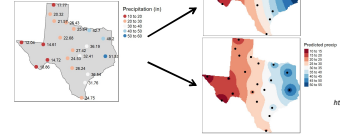


Základy geostatistiky

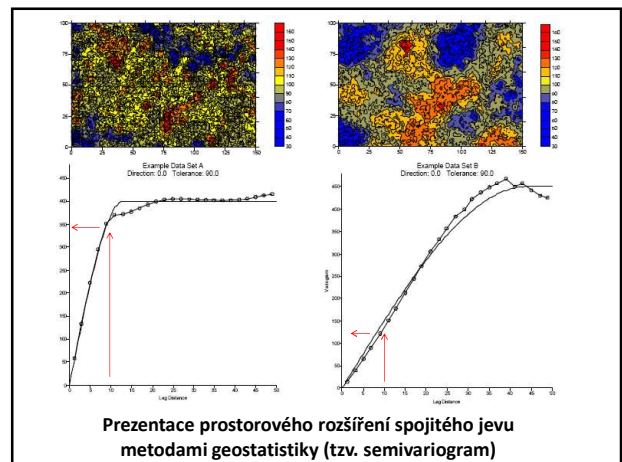
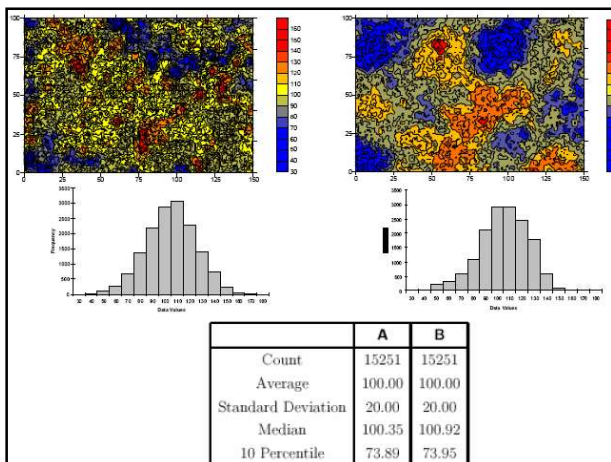


GEOSTATISTIKA – vymezení pojmu

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

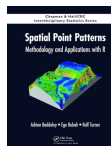


- V širším slova smyslu – **statistická analýza prostorově lokalizovaných dat**.
- Pomocí „klasických“ statistických metod lze vhodně analyzovat především atributová data – jejich kvantitativní či kvalitativní vlastnosti.
- Velmi omezeně však jimi lze charakterizovat prostorové vlastnosti objektů a jevů.
- Prostorové vlastnosti jako např. **spojitost jevů**, **prostorovou autokorelaci**, **prostorové uspořádání (strukturu)** lze charakterizovat právě pomocí geostatistických metod.

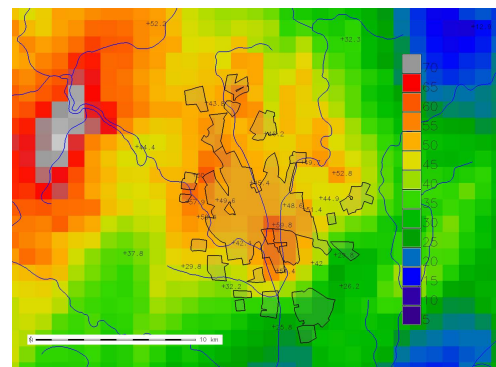


GEOSTATISTIKA – vymezení pojmu

- Deterministické metody interpolace
- Koncept prostorové autokorelace
- Strukturní analýza** a popis prostorové autokorelace strukturními funkcemi
- Konstrukce spojitých polí metodami **krigingu**
- Point pattern analysis** (Analýza bodových vzorů)
 - 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, ...)

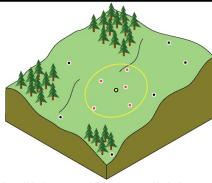


Metody prostorové interpolace



Metody prostorové interpolace

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

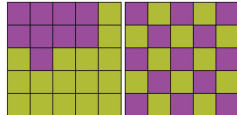


Prostorová interpolace – skupina metod, které slouží k vytváření spojitých povrchů (polí) z bodových měření.

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.

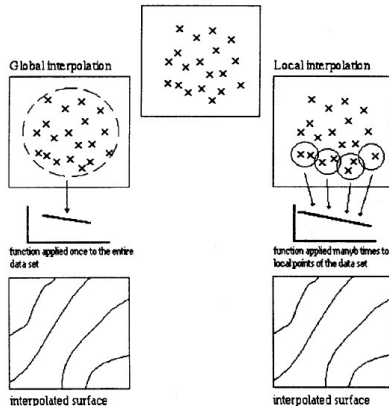
Naprostá většina interpolačních postupů je založena na principu **prostorové autokorelace** – tedy na předpokladu, že hodnoty odhadované veličiny v lokalitách blízkých si budou více podobné než hodnoty v lokalitách vzdálených.



