
UGOTAVLJANJE KAKOVOSTI PROSTORSKIH PODATKOV PRI DIGITALIZACIJI PROSTORSKEGA PLANA REPUBLIKE SLOVENIJE

mag. Matjaž Ivačič

MOP-Urad Republike Slovenije za prostorsko planiranje,
Ljubljana

Borut Pegan Žvokelj, Gregor Sever

Inštitut za geodezijo in fotogrametrijo FGG, Ljubljana

Prispelo za objavo: 1997-02-21

Pripravljeno za objavo: 1997-02-21

Izvleček

V članku sta opisana projekt digitalizacije kartografske dokumentacije za dolgoročni prostorski plan Republike Slovenije in postopek ugotavljanja kakovosti opravljene digitalizacije. Pri projektu sta sodelovala Urad Republike Slovenije za prostorsko planiranje in Inštitut za geodezijo in fotogrametrijo FGG, ki sta pri delu uporabila mednarodno primerljivo metodologijo za ugotavljanje kakovosti prostorskih podatkov. Rezultati projekta so pokazali, da je treba vključiti metode preverjanja kakovosti prostorskih podatkov v vse pomembnejše projekte digitalizacije.

Ključne besede: GIS, kakovost, prostorski podatki, prostorsko planiranje

Abstract:

This paper describes a method for spatial data quality determination. Two institutions were involved in the project, the Office for Physical Planning and the Institute for Geodesy and Photogrammetry. The first phase of the project was the digitalization of spatial data needed in the process of physical planning. The second phase was the spatial data quality determination. The results show that spatial data quality determination must be part of all digitalization projects.

Keywords: GIS, physical planning, quality, spatial data

1 UVOD

V prispevku je opisan projekt testiranja kakovosti prostorskih podatkov na primeru digitalizacije tematskih slojev kartografske dokumentacije za dolgoročni plan Republike Slovenije. Pri projektu sta sodelovala Urad Republike Slovenije za prostorsko planiranje, kot naročnik in Inštitut za geodezijo in fotogrametrijo FGG, kot izvajalec. Urad Republike Slovenije za prostorsko planiranje (v nadaljnjem besedilu Urad) je z Zakonom o organizaciji in delovnem področju ministrstev (Ur. l. RS, št. 71-2546/94) v 11. členu zadolžen med drugim tudi za vzpostavitev prostorskega informacijskega sistema. V okviru te naloge bo Urad vzpostavil digitalne zbirke prostorskega plana Republike Slovenije v merilu 1:250 000 in kartografske dokumentacije k temu planu v merilu 1:25 000. V kartografski dokumentaciji so strokovne službe za okolje in prostor zadolženih sektorjev na topografske karte v merilu 1:25 000 vrisale obstoječe in predvidene prostorske pojave (gozdove, kmetijske površine, območja naravne in kulturne dediščine ...). Prostorski plan je bil sprejet leta 1986 in dopolnjen leta 1989. Zadnje spremembe so bile na kartografske podlage vnešene leta 1989, zato lahko domnevamo, da nastala podatkovna zbirka opisuje stanje v Sloveniji za to leto.

Projekt digitalizacije sam ne bi bil zanimiv, če mu ne bi sledila faza preverjanja kakovosti digitalizacije. Zaradi različnih izvajalcev in različnih metod zajemanja podatkov je bilo treba ugotoviti in uskladiti natančnost postopka digitalizacije. Na podlagi opravljene analize smo dobili tudi informacije o primernosti nadaljnje uporabe posamezne podatkovne zbirke. Preverjeni so bili naslednji elementi, ki definirajo kakovost prostorskih podatkov: položajna in opisna (atributna) natančnost, logična doslednost in popolnost. Pri ugotavljanju kakovosti digitalizacije smo se omejili samo na kakovost prenosa podatkov iz analogne karte v digitalno obliko. Za celovito ugotavljanje kakovosti pa bo v prihodnosti treba ugotoviti še odstopanje med podatkovno zbirko in stanjem na terenu.

