

# UPORABA PROSTORSKE STATISTIKE ZA DOLOČEVANJE ZGOSTITEV PROMETNIH NESREČ

## IDENTIFICATION OF ROAD ACCIDENT HOT SPOTS USING SPATIAL STATISTICS

*Peter Lipar, Jure Kostanjšek, Marijan Žura*

UDK: 519.22:656.1.08:711

### POVZETEK

*V prispevku predstavljamo prilagojeno metodo za odkrivanje nevarnih odsekov in krajev, ki temelji na uporabi indeksa Getis-Ord  $G^*$ . Od klasične metode se razlikuje v tem, da poleg teže nezgod upoštevamo prometne obremenitve na kraju nezgode. Metoda je bila uporabljena na primeru omrežja slovenskih avtocest in hitrih cest. Prikazani so rezultati analiz z upoštevanjem prometnih obremenitev in brez njih oziroma primerjava z metodama določanja gostote jeder in lokalnega Anselin-Moranovega indeksa.*

Klasifikacija prispevka po COBISS-u: 1.03

### ABSTRACT

*This paper presents a modified spatial statistical method for identifying road accident hot spots. Our suggested method is based on the Getis-Ord  $G^*$  index. It differs from the standard method, because in addition to the number of accidents and their consequences, we also consider traffic at the locations of these accidents. The method was tested on the Slovene motorway network. The results are presented with and without considering traffic volumes and are compared with both Kernel density method and Anselin-Moran local index method.*

### KLJUČNE BESEDE

*cestnoprometne nezgode, črne točke, prostorska statistika*

### KEY WORDS

*road traffic accidents, black spots, spatial statistics*

## 1 UVOD

Identifikacija črnih točk je ključna za izboljšanje prometne varnosti na cestah (Cheng in Washington, 2008; Moons in sod., 2008; Anderson, 2009). Črne točke so kraji ali odseki, kjer se nezgode relativno pogosteje pojavljajo in so njihove posledice hujše. V literaturi lahko zasledimo različne pristope k odkrivanju nevarnih krajev. Več avtorjev uporablja enostavno empirično Bayesovo metodo (EB) z različnimi porazdelitvenimi funkcijami: Poisson, Poisson-gama, Poisson-lognormal. Poleg binomskega modela uporabljajo hierarhične modele (Sugihara in sod., 2009).

V zadnjem času se vedno bolj širi uporaba metod prostorske statistike (Páez in Scott, 2005). V študijah (Anderson, 2009; Xie in Yan, 2008) zasledimo uporabo gostote jeder (angl. kernel density estimation). Pogosteje se kot indikator prostorske asociacije uporablja lokalni Moranov indeks (Moons in sod., 2008 in 2009) in predvsem indeks Getis-Ord  $G^*$ . V prispevku predstavljamo

prilagojeno metodo za odkrivanje nevarnih odsekov in krajev, ki temelji na uporabi indeksa Getis-Ord  $G^*$ .

Želimo preizkusiti domnevo, da daje prilagojena metoda indeksa Getis-Ord  $G^*$ , ki se od klasične metode razlikuje v tem, da poleg teže nezgod upošteva tudi prometne obremenitve na kraju nezgode, zanesljivejše rezultate. Metoda je bila uporabljena na primeru analize zgostitev prometnih nesreč na omrežju slovenskih avtocest in hitrih cest. Prikazani so rezultati analiz z upoštevanjem prometnih obremenitev in brez njega ter primerjava z metodama določanja gostote jeder in lokalnega Anselin-Moranovega indeksa. Za izvedbo analiz in prikaz rezultatov smo uporabili orodje ArcGIS.

## 2 SEDANJA METODOLOGIJA ZA DRŽAVNE CESTE

### 2.1 Prostorsko načrtovanje in urejanje nevarnih krajev

Podatki o prometnih nesrečah in njihovih udeležencih se za potrebe prostorskega načrtovanja, vzdrževanja, upravljanja cest in urejanja nevarnih krajev vodijo v informacijskem sistemu Direkcije Republike Slovenije za ceste ProVaGIS. Upravljavca državnih cest jih uporablja za analize, s katerimi ugotavlja, na katerih od ugotovljenih lokacij se razlogi za prometne nesreče skrivajo v cesti, poteku ceste, stanju voziščne konstrukcije, prometni signalizacije in opremi ceste ali samem vodenju prometa.

