

**Univerzita Palackého v Olomouci**  
**Přírodovědecká fakulta**  
**Katedra geoinformatiky**

# **TESTOVÁNÍ DISAGREGAČNÍCH METOD V GIS**

**Bakalářská práce**

**Jan ZAPLETAL**

**Vedoucí práce RNDr. Jaroslav BURIAN, Ph.D.**

**Olomouc 2019**  
**Geoinformatika a geografie**

## **ANOTACE**

Práce je zaměřena na testování vybraných disagregačních metod za účelem přepočtu vybraných statistických ukazatelů do nižších územních jednotek. Hlavním cílem práce je pro vybrané statistické ukazatele – počet obyvatel, podíl nezaměstnaných osob a průměrnou hrubou mzdu provést přepočet v několika úrovních podrobnosti a výsledky zhodnotit na referenčních datech, která jsou dostupná na detailní úrovni. Pro každý statistický ukazatel navrhnout vhodná pomocná data, jejich váhy a vhodnou velikost buňky pro přepočet.

Disagregace statistických dat byla testována v rámci dvou testovacích území. Pro každý statistický ukazatel bylo v rámci testovacích území použito několik typů pomocných dat.

Výsledkem práce jsou vizualizace všech příkladů přepočtených dat. Tyto vizualizace jsou dostupné v digitální podobě v projektu ArcGIS for Desktop přiloženém na DVD. Dalšími výsledky jsou zhodnocení a porovnání přesnosti přepočtů pomocí jednotlivých typů pomocných dat. Vybrané vizualizace jsou ve formě map přiloženy v tištěné podobě. Výsledky jsou prezentovány na informačním posteru a na webových stránkách bakalářské práce.

## **KLÍČOVÁ SLOVA**

disagregace; pomocná data; srovnání; vizualizace

Počet stran práce: 54

Počet příloh: 20 (z toho 17 vázaných a 3 volné)

## **ANOTATION**

This thesis is focused on testing various data disaggregation methods in order to disaggregate selected statistical indicators to smaller spatial units. The main goal is to recalculate selected statistical indicators – population, unemployment rate and average gross wage in several levels of detail and to evaluate the results on the reference data available at the detailed level. For each statistical indicator, this thesis recommends convenient ancillary data, their weights and appropriate cell size for disaggregation.

Statistical data disaggregation was examined for two testing areas. Several types of ancillary data were used for each statistical indicator within the testing areas.

The results of the work are the visualisations of all example of the disaggregated data. These visualisations are available in digital form in the ArcGIS for Desktop project as a part of the DVD appendix. Other results are the evaluation and comparison of the accuracy of data disaggregation using several types of ancillary data. Selected examples of visualisations are provided as printed maps. The results of the thesis are presented on the poster and the website of the thesis.

## **KEYWORDS**

disaggregation; ancillary data; comparison; visualisation

Number of pages: 54

Number of appendixes: 20

**Prohlašuji, že**

- bakalářskou práci včetně příloh, jsem vypracoval(a) samostatně a uvedl(a) jsem všechny použité podklady a literaturu,

- jsem si vědom(a), že na moji bakalářskou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevýdělečně, ke své vnitřní potřebě, bakalářskou práci užívat (§ 35 odst. 3),

- souhlasím, aby jeden výtisk bakalářské práce byl uložen v Knihovně UP k prezenčnímu nahlédnutí,

- souhlasím, že údaje o mé bakalářské práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé bakalářské práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé bakalářské práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne

Jan Zapletal

---

podpis autora



Děkuji vedoucímu práce RNDr. Jaroslavu BURIANOVÍ, Ph.D. za cenné podněty, rady a připomínky během vypracovávání práce. Dále děkuji mé rodině a přátelům za podporu po celou dobu mého studia.

Za poskytnutá data děkuji Katedře geoinformatiky.

UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta

Akademický rok: 2017/2018

## ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Jan ZAPLETAL**

Osobní číslo: **R16413**

Studijní program: **B1301 Geografie**

Studijní obor: **Geoinformatika a geografie**

Název tématu: **Testování disagregačních metod v GIS**

Zadávající katedra: **Katedra geoinformatiky**

### Z á s a d y p r o v y p r a c o v á n í :

Cílem diplomové práce je otestovat vybrané disagregační metody za účelem přepočtu vybraných statistických ukazatelů do nižších územních jednotek. V rešeršní části práce student nastuduje a zhodnotí vybrané disagregační metody (např. disagregace, areálová interpolace, dasymetrické mapování) s důrazem na existující programová řešení v prostředí GIS. V praktické části práce aplikuje vybrané metody na několik statistických ukazatelů na několika úrovních podrobnosti (např. evropská, národní, regionální, lokální) a zhodnotí přesnost použitých metod na referenčních datech (ta která jsou dostupná na detailní úrovni). Výsledky budou vhodným způsobem vizualizovány a interpretovány.

