

GEOSTATISTIKA – vymezení pojmu

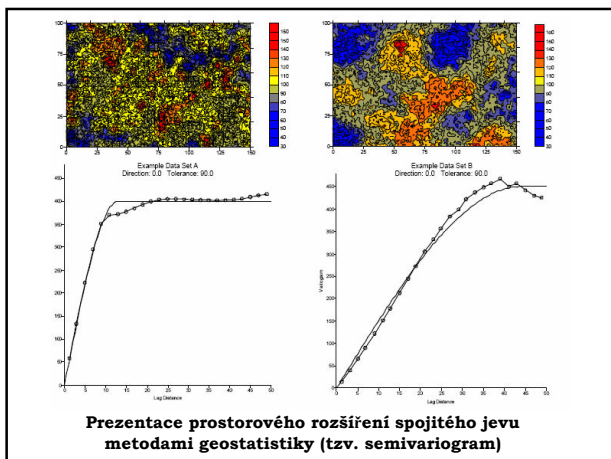
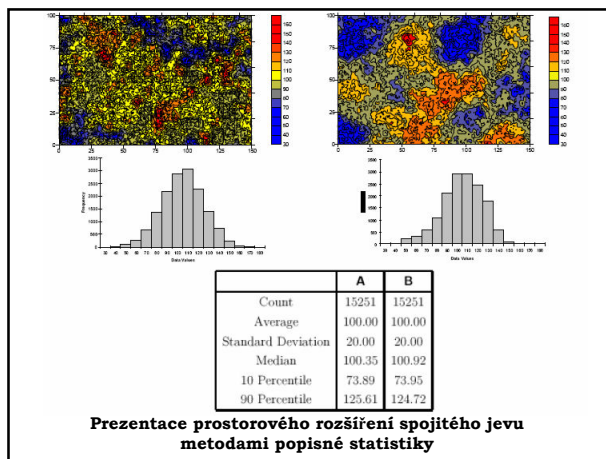
• V užším slova smyslu – skupina interpolačních algoritmů založených na metodě krigingu.

• V širším slova smyslu – statistická analýza prostorově lokalizovaných dat.

Pomocí „klasických“ statistických metod lze vhodně analyzovat především atributová data – jejich kvantitativní či kvalitativní vlastnosti.

Velmi omezeně však jimi lze charakterizovat prostorové vlastnosti objektů a jevů.

Prostorové vlastnosti jako např. spojitost jevů, prostorovou autokorelaci, prostorové uspořádání (strukturu) lze charakterizovat právě pomocí geostatistických metod.



GEOSTATISTIKA – vymezení pojmu

- Deterministické metody interpolace
- Koncept prostorové autokorelace
- Strukturní analýza a popis prostorové autokorelace strukturními funkcemi
- Konstrukce spojitých polí metodami krigingu
- Statistický popis prostorově lokalizovaných dat (geografických objektů) – „**point descriptors**“
- Statistický popis prostorového uspořádání objektů (bodů, linií, ploch) – „**pattern detectors**“
- Objektivní metody klasifikace jevů

Metody prostorové interpolace

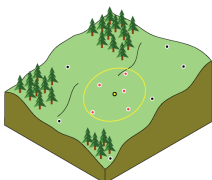
Interpolace – skupina metod, které slouží k odhadu neznámých hodnot proměnné v jistých bodech (neměřených) na základě hodnot proměnné v bodech měřených.

Prostorová interpolace – skupina metod, které slouží k vytváření spojitých povrchů (polí) z bodových měření.

Body mohou být lokalizovány v 1, 2 i 3 rozměrném prostoru. Interpolace se může týkat nejenom bodů, ale i linií a ploch.

Extrapolace – odhad hodnot proměnné vně oblasti definované krajními body měření.

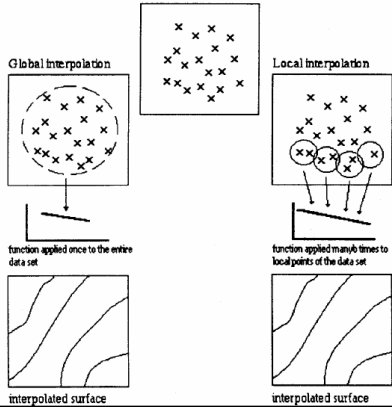
Naprostá většina interpolačních postupů je založena na principu **prostorové autokorelace** – tedy na předpokladu, že hodnoty odhadované veličiny v lokalitách blízkých si budou více podobné než hodnoty v lokalitách vzdálených.



