

Karol HUSÁR

PRESNOSŤ DIGITÁLNYCH PRIESTOROVÝCH ÚDAJOV

Husár, Karol: The Accuracy of Spatial Digital Data. Kartografické listy, 1996, 4, 12 refs.

Abstract: The aim of this contribution is to point at the factors influencing the accuracy of digital spatial data starting with the evaluation of the accuracy of their primary analogue carrier, an object of digitization, ending with form of their digital storage in a computer environment. Besides we also talk about the accuracy from the viewpoint of semantic contents of the map.

Keywords: Accuracy, spatial data, map data, digitization.

Úvod

Transformáciu analógových údajov do digitálnej, počítačovo akceptovateľnej formy, možno v súčasnosti priradiť k jednému z dominantných trendov v disciplínach, ktoré narábajú s priestorovými údajmi.

Ak sme na poli geografie, geografické údaje z hľadiska počítačového spracovania sa obyčajne delia na priestorové a nepriestorové (atribútové). Pod digitalizáciou budeme v ďalšom rozumieť proces, ktorý sa výlučne týka časti priestorových údajov.

Práve **priestorovosť** je špecifickým znakom toho, čo odlišuje geografický typ údajov a rovnako aj geografické informačné systémy od iných "negeografických" údajov, resp. od iných typov informačných systémov. Priestorovosť chápeme v tom zmysle, že daný geografický objekt alebo jav sa týka dvoj-, troj- alebo n-dimenzionálneho priestoru a že je bezprostredne zviazaný s povrchom planéty Zem. V tejto súvislosti sa hovorí o georeferencnosti údajov. To znamená, že napríklad v súvislosti s bunkou ľudského tela, knihou na polici a ani s hviezdou Alfa Centauri nemožno hovoriť o objektoch, ktoré sú predmetom geografického bádania.

Dôležitým aspektom digitálnych priestorových údajov je ich spoľahlivosť, *presnosť*. V prípade analógových údajov mapy a rovnako aj ich percepcie sa možno stretnúť s menej striktným chápaním presnosti. Na druhej strane, v prípade digitálnej priestorovej databázy je problém presnosti závažný a nemožno sa mu vyhnúť.

Presnosť digitálnych priestorových údajov nie je obmedzená iba na proces digitalizácie a následného spracovania a manipulácie v počítačovom prostredí, ale nevyhnutne je spojená aj s presnosťou jej primárneho analógového zdroja.

Cieľom tohto príspevku je pokiaľ možno prehľadnou formou poukázať na faktory, ktoré vplyvajú na presnosť digitálnych priestorových údajov, a to počnúc hodnotením presnosti ich primárneho analógového nosiča, ktorý je predmetom digitalizácie, až po formu ich digitálneho uloženia v prostredí počítača. Stručne sa dotkneme aj presnosti z hľadiska sémantického obsahu mapy. Ide teda o pomere širšie stanovený rozsah témy, ktorý z hľadiska vymedzeného priestoru nie je možné v jednom príspevku rozobrať detailnejšie. Preto na niektorých miestach problém naznačujeme iba v hrubších rysoch, bez toho, aby sme sa jeho analýze venovali podrobnejšie.

Presnosť na úrovni primárnych nosičov údajov

Polohová (geometrická) presnosť

K hlavným zdrojom údajov či už pre vektorový alebo rastrový proces digitalizácie patrí mapa a snímka diaľkového prieskumu Zeme (DPZ).

V prípade *snímky DPZ* presnosť jej údajov je primárne závislá od rozlišovacej schopnosti snímača. Pri satelitnom snímači sa pohybuje od 10 metrov (v prípade SPOT P) do 80 metrov (v prípade LANDSAT MSS), pričom platí, že dosiahnuteľná polohová presnosť je rovná polovici hodnoty rozlišovacej úrovne daného snímača, t.j. od 5 do 40 metrov (Galtier 1993). Okrem toho je potrebné brať v úvahu elimináciu centrálnej projekcie snímky DPZ. Jej odstránenie je predmetom geometrických a rádiometrických korekcií.

Najdôležitejším dátovým zdrojom na tvorbu digitálnej databázy stále zostáva *mapa*. Aj napriek niektorým svojim nedostatkom, ktoré vyplývajú z toho, že mapa reprezentuje istú abstrakciu reality a tiež z už uvedeného menej striktného chápania presnosti, analógová mapa stále zostáva jedinečným médiom priestorových údajov.

Chyby súvisiace s procesom tvorby mapy možno zjednodušene rozdeliť na chyby vznikajúce meraním priamo v teréne - ako chyby *meračské*, na chyby vyplývajúce z aplikácie určitých algoritmických postupov nad údajmi získanými z terénu - ako chyby *výpočtové* a chyby vznikajúce pri vlastnej konštrukcii mapy - ako chyby *konštrukčno-kresličské*. K týmto trom typom chýb, dotýkajúcich sa presnosti pri tvorbe mapy je potrebné zaradiť aj chyby z polygrafického spracovania a ďalšie chyby, ktoré vyplývajú z pôsobenia času a prostredia už na vy publikovanú mapu na papierovom médiu. Ide o tzv. *zrážku papiera* mapy.

