

Geostatistické zpracování dat v programovém produktu ISATIS

Jiří Paloch

Institut Geoinformatiky, Hornicko-geologická fakulta, Vysoká škola báňská - Technická
univerzita Ostrava, 17. Listopadu 15
708 33, Ostrava, Česká Republika
jiripaloch@email.cz

Abstrakt. Tato bakalářská práce demonstruje na příkladu postupy geostatistického lineárního modelování. Rozsah zkoumání a analýz pokrývá celou kapitolu předmětu Prostorové analýzy dat vyučovaného na VŠB-TU Ostrava. V současné době probíhá výuka v programu GeoEAS, který by měl ISATIS v budoucnu nahradit. Bakalářská práce je rozdělena do několika částí, jež například popisují teorii geostatistických analýz, data a práci v prostředí ISATIS. Dále je pak prezentován návod, který bude sloužit jako vodítko pro samostatnou práci studentů. Na závěr jsou porovnány výsledky modelování provedených v dalších programech umožňujících geostatistické zpracování dat.

Klíčová slova: geostatistika, návod, ISATIS, modelování, znečištění

Abstract. This work demonstrates geostatistics linear modeling procedures on example. The coverage of investigation and analysis extending the whole chapter of spatial data analysis course, which is taught at the Technical University of Ostrava. Current education is realized by software GeoEAS which should be replaced by ISATIS. The thesis is divided into several parts, for example, to describe the theory of geostatistical analysis, data and work in an ISATIS environment. Further is presented a guide which will serve as a lead for separate work of students. In conclusion are compared the results of other programs which are allowed for geostatistical data procesing.

Keywords: geostatistics, guide, ISATIS, modeling, polution

1 Úvod

Tato bakalářská práce je zaměřena na popis, jak zacházet s některými nástroji programu ISATIS a seznámit čtenáře s lineární geostatistikou, postupy tvorby a zpracování digitálního modelu znečištění půdy. Dále jsou prezentovány výsledky z dalších dvou různých programových produktů, jejich hodnocení a hodnocení uživatelského prostředí a možností těchto produktů.

První část této práce se věnuje vybraným matematickým charakteristikám a geostatistice. Bez jejich znalosti není možné porozumět k čemu samotné nástroje ISATISu slouží, natož např. výsledkům analýz. Další část je věnována představení použitého software při tvorbě této práce a ukázce dat. Data ukazují obsahy jedovatých kovů v místech vrtů na blíže neurčeném místě. Poslední částí je podrobný popis

jednotlivých kroků, vedoucích k vytvoření digitálního modelu znečištění půdy a prezentace výsledků této bakalářské práce.

2 Cíle práce

Cílem této práce je vytvoření návodu k potřebám cvičení z předmětu prostorové analýzy dat. Konkrétně se to týká podkapitoly geostatistika, kde jsou cvičení prováděna v dnes již archaickém produktu GeoEAS a který by měl ISATIS ve výuce nahradit. Části návodu obsahují, seznámení se s pracovním prostředím, exploratorní analýzu dat, transformaci dat, výběr vhodného teoretického modelu a vytvoření digitálního modelu.

Porovnání digitálních modelů půdního znečištění vytvořených v produktech GeoEAS, ISATIS a ArcGIS Geostatistical Analyst a porovnání možností těchto produktů .

3 Teoretická část práce

Zde jsou čtenáři seznámeni se základy matematické statistiky, které zahrnují:

- Základní pojmy - Statistický soubor s jedním argumentem
- Charakteristiky polohy, variability, šikmosti, špičatosti
- Vhodná rozdělení statistické distribuce sledované veličiny
- Základní pojmy - induktivní statistika

Dále je vysvětleno co je geostatistika, vysvětlení pojmů a jaké nástroje zahrnuje:

- Strukturální funkce
- Teoretické modely semivariogramu
- Interpretace teoretického modelu semivariogramu
- Lokální odhad - krigování

4 Data

Data jsou uložena ve formátu *.xls a obsahují naměřené hodnoty jedovatých kovů z vrtů na bližší neurčeném místě. Jedovatými kovy jsou Arzen, Kadmium a Olovo v jednotkách ppm (part per million).