Student vyplní údaje o všech datových sadách, které vytvořil nebo získal v rámci práce, do Metainformačního systému katedry geoinformatiky a současně vytvoří zálohu údajů ve formě validovaného XML souboru. Celá práce (text, přílohy, výstupy, zdrojová a vytvořená data, XML soubor) se odevzdá v digitální podobě na CD (DVD) a text práce s vybranými přílohami bude odevzdán ve dvou svázaných výtiscích na sekretariát katedry. O diplomové práci student vytvoří webovou stránku v souladu s pravidly dostupnými na stránkách katedry. Práce bude zpracována podle zásad dle Voženílek (2002) a závazné šablony pro diplomové práce na KGI.

Rozsah grafických prací: **dle potřeby**  
Rozsah pracovní zprávy: **max. 50 stran**  
Forma zpracování bakalářské práce: **tištěná**  
Seznam odborné literatury: **viz příloha**

Vedoucí bakalářské práce: **RNDr. Jaroslav Burian, Ph.D.**  
Katedra geoinformatiky

Datum zadání bakalářské práce: **11. května 2018**

Termín odevzdání bakalářské práce: **13. května 2019**

doc. RNDr. Martin Kubala, Ph.D.  
děkan

L.S.

UNIVERZITA PALACKÉHO V OLOMOUCI  
PŘÍRODOVĚDECKÁ FAKULTA  
KATEDRA GEONFORMATIKY  
17. listopadu 50, 771 46 Olomouc

prof. RNDr. Vít Voženílek, CSc.  
vedoucí katedry

V Olomouci dne 15. května 2018

# Příloha zadání bakalářské práce

Seznam odborné literatury:

Šablona KGI pro psaní bakalářských a magisterských prací

VOŽENÍLEK, V.: Diplomové práce z geoinformatiky, Vydavatelství UP, Olomouc, 2002, 60 s.

Areálová interpolace v ArcGIS

(<http://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/what-is-areal-interpolation.htm>)

STARÝ, Vlastimil. Prostorová disagregace s využitím dat LU/LC [online]. VŠB - TU

Ostrava, 2012 [cit. 2017-08-01]. Dostupné z WWW:

[http://gisak.vsb.cz/GISacek/GISacek\\_2012/papers/stary.pdf](http://gisak.vsb.cz/GISacek/GISacek_2012/papers/stary.pdf)

Zajícová, Hana. 2017. "Srovnání Prostorové Reprezentace Adresních Bodů - Grid/areály." Univerzita Palackého v Olomouci.

Zandbergen, Paul A. 2011. "Dasymetric Mapping Using High Resolution Address Point Datasets." *Transactions in GIS* 15 (SUPPL. 1): 5-27.

doi:10.1111/j.1467-9671.2011.01270.x.

Bielecka, Elzbieta. 2005. "A Dasymetric Population Density Map of Poland." *Proceedings of the 22nd International Cartographic Conference*, no. 48 22: 9-15.

Nagle, Nicholas N., Barbara P. Battenfield, Stefan Leyk, and Seth Spielman.

2014. "Dasymetric Modeling and Uncertainty." *Annals of the Association of American Geographers* 104 (1): 80-95. doi:10.1080/00045608.2013.843439.

Zandbergen, Paul A., and Drew A. Ignizio. 2010. "Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates." *Cartography and Geographic Information Science* 37 (3): 199-214.

doi:10.1559/152304010792194985.

Holt, James B., C.P. Lo, and Thomas W. Hodler. 2004. "Dasymetric Estimation of Population Density and Areal Interpolation of Census Data." *Cartography and Geographic Information Science* 31 (2): 103-21.

doi:10.1559/1523040041649407.

Holt, James B., and Hua Lu. 2011. "Dasymetric Mapping for Population and Sociodemographic Data Redistribution." In *Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment*, 195-210.

doi:10.1002/9780470979563.ch14.

