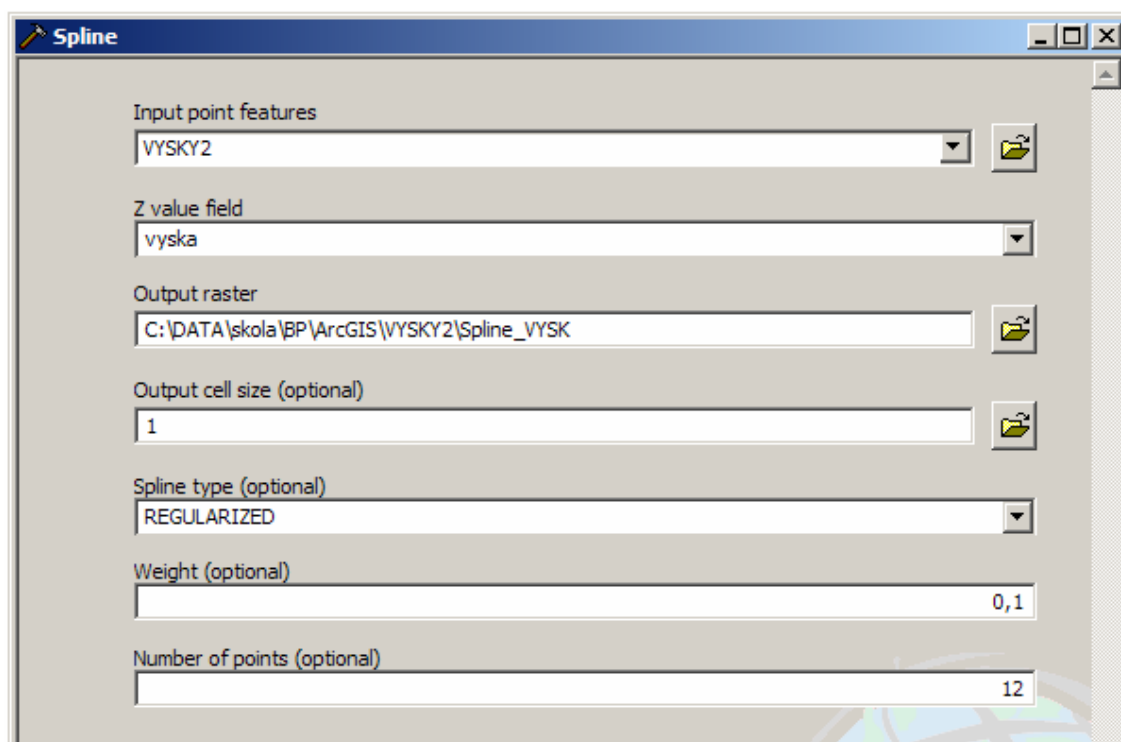
Váha (weight)

Metoda *regularized spline* – parametr váhy představuje váhu třetí derivace povrchu při minimalizaci křivosti. Váha tedy ovlivňuje křivost. Čím vyšší váha, tím vzniká menší křivost a výsledná struktura je hladší. Hodnota tohoto parametru musí být větší nebo rovna nule. Obvykle užívané hodnoty jsou 0; 0,001; 0,01; 0,1 a 0,5. Standardním nastavením (default) je hodnota 0,1.

Metoda *tension spline* – parametr váhy definuje váhu první derivace povrchu při minimalizaci křivosti. Parametr váhy představuje tenzi. Čím vyšší váha, tím hrubší je struktura modelovaného jevu a odhadované hodnoty se více přimykají intervalu vstupních dat. Hodnota váhy musí být opět větší nebo rovna nule. Nejčastěji používané hodnoty jsou 0, 1, 5 a 10.

Počet bodů (number of points)

Tímto zvolíme počet bodů, který bude použit pro výpočet odhadu každého určovaného bodu. Čím více bodů do výpočtu zahrneme, tím větší vliv budou mít vzdálenějšími body na interpolované body a výsledná struktura bude hladší. Zároveň ale s narůstajícím počtem bodů roste i délka výpočtu rastru.



Obr. 33 Nastavení parametrů *spline* v ArcGIS Spatial Analyst

Aplikace metody *spline* na vzorek nadmořských výšek neposkytuje optimální výsledky. Důvodem je struktura dat. Hodnoty nadmořských výšek bodů jsou poměrně rozdílné a v některých místech hodně blízko u sebe.

Použití *regularized spline* zaručuje vytvoření hladkého povrchu. Rozdíly ve výsledných rastroch při různém použití váhového parametru jsou vidět na obr. 34 a obr. 35. Při vyšším parametru došlo k většímu vyhlazení, které se nepříznivě projevilo v extrapolovaných oblastech. Typickým znakem *spline* je, že interpolované hodnoty leží mimo interval vstupních dat. Při porovnání maximální vstupní hodnoty nadmořské výšky s maximální vypočtenou hodnotou vypočtená výška převyšuje vstupní téměř o dvacet metrů. Tento rapidní rozdíl je samozřejmě způsoben tím, že maximální hodnota odhadu leží v extrapolované oblasti.

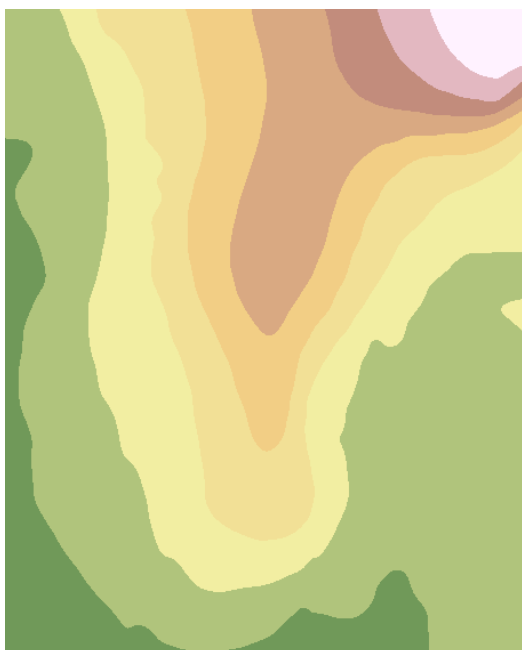
Z porovnání rastrů *tension spline* na obr. 37 a obr. 38 je vidět, že rastr s menší váhou je vyhlazenější. *Tension spline* vytváří odhad s hodnotami více se blížícími rozmezí vstupních dat. Toto platí nejvíce při vysoké hodnotě váhy. U váhy rovné deseti vypočtená výška převyšuje vstupní výšku o tři metry (v extrapolované části). To je oproti *regularized spline* velký pokles.



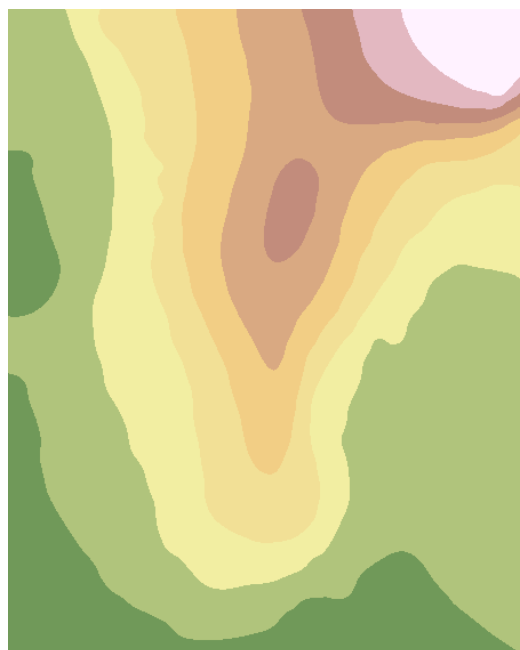
Obr. 34 *Regularized spline*, váha 0,001



