

SLOVENSKÁ ŠTATISTIKA a DEMOGRAFIA

SLOVAK STATISTICS
and DEMOGRAPHY

1/2024

ročník/volume 34

Recenzovaný vedecký časopis so zameraním na prezentáciu moderných štatistických a demografických metód a postupov.

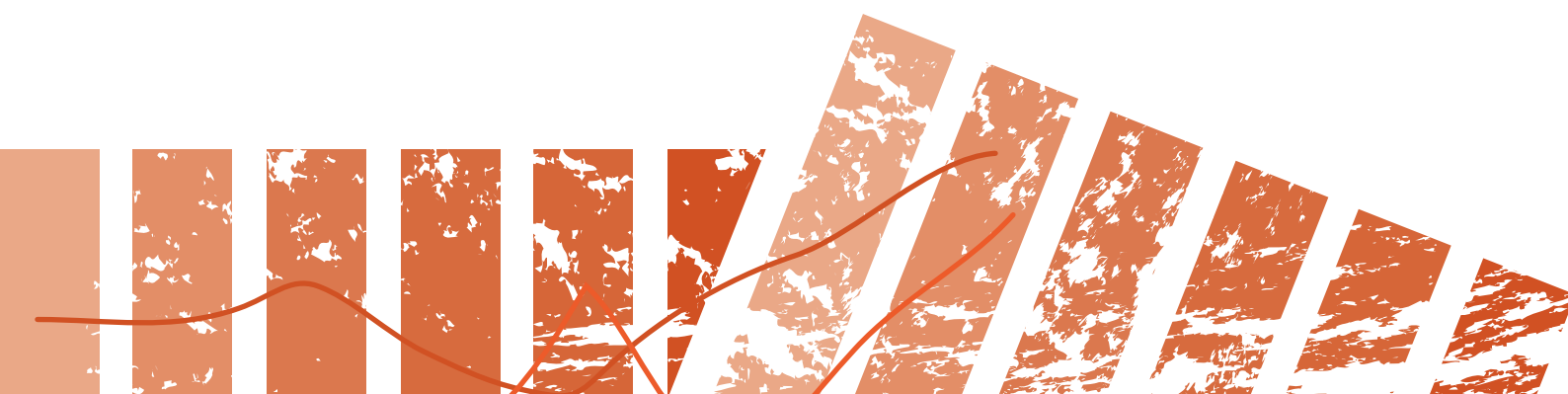
Scientific peer-reviewed journal focusing on the presentation of modern statistical and demographic methods and procedures.

Článok/Article: 4

Typ článku/Type of article: vedecký článok/scientific article

Strany/Pages: 27 – 51

Dátum vydania/Publication date: 15. január 2024/January 15, 2024



Martin ŠVEDA, Michala SLÁDEKOVÁ MADAJOVÁ
Prírodovedecká fakulta Univerzity Komenského, Geografický ústav SAV, v. v. i.
Pavol HURBÁNEK, Konštantín ROSINA
Geografický ústav SAV, v. v. i.

SPRACOVANIE PASÍVNYCH LOKALIZAČNÝCH ÚDAJOV MOBILNEJ SIETE NA POUŽITIE V EXPERIMENTÁLNEJ POPULAČNEJ ŠTATISTIKE

PROCESSING OF PASSIVE MOBILE POSITIONING DATA FOR USE IN THE EXPERIMENTAL POPULATION STATISTICS

ABSTRAKT

Mobilný telefón sa stal neoddeliteľnou súčasťou každodenného života a unikátnym zdrojom údajov o obyvateľstve, jeho priestorovom rozmiestnení, mobilite a aktivitách. Cieľom príspevku je opísať konceptuálny a metodický prístup spracovania pasívnych lokalizačných údajov mobilnej siete a pokúsiť sa odpovedať na nasledujúce výskumné otázky: Dokážu lokalizačné údaje z mobilnej siete poskytnúť relevantnú informáciu o priestorovom rozmiestnení populácie na Slovensku? Môžu byť tieto údaje relevantným doplnkom ku štatistickým údajom o obyvateľstve a aké sú ich limity? Výsledky spracovania lokalizačných údajov mobilnej siete priniesli robustné výsledky, ktoré môžu prispieť k hlbšiemu poznaniu priestorového rozmiestnenia obyvateľstva. Pri rešpektovaní viacerých limitov dostávame perspektívne dáta, vhodné na rozmanité analýzy na úrovni regiónov alebo lokalít.

ABSTRACT

The mobile phone has become an integral part of everyday life and a unique source of data on the population, including its spatial distribution, mobility, and activities. The aim of this paper is to describe the conceptual and methodological approach of processing passive mobile positioning data and to address the following research questions: Can mobile positioning data provide meaningful information about the spatial distribution of the Slovak population. Can data from the mobile network be a meaningful complement to conventional population statistics, and what their limits are. The results of processing mobile network location data have yielded powerful results that can contribute to a deeper understanding of the spatial distribution of the population. Despite several limitations, we obtain prospective data suitable for diverse analyses at the regional and local level.

KLÚČOVÉ SLOVÁ: mobilná sieť, lokalizácia, odhad populácie, Slovensko

KEY WORDS: mobile network, localization, population estimates, Slovakia

1. ÚVOD

Poznanie priestorovej distribúcie a mobility obyvateľov je kľúčovým vstupom pre množstvo vedeckých analýz a praktických úloh. Tradičným zdrojom týchto úloh sú údaje z evidencie obyvateľstva a z celoštátneho cenzu. Napriek tomu, že ide o koncepčne a metodicky prepracované zdroje údajov, vývoj spoločnosti prináša nové fenomény, na ktoré konvenčné zdroje údajov nemusia dostatočne spoľahlivo reagovať. Osobitne to platí pre spoločensky orientované vedné disciplíny, ktoré potrebujú aktuálne, presné a detailné údaje o rozmanitých fenoménoch časovo-priestorového správania populácie. Je zrejmé, že ak chceme reagovať na spoločenský

vývoj, potrebujeme viac participovať na možnostiach, ktoré prinášajú informačno-komunikačné technológie. Výsledkom adaptácie výskumných metód na nové podmienky by malo byť predovšetkým využitie informačno-komunikačných technológií, ktoré prinášajú nielen zmenu paradigmy priestorového správania, ale aj prístupu k jeho sledovaniu.

Motívov na použitie iných prístupov je viacero. Predovšetkým pracovné aktivity obyvateľov sa neviažu na jedno miesto tak pevne, ako to bývalo v minulosti a miesto trvalého pobytu, miesto bývania a miesto práce sa čoraz častejšie od seba odlišujú [30, 41]. Pohyby obyvateľstva sa stávajú pestrejšími, nadobúdajú nerutinný a nepravidelný charakter. V súčasnosti nie je neobvyklé bývanie v podnájme, sezónne bývanie na chate či dlhodobé pracovné stáže. Čoraz menšia viazanosť obyvateľstva na jedno miesto sa prejavuje v rôznych spoločenských kontextoch a relativizuje výsledky sociálnych, geografických či ekonomických analýz založených na tradičných štatistických zdrojoch (cenzy, registre obyvateľstva a pod.). Pre časť populácie je zložitá jednoznačne odpovedať na otázku miesta bydliska a miesta práce. Túto skutočnosť pritom nemôžeme vnímať ako dočasné „uvoľnenie“ viazanosti na základné priestorové súradnice nášho každodenného života, ale ako dlhodobé trendy prinášajúce zvýšenú fluktuáciu a mobilitu obyvateľstva, ktoré stimulujú viaceré fenomény postmodernej spoločnosti, predovšetkým informačno-komunikačné technológie [8]. Tie uľahčili nielen nástup práce z domu, ale aj ekonomiky postavenej na nezávislých dodávateľoch a externých pracovníkoch (*gig economy*), to všetko s významným vplyvom na spôsob práce a mobility obyvateľov, najmä v metropolitných regiónoch.

Aktuálnym fenoménom, ktorý pravdepodobne prispel k zmene priestorových vzorov správania, je pandémia ochorenia COVID-19. Potreba sociálnej separácie a obmedzenia každodenného života nás prinútili adaptovať sa na nové spôsoby práce, učenia sa a interakcií. Hoci obmedzenia boli len dočasné, môžu mať za následok dlhodobé zmeny v priestorových preferenciách pre prácu, bývanie a každodennú mobilitu [26]. Dnes je však ešte predčasné predpovedať charakter zmien v priestorových prejavoch populácie. Je však ale možné, že stojíme na prahu novej paradigmy, ktorá ich bude definovať najbližšie desaťročia. Je otázne, ako na uvedené zmeny v spoločnosti dokážeme reagovať prostredníctvom konvenčných zdrojov údajov.

V nových zdrojoch údajov, ktoré nám môžu pomôcť porozumieť trendom a fenoménom doby, má osobitné postavenie využitie údajov z mobilnej siete. Mobilný telefón sa stal neoddeliteľnou súčasťou každodenného života a unikátnym zdrojom údajov o obyvateľstve, jeho priestorovom rozmiestnení, mobilite a aktivitách. Myšlienky využitia týchto údajov v priestorových analýzach sa objavili súčasne s masovým rozvojom mobilnej komunikácie na báze GSM [4, 28, 35], avšak až v ostatnom desaťročí sme svedkami skutočne rozsiahleho rozvoja v tejto výskumnej oblasti.

Intenzívny záujem o údaje z mobilnej siete vyvolala nielen prirodzená „zvedavosť“ výskumníkov, ale aj viaceré praktické úlohy, pred ktorými stojí súčasná globalizovaná spoločnosť. Záujem o údaje z mobilnej siete prejavili viaceré medzinárodné organizácie, inštitúcie Európskej únie či národné štatistické úrady. Príkladom sú aktivity *Joint Research Centre* pri Európskej komisii, kde sa podrobne zaoberali využitím údajov z mobilnej siete pre spresnenie priestorového rozmiestnenia populácie

[36]. Spomedzi krajín EÚ sa azda najďalej dostalo Estónsko, kde výskumníci pod vedením R. Ahasa vybudovali komplexnú metodiku pre sledovanie pohybu zahraničných návštevníkov [2].

Lokalizačné údaje mobilnej siete majú v porovnaní s tradičnými údajmi o dochádzke niekoľko pozitívnych vlastností. V prvom rade poskytujú aktuálne informácie, ktoré sú pomerne rýchlo spracovateľné a nevyžadujú si aktívnu participáciu obyvateľov. Nie sú teda poznačené časovým odstupom, ako to spravidla býva v celoplošných cenzoch. Dostupnosť týchto údajov prakticky v akomkoľvek čase umožňuje pružne reagovať na vývoj spoločnosti a aktuálne zmeny v priestorovom rozmiestnení obyvateľstva (výstavba bytov, ciest, koncentrácia novej výroby a pod.). Navyše prostredníctvom mobilnej siete môžeme sledovať zmeny v dennej priestorovej mobilite spôsobené sezónnymi cyklami a flexibilne prispôbovať rozsah pozorovania spoločenským či výskumným potrebám, respektíve vytvárať dlhodobé pozorovania a sledovať tak reakcie spoločenských procesov na zmeny v intenzite a smerovaní mobility populácie [14, 46]. Skutočnosť, že lokalizačné údaje mobilných zariadení vznikajú na úrovni jednotlivých buniek mobilnej siete, umožňuje agregáciu podľa rozmanitých priestorových schém, ktoré nemusia rešpektovať územnosprávne členenie. To nám umožňuje získavať informácie o rozmiestnení a mobilite populácie na subregionálnej a lokálnej úrovni, mierke bežne nedostupnej s využitím údajov viazaných na administratívne členenie Slovenska. To má význam osobitne pri sledovaní procesov na intraurbánnej úrovni.

Perspektívy tohto nekonvenčného zdroja údajov si uvedomuje aj Štatistický úrad SR. V spolupráci s ním bol realizovaný projekt Socioekonomické aspekty Big data s cieľom preskúmať možnosti využitia údajov z mobilnej siete v experimentálnej populačnej štatistike. Cieľom príspevku je opísať konceptuálny a metodický prístup spracovania lokalizačných údajov mobilnej siete v tomto projekte a pokúsiť sa odpovedať na nasledujúce výskumné otázky:

- Dokážu údaje z mobilnej siete poskytnúť relevantnú informáciu o priestorovom rozmiestnení a mobilite populácie na Slovensku?
- Môžu byť údaje z mobilnej siete relevantným doplnkom ku štatistickým údajom o obyvateľstve a aké sú ich limity?