V tejto súvislosti, ako aj v súvislosti s nasledovne uvádzanými typmi chýb je dôležité rozlišovať, či ide o (napr. tvorcom) *kontrolovanú* chybu, ktorá je striktné definovaná alebo o chybu *nekontrolovanú* (*náhodnú*), ktorá nie je apriori uvažovaná a ani definovaná. Nekontrolovaný typ chyby sa môže vyskytnúť v ľubovoľnom kroku procesu tvorby alebo obnovy mapy, ako aj jej ďalšieho digitálneho spracovania. Je iba prirodzenou požiadavkou, aby tento typ chýb, ktorý je náročný na identifikáciu, bol eliminovaný na minimum. Určitým špecifickým prípadom tohto nekontrolovaného druhu chýb je vyššie spomenutá *zrážka papiera* mapy. V praxi chyby vyplývajúce z tejto *zrážky papiera* možno eliminovať transformáciou pomocou identických bodov.

Je potrebné zdôrazniť, že v tomto príspevku pod chybou budeme akcentovať predovšetkým typ kontrolovanej chyby, resp. predom definovanú mieru kvality údajov.

Dôležitú úlohu z hľadiska presnosti analógových údajov mapy zohráva jej *mierka* a na ňu naviazaná rozlišovacia úroveň. Možno povedať, že mierka je mierou presnosti. Mierka o.i. určuje minimálnu veľkosť zobrazovaného mapového objektu. To znamená, že areál s rozmerom menším ako definovaná minimálna veľkosť nie je na mape zobrazený, ale je priradený k susednému areálu s najväčším plošným obsahom. Tento spôsob agregovania areálov, vyplývajúci z definovanej mierky, má svoje konzekvencie, ktoré je potrebné rešpektovať v prípade interpretácie geografických javov. Na takto vytvorený areál potom vo všeobecnosti nemožno nazeráť ako na areál s výlučnou existenciou určitého javu, ale ako na areál s dominantnou existenciou daného javu, zahŕňajúci v sebe aj v rôznej miere fragmentárne prvky inej kvality. Pritom dominancia určitej kvality môže byť definovaná pomerom napríklad od 20, limitne až do 100% plošného obsahu celého areálu.

Na rozdiel od mierky mapy a snímky DPZ, digitálny priestorový súbor údajov je v určitom zmysle od mierky nezávislý. Pod uvedenou nezávislosťou rozumieme, že ho možno graficky zobrazovať v ľubovoľných mierkových úrovniach. Je však zrejmé, že jeho grafická (aj metrická) kvalita bude rôzna. Z uvedeného vyplýva, že digitálny súbor z hľadiska kvality jeho grafického obrazu je závislý od mierky digitalizovanej mapy. V tomto zmysle sa v praxi

udomácnil nepísaný úzus, že digitálny súbor priestorových údajov by nemal byť aplikovaný na účely, v rámci ktorých mierka je väčšia ako mierka digitalizovanej mapy, t.j. ako jeho primárneho zdroja údajov.

Určitým nedostatkom z hľadiska hodnotenia presnosti máp je skutočnosť, že v tomto smere neexistuje vo svete nejaký univerzálny, všeobecne uznávaný štandard. Existujú však rôzne národné štandardy, resp. normy. U nás sa polohovou presnosťou na mapách veľkých mierok (1:200, 1:250, 1:500, 1:1000, 1:2000 a 1:5000) zaoberala Slovenská technická norma 01 3410 (1990). Daná norma na určenie polohovej presnosti vychádza zo strednej súradnicovej chyby a pre každú z veľkých mierok definuje tzv. triedu presnosti pre podrobné polohové body, nivelačné body a pre body vrstevnice.

Žiaľ, obdobná norma vzťahujúca sa na mapy stredných mierok u nás neexistuje. Treba pritom uviesť, že v prípade využívania máp ako digitálnych súborov napr. v GIS-e je mierkový rad stredných mierok dominantný, ak nie vôbec výlučný. S niektorými parciálnymi charakteristikami dotýkajúcimi sa polohovej presnosti máp stredných mierok sa možno stretnúť v Inštrukcii na tvorbu, obnovu a vydávanie Základnej mapy 1:10 000 - ďalej len ZM 10 - (1:25 000, 1:50 000, 1:100 000, 1:200 000) -) ďalej len v Inštrukcii - (SÚGK 1985a) a v Metodickom návode na tvorbu, obnovu a vydávanie Základnej mapy 1:10 000 (1:25 000, 1:50 000, 1:100 000, 1:200 000) - ďalej len Metodický návod (SUGK 1985b). Metodický návod hovorí o maximálnej odchýlke v polohopisnom obsahu (1,0 mm), maximálnej odchýlke v bodoch vnútorného rámu základnej mapy (0,3 mm v stranách a 0,4 mm v uhlopriečkach), atď. Stredná súradnicová chyba $m_{x,y}$ pre ZM 10 v Inštrukcii je 8 m a krajná hodnota medzi dĺžkou odmeranou v teréne a v tlačových podkladoch pre jednoznačne identifikovateľné body je 33 m. Ak budeme vychádzať z tejto hodnoty a ak ju budeme extrapolovať na ďalšie stredné mierky, potom maximálne polohové chyby budú:

1:10 000	33.0 m,
1:25 000	82.5 m,
1:50 000	165.0 m,
1:100 000	330.0 m,
1:200 000	660.0 m.