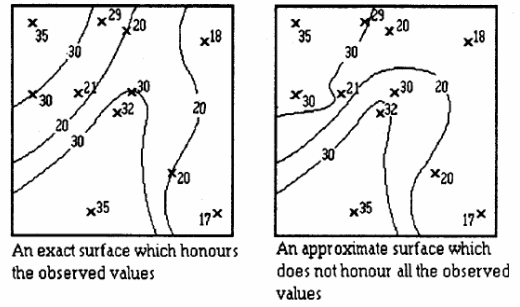
Rozdělení metod prostorové interpolace

- metody interpolace bodů, linií a ploch.
- metody lokální a globální
- metody exaktní a aproximující
- metody spojitě a zlomově (abrupt)
- metody deterministické a stochastické

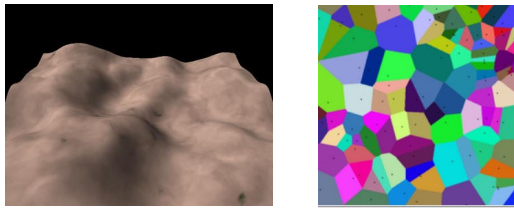
Globální a lokální metody interpolace



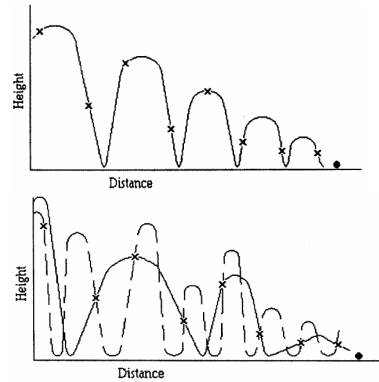
Exaktní a aproximující metody interpolace



Spojité a zlomové metody interpolace

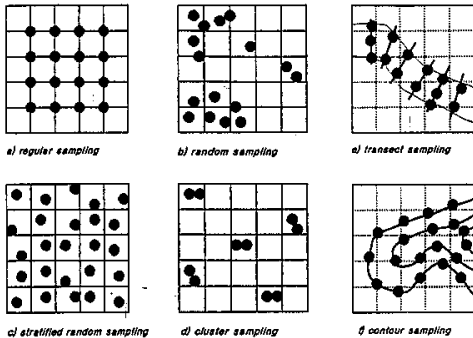


Deterministické a stochastické metody interpolace



Výběr reprezentativních vzorků (sampling)

Je důležitý pro výběr interpolačního algoritmu, úspěšnost vlastní interpolace a pro validaci výsledků



Aspekty ovlivňující úspěšnost interpolace

- > způsob prezentace spojitého pole (grid, TIN, izočáry, areály)
- > dostupné datové zdroje pro interpolaci
- > vymezení studované plochy – přirozené a administrativní hranice
- > dostupnost bodů měření vně studované plochy

Předpoklady úspěšné prostorové interpolace

- > existence dostatečně reprezentativního vzorku měřených dat
- > vhodné vlastnosti měřené veličiny a typ dat (ordinální, intervalová, poměrová)
- > teoretické i empirické znalosti o povaze prostorové diferenciace studovaného jevu
- > znalost podstaty použitelných interpolačních metod
- > znalost způsobu výběru nevhodnější metody

Průzkumová analýza prostorových dat (ESDA).

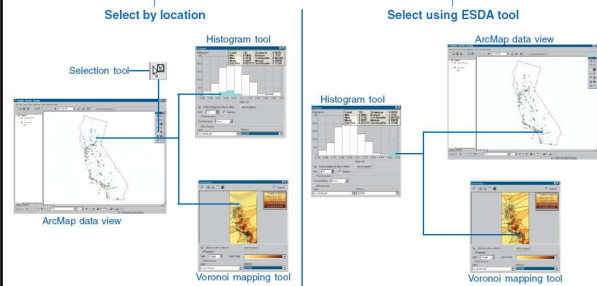
- **EDA** – Exploratory Data Analysis
- **ESDA** – Exploratory Spatial Data Analysis
- (**ESTDA** – Exploratory Spatio – Temporal Data Analysis)

- Množina statistických metod a speciálních nástrojů, zvláště grafických metod, používaných k lepšímu porozumění datům, k odhalení jejich důležitých vlastností.
- Jejím cílem je zjistit základní informace o charakteru vstupních dat v tomto případě za účelem následné interpolace.