2. VYUŽITIE ÚDAJOV MOBILNEJ SIETE NA SPRESNENIE PRIESTOROVÉHO ROZMIESTNENIA POPULÁCIE

Nekonzistencia medzi oficiálnym (štatistickým) počtom obyvateľov a skutočným stavom obyvateľstva v priestorových jednotkách rôznej úrovne sa v odborných kruhoch na Slovensku sleduje už približne od polovice 90. rokov minulého storočia (napr. [11, 13, 32, 41]). Napriek pokusom o aproximáciu reálneho počtu obyvateľov využitím zástupných ukazovateľov, ako je napríklad kapacita bytového fondu, musíme konštatovať, že v súčasnosti nepoznáme reálne priestorové rozmiestnenie obyvateľstva na Slovensku. Uvedomenie si tejto triviálnej skutočnosti má pritom kľúčový význam pre sociálno-ekonomické analýzy, hodnotenie transformačných procesov či prognózovanie populačného a hospodárskeho vývoja regiónov, ako aj pre ďalšie praktické opatrenia (rozpočtovanie daní, kreovanie obecných zastupiteľstiev a volebných okrskov a pod.).

Tento problém sa netýka len Slovenska, ale predstavuje globálny fenomén, ktorému sa vo svetovej literatúre venuje značná pozornosť. Súčasná spoločnosť je viac ako

kedykoľvek predtým tvorená tokmi ľudí, tovarov a informácií, avšak dáta, ktoré tieto toky zaznamenávajú, sú pomerne náročné na spracovanie využitím tradičných prístupov spoločenského výskumu [34, 38, 39]. Od prvých štúdií [18, 25, 45] až po najnovšie prístupy (napr. projekt ENACT – [6, 7], sa metódy mapujúce priestorové rozmiestnenie populácie sa postupne zdokonaľovali a poskytovali čoraz presnejší obraz rezidenčnej populácie v rôznych častiach sveta. Vďaka týmto aktivitám mohli byť tradičné mapy znázorňujúce hustotu obyvateľstva v administratívnych jednotkách, nahradené realistickejšími zobrazeniami distribúcie populácie prostredníctvom buniek pravidelných gridov označovaných ako populačné gridy [6]. Vysoké priestorové rozlíšenie populačných gridov (najčastejšie bunky o veľkosti 1 x 1 km) umožnilo širšiu integráciu týchto údajov v geografických informačných systémoch a tieto údaje sa stali nenahraditeľnými pre sociálny a environmentálny výskum či územné plánovanie.

Napriek značnému vedeckému pokroku sú populačné gridy vo svojej podstate naďalej „len“ statickými záznamami rezidenčnej populácie, ktoré zachytávajú rozmiestnenie obyvateľstva počas noci, a to za predpokladu, že väčšina ľudí zotrúva v rámci deklarovaného miesta pobytu. Napriek tomu, že populačné gridy poskytujú rozmanité využitie, opisujú len časť reality. Priestorové rozmiestnenie populácie počas dňa, respektíve v priebehu rôznych častí roka, je prakticky neznáme pre ľubovoľnú priestorovú mierku. Takéto informácie sú však nevyhnutné pre celý rad aplikácií. Distribúcia obyvateľstva počas dňa je determinovaná rozmiestnením hospodárskych, sociálnych a rekreačných zariadení, ktoré priťahujú obyvateľstvo z ich bydliska, predurčujú dochádzkové toky a iné formy dennej mobility. Denné rozmiestnenie populácie sa teda (značne) líši od rezidenčnej (nočnej) populácie. Je preto zrejmé, že modelovanie rozmiestnenia populácie nedokážeme efektívne realizovať len s využitím konvenčných zdrojov údajov.

Jedným z riešení, ako sa vyrovnat' s chýbajúcimi údajmi, je využitie údajov z mobilnej siete. So zvyšujúcim sa výskytom mobilných telefónov v populácii sa lokalizácia v mobilnej sieti stáva dôležitým zdrojom údajov, ktorý môže dopĺňať a rozširovať existujúcu populačnú štatistiku. Tento nový zdroj údajov nám umožňuje sledovať, ako sa rozmiestnenie populácie mení nielen v čase, ale aj v závislosti od rôznych cyklov (denný, týždenný, sezónny a pod.) či akejkoľvek udalosti, ktorá ovplyvňuje koncentráciu populácie (športové podujatia, prírodné katastrofy). Uplatnenie týchto údajov tak nachádzame nielen vo vedeckom výskume, ale čoraz častejšie aj v rozhodovacej praxi, kde prispievajú k zvyšovaniu kvality a efektívnosti územného plánovania a manažmentu.

Údaje z mobilnej siete majú potenciál vytvoriť automatizované odhady rozmiestnenia populácie prakticky v reálnom čase. Hoci nedosahujú presnosť a spoľahlivosť censov, ich porovnanie s konvenčnými zdrojmi (národnými censami či populačnými registrami) prináša pomerne vysoké korelácie [20, 24, 30]. Aktuálne metodické prístupy [16, 17, 27] využívajú lokalizačné údaje z mobilnej siete v úzkej synergii so satelitným snímkovaním a rozmanitými mapovými vrstvami (zastavané územie, využitie zeme, funkčné areály). Môžeme očakávať, že pri spracovaní a analýze časovo-priestorových záznamov z mobilnej siete sa budú čoskoro využívať aj rozšírené možnosti umelej inteligencie [21].

3. LOKALIZÁCIA V MOBILNEJ SIETI

Princíp lokalizácie v mobilnej sieti je prirodzenou nadstavbou základných vlastností mobilnej komunikácie. Z priestorového pohľadu môžeme územie rozdeliť do buniek, ktoré sú obsluhované jednotlivými anténami mobilnej siete. Každá anténa je schopná pokryť určité územie a obslúžiť určitý počet zákazníkov. Identifikačné údaje o aktuálne využívanej anténe, ako aj ďalšie doplnkové informácie sa môžu využiť na určenie približnej polohy mobilného zariadenia. Tieto lokalizačné údaje sú prirodzenou súčasťou mobilnej komunikácie, keďže identifikácia základňových staníc (*base transceiver station*, BTS), s ktorými komunikuje mobilný telefón, je nevyhnutná pre samotné fungovanie mobilnej siete.

Lokalizačné údaje, ktoré môžeme získať prostredníctvom lokalizácie v mobilnej sieti, principiálne rozdelujeme na aktívne a pasívne [3]. Kým pri pasívnom type ide o využitie existujúcich záznamov v systéme mobilného operátora, pri aktívnom type sa záznam vytvára na základe konkrétneho dopytu a s využitím špecializovaného softvéru. Pasívne lokalizačné údaje sú teda digitálne „stopy“, ktoré zanecháva mobilné zariadenie na infraštruktúre mobilnej siete. Tieto záznamy vznikajú buď pre potreby vyúčtovania hovorov, SMS a pod., alebo sú dôsledkom pravidelných aktualizácií polohy zariadenia v mobilnej sieti. Vzhľadom na skutočnosť, že k údajom o aktivite mobilného zariadenia vieme priradiť nielen približnú polohu (konkrétnu bunku mobilnej siete), ale aj niektoré základné informácie o používateľovi (vek, pohlavie či fakturačná adresa), získavame bohatú databázu, ktorej využitie (pri zachovaní anonymity používateľov, resp. pri čiastočnom agregovaní údajov) prináša nesmierne cenný analytický nástroj.

Lokalizácia mobilného zariadenia na úrovni buniek mobilnej siete umožňuje analyzovať priestorové rozmiestnenie obyvateľstva v relatívne detailnej mierke a v každom okamihu. Tieto vlastnosti predurčujú takýto zdroj údajov na široké multidisciplinárne uplatnenie. Svedčí o tom aj pestrá paleta aplikácií, ktorú nachádzame vo svetovej literatúre [40, 48]. Využitím údajov z mobilnej siete môžeme sledovať nielen statický obraz priestorového rozmiestnenia populácie na úrovni jednotlivých lokalít (napr. obcí, urbanistických obvodov a pod.), ale aj zmeny v koncentrácii obyvateľstva v ľubovoľných časových rezoch (napr. každú hodinu). To má zásadný význam predovšetkým pre lepšie porozumenie denných rytmov lokalít [43] a vo výskume priestorovej mobility [15, 22]. Bezprostredné využitie nachádzajú takéto údaje aj v krízovom manažmente, v ktorom pomáhajú distribuovať záchranné zložky a kapacity na evakuáciu obyvateľstva [9], prípadne pomáhajú predikovať vývoj epidémií [19, 49].

4. SPRACOVANIE PASÍVNYCH LOKALIZAČNÝCH ÚDAJOV MOBILNEJ SIETE

Základným krokom na relevantné využitie mobilnej lokalizácie v množstve priestorovo orientovaných analýz je určenie pravidelnej dennej a nočnej lokality používateľiek a používateľov mobilnej siete. Metodický a konceptuálny rámec vychádza z konceptu tzv. kotevných bodov [5]. Kotevné body sa chápu ako hlavné uzly aktivít človeka, ktoré vytvárajú kostru jeho každodenných pohybov. V praxi ide predovšetkým o identifikovanie základných kotevných bodov „domov“ a „práca“. Údaje o koncentrácii denných a nočných kotevných bodov nám umožňujú nielen spresniť priestorovú distribúciu obyvateľstva (napr. koľko ľudí býva v časti mesta), ale aj extrahovať údaje o predpokladaných denných dochádzkových tokoch (napr. koľko ľudí dochádza denne do danej časti mesta).

Originálna metodika využíva rozsiahle údaje z mobilnej siete na odhad rozmiestnenia populácie prostredníctvom identifikovania pravidelnej dennej a nočnej lokality (bunky mobilnej siete) používateľov mobilných zariadení. Cieľom štúdie bolo vypracovať metodický rámec na zhromažďovanie a spracovanie lokalizačných údajov mobilnej siete, ktorý je vhodný na použitie medzi viacerými poskytovateľmi služieb mobilnej komunikácie. Hlavnou výzvou tejto úlohy bolo navrhnúť metodiku, ktorá by bola všeobecne aplikovateľná aj v heterogénnych scenároch, kde niekoľko technických detailov konfigurácie sietí a organizácie dát zostáva špecifických pre jednotlivých poskytovateľov mobilných služieb.

S týmto cieľom sa snažíme o vytvorenie „odolného“ metodického rámca, pričom základný súbor funkcií sa nespolieha na žiadnu neštandardnú konfiguráciu špecifickú pre jednotlivých mobilných operátorov a zároveň je dostatočne flexibilný na využitie prípadných osobitných charakteristík mobilnej siete a/alebo údajov špecifických pre jednotlivých mobilných operátorov. Vďaka tejto flexibilitě sa navrhovaná metodika môže v budúcnosti rozšíriť a ďalej zdokonaľiť, pričom sa počíta s prirodzenou evolúciou infraštruktúry mobilnej siete (napríklad zmenšovanie plochy buniek a nárast ich počtu).

Vstupné dáta reprezentovali pasívne lokalizačné údaje mobilnej siete od troch najväčších mobilných operátorov na Slovensku (Slovak Telekom, Orange Slovensko, O2 Slovakia), ktoré boli spracované na základe vopred stanovenej metodiky s cieľom identifikovať pravidelnú nočnú a dennú lokalitu mobilných zariadení (SIM kariet) na úrovni buniek mobilnej siete a následne transformovať tieto údaje do siete populačného gridu 1x1 km alebo katastrálneho územia obcí. Súčasťou datasetov sú aj matice sumarizujúce vektory pravidelnej dennej a nočnej lokalizácie individuálnych používateľov mobilnej siete. V nasledujúcom texte predstavíme štruktúru údajov, spôsob extrahovania pravidelných lokalizácií a transformáciu údajov z úrovne buniek mobilnej siete do cieľových priestorových zón.

