

VYUŽITÍ STANOVENÍ PŘESNOSTI DIGITÁLNÍCH DAT K MODELOVÁNÍ PŘÍRODNÍCH JEVŮ V SOFTWARE R

(The error estimation used for the modelling of the natural hazards with help of software R)

Lukáš MAREK

Katedra geoinformatiky, Přírodovědecká fakulta,
Univerzita Palackého v Olomouci,
tř. Svobody 26, 771 46, Olomouc, Česká republika
lukas.mar@seznam.cz

Abstrakt: Úkolem této práce je napsat pojednání o přesnosti dat získaných měření pomocí digitálních měřících přístrojů a využití těchto zpřesněných dat k modelování přírodních jevů v open-source statistickém softwaru R. Měření bylo prováděno pomocí totální stanice a také pomocí GPS, tvorba datasetů a jejich zpracování potom v programu R. Vzhledem k tomu, že měření byla prováděna na území, které bývá postiženo sesuvy, tak se v závěr práce zabývá aplikací modelu vytvořeném v R na tento druh svahového pochodu.

Klíčová slova: Klíčová slova: software R, (geo)statistika, konfidenční elipsa, modely, sesuv

Abstract: Main deal of this bachelor's thesis is to write an essay on evaluation of precision of datas which was measured by digital measuring devices and application of these more accurate datas to simulate natural phenomenons in open source statistical software R. Measurement was done by total station as well as by GPS. Creation of datasets and all of analyses in statistical software R. Because of situation that all fo measurments was done on the area where is possibility of landslides, the end of thesis considers application of model created in software R on this kind of slumping.

Keywords: software R, (geo)statistics, confidence ellipses, model, landslide

SEZNÁMENÍ S R

R je ucelený balík programového vybavení pro manipulaci, výpočty a grafické zobrazení digitálních dat.

Zahrnuje:

- efektivní zacházení s daty a nástroje pro jejich uchování,
- balík operátorů pro výpočty řad a matic,
- obsáhlou, srozumitelnou a celistvou kolekci nástrojů pro analýzu dat
- grafické prostředky pro analýzu dat a jejich zobrazení na obrazovce či přípravu k tisku
- vyvinutý, jednoduchý a efektivní programovací jazyk, který obsahuje podmínky, cykly, uživatelem definované rekurzivní funkce a vstupní a výstupní prostředky.