Uvedené hodnoty polohovej chyby bodu sú v porovnaní s inými štandardami menej presné. Tak napríklad Tobler (1988) definuje maximálnu toleranciu pre presnosť:

1:10 000	5.0 m,
1:25 000	12.5 m.
1:50 000	25.0 m,
1:100 000	50.0 m,
1:200 000	100.0 m.

Ide teda o niekoľkonásobne vyššiu polohovú presnosť ako v prípade presnosti definovanej v Inštrukcii (1985a).

Podľa Thompsona (1988) Americký US National Map Accuracy Standard Bureau (US NMASB) definuje presnosť pre dve skupiny máp: pre mapy s mierkou väčšou ako M 1:20 000 a pre mapy s mierkou menšou alebo rovnou M 1:20 000. V prvom prípade,

a) pre mapy s mierkou väčšou ako 1:20 000 určuje, že maximálne 10% bodov môže byť s chybou väčšou ako 1/50" (t.j. asi 0,5 mm),

b) pre mapy s mierkou menšou alebo rovnou 1:20 000 platí, že maximálne 10% bodov môže byť s chybou väčšou ako 1/30" (t.j. asi 0,85 mm).

To znamená, že ak zanedbáme daných do 10% bodov, presnosť pre jednotlivé stredné mierky je:

1:10 000	5.08 m,
1:25 000	21.25 m.
1:50 000	42.50 m,
1:100 000	85.00 m,
1:200 000	170.00 m.

Hodnota tolerancie pre presnosť v prípade mapy M 1:10 000 v americkom US NMSB je v porovnaní s Toblerom približne rovnaká. V prípade mierky 1:25 000 a menších mierok US NMSB je však presnosť v porovnaní s hodnotami u Toblera (1988) takmer polovičná.

Určitou nevýhodou stanovenia presnosti v prípade US NMSB je skutočnosť, že maximálna tolerancia je striktno definovaná iba pre niečo viac ako 90% bodov. Pre zvyšné body, počtom limitne do 10% (t.j. komplement k vyššie spomenutým viac ako 90% bodom), nie je tolerancia chyby definovaná a teda ani nijak obmedzená. To u používateľa môže spôsobiť istú mieru neurčitosti.

Porovnávaním viacerých štandardov na definovanie presnosti na mapách logicky vyvstáva otázka, ktorý zo štandardov je "optimálny", všeobecne akceptovateľný? Odpoveď na túto otázku nie je celkom jednoznačná. Možno opodstatnene pochybovať, či vôbec možno v prísne striktnom zmysle uskutočniť korektnú komparáciu viacerých, vo svete používaných štandardov? Tieto štandardy sú obyčajne úzko naviazané na určitý druh mapy a teda odrážajú v sebe aj určité jej kartograficko-zobrazovacie a konštrukčno-technologické špecifiká, hoci v mnohých prípadoch o tom explicitne nie je v literatúre uvedená žiadna zmienka. Zdá sa dosť prirodzené, že tvorcovia týchto štandardov či už vedome alebo nevedome vychádzali zo svojich skúseností práve s určitými, konkrétnymi druhmi máp.

Aj napriek uvedenému vráťme sa k pôvodnej otázke "optimálneho" štandardu, abstrahujúc pritom špecifiká, vyplývajúce z používania konkrétnych druhov máp. Ak hovoríme o "optimálnosti" z hľadiska polohovej presnosti na mape, potom nám ide o taký štandard, ktorý má toleranciu chyby definovanú čo najmenšiu. Z tohto pohľadu za určitý štandard možno pokladať Merchantom (1987) definovanú presnosť pre dobre definované súradnice bodov, ktorý maximálnu toleranciu chyby pre mierky 1:5000, 1:10 000 a 1:20 000 definuje:

1:5000	1.25 m,
1:10 000	2.50 m,
1:20 000	5.00 m.

Ak túto Merchantovu logiku postupnosti extrapolujeme na ďalšie stredné mierky, potom nám výjde:

1:50 000	12.5 m,
1:100 000	25.0 m,
1:200 000	50.0 m.

Uvedené Merchantove hodnoty polohovej presnosti by mohol niekto označiť za maximalistické, prípadne za teoretické, ktoré neodrážajú skutočný stav vo vydávaní veľkej väčšiny máp stredných mierok. Tento druh námietok by bol opodstatnený, pokiaľ by sa výlučne týkal mapy ako analógového nosiča údajov. Situácia je však iná, pokiaľ sa nachádzame v pozícii používateľa digitálnej priestorovej databázy, ktorá vzišla či už z mapy alebo zo snímky DPZ. V tomto prípade existuje napríklad štandardne požiadavka minimalizovať polohovú chybu v identických bodoch pri numerickej transformácii súboru údajov z lokálneho súradnicového systému digitizéra do nejakého reálneho súradnicového systému alebo napríklad požiadavka pokiaľ možno "presne" na seba naložiť dve rôzne tematické vrstvy údajov, atď. V oboch naznačených prípadoch hrá polohová presnosť údajov významnú rolu a nemožno ju podceňovať. Naopak, snahou používateľa digitálneho priestorového súboru údajov je pokiaľ možno maximalizovať mieru polohovej presnosti dát. Z tohto pohľadu má pre používateľa digitálnej databázy Merchantom definovaná presnosť svoj význam.