4.1. Štruktúra pasívnych lokalizačných údajov mobilnej siete

Pasívne lokalizačné údaje z mobilnej siete sa principiálne skladajú z dvoch datasetov. Prvý predstavujú individuálne záznamy mobilného zariadenia na infraštruktúre mobilnej siete. Obsahujú identifikátor telefónneho čísla asociovaného so SIM kartou (MSISDN), dátum, časovú známku a identifikátor bunky mobilnej siete (tabuľka č. 1). Druhý dataset obsahuje priestorovú geometriu vyžarovacích polygónov pre jednotlivé bunky mobilnej siete (tabuľka č. 2). Vyžarovací polygón predstavuje mnohostranný plošný útvar opisujúci pokrytie signálom práve jednej BTS bunky. Vzhľadom na citlivosť oboch datasetov tieto záznamy používa len mobilný operátor na generovanie vopred definovaných časových a priestorových agregátov. Hoci sa údaje vzťahujú na individuálnu SIM kartu, pre jednoduchosť budeme v ďalšej časti príspevku uvažovať o ideálnej situácii, keď jedna SIM karta reprezentuje jedného používateľa mobilnej siete. Limitom tohto zjednodušenia sa budeme podrobnejšie venovať v 6. kapitole.

Tabuľka č. 1: Individuálne záznamy v mobilnej sieti

msisdn	dátum	čas	Id bunky mobilnej siete
ID_SIM 001	15.5.2023	18:02:05	BTS_0001

Zdroj: vlastné spracovanie autorov

Tabuľka č. 2: Priestorová geometria vyžarovacích polygónov BTS definovaná ako zlomové body multipolygónu

enodeb_id	Priestorová geometria vyžarovacieho polygónu
BTS_0001	17,0574; 48,1556 ; 17,0566, 48,1776 ;...

Zdroj: vlastné spracovanie autorov

4.2. Extrakcia pravidelnej dennej a nočnej lokalizácie mobilného zariadenia v sieťovej lokalizácii

Základným krokom na relevantné využitie mobilnej lokalizácie pri snahe o zachytenie priestorového rozmiestnenia populácie je určenie pravidelnej dennej a nočnej lokality užívateľov mobilnej siete. V praxi existuje celý rad techník (pozri napr. [5, 50, 51] na extrahovanie relevantných denných (pracovných), nočných (domácich) či iných periodických lokalít z množstva polohových záznamov, ktoré vznikajú v mobilnej sieti. Principiálne môžeme uvažovať o dvoch komplementárnych prístupoch identifikácie periodickej lokalizácie. Pri prvom prístupe určíme tú bunku mobilnej siete, ktorá zaznamenala najväčší počet lokalizácií (záznamov) v danom čase (napr. počas nočných hodín). V druhom prípade je cieľom identifikovať bunku, na ktorej strávilo mobilné zariadenie najdlhší čas. V oboch prípadoch je dôležité v čo najväčšej miere eliminovať náhodné a nerutinné lokalizácie. Na vyprofilovanie rutinných denných, nočných a iných pravidelných lokalizácií je dôležitá dostatočná dĺžka pozorovania, ktorú je potrebné prispôsobiť sledovaným javom, finančným možnostiam, ako aj výpočtovým kapacitám. Na základe predchádzajúcich skúseností [42, 43] boli algoritmy na identifikáciu pravidelných denných a nočných lokalít užívateľov mobilnej siete skonštruované zo 4-týždňového pozorovania takto:

Pravidelná nočná lokalita je územie pokrytia signálom práve jednej bunky mobilnej siete, v ktorom bola SIM karta najčastejšie lokalizovaná počas pracovného týždňa (PO – PIA) v čase od 23:00 do 5:59 hod. Táto lokalita bola identifikovaná zo súboru 20 pracovných dní tak, že za každú hodinu pozorovania sa identifikovala BTS stanica, kde mala SIM karta najväčší počet záznamov. Vzhľadom na 16 nocí pozorovania a 7 hodinových intervalov počas dňa ide o 112 časových intervalov. Bunka mobilnej siete s najväčším počtom lokalizácií (maximálne 112) bola určená ako pravidelná nočná lokalita danej SIM karty.

Pri väčšine používateľov sa pravidelná nočná lokalita extrahovala z rádovo stoviek záznamov. Mohli však nastať situácie, keď sa nočná lokalita určila len z malého počtu záznamov. V realite môže ísť napríklad o používateľov mobilnej siete, ktorí bývajú v zahraničí a na území Slovenska sa vyskytli len počas dňa. Vzhľadom na to sa aplikovala dodatočná podmienka, ktorá ponechala len tých používateľov mobilnej siete (SIM karty), ktorí strávili na infraštruktúre mobilnej siete minimálne 3 hodiny počas pozorovania v pracovnom týždni v čase od 23:00 do 5:59 hod. Túto podmienku bolo potrebné splniť minimálne 9 zo 16 nocí pozorovania.

Analogicky bola identifikovaná pravidelná denná lokalita. S cieľom eliminovať nerutinné vplyvy pondelkového a piatkového režimu sa použilo kratšie obdobie pozorovania, ktoré zahŕňalo UT – ŠT v čase od 9:00 do 14:59 hod. (tabuľka č. 3).

Tabuľka č. 3: Podmienky na extrahovanie pravidelnej nočnej a dennej lokality mobilného zariadenia (SIM karty) z pasívnych lokalizačných údajov mobilnej siete

	Dĺžka pozorovania	dni počas týždňa	čas	počet hodinových intervalov	Minimálna dĺžka záznamu
Pravidelná nočná lokalita SIM karty	28 dní	PO – PI	23:00 – 05:59	112	aspoň 3 hodiny počas 9 nocí pozorovania
Pravidelná denná lokalita SIM karty		UT – ŠT	09:00 – 14:59	72	aspoň 3 hodiny počas 7 dní pozorovania

Zdroj: vlastné spracovanie autorov

4.3. Transformácia lokalizačných záznamov z mobilnej siete do cieľových územných rámcov

Na praktické využitie uvedených kategórií v populačnej štatistike bolo potrebné tieto údaje transformovať do cieľových územných zón, v tomto prípade buniek gridu 1 x 1 km (50 661 buniek) a do katastrálnych území obcí (2 927 obcí). Pri transformácii údajov (počtu SIM kariet s pravidelnou nočnou/dennou lokalizáciou) do cieľových priestorových jednotiek sme použili nebinárnu dazymetrickú metódu¹ s využitím pomocnej vrstvy objemu budov. Výber tejto transformačnej metódy bol výsledkom testovania viacerých perspektívnych metód [44]. Pomocná vrstva vychádza zo spracovania polygónov budov databázy ZBGIS (2017) a bola skonštruovaná v dvoch verziách:

1) Na transformáciu údajov o počte SIM kariet s pravidelnou nočnou lokalitou sa použili polygóny rodinných domov a bytových domov. Na spresnenie interpolácie sa použili odlišné váhy pre objem bytových domov a rodinných domov v pomere 3 : 1. Tento pomer je výsledkom kalibrácie modelu na testovacej vzorke údajov, ktoré reprezentovali približne 1/50 územia Slovenska.

2) Na transformáciu údajov z dennej lokalizácie sa použil súbor rezidenčných aj nerezidenčných typov budov, v ktorých môžeme predpokladať prítomnosť obyvateľov počas dňa (napr. školy, administratívne budovy, priemyselné budovy a pod). Objem budov bol odhadnutý prostredníctvom rozlohy zastavanej plochy a maximálnej výšky budovy. Prístup používajúci objem budov je vhodný osobitne vo vysoko

¹ Dazymetrická metóda pracuje na princípe tzv. dazymetrického mapovania, keď sú dáta znázornené prostredníctvom hraníc, ktoré rozdeľujú mapované územie do zón relatívnej homogenity. Umožňuje nám teda jemnejšie prerozdelenie údajov v zdrojovej zóne, a to prostredníctvom rôznych pomocných informácií (ancillary data). Tieto pomocné informácie sa týkajú študovaného územia a zvyčajne sú to vrstvy využitia zeme či krajinej pokrývky. Prekryvom zdrojových jednotiek (bunky mobilnej siete) a vrstvami pomocných informácií (napr. vrstva zastavaných areálov) dostávame tzv. dazymetrické zóny, ktoré môžeme prepojiť na cieľovú zónu (napr. katastrálne územia obcí).

urbanizovanom území, kde je vertikálna dimenzia a objem budov kľúčovým atribútom, ktorý determinuje priestorovú distribúciu obyvateľov [12, 23, 47].

4.4. Vyhodnotenie modelu

Presnosť výsledkov sme hodnotili porovnaním s referenčnými údajmi z národného cenzu (SODB 2021) a pomocou niekoľkých korelačných mier, vrátane Pearsonovej korelácie (r), Spearmanovej korelácie (s), Kendallovej korelácie (τ), strednej kvadratickej odchýlky (RMSE) a relatívnej celkovej absolútnej chyby (RTAE). Použitie RTAE ponúka možnosť relativizovať výslednú celkovú absolútnu chybu k najlepšiemu dostupnému odhadu populácie v záujmovom území, ktorý môžeme vyjadriť ako súčet referenčných hodnôt všetkých buniek v sledovanom území (S) podľa vzťahu (1):

$$S = \sum_{i=1}^n P_i \quad (1)$$

Výsledkom tejto relativizácie je relatívna celková absolútna chyba (RTAE) vyjadrená ako podiel nesprávne alokovanej populácie vzhľadom na počet obyvateľov v území:

$$RTAE = \frac{\sum_{i=1}^n |X_i - P_i|}{\sum_{i=1}^n P_i} \times 100 (\%) = \frac{\sum_{i=1}^n |X_i - P_i|}{S} \times 100 (\%) \quad (2)$$

kde P_i je veľkosť referenčnej populácie, teda počet obyvateľov (počet obyvateľov evidovaných na súčasný pobyt v zóne i podľa údajov z cenzu 2021) a X_i je modelovaný počet obyvateľov v zóne i , definovaný podľa vzťahu (3) ako:

$$X_i = k \times U_i \quad (3)$$

kde U_i je odhadovaný počet používateľov mobilnej siete v zóne i a k je koeficient zabezpečujúci, že:

$$\sum_{i=1}^n P_i = \sum_{i=1}^n X_i = k \times \sum_{i=1}^n U_i \quad (4)$$

4.5. Anonymizácia

Kvôli citlivosti jednotlivých záznamov a dodržiavaniu platných zákonov o ochrane osobných údajov boli údaje anonymizované a agregované prostredníctvom infraštruktúry prevádzkovateľa mobilnej siete. Aj po zoskupení obsahovali niektoré zóny (bunky gridu/obce) len malý počet alokovaných používateľov. Aby sme predišli riziku práce s malými číslami, zónam s 1 až 3 používateľmi sme priradili priemernú hodnotu podľa ich štatistickej distribúcie.

5. VÝSLEDKY

Skôr ako sa pozrieme na výsledky modelu, je potrebné uviesť dôležitú interpretačnú poznámku. Spracované údaje z mobilnej siete tvorí počet alokovaných SIM kariet v priestorových jednotkách (grid, obce). Hoci počet SIM kariet nemôžeme spoľahlivo

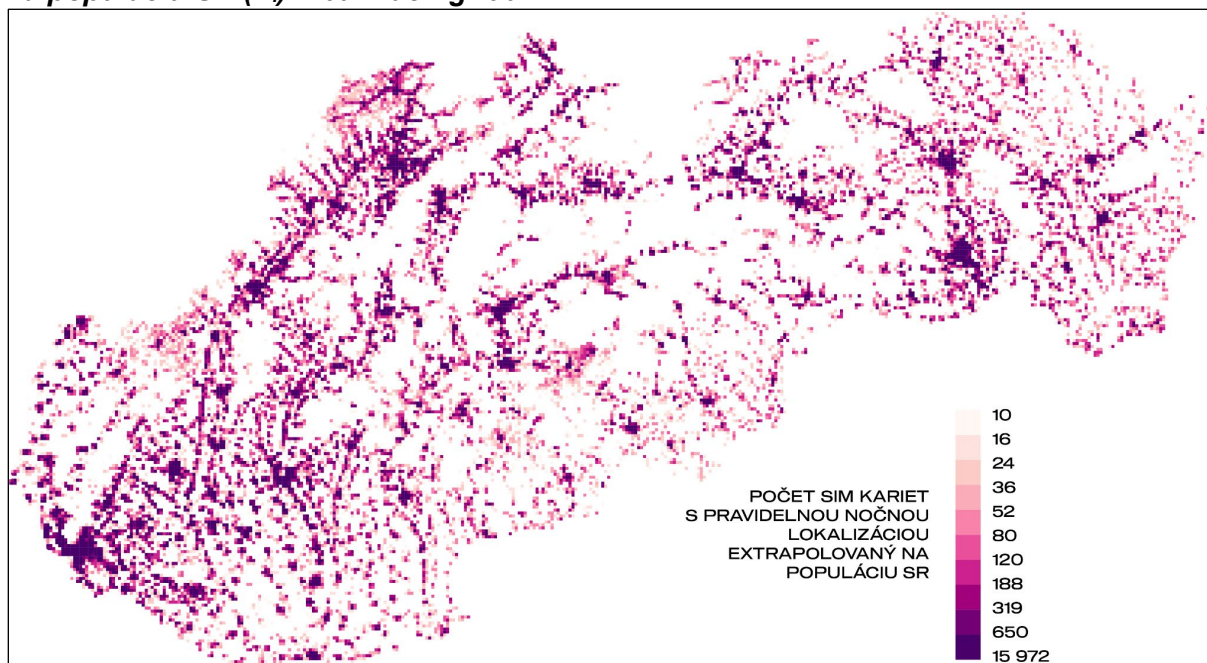
stotožniť s počtom obyvateľov, na účely vyhodnotenie globálnej štatistiky a presnosti modelu prijmem predpoklad, že jedna SIM karta reprezentuje jedného obyvateľa.