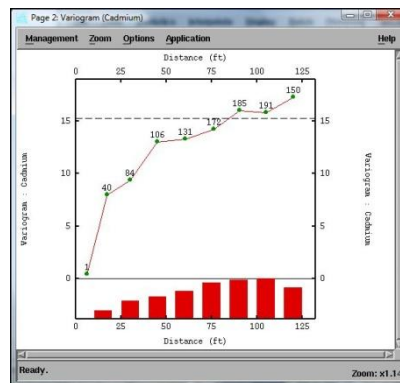
5 Praktická část práce

Tato část je věnována praktickému postupu, který v první části seznamuje čtenáře s prostředím programu ISATIS. Druhá část je věnována popisu dat prostřednictvím matematické statistiky a exploratorní analýzy. Ve třetí části je vybrán vzorek dat arzenu a ten je transformován do NNR. Čtvrtá část zahrnuje korelační a regresní

analýzu, kde je porovnáváno rozložení hodnot mezi kadmíem a arzenem před a po transformaci do NNR. Pátá část obsahuje Provedení strukturální analýzy nad daty arzenu a šestá část se věnuje bumerangovému testu (Cross Validation). Závěrečným krokem je provedení lokálního odhadu - krigování. Z důvodu omezeného prostoru v příspěvku se tato část věnuje pouze částem strukturální analýza, bumerangový test a krigování.

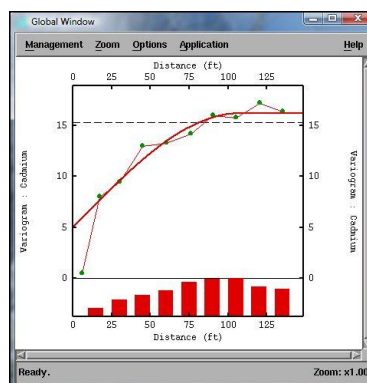
5.1 Strukturální analýza

Je provedena nad hodnotami obsahu kadmia ve studovaném území. Nejprve je vypočten experimentální všesměrný semivariogram s velikostí kroku 15ft, počtem kroků 9 a dosahem 135ft. Výsledek na obr. 1.



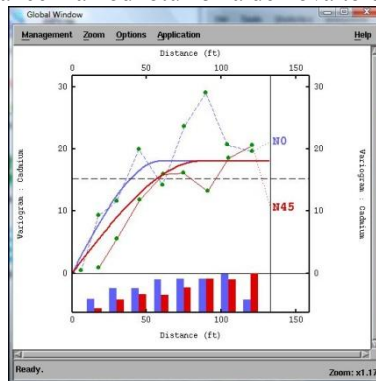
Obrázek 1: Experimentální semivariogram hodnot Cd

Dále je nad tímto semivariogramem vypočten teoretický model s hodnotami: dosah 107ft, práh 11.2 a zbytkový rozptyl 5.

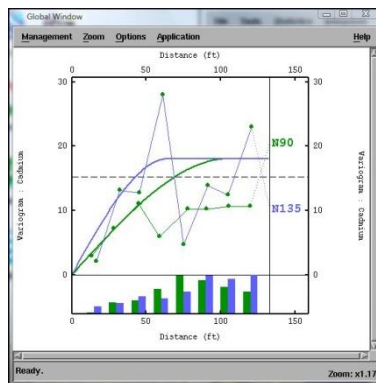


Obrázek 2: Experimentální semivariogram se sférickým teoretickým modelem

Dalším krokem je prověření možnosti anizotropie v poli pomocí směrových semivariogramů. Zde bylo pro výpočet použito parametrů všesměrného semivariogramu, tedy: počet kroků 9, délky 15ft. Byly voleny pouze významné směry, tedy 4, tolerance kroku byla nastavena na hodnotu 1 a min. počet párů bodů na buňku na 1. Úhlová tolerance má hodnotu 15° a délková tolerance 50%.



Obrázek 3: Zobrazení experimentálních směrových semivariogramů N0 a N45 a teoretických modelů



Obrázek 4: Zobrazení experimentálních směrových semivariogramů N90 a N135 a teoretických modelů

