

**M E T O D E - M E T H O D S - M E T H O D E**

UDK 910.1

**UČINKOVITOST PRIKAZOVANJA PODATKOV: NA IZBRANIH  
GEOGRAFSKIH PRIMERIH**

Marko Krevs\*

*Povzetek*

*V prispevku je prikazana problematika ustrezne preureditve številskih podatkov za prikaz v tabelah, na grafikonih ali na kartah. Na podlagi teoretične opredelitve učinkovitosti izkoristka izbranega števila razredov/kategorij za prikaz avtor izpeljuje dva koeficienta za merjenje te učinkovitosti. Uporaba opisanih metod je prikazana na primeru osončenja v delu Brkinov in prometa na prodajalno v trgovini na drobno po slovenskih občinah.*

Strokovno poročilo. Prikaz podatkov, linearno kontrastno povezovanje histogramsko izenačevanje, koeficient izkoriščenosti spektra.

**DATA PRESENTATION EFFICIENCY: ON CHOSEN GEOGRAPHICAL CASES**

*Abstract*

*The article deals with the problems of appropriate classification of numerical data for presentation in tables, graphs or maps. On the basis of theoretical definition of efficiency of usage of the chosen number of classes/categories for presentation the autor presents two coefficients to measure this efficiency. The use of the methods is shown on a case of insolation in a part of Brkini, and on a case of turnover on a shop in retail trade in Slovenia.*

Scientific report. Data presentation, Linear contrast stretching, Histogram equalization, Coefficient of usage of spectrum.

---

\* Dipl. geogr., mladi raziskovalec, Oddelek za geografijo, Filozofska fakulteta v Ljubljani, 61000 Ljubljana, Aškerčeva 12, Slovenija

## Uvod

Geografi smo pri raziskovalnem delu neredko izgubljeni med podatki, katerih zbiranje terja ogromno časa in so nazadnje še "obupno nekvalitetni", ter nepregledno množico metod, katerih večine ne razumemo ali jih ne znamo uporabiti. Tako se običajno odločimo za metode, ki so preprosto izvedljive in se zato zdijo tudi povsem razumljive. Pa vendarle ni vedno tako preprosto. Ustavimo se pri morda najpogosteje, v vseh fazah raziskovalnega dela uporabljani, lahko bi rekli "prvi in zadnji" izmed njih, ki je mogoče prav zaradi tega pred uporabo sploh več ne premislimo.

Kadar hočemo grafično ali številsko prikazati notranjo (tematsko) ali prostorsko strukturiranost preučevanega pojava, smo običajno v zadregi, kako konkretne vrednosti porazdeliti po ustreznih razredih, da bo rezultat - običajno slika - kar najbolj pokazal tisto, kar hočemo upodobiti. Ni predpisanih pravil, kako pridemo do najboljših rezultatov. Odločitve je odvisna od same tematike in seveda od presoje raziskovalca, kartografa,...

Med drugim gre za star kartografski problem, ki je ponovno dobil izredno velik pomen na področju analize podatkov ("slik"), zbranih s pomočjo daljinskega zaznavanja. Pojem "izboljšanje kvalitete" (ang. *enhancement*) pomeni pri analizi tovrstnih "slik" spremembo prikaza na tak način, da je informacija, ki jo vsebuje slika, bolj priročna za interpretacijo s strani opazovalca glede na njegove specifične potrebe (Mather, 1989, 154).

Nikakor pa to ni le problem (tematsko-) kartografskega prikazovanja - to je le najbolj očiten primer. S podobnimi težavami se srečamo tudi pri porazdeljevanju vrednosti pojava po razredih za potrebe "tabelarnega" prikaza (npr. frekvenčna porazdelitev), prikaza z grafikoni (npr. histogrami), ali pri razvrstitvi velikega števila podatkov v manjše število razredov, da jih lažje uporabimo v nadaljnji analizi ali pri razlagi pojava.

Gre torej za dve ravni presoje: "tematska" temelji na strokovnem znanju s področja, ki preučuje obravnavani pojav, "tehnična" pa na poznavanju principov in tehnik prikazovanja podatkov. Na tem mestu bomo obravnavali predvsem prvo.

Govorimo seveda o porazdelitvi "številskih" vrednosti, merjenih na ordinalni, intervalni ali razmernostni lestvici. Prikazovanje porazdelitve za nominalne, "neštevilske" podatke je določeno že s samo opredelitvijo kategorij in njihovo "smiselnostjo".

Meje razredov za "številске" podatke lahko določimo na tri načine:

- 1) na podlagi empirično ugotovljenih meja, pragov, ..., znotraj vrednosti za preučevan pojav (npr. pri prikazu naklonov zemljišča poznamo nekatere mejne vrednosti za določene rabe zemljišč ali dejavnosti, ki jih na njih opravljamo),
- 2) na podlagi preučitve nekaterih statističnih kazalcev o podatkih;
- 3) subjektivno, intuitivno - na podlagi le šibkega poznavanja pojava in na podlagi predvidevanj, izkušenj (običajno v začetnih delovnih fazah raziskave, ko še zbiramo podatke).

V tem sestavku se bomo zadržali predvsem pri drugem načinu. Ta je morda najpogostejši pri geografskem preučevanju. Situacija je običajno naslednja: imamo zbrane podatke, na katerih (lahko) opravimo nekatere izračune, pojava pa ne poznamo toliko, da bi z gotovostjo postavili "smiselne" meje razredov.

Učinkovitost prikaza geografskih podatkov sicer ni odvisna samo od slednjega. Z nerazumevanjem bistva tega postopka pa lahko zgrešimo cilj predstavitve podatkov, kar nas lahko pripelje celo do napačnega sklepanja.

