

AGREGACE DATOVÝCH ZDROJŮ PRO TVORBU REPREZENTATIVNÍ VRSTVY DISTRIBUCE OBYVATELSTVA OLOMOUCKÉHO KRAJE

Oldřich Rypl, Jaroslav Burian

Informace o rozložení obyvatelstva v prostoru jsou nezbytné pro popis naší společnosti i pro studium vzájemných vztahů mezi prvky humánně-geografické sféry – nejinak tomu je v oblasti územního plánování a při správě měst a regionů. Tradiční zdroje dat (obvykle národní registry) však poskytují pouze statický pohled vázaný na bydliště (respektive pracoviště) obyvatel, který je nepřesný a může být značně zavádějící. Při analýze jevů a jejich příčin, které jsou často na administrativních hranicích nezávislé, je žádoucí zjevnou nepřesnost eliminovat a poskytnout výzkumníkům, úředníkům či místním samosprávám přesná a dostatečně podrobná data, která se stanou podkladem pro jejich rozhodování a zajistí kvalitnější výstupy. Příspěvek popisuje postup agregace několika datových zdrojů s cílem poskytnout řešení pro tvorbu přesnějších datových sad o distribuci obyvatel v prostoru v čase. Postup je popsán jak na úrovni metodických kroků, tak na úrovni technické implementace v prostředí GIS. Výsledky jsou představeny na oblasti Olomouckého kraje.

Klíčová slova: agregace, obyvatelstvo, distribuce, reprezentativní vrstva, GIS

Úvod

Primární statistická data jsou často získávána ve formě bodových údajů, které jsou později rozpočítávány do územních jednotek – ať už z důvodu ochrany osobních údajů nebo proto, aby je bylo možné snáze interpretovat a vizualizovat. Tento jednoduchý způsob přiřazování statistické informace příslušným cílovým (územním) jednotkám se nazývá agregace, případně též „bottom-up“ metoda. Statistické agregáty původních dat jsou založeny na minimech, maximech, četnostech a průměrech, které jsou následně v mapách vizualizovány formou kartogramu nebo metodou plošných značek. Pro účely agregace se nejčastěji používají sčítací obvody, dojížďkové zóny nebo administrativní územní jednotky.

Wilson et al. [2021] považují agregaci, tzv. „big data“, do administrativních jednotek za jeden z možných přístupů k prostorovému výzkumu. Velký potenciál mj. skýtají geolokační data mobilních operátorů, která poskytují obraz o rozložení i mobilitě obyvatelstva a představují alternativu ke konvenčním statistickým zdrojům – jako jsou populační cenzy ČSÚ nebo sčítání dopravy prováděné ŘSD ČR [Dvořák, 2020]. Výhodou je jejich aktuálnost (populační census se provádí s desetiletou periodicitou, sčítání dopravy jedenkrát za pět let), na druhé straně však může být omezující proměnlivá přesnost, různá velikost agregačních jednotek, prováděná anonymizace a v neposlední řadě i vysoká pořizovací cena [viz například Kopáček et al., 2021]. Proto je při analýzách socioekonomické povahy s požadavkem na velký prostorový detail příhodné i nadále vycházet z tradičních konvenčních statistických zdrojů, jako jsou cenzy a registry. Ne vždy se však podaří naplnit jejich skutečný potenciál – například Registr obyvatel (jeden ze čtyř základních registrů České republiky, zřízený zákonem č. 111/2009 Sb., o základních registrech), který obsahuje aktuální data o obyvatelstvu (počet obyvatel, věková struktura) s podrobností na adresní místo, je přístupný pouze orgánům státní správy a samosprávám a je vázáný pouze na trvalé bydliště osoby.

Nespornou výhodou agregace statistických dat do větších prostorových jednotek je podle Buriana et al. [2022]

skutečnost, že data lze lépe studovat a kombinovat je s dalšími statistickými údaji vstupujícími do analýz, protože jsou si vzájemně prostorově kompatibilní. Na druhé straně existují i určitá omezení – zejména variabilita velikosti jednotek [několikanásobně více obcí v rámci jednoho kraje; viz např. Cebrykow, 2017], jejich časová nestabilita (změny hranic obcí nebo slučování či rozdělování) nebo vnitřní nehomogenita jednotek. S posledně jmenovaným fenoménem je spojena tzv. *arbitrárnost* administrativních nebo statistických jednotek [Goodchild et al., 1993; Rosina & Hurbánek, 2016] vymezených z historických, politických, fyzických nebo praktických (např. efektivita sčítání) důvodů, které mnohdy nemají žádný vztah k údajům, jež reprezentují, namísto toho, aby ctily skutečné diskontinuity (přirozené zlomové body) v prostorovém rozložení jevu [Martin, 1989].

Zjevnou nevýhodou agregace je snižování polohové přesnosti sumarizací vstupních dat, která může ztěžovat identifikaci a případné následné zkoumání prostorových procesů. Řešením, jak se vyhnout zpochybnění přesnosti prostorových analýz s příliš hrubým prostorovým rozlišením a se zachováním určité míry „pravdivosti“, je použití uměle vytvořených pravidelných jednotek – gridů [Burian et al., 2022; Rosina & Hurbánek, 2016].

Grid lze definovat jako pravidelnou síť, kde je zkoumaná oblast rozdělena pravidelnou mřížkou, jejímž základním prvkem je buňka (nejčastěji čtvercová nebo šestiúhelníková). Nespornou výhodou gridu je jeho stabilita v čase, možnost volby různých velikostí, vzájemná kompatibilita (v případě čtverců), rovnoměrnost pokrytí a možnost vizualizace absolutních hodnot (díky stejné velikosti všech buněk). Naopak hlavní nevýhodou je obtížné propojení se statistickými údaji, které jsou obvykle k dispozici na úrovni administrativních jednotek. Z tohoto důvodu je agregace dat v administrativních jednotkách běžným postupem při zpracování statistických dat, zejména u dat ze sčítání lidu [European Forum for Geostatistics, 2012; Eurostat, 2023; Klauđa, 2016].

Kromě výše uvedeného způsobu přiřazování statistické informace do buněk

sítě (tj. agregace) existuje druhý, procesně náročnější, způsob. Nejsou-li podrobnější statistická data k dispozici, je možné využít metody disagregace (tzv. „top-down“). Ta ke zpřesnění původní lokace statistické informace využívá např. letecké či satelitní snímky nebo topografická data o území, která reflektují příslušné rozložení statistického jevu v prostoru. Obě metody je možné vzájemně kombinovat. Výsledkem je pak „hybridní model grid“ [Klauđa, 2016].

Nejfrekventovanějším ukazatelem prezentovaným formou gridu je počet obyvatel (respektive hustota zalidnění). Ačkoliv některé státy již přistoupily k poskytování populačních dat v pravidelných mřížkách, v celoevropském kontextu se však stále jedná o minoritu. ČSÚ se v letech 2013–2014 podílel na celoevropském projektu GEOSTAT 1B, jehož výsledkem je celoevropská populační gridová síť a rozpracované agregační i disagregační metody [Klauđa, 2016]. V rámci projektu GEOSTAT 1B vznikla první verze celoevropské gridové datové sady [European Forum for Geostatistics, 2014]. V roce 2021 Eurostat na svých webových stránkách zpřístupnil aktualizaci populačního gridu GEOSTAT s přepočtenými daty za rok 2018; přitom stále nabízí i dřívější verze z let 2006 a 2011 [viz Eurostat, 2023], vše v rozlišení 1 km². Počátkem roku 2023 byla uvolněna provizorní verze populačního gridu „Eurostat Census Grid“ za rok 2021, opět v rozlišení 1 km² [viz tamtéž]. Populační gridy téhož rozlišení i téže aktuality jsou pro Česko nově dostupné i v rámci statistických georeportů a statistického atlasu na Statistickém geoportálu ČSÚ (viz <https://geodata.statistika.cz/as/rozcestnik/>).

V posledních letech převzaly iniciativu i odborné komunity, které začaly poskytovat aktuální údaje o počtu obyvatel v pravidelných sítích a mnohdy svou aktuálností i rozlišením převyšují produkty národních či evropských statistických úřadů a organizací. Za všechny zmíníme European Forum for Geostatistics [2023], POPGRID [2023] nebo WorldPop [2023].

Informace o rozložení obyvatelstva v prostoru jsou nezbytné pro popis naší společnosti i pro studium vzájemných vztahů

mezi prvky humánně-geografické sféry a nejinak tomu je v oblasti územního plánování a při správě měst a regionů. Využívání informačních technologií a jejich produktů (tedy i podpůrných dat o obyvatelstvu s vysokou mírou prostorového detailu) je základním pilířem konceptu *Smart City*. Tradiční datové zdroje (národní registry) poskytují pouze statický pohled vázaný na bydliště nebo pracoviště obyvatel, který je nepřesný a může být značně zavádějící. Slabinou je například běžné uvažování trvalého bydliště osoby, které nemusí dostatečně reflektovat skutečný stav – v praxi je běžné, že osoba žije na jiném místě, než je přihlášena k trvalému pobytu. Zároveň na národní úrovni nedostačuje popisovat obyvatelstvo pouze na základě jediného datového zdroje (kupříkladu prostřednictvím adresních míst z Registru sčítacích obvodů a budov ČSÚ). Je nutné si uvědomit, že rozložení obyvatelstva v prostoru je závislé i na denní době – tj. je proměnlivé ve dne, v noci, v pracovní dny i o víkendech. Potenciální vrstva distribuce obyvatelstva by proto měla vycházet z více datových zdrojů a měla by být vyhotovena ve více než jedné základní variantě. Vzhledem ke zmíněnému zatížení administrativních a statistických jednotek svojí arbitrárností je žádoucí operovat s pravidelnou mřížkou (gridem).