Obr. 35 *Regularized spline*, váha 0,5



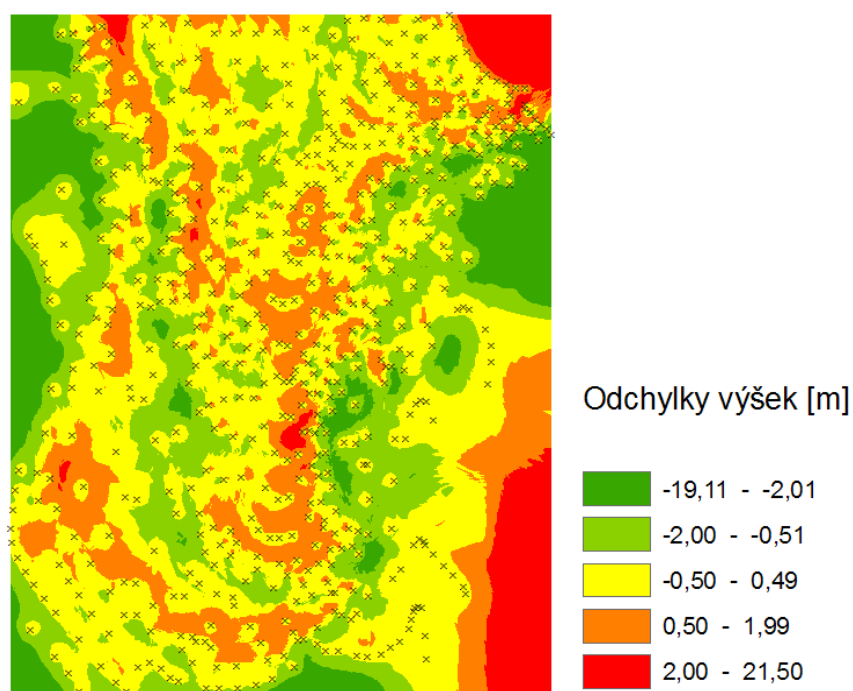
Obr. 37 *Tension spline*, váha 0,01



Obr. 38 *Tension spline*, váha 10

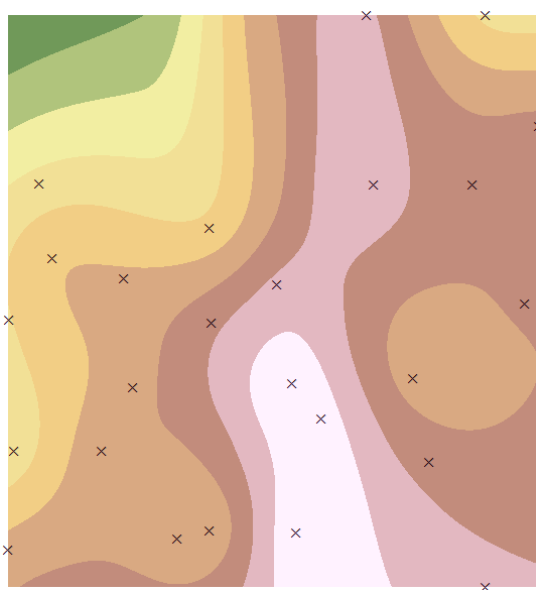
Následující obrázek (obr. 39) poskytuje porovnání metod *spline* a *IDW* u souboru nadmořských výšek. Rastr *IDW* s exponentem 2 byl odečten od rastru *regularized spline* s váhou 0,001. Výsledný rastr tedy zobrazuje odchylky odhadovaných výšek těchto dvou metod. Protože metoda *spline* není vhodná pro interpolaci tohoto souboru, odchylky jsou na mnoha místech vysoké. V místě měřených bodů nejsou odchylky žádné, protože obě metody jsou exaktními interpolátory. Většina

extrémních hodnot odchylek leží v extrapolovaných oblastech, kde se razantní rozdíly dají očekávat.

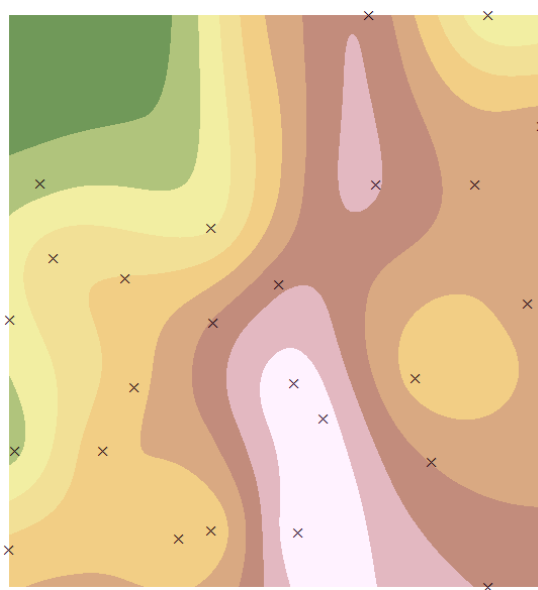


Obr. 39 Odchylky výšek *spline* a *IDW*

Použití *spline* na vzorek měřených teplot vzduchu vykazuje dobré výsledky. Teplota vzduchu je veličina, která se mění pozvolna, tato metoda je tedy pro ni optimální. Když porovnáme *regularized spline* s váhou 0,001 a *tension spline* s váhou 1 (obr. 40 a obr. 41), oba typy dosahují podobných výsledků, povrch *regularized spline* je hladší a *tension spline* se zas o trochu více přimyká rozsahu vstupních teplot.

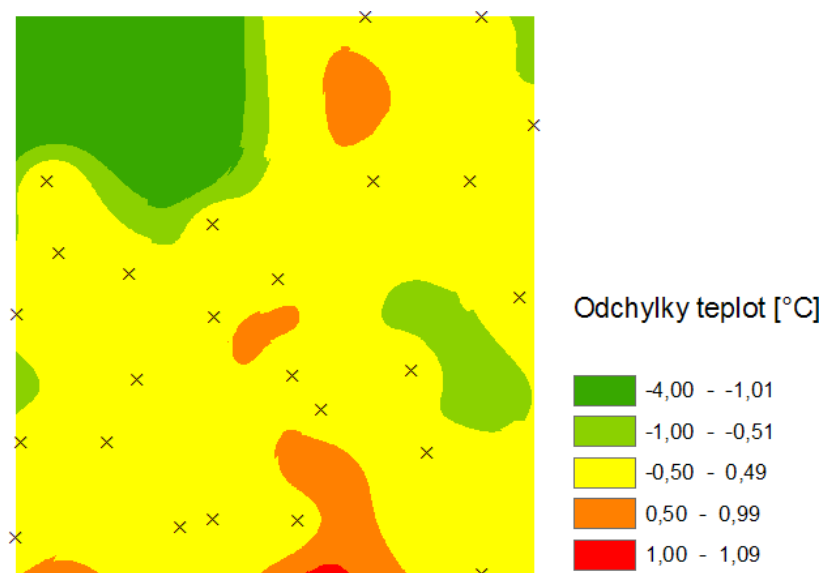


Obr. 40 *Regularized spline*, váha 0,001



Obr. 41 *Tension spline*, váha 1

Rastr na obr. 42 zobrazuje odchylky odhadovaných teplot metody *spline* a *IDW*. Vznikl tak, že od rastru *regularized spline* s váhou 0,001 byl odečten rastr *IDW* s parametrem 2. Kromě předpokládaných odchylek v extrapolovaných částech se objevily větší odchylky ještě v místech, které jsou dále od okolí měřených bodů. Zde se více projeví rozdíl toho, že *IDW* pracuje s váženým průměrem okolních bodů a *spline* se řídí podmínkou minimální křivosti.



Obr. 42 Odchylky teplot *spline* a *IDW*

Metoda *spline* je vhodná pro modelování pozvolna se měnících jevů, například teploty, výšky vodní hladiny, znečištění ovzduší nebo nadmořské výšky v málo členitém terénu. *Spline* není vhodnou metodou, když jsou vstupní body blízko u sebe a mají hodně rozdílné hodnoty, vede to k velkým oscilacím splinu.

Regularized spline vytváří hladký, pozvolna se měnící povrch převyšující rozsah vstupních dat. *Tension spline* vytváří méně hladké povrchy, více se však blíží vstupnímu rozsahu hodnot.

3.4 Trend

Nadstavba ArcGIS Spatial Analyst poskytuje globální odhad trendu v celém zájmovém území. Nastavení metody *trend* vyžaduje následující specifikace:

Stupeň polynomu (polynomial order)

Lineární polynom vyjadřuje rovinu, kvadratický polynom představuje plochu s jedním ohybem, dva ohyby popisuje kubický polynom a tak dále. Výběr stupně polynomu usnadňuje obecné pravidlo: jakýkoliv příčný řez polynomicke funkce řádu p může mít nejvýše $p - 1$ střídajících se maxim a minim. Stupeň polynomu musí být celé

číslo v intervalu od 1 do 12. Nejčastěji se používají hodnoty 1, 2 nebo 3. Standardním nastavením (default) je stupeň 1.

Typ regrese (type of regression)

V nabídce jsou dva druhy regrese – linear a logistic. Typ linear nastaví vybraný polynom na měřené body pomocí metody nejmenších čtverců.

Typ logistic je vhodný pro odhad pravděpodobnosti výskytu určitého jevu. Vstupní hodnoty mohou nabývat pouze dvou hodnot, 0 nebo 1. Tyto hodnoty mohou například vyjadřovat existenci ohroženého druhu na určitém místě. Výstupem je potom rastr pravděpodobnosti, který je spojitý. Hodnoty tohoto rastru leží v intervalu od 0 do 1.

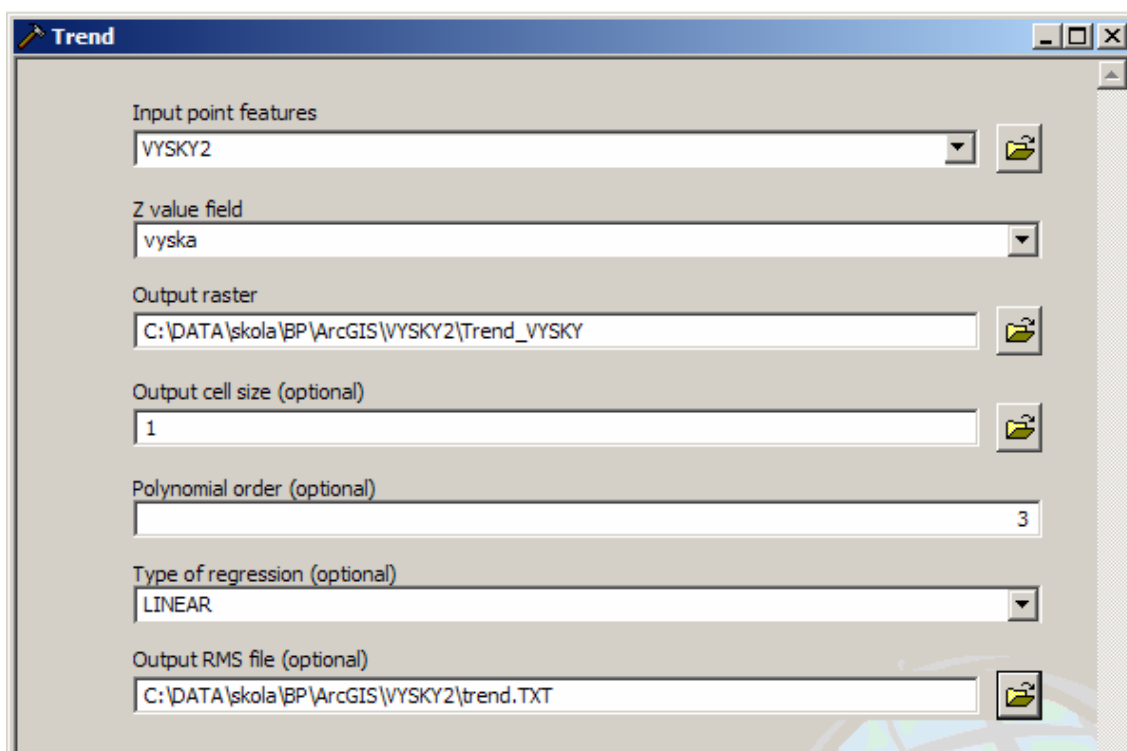
Výstupní soubor pro RMS (output RMS file)

Jedná se o soubor, do kterého bude uložena střední kvadratická chyba (RMS error), která vznikne porovnáním vstupních hodnot s interpolovanými hodnotami v místech vstupních bodů. Tato chyba může mít rozhodující vliv při volbě stupně polynomu. Vhodný stupeň polynomu je takový, u kterého je dosaženo nejmenší střední kvadratické chyby.

