

Základy geostatistiky



GEOSTATISTIKA – vymezení pojmu

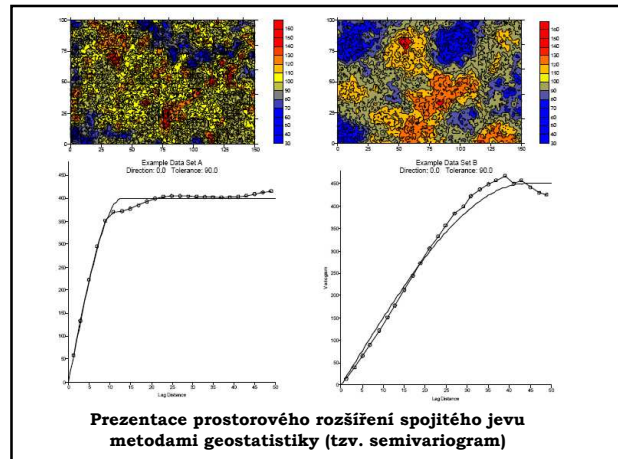
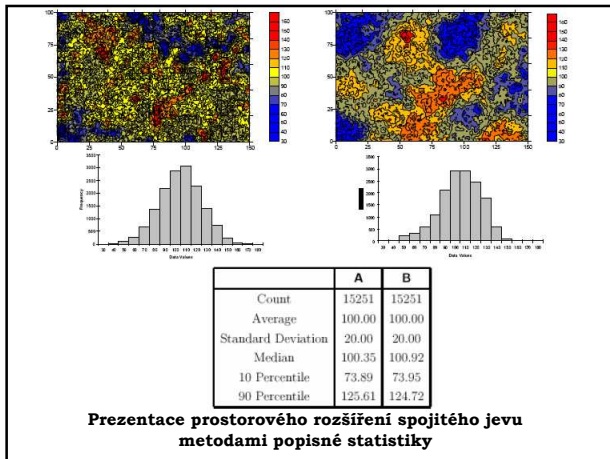
• V užším slova smyslu – skupina interpolačních algoritmů založených na metodě krigingu.

• 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.

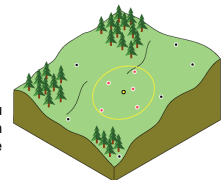
Prostorové vlastnosti jako např. spojitost jevů, prostorovou autokorelaci, prostorové uspořádání (strukturu) lze charakterizovat právě pomocí geostatistických metod.



GEOSTATISTIKA – vymezení pojmu

- Deterministické metody interpolace
- Koncept prostorové autokorelace
- Strukturální analýza a popis prostorové autokorelace strukturálními 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
- Aplikace v geografických disciplínách (indexy krajinné struktury, cílené hospodaření (precision farming), geografie obyvatelstva,

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.

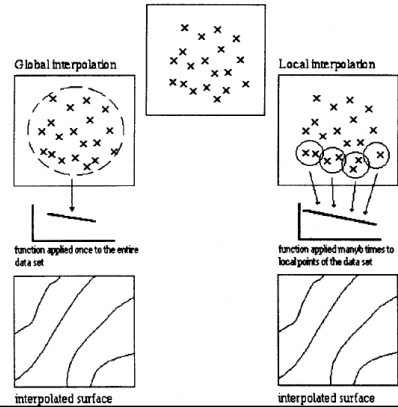
Extrapolace – odhad hodnot proměnné vně oblasti definované krajními body měření.

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.