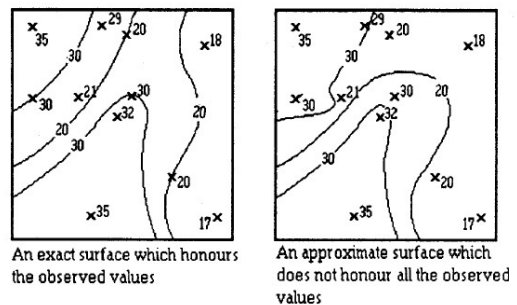
Rozdělení metod prostorové interpolace

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

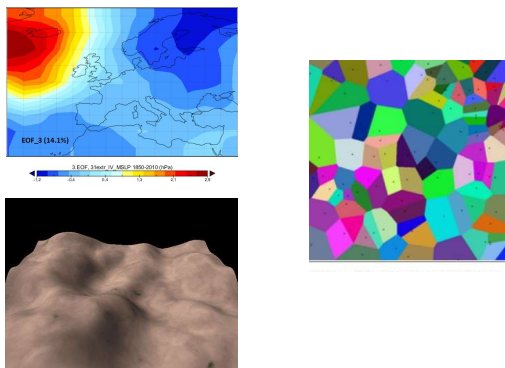
Globální a lokální metody interpolace



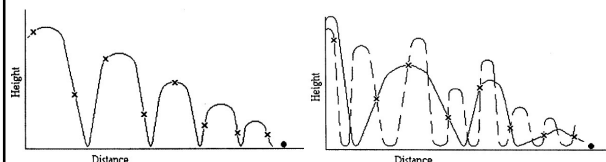
Exaktní a aproximující metody interpolace



Spojitě a zlomové metody interpolace



Deterministické a stochastické metody interpolace

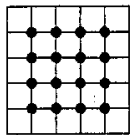


What Is the Difference Between Stochastic and Deterministic Models?

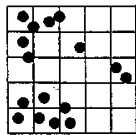
Unlike deterministic models that produce the same exact results for a particular set of inputs, stochastic models are the opposite; the model presents data and predicts outcomes that account for certain levels of unpredictability or randomness.

Výběr reprezentativních vzorků (sampling)

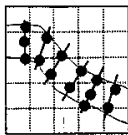
Je důležitý pro výběr interpolačního algoritmu, úspěšnost vlastní interpolace a pro validaci výsledků



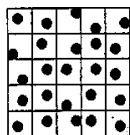
a) regular sampling



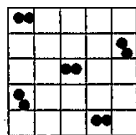
b) random sampling



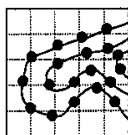
c) transect sampling



d) stratified random sampling



e) cluster sampling



f) contour sampling

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

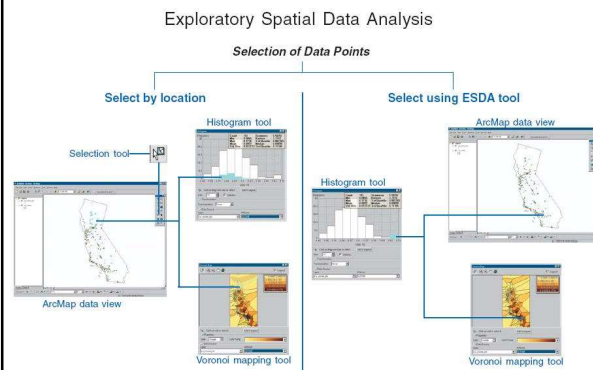
- existence dostatečně reprezentativního vzorku měřených dat
- vhodné vlastnosti měřené veličiny a typ dat (ordinální, intervalová, poměrová)
- teoretické i empirické znalosti o povaze prostorové diferenciace studovaného jevu
- znalost podstaty použitelných interpolačních metod
- znalost způsobu výběru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

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).



Základní postupy průzkumové analýzy prostorových dat

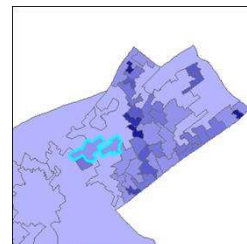
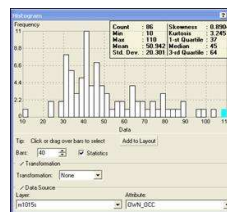
- výpočet základní popisné statistiky včetně momentů vyššího řádu (asymetrie a špičatosti)
- prověření požadavků normality a stacionarity
- analýza rozdělení hodnot - analýza histogramu
- analýza kvantilového grafu (Q-Q grafu)



- zkoumání odlehklých 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

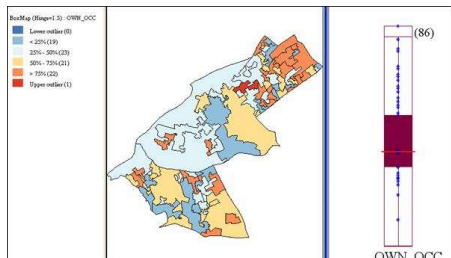


Hodnocení polohy a prostorového uspořádání typických resp. extrémních hodnot.