Dále tento soubor obsahuje hodnoty koeficientů polynomu. Ukázka souboru je na následujících řádcích:

```
coef #          coef
-----
0 -12437396,8735282
1 -18,7855715479551
2 -25,2687617034003
3 -2,62343930487412e-005
4 -7,48100817287057e-006
5 -2,1920787356535e-005
6 -1,25757264156586e-011
7 -4,61228389565324e-012
8 -1,08633939941133e-012
9 -6,72147611008278e-012
-----
RMS Error   = 0,758527074425125
Chi-Square  = 14,3840830919873
```

Jedná se o výstupní charakteristiky rastru teplot vzduchu. Podle počtu koeficientů je zřejmé, že byl zvolen polynom třetího stupně.



Obr. 43 Nastavení metody *trend* v ArcGIS Spatial Analyst

Aplikací metody *trend* na soubor nadmořských výšek zjistíme hlavní charakteristiku povrchu. Z výsledných rastrů (obr. 44 - 46) je vidět, že terén stoupá směrem k severovýchodu. Pro srovnání byl volen stupeň polynomu postupně 1, 2 a 3. Kvadratický a kubický *trend* poskytují téměř totožné výsledky. Potvrzují to i výsledné hodnoty střední kvadratické chyby obou interpolací, které se liší jen nepatrně.



Obr. 44 Lineární *trend*



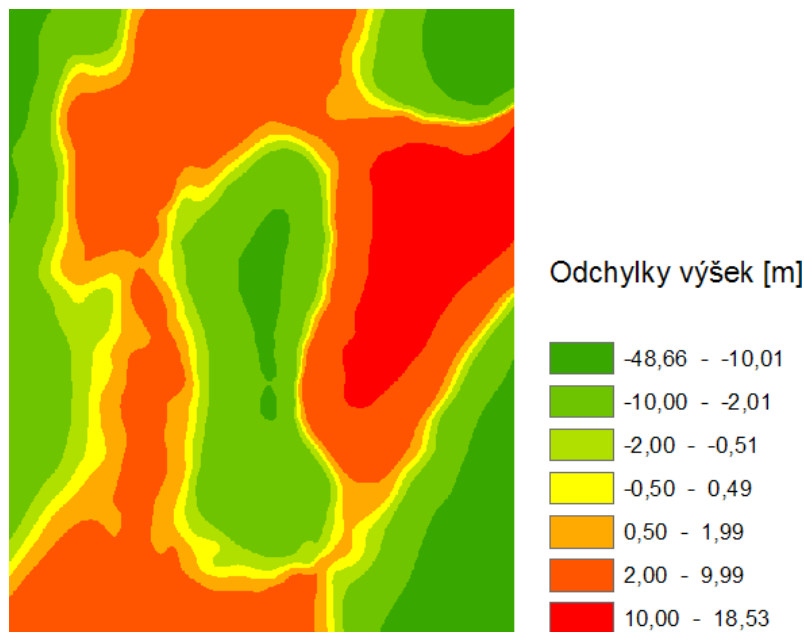
Obr. 45 Kvadratický *trend*



Obr. 46 Kubický *trend*

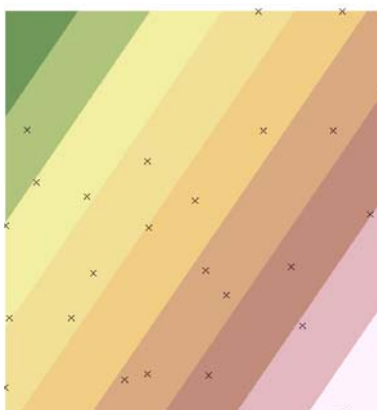
Pro porovnání metody *trend* byla vybrána metoda *spline*, která také používá pro interpolaci polynomické funkce a vytváří hladké povrchy. Od rastru kubického *trendu* byl odečten rastr *regularized spline* s parametrem váhy 0,001 (obr. 47). Vypočtené

odchylky nadmořských výšek dosahují vysokých hodnot. Pásmo s malými odchylkami je velice úzké, vzniklo v okolí průniku trendového a splinového povrchu. Tato pásma vznikla dvě, když se trendový povrch dostal nad splinový a pak když se *trend* opět dostal pod *spline*.



Obr. 47 Odchylky výšek *trend* a *spline*

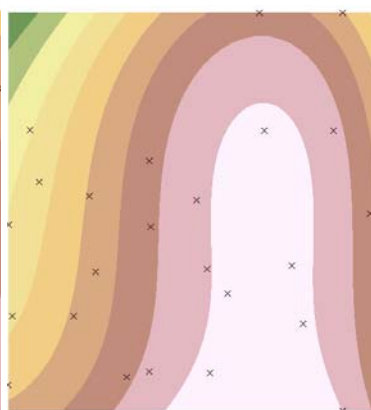
Dále byla metoda vyzkoušena na odhad trendu teploty vzduchu. Na obr. 48 - 50 jsou použity různé typy *trendu*. Charakter teploty nejvíce vystihuje kubický *trend*, který dosahuje i nejmenší hodnoty střední kvadratické chyby. Tento kubický *trend* byl porovnán s *regularized spline* s váhou 0,001. Výsledný rozdíl mezi *trend* a *spline* je znázorněn na obr. 51. *Trend* odstraňuje z odhadu lokální detaily a v těchto místech se tedy objevily větší odchylky teplot. Dále jsou hodnoty odchylek vyšší v okrajových extrapolovaných oblastech, kde *trend* obecně nevykazuje dobré výsledky.



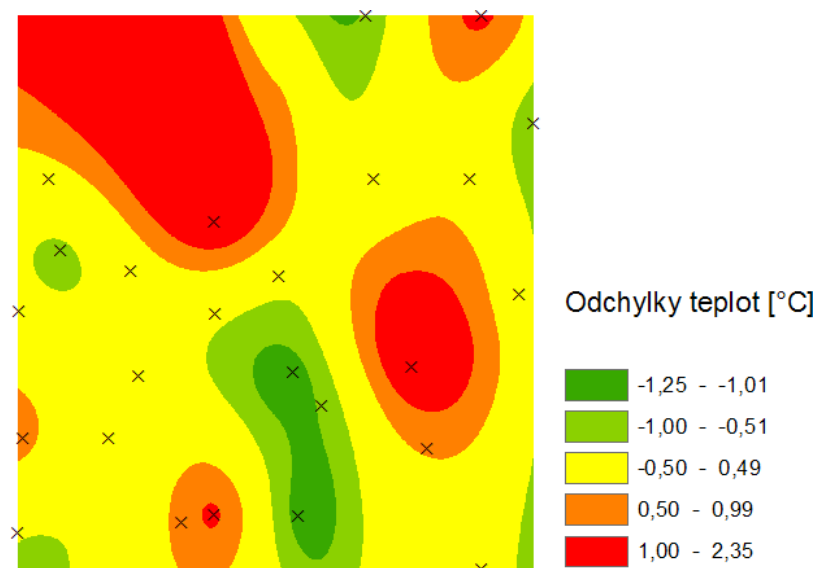
Obr. 48 Lineární *trend*



Obr. 49 Kvadratický *trend*



Obr. 50 Kubický *trend*



Obr. 51 Odchylky teplot *trend* a *spline*

Model vytvořený metodou *globálního trendu* je hladký, pozvolna se mění, vyhlazuje detaily a zachycuje hlavní charakteristiku určitého jevu. Zřídka prochází přesně měřenými body, k největším odchylkám od těchto bodů dochází na kraji interpolovaného území.

Metodu je vhodné použít při hrubém odhadu prostorového trendu. Typickým příkladem je sada bodů nadmořských výšek, ze kterých zjistíme, zda území má například stoupající nebo klesající charakter. Výsledné povrchy jsou také používány k popisu a odstranění hrubých rysů datových sad.

V nadstavbě ArcGIS Geostatistical Analyst je výše popsaná metoda *globálního trendu* označena jako globální polynomická interpolace. Mimo ní je v této nadstavbě možnost výběru i lokální polynomické interpolace, neboli *lokálního trendu*. *Globální trend* modeluje celé zájmové území pomocí jednoho polynomu, zatímco *lokální trend* používá více polynomů, každý pro určitou malou část území. Ve výsledném odhadu je pak vystiženo více lokálních detailů a dosaženo také vyšší přesnosti.

3.5 Kriging

Metoda *kriging* bude ukázána v prostředí nadstavby Geostatistical Analyst. Nadstavba Spatial Analyst také umožňuje použití této metody, ale s určitou omezenou volbou parametrů. ArcGIS Geostatistical Analyst je v tomto případě vhodnější hlavně z důvodu grafického zobrazení vybraných parametrů, které jsou aplikovány na soubor vstupních dat. Vzhledem ke složitosti metody je možnost tohoto zobrazení téměř nepostradatelná a usnadňuje výběr správného nastavení. Tato nadstavba disponuje i

dalšími funkcemi mimo rámec nástrojů pro interpolaci, které mohou být použity k prozkoumání vlastností vstupních dat před samotnou interpolací. Použití těchto funkcí (explore data) ale nebude dále specifikováno z důvodu velkého rozsahu tématu.