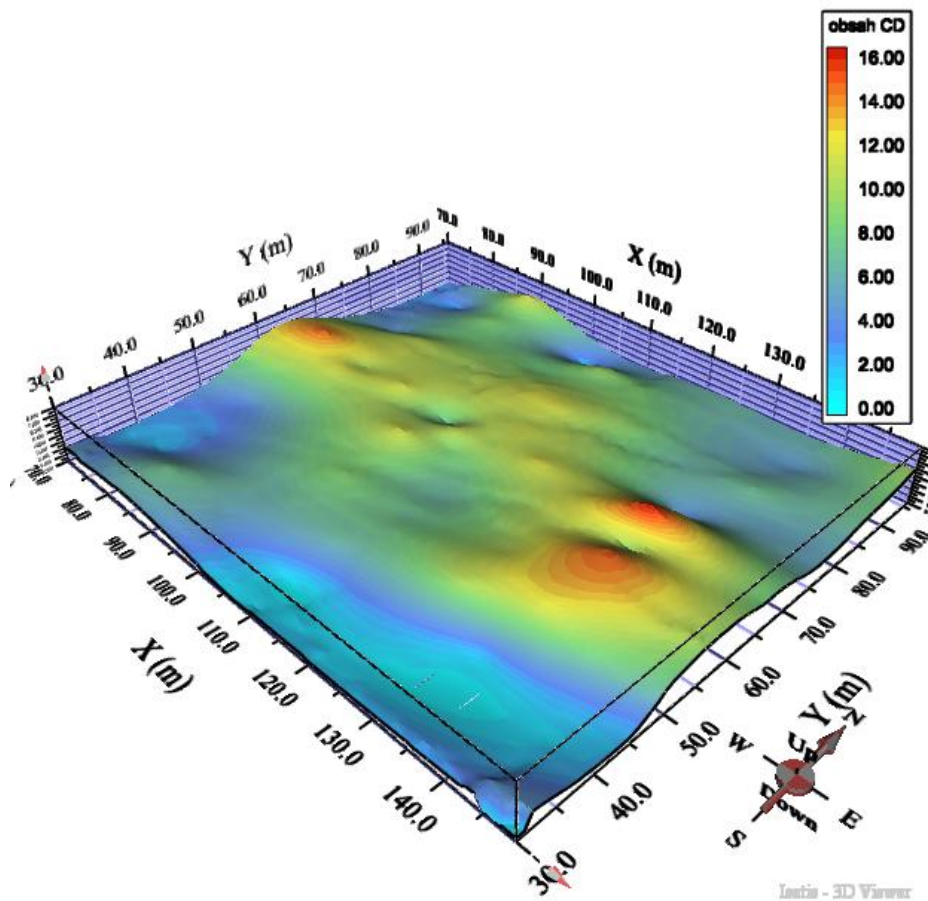
Podle dosahů, SJ 58ft a VZ 114ft s natočením 15° od kladné osy Y, je v poli výrazná anizotropie, výsledky obou strukturálních analýz jsou podrobeny bumerangovému testu v další podkapitole.

5.2 Bumerangový test - Cross Validation

Testům byly podrobeny modely všesměrného a směrového teoretického modelu, lepší výsledky pak dával směrový teoretický model, který zahrnoval anizotropii. Pro výběr okolí dávalo nejlepší výsledek nastavení elipsy s hodnotami dosahů anizotropie bez natočení.

5.3 Lokální odhad - krigování

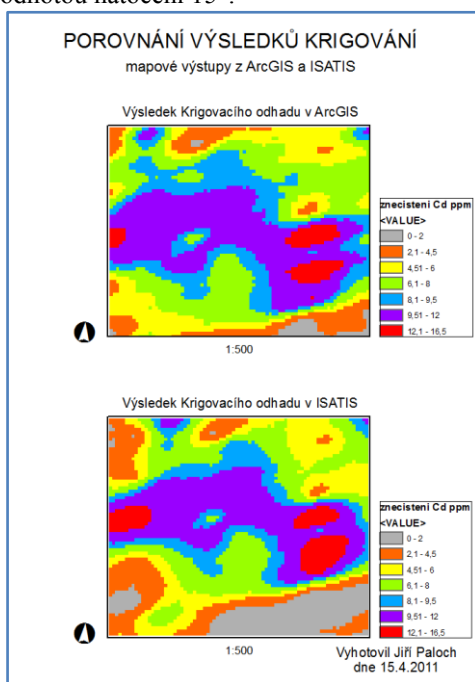
Výsledek lokálního odhadu.



Obrázek 5: Výsledek odhadu vývoje hodnoty Cd(ppm) v poli

6 Porovnání výsledků lokálního odhadu a možností ISATIS, ArcGIS a Geo-EAS

V programu ArcGIS byl rovněž proveden lokální odhad hodnot Cd v půdě. K tomu sloužila extenze Geostatistical Analyst, kde byly při výpočtech zadávány, do jisté míry, stejné vstupní parametry, jako při odhadu v ISATIS: Podle výsledků Cross-Validation vyšel nejlépe sférický teoretický model s dosahem anizotropie VZ 168 stop a SJ 74 stop s hodnotou natočení 15°.