Presnosť tematického obsahu

V literatúre sa zriedkavejšie možno stretnúť s problematikou presnosti na úrovni tematického obsahu mapy. Vyplyva to okrem iného aj z toho, že viacerí z klasických expertov v jednotlivých vedných odboroch pri tvorbe mapových diel sa v značnej miere riadili tzv. "expertným citom", bez toho, aby tento expertný cit bol striktné a jasne definovaný.

Nemusí to ešte znamenať a ani to neznamená, že takto vzniknuté diela sú menej hodnotné alebo nekvalitné. Príkladom toho je aj Luknišova a Mazúrova mapa geomorfologických jednotiek Slovenska (1980), ktorá je všeobecne akceptovaná, a to nielen na poli geografie, ale aj širokou obcou pracovníkov prírodovedných disciplín.

Na druhej strane dnešná doba si už od tvorcov tematických máp vyžaduje jasnejšie definovanie kritérií a pravidiel, na základe ktorých je tematická mapa generovaná. Núti ich k tomu aj technológia počítačového spracovania.

Ako príklad stanovenia kritérií pri tvorbe tematického obsahu mapy uvedme definovanie presnosti dvoch amerických inštitúcií z oblasti interpretácie foriem využitia krajiny (land use) a krajinej pokrývky (land cover) a z oblasti pedologickej regionalizácie.

V prvom prípade USGS Land Use and Land Cover maps presnosť špecifikuje nasledovne (Anderson et al. 1976):

- 1) Minimálne 85% kategórií (land use/land cover) musí byť identifikovaných správne.
- 2) Rôzne kategórie by mali mať približne rovnakú úroveň presnosti.
- 3) Presnosť by mala byť udržiavaná medzi interpretátormi a časom snímania.

Tieto kritériá sú príkladom nejasného vymedzenia a zrejme aj poplatné dobe svojho vzniku. Osobitne kritériá bodu 2) a 3) sú určené pomere vágne.

Na druhej strane uvedme príklad jasne stanovených kritérií v prípade pedologickej regionalizácie USDA SCS, kde presnosť je definovaná nasledovne (SCS 1984):

- 1) V prípade, že daný pôdny typ nie je z hľadiska územného plánovania významný, do 25% pedónov môže byť iného typu než daný pôdny typ.
- 2) V prípade, že daný pôdny typ je z hľadiska územného plánovania významný, do 10% pedónov môže byť iného typu než daný pôdny typ.
- 3) Žiadny pôdny typ nemôže zaberáť viac ako 10% plošného obsahu študovaného územia.

Čas a mapa

Dôležitým faktorom z hľadiska presnosti mapy a rovnako aj snímky DPZ je čas. Vyššie sme už uviedli zrážku papiera, ktorá o.i. je aj funkciou času. Teraz spomeňme ďalší aspekt vplyvu času na presnosť mapy. Mapa ako aj snímka DPZ reprezentuje statický obraz sveta, platný výlučne k určitému časovému okamžiku. Zmeny objektov v krajine však prebiehajú kontinuálne. Obyčajne priamo úmerne s časom vzniká určitá disproporcía medzi reálnym stavom a stavom na mape, ktorú možno eliminovať jej reedíciou. Vždy však počnúc od mapovania v teréne a ďalších prípravných prác predchádzajúcich tvorbe alebo obnove mapy, cez jej publikovanie a reálne využitie používateľom uplynie určitý čas, ktorý možno skrátiť, ale nie úplne odstrániť.

Priebežná aktualizácia obsahu mapy nie je teda možná. V tomto smere jedinečnou prednosťou digitálnej bázy údajov pred mapou je skutočnosť, že aktualizácia jej stavu je možná v ľubovoľnom čase.

Nadväznosť mapových listov

S problémom nadväznosti mapových listov sa stretneme vždy vtedy, keď zájmové študované územie sa rozprestiera na viac ako jednom mapovom liste. **Možno sa dokonca**

stretnúť s názorom, že je to najväčší problém digitalizácie.

Príčiny tohto stavu vyplývajú z viacerých faktorov. Ak vylúčime chyby vyplývajúce z meračských, výpočtových a konštrukčných prác, tento stav je v menšej miere spôsobený ručným spôsobom kreslenia a zrážkou papiera. Hlavnou príčinou tohto stavu je však rôzny čas vydávania máp. Nie je výnimočnou situáciou, ak časť objektu na jednom mapovom liste nemá svoje pokračovanie na susednom mapovom liste, pritom obsahy daných máp môžu úplne korešpondovať s reálnym stavom v čase, ku ktorému boli vydané.