### Opredelitev dveh pristopov

Problematika je "živa" na področju analize satelitskih slik, ki je tesno povezana z zdaj morda razvojnno najživahnjšim področjem geografskega dela in raziskovanja - geografskimi informacijskimi sistemi. Zato bo v besedilu uporabljeno "njihovo" poimenovanje tovrstnih tehnik, zaradi nazornosti pa tudi nekaj druge "grafične terminologije".

Osnovni dve vprašanji, na kateri moramo odgovoriti, sta:

- 1) na koliko razredov bomo porazdelili vrednosti in
- 2) na kakšen način hočemo izkoristiti nastavljeni "spekter".

"Spekter" naj nam predstavljajo številke razredov, po katerih porazdelimo vrednosti (npr. vrednosti za nek pojav za posamezna obravnavana območja - občine, celice iz mreže celic in podobno), oziroma barve, šrafure, odtenki, s katerimi bomo prikazali vrednosti, ki pripadajo posameznim razredom.

Na prvo vprašanje odgovorimo na podlagi zahtev glede natančnosti prikaza. Te zahteve lahko izhajajo iz notranje heterogenosti pojava, iz uporabljenih metod v naslednjih fazah dela, iz potreb ali želja naročnika raziskave ali samega raziskovalca, pa tudi iz zahtev same delovne faze.

Pri odgovoru na drugo vprašanje je koristno poznati princip vrednotenja izkoriščenosti "spektra" in s tem dva ekstremna primera. *Minimalno* ga izkoristimo, če so vse vrednosti pojava (npr. za vse prostorske enote, ki jih obravnavamo) v enem samem razredu, vsi preostali razredi pa so "prazni". Karta bi bila v tem primeru cela pobarvana z enako barvo. *Maksimalno* pa "spekter" izkoristimo tako, da je število vrednosti (npr. uvrščenih prostorskih enot) v vseh posameznih razredih enako "pričakovani frekvenci",  $f$ , in s tem v vseh razredih enako. Na karti bi bile vse barve izbranega "spektra" številčno enako zastopane (po površini pa le, če bi bila prikazana območja enako velika - tak primer so slike, sorodne sliki 2).

V nadaljnjem sta prikazani metodi za razvrščanje, ki se pogosto uporabljata za izboljšanje kontrasta slik satelitskih ali letalskih posnetkov. Tukaj se ne omejujemo samo na tovrsten prikaz podatkov. Z uporabnikovega vidika sta obe namenjeni čimbolj nazornemu, prepričljivemu prikazu pojava. Z vidika zgornje opredelitve "izkoristka izbranega spektra" pa sta - kot bodo pokazali tudi izbrani primeri - usmerjeni vsaka v svojo

smer. Katero izmed njiju bo raziskovalec uporabil, je odvisno predvsem od njegovega poznavanja pojava, uporabljenih metod in samega namena prikaza.

Prva tehnika je *linearno kontrastno raztezanje* (ang. *linear contrast stretching*) (Mather 1989, 158-164). Izmerjene oziroma podane vrednosti med  $P_{\min}$  in  $P_{\max}$  preslika na celoten interval, ki je neposredno povezan z zmožnostmi prikazovalnika (npr. monitorja; običajno 256 sivih tonov) ali sposobnostmi razpoznavanja in kriteriji nazornosti pri tematski kartografiji (npr. do 8-10 različnih šrafurnih vzorcev). Podrazredi na razponu med  $P_{\min}$  in  $P_{\max}$  so enako veliki, ustrezne frekvence po odgovarjajočih razredih pa običajno zelo različne. Na sliki prevladujejo (so najštevilčnejše zastopane) barve razredov, ki so blizu povprečja, ekstremne vrednosti pa so bolj na redko posejane. Z drugimi besedami: kontrastnost slike je okoli povprečja nizka, proti skrajnim razredom pa se veča.

Prednost te tehnike pred npr. "slepim" določanjem meja razredov ali pred "avtomatičnim" določanjem teh meja z uporabo razpona med 3 standardnimi odkloni (SD) pod povprečjem in 3 SD nad njim je, da upošteva dejanska ekstrema, ki ju potem namesti tudi na ekstremna konca spektra prikazovanja. To pomeni, da je ta spekter izkoriščen "v celoti". Za nekatere naloge je koristno, da so razredi enako veliki (glede na mersko lestvico) in razlike med razredi odsevajo povprečno razliko med vrednostmi, ki so v različnih razredih.

Pomanjkljivosti te osnovne variante prikazane tehnike so predvsem naslednje: Če skrajne vrednosti izrazito odstopajo od glavnine vrednosti, potem se lahko pojavljajo "praznine" (razredi s frekvenco=0) tudi znotraj spektra prikazovanja. V primeru izrazito koničaste frekvenčne porazdelitve dobimo bolj malo vpogleda v razlike znotraj glavnine vrednosti, ki so nagnetene v enem, dveh, morda treh razredih, medtem ko so drugi razredi skorajda ali povsem prazni.

Druga tehnika je *histogramsko izenačevanje* (ang. *histogram equalization*) (Mather 1989, 164-166), ki deluje skorajda nasprotno, kot linearno kontrastno raztezanje. Vsak razred v spektru prikazovanja (oziroma v histogramu tega spektra) ima približno enako frekvenco. S tem se poveča entropija slike, t.j. mera informacijske vsebine slike (Mather 1989, 164, 45). V primerjavi s prej omenjeno tehniko se tukaj razredi, ki so imeli tam relativno nizko frekvenco, "okrepijo", medtem ko so tisti z višjimi frekvenca mi porazdeljeni na več razredov. Rezultat je povečanje kontrasta v osredju spektra prikazovanja (t.j. okoli povprečja) in zmanjšanje na njegovem obrobju.