Průzkumová analýza prostorových dat (ESDA).

Exploratory Spatial Data Analysis

Selection of Data Points

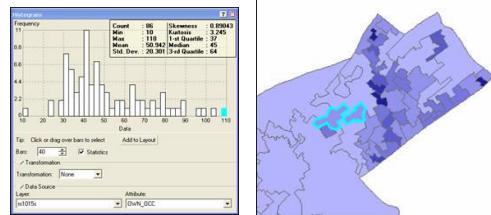


Základní postupy průzkumové analýzy prostorových dat

- výpočet základní popisné statistiky včetně momentů vyššího řádu (asymetrie a špičatosti)
- prověření požadavků normality a stacionarity
- analýza rozdělení hodnot - analýza histogramu
- analýza kvantilového grafu (Q-Q grafu)
- zkoumání odlehých hodnot a jejich případné odstranění
- analýza trendu a jeho případné odstranění
- případná transformace vstupních dat (log)

Základní nástroje průzkumové analýzy prostorových dat

Popisná statistika a „mapped histogram“ - propojení mapy a grafu



Hodnocení polohy a prostorového uspořádání typických resp. extrémních hodnot.

Základní nástroje průzkumové analýzy prostorových dat

Voronoi map

- definování přirozených sousedů k vyšetřovanému bodu
- výpočet lokální statistiky (od měr úrovně až po míry entropie)

Základní nástroje průzkumové analýzy prostorových dat

Voronoi map

Entropie – je počítána z hodnot daného polygonu a všech polygonů sousedních. Nejprve jsou všechny polygony rozříděny do pěti tříd.

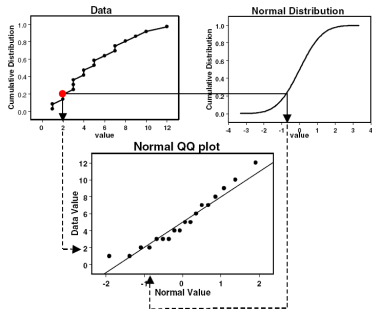
$$Entropie = -\sum p_i * \log_{-} p_i$$

kde p_i je poměr polygonů náležejících do dané třídy z celkového počtu polygonů

Minimální entropie – všechny buňky patří do stejné třídy
Maximální entropie – každá z buněk náleží k jiné třídě.

Základní nástroje průzkumové analýzy prostorových dat

Kvantilové grafy - grafy zobrazující kvantily dvou rozdělení

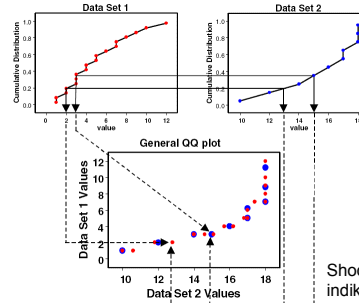


Normální Q-Q graf

Slouží jako nástroj k posouzení normality vstupních dat.

Základní nástroje průzkumové analýzy prostorových dat

Obecný Q-Q graf – testuje se podobnost rozdělení dvou datových souborů, vynáší se odpovídající si hodnoty kvantilů



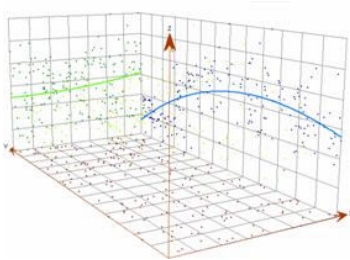
Shodu v obou případech indikují v grafech body přimykající se k přímce.

Základní nástroje průzkumové analýzy prostorových dat

Analýza trendu

Definování globálního trendu v datech, jeho odhalení a eventuálního odstranění.

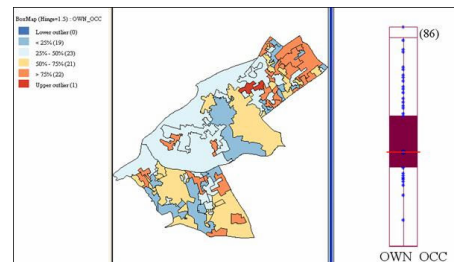
Spočívá v projekci hodnot vyšetřovaných bodů do rovin xz a yz a jejich proložení polynomem n-tého řádu.