Záměrem autorů je proto eliminovat nastíněné nepřesnosti a slabiny existujících pohledů na distribuci obyvatelstva. V příspěvku je popsán postup agregace několika datových zdrojů s cílem poskytnout výzkumníkům, úředníkům či místním samosprávám řešení pro tvorbu přesnějších datových sad o distribuci obyvatel v prostoru v čase, které se stanou podkladem pro jejich rozhodování a zajistí kvalitnější výstupy. Postup je popsán jak na úrovni metodických kroků, tak i na úrovni technické implementace v prostředí GIS. Výsledky jsou představeny na oblasti Olomouckého kraje.

Datové zdroje a postup zpracování

Datové zdroje

Metodický postup tvorby reprezentativní vrstvy distribuce obyvatelstva, jenž je předmětem tohoto příspěvku, se opí-

rá o celkem šest datových zdrojů. Prvním z nich je **Registr sčítacích obvodů a budov (RSO)**, který spravuje a poskytuje Český statistický úřad (ČSÚ). Ten v rámci tematické sady „Budovy s číslem domovním a vchody“ obsahuje definiční body budov, respektive všech jejich evidovaných částí (vchodů) spolu s vybranými technicko-ekonomickými parametry, včetně počtu obyvatel. Prohlížení a filtrování s omezením počtu záznamů i poskytovaných údajů je možné na adrese <https://apl.czso.cz/irso4/>.

Druhým zdrojem, opět od ČSÚ, jsou dílčí údaje ze **Sčítání lidu, domů a bytů 2021 (SLDB 2021)**, které jsou volně distribuovány ve formě souborů tabulkového procesoru nebo ve formátu CSV (*Comma-separated Values*). V našem případě byly pro výpočet podílu ekonomicky neaktivních obyvatel a nezaměstnaných využity údaje o obyvatelstvu podle ekonomické aktivity na úrovni dílů ZSJ, které jsou veřejnosti nedostupné. Alternativně lze využít volně dostupné údaje za kraj (méně přesné ve srovnání s předchozím). Vzhledem k dílčí nedostupnosti dat tento zdroj v tabulce 1 kvalifikujeme jako „částečně dostupný“. Rovněž bylo vycházeno z **Registru ekonomických subjektů (RES)**, který obsahuje vybrané informace o všech ekonomických subjektech sídlících v České republice (např. kategorie dle počtu pracovníků, adresa). Data z registru ČSÚ poskytuje zdarma ve formátu CSV (viz <https://apl.czso.cz/res/>).

Dalším podpůrným zdrojem dat pro tvorbu reprezentativní vrstvy distribuce obyvatelstva je datová sada o **obsazenosti zařízení poskytující preprimární, primární a sekundární vzdělání**, kterou na vyžádání ve formě tabulky poskytl Krajský úřad Olomouckého kraje (KÚ OLK). Obsahuje mj. adresu zařízení a počet dětí/žáků. Alternativně lze využít údaj o počtu žáků z výkazů Ministerstva školství, mládeže a tělovýchovy České republiky (MŠMT ČR) – výkaz o základní škole M 3 a výkaz o střední škole M 8. Jedná se však o identická data. Většina informací o vzdělávacích institucích je k dispozici ve veřejnosti přístupném Rejstříku škol a školských zařízení, který spravuje MŠMT ČR (viz <https://rejstriky.msmt.cz/rejskol/>) – potřebná informace o počtu dětí/žáků však uvedena není.

Registr poskytovatelů sociálních služeb, který provozuje Ministerstvo práce a sociálních věcí České republiky (MPSV ČR), byl zřízen na základě zákona č. 108/2006 Sb., o sociálních službách, s účinností od 1. 1. 2007. Veřejně dostupné jsou zde informace jako např. forma poskytování sociálních služeb (pobytové, ambulantní, terénní), kapacita zařízení, cílová i věková skupina klientů (např. senioři) a adresa. Data z registru jsou dostupná na adrese <https://data.mpsv.cz/web/data/rpss> ve formátu JSON (*JavaScript Object Notation*).

Hranice zájmového území (v tomto případě Olomouckého kraje) byly získány z **Registru územní identifikace, adres a nemovitostí (RÚIAN)** ve výměnném formátu založeném na XML.

Užité zdroje jsou shrnuty v tabulce 1. Pro realizaci záměru bylo využito i běžně nedostupných datových zdrojů (RSO a Obsazenost MŠ, ZŠ, SŠ) či zdrojů částečně dostupných (detailní data ze SLDB). Pouze pro dva datové zdroje (RSO a RÚIAN) je imanentní geoprostorová podoba dat s přímou aplikovatelností v prostředí GIS.

Předzpracování dat

Prvotní a časově nejnáročnější fází, která předcházela vlastnímu agregáčnímu řešení, byla fáze předzpracování zdrojových dat. Úvodním krokem (vzhledem ke geoprostorové povaze dat) bylo provedení ořezu dat ze sady RSO „**Budovy s číslem domovním a vchody**“ pouze na území Olomouckého kraje, a to prostřednictvím geoprocessingového nástroje *Clip* v prostředí desktopového GIS. V široké paletě technicko-ekonomických parametrů vztažených k adresním místům byl následně dohledán klíčový ukazatel počtu obyvatel, přičemž upřednostněn byl počet obyvatel s obvyklým pobytem vč. cizinců dle SLDB 2021 (označený jako „*BOBYOSL21*“, aktualizovaná obdoba atributu „*BUDOBYOSL*“). Současně je dostupný i počet obyvatel s trvalým pobytem „*BUDOBYTSL*“, nicméně tento ukazatel nemusí dostatečně reflektovat skutečný stav distribuce populace. V případě nedostupnosti nebo zastarávání dat je potenciálně možné při-

Registr / datová sada	Poskytovatel	Použitý údaj	Geodata	Formát	Dostupnost	
1	Registr sčítacích obvodů a budov	ČSÚ	Počet obyvatel	Ano	Esri Shapefile (*.shp, *.shx, *.dbf)	Veřejně nedostupné
2	Sčítání lidu, domů a bytů	ČSÚ	Počet zaměstnaných a počet žáků/studentů	Ne	Excel Spreadsheet (*.xls/*.xlsx)	Částečně dostupné
3	Registr ekonomických subjektů	ČSÚ	Počet pracovníků, zaměstnanců	Ne	Comma-separated Values (*.csv)	Volně dostupné
4	Obsazenost MŠ, ZŠ, SŠ	KÚ OLK	Počet dětí/žáků	Ne	Excel Spreadsheet (*.xls/*.xlsx)	Veřejně nedostupné
5	Registr poskytovatelů sociálních služeb	MPSV ČR	Počet důchodců	Ne	JSON (JavaScript Object Notation)	Volně dostupné
6	Registr územní identifikace, adres a nemovitostí	ČÚZK	Zájmové území (polygon)	Ano	Výměnný formát RÚIAN (*.xml)	Volně dostupné

Tab. 1: Využitě registry / datové sady

stoupit k přepočtu počtu obyvatel na adresní místo násobením příslušného počtu bytů tzv. obložností (počtem osob na 1 trvale obydlený byt; dle SLDB 2021 *obložnost* v Olomouckém kraji odpovídá hodnotě 2,3). Poté byl na základě struktury obyvatelstva Olomouckého kraje podle ekonomické aktivity s podrobností na díly ZSJ (základní sídelní jednotky) od výchozí hodnoty počtu obyvatel s obvyklým pobytem odečten příslušný poměr pracujících a studujících, čímž vznikl atribut zohledňující stav (pouze) ekonomicky neaktivní populace a nezaměstnaných. Tento atribut odstraňuje souběh informací o populaci z prezentovaných datových zdrojů, které se vzájemně překrývají.