Prednost te tehnike je enakomernejša porazdelitev meritev po razredih ali z drugimi besedami - približno enaka zastopanost posameznih barv oziroma šrafur na sliki.

Njena osnovna slabost pa izhaja iz dejstva, da temelji na rangih in ne na osnovnih vrednostih. Zaradi tega razlike med razredi ne odražajo dejanskega merila razlik med vrednostmi, ki so v njih uvrščene. Z drugimi besedami: intervali (glede na mersko lestvico) so okoli povprečja zelo majhni in so razlike med posameznimi uvrščenimi enotami tudi zelo majhne. Obratno pa je na robovih porazdelitve, kjer imajo skrajni razredi zelo velik razpon, razlike med uvrščenimi vrednostmi pa so lahko zelo velike.

To sta le dve osnovni tehniki. Obstaja še vrsta variant, ki bolj ali manj uspešno od-

pravljajo omenjene napake teh dveh, pa tudi drugačne tehnike, ki na primer poskušajo obstoječo frekvenčno porazdelitev preoblikovati v katero izmed znanih porazdelitev (npr. v normalno porazdelitev, t.i. Gaussiansko raztezanje) ali za različna makro-območja na sliki (npr. za morje/kopno ali za reliefno razgibano/ravninsko območje) uporabijo različne kriterije porazdeljevanja (Mather 1989, 166-173).

### **Koeficienta izkoriščenosti spektra**

Že na podlagi preučitve tabele, ki prikazuje frekvenčno porazdelitev vrednosti nekega pojava po razredih, ali ustreznega histograma, morda tudi že same slike (karte), lahko do neke mere ocenimo, koliko smo izkoristili izbrano število razredov oziroma kategorij. Za natančnejšo presojo, ki naj ima tudi bolj splošno veljavo, pa potrebujemo ustrezen kazalec. V ta namen je avtor oblikoval dva koeficienta, ki naj ju preizkusi in izpopolni uporaba pri praktičnem delu.

Koeficienta sta zasnovana na zgoraj opredeljenih pogledih na izkoristek uporabljene "spektra", njeni vrednosti pa sta korigirani tako, da na njuni podlagi lahko primerjamo učinkovitost porazdelitve različnih pojavov ne glede na število obravnavanih podatkov in razredov. Razlikujeta se po tem, da prvi obravnava absolutne odklone frekvenc v posameznih razredih od pričakovanih frekvenc (nekakšna analogija povprečnega absolutnega odklona), drugi pa kvadrate teh odklonov (analogija standardnega odklona).

Pomembno je, da se uporabnik metod zaveda, da sama koeficienta ne kažeta neposredno učinkovitosti porazdelitve "nasploh", ampak le glede na zgoraj podano teoretično opredelitev te učinkovitosti!

Prvi "koeficient odstopanja od maksimalnega izkoristka spektra",  $W$ , je izračunan tako:

$$W = \sum_{i=1}^r \frac{\text{abs} ( f_i - f' )}{N}$$

kjer je  $r$  število razredov (npr. barv v "spekttru"),  $\text{abs}$  pomeni absolutno vrednost izraza v oklepaju,  $f_i$  dejansko frekvenco v razredu  $i$ ,  $N$  skupno število preučevanih enot (vsota vseh razrednih frekvenc) ter  $f'$  pričakovano frekvenco. Slednja je izračunana takole:

$$f' = \frac{N}{r}$$

Minimalno vrednost (=0) ima koeficient  $W$  takrat, kadar so vse dejanske frekvence enake pričakovanim frekvencam - kadar je v vsakem razredu enako število uvrščenih enot.

Videli bomo, da je ta meja v praksi le težko dosegljiva. Sama metoda histogramskega izenačevanja na primer teži k uresničitvi zgornje zahteve. Pri tem jo ovira dvojje: pričakovana frekvenca je celo število (in jo dejanska frekvenca sploh lahko doseže) le v primerih, ko je  $N$  deljiv z  $r$ . Tudi pri postopku porazdeljevanja po razredih lahko pride do nekaterih "netočnosti". Na koncu tega poglavja sta postopka za obe metodi na kratko opisana.

Največjo mogočo vrednost ima koeficient takrat, ko so vse enote v enem razredu, ostali razredi pa so prazni (njihove  $f_i=0$ ). Pri izbranem številu razredov  $r$  izračunamo  $W_{\max}$  na naslednji način:

$$W_{\max} = \frac{N - f'}{N} + (r - 1) \frac{f'}{N}$$

po manjši preureditvi pa se izkaže naslednje:

$$W_{\max} = 1 + \frac{(r - 2)}{r} = 2 - \frac{2}{r}$$

Če je razred samo en, je  $W_{\max}$  enak 0. Z večanjem števila razredov pa se  $W_{\max}$  približuje 2. To povejmo še na naslednji način:

$$W_{\max}(r = 1) = 0$$

$$\lim_{r \rightarrow \infty} W_{\max} = 2$$

Ker lahko doseže koeficient  $W$  ob različnem številu razredov različne maksimalne vrednosti (med 0 in 2), ga korigiramo tako, da je končni rezultat koeficient  $W_{\text{kor}}$ , ki pokaže delež, ki ga doseže  $W$  od največjega mogočega  $W_{\max}$  pri tem številu razredov:

$$W_{\text{kor}} = \frac{W}{W_{\max}}$$

za katerega velja:

$$0 \leq W_{\text{kor}} \leq 1$$

Tako lahko korigirane koeficiente  $W_{\text{kor}}$  primerjamo med seboj ne glede na to, za kakšen način porazdeljevanja gre, in ne glede na to, po koliko razredih so bile enote porazdeljene ter kolikšno je bilo njihovo skupno število  $N$ .