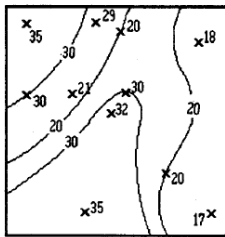
Rozdělení metod prostorové interpolace

- > metody interpolace bodů, linií a ploch.
- > metody lokální a globální
- > metody exaktní a aproximující
- > metody spojité a zlomové (abrupt)
- > metody deterministické a stochastické

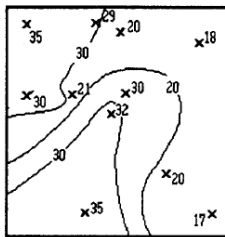
Globální a lokální metody interpolace



Exaktní a aproximující metody interpolace

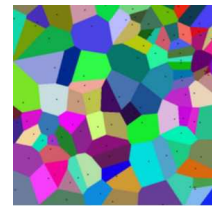
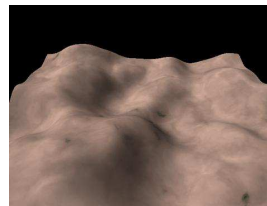


An exact surface which honours the observed values

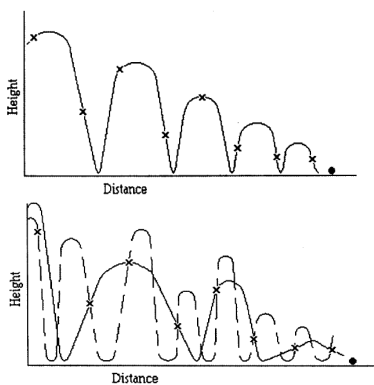


An approximate surface which does not honour all the observed values

Spojité a zlomové metody interpolace

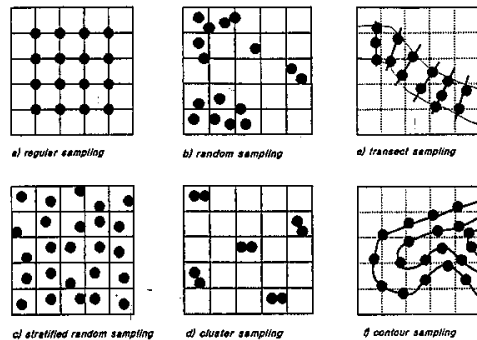


Deterministické a stochastické metody interpolace



Výběr reprezentativních vzorků (sampling)

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



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.

$$Entropie = -\sum p_i * \text{Log}_2 p_i$$

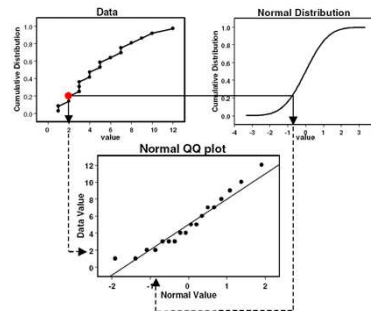
kde p_i je poměr polygonů náležejících do dané třídy z celkového počtu polygonů

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

Def. **entropie** jako veličiny udávající "míru neuspořádanosti" zkoumaného systému.

Základní nástroje průzkumové analýzy prostorových dat

Kvantilové grafy - grafy zobrazující kvantily dvou rozdělení

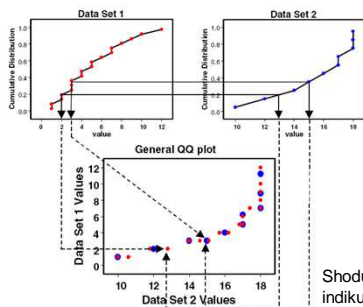


Normální Q-Q graf

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, vynášejí se odpovídající si hodnoty kvantilů

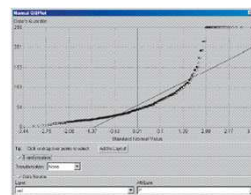


Shodu v obou případech indikují v grafech body přímkykající se k přímce.