Ačkoliv byl přepočet realizován s využitím údajů za díly základních sídelních jednotek, pro větší názornost a přímoučarost jej budeme demonstrovat na příkladu dat za celý kraj, neboť obecný princip výpočtu zůstává zachován. Jak ukazuje tabulka 2, v Olomouckém kraji v roce 2021 (na základě druhé datové sady souhrnné tabulky 1) žilo 619 788 obyvatel. Počet žáků/studentů (87 418) a počet

zaměstnaných poníženy o počet osob na mateřské dovolené (299 467) dohromady činí 386 885 obyvatel. Tato hodnota reflektující pracující a studující populaci představuje cca 62 % z populace celkové (P_{celkem}). Počet ekonomicky neaktivních obyvatel a nezaměstnaných (P_{ENN}) pak lze získat výpočtem dle následujícího vzorce:

$$P_{ENN} = P_{Celkem} - (P_{celkem} \times 0,624)$$

Na základě příslušnosti adresního místa ke konkrétnímu dílu ZSJ byla každému bodově lokalizovanému adresnímu místu přiřazena odpovídající hodnota představující podíl pracujících a studující populace. Vlastní výpočet pak probíhal podle upraveného vzorce, kde již tento podíl vystupuje ve formě proměnné A_{PS} :

$$P_{ENN} = P_{Celkem} - (P_{celkem} \times A_{PS})$$

V případě dat z Registru ekonomických subjektů ČSÚ však byla situace složitější. Nejprve byly odstraněny subjekty, jejichž působnost sahá mimo území Olomouckého kraje, poté byly odstraněny neplatné (neexistující) subjekty a vzhledem k souběhu dat

i všechny mateřské, základní a střední školy. Naproti tomu vyšší odborné a vysoké školy byly ponechány, sehrávají toliko roli pracoviště s odpovídajícím počtem pracovních míst. Z celkového počtu bezmála 137,5 tisíc subjektů bylo 2 tisíce neexistujících a po odfiltrování příslušných školských zařízení (které byly vyděleny pro tvorbu obsazenosti MŠ, ZŠ a SŠ včetně pedagogů a školského personálu – popsáno dále) počet subjektů čítal okolo 135 tisíc. Následně bylo provedeno časově náročné a rozsáhlé geokódování, protože datová sada není poskytována ve formě geodat s možností přímého nasazení v GIS, ale ve formě souboru spustitelném např. v prostředí tabulkového procesoru, kde je jediným lokalizačním atributem poštovní adresa. Nejprve bylo využito komerčního geokódovacího nástroje založeného na HERE API, který však nedokázal geokódovat adresy v menších sídlech (typicky vesnice) a u nových stavebních objektů. Proto bylo nutné geokódování u značné části záznamů dokončit pomocí volně dostupných (limitovaných) řešení – jmenovitě ezGeocodeaGeocodebyAwesomeTable v prostředí Google Tabulky. Samot-

Obyvatelstvo Olomouckého kraje podle SLDB 2021										
v tom podle ekonomické aktivity										
Obyvatelstvo celkem	v tom					z toho				nezjištěno
	pracovní síla	zaměstnaní	z toho		nezaměstnaní	mimo pracovní sílu	nepracující důchodci	na rodičovské dovolené	žáci/studenti	
			pracující důchodci	na mateřské dovolené						
619 788	317 949	301 779	29 284	2 312	16 170	295 909	146 073	13 059	87 418	5 930

Tab. 2: Obyvatelstvo Olomouckého kraje podle ekonomické aktivity za rok 2021 (zdroj: Český statistický úřad, 2021; upraveno)

Vstupní datová sada (vrstva)		Popis	Úroveň podrobnosti	Rok	Zdroj dle tab. 1
A	Počet obyvatel	Počet obyvatel s obvyklým pobytem dle SLDB 2021	Adresní místo	2021	(1)
B	Počet ekonomicky neaktivních obyvatel a nezaměstnaných	Na základě struktury obyvatelstva podle ekonomické aktivity s podrobností na díly ZSJ odečten příslušný poměr pracujících a studujících od původní hodnoty počtu obyvatel s obvyklým pobytem	Adresní místo	2021	(1), (2)
C	Počet pracovníků, zaměstnanců	Průměrný počet pracovníků; bez subjektů MŠ, ZŠ a SŠ	Adresní místo	2022	(3)
D	Počet dětí, žáků; počet pedagogů a školského personálu	Počet pedagogů a školského personálu odvozen z RES (průměrný počet pracovníků)	Adresní místo	2022	(4), (3)
E	Počet seniorů	Odvozeno z kapacity bytového sociálního zařízení, kde jsou seniori primární cílovou skupinou	Adresní místo	2022	(5)
F	Zájmové území	Území Olomouckého kraje	VÚSC (kraj)	2022	(6)

Tab. 3: Přehled datových sad (vrstev) vstupujících do agregačního procesu

ný proces geokódování byl ztížen také kvůli:

- nejednotnému formátu adres v RES;
- chybějícímu číslu popisnému/evidenčnímu, číslu orientačnímu;
- výskytu slov, které do adresní části nepatří (jméno subjektu, orientační popis, umístění vchodu apod.);
- zastaralému názvu ulic nebo číslování budovy;
- nedostatečné úrovni podrobnosti adresy (např. uvedení pouze názvu obce a PSČ).

I proto nakonec bylo nutné některým adresám přiřadit zeměpisné souřadnice manuálně. Při ohledání již geokódovaných dat bylo zjištěno, že u více než 72 tisíc záznamů (tj. 54 % subjektů v Olomouckém kraji) není uveden klíčový atribut „Kategorie dle počtu pracovníků“, který udává interval počtu pracovníků příslušného subjektu (rozpětí od–do) a je pro tvorbu reprezentativní vrstvy distribuce obyvatelstva Olomouckého kraje klíčovým nositelem informace. Nejčastěji tento údaj chybí u fyzických osob podnikajících dle živnostenského zákona (více než 48 tisíc záznamů), společností s ručením omezeným (cca 9 700), spolků (více než 4 800), společenství vlastníků jednotek (3 900) a pobočných spolků (okolo 1 450 záznamů). Chybějící údaje lze snadno doplnit imputací. Druhým specifickým případem jsou subjekty vykázané jako bez zaměstnanců (46 tisíc záznamů, což odpovídá 34 % z celkového počtu subjektů v kraji). Opět se jedná převážně o fyzické osoby podnikající (bezmála

43 700 záznamů), případně ostatní fyzické osoby (cca 2 300), zemědělské podniky (1 600) nebo zahraniční fyzické osoby (500).

V případě, že byl u subjektu uveden atribut „Kategorie dle počtu pracovníků“, byl tento atribut převeden na průměrnou hodnotu stanoveného rozpětí. Údajům u subjektů, kde byl atribut kódován jako „neuvedeno“, byla provedena imputace na základě průměru příslušného typu subjektu ze souboru dat s úplným atributem „Kategorie dle počtu pracovníků“. Subjektům bez zaměstnanců byla přiznána hodnota „1“.

Kromě získaných údajů o počtu dětí v MŠ, respektive žáků na ZŠ a SŠ – dostupných v rámci sestavy **obsazenosti zařízení poskytující preprimární, primární a sekundární vzdělání** – bylo nutné k těmto školským zařízením doplnit i informace o počtu zaměstnanců a pedagogů. To bylo řešeno navázáním (funkce *Join*) vyděleného souboru dat z RES a využití atributu „Kategorie dle počtu pracovníků“. Zpracování bylo obdobné jako v případě dat z RES – rozpětí počtu pracovníků bylo převedeno na průměrnou číselnou hodnotu. U zanedbatelného počtu školských zařízení byly potřebné údaje získány z výročních zpráv. Následně bylo třeba ošetřit situaci, kdy jsou institucionálně, v různých kombinacích, sloučeny MŠ, ZŠ a SŠ. Bylo totiž nutné zajistit to, aby se počet pracovníků v rámci instituce započítával pouze jedenkrát – údaje o počtech dětí, žáků jsou

poskytovány zvlášť (na úrovni druhů poskytované úrovně vzdělání – odděleně pro MŠ, ZŠ a SŠ), naproti tomu v RES je v případě případné kombinace preprimárního, primárního a sekundárního vzdělávání informace o počtu pracovníků vedena sdruženě, na institucionální úrovni – jeden subjekt, jeden údaj. Po zkompletování atributových informací bylo nutné provést geokódování. V Olomouckém kraji se v roce 2022 nacházelo 484 mateřských škol, 329 škol základních a 96 středních škol.

Registr poskytovatelů sociálních služeb MPSV ČR rozlišuje celkem 32 druhů zařízení. Pro naše účely však bylo nutné vyfiltrovat pouze ty, kde jsou seniori dominantní cílovou skupinou a které jsou bytové (rozumějme zařízení, kde seniori dlouhodobě žijí a zdržují se – nikoliv ta, kam nepravdělně docházejí). Takové znaky splňovaly tři druhy zařízení – domovy pro seniory, domovy pro osoby se zdravotním postižením a domovy se zvláštním režimem. Tato zařízení byla následně vyhledána pro území Olomouckého kraje (celkem 84 zařízení) včetně příslušných popisných informací (kód zařízení, název, adresa). Kapacita příslušného zařízení byla chápána jako uvažovaný počet dlouhodobě žijících seniorů v zařízení. Validace nad výročními zprávami potvrdila plně obsazené nebo téměř obsazené kapacity. Vzhledem k neprostorové povaze registru a nově vzniklého tabelárního soupisu bylo i v tomto případě nutné provést geokódování.

Vstupní (modifikované nebo nově vzniklé) datové sady pro proces agregace, včetně klíčových údajů o populaci, shrnuje tabulka 3.

Vlastní řešení agregace dat

Vlastní agregace pro tvorbu reprezentativní vrstvy distribuce obyvatelstva demonstrovaná na příkladu Olomouckého kraje byla realizována ve dvou variantách, reflektující:

- noční stav obvyklého pracovního dne;
- denní stav obvyklého pracovního dne.