In še dodatna pomoč pri razlagi pomena koeficientov.  $W_{\text{kor}}$  je kazalec stopnje odstopanja od maksimalnega izkoristka "spektra" oziroma izbranega števila razredov, in sicer kaže delež (če ga pomnožimo s 100, kaže število odstotkov) največjega možnega odstopanja od maksimalnega (popolnega) izkoristka "spektra".

Nasprotno pa lahko naslednji koeficient:

$$1 - W_{kor}$$

razložimo kot stopnjo odstopanja od minimalnega izkoristka "spektra". Tudi ta ima praktičen pomen. Če bi na primer radi, da na sliki izstopa le en ali nekaj razredov, preostali pa so čim šibkeje zastopani, nam ta koeficient pokaže stopnjo uspešnosti pri tem. Po drugi strani pa lahko ta koeficient razložimo tudi kot kazalec deleža od maksimalnega možnega izkoristka "spektra" pri danem številu razredov. Zato je morda še bolj uporaben v razlagi značilnosti porazdelitve, kot sam  $W_{kor}$ !

Drugi "koeficient odstopanja od maksimalnega izkoristka spektra",  $W_2$ , je izračunan tako:

$$W_2 = \sqrt{\sum_{i=1}^r \frac{(f_i - f')^2}{N}}$$

Tudi pri tem se maksimalna možna vrednost spreminja s številom razredov, a tudi s številom enot (N):

$$W_{2_{max}} = \sqrt{\frac{(N - f')^2}{N - 1} + (r - 1) \frac{f'^2}{N - 1}}$$

oziroma po manjšem preurejanju:

$$W_{2_{max}} = \sqrt{N - \frac{N}{r}} = \sqrt{N \left(1 - \frac{1}{r}\right)}$$

Za to vrednost velja naslednje:

$$\lim_{r \rightarrow 1} W_{2_{max}} = 0$$

če je N naravno število, potem je

$$\lim_{r \rightarrow \infty} W_{2_{max}} = \sqrt{N}$$

in če je r naravno število, velja

$$\lim_{r \rightarrow \infty} W_{2_{max}} = \infty$$

Korigirano vrednost tega koeficienta,  $W_{2_{kor}}$ , izračunamo na podoben način:

$$W_{2_{kor}} = \frac{W_2}{W_{2_{max}}}$$



in zanj prav tako velja:

$$0 \leq W_{2_{kor}} \leq 1$$

Ta koeficient je - podobno kot standardni odklon v primerjavi s povprečnim absolutnim odklonom - bolj občutljiv na bolj ekstremna odstopanja od pričakovane frekvence, zato je  $W_2$  praviloma večji kot  $W$ . Enako velja tudi za koeficienta  $W_{kor}$  in  $W_{2_{kor}}$ . Primerjava korigiranih koeficientov nam pove naslednje: čim večja je razlika med njima, tem intenzivnejši so prispevki odstopanj ekstremnih frekvenc  $f_i$  od pričakovane frekvence  $f^*$  k skupni oceni odstopanj, izraženi s koeficientoma. Sicer pa je obrazložitev vrednosti koeficienta  $W_{2_{kor}}$  in koeficienta

$$1 - W_{2_{kor}}$$

enaka, kot za  $W_{kor}$  oziroma  $1 - W_{kor}$ .

Na kratko si oglejmo še postopka porazdeljevanja po razredih, kakršna uporabljata tudi avtorjeva programa. V obeh primerih razvrščanje po razredih deluje s celimi števili, zato podatke najprej pomnožimo z ustreznim faktorjem (z 10 na ustrezno potenco), da je "ločljivost" med vrednostmi dovolj velika za uspešno porazdeljevanje.

Pri linearnem kontrastnem raztezanju razpon med največjo in najmanjšo vrednostjo preprosto razdelimo na izbrano število razredov. Program omogoča, da "popravimo" tako določitev meja razredov tako, da je ena izmed meja "vsiljena" s strani uporabnika (v tabelah 1 in 2 t.i. fiksna točka). To je lahko poljubno izbrana vrednost, ali pa povprečje, ki ga v ta namen izračuna. Primer: če vzamemo kot "fiksno točko" povprečje, bo ena izmed meja enega od razredov enaka tej vrednosti. Tako bomo neposredno videli, kateri razredi (in na sliki barve) ter koliko pojavov je pod oziroma nad povprečjem.

Pri histogramskem izenačevanju pa je drugače. Porazdeljevanje lahko poteka na dva različna načina: na podlagi rangov vrednosti (kar je smotno le pri relativno majhnem številu obravnavanih podatkov), ali pa na podlagi "vmesnega", podrobnejšega histograma (ta predstavlja nekakšen približek rangov). V slednjem primeru najprej porazdeli vse vrednosti v "r krat p" enako velikih podrazredov (v bistvu opravi linearno raztezanje). Faktor p (v programu med 10 in 30) pomeni podrobnost razreza podatkov, neposredno pa pove, kolikokrat več bo teh podrazredov, kot je izbrano število razredov. V naslednjem koraku iz teh podrazredov ali na podlagi rangov "nabira" frekvence, dokler ne doseže kumulativne frekvence, ki je enaka ali večja (tukaj je vir napak - podrobnejši ko je razrez na podrazrede, manjše so možne napake) kot bi morala biti ob predpostavki, da je v vsakem razredu enako število podatkov.