Langford, Mitchel. 2006. "Obtaining Population Estimates in Non-Census Reporting Zones: An Evaluation of the 3-Class Dasymetric Method." *Computers, Environment and Urban Systems* 30 (2): 161-80.

doi:10.1016/j.compenvurbsys.2004.07.001.

Gallego, Francisco Javier. 2010. "A Population Density Grid of the European Union." *Population and Environment* 31 (6): 460-73.

doi:10.1007/s11111-010-0108-y.

Batista e Silva, Filipe, Javier Gallego, and Carlo Lavalle. 2013. "A

High-Resolution Population Grid Map for Europe." *Journal of Maps* 9 (1): 16-28. doi:10.1080/17445647.2013.764830.

Li, Tiebei, David Pullar, Jonathan Corcoran, and Robert Stimson. 2007. "A Comparison of Spatial Disaggregation Techniques as Applied to Population Estimation for South East Queensland (SEQ), Australia." *Applied GIS* 3 (9).

doi:10.4225/03/57E9AECEBA789.

Huang, Zhengdong, Henk Ottens, and Ian Masser. 2007. "A Doubly Weighted Approach to Urban Data Disaggregation in GIS: A Case Study of Wuhan, China." *Transactions in GIS* 11 (2): 197-211.

# OBSAH

<b>SEZNAM POUŽITÝCH ZKRATEK .....</b>	<b>10</b>
<b>ÚVOD .....</b>	<b>11</b>
<b>1 CÍLE PRÁCE.....</b>	<b>12</b>
<b>2 METODY A POSTUP ZPRACOVÁNÍ.....</b>	<b>13</b>
2.1 Použité metody .....	13
2.2 Použitý software a data .....	14
2.3 Postup zpracování.....	15
<b>3 REŠERŠE STUDOVANÉ LITERATURY .....</b>	<b>17</b>
3.1 Metoda disagregace.....	17
3.2 Možnosti software .....	23
3.2.1 Možnosti přepočtu dat.....	24
3.2.2 Možnosti importu dat z RÚIANu .....	26
<b>4 PŘÍPRAVA A ZPRACOVÁNÍ DAT .....</b>	<b>28</b>
4.1 Výběr statistických ukazatelů .....	28
4.2 Testovací území .....	29
4.3 Výběr pomocných dat .....	29
4.3.1 Počet obyvatel .....	30
4.3.2 Průměrná hrubá mzda .....	32
4.3.3 Podíl nezaměstnaných osob.....	33
<b>5 PŘEPOČET DAT .....</b>	<b>34</b>
5.1 Průběh zpracování.....	34
5.1.1 Počet obyvatel .....	34
5.1.2 Průměrná hrubá mzda .....	38
5.1.3 Podíl nezaměstnaných osob.....	39
5.2 Zhodnocení metod .....	40
5.2.1 Výběr velikosti buňky .....	40
5.2.2 Zhodnocení přesnosti .....	42
5.2.3 Vizualizace.....	48
<b>6 VÝSLEDKY .....</b>	<b>50</b>
6.1 Srovnání metod .....	50
6.2 Vizualizace .....	51
<b>7 DISKUZE .....</b>	<b>52</b>
<b>8 ZÁVĚR .....</b>	<b>54</b>
<b>POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE</b>	
<b>PŘÍLOHY</b>	

## SEZNAM POUŽITÝCH ZKRATEK

<b>Zkratka</b>	<b>Význam</b>
ČSÚ	Český statistický úřad
ESRI	Environmental System Research Institute
GIS	geografický informační systém
LPIS	Land Parcel Identification System
RÚIAN	Registr územní identifikace, adres a nemovitostí
SHP	Shapefile
SLDB	Sčítání lidu, domů a bytů
VBA	Visual Basic for Applications
VFR	Výměnný formát RÚIAN
VVP	vojenský výcvikový prostor
XML	eXtensible Markup Language
ZSJ	základní sídelní jednotka

## ÚVOD

V současné době je k dispozici velké množství bodových či polygonových statistických dat v různých úrovních podrobnosti. Tato data jsou použitelná pro různé populační analýzy v daném území. V některých případech je však pro analýzy potřeba mít data v územních jednotkách, pro které data buď neexistují nebo jsou neveřejná či finančně nedostupná. V tomhle případě lze potřebná data získat jinými způsoby. Prvním možným způsobem je vytvoření vlastních primárních dat za námi zvolené území. Tato možnost však vyžaduje velkou časovou a finanční investici. Pomocí dotazníkových šetření se většinou podaří získat jen vzorek údajů o obyvatelstvu, který může být více či méně reprezentativní. Druhou a zároveň dostupnější metodou je získání potřebných dat z již existujících sekundárních dat pomocí přepočtu statistických informací do námi zvolených územních jednotek.