Na podlagi rezultatov analiz upravljavec načrtuje svoje postopke in ukrepe, slednje prednostno razvršča, pri čemer upošteva investicijsko vrednost in razpoložljiva proračunska sredstva.

### 2.2 Sedanja metodologija

Študije lokacij prometnih nesreč se za državne ceste v Sloveniji po naročilu Direkcije Republike Slovenije za ceste izdelujejo vsako leto, in sicer na podlagi podatkov o prometnih nesrečah preteklega triletnega obdobja.

Metodologija določitve nevarnih krajev temelji na izračunu kritične stopnje nezgod. Zadnja študija, za obdobje 2006–2008, je bila izdelana v letu 2009 (Direkcija RS za ceste in Omega Consult, d. o. o., 2009). Poglavitna slabost obstoječe metodologije, ki temelji na statistični analizi atributnih podatkov, je, da je analiza podatkov na mejah med sosednjimi odseki bolj zahtevna. Težave nastopijo pri križiščih cest, kjer se odseki končujejo oziroma začenjajo, saj je konec odseka tudi konec območja obdelave oziroma analize podatkov. Vzpostaviti in voditi je treba vzporedno evidenco križišč in pripadajočih krakov ter stacionaž krakov na odsekih cest, ki sestavljajo križišče, za ločeno analizo območij križišč.

## 3 METODE

### 3.1 Podatki

Prometne nesreče se vodijo kot točkovni podatki v atributni bazi podatkov Prometne nezgode (vir podatkov: Ministrstvo za notranje zadeve – Policija, obdobje 2006–2008), pri čemer se

lokacija opredeli z enolično številko odseka državne ceste (ODSEK) ter stacionaže v metrih (STAC), kar pomeni razdaljo v metrih od začetka odseka do kraja dogodka oziroma prometne nesreče, oziroma z navedbo naselja, ulice in hišne številke za prometne nesreče v naseljih. Žal ni neposredne povezave med naseljem, ulico in hišno številko ter evidenco hišnih številk EHIŠ Geodetske uprave RS, kar pri nesrečah na državnih cestah pomeni dodatno delo pri čiščenju podatkov oziroma povezovanju nesreč z odsekom in stacionažo.

Uradni podatkovni vir so podatki o prometnih nesrečah, zabeleženi s policijskim zapisnikom. Vodijo se v dveh osnovnih, relacijsko povezanih tabelah NEZGODE in UDELEŽENCI.

V tabeli NEZGODE se vodijo podatki o sami prometni nesreči, in sicer enolična šifra nesreče, lokacija (odsek, stacionaža, kraj opisno, lega glede na naselje, opis kraja), čas (datum, ura, dan), tip, vzrok, promet, vreme, površje in klasifikacija nesreče glede na najhuje poškodovanega udeleženca.

V tabeli UDELEŽENCI se vodijo podatki o vseh udeleženi oseb. Relacijska povezava s tabelo nesreč poteka prek enoličnega atributa šifra, ki se v tabeli udeležencev ponovi tolikokrat, kolikor je zapisov o udeleženci posamezne prometne nesreče. Podatki, ki se vodijo o udeleženci, so: poškodba (B - brez, L - lahka, H - huda, S - smrt), vrsta udeležbe (voznik, pešec, kolesar, sopotnik ...), starost, povzročitelj (da/ne), zaščita (uporaba varnostnega pasu ali čelade), vozniški staž, spol, rezultati alkotesta (če je bil opravljen) in drugi.

Z dinamično segmentacijo se generira vektorska točkovna podatkovna plast prometnih nesreč. Podlaga je ustrezen sistem odsekov omrežja državnih cest (angl. route system), ki omogoča generiranje dogodkovnih plasti (angl. event themes) oziroma dogodkov (angl. route events).