5.1. Vyhodnotenie validity údajov pravidelnej nočnej lokalizácie SIM kariet na úrovni gridu 1 x 1 km

V rámci zvolenej metodiky sme hľadali pravidelné nočné a denné lokalizácie individuálnych používateľov mobilnej siete (SIM kariet), ktoré by mohli slúžiť na odhad prítomnej populácie. V prvom priblížení sme alokovali pravidelnú dennú alebo nočnú polohu mobilného zariadenia do populačného gridu 1 x 1 km, ktorý predstavuje mriežku zloženú z pravidelnej siete štvorcových buniek obsahujúcich počet obyvateľov v Európskom štatistickom systéme. Ako zachytáva obrázok č. 1, rozmiestnenie používateľov mobilnej siete (SIM kariet) prináša očakávaný priestorový obraz s koncentráciou populácie do urbanizovaného územia Slovenska.

Na vyhodnotenie presnosti použitého prístupu sme skonštruovali lineárny regresný model na predikciu počtu obyvateľov, kde počet používateľov s pravidelnou nočnou lokalizáciou v obci je nezávislou premennou a počet obyvateľov zo sčítania obyvateľstva je závislou premennou. Použili sme údaje z posledného cenzu z roku 2021 a z dvoch dostupných kategórií sme zvolili kategóriu súčasný pobyt, ktorý by mal lepšie opisovať skutočné priestorové rozmiestnenie obyvateľstva. Všetky testované datasey priniesli štatisticky významné výsledky ($p < 0,001$) a potvrdili očakávania o schopnostiach odhadnúť priestorové rozmiestnenie populácie. Tesnosť závislosti medzi modelovanými a referenčnými údajmi zachytáva tabuľka č. 4.

Obrázok č. 1: Počet SIM kariet s pravidelnou nočnou lokalizáciou extrapolovaný na populáciu SR (X_i) v bunkách gridu 1 x 1 km



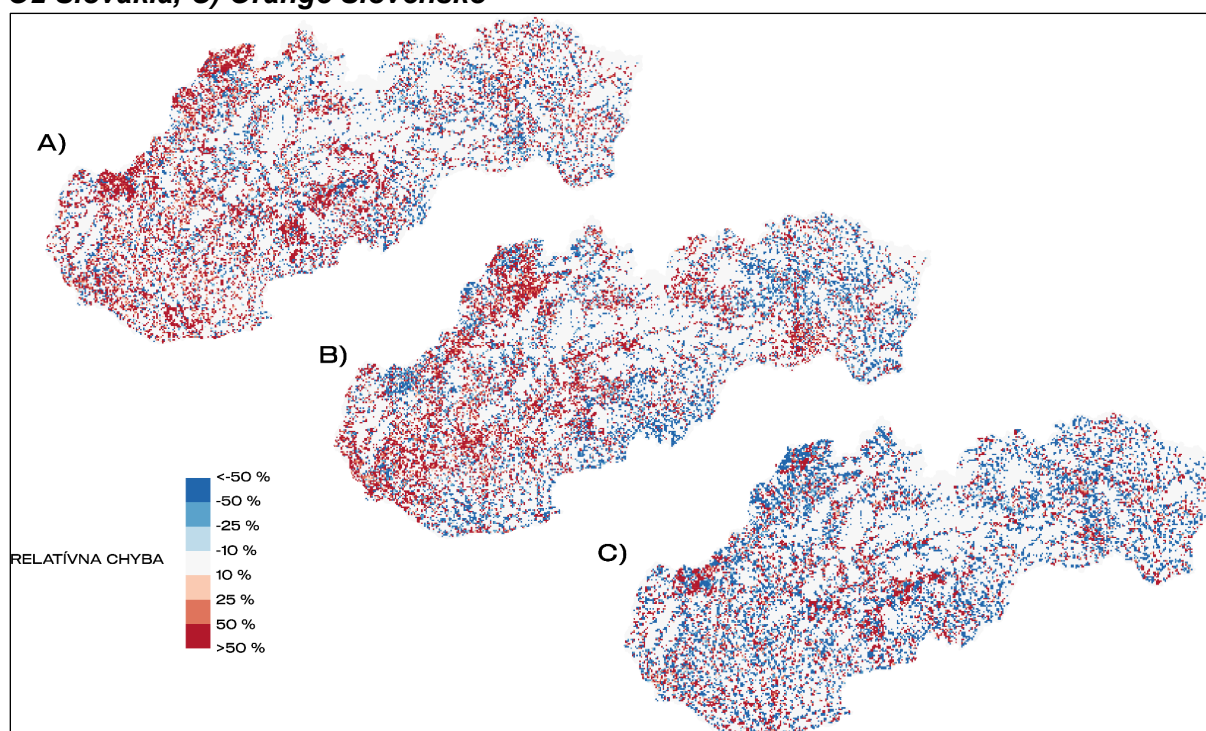
Zdroj údajov: Slovak Telekom (2023)

Tabuľka č. 4: Lineárna regresia medzi modelovanými údajmi (počet SIM kariet s pravidelnou nočnou lokalizáciou) a referenčnými údajmi (súčasný pobyt) v bunkách gridu 1 x 1 km

Dataset	r	s	τ	RMSE	Relatívna celková absolútna chyba [%]
Slovak Telekom	0,92	0,83	0,77	198	43,63
O2 Slovakia	0,95	0,88	0,82	163	36,95
Orange Slovensko	0,89	0,82	0,75	242	62,50

Zdroj údajov: Slovak Telekom (2023), O2 Slovakia (2023), Orange Slovensko (2023), SODB (2021)

Obrázok č. 2: Relatívna chyba medzi modelovaným počtom obyvateľov (X_i) a referenčným počtom obyvateľov (P_i) v bunkách gridu 1 x 1 km. A) Slovak Telekom, B) O2 Slovakia, C) Orange Slovensko



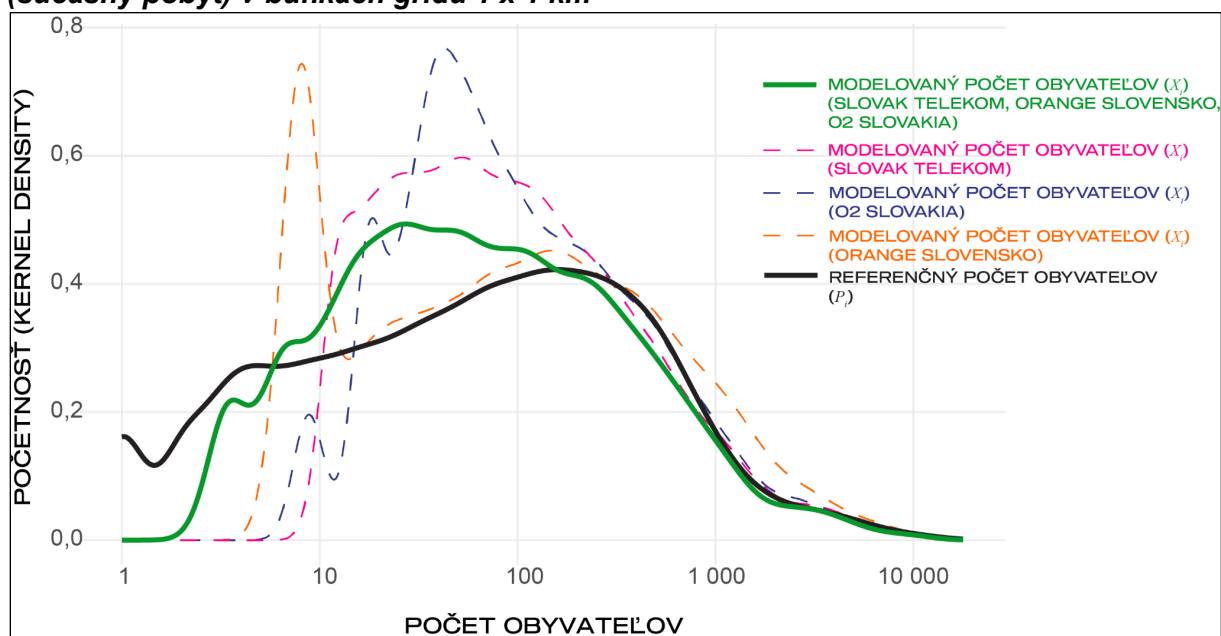
Zdroj údajov: Slovak Telekom (2023), O2 Slovakia (2023), Orange Slovensko (2023), SODB (2021)

Na detailnejšie vyhodnotenie odhadu populácie na úrovni populačného gridu je potrebné venovať pozornosť aj priestorovej lokalizácii buniek s významným nadhodnotením alebo podhodnotením modelovaného počtu rezidentov. Priestorový priemet relatívnej chyby (obrázok č. 2) zachytáva rozdielne oblasti podhodnotenia a nadhodnotenia rozličných operátorov. Táto skutočnosť je pravdepodobne dôsledkom priestorovo diferencovaných trhových podielov mobilných operátorov. Početné nadhodnotené bunky signalizujú nesprávne alokovaných používateľov mobilnej siete do buniek gridu bez evidovanej populácie. Detailný pohľad na nadhodnotené bunky ukazuje, že prevažná väčšina z nich sa nachádza mimo urbanizovaného územia. Priradenie používateľov s pravidelnou nočnou lokalitou do buniek populačného gridu, ktoré evidujú malý alebo žiadny počet obyvateľov, reflektuje limity použitého prístupu, ktorý prerozdeľuje rezidentov podľa objemu rodinných a bytových domov. Ak sa v bunke nachádza čo len jedna rezidenčná budova, model prerozdelí časť používateľov

mobilnej siete. V praxi môže ísť o osamelé budovy, napr. časti hospodárskych dvorov. V niektorých prípadoch však nadhodnotené bunky indikujú lokality, ktoré koncentrujú väčší počet obyvateľov než udávajú referenčné dáta. V praxi sú to napr. lokality novej rezidenčnej výstavby, chalupárske či rekreačné lokality.

Pri celorepublikovom pohľade prostredníctvom spojitého odhadu početnosti (obrázok č. 3) môžeme pozorovať vyšší počet málo saturovaných buniek gridu, ktoré zodpovedajú spomínanej vlastnosti modelu prerozdeľovať používateľov mobilnej siete do rezidenčných budov (bez ohľadu na to, či sú obývané). V prípade údajov zo Slovak Telekomu (2023) je počet nenulových buniek do 10 obyvateľov 5298, naproti tomu buniek s počtom rezidentov evidovaných na súčasný pobyt je len 4184. Výsledkom integrácie údajov od všetkých troch operátorov je rozdelenie početnosti bližšie k referenčným údajom. V grafe môžeme pozorovať podobný priebeh krivky najmä v populačne stredne veľkých a veľkých bunkách gridu.

Obrázok č. 3: Spojitý odhad početnosti (kernel density estimation) údajov z mobilnej siete (pravidelná nočná lokalita SIM kariet) a referenčných údajov z národného cenzu (súčasný pobyt) v bunkách gridu 1 x 1 km



Zdroj údajov: Slovak Telekom (2023), O2 Slovakia (2023), Orange Slovensko (2023), SODB (2021)

5.2. Vyhodnotenie validity údajov pravidelnej nočnej lokalizácie SIM kariet na úrovni obcí

Pri spracovaní a vyhodnotení odhadu rezidenčnej populácie na úrovni obcí sme postupovali analogicky ako v prípade gridu. Vzhľadom na nižší počet cieľových zón (2 927 obcí) než buniek gridu (50 661) sú výsledky globálnej štatistiky ešte spoľahlivejšie, než v prípade gridu (tabuľka č. 5).