Prezentované řešení bylo navrženo a sestaveno v prostředí desktopového GIS, jmenovitě v softwaru ArcGIS Pro. Navržené postupy je možné opakovaně aplikovat nad novými daty, na jiném území, s jiným prostorovým rozlišením.

Agregační technika prvně jmenované varianty – **noční stav obvyklého pracovního dne** – je z hlediska technického podstatně jednodušší a přímočařejší. Jak vyplývá ze schématu na obrázku 1, po vytvoření pravidelné mřížky (nástroj *Create Fishnet*) s prostorovým rozlišením 100 m × 100 m dochází na úrovni jednotlivých buněk mřížky k sumaci bodových záznamů vstupní vrstvy A (nástroj *Summarize Within*, atribut „BOBYOSL21“). Následně je – respektive může být – převeden vektorový výstup agregovaných dat na rastr (*Feature to Raster*).

V případě **denního stavu obvyklého pracovního dne** pak všechny vrstvy

B–F z tabulky 3 vstupují do agregačního procesu dle následujícího schématu na obrázku 2. Úvodním krokem je, obdobně jako u předchozího agregčního schématu, vygenerování pravidelné mřížky s prostorovým rozlišením 100 m × 100 m (modifikovatelné; nástroj *Create Fishnet*) pro zájmové území F. Poté dochází k dílčím agregacím vrstev B–E prostřednictvím *Summarize Within*, kdy je pro každou z vrstev na úrovni buněk gridu provedena sumace příslušných bodových záznamů. Tyto agregáty jsou následně sjednoceny (*Union*) a opět sečteny (*Calculate Field* s aktivovanou funkcí součtu), čímž vznikne výsledná agregovaná vrstva ve vektorové podobě. Ta je závěrem převedena na rastr pomocí nástroje *Feature to Raster*.

Představená schémata na obrázcích 1 a 2 počítají s možností modifikace rozlišení, jakož i volbou tvaru buňky gridu (čtverec, hexagon). Ořez vstupních vrstev A–E na zájmové území je řešen na úrovni předzpracování dat – v případě jeho neprovedení budou zohledněny i body ležící mimo polygon zájmového území (uvnitř území přesahujícího polygonu), přičemž body ležící mimo vygenerované buňky gridu se automaticky nepoužijí, protože princip kontrolního ořezu je zajištěn nativně samotným nástrojem *Summarize Within*.

Použití cílových jednotek ve formě gridu však není podmínkou. Kritéria jako velikost, tvar a rozsah cílových jednotek je možné volit na základě osobního požadavku a účelu příslušné studie. Například z hlediska urbanis-

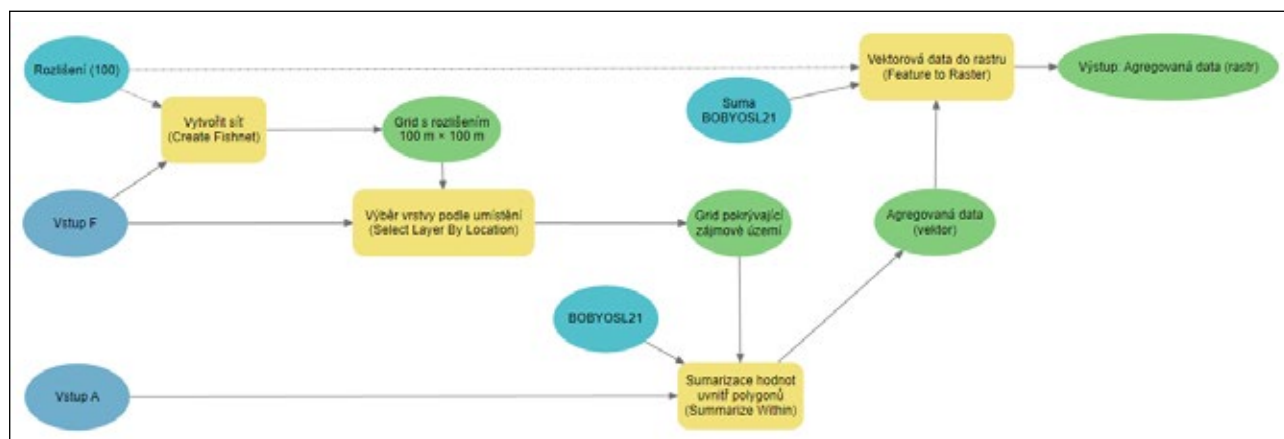
tické teorie i územně plánovací praxe mohou být agregovaná data vztažena k urbanistické struktuře, skupinám objektů nebo volnému prostoru mezi objekty.

Výsledek ve formě reprezentativních vrstev distribuce obyvatelstva

Výsledkem agregace prostorových datových sad z tabulky 3 podle schémat na obrázku 1 a 2 jsou rastrová data v pravidelné čtvercové mřížce s rozlišením 100 m × 100 m reflektující distribuci populace v Olomouckém kraji pro denní a noční stav obvyklého pracovního dne.

Základní představu o distribuci obyvatelstva na úrovni Olomouckého kraje je možné si udělat prostřednictvím obrázku 3, kde jsou na základě směrodatné odchylky zvýrazněny oblasti s vysokým a nízkým stavem populace. Navzdory různému rozpětí hodnot je díky tomuto způsobu vizualizace možné stavy vzájemně „opticky“ porovnávat – vlevo se nachází mapa reflektující denní stav, vpravo je nastíněna situace pro stav noční. Záměrem těchto map je prozatím pouze poukázat na to, v jakých místech se v různých částech dne vyskytují/koncentrují lidé, nikoliv poskytnout konkrétní kvantifikaci těchto stavů.

Na první pohled je možné si povšimnout, že za denního stavu dochází k předpokládanému poklesu obyvatelstva ve venkovských oblastech či menších sídlech a jeho koncentrace ve větších sídlech. Naproti tomu ven-



Obr. 1: Schéma agregace dat pro variantu reflektující noční stav obvyklého pracovního dne

kovská sídla mnohem více vynikají za nočního stavu, kdy není uvažována dojíždka do zaměstnání a za vzděláním a pobyt na pracovišti a ve vzdělávacích zařízeních. V červeně ohraničeném výřezu na obrázku 3 je pak v detailu přiblíženo území Olomouce s nenulovou koncentrací obyvatelstva – opět se jedná o zvýrazněnou vizualizaci s cílem odlišit městské jádro od okolního neosídleného území.

Po představení globálního pohledu na distribuci obyvatelstva v Olomouckém kraji je příhodné přistoupit ke konkrétní kvantifikaci populace a přiblížit rozdíly mezi denním a nočním stavem na úrovni konkrétních obcí. Za tímto účelem byly pro tři nejlidnatější města v kraji vytvořeny podrobnější mapy, které vsazují výsledná data o populaci v pravidelné mřížce do kontextu s katastrálními územními a hlavní silniční infrastrukturou. Kontinuální proměnná předpokládaného počtu obyvatel je v těchto mapách rozdělena tzv. *metodou přirozených zlomů* do sedmi skupin (intervalů), vyznačujících se minimální vnitroskupinovou diferencovaností a naopak co největšími rozdíly mezi skupinami samotnými.

mi. Tato konfigurace zajistila odlišení oblastí s vysokým stavem populace od oblastí se stavem nízkým, což je klíčové pro interpretaci výsledných dat a další vyvozování závěrů.

Situace v nejlidnatějším městě kraje a hanácké metropoli, Olomouci, je znázorněna na obrázcích 4 (pro denní stav pracovního dne) a 5 (pro noční stav pracovního dne). Z těchto map je patrné, že při denní době vyvstávají jakási „ohniska“ roztržštěná napříč celým územím města, které reprezentují pracoviště a školská zařízení s vysokou koncentrací populace, typicky rezidenční části města jsou naopak utluštěny. Noční stav naopak předpokládá přítomnost populace v lidských obydlích, proto je možné sledovat zvýraznění rezidenčních částí města, zejména pak sídlišť (Neředín, Nová Ulice, Povel, Nové Sady či Lazce).