Vir podatkov za državno cestno omrežje: digitalna os državnih cest, stanje 1. 1. 2009, Direkcija Republike Slovenije za ceste.

### 3.2 Izračun teže nezgod

Iz osnovnih tabel smo za vsako nesrečo izračunali število mrtvih  $m_i$ , težko ranjenih  $t_i$ , lažje ranjenih  $l_i$ , udeležencev brez poškodb  $b_i$ .

Nato smo izračunali težo za vse nesreče  $p_i$ .

$$p_i = 5m_i + 3t_i + 3l_i + b_i \quad i = 1 \dots n \text{ (skupno število nezgod).} \quad (1)$$

### 3.3 Izračun relativne teže nezgod

Za izračun relativne teže nesreč upoštevamo povprečni letni dnevni promet na kraju nesreče. Promet na kraju  $i$ -te nesreče  $v_i$  ugotovimo s prostorskim prekrivanjem (angl. spatial intersect) podatkov o nesrečah in podatkov o prometu.

Nato smo izračunali relativno težo nesreč  $x_i$

$$x_i = \frac{p_i}{v_i} \quad i = 1 \dots n \text{ (skupno število nezgod),} \quad (2)$$

kjer je

$p_i$         teže  $i$ -te nezgode,

$v_i$         povprečni letni dnevni promet na kraju  $i$ -te nezgode (vozil/dan).

### 3.4 Izračun indeksov

#### 3.4.1 Getis-ord $G^*$

Analiza krajev s povečano frekvenco dogodkov z visoko vrednostjo izbranega parametra (angl. hot spots) za vse element obravnavane množice izračuna »Gi\*-statistiko« za izbrani parameter. Gi\*-statistika pove, ali se elementi z visoko ali nizko vrednostjo parametra pogosteje pojavljajo na nekaterih krajih znotraj obravnavanega območja. Če ima parameter izbranega elementa visoko vrednost in imajo bližnji elementi prav tako visoko vrednost, leži obravnavani element znotraj kraja s povečano gostoto elementov z visoko vrednostjo parametra. Delna vsota vrednosti parametra elementov znotraj kraja s povečano gostoto se primerja s skupno vsoto vrednosti parametra vseh elementov: če delna vsota bistveno odstopa od pričakovane delne vsote in je razlika odstopanja večja, kot bi lahko bila zaradi naključja, je ta razlika podana kot statistično značilna vrednost  $z$ .

$Z$ -vrednost (tudi standardna vrednost) je statistični kazalec položaja posameznega statističnega elementa v statistični populaciji (obravnavanem območju) glede na aritmetično sredino, v merski enoti standardnega odklona  $\sigma$ . Opredeljena je kot razlika vrednosti konkretne enote in aritmetične sredine populacije v razmerju s standardnim odklonom.  $Z$ -vrednost je negativna, kadar vrednost enote leži pod povprečjem populacije, in nasprotno, kadar ta vrednost leži nad povprečjem.

Računanje  $z$ -vrednosti tako kot množica drugih statističnih kazalcev temelji na teoriji verjetnosti oziroma predpostavlja normalno porazdelitev enot v statistični populaciji. Ker je mnogokrat težko ali nemogoče obravnavati celotno populacijo, se standardni odklon, ki je potreben za računanje  $z$ -vrednosti, izračunava z vzorčenjem.

Kazalec  $z$  je mogoče izračunati z naslednjo enačbo:

$$G_i^* = \frac{\sum_{j=1}^n w_{ij} x_j - \bar{X} \sum_{j=1}^n w_{ij}}{S \sqrt{\frac{n \sum_{j=1}^n w_{ij}^2 - \left(\sum_{j=1}^n w_{ij}\right)^2}{n-1}}} \quad (3)$$

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (4)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - \bar{X}^2}, \quad (5)$$

kjer je

- $x_j$  vrednost atributa nezgode  $j$  (teža ali relativna teža),
- $w_{ij}$  utež oddaljenosti med nezgodo  $i$  in nezgodo  $j$ ,
- $n$  skupno število nezgod,
- $\bar{X}$  povprečna vrednost atributa nezgode  $j$  (teža ali relativna teža),
- $S$  standardni odklon vzorca.