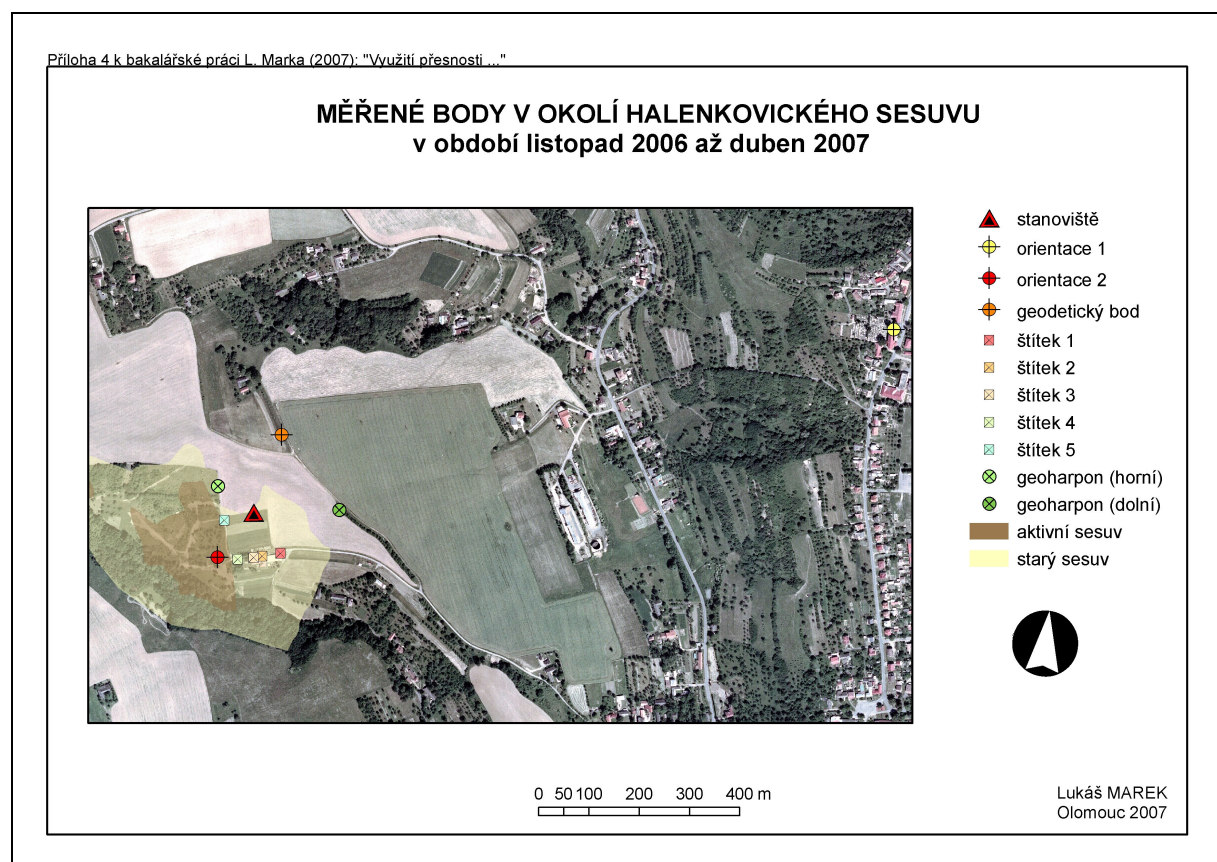
CÍLE PRÁCE

- napsat pojednání o přesnosti dat získaných digitálními měřícími přístroji

- využití modelů ve statistickém R, které mohou být využity pro zpřesňování dat získaných digitálními měřicími přístroji
- Návrh metodiky postupu zpracování měření GPS ze statistického hlediska (tzn. určení co nejpřesnějšího odhadu chyby budoucího měření)
- analýza možného sesuvu a možné statické poruchy na sledovaných objektech

ÚZEMÍ

Všechna měření byla prováděna na území nedaleko obce Halenkovice, která se nachází v západní části bývalého okresu Zlín asi 4 km od Napajedel. Toto území je oblastí náchylnou k sesuvům. Samotná měření se konala na území starého, pravděpodobně již neaktivního sesuvného území. Z toho vyplývá, že v závěru mé práce se vyskytuje také pojednání o tom, zda je sesuv opravdu neaktivní či nikoliv.



Obr.1 Rozmístění měřených bodů

POSTUP A METODY

Nejdříve bylo nutné v terénu zanalyzovat situaci, aby mohl být stanoven návrh postupu měření, tzn. počet měřených bodů, jejich poloha, počet samotných měření bodu.

Po zhodnocení situace probíhal sběr dat v terénu (měření listopad, leden, únor, březen, duben) a to jak pomocí totální stanice (všechny body), tak pomocí GPS (body umístěné v zemi).

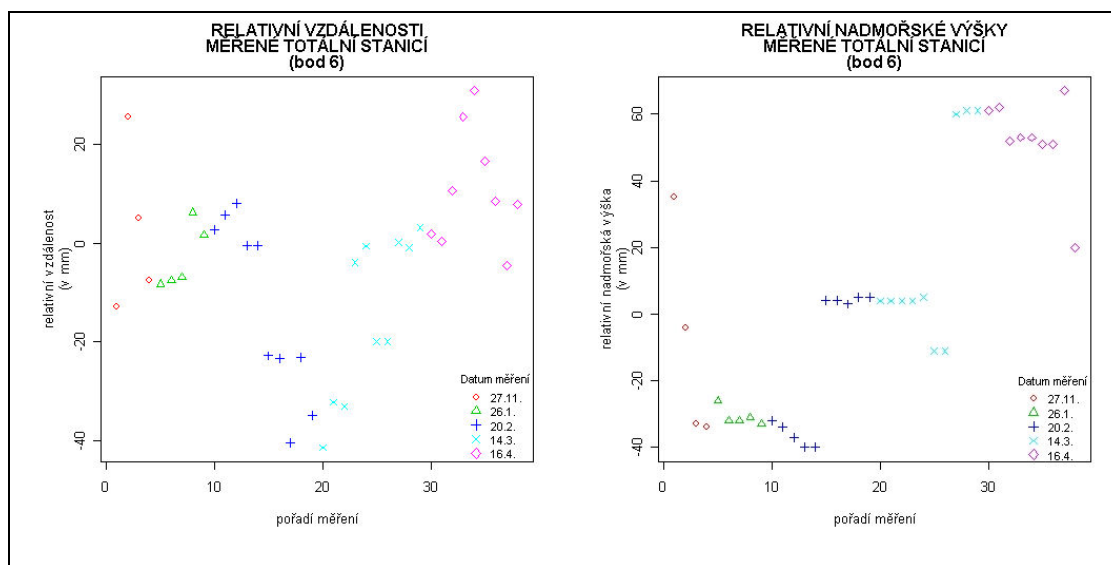
Jakmile byla data naměřena muselo proběhnout jejich stažení z mobilních zařízení do PC, aby mohla být provedena jejich konverze z interního formátu daného přístroje do požadovaného

formátu, ve kterém byly body z obou přístrojů v jednotném formátu, který obsahoval souřadnice bodů v souřadnicovém systému S-JTSK.