Koeficienta v tej obliki kažeta "učinkovitost" porazdelitve po razredih in z vidika prostorskega (kartografskega) prikaza - strogo vzeto - veljata samo v primeru, če so obravnavana območja enako velika. Če bi hoteli, da upoštevata poleg številčne zastopnosti posameznih razredov tudi morebitne različne površine posameznih obravnavanih prostorskih enot na karti (npr. občin), bi morali frekvence v formulah pomnožiti z ustreznimi površinami (podobno velja za N).



### Izbrana primera

Prvi primer prikazuje dekadne vsote energije sončnega obsevanja po kvadratnih celicah z osnovnico velikosti 50 metrov v delu Brkinov, za dvajseto dekada v letu. Podatke (digitalni model reliefa) je zbral Igor Šebenik s sodelavci (1990). Količine sončnega obsevanja so bile ocenjene s pomočjo programa SONČEK (Gabrovec in Krevs, 1989). Za lažje razumevanje slike sončnega obsevanja je na sliki 1 ilustrativen prikaz reliefa na obravnavanem območju.

Poleti je sonce tako visoko, da je pretežni del površja cel dan osončen. Zato površine, ki so večji del dneva vseeno v senci, toliko intenzivneje odstopajo od povprečja.

V tabeli 1a in na sliki 4a vidimo, da so pri prvi metodi - linearnem raztezanju - nadpovprečne vrednosti le v dveh razredih (izmed osmih), v katerih pa je skoraj 60% vseh prostorskih enot (celic 50 x 50 metrov). Prvi trije razredi so na sliki 2a komajda opazni, glede na frekvenčno porazdelitev v tabeli pa tudi pričakujemo, da na sliki prevladujejo temnejše "barve". Če hočemo prikazati prevlado močnejše osončenih površin in hkrati ohraniti povprečno razliko med vrednostmi v sosednjih razredih (za sliko 1a znaša 219), bomo s to porazdelitvijo kar zadovoljni. Koeficienta (tabela 1a) kažeta, da smo kljub morda drugačnemu prvemu vtisu bliže maksimalnemu, kot pa minimalnemu izkoristku "spektra". Prvi kaže na 58.6%, drugi pa celo na 64.7% delež njegovega popolnega izkoristka.

V tabeli 1b in na sliki 4b, ki prikazujeta rezultate druge metode, so frekvence po razredih približno enake, zato je tudi slika 2b "barvno uravnotežena". Če smo hoteli pokazati (okolici) 48 najbolj osončenih celic, in naslednjih 48,... in 48 najmanj osončenih (npr. da iščemo območje, ki bi bilo najprimernejše za določeno rastje), bo ta način porazdelitve za prikaz ustrežnejši. Kot vidimo (tabela 1b), se razponi vrednosti v posameznih razredih zdaj izrazito razlikujejo. Oba koeficienta kažeta, da je porazdelitev popolnoma (99 odstotno) izkoristila izbrano število razredov.

Slika 1

Pogled na obravnavano območje v Brkinih z juga in z vzhoda.

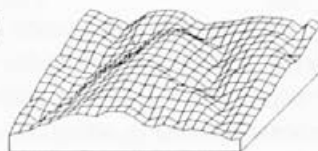
Figure 1

Digital elevation model of a part of Brkini - view from the a) south and b) east.

a) Pogled z juga



b) Pogled z vzhoda

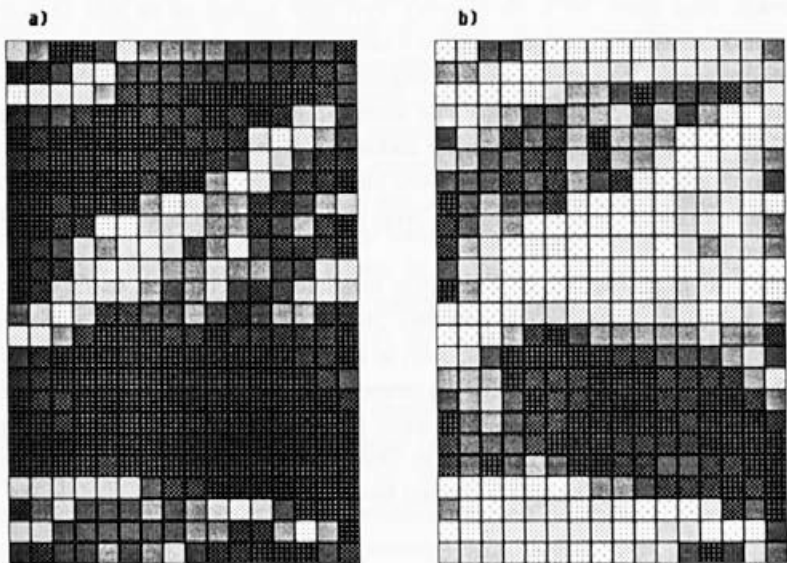


## Slika 2

Dekadne vsote količine sončnega obsevanja za del Brkinov po celicah 50 x 50 metrov, za 20. dekada v letu. Lepo je vidna razlika med rezultatoma uporabe različnih tehnik porazdeljevanja po razredih: a) linearno kontrastno raztezanje, b) histogramsko izenačevanje.

## Figure 2

The ten-days-sums of amount of insolation for a part of Brkini (20<sup>th</sup> decade in a year), cells 50 x 50 meters. Easily recognizable differences between the figures are result of the two techniques: a) linear contrast stretching and b) histogram equalization.