Ob normalni porazdelitvi vrednosti je na podlagi  $z$ -vrednosti mogoče določiti tudi kvantilni rang enote.

### 3.4.2 Jedrna gostota

Z uporabo funkcije jedrne gostote (angl. kernel density) izračunamo gostoto nezgod v mreži celic. Pri svoji analizi smo uporabili mrežo kvadratnih celic v velikosti 500 metrov.

Pri tej metodi najprej nad lokacijo vsake nezgode napnemo ukrivljeno ploskev. Če označimo  $z$  oddaljenost od lokacije nezgode,  $s$   $h$  radij ploskve, je v našem primeru to ploskev naslednje oblike:

$$x = \begin{cases} \frac{3}{k}(1-t^2)^2, & t \leq 1 \\ 0, & t > 1 \end{cases} \quad (6)$$

$$t = \frac{d}{h}, \quad (7)$$

kjer je

- $x$  vrednost ploskve,
- $k$  konstanta oziroma parameter.

Ploskev je najvišja nad točko nezgode in pada z oddaljenostjo od točke ter doseže vrednost 0 na oddaljenosti  $d=500$  m, ki smo jo vnesli kot vhodni podatek. Volumen lika pod ploskvijo je enak vrednosti atributa (teža ali relativne teža). Gostota v vsaki celici mreže se izračuna s seštevanjem vrednosti vseh ploskev, ki prekrivajo neko celico.

### 3.4.3 Lokalni Anselin-Moranov indeks

Z orodjem za razvrščanje v skupine in odkrivanje odstopajočih vrednosti (angl. cluster and outlier analysis) opredelimo območja s podobnimi in različnimi vrednostmi teža nezgod.

Anselin-Moranov indeks  $I$  predstavlja prostorsko korelacijo elementov izbrane množice. Visoka vrednost  $I$  pomeni, da se obravnavna lokacija nahaja v bližini lokacij, ki imajo podobno visoke vrednosti teže nezgod. To lahko pomeni, da je teža neke nezgode večja kot povprečna teža nezgod in imajo tudi sosednje nezgode večjo težo od povprečja, ali pa, da je teža nezgode manjša od povprečja in imajo tudi sosednje nezgode manjšo težo.

Indeks se izračuna po naslednji enačbi

$$I_i = \frac{z_i}{m_2} \sum_{j=1}^n W_{ij} z_j \quad (8)$$

$$z_i = x_i - \bar{x} \quad (9)$$

$$m_2 = \frac{1}{n} \sum_{i=1}^n z_i^2, \quad (10)$$

kjer je

$x_j$  vrednost atributa nezgode  $i$  (teža ali relativna teža),

$W_{ij}$  utež oddaljenosti med nezgodo  $i$  in nezgodo  $j$ .

#### 4 REZULTATI

Rezultati analize, ki so prikazani na sliki 1, potrjujejo zastavljeno domnevo, da prometne obremenitve bistveno vplivajo na število in lokacijo odkritih nevarnih krajev. Če se ne upoštevajo prometne obremenitve, je na slovenskih avtocestah 174 lokacij, kjer indeks Getis-Ord presega 1,96. Če se prometne obremenitve upoštevajo, pa je teh lokacij le 30. Za slednje lahko s 95-odstotno gotovostjo trdimo, da so nevarni kraji.