Prvním krokem je volba typu metody *kriging*. Jak bylo uvedeno dříve, druhy *kriging* se od sebe liší typem střední hodnoty neboli trendu μ . U *simple kriging* je μ známou konstantou, u *ordinary kriging* je μ neznámou konstantou. V případě *universal kriging* je trend $\mu(s)$ deterministická funkce, většinou polynom prvního až třetího řádu. Trend může být z datové sady na začátku analýzy dat odstraněn a pak přidán zpět před tvorbou výsledného odhadu. Nejčastěji používaným a standardním nastavením (default) je *ordinary kriging*.

Část, které se musí během nastavení věnovat náležitá pozornost, je výběr modelu semivariogramu. Cílem je určit nejlepší model, který prochází body empirického semivariogramu.

Velikost intervalu a počet intervalů empirického semivariogramu

Velikost intervalu (lag size) a počet intervalů (number of lags) pro seskupení hodnot (binning) se nastaví automaticky na určité hodnoty. Tyto hodnoty lze upravit dle struktury vstupních dat. Mělo by být dodrženo pravidlo, že násobek velikosti intervalu a počtu intervalů by měl být menší než polovina největší délky v datové sadě. Zmenšení velikosti intervalu znamená zaměření na modelování detailů mezi sousedními měřenými body. Při menší velikosti intervalu model semivariogramu roste strměji, pak se ustálí. Vzdálenost, ve které se ustálí, je hodnota range, za kterou už nedochází k autokorelaci. Pokud by byla velikost intervalu příliš malá, mohly by existovat intervaly bez zastoupení a empirický semivariogram by byl zkreslen. Příliš velké intervaly vliv autokorelace mohou skrýt.

Model semivariogramu

Na základě standardního nastavení (default) se empirickým semivariogramem proloží sférický model. Automaticky se vypočtou optimální hodnoty jeho charakteristik range, nugget a partial sill. Kromě základního sférického modelu je k dispozici celá řada modelů, z nichž dalším nejčastěji používaným je exponenciální model. Je zde i možnost kombinace až tří modelů dohromady. Při výběru mezi mnoha podobnými modely může pomoci porovnání výsledků cross-validation, která bude popsána později.

Anizotropie

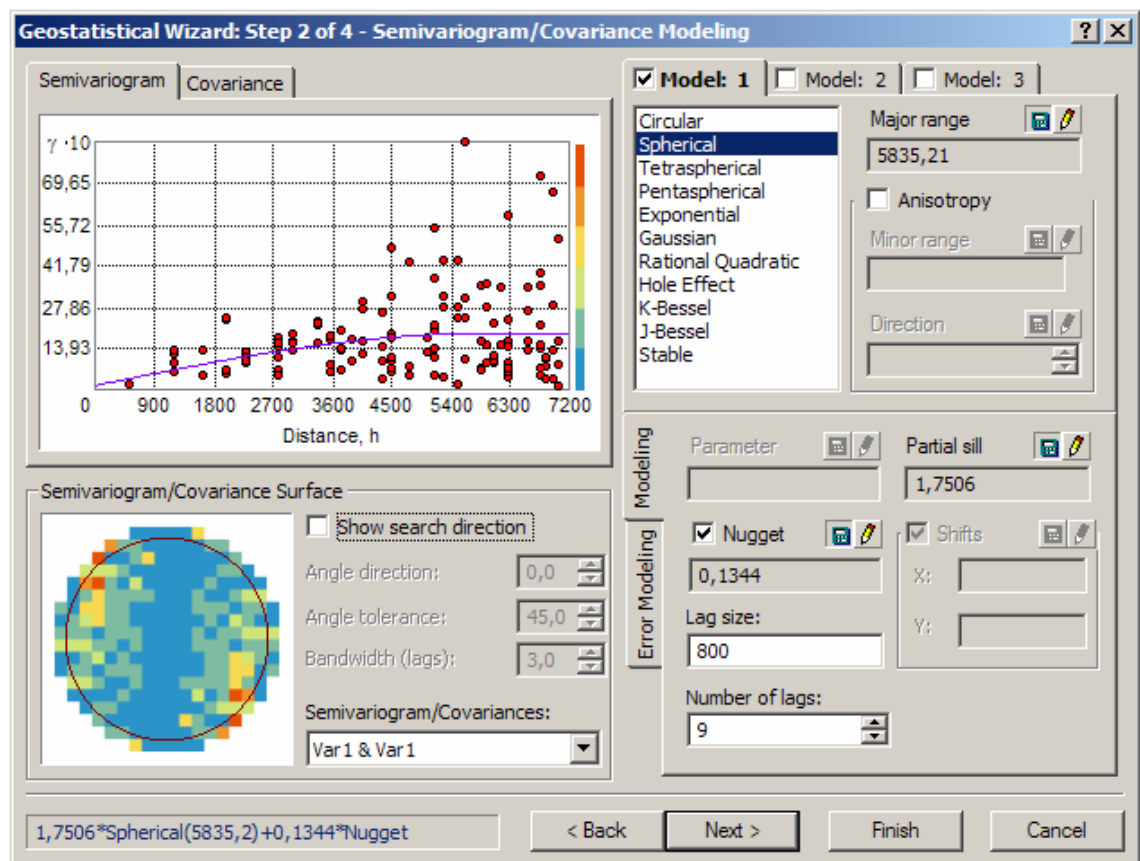
V určitých směrech si jevy mohou být více podobné než v jiných směrech. Směrové působení se nazývá anizotropie. Může být způsobeno větrem, proudem,

geologickou strukturou atd. Toto směrové působení může být zahrnuto do tvorby modelu.

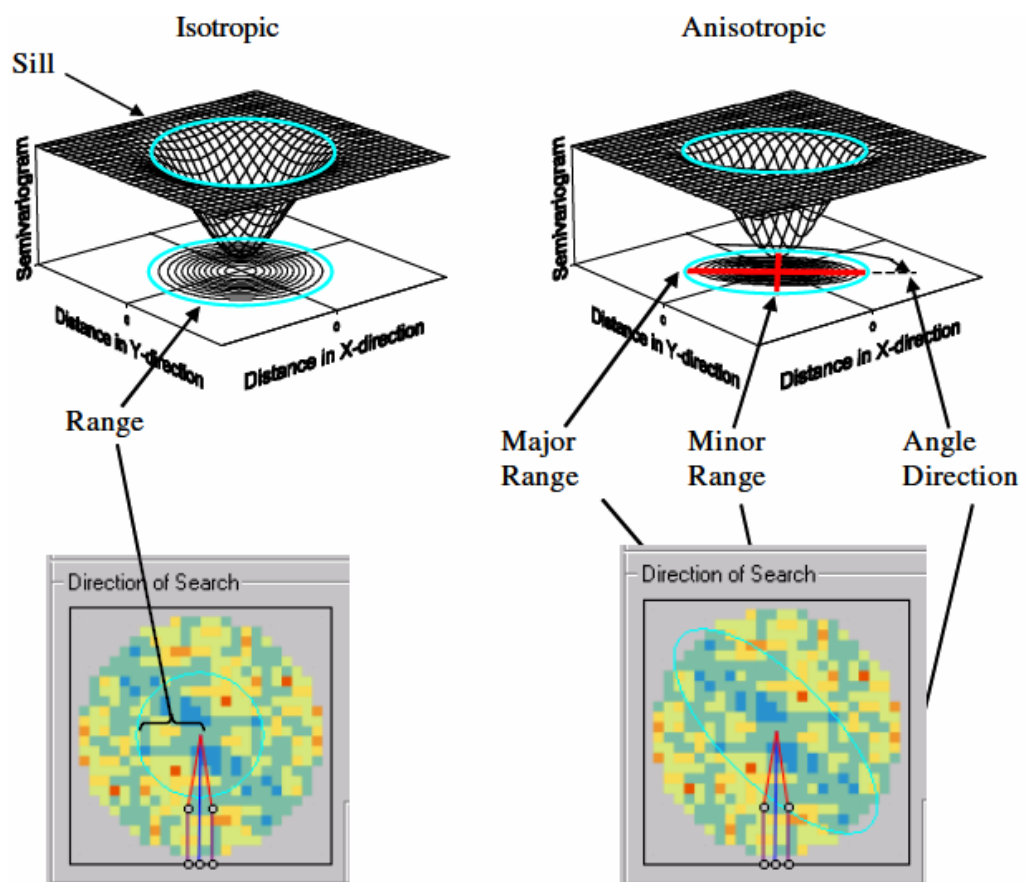
Odlišnost datové sady v určitém směru může být zjištěna použitím nástroje *search direction*. Tato funkce způsobuje, že v závislosti na vybraném směru se mění i empirický semivariogram, který obsahuje pouze hodnoty ve zvoleném směru. Tento nástroj má charakter analýzy, výsledky se pouze zobrazí, ale neovlivní výsledný model.

Pokud byla zjištěna anizotropie v určitém směru, aplikuje se na model pomocí hodnot *major range*, *minor range* a směru anizotropie. Tím je definován anizotropický model semivariogramu (obr. 53). Křivka modelu je nahrazena více křivkami, které zobrazují model semivariogramu v různých směrech. Pro kontrolu může být opět použit nástroj *search direction*. Pokud je model vybrán správně, jeho křivka se musí měnit v každém směru tak, aby dobře vystihovala rozptýlené hodnoty empirického semivariogramu.