Riešenie tohto problému je často sofistikované, častokrát však pre bežného používateľa nie je ho možné z hľadiska presnosti dostupnými prostriedkami optimálne vyriešiť.

Na rozdiel od mapy s problémom nadváznosti sa v prípade jednotlivých snímok DPZ nestretáme. Jednak preto, že snímky nejakého územia je možné vyhotoviť buď k určitému termínu alebo behom krátkeho obdobia niekoľkých dní a jednak preto, že snímka na všetkých svojich stranách spolu so susednými snímkami vytvára pomerne širokú tzv. prekrytovú zónu. Nie je teda možné, aby sa časť objektu zobrazila na jednej snímke a zároveň aby sa nezobrazila na príslušnej susednej snímke. Vyskytuje sa tu však iný typ problému, ktorý súvisí s deformáciou z centrálného zobrazenia snímky, ktorá spôsobuje to, že ten istý objekt na jednej snímke nie je geometricky identický s tým istým objektom na druhej snímke. Tento druh deformácie sa rieši vhodnou transformáciou.

Presnosť v procese digitalizácie

Akokoľvek má zmysel hovoriť o chybách na mape, ktoré možno registrovať, brať v úvahu, hodnotiť, avšak jej používateľ je odkázaný do role jej trpného konzumenta. Na presnosť digitálnych priestorových údajov významnou mierou vplýva presnosť v procese digitalizácie, ktorý, na rozdiel od chýb v mape, je v rozhodujúcej miere pod kontrolou človeka. V prípade manuálnej vektorovej digitalizácii sám operátor rozhoduje, či a v akej miere budú k chybám analógových údajov mapy prirátané aj chyby, ktoré vzniknú pri jej digitalizácii.

Presnosť v procese vektorovej digitalizácie

Vo všeobecnosti možno povedať, že významnou mierou na presnosť v procese digitalizácie vplýva príprava grafického podkladu na digitalizáciu. Možno tvrdiť, že od kvality týchto prípravných prác, ktoré predchádzajú vlastnej digitalizácii zásadne závisí kvalita získaného digitálneho súboru údajov (Husár 1995). V rámci uvedenej prípravy je z hľadiska presnosti dôležitý najmä výber identických bodov a voľba vhodnej numerickej transformácie digitalizovaného súboru bodov z lokálneho súradnicového systému digitizéra do nejakého reálneho súradnicového systému (napríklad S-JTSK, S-42).

Pri výbere identických bodov na transformáciu sú požadované teoreticky tri body, v skutočnosti je však vhodné, brať v úvahu aspoň štyri body. Tieto body by mali byť rovnomerne rozložené, nekolineárne a ani blízke kolineárnosti. Ďalšou požiadavkou v súvislosti s výberom identických bodov je, aby primárne boli vybrané buď priamo alebo v blízkosti konvexného obalu daného študovaného územia. Elementárnym príkladom takéhoto rozloženia sú identické body zvolené v štyroch rohoch mapového listu. Za predpokladu, že poznáme dostatočne presne ich skutočné súradnice a že boli dostatočne presne zdigitalizované, možno konštatovať, že voľba štyroch identických bodov v rohoch mapového listu je pre veľkú väčšinu úloh pracujúcich s priestorovou informáciou postačujúca.

Konkrétna voľba metódy numerickej transformácie je obyčajne daná dispozíciou príslušného softveru používateľovi. Z dvoch v praktických aplikáciách najfrekventovanejších metód Helmertovej a afinnej transformácie možno z hľadiska presnosti jednoznačne preferovať metódu afinnej transformácie, ktorá zahŕňa v sebe riešenie všetkých možných druhov geometrických deformácií.

Z krokov prípravy grafického podkladu na digitalizáciu spomeňme ešte tie, ktoré majú bezprostredný vplyv na presnosť digitálnej databázy. Je to vyznačenie štartovacích bodov na jednoduchých ostrovoch, t.j. polygónoch, ktorých kontúra je spoločná s práve jedným areálom (Husár 1995) a stabilná poloha grafického podkladu na snímačej ploche digitizéra. V prípade, že táto poloha sa v priebehu digitalizácie zmení, je potrebné uskutočniť rekalibráciu.

V striktnom zmysle do presnosti dátového súboru hovorí aj rozlišovacia presnosť digitizéra. Tá sa však u väčšiny digitizérov pohybuje od 0,25 mm až po 0.025 mm, obyčajne pod hranicu grafickej presnosti mapy (0,1 mm). Preto rozlišovaciu schopnosť väčšiny digitizérov nemožno z hľadiska presnosti považovať za relevantný problém.

Ťažiskom na získanie kvalitného, polohovo postačujúco presného súboru dát je vlastná manuálna digitalizácia. Keďže v tomto prípade ide o výber diskretných bodov z nekonečného bodového poľa mapy alebo snímky, má zmysel hovoriť o reprezentatívnosti bodov, teda, aby rozsah dát bol minimálny a zároveň zodpovedajúci stanovenému účelu. Problém však vyplýva aj z vlastného pojmu reprezentatívnosť. Hoci sú známe určité zásady napríklad pri digitalizácii reliéfu Zeme (Krcho 1985) a pri niektorých iných druhoch vektorovej digitalizácie, zväčša sú pomerne všeobecné a nie dosť konkrétne. Navyiac, každý druh vektorovej digitalizácie má svoje špecifiká a aj reprezentatívnosť tu môže nadobúdať svoj špecifický obsah.