2 KAKOVOST PROSTORSKIH PODATKOV IN SMISEL NJENEGA UGOTAVLJANJA

Intenzivnejša pretvorba prostorskih podatkov v digitalno obliko se je začela šele z razvojem računalnikov, ki lahko hitreje obdelajo večje količine podatkov. S popularizacijo programskih orodij CAD in GIS je naraščalo tudi zanimanje za digitalne prostorske podatke. Cilj digitalizacije v večini primerov ni le shranjevanje, ampak predvsem analitična obdelava prostorskih podatkov. Prav pri prostorskih analizah je zato še posebej pomembna kakovost podatkov. Kakovost se razume kot zbir lastnosti proizvoda ali storitve, ki se nanašajo na sposobnost zadovoljlitve izražene ali pričakovane potrebe (SLS ISO 9001, 1992). V tem smislu je tudi kakovost prostorskih podatkov tista lastnost, ki omogoča zadovoljive prostorske analize. To je bilo potrjeno tudi v primerih digitalizacije analognih kart, ki so se v preteklosti že uporabljale, ob digitalizaciji in pri prostorskih analizah pa so se pojavile nove napake in nenatančnosti.

Kakovost prostorskih podatkov obravnavajo tudi različni standardi (SDTS, FFGDC, CEN/TC 287) za področje prostorskih podatkov, ki kakovost razčlenijo na položajno natančnost, natančnost opisnih podatkov, logično doslednost, popolnost in poreklo. Poleg teh elementov je v standardu CEN/TC 287 dodana še časovna

natančnost z informacijo o ažurnosti podatkov. Kakovost prostorskih podatkov se ugotavlja s primerjavo določene podatkovne zbirke in podatkovne zbirke večje natančnosti. Na tak način se ugotavlja natančnost kart pri Ameriškem geološkem združenju (USGS), pri standardu EMAS/ASPRS, ki v ZDA velja za topografske načrte večjih meril, in pri ugotavljanju vertikalne natančnosti digitalnega modela reliefa (Giordano et al., 1994).

V članku opisani primer se nanaša na ugotavljanje položajne natančnosti, opisne natančnosti in popolnosti. Ugotovljali smo razlike med analognim kartografskim virom in digitalno zbirko podatkov. V kolikor pa se bo izkazalo, da so digitalizirane zbirke lahko osnova pri vzpostavitvi podatkovnih zbirk za kmetijstvo, gozdarstvo, naravno dediščino in drugo, pa priporočamo predhodno ugotavljanje kakovosti prostorskih podatkov tudi v odnosu do stvarnega stanja v prostoru.

3 POTEK PROJEKTA

Projekt digitalizacije kartografske dokumentacije za dolgoročni plan Republike Slovenije je Urad Republike Slovenije za prostorsko planiranje začel izvajati spomladi leta 1995. Glavni cilj projekta je bil pridobitev primerne digitalne zbirke prostorskih podatkov za upravne in strokovne naloge. Kljub razmeroma dobri opremljenosti državnih organov z računalniško opremo večkrat primanjkujejo primerni podatki o okolju in prostoru. Prostorski podatki, ki jih vsebuje kartografska dokumentacija za prostorski plan, so pomembni tudi za ostale službe znotraj Ministrstva za okolje in prostor. Iz te dokumentacije so bile izbrane naslednje vsebine, ki so navzoče v prostoru in hkrati primerne za digitalizacijo:

- ☐ gozdarstvo
 - lesnoproizvodni gozdovi
 - varovalni gozdovi
 - gozdovi s posebnim namenom
- ☐ kmetijstvo
 - kmetijska zemljišča
 - agrooperacije
- ☐ energetika
- ☐ rudniki
 - plin
 - nafta
- ☐ naravna dediščina
- ☐ vodno gospodarstvo.