V další části nastavení se volí okolí určovaného bodu, ve kterém budou vyhledávány body vlivu na odhad v tomto místě. Možnosti nastavení jsou shodné jako u metody *IDW* v *Geostatistical Analyst*. Okolí může mít tvar kruhu nebo elipsy. Zkreslení v určitém směru se lze vyhnout rozdělením okolí na sektory. Z každého sektoru je potom vybrán požadovaný počet bodů. Počet bodů je stanoven minimálním a maximálním počtem. Standardní nastavení (default) tvaru okolí je převzato z předchozího výpočtu charakteristik modelu semivariogramu.



Obr. 52 Modelování semivariogramu v ArcGIS Geostatistical Analyst

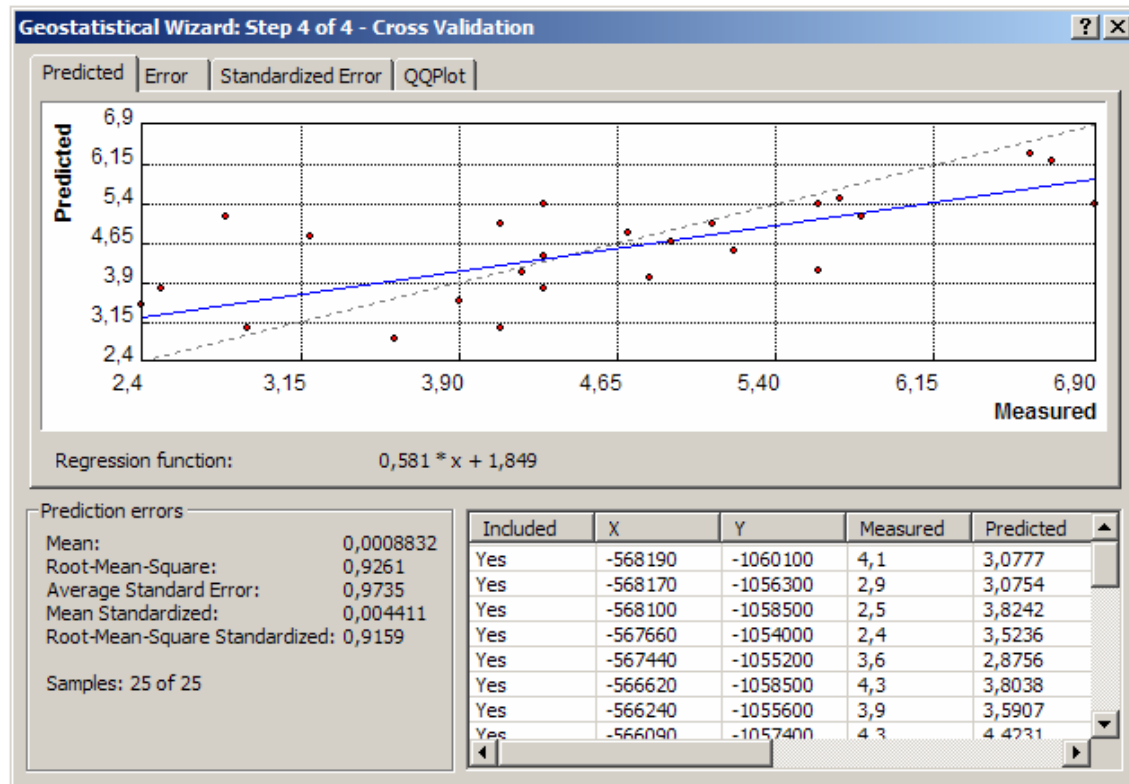


Obr. 53 Izotropický a anizotropický model semivariogramu

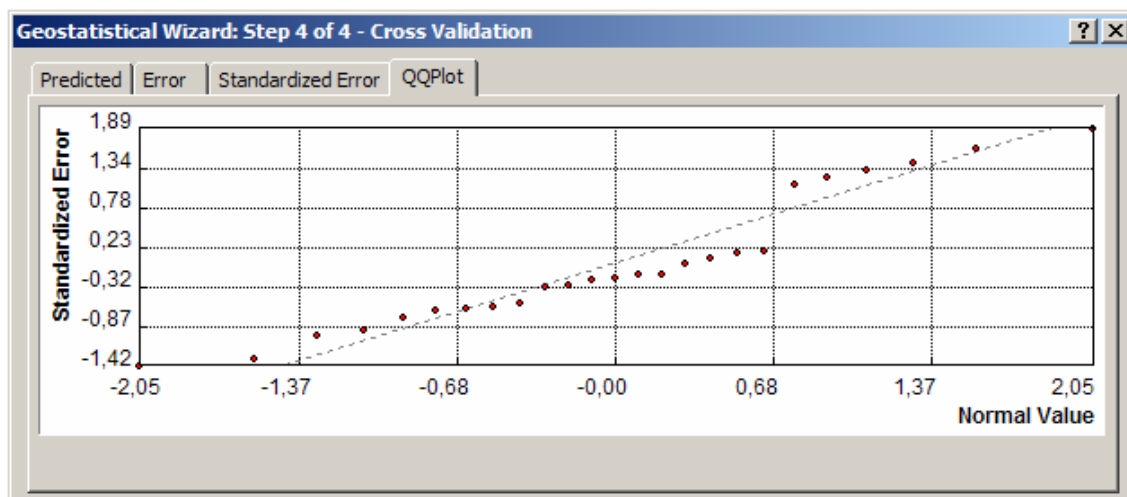
Zvolené okolí a počet bodů ve spojení s vybraným modelem semivariogramu umožňuje určení vah měřených bodů a výpočet výsledného odhadu. Kontrolou správnosti je tzv. cross validation.

Cross validation

Cross validation vypovídá o tom, jak dobře vytvořený model předpovídá neznámé hodnoty. Cross validation vypustí bod (jeden po druhém) a odhadne jeho hodnotu použitím zbývajících bodů. Potom porovná měřenou a odhadnutou hodnotu vynechaného bodu. Toto srovnání je i graficky zobrazeno jako graf odhadnutých hodnot ku měřeným. V ideálním případě by tyto hodnoty měly být rozptýleny kolem přímky 1:1 (přerušovaná čára na obr. 54). Nicméně většinou je sklon přímky menší než jedna (modrá plná čára na obr. 54). Rovnice proložené přímky rozptýlenými body je uvedena pod grafem (regression function). Kdyby byla data prostorově nezávislá (bez autokorelace), proložená přímka by byla vodorovná. V dalším grafu (error) je zobrazen přímo rozdíl mezi odhadnutou a měřenou hodnotou, který se označuje jako chyba odhadu. Vedlejší graf (standardized error) uvádí normalizované chyby odhadu. Jestliže normalizované chyby odhadu mají normální rozdělení, hodnoty výsledného grafu (QQ plot) by měly ležet přibližně podél přerušované čáry (obr. 55).



Obr. 54 Cross validation



Obr. 55 QQ plot

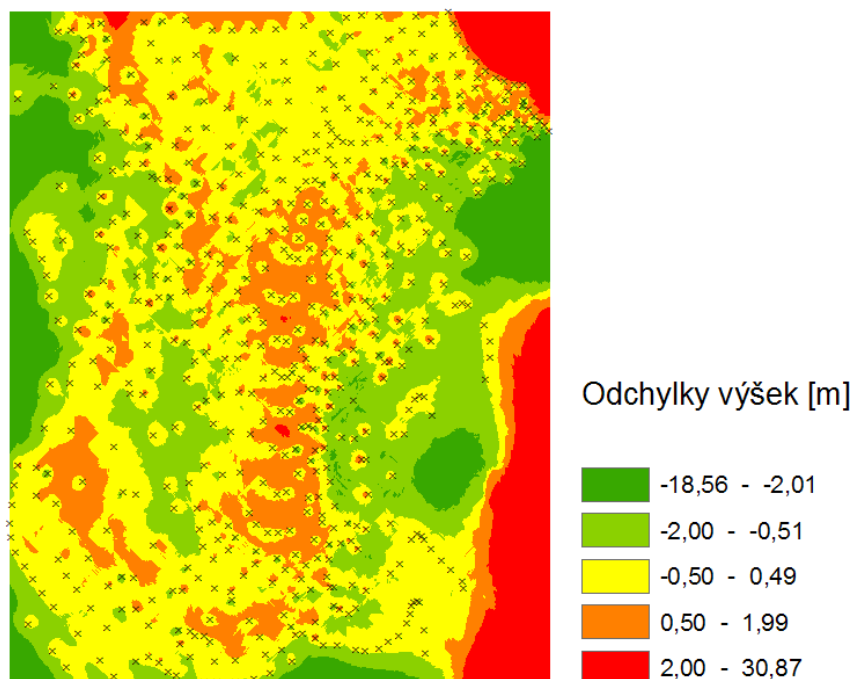
Model, který poskytuje přesné předpovědi, by měl mít střední chybu odhadu (mean prediction error) blízkou nule. Pokud tomu tak je, odhady jsou nezkreslené. Normalizovaná střední chyba odhadu (mean standardized) by také měla být blízká nule. Chyba odhadu RMS by měla být co nejmenší, čím víc je to splněno, tím víc se odhady blíží měřeným hodnotám. Chybu RMS je dobré sledovat při porovnání, jak přesně různé modely odhadují měřené hodnoty. Chyba RMS se vypočte jako odmocnina z průměru čtverců rozdílu mezi měřenou a odhadnutou hodnotou. Normalizovaná chyba odhadu RMS (RMS standardized) by měla mít hodnotu blízkou jedné.

Před tvorbou modelu semivariogramu byl ze souboru dat nadmořských výšek odstraněn trend druhého řádu. Empirický semivariogram byl rozdělen do deseti intervalů o délce deset metrů. Model, který tento empirický semivariogram nejlépe vystihuje, je typ gaussovský. Hodnota nugget vybraného modelu je větší než nula, což znamená, že interpolace bude aproximující. Okolí určovaného bodu, ze kterého jsou vybírány body pro výpočet vah, má tvar kruhu. Kruh byl rozdělen na čtyři sektory, aby se zamezilo zkreslení v určitém směru. Výsledný odhad je na obr. 56.