Základní nástroje průzkumové analýzy prostorových dat

Transformace dat

Data, která se odchylují od požadovaného rozdělení vyžadují transformaci



Základní typy transformací:

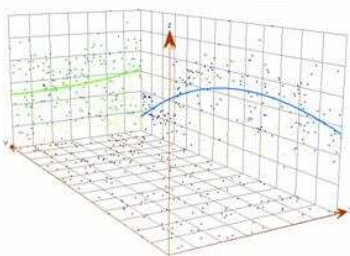
- Logaritická
- Box-Cox (mocnná)
- Arcsine

Základní nástroje průzkumové analýzy prostorových dat

Analýza trendu

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

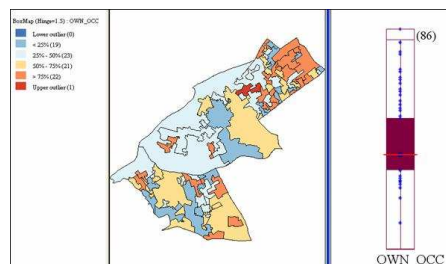
Počívá v projekci hodnot vyšetřovaných bodů do rovin xz a yz a jejich proložení polynomem n-tého řádu.



Základní nástroje průzkumové analýzy prostorových dat

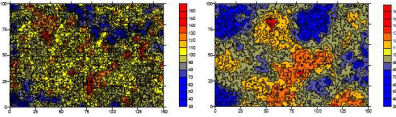
Krabicové grafy (box plots)

- detekce odlehlých či extrémních hodnot
- lokální a globální odlehlé hodnoty



Základní nástroje průzkumové analýzy prostorových dat

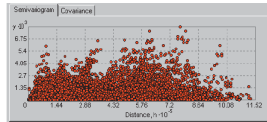
Detekce míry prostorové autokorelace



• **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.

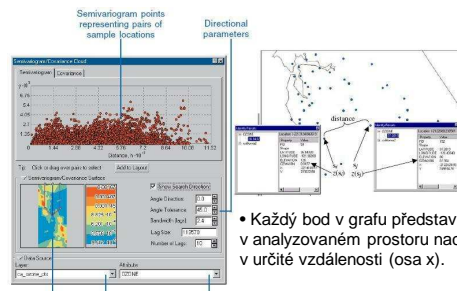
Semivariance je míra nepodobnosti

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

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

Základní nástroje průzkumové analýzy prostorových dat

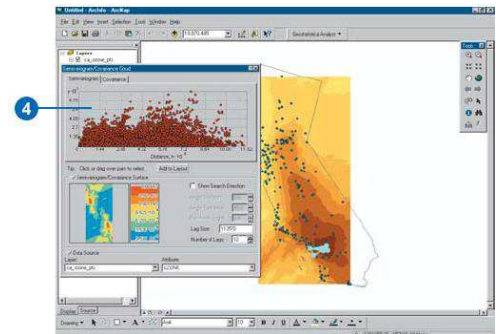
Vykreslení množiny hodnot semivariance či covariance



• 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



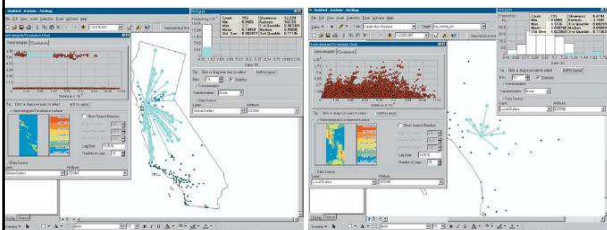
Příklad identifikace bodů, které se odlišují od obecného modelu prostorové autokorelace spojitě veličiny. Tento model je vyjádřen grafem (**semivariogram**)

Základní nástroje průzkumové analýzy prostorových dat

Detekce odlehých hodnot (outliers)

Základní nástroje:

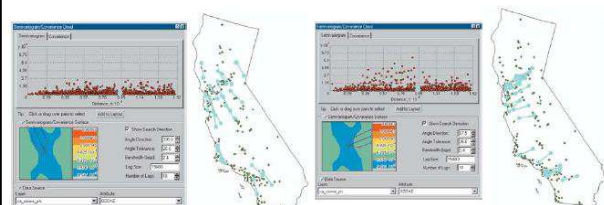
- histogram
- semivariogram/ covariance cloud
- Voronoi map



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)

To indikuje, že semivariance jako míra podobnosti závisí na směru, kterým je měřena – tzv. izotropní povrch. Okolí bodu bude potřeba definovat jako asymetrické.