Z troch možných režimov snímania (časový, dĺžkový a voľbou operátora) preferujeme režim snímania voľbou operátora. Je však potrebné, aby operátor mal okrem základných znalostí a technickej zručnosti z vektorovej digitalizácie rovnako aj základné poznatky o sémantike predmetného digitalizovaného javu.

Väčšina chýb vznikajúcich pri vlastnej digitalizácii je náhodného charakteru. Tým je daná aj obťažnosť ich detekovania a odstraňovania. Na druhej strane však, v prípade odstraňovania polohových chýb, výrazným pomocníkom je možnosť kontroly prostredníctvom vizualizácie na obrazovke počítača a použitie k tomu určených programových produktov.

Presnosť v procese rastrovej digitalizácie

V prípade rastrovej digitalizácie, na rozdiel od manuálneho spôsobu vektorovej digitalizácie, je celý proces automatizovaný. Tým je dané, že s väčšinou chýb, ktoré vyplývajú z charakteru manuálnej vektorovej digitalizácie sa tu nestretáme. Jediným faktorom, ktorý priamo vplyva na presnosť digitálneho rastrového súboru údajov je rozlišovacia schopnosť skenovacieho zariadenia. V súčasnosti sa pohybuje asi od 400 dpi až po 2000 dpi. Treba mať však na pamäti, že s rozlišovacou schopnosťou skenera bezprostredne súvisí rozsah údajov. Tak napríklad z analogovej mapy s rozmerom 50x50 cm, ktorá je digitalizovaná skenerom s rozlišovacou schopnosťou 500 dpi (20 bodov na mm) sa vytvorí 100 miliónov pixlov, ktoré v binárnom formáte zaberajú 12,5 Mb, v sivosti 100 Mb a vo farbe až 400 Mb pamäti.

Z kartografického hľadiska podmienkou dobrej akceptovateľnosti rastrových údajov je, aby vektor (čiarový prvok) bol reprezentovaný minimálne dvoma až troma pixlami. Z uvedeného vyplýva, že pri rozlišovacej úrovni 400 dpi z kartografického hľadiska možno dobre rastrovo digitalizovať čiaru s minimálnou hrúbkou 0.15 mm až 0.2 mm).

Nedostatkom rastrovej digitalizácie okrem potreby veľkej pamäti je aj skutočnosť, že rozmer (formát) skenera v porovnaní s vektorovým digitizérom je obyčajne viac limitovaný, a to z hľadiska dostupnosti pre používateľa. Z toho potom často vyplýva nevyhnutnosť mapu na skeneri digitalizovať po segmentoch a spojiť ich následne na úrovni digitálneho súboru. K tomu je rovnako potrebné vybrať vhodné identické body, pomocou ktorých sa uskutoční numerická transformácia a následné spájanie čiastkových súborov. Tieto operá-

cie zároveň nevyhnutne vnášajú do procesu digitalizácie určité iteračné postupy, ktoré sa sebou nesú aj zvýšenie pravdepodobnosti nárastu chýb. Možno teda konštatovať, že limitovaný rozsah skenera nepriamo vplýva na presnosť digitálneho súboru priestorových údajov.

Presnosť v procese rastovo-vektorovej digitalizácie

V poslednej dobe možno badať nástup špeciálneho, kombinovaného spôsobu digitalizácie. Ide o vektorový spôsob digitalizácie grafického podkladu na obrazovke, pričom tento grafický podklad bol vopred skenerom rastovo digitalizovaný a vizualizovaný. V prípade klasickej vektorovej digitalizácie operátor musí priebeh digitalizácie kontrolovať na obrazovke a teda jeho pohľad musí striedavo registrovať grafický podklad a obrazovku počítača. Práve v tomto striedavom sústreďovaní pohľadu operátora je nebezpečie straty kontextu v pokračovaní a tým aj riziko väčšej pravdepodobnosti vzniku chýb.

Výhodou rastovo-vektorovej digitalizácie je, že na rozdiel od vektorovej digitalizácie operátor pri tomto spôsobe digitalizácie nestráca kontext postupu, pretože digitalizovaná predloha a aj kontrola zo strany operátora sa uskutočňujú na tom istom médiu - obrazovke. Je zrejme, že to nevyhnutne má dosah nielen na rýchlosť digitalizácie, ale aj na redukciu chýb.

Ďalšou nezanedbateľnou výhodou rastovo-vektorovej digitalizácie, ktorá má vplyv aj na presnosť digitalizácie, je podľa potreby možnosť zväčšovania a zmenšovania grafickej predlohy. (Samozrejme za predpokladu, že to príslušný softver umožňuje. Ide však o bežnú elementárnu operáciu počítačovej grafiky.)