Legenda:



Številke v legendi ustrezajo številkam razredov v tabeli 1.

Numbers in legend correspond to numbers of classes in table 1.

Drugi primer prikazuje rezultate uporabe multiplega regresijskega modela, namenjenega za ocenjevanje prometa na prodajalno v trgovini na drobno po občinah v Sloveniji leta 1987 (Krevs 1991). Model, ki upošteva 12 različnih podatkov o razmerah v vsaki občini, je avtor uporabil za ocenjevanje (napovedovanje) sprememb v višini tega prometa ob predvidenih drastičnih spremembah zaposlitvenih in gospodarskih razmer.

Tabela 1

Frekvenčna porazdelitev dekadnih vsot količine sončnega obsevanja za del Brkinov po celicah 50 x 50 metrov in "koeficienta odstopanja od popolnega izkoristka spektra". Tehniki porazdeljevanja po razredih: a) linearno kontrastno raztezanje, b) histogramsko izenačevanje.

Table 1

Frequency distribution of ten-days-sums of amount of insolation for a part of Brkini for 20<sup>th</sup> decade in the year, and "coefficients of deviation from total usage of spectrum". The techniques for distribution: a) linear contrast stretching and b) histogram equalization.

a)

Fiksna točka: AS = 5287.667  
 Faktor povečave podatkov za razvrščanje: 100.000  
 Razredov: 8 Interval: 219.286

	Meje intervalov	Frekvence
1. razred:	3972 – 4191	1
2. razred:	4192 – 4411	8
3. razred:	4412 – 4630	19
4. razred:	4631 – 4849	36
5. razred:	4850 – 5068	37
6. razred:	5069 – 5288	57
7. razred:	5289 – 5507	77
8. razred:	5508 – 5726	149
Skupaj:	3972 – 5726	384

Pričakovana frekvenca: 48.000 (% od N = 12.5)

W = 0.72395833	W2 = 6.47106434
W <sub>max</sub> = 1.75000000	W2 <sub>max</sub> = 18.35421708
W <sub>kor</sub> = 0.41369048	W2 <sub>kor</sub> = 0.35256553
1-W <sub>kor</sub> = 0.58630952	1-W2 <sub>kor</sub> = 0.64743447

b)

Faktor povečave podatkov za razvrščanje: 1000.000  
 Razredov: 8 Povprečna frekvenca: 48.000

	Meje intervalov	Razpon	Frekvence
1. razred:	40880 – 47481	6601	48
2. razred:	47482 – 50562	3080	50
3. razred:	50563 – 52469	1906	48
4. razred:	52470 – 53936	1466	48
5. razred:	53937 – 55256	1319	49
6. razred:	55257 – 56283	1026	48
7. razred:	56284 – 56870	586	48
8. razred:	56871 – 57310	439	45
Skupaj:	40880 – 57310		384

Pričakovana frekvenca: 48.000 (% od N = 12.5)

W = 0.01562500	W2 = 0.19118976
W <sub>max</sub> = 1.75000000	W2 <sub>max</sub> = 18.35421708
W <sub>kor</sub> = 0.00892857	W2 <sub>kor</sub> = 0.01041667
1-W <sub>kor</sub> = 0.99107143	1-W2 <sub>kor</sub> = 0.98958333

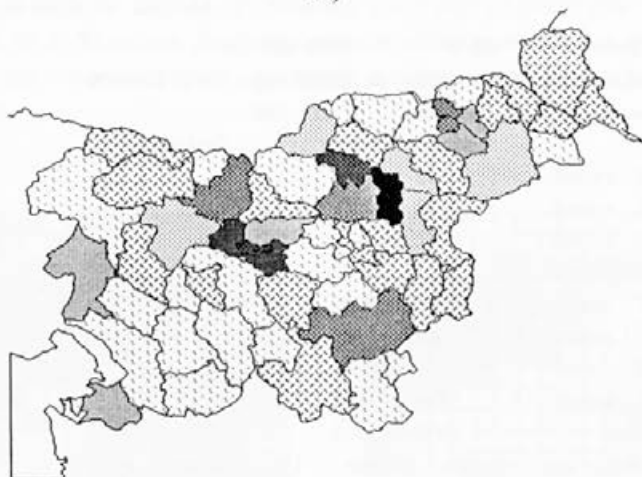
Slika 3

Ocene prometa na prodajalno v trgovini na drobno po slovenskih občinah s pomočjo multiplega regresijskega modela za leto 1987. Tehniki porazdelitve po razredih: a) linearno raztezanje; b) histogramsko izenačevanje.

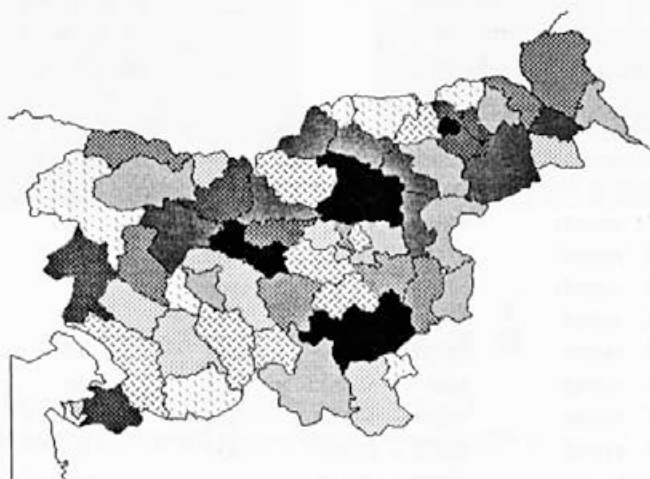
Figure 3

Multiple regression model's estimates of turnover on a shop in retail trade by Slovenian communes in 1987. The techniques for distribution: a) linear contrast stretching and b) histogram equalization.