V zgornjem seznamu manjkajo še nekatere dejavnosti, navzoče v prostoru. Podatki o cestnem in železniškem omrežju so bili digitalizirani že pred tem, prav tako tudi lokacije posameznih kulturnih spomenikov. Manjkajo nam še urbanizirana območja, ki bodo predvidoma zajeta v prihodnosti. Za prenos podatkov iz analognih načrtov dejavnosti v digitalno obliko so bili izbrani trije izvajalci. Vsi med njimi so že imeli reference na področju digitalizacije različnih kartografskih materialov. Izvajalci so sami izbrali najprimernejši način digitalizacije. Nekateri so digitalizirali ročno, drugi pa so skanirali karte in jih vektorizirali. Osnovne zahteve glede natančnosti, ki so jih od naročnika dobili izvajalci, so bile naslednje:

- natančnost digitalizacije ne sme biti manjša od grafične natančnosti vsebine na topografskih kartah
- izvesti je treba spajanje robov
- vnesti je treba vse opise iz kartografske dokumentacije
- opažene napake in pomanjkljivosti je treba označiti tudi v podatkovni zbirki.

Med delom so izvajalci že opozorili na nekatere napake, ki so bile na osnovnem viru in do tedaj prezrte. Med temi so bile naslednje: poligoni brez opisnega podatka, poligon se na sosednjem listu ne nadaljuje, na sosednjem listu je drugačen opisni podatek za isti poligon.

Ker so izvajalci označili le napake, ki so jih sami zabeležili na analognem viru podatkov, za Urad pa je pomembno tudi vedenje o tem, kako natančno je bila opravljena digitalizacija, se je projekt nadaljeval s preverjanjem kakovosti digitalizacije. Z Inštitutom za geodezijo in fotogrametrijo FGG je bila sklenjena pogodba o preverjanju opravljene digitalizacije po metodologiji, ki se uporablja na področju geografskih informacijskih sistemov prav za ugotavljanje kakovosti prostorskih podatkov (Ivačič, 1996). Da bi znižali stroške preverjanja, je bilo izbranih pet listov za vsako vsebino. Stroški preverjanja kakovosti so znašali približno 10 % celotnih stroškov digitalizacije. Listi so bili izbrani po načelu geografske razčlenjenosti, in sicer po en list z območja Gorenjske, Štajerske, Notranjske in Primorske in en list z območja Ljubljane. Kljub majhnemu vzorcu so bili rezultati zelo pomembni za nadaljnje delo.

4 POSTOPEK PREVERJANJA KAKOVOSTI

Preverjanje kakovosti najlažje opišemo z modelom kakovosti, ki ga lahko opredelimo z numeričnimi, kvalitativnimi in opisnimi elementi (Pegan, 1995): položajna natančnost, višinska natančnost, pravilnost opisnih podatkov, ločljivost in popolnost, logična doslednost. Položajna in višinska natančnost sta odvisni od natančnosti vira in načina zajema. Natančnost vira je odvisna od metode kartiranja, metode fotogrametričnega izvedenja, deformacij medija, značilnosti zajetih objektov (npr. določitev oblike terena je v gozdnih predelih precej manj natančna kot na travnatih območjih) itd. Na položajno in višinsko natančnost vpliva tudi način zajema, ki zajete grafične entitete obremeni z dodatno napako.

S preverjanjem pravilnosti opisnih podatkov želimo odpravljati napake, ki so posledica napačnega dodeljevanja vrednosti posameznim entitetam. Pri ločljivosti in popolnosti preverjamo popolnost zajetih grafičnih entitet. Ločljivost in popolnost največkrat primerjamo vizualno tako, da preverjamo zajete grafične entitete s kartografskimi in ostalimi viri. S preverjanjem logične doslednosti skušamo odkrivati in odpravljati napake, kot so: dvojne linije, slepi poligoni, napake vozlišč, napake spojev listov, topološke napake itd. V nadaljevanju je prikazan praktičen primer preverjanja kakovosti digitaliziranih tematskih slojev kartografske dokumentacije za dolgoročni plan Republike Slovenije v merilu 1:25 000.