Tabuľka č. 5: Lineárna regresia medzi modelovanými údajmi (počet SIM kariet s pravidelnou nočnou lokalizáciou) a referenčnými údajmi (súčasný pobyt) v obciach SR

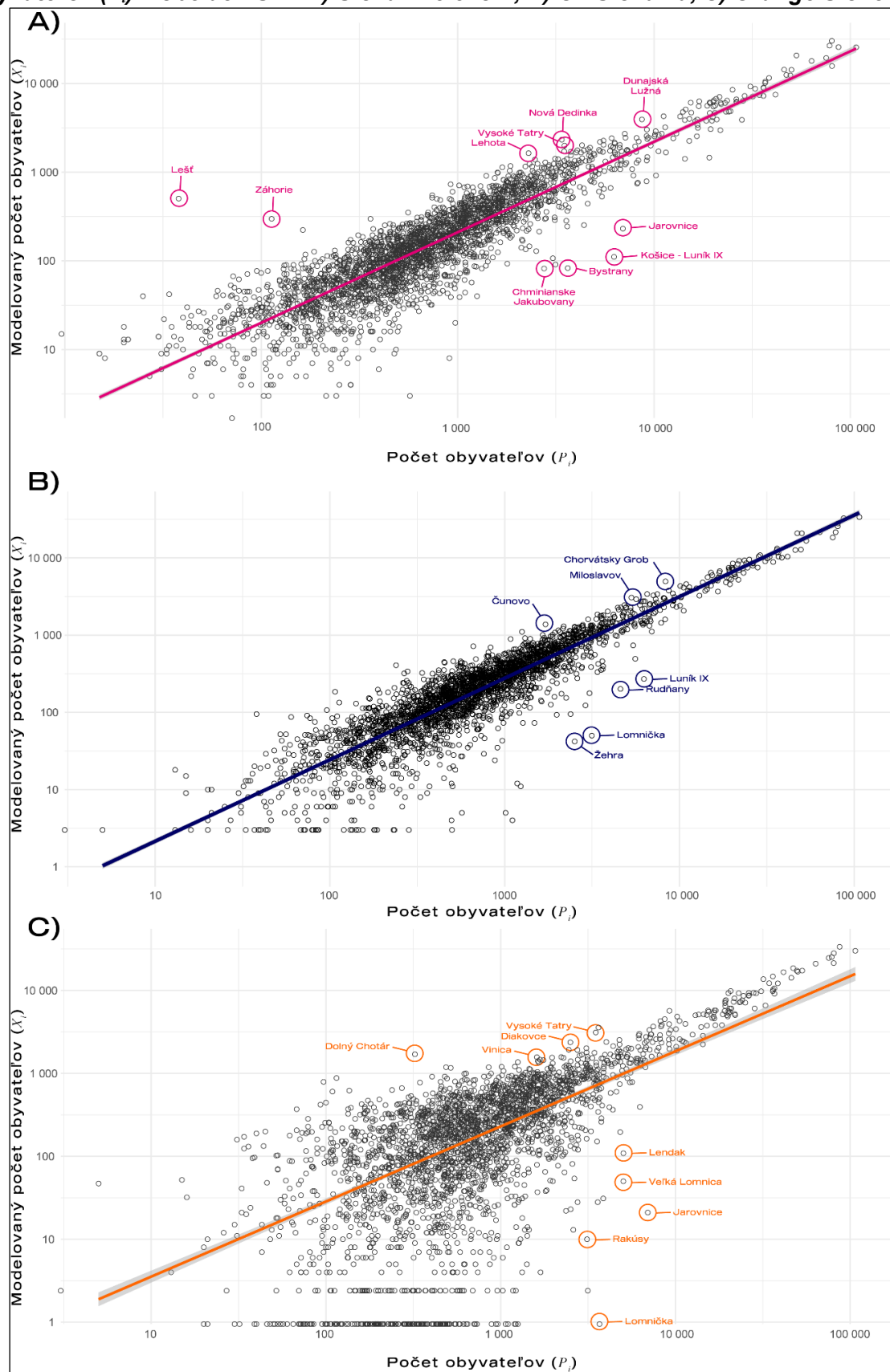
Dataset	r	s	τ	RMSE	Relatívna celková absolútna chyba [%]
Slovak Telekom	0,97	0,73	0,90	1576	28,74
O2 Slovakia	0,98	0,74	0,90	1036	22,49
Orange Slovensko	0,96	0,49	0,67	1288	37,86

Zdroj údajov: Slovak Telekom (2023), O2 Slovakia (2023), Orange Slovensko (2023), SODB (2021)

Hoci regresia priniesla veľmi uspokojivé výsledky, na hlbšie pochopenie použitého modelu odhadu populácie je potrebné preskúmať odľahlé hodnoty. Pri ich interpretácii budeme pozornosť venovať len stredne veľkým a veľkým obciam, keďže pri malých obciach je väčší rozptyl prirodzený a spôsobený nerovnomerným pokrytím územia mobilným signálom, ako aj rozkolísaným trhovým podielom mobilných operátorov v malých územných jednotkách. Ako zachytáva obrázok č. 4, v podhodnotených obciach môžeme pozorovať niekoľko spoločných znakov. V početných prípadoch ide o obce s marginalizovanými rómskymi komunitami (Lomnička, Stráne pod Tatrami, Jarovnice, Chminianske Jakubovany a pod.). Podhodnotenie modelovanej populácie je pravdepodobne spôsobené špecifickým využívaním služieb mobilnej siete (predplatené karty), nižšou penetráciou mobilných telefónov, ako aj štruktúrou populácie s početnou detskou zložkou (deti, ktoré mobilný telefón nemajú). Svoju úlohu môže zohrať aj skutočnosť, že segregované rómske osady sú často vytvorené z budov, ktoré nemusia byť evidované v mapových vrstvách, prípadne nemajú atribút rodinného domu. Pri transformácii údajov z vyžarovacích polygónov sa používatelia neprerozdelia do týchto území, keďže tieto budovy neboli súčasťou pomocných vrstiev. Závažnosť uvedených limitov nie je možné spoľahlivo vyhodnotiť bez podrobnej hĺbkovej analýzy.

V nadhodnotených obciach nie je možné identifikovať jednoznačného spoločného menovateľa. Ukazuje sa však, že nadhodnotené obce sa často nachádzajú v blízkosti Bratislavy (Dunajská Lužná, Nová Dedinka, Kvetoslavov), kde môžeme predpokladať vplyv väčšej skupiny neprihlásených rezidentov na trvalý (resp. súčasný) pobyt. Ďalšiu skupinu tvoria obce s dominantou rekreačnou funkciou (Vysoké Tatry, Demänovská dolina), kde je väčší počet pravidelne nocujúcich SIM kariet spôsobený nezanedbateľným počtom ľudí pracujúcich v turizme. Svoju úlohu môžu zohrať i rezidenti dlhodobo bývajúcich v rekreačných objektoch. Medzi nadhodnotenými obcami nachádzame aj viaceré nešpecifické obce, ktorých vyšší počet trvalo nocujúcich používateľov mobilnej siete môže byť výsledkom rozmanitých dôvodov, predovšetkým však nadpriemerného trhového podielu daného mobilného operátora.

Obrázok č. 4: Porovnanie modelovaného počtu obyvateľov (X_i) s referenčným počtom obyvateľov (P_i) v obciach SR. A) Slovak Telekom, B) O2 Slovakia, C) Orange Slovensko



Zdroj údajov: Slovak Telekom (2023), O2 Slovakia (2023), Orange Slovensko (2023), SODB (2021)

5.3. Vyhodnotenie validity integrovaných údajov pravidelnej nočnej lokalizácie na úrovni obcí

Výsledky regresie integrovaných údajov ponúka tabuľka č. 6. Dosiahnuté skóre relatívnej celkovej absolútnej chyby na úrovni 21 % môžeme považovať za veľmi dobrý výsledok, na ktorom sa významne podieľa výber vhodnej pomocnej vrstvy na interpoláciu údajov, ako aj skutočnosť, že údaje od jednotlivých operátorov sa vzájomne vhodne dopĺňajú, čím prispievajú k zvýšeniu presnosti a spoľahlivosti modelu nočnej populácie. Domnievame sa však, že existuje ešte nezanedbateľný priestor na zlepšenie presnosti modelu. Rezervy vidíme najmä v konštrukcii spoľahlivejšej pomocnej vrstvy, ako aj v odstránení „šumu“ vo vstupných údajoch. Podrobnejšie sa tejto téme venujeme v 6. kapitole.

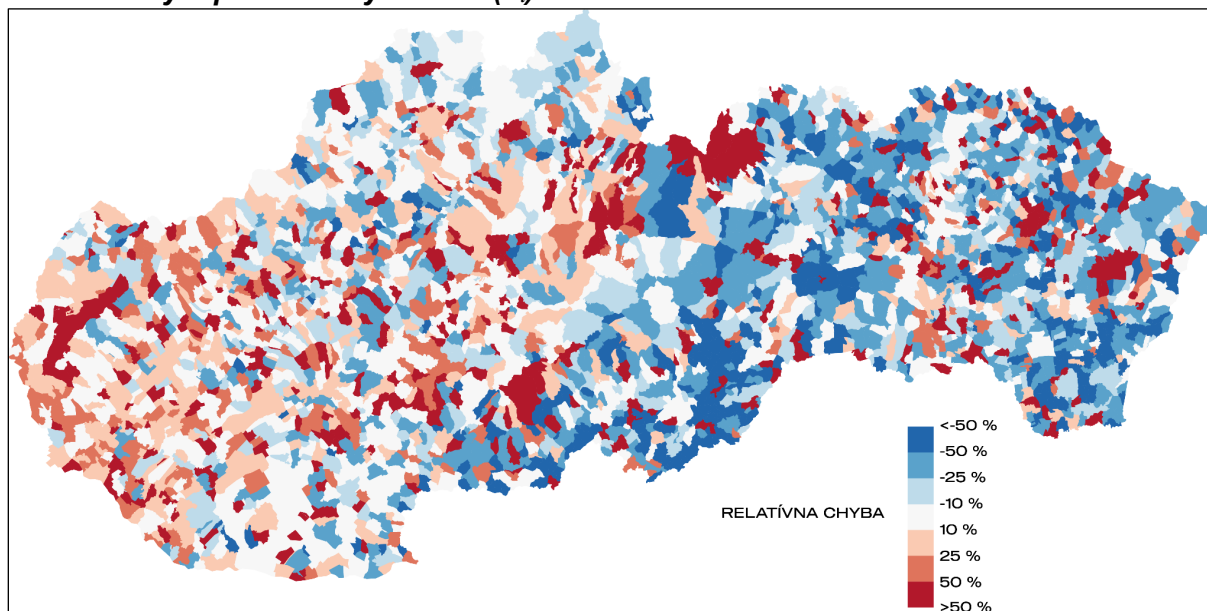
Tabuľka č. 6: Lineárna regresia medzi modelovanými údajmi (počet SIM kariet s pravidelnou nočnou lokalizáciou) a referenčnými údajmi (súčasný pobyt) v obciach SR

Dataset	r	s	τ	RMSE	Relatívna celková absolútna chyba [%]
Slovak Telekom + O2 Slovakia + Orange Slovensko	0,98	0,73	0,89	1078	21,23

Zdroj údajov: Slovak Telekom (2023), O2 Slovakia (2023), Orange Slovensko (2023), SODB (2021)

Výsledný pohľad na priestorovú diferenciáciu nadhodnotených a podhodnotených obcí z hľadiska modelu nočnej populácie (obrázok č. 5) prináša známy obraz o Slovensku, ktorý poznáme z analýzy rozmanitých sociálno-ekonomických ukazovateľov. Interpretácia však nie je jednoduchá, keďže do výsledku môže vstupovať množstvo rozličných faktorov, ktorých povahu a vplyv nie vždy vieme spoľahlivo vyhodnotiť. K interpretačnej zdržanlivosti nás nútia aj početné limity použitého modelu nočnej populácie. Napriek uvedenému sa pokúsime aspoň naznačiť niektoré súvislosti. Východiskom interpretácie môže byť analýza diferencí medzi súčasným a trvalým pobytom a z toho vyplývajúcich disproporcií medzi evidovaným a reálnym priestorovým rozmiestnením obyvateľstva (podrobnejšie sa téme venuje [30]). Vysokú mieru prisťahovaných na súčasný pobyt zaznamenali najmä obce v zázemí Bratislavy a Košíc, v celoslovenskom pohľade však môžeme jednoznačne identifikovať východo-západný gradient, ktorý pôsobí v prospech vyššej koncentrácie obyvateľov na západnom Slovensku. Výsledky modelu nočnej populácie by v tomto kontexte mohli naznačovať podobný trend. Pod vplyvom suburbanizácie a rozmanitých metropolizačných efektov sa môže dosahovať v Bratislave a jej zázemí ešte vyššia koncentrácia populácie, než zachytávajú oficiálne údaje. V prospech tejto argumentácie pôsobí aj veková štruktúra obyvateľstva, ktorá do suburbanizovaných obcí prináša najmä mladé rodiny [29]. Protiargumentom je však skutočnosť, že v tomto sociálne a ekonomicky silnom regióne môžeme očakávať vyššiu penetráciu mobilných zariadení než vo zvyšku Slovenska a teda aj početné duplicity spôsobené vlastníctvom dvoch či viacerých mobilných telefónov jednej osoby. Uvedené argumenty nie je možné v súčasnosti spoľahlivo vyhodnotiť a v ďalšom postupe pri spracovaní týchto unikátnych časopriestorových údajov by sme odporúčali podrobne analyzovať možné vplyvy vybraných sociálnych, ekonomických či sídelných faktorov na identifikované diferencie medzi evidovaným počtom obyvateľov a počtom používateľov mobilnej siete s pravidelnou nočnou lokalizáciou.