Základní nástroje průzkumové analýzy prostorových dat

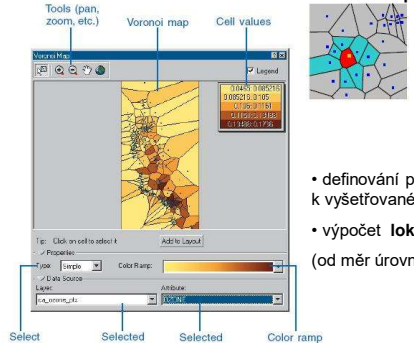
Krabicové grafy (box plots)

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



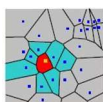
Základní nástroje průzkumové analýzy prostorových dat

Voronoi map



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

Základní nástroje průzkumové analýzy prostorových dat



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

$$Entropie = -\sum p_i * \log p_i$$

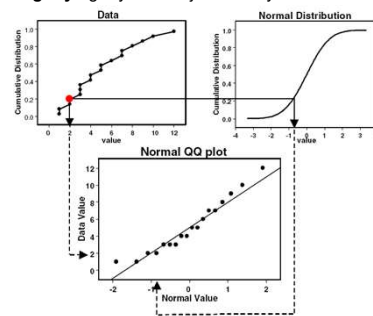
kde p_i je poměr polygonů náležejících do dané třídy z celkového počtu polygonů

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

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

Základní nástroje průzkumové analýzy prostorových dat

Kvantilové grafy - grafy zobrazující kvantily dvou rozdělení

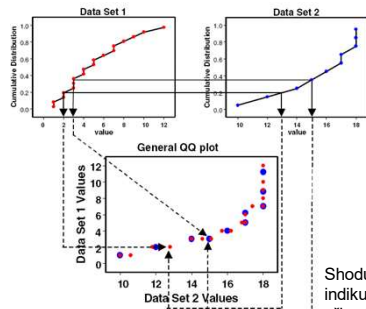


Normální Q-Q graf

Slouží jako nástroj k posouzení normality vstupních dat.

Základní nástroje průzkumové analýzy prostorových dat

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

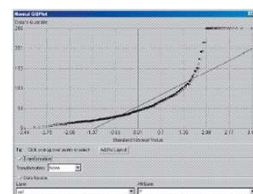


Shodu v obou případech indikují v grafech body přibližující se k přímce.

Základní nástroje průzkumové analýzy prostorových dat

Transformace dat

Data, která se odchylují od požadovaného rozdělení vyžadují transformaci



Základní typy transformací:

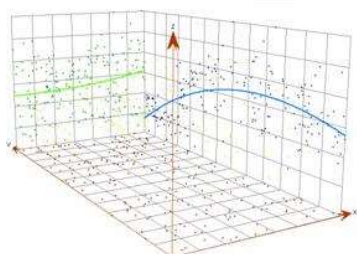
- Logaritmická
- Box-Cox (mocnná)
- Arcsine

Základní nástroje průzkumové analýzy prostorových dat

Analýza trendu

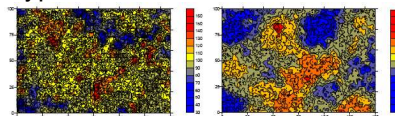
Definování globálního trendu v datech, jeho odhalení a eventuálního odstranění.

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

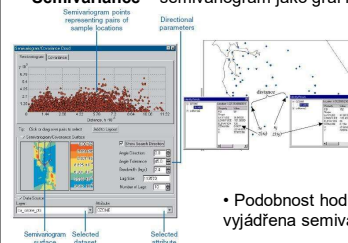


Základní nástroje průzkumové analýzy prostorových dat

Detekce míry prostorové autokorelace



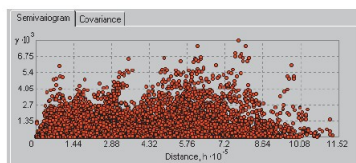
- **Semivariance** – semivariogram jako graf míry nepodobnosti.



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

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

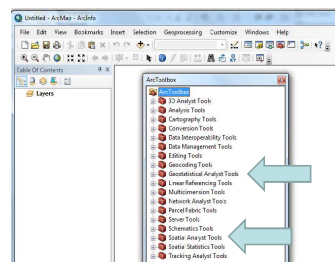
Strukturní analýza jako nástroj průzkumové analýzy prostorových dat



V explorační analýze slouží strukturní analýza především k

- detekování míry prostorové autokorelace
- vystižení míry anizotropie
- odhalení odlehklých hodnot

SW nástroje průzkumové analýzy a geostatistických metod



- R gstat, sp, ..., spatstat (point pattern analysis)
- Geoda <https://spatial.uchicago.edu/geoda>
- CrimeStat <https://www.icpsr.umich.edu/CrimeStat/>