4.1 Preverjanje položajne natančnosti z rezultati

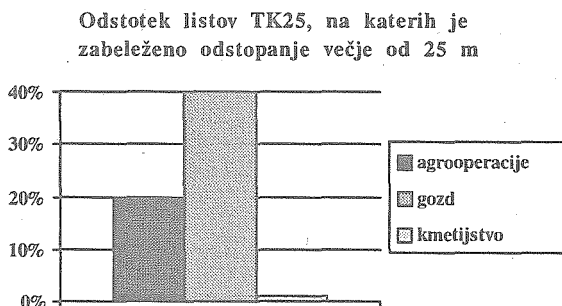
Pri testiranju položajne natančnosti smo najprej preverili število decimalnih mest za decimalno piko, s katerimi so zapisane koordinate testiranega sloja. Število je

preverjeno v Arc/Info-vih tabelah *.TIC in *.BND. V teh tabelah smo preverili tudi upoštevanje cone pri koordinati Y, t.j., ali je št. cone sestavni del koordinate Y. Istočasno so zabeležene še morebitne redukcije koordinat (translacija koordinate Y ali X v Gauss-Kruegerjevem koordinatnem sistemu). Vse vrednosti so ustrezno opisane v obrazcu, katerega se vodi za vsak testirani sloj.

Za vsak sloj smo opravili kontrolo oslonilnih točk (oslonilne točke ali tic-i predstavljajo v Arc/Inf-ovem podatkovnem modelu transformacijske točke). Preverili smo, ali so oslonilne točke zajete enako kot vogalne točke testiranega lista TK 25. V primeru, da niso zajete kot vogalne točke lista TK 25 in da jih ni več kot 10, smo jih v poročilu izpisali skupaj s koordinatami. Če pa je bilo v posameznem sloju zajetih več kot 10 oslonilnih točk, smo v poročilu navedli le njihovo število.

Najobsežnejše od vseh kontrol pri testiranju položajne natančnosti je bilo prekritje digitalizirane vsebine s skanogramom TK 25. Ker je vsebina, katero so izvajalci digitalizirali, vrisana na kopijah TK 25, smo kot podlago vektorske vsebine uporabili ustrezen skanogram. Situacija digitaliziranih točk, linij in poligonov na skanogramu bi morala biti enaka situaciji točk, linij in poligonov, vrisanih na papirnatih podlagah. Preverjali smo vse dele lista, t.j. od 5 do 7 kontrolnih območij na vsakem listu. Za grobi pogrešek smo označili, če je bilo položajno odstopanje od kartografskega vira (skanograma) večje od 1mm na karti (t.j. 25 m v naravi). List, na katerem je bila takšna napaka, smo označili posebej.

Rezultati:



4.2 Ugotavljanje opisne natančnosti in rezultati testa

Pri ugotavljanju opisne natančnosti smo uporabili test Cohenov kapa, s katerim ugotavljamo verjetnost, da je opisni podatek v podatkovni zbirki enak podatku na kartografskem viru. Preverjali smo zalogo vrednosti oziroma omejenost intervala posameznih opisov. Od naročnika smo dobili legende za vse testirane zbirke podatkov, tako da smo vedeli, katere vrednosti opisov se smejo pojaviti pri posameznih slojih. V obrazec smo vpisali vrednosti, ki se pojavijo pri posameznem opisu in ne bi smele biti v zalogi vrednosti tega opisa. Preverjanje je izvedeno z logičnim izločanjem pravih vrednosti.

Tudi kontrolo popolnosti opisov smo izvedli z logičnim izločanjem vrednosti 0 v tabelah:

- ☐ določili smo število polj v tabeli, ki nimajo vnešenih vrednosti
- ☐ določili smo število zapisov znotraj posameznega polja, ki nimajo vnešenih vrednosti.