Obrázok č. 5: Relatívna chyba medzi modelovaným počtom obyvateľov (X_i) a referenčným počtom obyvateľov (P_i) v obciach SR



Zdroj údajov: Slovak Telekom (2023), O2 Slovakia (2023), Orange Slovensko (2023), SODB (2021)

6. ZHODNOTENIE PREDNOSTÍ A LIMITOV MODELU

Lokalizačné údaje mobilných telefónov prinášajú nový zdroj údajov o populácii, ktorý môže poskytovať významné informácie k už existujúcim zdrojom údajov, akými sú celoplošné cenzy a výberové zisťovania. Na rozdiel od týchto „tradičných“ zdrojov však prinášajú viaceré výhody, ktoré charakterizuje veľkosť vzorky, rýchlosť a frekvencia zberu údajov, ako aj finančné náklady potrebné na ich obstaranie.

6.1. Prednosti modelu

Metodika vyvinutá v tejto štúdii prináša niekoľko dôležitých inovácií, ktoré rozvíjajú oblasť využitia údajov z mobilnej siete aj v širšom (európskom) kontexte. Ide o nasledujúce charakteristiky:

- Model je založený na spracovaní signalizačných údajov z mobilnej siete. Ide o automaticky generované záznamy, ktoré produkuje mobilná sieť pri pravidelných kontrolách pripojených zariadení. Vzhľadom na veľký počet takýchto záznamov je ich spracovanie pomerne náročné a ich využitie v obdobných analýzach bolo dosiaľ iba sporadické. Napriek tomu ich význam v priestorových analýzach bude pravdepodobne narastať. Dôvodom sú zmeny vo využívaní mobilného telefónu. Kým v minulosti sme ho využívali najmä na volania a SMS správy, v súčasnosti toto využitie ustupuje v prospech dátových služieb (sociálne siete, odosielanie a prijímanie okamžitých správ a pod.) a pasívneho využívania telefónu (napr. rôzne notifikácie v aplikáciách si nevyžadujú aktívnu interakciu).
- Model pracuje s podrobnou topológiou mobilnej siete, kde využíva detailný model vyžarovacích polygónov. Dosiaľ bežne používaným prístupom bolo použitie polohy antény mobilnej siete, alebo využitie jednoduchšej teselácie (vyplnenie roviny pomocou jedného alebo viacerých geometrických útvarov bez prekryvania a medzier) založenej na Voronoiových polygónoch (napr. [10, 31]. Na transformáciu údajov z vyžarovacích polygónov buniek mobilnej siete

do cieľových priestorových rámcov (obce, grid) model využíva dazymetrickú interpoláciu mapujúcu alokovaných používateľov s pravidelnou nočnou/dennou lokalitou do objemu budov. Tým sa zásadne spresňuje odhad priestorového rozmiestnenia populácie, keďže koncentrácia obyvateľov je prirodzene viazaná na sídelnú zástavbu.

- Cieľovou priestorovou zónou modelu nie sú len katastrálne územia obcí, ale aj bunky gridu 1x1 km. Odhad populácie v tomto detaile nie je bežný a prináša mnohé úskalia. Napriek tomu model priniesol veľmi uspokojivé výsledky aj v tejto mierke. Konštrukcia vektorov medzi pravidelnou nočnou lokalizáciou a pravidelnou dennou lokalizáciou používateľa mobilnej siete prináša unikátny pohľad najmä na intraurbánnu mobilitu. Tento typ informácií bol doposiaľ možný len prostredníctvom nákladných prieskumov mobility a výberových zisťovaní.

Hoci lokalizačné údaje z mobilnej siete prinášajú bezprecedentný rozsah informácií o pobyte a pohybe veľkej časti populácie, je potrebné si uvedomiť, že ich charakter a kvalitu ovplyvňuje použitá technológia, kontext ich vzniku a spôsob formalizovania dátového modelu (entity, kategórie, atribúty, väzby). Práca s týmito údajmi si navyše vyžaduje opatrné prehodnocovanie postavené na kontextuálnych znalostiach s ohľadom na charakter analýzy a spôsob interpretácie.

6.2. Limity modelu

Pri interpretácii pasívnych lokalizačných údajov mobilných zariadení v sieťovej lokalizácii je dôležité si uvedomiť viaceré limity, ktoré vyplývajú z princípov prevádzky mobilnej siete a z právnych podmienok spracovania údajov jej používateľov. Vo všeobecnosti ide najmä o obmedzenia spojené s ochranou súkromia používateľov mobilnej siete, s limitovanými socio-demografickými štruktúrami, ako aj so špecifikami mobilnej komunikácie. V stručnosti sa pokúsime upozorniť na tieto kľúčové limity:

- Počet SIM kariet nemožno spoľahlivo stotožniť s počtom osôb (individuálnych používateľov). Dôvodom je skutočnosť, že nemôžeme predpokladať, že každý obyvateľ disponuje mobilným zariadením. Takisto nedokážeme spoľahlivo vylúčiť osoby využívajúce viaceré SIM karty.
- Presnosť lokalizácie závisí od architektúry mobilnej infraštruktúry, typu antén (2G/3G/4G/5G) a reliéfu. V mestskom prostredí sa pohybuje rádovo v stovkách metrov, vo voľnej krajine dosahuje rádovo kilometre. Presnosť lokalizácie ovplyvňuje aj spoľahlivosť modelu vyžarovacích polygónov. Ich tvar nemusí presne zodpovedať skutočnému pokrytiu mobilným signálom a v čase sa môže meniť (napr. vplyvom počasia). Prístupy, akým sa modeluje topológia mobilnej siete sú pri jednotlivých prevádzkovateľoch mobilnej siete rozdielne. Príkladom je model vyžarovacích polygónov Orange Slovensko, ktorý využíva polohu BTS, azimut (smer) a silu signálu. Výsledkom je vyžarovací polygón v tvare kruhovej výseče. Je zrejmé, že takto skonštruovaný polygón len veľmi voľne opisuje skutočný tvar územia, ktorý pokrýva anténa mobilnej siete.
- Mobilné zariadenia sa spravidla pripájajú na najbližšiu základňovú stanicu. V systéme priradenia konkrétnej bunky k telefónu však zohrávajú úlohu desiatky faktorov (napr. intenzita signálu, atmosférické podmienky, zaťaženie infraštruktúry, plán údržby a pod.). Je preto možné, že používateľ na rovnakom mieste, ktorý uskutoční napr. tri hovory, sa pripojí k trom rôznym anténam a teda k fyzickému presunu používateľa nedôjde, ale v lokalizačných záznamoch pohyb nastane. Táto skutočnosť môže ovplyvniť aj identifikáciu pravidelnej dennej/nočnej lokality SIM karty.

- Pasívne lokalizačné údaje mobilnej siete nedokážu poskytnúť informácie o type a trvaní konkrétnych aktivít. Príkladom je odhad dochádzkových tokov medzi pravidelnou dennou a nočnou lokalitou používateľa mobilnej siete. Je potrebné poznamenať, že nemusí ísť nevyhnutne o dochádzku do zamestnania či školy. V realite môže takáto väzba pokrývať veľmi rozdielne situácie. Pri použití údajov o využití zeme či type urbanizovaného prostredia sa môžeme pokúsiť niektoré aktivity zmysluplne odhadnúť.
- Jednotliví operátori mobilnej siete môžu byť zacielení na špecifickú skupinu používateľov či používateľiek (napr. na mladších používateľov, podnikateľov, seniorov a pod.). Extrapolácia údajov na celú populáciu je tak značne obmedzená. Pri spracovaní údajov od viacerých operátorov, ktorí spoločne pokrývajú väčšinu populácie, sa však tento problém do značnej miery eliminuje. Napriek tomu niektoré vekové kategórie ostávajú poddimenzované (najmä malé deti a seniori). Ako sme naznačili v predchádzajúcich častiach príspevku, reprezentatívnosť údajov mobilnej siete je problematická najmä v obciach s marginalizovanými rómskymi komunitami. V týchto skupinách obyvateľstva môžeme predpokladať nielen nižšiu penetráciu mobilných zariadení (štruktúra populácie s veľkou detskou zložkou), ale aj špecifické využívanie mobilnej siete (predplatené karty minoritných prevádzkovateľov mobilnej komunikácie).
- Na zabezpečenie anonymity a bezpečnosti údajov je potrebná eliminácia záznamov pod určitú hranicu. V predkladanej analýze to bola hranica na úrovni 3 SIM kariet v danej územnej jednotke alebo toku. Hoci interpretačný dôraz sa kladie najmä na veľké koncentrácie používateľov alebo veľké toky SIM kariet v rámci infraštruktúry mobilnej siete, vynechaním málo početných lokalizácií sa môžu závažne skresliť výsledky, najmä pokiaľ je zámerom extrapolovať počty používateľov mobilnej siete na celú populáciu. Takýto prístup nežiaduco skresľuje aj vektorové dáta, ktoré majú prirodzene výrazné zošikmenie distribúcie (prevažujú málopočetné toky).
- Algoritmus na identifikáciu pravidelnej nočnej a dennej lokality nezohľadňuje rozdielnu hustotu záznamov jednotlivých mobilných operátorov. Tento limit bol identifikovaný až počas implementačnej fázy a vzhľadom na vopred stanovenú metodiku (postavenú na jednotnom algoritme) nebolo možné vykonať dodatočné zmeny. Ako perspektívna možnosť sa javí individuálne „vyladenie“ algoritmu vzhľadom na špecifiká jednotlivých mobilných operátorov.
- Použitá metodika spracovania lokalizačných údajov z mobilnej siete je založená na spracovaní surových záznamov (signalingových, CDR). Vzhľadom na dlhšie obdobie pozorovania však pri niektorých mobilných operátoroch došlo k čiastočnej agregácii starších údajov. Výsledkom tejto semiagregácie je „zriedenie“ záznamov individuálnych používateľov, čo môže mať vplyv na spoľahlivosť identifikovania pravidelnej nočnej a dennej lokality SIM karty (a z toho odvodených vektorov dochádzky). Vzhľadom na to, že nám neboli známe detaily uvedenej semiagregácie, nebolo možné dodatočne prispôsobiť algoritmy.
- Pri transformácii údajov z vyžarovacích polygónov mobilnej siete do cieľových priestorových jednotiek (obec/grid) sme použili dazymetrickú transformáciu s využitím objemu budov. Hoci táto metóda priniesla veľmi uspokojivé výsledky, má nepochybne aj limity. Tým prvým je skutočnosť, že použitá vrstva budov pochádza z databázy ZBGIS (2017), takže neodráža aktuálny stav urbanizovaného územia. V praxi ide najmä o lokality s dynamickou výstavbou okolí veľkých miest. V databáze sú aj početné chyby, resp. neúplné

atribúty o stave a type budov. Druhým limitom je použitie len niektorých typov budov na konštrukciu pomocnej vrstvy. Na nočnú lokalizáciu to boli rodinné domy, bytové domy a polyfunkčné domy. V praxi však množstvo ľudí trvalo obýva aj iné typy budov, napr. chaty. Na túto skutočnosť je vhodné prihliadať najmä v oblastiach s rozptýleným osídlením, v obciach s marginalizovanými rómskymi komunitami a v záhradkárskych či rekreačných lokalitách.