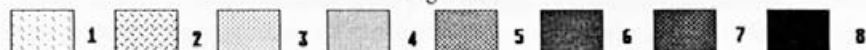
3a)



3b)



Legenda:



Vir: Krevs 1991, 14, 15.

Številke v legendi ustrezajo številkam razredov v tabeli 2.

Numbers in legend correspond to numbers of classes in table 2.

Tabela 2

Frekvenčna porazdelitev ocen prometa na prodajalno v trgovini na drobno leta 1987 po slovenskih občinah in "koeficienta odstopanja od popolnega izkoristka spektra". Tehniki porazdeljevanja po razredih: a) linearno kontrastno raztezanje, b) histogramsko izenačevanje.

*Table 2*

*Frequency distribution of estimates of turnover on a shop in retail trade by Slovenian communes in 1987, and "coefficients of deviation from total usage of spectrum". The techniques for distribution: a) linear contrast stretching and b) histogram equalization.*

a)

Fiksna točka:	AS =	43708.380		
Razredov:	8	Interval:	9318.429	
		Meje intervalov		Frekvence
1. razred:		25072 – 34390		24
2. razred:		34391 – 43709		20
3. razred:		43710 – 53027		5
4. razred:		53028 – 62346		6
5. razred:		62347 – 71664		5
6. razred:		71665 – 80983		2
7. razred:		80984 – 90301		2
8. razred:		90302 – 99619		1
Skupaj:		25072 – 99619		65
Pričakovana frekvenca:	8.125	(% od N = 12.5)		
W	= 0.79411765	W2	= 2.75355943	
W <sub>max</sub>	= 1.75000000	W2 <sub>max</sub>	= 7.77097547	
W <sub>kor</sub>	= 0.45378151	W2 <sub>kor</sub>	= 0.35433897	
1-W <sub>kor</sub>	= 0.54621849	1-W2 <sub>kor</sub>	= 0.64566103	

b)

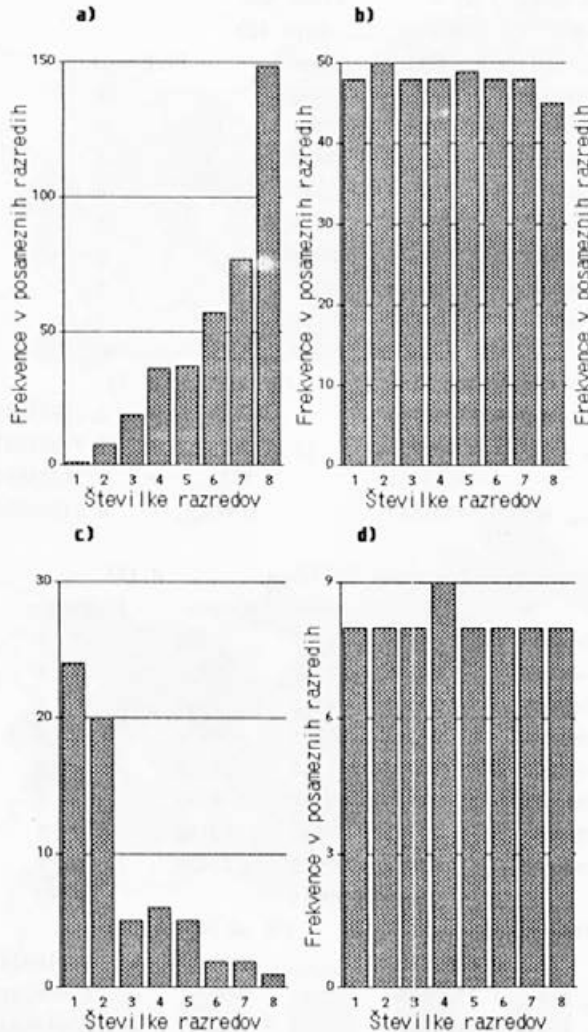
Razredov:	8	Povprečna frekvenca:	8.125	
		Meje intervalov	Razpon	Frekvence
1. razred:		25694 – 30106	4412	8
2. razred:		30107 – 32259	2152	8
3. razred:		32260 – 34322	2062	8
4. razred:		34323 – 38295	3972	9
5. razred:		38296 – 41564	3268	8
6. razred:		41565 – 50307	8742	8
7. razred:		50308 – 63490	13182	8
8. razred:		63491 – 90919	27428	8
Skupaj:		25694 – 90919		65
Pričakovana frekvenca:	8.125	(% od N = 12.5)		
W	= 0.02692308	W2	= 0.11692679	
W <sub>max</sub>	= 1.75000000	W2 <sub>max</sub>	= 7.60024157	
W <sub>kor</sub>	= 0.01538462	W2 <sub>kor</sub>	= 0.01538462	
1-W <sub>kor</sub>	= 0.98461538	1-W2 <sub>kor</sub>	= 0.98461538	

Slika 4

Frekvenčne porazdelitve po razredih za prikazana dva primera: osončenje v Brkinih (a,b) in promet na prodajalno v trgovini na drobno v Sloveniji (c,d). Tehniki porazdeljevanja po razredih: linearno kontrastno raztezanje (a,c) in histogramsko izenačevanje (b,d). Dejanski intervali za posamezne razrede so razvidni iz tabel 1 in 2.

Figure 4

Frequency distributions for the two cases: insolation in Brkini (a,b) and turnover on a shop in retail trade in Slovenia (c,d). The techniques for distribution: linear contrast stretching (a,c) and histogram equalization (b,d). Actual intervals for different classes can be found in tables 1 and 2.