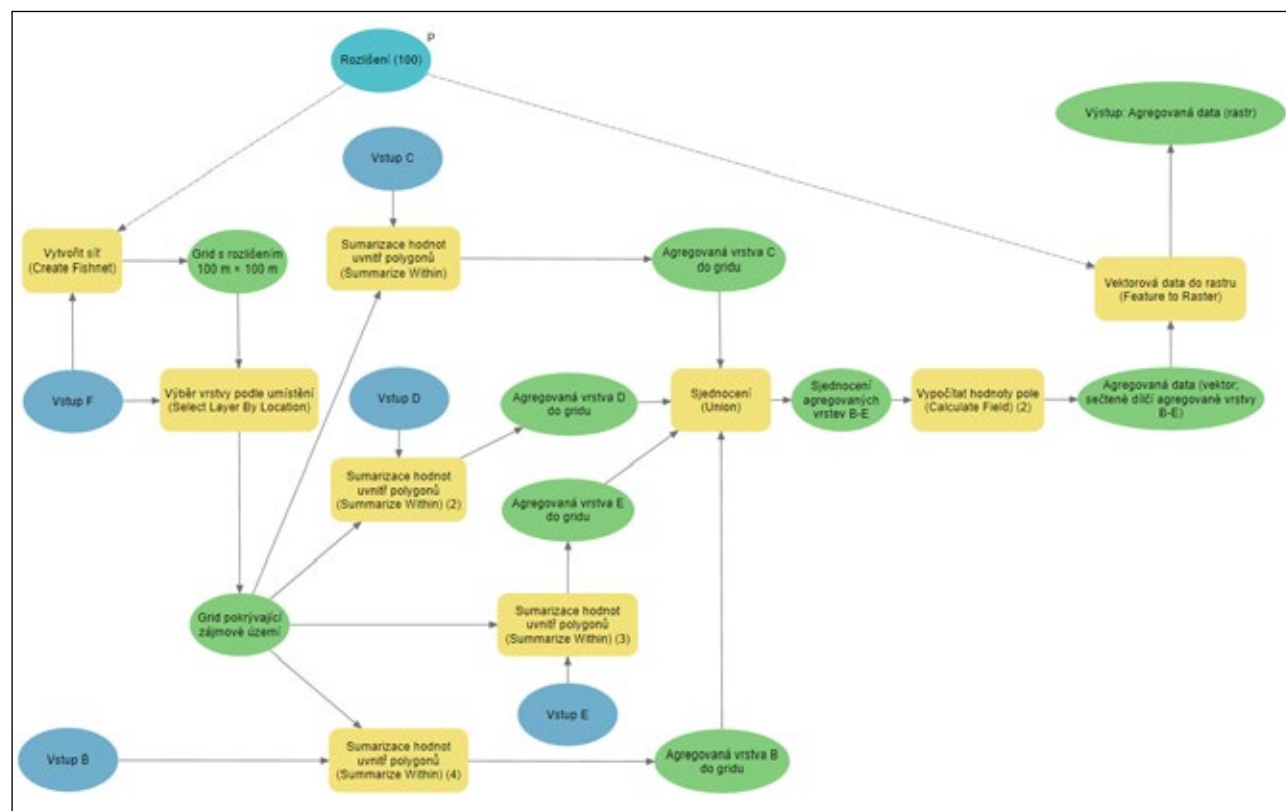
Obdobné projevy lze sledovat i ve druhém největším městě v kraji – Prostějově. I zde za uvažovaného denního stavu pracovního dne (viz obr. 6) vyvstávají pracoviště, vzdělávací a sociální zařízení a jiná další místa, kde se lidé obvykle setkávají.

Ryze rezidenční části Prostějova, které přirozeně představují oblasti s vysokou koncentrací obyvatel, je možné identifikovat prostřednictvím mapy na obrázku 7.

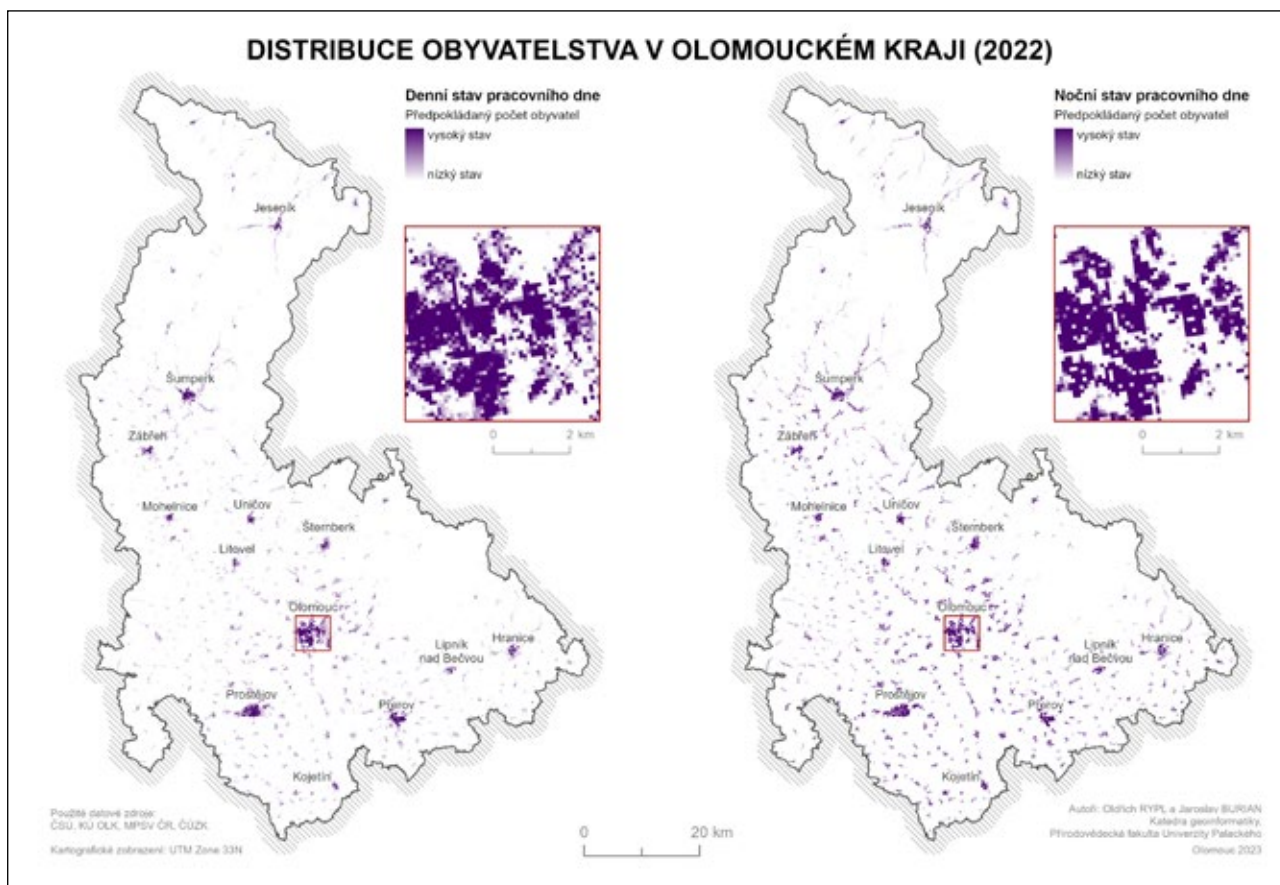
Dvojice podrobnějších map zachycující denní a noční stav pracovního dne byly vytvořeny i pro třetí nejlidnatější sídlo, Přerov – viz následující obrázky 8 a 9.

Uvedené mapy lze rovněž doplnit i souhrnným tabulárním přehledem (viz tabulka 4), který pro 13 nejvíce osídlených měst Olomouckého kraje uvádí maximální (max.), průměrnou (μ) a celkovou (Σ) hodnotu obyvatelstva při denním a nočním stavu pracovního dne. Celkovou hodnotu příslušného stavu je navíc možné srovnat s oficiální hodnotou počtu obyvatel z roku 2022.

Na první pohled je možné si všimnout, že celková hodnota obyvatelstva za denního stavu mnohdy dosahuje více než dvojnásobku odhadu počtu obyvatel za rok 2022. To je dáno tím, že metodika v jeden okamžik předpokládá stoprocentní přítomnost všech



Obr. 2: Schéma agregace dat pro variantu reflektující denní stav obvyklého pracovního dne



Obr. 3: Distribuce obyvatelstva v Olomouckém kraji

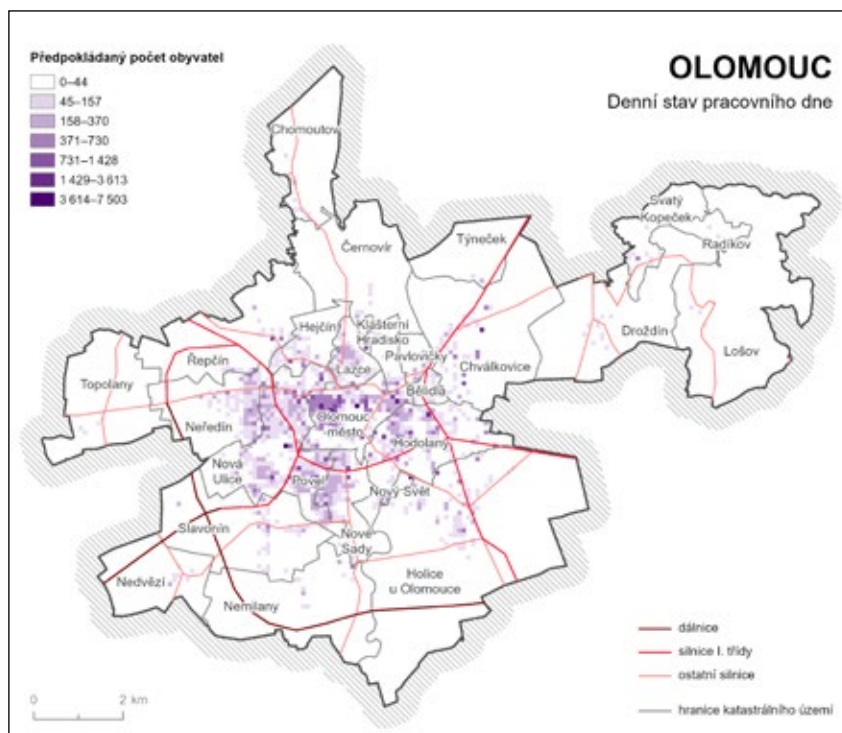
pracovníků na pracovišti, všech dětí a žáků ve vzdělávacích institucích, kterého ve skutečnosti není nikdy dosaženo. Dále je nutné si uvědomit, že do tohoto čísla se nepromítají pouze obyvatelé příslušné obce, ke které je vztaheno, ale i populace jak z ostatních obcí kraje (např. v Olomouci studují, pracují, pobývají v sociálních zařízeních i obyvatelé z jeho okolí, kteří běžně nejsou započteny mezi obyvateli), tak i krajů okolních, kteří do Olomouckého kraje za práci dojíždí. Hodnota počtu obyvatel je rovněž zatížena skutečností, že mohou být započítáni i pracovníci poboček/provozoven ležících mimo území Olomouckého kraje – v RES totiž není možné spolehlivě odlišit sídlo subjektu od jeho dílčí provozovny. Je tedy zřejmé, že rozdíly mezi hodnotami jsou bezesporu dány povahou vstupních dat samotných (řešeno v kapitole „Diskuse“).