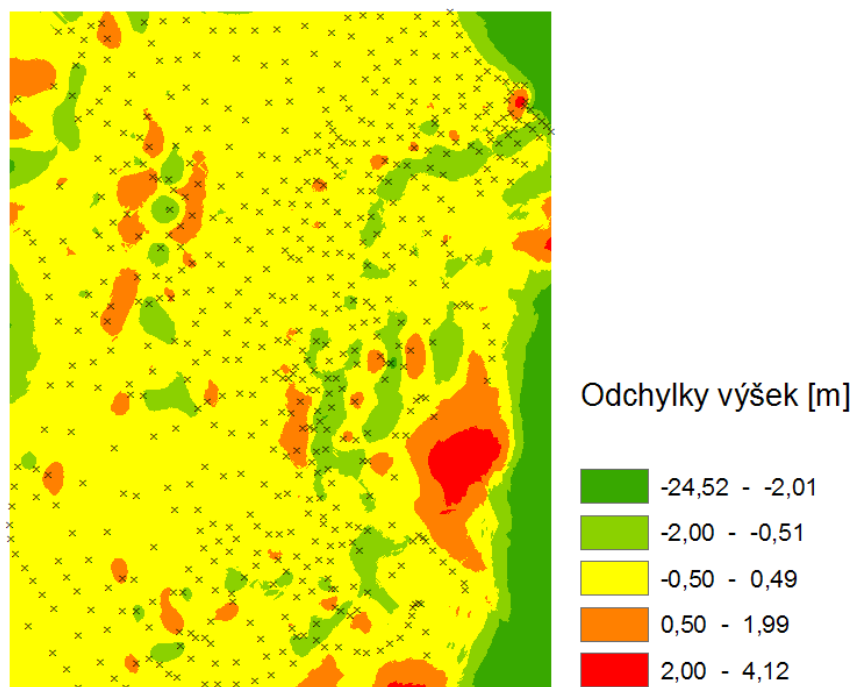
Obr. 56 *Kriging*

Interpolace nadmořských výšek metodou *kriging* byla porovnána s interpolací *IDW* s parametrem 2. *IDW* byla vypočtena v Geostatistical Analyst se stejným nastavením okolí pro vyhledávání bodů vlivu na určovaný bod, jako u *kriging*. Rastr na obr. 57 je výsledek odečtení *IDW* od *kriging*. Protože *IDW* je exaktní a *kriging* aproximující, odchylky výšek jsou i v místech vstupních bodů. *Kriging* oproti *IDW* dokáže vypočítat i výšky přesahující rozsah měřených výšek, což se projevilo na mnoha místech výsledného rastru.



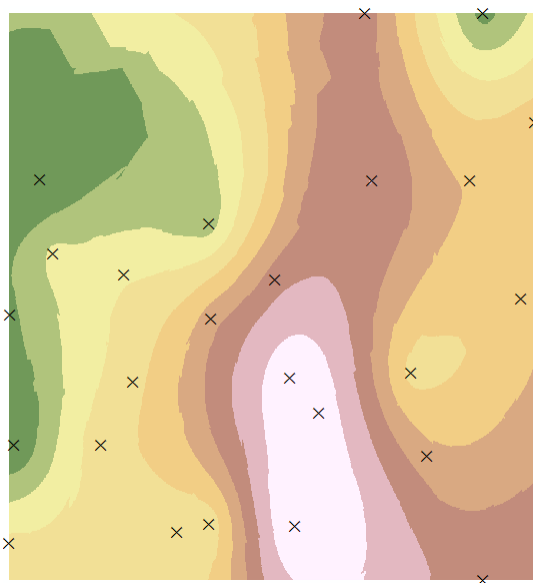
Obr. 57 Odchylky výšek *kriging* a *IDW*

Pro další porovnání s metodou *kriging* byl zvolen *regularized spline* s váhou 0,001. Rozdíl *spline* a *kriging* je na obr. 58. Jedná se opět o rozdíl exaktní a aproximující metody, proto jsou znatelné odchylky v místech měřených výšek. Obě metody přesahují minimální a maximální hodnoty měřených bodů. V tomto případě je tedy oblastí s většími odchylkami výšek mnohem méně.



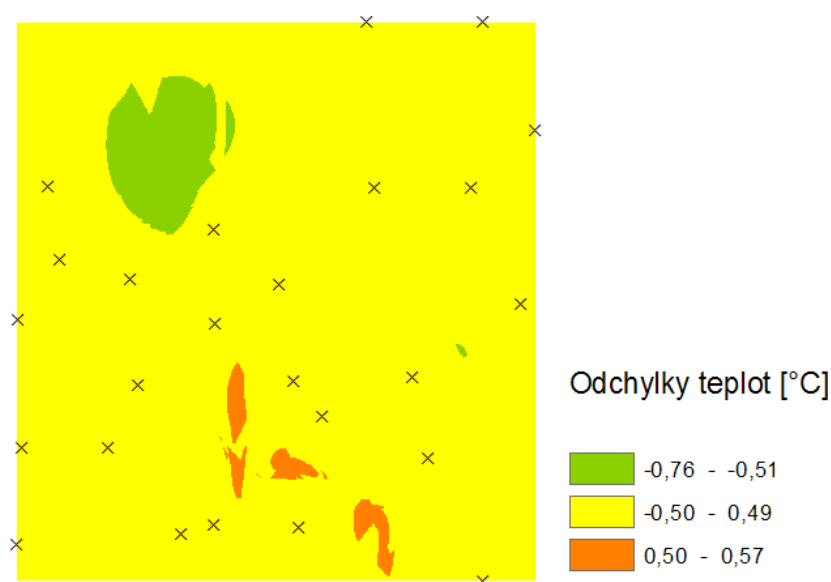
Obr. 58 Odchylky výšek *spline* a *kriging*

Empirický semivariogram znázorňující prostorovou závislost teplot vzduchu byl rozdělen na sedm intervalů o délce 800 m. Pro tento empirický semivariogram byl zvolen sférický anizotropický model. Směrový vliv byl vyjádřen mírným natočením major range modelu (od směru severu o $5,5^\circ$ vpravo). Charakteristika modelu nugget je rovna nule, interpolace tedy bude exaktní. Okolí určovaného bodu, ze kterého jsou vybírány body pro výpočet vah, má tvar elipsy. Parametry elipsy (hlavní, vedlejší poloosa a úhel stočení) jsou rovny charakteristikám modelu semivariogramu (major range, minor range, direction). Elipsa je rozdělena na čtyři sektory, ze kterých je vybírán požadovaný počet bodů. Výsledný odhad je na obr. 59.



Obr. 59 Kriging

Odchyly výsledných teplot metody *kriging* a *IDW* s parametrem 2 jsou na obr. 60. U *IDW* i *kriging* bylo shodně nastaveno okolí pro vyhledávání bodů vlivu. Odhady obou metod jsou si hodně podobné, přispívá tomu skutečnost, že obě metody jsou exaktní. Obě metody používají k výpočtu odhadu vážený průměr, u *IDW* i *kriging* jsou váhy závislé na vzdálenostech mezi měřeným a určovaným bodem, u *kriging* jsou ale ještě závislé na vzájemných prostorových vztazích mezi měřenými body okolo místa určovaného bodu.



Obr. 60 Odchyly výšek *kriging* a *IDW*

Metoda *kriging* vytváří odhad, který může, ale také nemusí procházet vstupními body. V závislosti na tom potom přesahuje nebo nepřesahuje minimální a maximální

hodnoty měřených bodů. Výhodná je možnost ovlivnění průběhu interpolace. Zároveň je ale dosti náročné vytvořit správný model, který vyjadřuje prostorovou autokorelaci sledovaného jevu.

Kriging je vhodnou metodou pro data prokazující zřetelný trend. Používá se zejména pro geologické aplikace jako průzkum ložisek nerostných surovin, někdy se však také hodí pro interpolaci výšek, meteorologických dat nebo znečištění ovzduší.

3.6 Topo to raster

Topo to raster je specifickou metodou programu ArcGIS, která je určena pro tvorbu hydrologicky korektního digitálního modelu terénu. Metoda je založena na programu ANUDEM verze 4.6.3. Je modifikací metody spline *TPS*, přičemž dovoluje modelovat náhlé změny povrchu. *Topo to raster* je metoda, která používá iterativní způsob výpočtu. Je optimalizována tak, aby byla výpočetně účinná jako lokální interpolační metody (např. *IDW*) a zároveň aby výsledný povrch neztratil spojitost, jakou mají globální interpolační metody (např. *spline*).

Hydrologické procesy významně ovlivňují tvar povrchu. Proto metoda *topo to raster* kombinuje při výpočtu nejen výšková data, ale i data vodních toků, vodních ploch a dalších. Výšková data, která jsou v rozporu s daty vodních toků a ploch, jsou ignorována. Použití vodních toků a ploch při tvorbě povrchu je velice efektivní, protože je při interpolaci zohledněna topografická informace. Metoda tedy používá ojedinělý algoritmus, který dokáže velmi zpřesnit výsledný povrch, samozřejmě za předpokladu správného definování všech parametrů:

Vstupní data (input feature data)

Vstupní data mohou být různého druhu:

- výškové body (point elevation) – bodová vrstva, která má uložené výšky jako atributy
- vrstevnice (contour) – liniová vrstva, výšky vrstevnic jsou uloženy jako atributy
- vodní toky (stream) – liniová vrstva, linie musí být orientovány směrem po proudu a každý tok může být vyjádřen pouze jednou linií. Vodní toky mají prioritu před výškovými body a vrstevnicemi.
- deprese (sink) – bodová vrstva reprezentující topografický pokles
- hranice (boundary) – polygonová vrstva, buňky výstupního rastru za touto hranicí budou mít hodnotu NoData

- jezera (lake) – polygonová vrstva, buňky výsledného rastru uvnitř jezera budou mít hodnotu minimální výšky na břehu jezera

Pokud je k dispozici i známá výška hladiny jezera, může být použita linie břehové čáry s výškou jako vstupní vrstva typu vrstevnice (contour). Pořadí, ve kterém jsou vstupní data přidávána, nemá žádný vliv na výsledek.