Na druhej strane, nedostatkom tohto spôsobu digitalizácie je, tak ako to už bolo uvedené vyššie v súvislosti s rastovou digitalizáciou, spravidla menší rozmer skenera a potreba veľkej pamäti, ktorá však v tomto prípade môže byť iba dočasná¹⁾, do doby skončenia vektorovej digitalizácie a následnej kontroly dátového súboru.

Z toho, čo sme o vektorovej a rastovej digitalizácii doteraz uviedli, možno urobiť pre nás určité zhnutie a konštatovať, že vektorová digitalizácia je vhodná na digitalizáciu máp väčších rozmerov, presnejšie vtedy, ak rozmery študovaného územia sú menšie ako rozmery účinnej snímačej plochy digitizéra (napr. formátu A0), avšak väčšie ako rozmery disponibilného skenera (napr. formátu A4).

V prípade, ak rozmery študovaného územia presahujú rozmery účinnej snímačej plochy skenera aj digitizéra, študované územie na mape je potrebné rozdeliť na jej menšie segmenty. Uprednostnenie vektorovej alebo rastovej digitalizácie v tomto prípade môže byť závislé od porovnania množstva potrebnej práce a s tým súvisiacej potenciálnej chybovosti, a to na jednej strane v prípade spájania dielčích vektorových súborov (spravidla menšieho počtu) a na druhej strane v prípade spájania rastových súborov (spravidla väčšieho počtu).

Ak rozmery študovaného územia sú menšie ako rozmery digitizéra a skenera, je možné (za predpokladu dostatočnej rozlíšovacej úrovne skenera a existencie príslušného softveru na ďalšie spracovanie) doporučiť rastovú digitalizáciu. V prípade, že je potrebné získať vektorový digitálny súbor, je vhodné realizovať vektorovú digitalizáciu na obrazovke počítača.

Záver

Na to, aby bolo možné o nejakom digitálnom priestorovom súbore údajov povedať, či je z hľadiska presnosti údajov vyhovujúci alebo nevyhovujúci, závisí od toho, či spĺňa alebo nespĺňa špecifické požiadavky účelu, na ktorý by mal byť použitý. V konečnom dôsledku o hodnotení kvality dátového súboru rozhoduje splnenie alebo nespĺnenie očakávaní

¹⁾ O dočasný nárast potreby pamäti ide iba v prípade, ak sa rastový súbor po uskutočnení vektorovej digitalizácie na pamäťovom médiu nearchivuje.

užívateľa. Akokoľvek, takýto spôsob hodnotenia kvality digitálneho súboru je v mnohých smeroch subjektívny. Korekciu tejto subjektivity by mohla byť akceptovateľnosť dátového súboru širšou komunitou používateľov. Presnosť a chyby dát sú imanentnou súčasťou digitálnych priestorových dát. Táto skutočnosť nás však nemá demobilizovať v snahe o vylúčenie chýb všade tam, kde je to možné. Osobitne je potrebné minimalizovať vplyv náhodných, človekom nekontrolovaných chýb. Na druhej strane je potrebné presne zadefinovať presnosť a chyby vyplývajúce z mierky a rozlišovacej úrovne mapy, ako typ predom uvažovaných a kontrolovaných chýb.

Príspevok vznikol ako súčasť riešenia vedeckého projektu č. 1066 "Analýza zmien krajiny aplikáciou údajov DPZ", ktorý bol financovaný grantovou agentúrou VEGA.

LITERATÚRA

- ANDERSON, J.R., HARDY, E.E., ROACH, J.T., WITMER, R.E. (1976). A land use and land cover classification system for use with remote sensor data. Professional Paper 964. US Geological Survey, Reston.
- GALTIER, B. (1993). La télédétection aux fins de la cartographie topographique et thématique (manuscript).
- HUSÁR, K. (1994). Vektorová digitalizácia a morfometrická analýza areálov na príklade foriem land cover JZ Slovenska. Bratislava, Geografický ústav SAV, Kandidátska dizertačná práca.
- KRCHO, J. (1985). Automatizovaná tvorba vrstevníc a primárnych trojuholníkových sietí na základe Komplexného digitálneho modelu reliéfu. Záverečná správa pre Kovoprojekt Bratislava v rámci úlohy 600-77/84-VČ. Bratislava
- MAZÚR, E., LUKNIŠ, M. (1978). Regionálne geomorfologické členenie SSR. Geografický časopis, 30, 101-125.
- MERCHANT, D., C. (1987). Spatial accuracy specification for large scale topographic maps. Photogrammetric Engineering and Remote Sensing 53, 958-961.
- SCS (1984). Technical Specifications for Line Segment Digitizing of Detailed Soil Survey Maps. Government Printing Office, Washington.
- STN 01 3410 (1990). Mapy veľkých mierok. Základné a účelové mapy.
- SÚGK (1985a). Inštrukcia na tvorbu, obnovu a vydávanie Základnej mapy ČSSR 1:10 000. Bratislava.
- SÚGK (1985b). Metodický návod na tvorbu a vydávanie Základnej mapy ČSSR 1:10 000. Bratislava.
- THOMPSON, M.M. (1988). Maps for Amerika, 3rd ed. US Geological Survey, Reston.
- TOBLER, W.R. (1988). Resolution, resampling and all that. In: Mounsey H M (ed.) Building Databases for Global Science. Taylor & Francis, London, pp. 129-137.