Gostota frekvenc je pri uporabi prve metode (tabela 2a, sliki 3a in 4c) največja na nasprotnem koncu "spektra", kot v prvem primeru. Tudi tukaj sta najštevilčnejše zastopana dva razreda (s skoraj 70% vseh občin), ki pa sta pod povprečjem. Preostala razla-

ga je podobna, kot za prvi primer. Opazimo le, da je razlika med koeficientoma ( $W_{kor}$  in  $W2_{kor}$ ) tukaj znatnejša, kot tam. To kaže na intenzivnejša ekstremna odstopanja frekvenc po razredih od pričakovane frekvence.

Predstavljajmo si hipotetično situacijo: iščemo nekaj (v tem primeru prvih 8) občin, v katerih bi bil promet na prodajalno, ki tujemu investitorju predstavlja pomemben makrolokacijski kazalec, najvišje rangiran v republiki (računa npr. na možnost, da v najvišje rangiranem Celju ne bi dobil ustreznega lokala). Ali pa nasprotno: iščemo tistih nekaj občin, v katerih so razmere najbolj neugodne. V ta namen bi uporabili drugo metodo. Tabela 2b ter sliki 3b in 4d razložimo podobno, kot pri prejšnjem primeru. Morda lahko le še enkrat opozorimo na tukajšnje neupoštevanje različnih površin občin (glej konec prejšnjega poglavja).

### **Sklep**

Prikazana metodologija lahko znatno prispeva k učinkovitosti prikaza geografskih podatkov ter njenemu razumevanju. Uporabimo jo lahko tako pri strokovni, kot pri tehnični presoji prikazovanja podatkov.

Odločitev o metodi porazdeljevanja po razredih je v veliki meri odvisna od tega, kaj hočemo poudariti na prikazu. Za to je nujno potrebno vsaj osnovno poznavanje pojavnosti in razumevanje samih metod. Odločitev o tem, kaj je cilj prikaza, katera je optimalna porazdelitev vrednosti za preučevan pojav po izbranem številu razredov, je namreč še vedno v rokah raziskovalca. Izdelana računalniška programa ne svetujeta neposredno pri izboru tehnike prikaza, temveč sta v pomoč pri zasledovanju postavljenega cilja. S tem, da omogočata preprosto uporabo obeh opisanih tehnik, se lahko na podlagi preizkušanj in skromnega razumevanja metodologije lažje odločimo za postopek razvrščanja po razredih in s tem za čimbolj učinkovit prikaz podatkov ali rezultatov svojega dela.

### **Literatura in viri**

- Andrew, A.M. 1985: Computational techniques in operations research. Abacus press, Tunbridge Wells & Cambridge.
- Cesario, F.J., 1975: Linear and nonlinear regression models of spatial interaction. *Econ. Geogr.*: 51(1975)-1.
- Clark, W.A.V. & Hosking, P.L., 1986: Statistical Methods for Geographers. John Wiley & Sons, New York.
- Clements, D.W., 1978: Utility of linear models in retail geography. *Progress in human geography* 54(1978)-1.
- Cowlard, K.A., 1990: Decision-making in geography. A manual of method and



- practice, Hodder & Stoughton, London.
- Gabrovec M. & Krevs M., 1989: Računalniški program SONČEK\_ST, verzija, ki teče na računalniku Atari ST.
- Johnston, R.J., 1989: Multivariate statistical Analysis in Geography. Longman Scientific & Technical, with John Wiley Sons, Essex.
- Krevs M., 1991: Pisni izdelek za izpit pri predmetu Kvantitativne metode v geografiji. Podiplomski študij, Filozofska fakulteta v Ljubljani, Oddelek za geografijo, mentor: dr. Andrej Černe.
- Mather P.M., 1989: Computer processing of remotely - sensed images. John Wiley & Sons, Chichester.
- Popis delavcev v združenem delu 31.12.1986; 1988; Rezultati raziskovanj 432. Zavod SR Slovenije za statistiko, Ljubljana.
- Popis prebivalstva, gospodinjstev in stanovanj v SR Sloveniji 31.3.1981; 1981. Zavod SR Slovenije za statistiko, Ljubljana.
- Prodajne zmogljivosti v trgovini na drobno 1986; 1989; Rezultati raziskovanj 458/1989. Zavod SR Slovenije za statistiko, Ljubljana.
- Statistični letopis SR Slovenije; 1989. Zavod SR Slovenije za statistiko, Ljubljana.
- Šebenik, Igor et al., 1990: Digitalni model reliefa za del Brkinov, rezultat dela na raziskovalnem taboru v Brkinih 1990, na disketi, osnovnica kvadratne celice je 50 metrov.

## **DATA PRESENTATION EFFICIENCY: ON CHOSEN GEOGRAPHICAL CASES**

Marko Krevs

Summary

Collection of data usually represents a great amount of geographical research. In some phases of work we can use computer as adviser or as a guide in decision process about method to use, or at least as a "guide" through cases, the examination of which should help us to decide.

That kind of work-phase is also the distribution of values of studied phenomenon into classes, for presentation in tables, on graphs or on maps. Help of the computer is particularly welcome when we have data collected but we do not know enough about the phenomenon to determine sensible borders of classes. In the article the autor shows two opposite approaches to distribution of values into such classes, and define two coefficients to indicate the degree of deviation from maximal or minimal usage of chosen "spectrum". The techniques of distribution are easy to understand, and their application is simple with the use of computer programs. Results are directly usable for thematic-cartographic purposes and for further analysis.