Sledilo je vizualno preverjanje opisnih podatkov s kartografsko podlago. Izvedli smo ga s programom ArcView s pomočjo primerne izbire kartografskega znaka za posamezne grafične elemente (točke, linije, poligoni). Vsa neskladja pri opisovanju smo vpisali v obrazec. Izračun faktorja Cohenov kapa (Goodchild, 1990) v vrednosti med 0 in 1 nam pove, kakšna je opisna natančnost. Izračunali smo ga na podlagi vzorca za vsak opis posebej. Za izračun Cohenovega faktorja je izbran vzorec z najmanj tridesetimi elementi posameznega sloja. Na tem vzorcu smo preverili vrednost zapisa v tabeli z vrednostjo elementa v podatkovni zbirki. V primeru, da so vsi elementi vzorca vnešeni pravilno, je vrednost faktorja 1. V kolikor faktor ni enak ena (da obstaja vsaj eno nepravilno opisovanje), je pripisana tudi klasifikacijska matrika, po kateri se izračuna faktor Cohenov kapa. Vpisani sta tudi velikost vzorca (n) in velikost testiranega pojava (N = št. vseh elementov testiranega lista). V kolikor je testirani sloj združen za območje več listov TK 25, je N enak seštevku vseh elementov, ki so na testnih listih. Matematična enačba, po kateri se izračuna Cohenov kapa faktor:

$$k = \frac{d - q}{N - q}$$

kjer je d vsota diagonalnih vrednosti matrike. Vrednost q je vsota produktov seštevcev stolpcev in vrstic, deljena s številom vseh elementov, ki nastopajo v vzorcu (N). Primer izračuna Cohenovega faktorja:

Klasifikacijska matrika za podatkovno zbirko agrooperacij:

		vrednosti v naravi			
		K	O	KO	vsota vrstic
vrednosti v podatkovni zbirki	komasacija - K	1	1		2
	osuševanje - O	1	11		12
	komasacija in osuševanje - KO			3	3
	vsota stolpcev	2	12	3	17

Izračun:

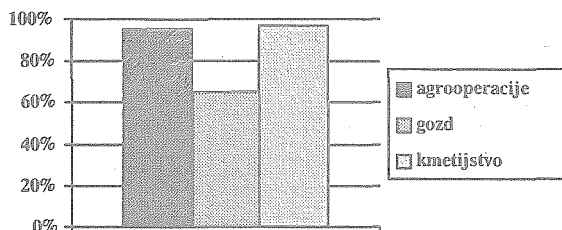
$$q = [(2 * 2) + (12 * 12) + (3 * 3)] / 17 = 9,24, \quad N = 17, \quad d = 15$$

$$k = [15 - 9,24] / [17 - 9,24] = 0,74$$

Preverjanje je izvedeno v programu ArcView.

Rezultati:

Povprečen Cohenov faktor



4.3 Preverjanje logične doslednosti in rezultati testa

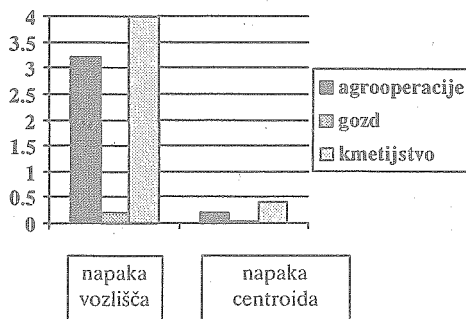
Izraz logična doslednost je opredeljen kot doslednost povezav med grafičnimi in opisnimi podatki. Pri testiranju logične doslednosti smo s pomočjo programa Arc/Info odkrili topološke napake v grafičnih in opisnih podatkih:

- ugotovili smo število napak v vozliščih. Pri poligonski strukturi podatkov smo preverili število visečih linij, zaradi katerih so bili poligoni nezaključeni.
- Ugotovili smo število napak pri centroidih poligonov. Kontrola pride v poštev pri poligonski strukturi podatkov, kjer smo podali število poligonov, ki so brez centroida.