Z uvedených dôvodov je zrejmé, že môžu existovať značné rozdiely medzi kvantitatívnymi charakteristikami populácie vypočítanými na základe oficiálnych dát (napr. počet dochádzajúcich do obce) a údajmi z mobilnej siete. Vo všeobecnosti je potrebné sa vyhnúť práci na úrovni absolútnych hodnôt a preferovať skôr prístup využívajúci relativizované údaje (napr. podiel dochádzajúcich do obce namiesto počtu dochádzajúcich do obce).

Napriek uvedeným limitom priniesli výsledky spracovania lokalizačných údajov mobilnej siete veľmi solídne výsledky, ktoré môžu prispieť k hlbšiemu poznaniu priestorového rozmiestnenia a dynamiky obyvateľstva. Pri rešpektovaní uvedených limitov dostávame perspektívne dáta, vhodné na rozmanité analýzy na úrovni regiónov (obce) alebo lokalít (grid).

Hoci navrhnutá metodika priniesla uspokojivé výsledky, z nášho pohľadu ide len o východiskový bod, na základe ktorého chceme stimulovať ďalšiu odbornú diskusiu medzi poskytovateľmi služieb mobilnej komunikácie a verejnými inštitúciami, s cieľom zosúladiť predstavy a možnosti implementácie týchto unikátnych údajov. Predpokladáme, že navrhnutá metodika budú ďalej zdokonalená a prinesie ešte spoľahlivejšie výsledky.

7. ZÁVER

Lokalizačné údaje mobilnej siete predstavujú perspektívny zdroj údajov o priestorovom rozmiestnení a mobilite obyvateľov a sľubný výskumný smer s bohatým využitím, osobitne v riadiacej a plánovacej praxi. Stále sme však ďaleko od toho, aby sa tieto údaje stali bežnou súčasťou priestorových analýz. Je však zrejmé, že rozvoj v tejto oblasti výskumu dospel dostatočne ďaleko, aby prekročil prah experimentálneho používania a stal sa súčasťou aplikovanej praxe. Ako dokumentuje narastajúci počet štúdií, údaje z mobilnej siete poskytujú porovnateľnú presnosť a spoľahlivosť ako štatistické údaje o obyvateľstve [1, 37, 42]. Konceptuálny a metodický aparát, ktorý vyprodukovala geografia a iné vedné disciplíny v oblasti spracovania údajov z mobilnej siete, je skutočne rozsiahly a prináša praktické riešenia a významné výsledky. Najväčšou prekážkou pre rozvoj tejto perspektívnej oblasti výskumu je v súčasnosti dostupnosť údajov, najmä možnosť analýzy na úrovni buniek mobilnej siete a so zohľadnením doplnkových údajov potrebných na korektnú interpoláciu (tvar vyžarovacích polygónov) a extrapoláciu (poznatie regionálne diferencovaného trhového podielu) údajov. Tento typ údajov však nemôžu poskytnúť spoločnosti spracúvajúce údaje mobilnej siete pre koncových zákazníkov. Princiipiálne ide o citlivé údaje, a tak je zrejmé, že ďalší rozvoj analýz založených na údajoch mobilnej siete sa nezaobíde bez úzkej spolupráce s mobilnými operátormi. Kľúčovou úlohou bude nájsť taký model spolupráce, ktorý minimalizuje riziká spojené so spracovaním mikroúdajov a zároveň prináša potenciálne inovácie do spracovania údajov mobilnej siete. Je teda dôležité, aby akademický či verejný sektor prichádzal s takými návrhmi spolupráce, ktoré neohrozia kredibilitu mobilných operátorov,

a súčasne prinesú inovatívne metódy spracovania údajov, ktoré v konečnom dôsledku môžu byť prínosné aj pre mobilných operátorov. Prekonanie bariér bude nepochybne výhodné pre obe strany. Ak sa nájde model dlhodobej spolupráce, môžu verejné či výskumné inštitúcie získať spoľahlivé údaje na pravidelnej báze (mesiac/rok). Mobilní operátori zas môžu využiť príležitosť ako monetizovať „vedľajší produkt“ fungovania mobilnej siete.

Lokalizačné údaje z mobilnej siete budú nepochybne dôležitým zdrojom informácií na prijímanie kvalifikovaných rozhodnutí v početných krízach, ktorým čelí a bude čeliť naša spoločnosť. Bolo by chybou tento cenný zdroj priestorových údajov nevyužiť.

LITERATÚRA

- 1] AASA, A. – KAMENJUK, P. – SALUVEER, E. – ŠIMBERA, J. – RAUN, J.: Spatial interpolation of mobile positioning data for population statistics. In: *Journal of Location Based Services*, 2021, č. 4, p. 239 – 260.
- [2] AHAS, R. – AASA, A. – MARK, Ü. – PAE, T. – KULL, A.: Seasonal tourism spaces in Estonia: Case study with mobile positioning data. In: *Tourism Management*, 2007a, č. 3, s. 898 – 910.
- [3] AHAS, R. – AASA, A. – SILM, S. – AUNAP, R. – KALLE, H. – MARK, Ü.: Mobile positioning in space–time behaviour studies: social positioning method experiments in Estonia. In: *Cartography and Geographic Information Science*, 2007b, č. 4, s. 259 – 273.
- [4] AHAS, R. – MARK, Ü: Location based services – new challenges for planning and public administration? In: *Futures*, 2005, č. 6, s. 547 – 561.
- [5] AHAS, R. – SILM, S. – JÄRV, O. – SALUVEER, E.– TIRU, M.: Using mobile positioning data to model locations meaningful to users of mobile phones. In: *Journal of Urban Technology*, 2010, č. 1, s. 3 – 27.
- [6] BATISTA E SILVA, F. – CRAGLIA, M. – FREIRE, S. – ROSINA, K. – LAVALLE, C. – MARIN, M. – SCHIAVINA, M.: Enhancing activity and population mapping. European Commission: Joint Research Centre, 2016.
- [7] BATISTA E SILVA, F. – FREIRE, S. – SCHIAVINA, M. – ROSINA, K. – MARÍN-HERRERA, M. A. – ZIEMBA, L. – CRAGLIA, M. – KOOMEN, E. – LAVALLE, C.: Uncovering temporal changes in Europe’s population density patterns using a data fusion approach. In: *Nature Communications*, 2020, č. 1, s. 1 –11.
- [8] BATTY, M.: *The new science of cities*. MIT press, 2013
- [9] BENGTTSSON, L. – LU, X. – THORSON, A. – GARFIELD, R. – VON SCHREEB, J.: Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in Haiti. In: *PLoS Medicine*, 2011, č. 8, e1001083.
- [10] BERGROTH, C. – JÄRV, O. – TENKANEN, H. – MANNINEN, M. – TOIVONEN, T.: A 24-hour population distribution dataset based on mobile phone data from Helsinki Metropolitan Area, Finland. In: *Scientific data*, 2022, č. 1, s. 39.
- [11] BEZÁK, A.: Vnútné migrácie na Slovensku: súčasné trendy a priestorové vzorce. In: *Geografický časopis*, 2006, č. 1, s. 15 – 44.
- [12] BILJECKI, F. – ARROYO OHORI, K. – LEDOUX, H. – PETERS, R. – STOTER, J.: Population estimation using a 3D city model: A multi-scale country-wide study in the Netherlands. In: *PLoS one*, 2016, č. 6, e0156808.
- [13] BLEHA, B.– POPJAKOVÁ, D.: Migrácia ako dôležitý determinant budúceho vývoja na lokálnej úrovni–príklad Petržalky. In: *Geografický časopis*, 2007, č. 3, s. 265 – 291.

- [14] CASTELLS, M. – FERNANDEZ-ARDEVOL, M. – QIU, J. L. – SEY, A.: Mobile communication and society: A global perspective. MIT Press, 2009.
- [15] CSÁJI, B. C. – BROWET, A. – TRAAG, V. A. – DELVENNE, J. C. – HUENS, E. – VAN DOOREN, P. – SMOREDA, Z. – BLONDEL, V. D.: Exploring the mobility of mobile phone users. In: *Physica A: Statistical Mechanics and its Applications*, 2013, č. 6, s. 1459 – 1473.
- [16] DEVILLE, P. – LINARD, C. – MARTIN, S. – GILBERT, M. – STEVENS, F. R. – GAUGHAN, A. E. – BLONDEL, V. D. – TATEM, A. J.: Dynamic population mapping using mobile phone data. In: *Proceedings of the National Academy of Sciences*, 2014, č. 45, s. 15888 – 15893.
- [17] DOUGLASS, R. W. – MEYER, D. A. – RAM, M. – RIDEOUT, D. – SONG, D.: High resolution population estimates from telecommunications data. In: *EPJ Data Science*, 2015, č. 4, s. 1 – 13.
- [18] GOODCHILD, M. F., ANSELIN, L., DEICHMANN, U.: A framework for the areal interpolation of socioeconomic data. In: *Environment and planning A: Economy and Space*, 1993, č. 3, s. 383 – 397.
- [19] GRANTZ, K. H. – MEREDITH, H. R. – CUMMINGS, D. A. – METCALF, C. J. E. – GRENFELL, B. T. – GILES, J. R. – MEHTA, S. – SOLOMON, S. – LABRIQUE, A. – KISHORE, N. – BUCKEE, C. – WESOLOWSKI, A.: The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. In: *Nature Communications*, 2020, č. 1, s. 1 – 8.
- [20] HORANONT, T.: A Study on Urban Mobility and Dynamic Population Estimation by Using Aggregate Mobile Phone Sources. CSIS Discussion Paper No. 115, 2012.
- [21] CHEN, M. – CLARAMUNT, C. – ÇÖLTEKIN, A. – LIU, X. – PENG, P. – ROBINSON, A. C. – LÜ, G.: Artificial intelligence and visual analytics in geographical space and cyberspace: Research opportunities and challenges. *Earth-Science Reviews*, 2023, 104438.
- [22] JÄRV, O. – AHAS, R. – WITLOX, F.: Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. In: *Transportation Research Part C: Emerging Technologies*, 2014, s. 122 – 135.
- [23] JÄRV, O., H. – TENKANEN, T. TOIVONEN.: "Enhancing Spatial Accuracy of Mobile Phone Data Using Multi-temporal Dasymetric Interpolation." In: *International Journal of Geographical Information Science*, 2017, č. 8, s. 1630 – 1651. Taylor & Francis
- [24] KRINGS, G. – CALABRESE, F. – RATTI, C. – BLONDEL, V. D.: Urban gravity: a model for inter-city telecommunication flows. In: *Journal of Statistical Mechanics: Theory and Experiment*, 2009, L07003.
- [25] LANGFORD, M. – MAGUIRE, D. J. – UNWIN, D. J.: The areal interpolation problem: Estimating population using remote sensing in a GIS framework. In: Masser, I., Blakemore M. B. (Eds.). *Handling geographic information*. Essex (Longman Scientific & Technical), 1991, s. 55 – 77.
- [26] LONG, A. – CARNEY, F. – KANDT, J.: Who is returning to public transport for non-work trips after COVID-19? Evidence from older citizens' smart cards in the UK's second largest city region. In: *Journal of Transport Geography*, 2023, 103529.
- [27] MARTIN, D. – COCKINGS, S. – LEUNG, S.: Developing a flexible framework for spatiotemporal population modeling. In: *Annals of the Association of American Geographers*, 2015, č. 4, s. 754 – 772.
- [28] MOUNTAIN, D. – RAPER, J.: Modelling human spatio-temporal behaviour: a challenge for location-based services, In: *Proceedings of the Sixth International*

Conference on GeoComputation, University of Queensland, Brisbane, Australia, 24 – 26 September, 2001.