Rozsah výstupního rastru (output extent)

Výstupní rastr zahrnuje i hodnoty NoData. Minimální souřadnice X je standardně nastavena (default) na nejmenší souřadnici X všech vstupních dat. Maximální souřadnice X je standardně nastavena (default) na největší souřadnici X všech vstupních dat. Stejným způsobem je určena i minimální a maximální souřadnice Y.

Lem (margin in cells)

Je to vzdálenost od hranice (boundary) zadaná počtem buněk. V takto stanoveném lemu ještě proběhne interpolace. Pokud by tento lem nebyl vytvořen, interpolované hodnoty podél hranice by nenavazovaly na interpolované hodnoty případného sousedícího rastru. Pro interpolaci by bylo použito poloviční množství dat, které využívají body uvnitř rastru. Vnitřní body jsou totiž obklopeny vstupními daty ze všech stran. Parametr tedy umožní, že vstupní data ve specifikovaném lemu budou také použita pro interpolaci. Hodnota parametru musí být větší nebo rovna nule, standardní hodnota (default) je 20.

Minimální hodnota souřadnice Z (the smallest Z value)

Minimální interpolovaná hodnota souřadnice Z je standardně nastavena (default) na nejmenší souřadnici Z vstupních souborů sniženou o dvacet procent.

Maximální hodnota souřadnice Z (the largest Z value)

Maximální interpolovaná hodnota souřadnice Z je standardně nastavena (default) na největší souřadnici Z vstupních souborů zvýšenou o dvacet procent.

Zdůraznění odtokové sítě (drainage enforcement)

Typy zdůraznění odtokové sítě jsou následující:

- enforce – všechny bezodtoké oblasti jsou vyplněny, ať jsou pravé či nepravé, výsledkem tedy bude hydrologicky korektní model
- no enforce – žádné bezodtoké oblasti nejsou vyplněny
- enforce with sink – body vstupní vrstvy charakterizující deprese (sink) budou zachovány, ostatní deprese budou považovány za nepravé a budou vyplněny.

V případě, že počet nalezených nepravých depresí přesáhne osm tisíc, výpočet skončí chybou.

Základní typ vstupních dat (primary type of input data)

Jedná se o typ výškových vstupních dat, který je dominantní. Mohou to být buď výškové body nebo vrstevnice. Standardním typem (default) jsou vrstevnice.

Maximální počet iterací (maximum number of iterations)

Počet iterací musí být větší než nula. Standardní počet (default) je 40. Menší počet iterací vyplní méně bezodtokých oblastí. Vzácně se používá i větší počet iterací, kdy dochází k vyplnění více depresí.

Roughness penalty

Parametr charakterizující míru nerovnosti musí být větší nebo roven nule. Pokud jsou základním datovým typem vrstevnice, standardní hodnota (default) je nastavena na nulu. Pro datový typ výškové body je to hodnota 0,5. Větší hodnoty nejsou doporučeny.

Discretisation error factor

Tento parametr se používá k regulaci hladkosti při převodu vstupních dat na rastr. Hodnota parametru musí být větší než nula. Běžně se používají hodnoty z intervalu od 0,5 do 2, standardní hodnota (default) je 1. Čím vyšší je parametr, tím větší vyhlazení.

Vertikální směrodatná odchylka (vertical standard error)

Vertikální směrodatná odchylka může být nastavena na malou kladnou hodnotu, v případě že vstupní výšková data vykazují velkou náhodnou chybu v souřadnici Z. Pro většinu výškových dat by měla být vertikální směrodatná odchylka nulová, což je také standardním nastavením.

Tolerance 1 a Tolerance 2

Hodnota tolerance 1 vyjadřuje hustotu výškových dat ve vztahu k odtokové síti. U výškových bodů se tato hodnota přebírá z hodnoty vertikální směrodatné odchylky. U vrstevnic se doporučuje použít poloviční hodnota průměrného kroku vrstevnic. Tolerance 2 zamezuje vyplňování bezodtokých oblastí nereálně vysokými bariérami. Hodnoty obou tolerancí musí být větší nebo rovny nule. Následující tabulka poskytuje určitý přehled o volbě hodnot.

Typ vstupních dat	Měřítko/krok	Tolerance 1	Tolerance 2	Default	
				Tol 1	Tol 2
Výškové body	1 : 100 000	5	200	0	200
	1 : 500 000	10	400		
Vrstevnice	Krok 10	5	100	2,5	100

Výstupní vrstva vodních toků (output stream polyline feature)

Výstupní liniová vrstva vodních toků je vytvořena na během interpolačního procesu. Porovnáním se vstupní vrstvou vodních toků může být ověřena správnost vytvořené odtokové sítě.

Výstupní vrstva neodstraněných depresí (output remaining sink point feature)

Výstupní bodová vrstva neodstraněných depresí nezahrnuje body vstupní vrstvy charakterizující deprese. Pokud je počet neodstraněných depresí vysoký, lze snížit změnou hodnot tolerancí 1 a 2. Zbývající deprese většinou vyjadřují chyby ve vstupních datech, které proces zdůraznění odtokové sítě (drainage enforcement) nebyl schopen odstranit. Toto je tedy účinný nástroj pro odhalení chyb ve vstupních datech.

Výstupní protokol (output diagnostic file)

Výstupní protokol je textový soubor s výpisem všech vstupních dat, zvolených parametrů a počtem odstraněných depresí při každé iteraci.

Soubor parametrů (output parameter file)

Soubor parametrů je textový soubor s výpisem všech vstupních dat a zvolených parametrů, který může být použit pro zopakování interpolace pomocí nástroje *topo to raster by file*.

Metoda *topo to raster* umožňuje použití hodně typů vstupních dat a nastavení mnoha parametrů. Pro jednodušší zkoušení různého nastavení parametrů je vhodné využít nástroj *topo to raster by file*. Vstupní soubor parametrů se vytvoří nástrojem *topo to raster*, v tomto souboru lze dále modifikovat nastavení parametrů v jakémkoliv textovém editoru. Ukázka souboru parametrů je na následujících řádcích:

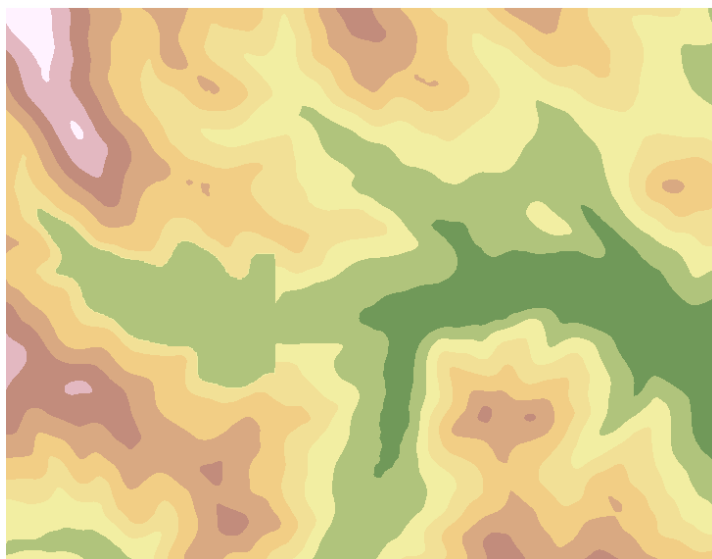
```
POINT C:\DATA\TTR\vysky.shp VYSKA
STREAM C:\DATA\TTR\vodni_toky.shp
LAKE C:\DATA\TTR\vodni_plochy.shp
BOUNDARY C:\DATA\TTR\hranice.shp
ENFORCE ON
DATATYPE SPOT
ITERATIONS 40
ROUGHNESS_PENALTY 0,5000000000
DISCRETE_ERROR_FACTOR 1,0000000000
```

```

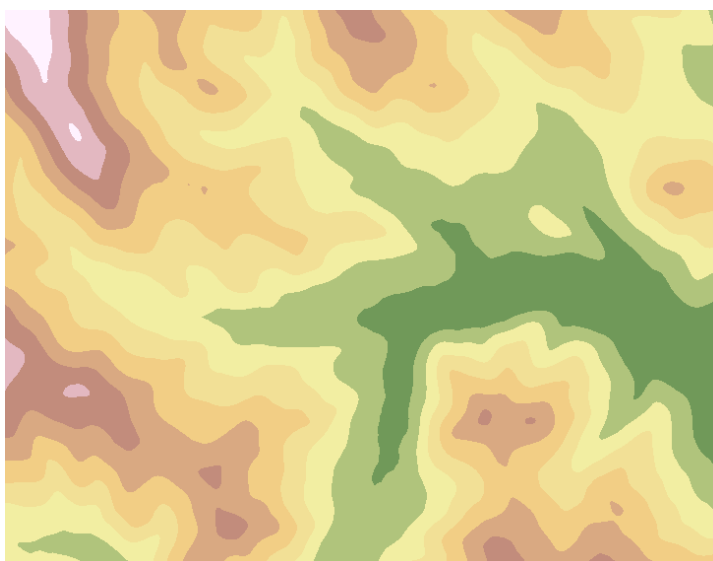
VERTICAL_STANDARD_ERROR 0,000000000000
TOLERANCES 0,000000000000 200,000000000000
EXTENT 672500,00012206926 5556500,00012206940
683500,00012207020 5567500,00012207030
CELL_SIZE 10,000000000000
MARGIN 20
OUT_STREAM
OUT_SINK
OUT_DIAGNOSTICS C:\DATA\TTR\diagnostic.TXT

```

Pro ukázkou metody *topo to raster* bylo zvoleno území, pro které jsou k dispozici následující data: body s nadmořskými výškami, linie vodních toků, polygon jezera a hranice území. Na obr. 61 je výsledný odhad, který vznikl interpolací s použitím všech dostupných dat. Při porovnání vstupní a výstupní sítě vodních toků zjistíme, že obě sítě jsou si velice podobné, což vypovídá o dobrém výsledku interpolace. Odtoková síť vytvořeného povrchu je v souladu se vstupními daty.