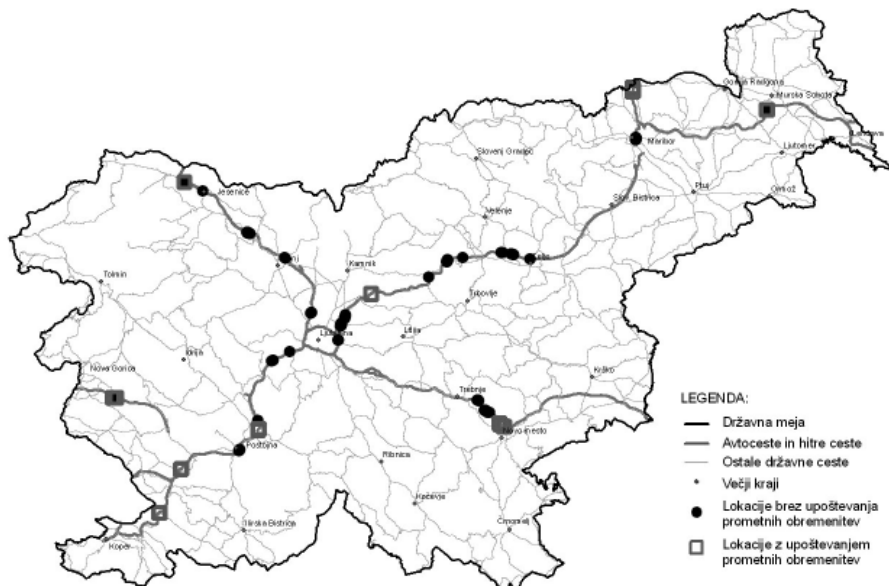
Zanimivo je, da se lokacije, kjer oba izračuna kažeta na nevaren kraj, le ponekod prekrivajo.

Ker je pri analizi nevarnih krajev smiselno upoštevati odseke v dolžini vsaj 500 metrov, lahko z uporabo funkcije določitve vplivnega območja (angl. buffer) te lokacije združimo. Tako nam ostane le 14 nevarnih krajev.

Z Anselin-Moranovim indeksom  $I$  odkrijemo samo 15 nevarnih krajev. Primerjava je prikazana na sliki 2.

Na sliki vidimo, da se lokacije večinoma prekrivajo, obstaja pa tudi lokacija, ki je bila ugotovljena samo z Anselin-Moranovim indeksom.

Za primerjavo z rezultati analize jedrne gostote smo pretvorili rezultate indeksa Getis-Ord  $G^*$  ter Anselin-Moranovega indeksa v rastrski zapis, slika 3 prikazuje rezultat analize jedrne gostote v rastrskem zapisu.



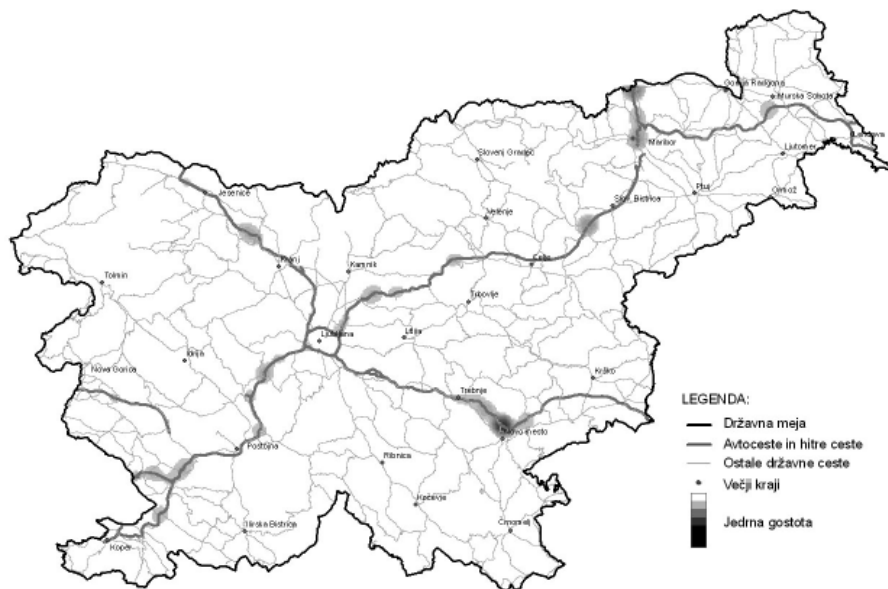
Slika 1: Primerjava ugotovljenih črnih točk z indeksom Getis-Ord  $G^*$  z upoštevanjem prometa in brez njega



Slika 2: Primerjava indeksa Getis-Ord  $G^*$  in Anselin-Moranovega indeksa

Primerjava rezultatov indeksa Getis-Ord  $G^*$  z rezultati jedrne gostote pokaže, da je od 30 lokacij, opredeljenih po Getis-Ord  $G^*$ , 21 (70 %) takšnih, ki jih prepoznamo z metodo jedrne gostote.