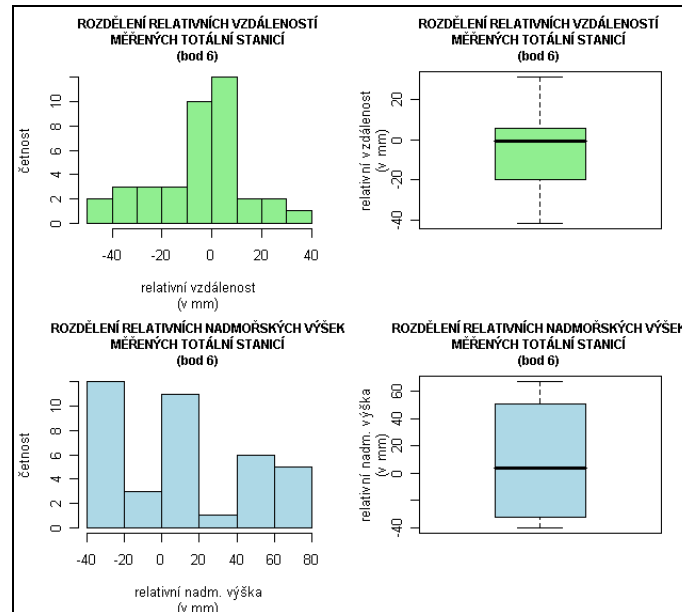
Po získání dat z přístrojů a jejich převodu do jednotného formátu byly sestaveny datasety měření pro jednotlivé dny a přístroje zvlášť, aby mohla být provedena jejich následné statistické vyhodnocení (tzn. základní numerické charakteristiky jednotlivých datasetů a jejich grafická podoba). Posledním krokem bylo potom modelování v R, kdy se jednotlivé datasety měřených bodů nechávaly vstoupit do vytvořeného modelu v R.

Tab. 1 Základní statistické charakteristiky relativní vzdálenosti a relativní nadmořské výšky bodu 6

	Relativní vzdálenost		Relativní nadmořská výška	
	GPS (m)	Totální stanice (mm)	GPS (m)	Totální stanice (mm)
Aritmetický průměr	0,002	-4,865	0,000	7,868
Maximum	39023,489	30,802	91499,010	67,000
Minimum	-29033,945	-41,323	-65005,990	-40,000
Medián	-192,600	-0,560	-693,990	4,000
1. kvartil	-1533,541	-18,133	-2422,990	-31,750
3. kvartil	977,947	5,470	1482,010	51,000
Koef. asymetrie	3,055	-0,293	4,056	0,271
Koef. špičatosti	42,366	-0,435	47,582	-1,429
Standardní odchylka	3786,597	17,675	8326,099	36,797
Rozptyl	14338314,187	312,410	69323920,717	1354,009
Kvartilová odchylka	2511,488	23,604	3905,000	82,750



Obr. 2 Poloha a nadmořská výška měření bodu 6 totální stanicí



Obř. 3 Rozdělení relativních vzdáleností a výřek měření bodu 6 totální stanicí

TEORIE A MODEL

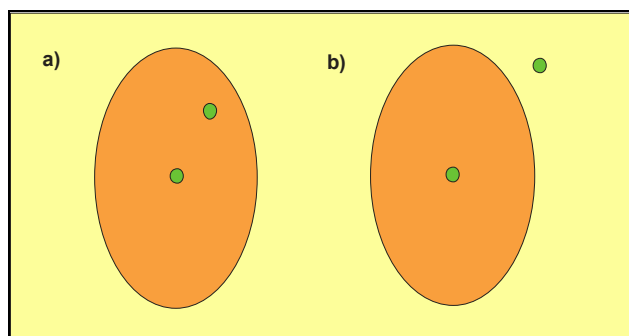
Celá práce vychází ze zcela jednoduchého předpokladu. Základem je vícenásobné měření bodu z jednoho stanoviřte v několika obdobích. (V mém případě bylo měření provedeno v 5 různých dnech v průběhu měsíců listopad až duben, kdy při každé návštěvě území byly jednotlivé body zaměřeny desetkrát.)

Potom se při zobrazení bodů v plořných souřadnicích vykreslují konfidenční elipsoidy kolem bodů. Tyto elipsoidy určují, zda se jedná pouze o nepřesnost měření, či již jde o posun měřeného bodu.