Rozdíly mezi udávaným počtem obyvatel z roku 2022 a nočním stavem nejsou tak markantní, jako tomu bylo v případě stavu denního – odchylky se

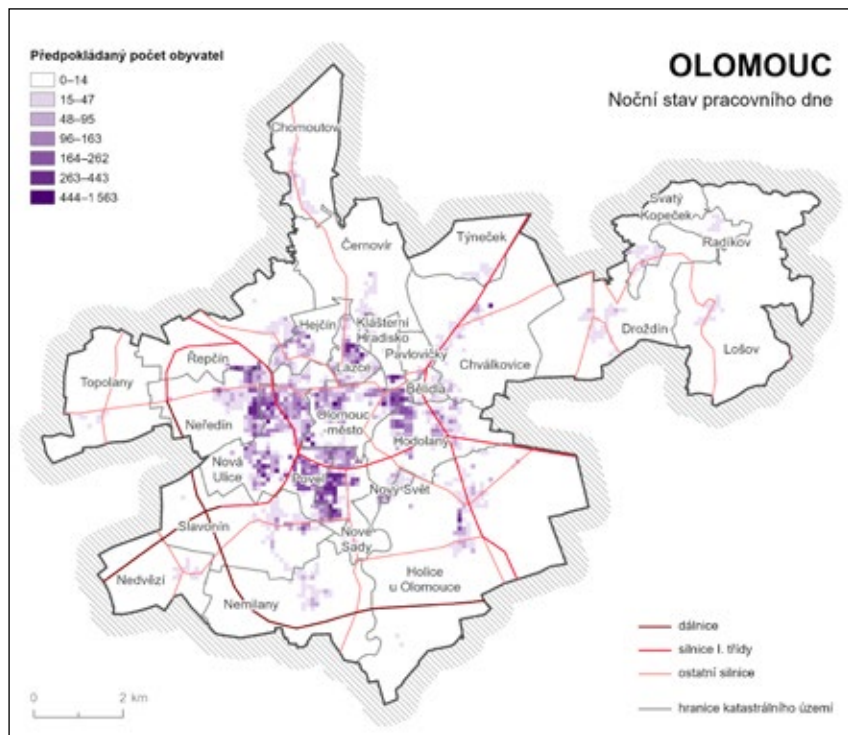
pohybují v jednotkách tisíců až stovkách obyvatel. Jsou způsobeny užitím údajů o obvyklém bydlišti včetně cizinců místo bydliště trvalého.

Diskuse

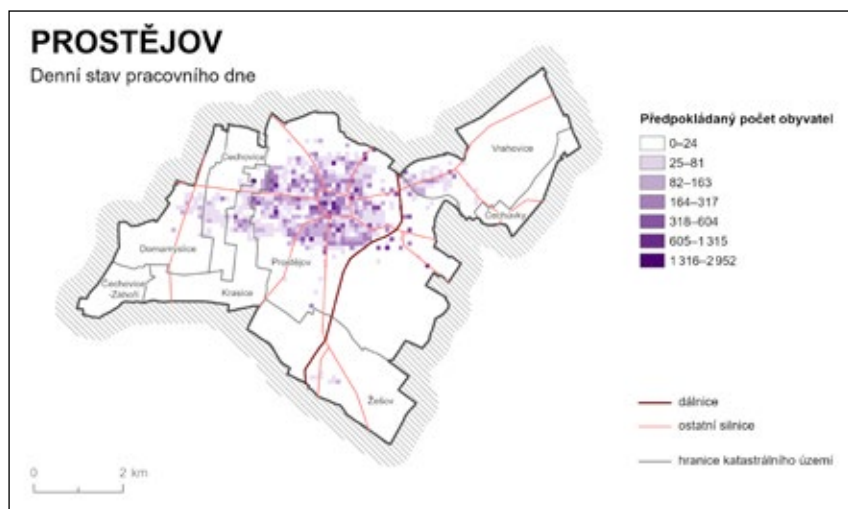
Jak již bylo v různých částech příspěvku zmíněno, představená metodika



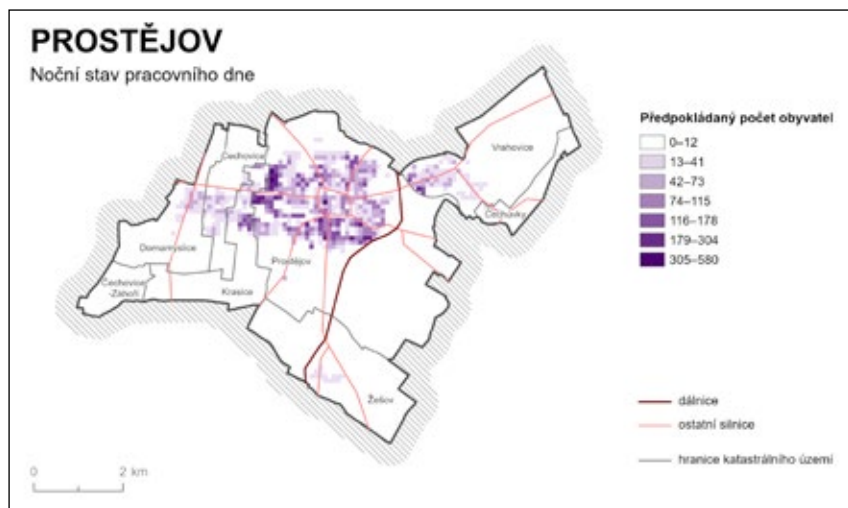
Obr. 4: Distribuce obyvatelstva v Olomouci při denním stavu pracovního dne



Obr. 5: Distribuce obyvatelstva v Olomouci při nočním stavu pracovního dne



Obr. 6: Distribuce obyvatelstva v Prostějově při denním stavu pracovního dne



Obr. 7: Distribuce obyvatelstva v Prostějově při nočním stavu pracovního dne

naráží na problémy související s nedostatečnou podrobností a neurčitostí vstupních dat. Zásadní slabinu prezentovaného metodického postupu představuje atribut „Kategorie dle počtu pracovníků“ z Registru ekonomických subjektů ČSÚ. Počet pracovníků a zaměstnanců v podnicích v podobě intervalu (rozpětí od–do), namísto exaktní číselné hodnoty, způsobuje strmý nárůst hodnot v denním stavu distribuce populace (celková hodnota obyvatelstva za denního stavu mnohdy dosahuje více než dvojnásobku odhadu počtu obyvatel za rok 2022). Přesnější dostupnější zdroj, který by poskytl tuto informaci, však neexistuje. Omezením je rovněž skutečnost, že v registru není možné spolehlivě odlišit sídlo subjektu a jeho dílčí provozovny. Proto lze připustit i situaci, že příslušná hodnota počtu pracovníků subjektu sídlícího v Olomouckém kraji může zahrnovat i pracovníky z poboček/provozoven ležících mimo jeho území.

Při odečítání poměru studujících a pracujících ve snaze odstranit souběh datových sad jsou odečítáni i studenti vyšších odborných škol a vysokých škol, se kterými v gridu není kalkulováno. Studenti této formy studia (bezmála 16 tisíc s trvalým bydlištěm v Olomouckém kraji) však ve srovnání se skupinou populace v preprimárním, primárním a sekundárním vzdělávání (okolo 110 tisíc) tvoří výraznou menšinu. Mimoto, postihnout s odpovídající přesností bydliště studenta vyšší odborné a vysoké školy je problematické, s vědomím toho, že tyto typy vzdělávacích institucí jsou mnohdy nadregionálního významu, což tuto situaci ještě více komplikuje. Rovněž by bylo příhodné odfiltrovat z denního stavu dojížděku za prací a studiem z ostatních krajů.