[29] NOVOTNÝ, L. – PREGI, L.: Selective migration of population subgroups by educational attainment in the urban region of Bratislava. In: *Geografický časopis*, 2017, č. 1, s. 21 – 39.

[30] ÓVÁRI, K. – KOČIŠ, M.: Priestorová diferenciácia rozmiestnenia obyvateľstva v kontexte súčasného pobytu v Slovenskej republike k 1. 1. 2021. In: *Slovenská štatistika a demografia*, 2023, č. 2, s. 5 – 29.

[31] PENG, Z. – WANG, R. – LIU, L. – WU, H.: Fine-Scale Dasymetric Population Mapping with Mobile Phone and Building Use Data Based on Grid Voronoi Method. In: *ISPRS International Journal of Geo-Information*, 2020, roč. 9, č. 6, 344.

[32] PODOLÁK, P.: Centre and Hinterland-Migration Relations. In: *Folia Geographica*, 2002, 5, s. 143 – 145.

[33] RATTI, C. – FRENCHMAN, D. – PULSELLI, R. M. – WILLIAMS, S.: Mobile landscapes: using location data from cell phones for urban analysis. In: *Environment and Planning B: Planning and Design*, 2006, č. 5, s. 727 – 748.

[34] READES, J. – CALABRESE, F. – RATTI, C.: Eigenplaces: analysing cities using the space–time structure of the mobile phone network. In: *Environment and Planning B: Planning and Design*, 2009, č. 5, s. 824 – 836.

[35] READES, J. – CALABRESE, F. – SEVTSUK, A. – RATTI, C.: Cellular census: Explorations in urban data collection. In: *IEEE Pervasive computing*, 2007, č. 3, s. 30 – 38.

[36] RICCIATO, F. – WIDHALM, P. – CRAGLIA, M. – PANTISANO, F.: Estimating population density distribution from network-based mobile phone data. Luxembourg (Publications Office of the European Union), 2015.

[37] RICCIATO, F. – WIDHALM, P. – PANTISANO, F. – CRAGLIA, M.: Beyond the “single-operator, CDR-only” paradigm: An interoperable framework for mobile phone network data analyses and population density estimation. In: *Pervasive and Mobile Computing*, 2017, s. 65 – 82.

[38] SHELLER, M.: Mobile publics: beyond the network perspective. In: *Environment and Planning D: Society and Space*, 2004, č. 1, s. 39 – 52.

[39] SHOVAL, N.: Sensing human society. In: *Environment and Planning B: Planning and Design*, 2007, č. 2, s. 191 – 195.

[40] STEENBRUGGEN, J. – BORZACCHIELLO, M. T. – NIJKAMP, P. – SCHOLTEN, H.: Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: a review of applications and opportunities. In: *Geo Journal*, 2013, č. 2, s. 223 – 243.

[41] ŠVEDA, M. – PODOLÁK, P.: Fenomén neúplnej evidencie migrácie v suburbánnej zóne (na príklade zázemia Bratislavy). In: *Geografický časopis*, 2014, č. 2, s. 115 – 132.

[42] ŠVEDA, M. – SLÁDEKOVÁ MADAJOVÁ, M.: Estimating distance decay of intra-urban trips using mobile phone data: The case of Bratislava, Slovakia. In: *Journal of Transport Geography*, 2023, 103552.

[43] ŠVEDA, M. – SLÁDEKOVÁ MADAJOVÁ, M. – BARLÍK, P. – KRIŽAN, F. – ŠUŠKA, P.: Mobile phone data in studying urban rhythms: towards an analytical framework. In: *Moravian Geographical Reports*, 2020, č. 4, s. 248 – 258.

[44] ŠVEDA, M. – SLÁDEKOVÁ MADAJOVÁ, M. – ROSINA, K. – HURBÁNEK, P. – FORSTL, F. – ZÁBOJ, P. – VÝBOŠŤOK, J.: When spatial interpolation matters: seeking an appropriate data transformation from the mobile network for population estimates, In: *Computers, Environment and Urban Systems*, 2023, rukopis.

- [45] TOBLER, W. – DEICHMANN, U. – GOTTSEGEN, J. – MALOY, K.: World population in a grid of spherical quadrilaterals. In: *International Journal of Population Geography*, 1997, č. 3, s. 203 – 225.
- [46] TOWNSEND, A.: Life in the real-time city: mobile telephones and urban metabolism. In: *Journal of Urban Technology*, 2000, s. 85 – 104.
- [47] URAL, S. – HUSSAIN, E. – SHAN, J.: Building population mapping with aerial imagery and GISdata. In: *International Journal of Applied Earth Observation and Geoinformation*, 2011, č. 6, s. 841 – 852.
- [48] WANG, Z. – HE, S. Y. – LEUNG, Y.: Applying mobile phone data to travel behaviour research: A literature review. In: *Travel Behaviour and Society*, 2018, s. 141 – 155.
- [49] WESOLOWSKI, A. – EAGLE, N. – TATEM, A. J. – SMITH, D. L. – NOOR, A. M. – SNOW, R. W. – BUCKEE, C. O.: Quantifying the impact of human mobility on malaria. In: *Science*, 2012, č. 6112, s. 267 – 270.
- [50] YANG, Y. – XIONG, C. – ZHUO, J. – CAI, M.: Detecting home and work locations from mobile phone cellular signaling data. In: *Mobile Information Systems*, 2021, s. 1 – 13.
- [51] YOO, B. – KANG, S. – CHON, K. – KIM, S.: Origin–destination estimation using cellular phone BS information. In: *Journal of the Eastern Asia Society for Transportation Studies*, 2005, s. 2574 – 2588.

RESUMÉ

Súčasný trendy, ktoré stimuluje postmoderná spoločnosť a informačno-komunikačné technológie, kladú nové požiadavky na získavanie aktuálnych a detailných údajov o časovo-priestorovom správaní populácie. Príspevok sa zaoberá možnosťami využitia lokalizačných údajov z mobilnej siete na poznanie priestorovej distribúcie a mobility obyvateľov. Projekt Socioekonomické aspekty Big data Štatistického úradu SR reflektuje túto potrebu a sústreďuje sa na využitie údajov z mobilnej siete v experimentálnej populačnej štatistike. Kľúčovou otázkou je, či údaje z mobilnej siete môžu poskytnúť relevantnú informáciu o priestorovom rozmiestnení a mobilite populácie na Slovensku. Okrem toho sa skúma aký je prínos a limity údajov z mobilných sietí v porovnaní s bežnými štatistickými údajmi o obyvateľstve.

Príspevok prezentuje originálnu metodiku, ktorá bola použitá na extrahovanie pravidelných denných a nočných lokalít používateľov mobilnej siete (SIM kariet). Výsledky naznačujú vysokú koreláciu medzi údajmi získanými z mobilnej siete a referenčnými údajmi o počte obyvateľov z národného cenzu. Prostredníctvom údajov z mobilnej siete tak dostávame relevantnú informáciu o priestorovom rozmiestnení a mobilite populácie, hoci pri interpretácii výsledkov je dôležité si uvedomiť viaceré limity, ktoré vyplývajú z princípov prevádzky mobilnej siete a z právnych podmienok spracovania údajov jej používateľov. Vo všeobecnosti ide najmä o špecifika vyplývajúce z princípov mobilnej komunikácie, priestorovo diferencovaných trhových podielov mobilných operátorov a nedostatočného pokrytia niektorých sociálno-demografických štruktúr populácie. V praxi ide napríklad o marginalizované rómske komunity. Napriek početným limitom priniesli výsledky spracovania lokalizačných údajov mobilnej siete pomerne spoľahlivé výsledky, ktoré môžu prispieť k hlbšiemu poznaniu priestorového rozmiestnenia a dynamiky obyvateľstva. Pri rešpektovaní limitov tak dostávame perspektívne dáta, vhodné na rozmanité analýzy na úrovni regiónov alebo obcí.

RESUME

Current trends stimulated by postmodern society and information and communication technologies impose new requirements for obtaining up-to-date and detailed data on the spatiotemporal behavior of the population. The contribution focuses on the possibilities of utilizing location data from the mobile network to understand the spatial distribution and mobility of the population. The project Socioeconomic Aspects of Big Data by the Statistical Office of the Slovak Republic reflects this need and concentrates on the use of mobile network data in experimental population statistics. The key question is whether mobile network data can provide meaningful information about the spatial distribution and mobility of the population in Slovakia. Additionally, we aim to explore the benefits and limits of mobile network data compared to conventional population statistics.

The contribution presents an original methodology used to extract regular day and night locations of mobile phone users (SIM cards). The results suggest a high correlation between the data obtained from the mobile network and reference data on population from the national census. Through mobile network data, we thus obtain meaningful information about the spatial distribution and mobility of the population. However, when interpreting the results, it is important to be aware of several limitations that resulted from the principles of mobile network operation and the legal conditions for processing user data. In general, these limitations stem from the specifics of mobile communication principles, spatially differentiated market shares of operators, and underrepresentation of certain socio-demographic population structures. In practice, this includes, for example, the marginalized Roma communities. Despite numerous limitations, the results of processing mobile network location data have yielded relatively powerful results that can contribute to a deeper understanding of the spatial distribution and dynamics of the population. By respecting these limitations, we obtain prospective data suitable for diverse analyses at the regional or municipal level.

PROFESIJNÝ ŽIVOTOPIS

Mgr. Martin Šveda, PhD., absolvoval magisterské štúdium v odbore geografia a kartografia (2007) a doktorandské štúdium v odbore regionálna geografia na Prírodovedeckej fakulte Univerzity Komenského v Bratislave (2011). Od roku 2017 pôsobí ako odborný asistent na Katedre regionálnej geografie a rozvoja regiónov Prírodovedeckej fakulty Univerzity Komenského v Bratislave. Súčasne je samostatným vedeckým pracovníkom na Geografickom ústave SAV. Vo svojej výskumnej činnosti sa zameriava predovšetkým na procesy suburbanizácie a ich dopady na transformáciu prímestských sídiel. Venuje sa aj sledovaniu časovo-priestorových vzorov správania obyvateľov prostredníctvom lokalizačných údajov mobilnej siete.

Mgr. Michala Sládeková Madajová, PhD., absolvovala magisterské štúdium v odbore geografia a kartografia (2006) a doktorandské štúdium v odbore regionálna geografia na Prírodovedeckej fakulte Univerzity Komenského v Bratislave (2010). V súčasnosti pôsobí ako samostatná vedecká pracovníčka v oddelení humánnej a regionálnej geografie Geografického ústavu SAV a od roku 2019 vyučuje na Katedre regionálnej geografie a rozvoja regiónov Prírodovedeckej fakulty Univerzity Komenského v Bratislave. Venuje sa rôznym témam z oblasti humánnej a regionálnej geografie, ale predovšetkým problematike harmonizácie geografických dát a možnostiam využitia štatistických metód v geografii.

Mgr. Pavol Hurbánek, PhD., absolvoval magisterské štúdium v odbore geografia a kartografia (2000) a doktorandské štúdium v odbore humánna geografia na Prírodovedeckej fakulte Univerzity Komenského v Bratislave (2007). Pôsobil ako odborný asistent na Katedre humánnej geografie a demogeografie na Prírodovedeckej fakulte Univerzity Komenského v Bratislave (2005 až 2008) a na Katedre geografie na Pedagogickej fakulte Katolíckej univerzity v Ružomberku (2011 až 2022). Od roku 2016 je vedeckým pracovníkom

Geografického ústavu SAV v Bratislave. Venuje sa výskumu tematickej presnosti mapovania zastavaného územia na odhad počtu obyvateľov, vymedzeniu mestských a vidieckych oblastí s využitím sídelných sietí či regionálnej taxonómii.

Mgr. Konštantín Rosina, PhD., absolvoval magisterské štúdium na Prírodovedeckej fakulte Univerzity Komenského v Bratislave v odbore geografia (2010). V roku 2015 ukončil doktorandské štúdium v oddelení geoinformatiky Geografického ústavu SAV. V rokoch 2016 – 2018 pôsobil ako výskumný pracovník Spoločného výskumného centra Európskej Komisie v Ispre (IT). Od roku 2019 pracuje ako špecialista na diaľkový prieskum Zeme v spoločnosti Solargis s. r. o. Od roku 2020 je aj výskumným pracovníkom Geografického ústavu SAV.

KONTAKT

martin.sveda@uniba.sk

geogmada@savba.sk

pavolhurbanek@gmail.com

konstantin.rosina@savba.sk