S u m m a r y

The accuracy of spatial digital data

The primary subject of the study the situational (geometrical) accuracy of digital data and partially also thematic accuracy. We are treating the situational accuracy at several levels starting by the accuracy at the level of the primary data carriers (map, photograph of earth's surface), accuracy at the vector level, screen and combined raster-vector digitization. In case of a photograph the accuracy of its data is primarily dependent on a discerning ability of the scanner. Mistakes associated with the process of map creation can be divided in a simplified way to the mistakes originating by measuring in the field as measuring errors, mistakes caused by the application of certain algorithmic procedures to the data obtained in the field as errors in calculation and mistakes originating at the proper construction of the map - drawing errors. These three types of mistakes concerning accuracy in map creation include also the mistakes caused by the action of time and environment upon the already published map on paper medium. It is a so-called shrinkage of map paper.

An important role from the point of view of accuracy of analogue map data is played by its scale and the related level of distinguishing. Scale beside other determines the minimum size of the projected map object. Certain drawback from the point of view of evaluation of accuracy of the maps is the fact that no universal, generally accepted standard exists in the world as yet. There exist though, national standards.

Situational accuracy of the basic maps of large scales in our country is dealt with by the Czechoslovak standards. Unfortunately there are no such standards for the medium scale maps (1:10 000 - 1:200 000). The study also deals with the comparison of the situational accuracy standards of some important foreign authors and centres.

As an example of establishing criteria for the creation of the thematic contents of the map we present some defined accuracies of two American institutions in the field of land use and land cover use form interpretation and that of the pedological regionalization.

The problem of the follow-up of the single map sheets is run into when the study area is extended over more than one map sheet.

The main cause of the errors found in this area, is the different time of map publishing. It is no exception if a part of object depicted on one map sheet does not have its continuation at the neighbouring one, while the contents of the given maps can correspond to the real state in time when they were published. Solution of this problem on the level of digital data base is often sophisticated.

The accuracy of digital spatial data is much influenced by the accuracy in the process of digitizing, that in difference to the map errors is in a decisive manner under the control of man. Also the preparation of the graphic source for digitizing is important. Within the mentioned preparation the choice of identical points and of an adequate numerical transformation of the digitized point set from the local systems of coordinates of digitizer into some real system of coordinates. Of the two most in practical use most frequently used methods of transformation - Helmert and affine transformation it is possible to prefer from the point of view of accuracy the method of affine transformation.

Accuracy of the data set has to do also with the distinguishing accuracy of digitizer. That oscillates with most of them from 0.25 mm to 0.025 mm i.e. under the limit of the graphic accuracy of a map.

In case of raster digitizing in difference from the manual way of vector digitization the entire process is automated. The only factor directly influencing the accuracy of digital raster data set is the distinguishing ability of the scanning facility. At present it oscillates between 400 and 2000 dpi. It is related also to the extension of the digital data.

The drawback of the raster digitization besides the need of a large memory is also the fact that the dimension (format) of the scanner in comparison to the vector digitizer is usually more limited in the aspect of accessibility to the user. That is what often influences the need to digitize map on scanner by segments and to connect them subsequently on the level of digital set. We may then say that the limited extension of the scanner indirectly determines the accuracy of the digital set of spatial data.

In the recent time we have observed the onset of a special combined way of digitization. It is a vector way of digitizing of a graphic material on monitor while this graphic material was raster digitized beforehand by a scanner and visualized. The advantage of the raster-vector digitizing is that in difference from the vector digitizing the operator does not lose the context of the procedure as the digitized material and the control by operator are carried out at the same medium -- monitor. It is obvious that it inevitably modifies the speed of digitizing, as well as the occurrence of the errors. Another and not the least advantage of the raster-vector digitization influencing its accuracy is the possibility to enlarge or reduce the graphic material.

We can conclude that the vector digitizing is suitable for digitizing the maps of bigger dimensions, particularly, when the dimensions of the study area are smaller than the dimensions of the effective scanning area of the scanner (for instance the format A0 - 84x120 cm) but larger than the dimensions of the available scanner (for instance the format A4 - 21x30 cm).

In case the dimensions of study areas surpass the dimensions of the effective scanning area of the scanner and the digitizer, the study area will have to be divided into smaller segments.

Preference of the vector or raster digitization in this case can depend on comparison of amount of work necessary and the related potential faults either while connecting the partial vector sets (less in number, as a rule) or while connecting the raster set (more numerous).

If the dimensions of the study are smaller than the dimensions of digitizer and scanner, it is possible (provided a sufficient discerning level of the scanner and existence of a corresponding software for further processing are available) to recommend the raster digitizing. In case of the necessity to obtain vector digital set it is advisable to realize the vector digitizing on a PC monitor.

Lektoroval:
Doc. Ing. Jozef. ČIŽMÁR, CSc.,
Stavebná fakulta STU,
Bratislava