Primerjava rezultatov Anselin-Moranovega indeksa z rezultati jedrne gostote pa pokaže, da z metodo jedrne gostote odkrijemo 11 (73 %) lokacij izmed 15, ki jih sicer ugotovimo z uporabo Anselin-Moranovega indeksa.



Slika 3: Analiza nevarnih krajev z jedrno gostoto

## 5 SKLEP

Sodobne prostorske statistične metode omogočajo analize, ki jih z analizo atributnih podatkov ni bilo mogoče izvajati oziroma je bila obdelava podatkov veliko bolj zapletena. Z upoštevanjem prometnih obremenitev in uporabo metod prostorske statistike se število in lokacije odkritih nevarnih krajev precej spremenijo. To velja tudi za območja križišč, kjer se stika več odsekov državnih cest, saj se hkrati in enotno obravnavajo križišča in odprti odseki državnih cest.

S primerjavo indeksov je bilo dokazano, da je prilagojena prostorska statistika Getis-Ord  $G^*$  primerna metoda za določevanje zgostitev nevarnih krajev na cestah. Metodologijo bi bilo smiselno uvesti kot nadgradnjo obstoječe metodologije za potrebe rednega letnega določanja nevarnih krajev oziroma »črnih točk« v slovenskem državnem cestnem omrežju.

## Literatura in viri:

Anderson, T. K. (2009). *Kernel density estimation and K-means clustering to profile road accident hotspots*. *Accident Analysis and Prevention*, 41(3), 359–364.

Cheng, W., Washington, S. (2008). *New Criteria for Evaluating Methods of Identifying Hot Spots*. *Journal of the Transportation Research Board*, (2083), 76–85.

Direkcija RS za ceste, Omega Consult d. o. o. (2009). *Določitev nevarnih mest na državnem cestnem omrežju v Republiki Sloveniji*. *Končno poročilo*.

Moons, E., Brijis, T., Wets, G. (2008). *Hot Spot Analysis: Improving a Local Indicator of Spatial Association for*



- Application in Traffic Safety. Lecture Notes in Computer Science: Computational Science and Its Applications – ICCSA.*
- Moons, E., Brijs, T., Wets, G. (2009). *Improving Moran's Index to Identify Hot Spots in Traffic Safety. Studies in Computational Intelligence Geocomputation and Urban Planning.*
- Páez, A., Scott, D. M. (2005). *Spatial statistics for urban analysis: A review of techniques with examples. GeoJournal, 61(1), 53–67.*
- Sugihara, K., Okabe, A., Satoh, T. (2009). *Computational method for the point cluster analysis on networks. Geoinformatica Online.*
- Xie, Z., Yan, J. (2008). *Kernel density estimation of traffic accidents in a network space. Computers, Environment and Urban Systems, 32, 396–406.*

**Prispelo v objavo: 20. januar 2010**

**Sprejeto: 16. marec 2010**

**dr. Peter Lipar, univ. dipl. inž. grad.**

UL FGG, Prometnotehniški inštitut, Jamova 2, SI-1000 Ljubljana  
e-pošta: peter.lipar@fgg.uni-lj.si

**mag. Jure Kostanjšek, univ. dipl. inž. grad.**

UL FGG, Prometnotehniški inštitut, Jamova 2, SI-1000 Ljubljana  
e-pošta: jure.kostanjsek@fgg.uni-lj.si

**dr. Marijan Žura, univ. dipl. inž. grad.**

UL FGG, Prometnotehniški inštitut, Jamova 2, SI-1000 Ljubljana  
e-pošta: marijan.zura@fgg.uni-lj.si