V primeru, da je bil testni sloj združen za večje, smo izvedli vizualno kontrolo uskladitve stikov robov. Kontrola je narejena na prehodu linijskih in poligonskih elementov z enega lista na drugega. V kolikor je odstopanje večje od grafične natančnosti 25 m, je to zabeleženo. Preverjanje je bilo izvedeno s primerjavo testiranega sloja, skanogramov in mreže kartografskega sistema TK 25.

Rezultati:

Povprečno število napak na listu
TK25

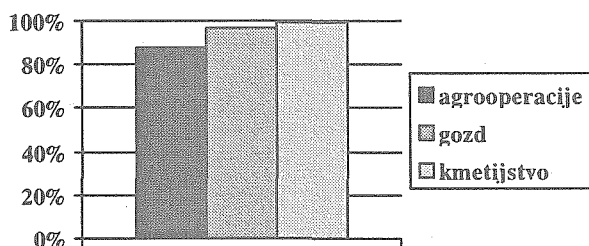


4.4 Preverjanje popolnosti in rezultati testa

Popolnost smo v okviru projekta opredelili kot razmerje med številom objektov na kartografskem viru in v digitalni zbirki. V primeru, da so manjkali elementi posameznega sloja, smo vpisali približno lokacijo in tudi oceno števila manjkajočih elementov (v odstotkih). Preverjanje popolnosti se izvaja pri zajemanju vzorcev za izračun faktorja kapa ter ob testiranju položajne natančnosti.

Rezultati:

Povprečen odstotek popolnosti na testnih listih



5 ZAKLJUČEK

V članku opisano preverjanje kakovosti digitalizacije je eden od prvih primerov v Sloveniji, ko je projektu digitalizacije sledila tudi faza preverjanja opravljene digitalizacije. Rezultati so pokazali, da je naloga dosegla svoj namen, saj je iz naloge razvidno, da izvajalci dela niso izvedli popolnoma brez napak. Naročnik bo lahko uporabil rezultate preverjanja kakovosti na več področjih. Prvič se bodo rezultati uporabljali pri dopolnitvi in odpravi napak v podatkovni zbirki. Drugič bomo pri vseh nadaljnjih pogodbah upoštevali rezultate projekta in zahtevali preverjanje kakovosti po eni od uveljavljenih metodologij. Tretjič bomo uporabili rezultate pri določanju stopnje zaupanja v rezultate prostorskih analiz, ki so bile in bodo opravljene z digitalnimi zbirkami prostorskega plana Republike Slovenije. Predlagamo tudi, da v pomoč vsem naročnikom in izvajalcem podobnih projektov sprejmemo okvirna navodila za zagotovitev ustrezne kakovosti obstoječih in nastajajočih prostorskih podatkovnih zbirk.

Literatura:

- Giordano, A., Veregin, H., *Il Controllo di Qualita nei Sistemi Informativi Territoriali. Il Cardo*, Benetke, 1994
- Ivačič, M., *Izbrane metode ugotavljanja kakovosti prostorskih podatkov v geografskih informacijskih sistemih. Magistrska naloga. FGG-Oddelek za geodezijo*, 1996
- SLS ISO 9001, *Sistemi kakovosti - model zagotavljanja kakovosti v načrtovanju/razvoju, proizvodnji in servisiranju (identičen z ISO 9001,1987). Urad Republike Slovenije za standardizacijo in meroslovje pri Ministrstvu za znanost in tehnologijo*, 1992
- Pegan, Ž.B., Radovan, D., *Pregled in ocena stanja na področju standardov prostorskih podatkov. FGG-Oddelek za geodezijo, Univerza v Ljubljani, Ljubljana*, 1995, 48 str.
- Goodchild, M.F., Kemp, K.K., *Technical issues in GIS. NCGLA*, 1990

Recenzija: mag. Božena Lipej
doc.dr. Radoš Šumrada