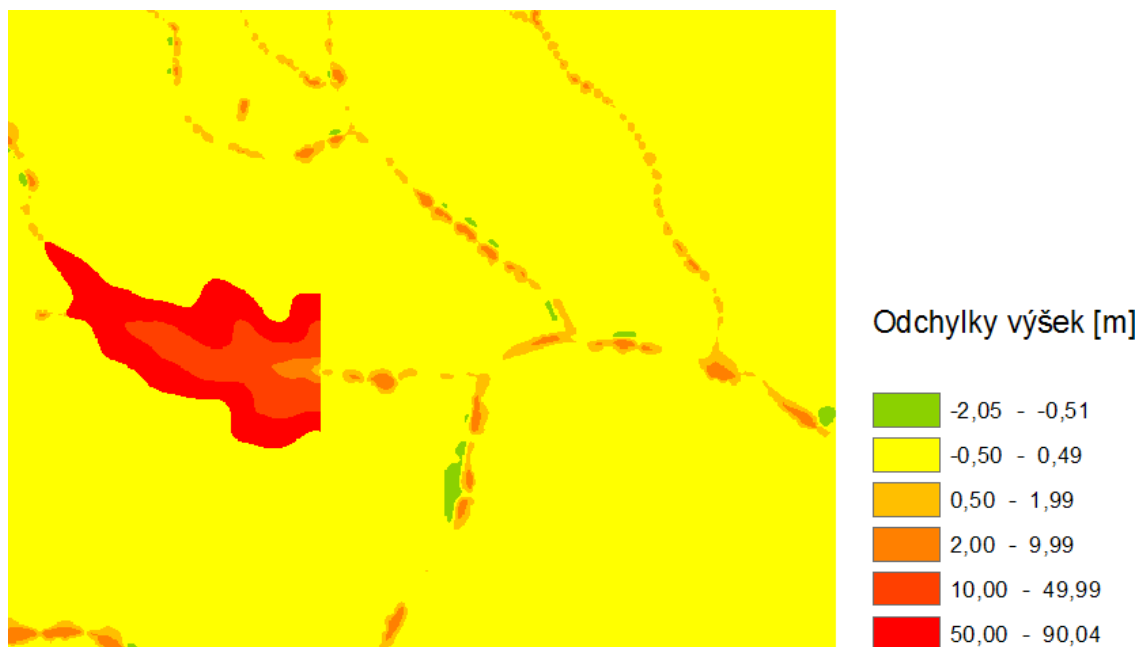


Obr. 61 *Topo to raster 1*



Obr. 62 *Topo to raster 2* (bez použití hydrologických dat)

Dále byla vyzkoušena interpolace povrchu stejného území bez použití hydrologických dat. Obr. 62 zobrazuje výsledek, kdy vstupními daty *topo to raster* jsou jen výškové body a hranice území. Vytvořená odtoková síť je v tomto případě více odlišná od té vstupní. Výsledný povrch dokonce obsahuje tři neodstraněné deprese. Nejvýraznější rozdíl mezi tímto odhadem a odhadem vytvořeným na základě všech dostupných dat je v oblasti plochy jezera v levé části rastru uprostřed. Při interpolaci s hydrologickými daty má vodní plocha přednost před výškovými daty a hladině jezera byla přiřazena nejmenší výška na břehu jezera. Pro lepší zobrazení rozdílů mezi těmito dvěma odhady byl odečten rastr *topo to raster* 1 od *topo to raster* 2 (obr. 63). Na obr. 63 lze kromě razantních odchylek v oblasti jezera pozorovat i odchylky podél všech vodních toků.



Obr. 63 Odchylky výšek *topo to raster* 2 a *topo to raster* 1

Metodu *topo to raster* je vhodné použít, když jsou k dispozici kromě výškových dat ještě data popisující hydrologickou strukturu povrchu. Výsledkem *topo to raster* je hydrologicky korektní model terénu.

3.7 Hodnocení kvality interpolace

Přehled o kvalitě výsledku interpolace lze získat díky již zmíněné cross validation nebo pomocí validace (validation). Pro validaci je potřeba mít k dispozici část dat nepoužitých pro výpočet odhadu, kterými se výsledný odhad otestuje. V mnoha případech je ale množství měřených dat příliš malé na to, aby bylo možné postupovat

tímto způsobem. Často se tedy k hodnocení interpolace používá spíše metoda cross validation. ArcGIS umožňuje provedení cross validation i validace (validation) v nadstavbě Geostatistical Analyst.

Závěr

Žádná interpolační metoda není univerzální a nedá se aplikovat na všechny jevy. Výběr metody závisí na charakteru sledovaného jevu a struktuře vstupních dat. I v rámci jednoho jevu může být volba metody rozdílná, například u nadmořských výšek záleží na členitosti území, dále také na struktuře měřených dat. Proto neexistuje žádná striktní poučka pro použití interpolační metody.

Při výběru metody je vhodné řídit se výsledky kvality interpolace (validation, cross validation). Tímto postupem může být nejen usnadněna volba samotné metody, ale i porovnání různých nastavení v rámci jedné metody, a tak vybráno nejvhodnější nastavení pro konkrétní úlohu.

Při porovnání nadstaveb ArcGIS pro využití při interpolaci dat se zdá být uživatelsky nejjednodušší prostředí 3D Analyst a Spatial Analyst (mají shodné prostředí). Tyto nadstavby nenabízejí při interpolaci tak širokou volbu parametrů jako Geostatistical Analyst, proto se zdá jejich použití snadnější. Pokročilejší nastavení Geostatistical Analyst je ale výhodou a práci s větším množstvím parametrů usnadňuje okno s náhledem výsledku interpolace. Geostatistical Analyst také umožňuje ověření kvality interpolace (validation, cross validation), což je velkou výhodou.

Cílem této práce bylo vysvětlení podstaty jednotlivých interpolačních metod a možností jejich provedení v ArcGIS. Často se s nutností použití interpolace setkáváme a jsme nuceni k výběru jedné určité metody. Tento text by měl pomoci porozumět principu jednotlivých metod a usnadnit tak rozhodování při výběru optimální interpolační metody.

V rámci této práce samozřejmě nebylo možné obsáhnout veškerou variabilitu některých metod. Daný rozsah neumožnil komplexní zpracování metod *spline* a *kriging*. I přesto by měl být popis těchto metod postačující.

Literatura

HLÁSNY, Tomáš. *Geografické informačné systémy : Priestorové analýzy*. 1. vyd. Poniky : Čižmarová - Partner, 2007. 160 s. ISBN 978-80-8093-029-5.

GEOBUSINESS. Springwinter, s.r.o. 2008 , roč. 7, č. 11.

Environmental Systems Research Institute, Inc.. *ArcGIS 9.2 Desktop Help* [online]. [2007] , Last Modified March 15, 2007 [cit. 2009-04-02]. Dostupný z WWW: <<http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=welcome>>.

JOHNSTON, Kevin, et al. *ArcGIS 9 : Using ArcGIS Geostatistical Analyst*. 2nd edition. USA : ESRI, c2003. 300 s. Dostupný z WWW: <http://dusk2.geo.orst.edu/gis/geostat_analyst.pdf>.

JOHNSTON, Kevin, et al. *Using ArcGIS Geostatistical Analyst*. 1st edition. USA : ESRI, c2001. 48 s. Dostupný z WWW: <http://www.ci.uri.edu/projects/geostats/Using_ArcGIS_Geostat_Anal_Tutor.pdf>.

MCCOY, Jill, et al. *ArcGIS 9 : Using ArcGIS Spatial Analyst*. 1st edition. USA : ESRI, c2001. 232 s.

SÁRKOZY, Ferenc. *Gis functions - interpolation* [online]. TU Budapest, 1998 [cit. 2009-03-10]. Dostupný z WWW: <http://www.agt.bme.hu/public_e/funcint/funcint.html>.

Surface interpolation and geostatistics [online]. GISCA, [cit. 2009-03-10]. Dostupný z WWW: <http://www.gisca.adelaide.edu.au/education_training/lectures/sdvis5014/lectures/lecture1.html>.

KRAUS, Jaroslav. *Geostatistika jako prostorové modelování statistických jevů* [online]. Praha : Český statistický úřad, [cit. 2009-03-15]. Dostupný z WWW: <<http://panda.hyperlink.cz/cestapdf/pdf07c6/kraus.pdf>>.

HORÁK, Jiří. *Prostorová analýza dat* [online]. VŠB, 2002 , Tato stránka byla naposledy aktualizována: 08 prosince, 2002 21:37 [cit. 2009-04-07]. Dostupný z WWW: <<http://gis.vsb.cz/pad/index.htm>>.

ROST, Michael, KLUFOVÁ, Renata. *Matematicko - statistické základy precizního zemědělství* [online]. JCU, [cit. 2009-03-21]. Dostupný z WWW: <http://www2.zf.jcu.cz/public/departments/kmi/MSMT_05/precision_farming_primer.pdf>.