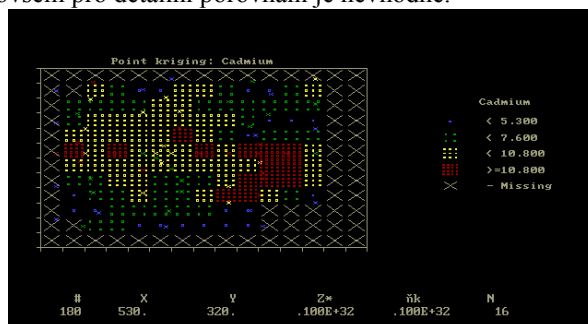


Obrázek 6: Vizualizace výsledků lokálních odhadů v prostředí ArcGIS

Pro vizualizaci byl záměrně volen grid s velikostí buňky 2,97stopy, aby tak lépe vynikly rozdíly v odhadech. Barevná škála byla zvolena ze stejného důvodu. Rozdíl v JV rohu je zřetelný na první pohled. Zatímco v ArcGIS jsou v této části zastoupeny intervaly hodnot až po interval 6,1-8 (ppm), ISATIS zde odhadnul hodnoty v rozmezí 0-2(ppm). Tato nestabilita je zapříčiněna, jednak nedostatečným počtem vzorků v této oblasti a také poněkud vyššími hodnotami dosahu anizotropie v ArcGIS.

Velmi podobné rozdíly je možné pozorovat také v JZ a SZ rohu studované oblasti. I zde jsou, opět ve větší míře, zastoupeny nízké hodnoty na odhadu z ISATIS. Oproti tomu ve středu studovaného území je vývoj hodnot velmi podobný u obou z výše jmenovaných odhadů.

Pokud se srovnají předešlé odhady vývoje obsahu Cd v poli s vizualizací v prostředí Geo-EAS, tak je na první pohled zřetelná podoba ve vývoji hodnot v pásu táhnoucím se ve směru VZ. Toto schéma může dobře posloužit pro hrubou představu o vývoji hodnot v poli, ovšem pro detailní porovnání je nevhodné.



Obrázek 7: Vizualizace krigovacího odhadu v prostředí GeoEAS

Z hlediska **uživatelsky přívětivého rozhraní (Human-Computer Interface)** není jednoduché tyto tři specializované programy srovnat. Každý má svá specifika, pokud půjde o srovnání na základě přívětivosti k uživateli, zde bude na prvním místě pravděpodobně ArcGIS. Zde může prakticky kdokoliv vytvořit např. digitální půdní model pomocí 10-ti kliknutí myši, v takovém případě tomu bude samozřejmě odpovídat kvalita výsledku. Pokud je tedy zanedbána kvalita lokálního odhadu a jde jen o náročnost na uživatele při této operaci, je na druhém místě Geo-EAS. ISATIS pak klade největší nároky na seznámení se uživatele s jeho prostředím a pochopení hierarchie procesů vedoucích k výsledku lokálního odhadu.

Ovládání je z pohledu dnešní doby nejsložitější v programu Geo-EAS, to mu však neubírá na funkčnosti. Pracovní prostředí z r. 1991, stejně jako např. v textovém editoru T602, je pro dnešní uživatele nepraktické. GUI (Graphical user interface) jsou v ISATIS i ArcGIS na velmi vysoké úrovni, ikdyž modely řízení jsou pojaty rozličně. Lze říci, že ISATIS nabízí širší možnosti nastavení a tak může ArcGIS vypadat, např. při exploratorní analýze, méně náročnější na ovládání.

Počet **datových formátů** podporovaných v jednotlivých programech se značně liší. S nejvyšším počtem vyčnívá ArcGIS, ovšem ISATIS a Geo-EAS jsou oproti němu specializované programy v oblasti geostatistiky. Všechny tři programy pak mají společný pouze jediný formát a to ASCII.

Tabulka 1. přehled software a počtů formátů se kterými operují

Software	Čtení	Import	Export
ISATIS	20	20	18
Geo-EAS	10	10	10
ArcGIS	94	97	76

7 Závěr

Cílem této práce bylo seznámit se s prostředím ISATIS a vytvořit půdní digitální model nad daty znečištění půdy. Dále vytvořit a porovnat modely z programových produktů GeoEAS a ArcGIS nad stejnými daty a shodnotit možnosti těchto produktů. První teoretická část práce byla zaměřena na vybrané matematické statistické charakteristiky a geostatistiku, kde jsou vysvětleny postupy a metody aplikované v další části této práce.