Základní nástroje průzkumové analýzy prostorových dat

Krabicové grafy (box plots)

- detekce odlehých či extrémních hodnot
- lokální a globální odlehle hodnoty



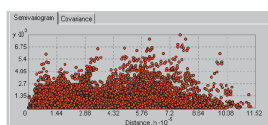
Základní nástroje průzkumové analýzy prostorových dat

Detekce míry prostorové autokorelace

• **Semivariance (semivariogram)** – empirický semivariogram jako graf míry nepodobnosti.

• V explorační analýze slouží k **vystižení míry anizotropie, odhalení odlehých hodnot.**

• V úlohách interpolace je tato veličina důležitá pro objektivní definování **velikosti a tvaru okolí** vyšetřovaného bodu.



Základní nástroje průzkumové analýzy prostorových dat

Vykreslení množiny hodnot semivariance či covariance

Polovina ze sumy čtverců rozdílů hodnot všech dvojic vyšetřovaných bodů vzdálených o určitou hodnotu.

Semivariance je míra nepodobnosti

Hodnota empirické **semivariance** proměnné z pro dvojici bodů v poloze x_i a x_j :

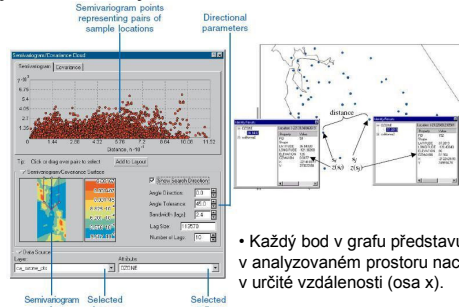
$$0,5 * (z(x_i) - z(x_j))^2$$

Hodnota empirické **covariance** jako míry podobnosti

$$(z(x_i) - \bar{z})(z(x_j) - \bar{z})$$

Základní nástroje průzkumové analýzy prostorových dat

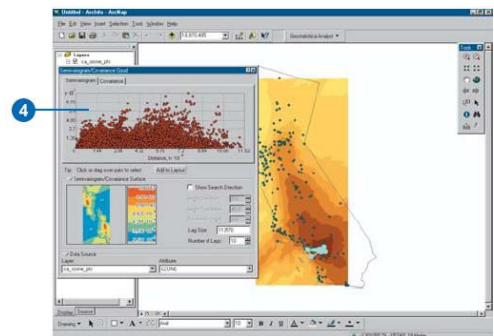
Vykreslení množiny hodnot semivariance či covariance



• Každý bod v grafu představuje dvojici bodů v analyzovaném prostoru nacházejících se v určité vzdálenosti (osa x).

• Podobnost hodnot interpolované veličiny je vyjádřena semivariací (osa y).

Základní nástroje průzkumové analýzy prostorových dat



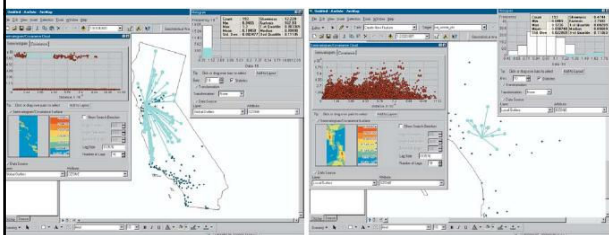
Příklad identifikace bodů, které se odlišují od obecného modelu prostorové autokorelace spojité veličiny. Tento model je vyjádřen grafem (semivariogram)

Základní nástroje průzkumové analýzy prostorových dat

Detekce odlehých hodnot (outliers)

Základní nástroje:

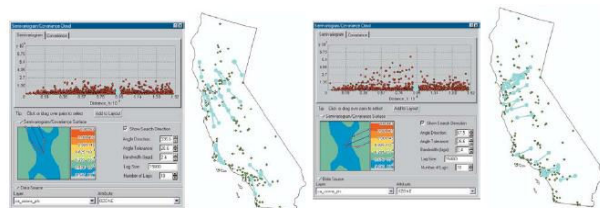
- histogram
- semivariogram/ covariance cloud
- Voronoi map



Detekce globální (vlevo) a lokální (vpravo) odlehle hodnoty.

Základní nástroje průzkumové analýzy prostorových dat

Vyšetřování tvaru okolí – izotropní a anizotropní povrch



Hodnoty semivariance pro směr definovaný na obr. vlevo jsou menší (tedy více podobné) než hodnoty semivariance bodů ve směru definovaném na obr. vpravo - izotropie