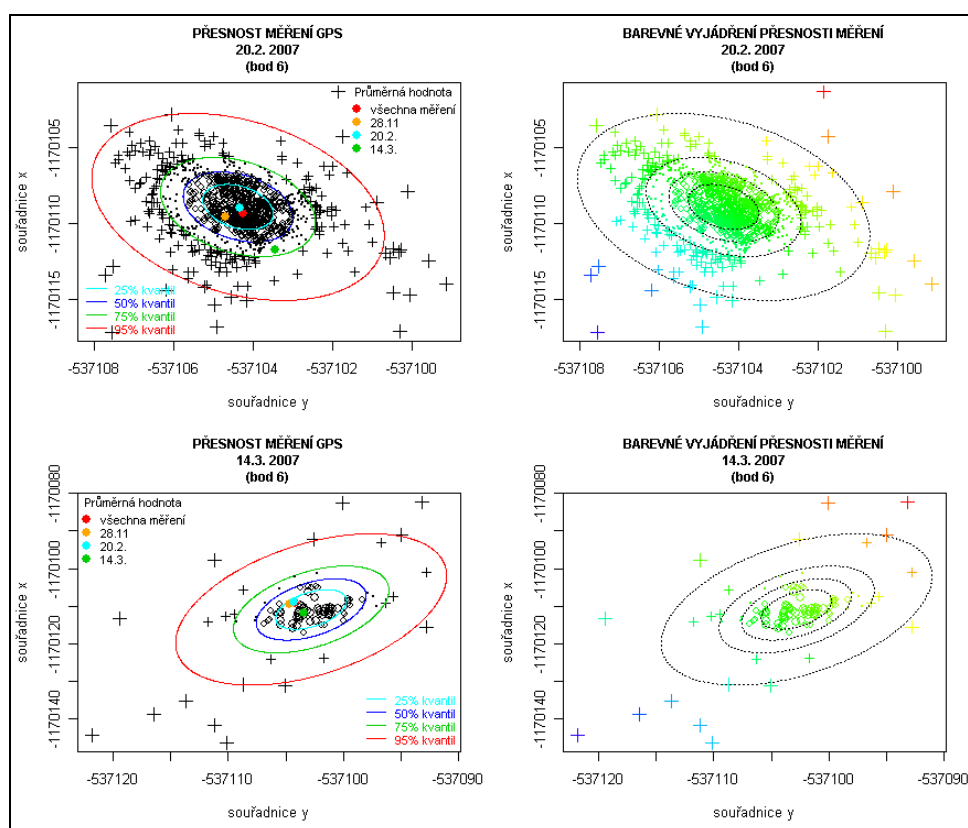
Při tvorbě elipsoidů je ovšem nutné data nejdřívě "očistit" od outlierů, tedy hodnot měření, které jsou jednoznačně chybné a mohou negativně ovlivnit další práci. Tento postup může být použit pouze pro plořné souřadnice.

Pro zjiřtění přesnosti či nepřesnosti měřených dat se využívá model implementovaný v R. Tyto implementované algoritmy sestavují konfidenční elipsy nad prostorovými daty na základě matice souřadnic, počtu a kvality vstupních dat a výpočtu vzdálenosti založené na Mahalanobisově mtrice, což není klasická Eukleidovská metrika, nýbrž její statistická obdoba.

I několikacentimetrový rozdíl při měření totální stanicí lze statisticky vysvětlit jako chybu, nikoliv jako signifikantní změnu. Pouze v případě, kdy rozdíly nemohou být vysvětleny na základě směrodatné odchylky přístroje (ne certifikované, ale v daný okamžik zjiřtěné), můžu tvrdit, že se jedná o podezřelý bod.



Obr. 4 Příklady měření bodu a) obraz stabilního bodu b) podezření na posun



Obr. 5 Konfidenční elipsy naměřených dat

MOŽNÉ APLIKACE

Využitím těchto poznatků můžeme:

- stanovit stupeň ohrožení v případě protržení přehrady
- stanovit stupeň ohrožení v případě sesuvů půdy
- stanovit stupeň ohrožení v případě porušení statiky mostu
- ověření stability velkých inženýrských staveb resp. jejich deformace

LITERATURA K TÉMATU

Tuček, P., Marek J. (2005): On estimation of dispersion in GPS data processing. Austrian Journal of Statistics

Voženílek, V. a kol. (2001): Integrace GPS/GIS v geomorfologickém výzkumu. Vydavatelství Univerzity Palackého, Olomouc, UP, 186 s.

Kubáček, L., Kubáčková, L. (1987): Pravděpodobnost a statistika v geodézii, Elsevier

Použitý software: R, TerraSync, Trimble GPS Pathfinder Office, Groma v.8, ArcView 3.1 / ArcGis 9.1, CorelDRAW 11

Použité nástroje: Totální stanice Trimble 5503 Direct Reflex Standard, Trimble GPS Pathfinder Pocket, HP iPAQ Pocket PC h5400

Použitý hardware: Hewlet Packard Compaq nx 6125 (Semprom 3100+, 60 GB, 512 MB)