Druhá část seznamuje čtenáře s použitým software při tvorbě této práce a demonstruje data, která byla vybraná pro tuto práci.

Třetí část byla věnována zpracování dat v programovém produktu ISATIS 2011, kde byly použity postupy, pojmy a metody z matematické statistiky a geostatistiky. Toto zpracování je doplněno vysvětlujícím textem, tak aby mohlo čtenářům sloužit jako vodítko při samostatné práci. V posledním kroku byly graficky vizualizovány také výsledky z programových produktů Geo-EAS a ArcGIS a bylo provedeno porovnání možností těchto produktů z uživatelského hlediska.

8 POUŽITÁ LITERATURA A INTERNETOVÉ ZDROJE

- [1] PRAVDĚPODOBNOST A STATISTIKA [online]. 2010 [cit. 2010-04-14]. Esf. Dostupné z WWW: <<http://homen.vsb.cz/~oti73/cdpast1/>>
- [2] Horák, J.: Prostorové analýzy dat. *Skriptum, VŠB-TU Ostrava, 2006, 151 s.*, Dostupné na [www:<http://gislinb.vsb.cz/~hor10/PAD/Skripta/PAD_skriptaF.pdf>](http://gislinb.vsb.cz/~hor10/PAD/Skripta/PAD_skriptaF.pdf)
- [3] Vizi, L. Kovanič, L...: Geoštatistický prístup k modelovaniu priestorovo-časových údajov na príklade prejavov banskej činnosti na ložisku Solivary, *Akta Montanistica Slovaca, ročník 13, 2008, číslo 3, 14 s.*
- [4] Schejbal, C.: Matematické metody a postupy ložiskového průzkumu. In *Sb.vědeckých prací VŠB – TU Ostrava, monografie 5, rok 2001, roč. XLVII, řada hornicko – geologická, 222 s.*
- [5] ŠEDIVÁ, Blanka. *Pravděpodobnost a statistika pro FEL* [online]. Plzeň : 2004-2005 [cit. 2010-05-08]. Regresní a korelační analýza, s. . Dostupné z WWW: <<http://home.zcu.cz/~sediva/>>
- [6] Schejbal, C.: Aplikovaná geostatistika III. *Příbram, GŘ ČSUP, 1985, 207 s.*
- [7] Schejbal, C.: Úvod do geostatistiky. *Skriptum, VŠB-TU Ostrava, 1996, 52 s.*
- [8] *Geovariances* [online]. 2010 [cit. 2010-05-01]. Geovariances - The most complete solution in geostatistics. Dostupné z WWW: <<http://www.geovariances.com/en>>
- [9] *US Environmental Protection Agency* [online]. 2010 [cit. 2010-04-21]. EPA GWERD Software - GEOEAS for CSMoS Modeling Software. Dostupné z WWW: <<http://www.epa.gov/ada/csmos/models/geoeas.html>>
- [10] *ARCDATA PRAHA* [online]. 2010 [cit. 2010-04-21]. Geografické informační systémy - ARCDATA PRAHA. Dostupné z WWW: <<http://www.arcdata.cz>>
- [11] *Root.cz* [online]. 2010 [cit. 2010-04-21]. Licence - Public domain - Root.cz. Dostupné z WWW: <<http://www.root.cz/specialy/licence/public-domain/>>
- [12] *ESRI - The GIS Software Leader* [online]. 2010 [cit. 2010-04-22]. ESRI Info. Dostupné z WWW: <<http://www.esri.com/about-esri/index.html>>
- [13] ISATIS Case Studies 2011
- [14] Armstrong, M.: Basic Linear Geostatistics. Berlin, Springer, 1998, 154s.

- [15] Vizi, L.: Regionálne zhodnotenie ložiskových vzťahov rudných polí Spišsko-Gemerského Rudohoria za použitia geoinformatiky [Disertační práce], Technická Universita v Košiciach, 2002
- [16] Staněk, F.: Geostatistika_přednáška_3, *výukový materiál, VŠB-TU Ostrava, 2007, 10s.*
- [17] Staněk, F.: Tvorba modelu ložiska uhlí a způsoby jeho hodnocení. Sb. vědeckých prací VŠB – TU Ostrava, rok 2005, roč. LI, řada hornicko - geologická, monografie 14, 125 s.