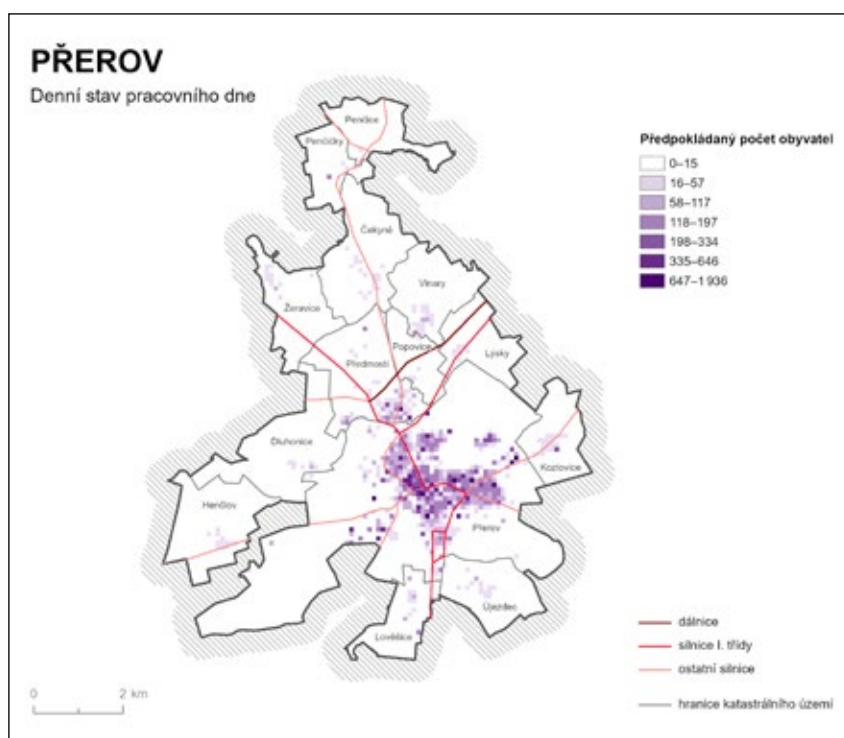
Zásadním aspektem, který nebývá při popisu distribuce obyvatelstva v prostoru zohledňován, je závislost na denní době – rozmístění obyvatelstva se liší ve dne, v noci, v pracovní dny i o víkendech. Na rozdíl od tradičních přístupů (např. prostého využití adresních míst z RSO) byly představeny přístupy k realizaci reprezentativních vrstev pro denní a noční stav pracovního dne. Realizace víkendového stavu je problematická, protože chování

obyvatel je odvislé mj. od turistické atraktivity a infrastruktury území, tedy je dáno samotným charakterem území a dalšími skrytými specifiky. Jejich popis je sám o sobě ztížený tím, že neexistuje žádný konvenční registr nebo datová sada, která by jej standardizovaně hodnotila. Technicky náročné je i zabezpečit samotné „sčítání“ osob mimo zalidněné oblasti.

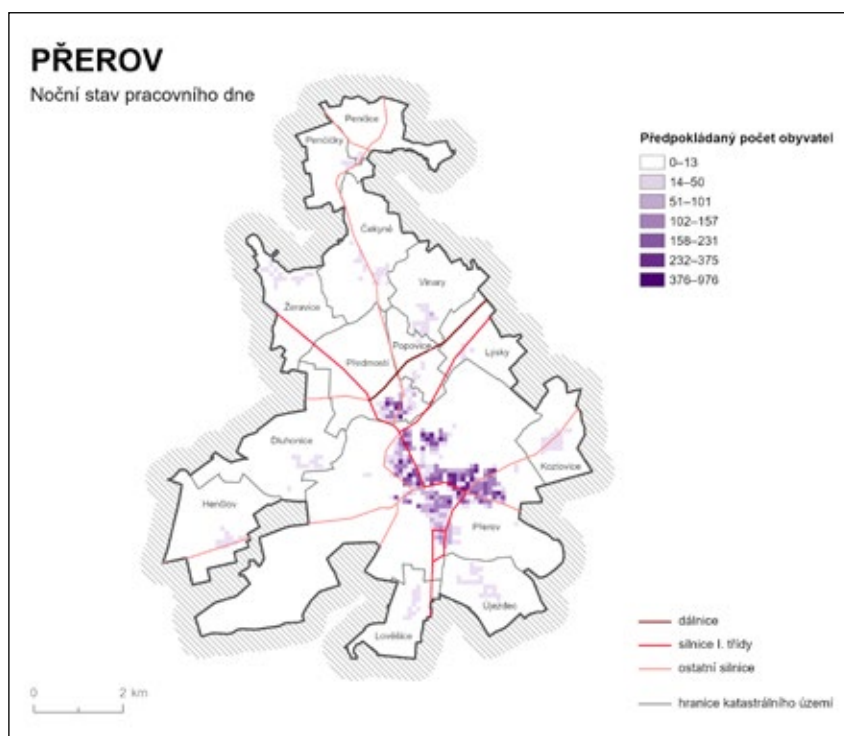
Odchylky předpokládaného denního a nočního stavu (S_p) od oficiálního odhadu počtu obyvatel pro rok 2022 (S_o) ukazuje tabulka 5, kde kladná hodnota (+) odchylky znamená nadhodnocení a záporná (–) podhodnocení. Zatímco rastr reflektující noční stav obyvatelstva dosahuje odpovídající přesnosti, v případě denního stavu se odchylka výrazně zvyšuje s rostoucí velikostí sídla – největší nadhodnocení nastává u nejlidnatějších sídel v kraji, menší sídla do tisíce obyvatel jsou stanovená celkem věrně. Je nutné si uvědomit, že je srovnáván obvyklý pobyt vč. cizinců s oficiálním odhadem populace založeném na trvalém pobytu – odchylky na přiměřené úrovni jsou tak naprosto přirozené.

Věrohodnost geoinformatických výstupů do jisté míry předurčuje kvalita a podrobnost dat, ze kterých tyto výstupy vychází. Kvalita prezentovaných výsledků sice odpovídá současným možnostem a je zatížena dílčí mírou nejistoty a nepřesnosti (více či méně jako každý z existujících analytických výstupů socioekonomických jevů), nicméně pro přenesení z teoretické roviny do roviny aplikační je jejich úroveň dostatečná. Představená metodika a uvedená agregační schémata jsou univerzálně platná a mohou být aplikována opětovně, tedy i na datech, které vykazují lepší určitost a přesnost a zajistí lepší výsledky, než je tomu v tomto příspěvku.

Velký potenciál pro validaci výsledků spojených s distribucí obyvatelstva a jejich mobilitou přinášejí data mobilních operátorů, která by mohla pomoci eliminovat neurčitost a nepřesnost existujících datových sad popisujících obyvatelstvo a další jevy socioekonomické povahy. Na druhé straně tato data už ze své povahy vzniku nemohou být dostupná na tak vysoké úrovni prostorové



Obr. 8: Distribuce obyvatelstva v Přerově při denním stavu pracovního dne



Obr. 9: Distribuce obyvatelstva v Přerově při nočním stavu pracovního dne

přesnosti (gridová síť 100 × 100 m). Nevýhodou pramenící z povahy samotné metody agregace je určitá ztráta polohové přesnosti, která může ztížit identifikaci a následné zkoumání prostorových procesů. Cílové prostorové jednotky (tj. územní jednotky, do kterých jsou data agregována) jsou

rovněž zatíženy *problémem měnitelné plošné jednotky*, známým pod zkratkou MAUP (*Modifiable Areal Unit Problem*) [Openshaw, 1983]. Úroveň agregace a volba měřítka sledovaného jevu může mít zásadní vliv na výsledky příslušné analýzy nebo výstupu, neboť hranice prostorových jednotek

Obec	Počet obyvatel (2022)	Denní stav obyvatelstva			Noční stav obyvatelstva		
		max.	μ	Σ	max.	μ	Σ
Olomouc	99 496	7 503	25,72	265 860	1 563	10,20	105 419
Prostějov	43 055	2 952	23,64	92 437	626	11,16	43 664
Přerov	41 404	1 936	12,46	72 878	976	7,00	40 950
Šumperk	24 910	1 300	18,83	52 539	483	9,02	25 165
Hranice	17 495	2 088	6,81	33 937	358	3,50	17 459
Zábřeh	13 281	733	6,17	21 370	394	3,89	13 458
Šternberk	13 144	786	4,34	21 195	493	2,67	13 005
Uničov	11 066	2 259	3,84	18 520	266	2,27	10 963
Jeseník	10 658	599	6,01	23 013	252	2,75	10 551
Litovel	9 567	768	3,39	15 758	297	2,04	9 477
Mohelnice	9 103	3 503	3,61	16 701	309	1,98	9 143
Lipník nad Bečvou	7 826	480	3,93	11 992	343	2,49	7 617
Kojetín	5 837	515	2,38	7 398	196	1,83	5 680

Tab. 4: Souhrnný přehled denních a nočních stavů distribuce populace u vybraných obcí Olomouckého kraje

jsou mnohdy vymezovány uměle a lze je proto v jisté míře modifikovat. Autoři si jsou těchto skutečností vědomi, v metodice užívají pravidelné čtvercové sítě s rozlišením 100 m × 100 m, které svojí podrobností považují za optimální pro účely územního plánování. V případě potřeby (využití metodiky pro jiné účely) je možné tvar a velikost (rozlišení) cílových jednotek měnit.

Závěr

V příspěvku byla na příkladu Olomouckého kraje představena metodika agregace několika datových zdrojů pro tvorbu reprezentativní vrstvy distribuce obyvatelstva, která je alternativou k tradičním přístupům založených na prosté informaci o počtu obyvatel na úrovni adresních míst nebo budov. Autoři vycházeli z premisy, že rozmístění obyvatelstva se liší jak ve dne a v noci, tak i o pracovních dnech či o víkendech. Pro popis distribuce tak

nepostačuje produkt ve formě vrstvy o jedné základní variantě, nýbrž o variantách několika. Informace o rozložení obyvatelstva ve formě mapy nebo prostorových dat, a to ve vysoké míře prostorového detailu, by měly být jedním z hlavních podkladů samospráv při chytrém řízení města (tzv. koncept *Smart City*) vedoucí ke správným rozhodnutím a dynamickému rozvoji města.

Metodický postup byl uveden soupisem užitých datových zdrojů a příslušných atributů, které jsou klíčovými nositeli informace o populaci. Na to bylo navázáno popisem důležité fáze předpracování dat, v rámci níž byla řešena eliminace souběhu informací (laicky řečeno „aby v důsledku více datových zdrojů užitých datové sady neřikaly to samé“), transformace dat do geoprostorové podoby, imputace chybějících údajů, jakož i další výpočty řešící neurčitost některé ze vstupních dato-

Kategorie sídla dle počtu obyvatel	Průměrná odchylka $S_p - S_o$	
	Denní stav	Noční stav
< 1 000	+ 100,57	- 1,55
<1 000; 2 000)	+ 515,10	- 31,03
<2 000; 5 000)	+ 1 336,16	- 88,87
<5 000; 10 000)	+ 4 879,00	- 104,00
> 10 000	+ 36 360,00	+ 680,56

Tab. 5: Odchylka předpokládaného stavu populace od oficiální hodnoty počtu obyvatel z roku 2022

vých sad. Popis vlastního agregčního řešení obsahoval dvě schémata a jejich popis jak na úrovni metodických kroků, tak i na úrovni technické implementace v prostředí GIS. Tyto postupy pak byly aplikovány na data za Olomoucký kraj. Jako výsledek agregace byla představena rastrová data v pravidelné čtvercové mřížce s rozlišením 100 m × 100 m reflektující denní a noční stav obvyklého pracovního dne, prezentovaná jak ve formě přehledové mapy za celý kraj, tak i formou podrobnějších map kvantifikujících populaci ve třech největších městech kraje. To bylo vhodně doplněno souhrnnou tabulkou se základním statistickým přehledem za 13 nejlidnatějších měst v Olomouckém kraji.

V diskusi byly řešeny limity a omezení vstupních datových sad (neurčitost, nedostatečná podrobnost), které mají nepříznivý vliv na výsledek. Předpokládaný denní a noční stav obyvatelstva byl konfrontován s oficiálním odhadem ČSÚ, vypočtena byla i odchylka znázorňující přesnost výstupů. Navzdory zjevným nepřesnostem (zejména denního stavu populace), představená metodika a uvedená agregční schémata zůstávají obecně platnými a mají potenciál pro opakované využití při analýzách socioekonomické povahy jakéhokoliv rozsahu a stupně podrobnosti. Ačkoliv se prvotní záměr autorů poskytnout výzkumníkům, úředníkům či místním samosprávám řešení pro tvorbu přesnějších datových sad o distribuci obyvatel v prostoru v čase nepodařilo uspokojivě naplnit, byl navržen potenciální metodický postup řešení a bylo upozorněno na skutečnosti, které v budoucnu mohou přispět k jeho naplnění.

Použité zdroje:

BURIAN, J., ZAPLETAL, J. & PÁSZTO, V. (2022): Disaggregator – a tool for the aggregation and disaggregation of spatial data. In: *Earth Science Informatics*, 15(2), 1323–1339. Dostupné z: <https://doi.org/10.1007/s12145-021-00737-9>.

CEBRYKOW, P. (2017): Elaboration of topographic bases for statistical maps, their contents and importance. In: *Polish Cartographical Review*, 49(3), 97–106. Dostupné z: <https://doi.org/10.1515/per-2017-0008>.

ČESKÝ STATISTICKÝ ÚŘAD (2021): *Ekonomická aktivita*. Sčítání 2021. Dostupné z: <https://www.scitani.cz/ekonomicka-aktivita>.

DVOŘÁK, M. (2020): *Geolokační data mobilních operátorů – principy, příklady, otázky*. Portál otevřených dat České republiky. Dostupné z: <https://data.gov.cz/články/geolokační-data-mobilních-operátorů-principy-příklady-otázky>.

EUROPEAN FORUM FOR GEOSTATISTICS (2012): *GEOSTAT 1A – Representing Census data in a European population grid*. Dostupné z: <https://www.efgs.info/wp-content/uploads/geostat/1a/GEOSTAT1A-final-report.pdf>.

EUROPEAN FORUM FOR GEOSTATISTICS (2014): *GEOSTAT 1B Final Report*. Dostupné z: <https://www.efgs.info/wp-content/uploads/geostat/1b/GEOSTAT1B-final-technical-report.pdf>.

EUROPEAN FORUM FOR GEOSTATISTICS (2023): EFGS Data. Dostupné z: <https://www.efgs.info/data/>.

EUROSTAT (2023): *GEOSTAT – GISCO*. Dostupné z: <https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography/geostat>.

GOODCHILD, M. F., ANSELIN, L. & DEICHMANN, U. (1993): A Framework for the Areal Interpolation of Socioeconomic Data. In: *Environment and Planning A: Economy and Space*, 25(3), 383–397. Dostupné z: <https://doi.org/10.1068/a250383>.

KLAUDA, P. (2016): Prostorově určená statistická data. In: *Statistika&My*, 6(5), 18–19.

KOPÁČEK, G., VAISHAR, A., & ŠIMARA, E. (2021): The dynamics of population distributions in cities based on daily mobile phone operations: A case study of some Moravian cities. In: *Moravian Geographical Reports*, 29(1), 71–86. Dostupné z: <https://doi.org/10.2478/mgr-2021-0006>.

MARTIN, D. (1989): Mapping Population Data from Zone Centroid Locations. In: *Transactions of the Institute of British Geographers*, 14(1), 90. Dostupné z: <https://doi.org/10.2307/622344>.

OPENSHAW, S. (1983): *The modifiable areal unit problem*. Geo Books.

POPGRID (2023): *POPGRID Data Collaborative*. Dostupné z: <https://popgrid.org/>.

ROSINA, K. & HURBÁNEK, P. (2016): *Spatial Disaggregation of Population Density Using Land Cover and Remote Sensing Data: Vol. Geographia Slovaca* (31st ed.). Geografický ústav SAV.

WILSON, B., WILSON, N. & MARTIN, S. (2021): Using GIS to Advance Social Economics Research: Geocoding, Aggregation, and Spatial Thinking. In: *Forum for Social Economics*, 50(4), 480–504. Dostupné z: <https://doi.org/10.1080/07360932.2018.1509798>.

WORLDPOP (2023): *Open Spatial Demographic Data and Research*. Dostupné z: <https://www.worldpop.org/datacatalog/>.

Tento článek vznikl v rámci projektu „MOSPREMA: Predikce a management kalamitních stavů komárů pro zachování biodiverzity v lužních lesích“ (SFŽP ČR) a projektu „Analýza, modelování a vizualizace prostorových jevů pomocí geoinformačních technologií II.“ (IGA_PrF_2023_017) za podpory interní grantové agentury Univerzity Palackého v Olomouci.

Mgr. Oldřich Ryppl

✉ oldrich.rypl01@upol.cz

doc. RNDr. Jaroslav Burian, Ph.D.

✉ jaroslav.burian@upol.cz

Katedra geoinformatiky

Přírodovědecká fakulta

Univerzita Palackého v Olomouci

ENGLISH ABSTRACT

Data Sources Aggregation for the Creation of a Representative Layer of the Population Distribution of the Olomouc Region, by Oldřich Ryppl & Jaroslav Burian

Information about the distribution of population in space is essential for describing our society and for studying the interrelationships between elements of the human-geographical sphere - this is so also in the field of spatial planning and the management of cities and regions. However, traditional data sources (usually national registers) provide only a static view linked to the place of residence (or workplace) of the population, which is inaccurate and can be highly misleading. When analysing phenomena and their causes, which are often independent of administrative boundaries, it is desirable to eliminate the apparent inaccuracy and provide researchers, officials or local governments with accurate and sufficiently detailed data for informed decision-making and obtaining better quality outputs. This paper describes a procedure for aggregating several data sources to provide a solution for producing more accurate datasets on the spatial distribution of population over time. The procedure is described both at the level of methodological steps and at the level of technical implementation in a GIS environment. The results are presented for the Olomouc region.