Podle Klaudy (2016) lze přiřazovat statistické informace do nově zvolených územních jednotek dvěma způsoby. Prvním z nich je agregace detailnějších statistických dat do vyšších územních jednotek (bottom-up přístup). Při agregaci dochází k sečtení všech bodových nebo polygonových dat v rámci vyššího územního celku. Druhým možným způsobem je metoda disagregace (top-down přístup). Tato metoda převádí statistické informace z vyšších územních celků do nižších, popřípadě do odlišného typu územních jednotek. Na rozdíl od agregace není disagregace jednoduchým součtem hodnot, pro přepočet je nutné modelovat rozložení statistického jevu v daném území. Přesnost výpočtu disagregace se pak odvíjí od přesnosti vytvořeného modelu.



# 1 CÍLE PRÁCE

Cílem bakalářské práce je provést testování vybraných disagregačních metod za účelem přepočtu vybraných statistických ukazatelů do nižších územních jednotek. Jako testované statistické ukazatele byl vybrán počet obyvatel, dále hustota zalidnění a průměrná hrubá mzda. Následně bude navržen postup přepočtu statistických dat do nižších územních jednotek a do gridové sítě. Pro automatizaci celého procesu bude vytvořen nástroj v prostředí GIS. Na testovacím území Olomouckého kraje, popřípadě území České republiky, bude vybráno několik druhů pomocných dat pro disagregaci a zhodnocena přesnost přepočtu na referenčních datech. V rámci testovacího území bude vybráno několik úrovní podrobnosti na základě dostupnosti zdrojových a referenčních dat. Na závěr bude pro zvolená testovací území proveden přepočet vybraných témat. Výsledky pak budou vhodným způsobem vizualizovány a interpretovány.

Výsledky práce doporučí nejvhodnější metodu pro přepočet vybraných statistických ukazatelů do nižších územních jednotek v rámci testovaných území.



## 2 METODY A POSTUP ZPRACOVÁNÍ

Pro řešení práce byla použita zejména analytická metoda disagregace, pomocí níž byla statistická data přepočítávána z vyšších územních jednotek do nižších a do hexagonální sítě. Výsledné disagregované údaje byly poté vizualizovány konvenčními vizualizačními metodami tematické kartografie. Pro vizualizaci hustoty zalidnění byla použita metoda kartogramu a metoda gridu, pro ostatní statistické ukazatele poté areálová metoda. Areálová metoda je podle Zajícové (2017) jedním z nepoužívanějších způsobů prezentace statistických dat.

### 2.1 Použité metody

#### **Areálová metoda (metoda plošných znaků)**

Dle Voženílek a Kaňok (2011) je metoda plošných znaků jednou z nejčastěji využívaných metod tematické kartografie. Plošný znak se používá pro znázornění jevů jako samostatný vyjadřovací prostředek v areálové metodě nebo jako součást složitějších vizualizačních metod (kartogram, kartografická anamorfóza, dasymetrická metoda, tečková metoda aj.). Areálová metoda bývá využita pro vyjádření kvalitativních i kvantitativních jevů.

Podle autorů využívá metoda plošných znaků pro vyjádření jevů pouze dva parametry plošných znaků – výplň a obrys. Výplň se rozumí barevné nebo rastrové zaplnění plochy areálu sloužící k vyjádření kvalitativní nebo kvantitativní vlastnosti znázorňovaného jevu. Podle autorů také platí, že shodné (popř. podobné) jevy se znázorňují shodnými (podobnými) barvami, odlišné jevy poté odlišnými barvami. Při znázorňování kvantitativního jevu pomocí kvantitativní stupnice barev platí, že čím vyšší intenzita jevu, tím intenzivnější barva. Obrys areálu je méně výrazným parametrem pro vizualizaci, může znázorňovat míru jistoty výskytu jevu nebo znázornění nadřazenosti areálů (Voženílek a Kaňok, 2011).

#### **Metoda kartogramu**

Metoda kartogramu je speciální typ areálové metody sloužící k vizualizaci kvantitativních dat. Voženílek a Kaňok (2011) popisují metodu kartogramu jako metodu znázornění kvantitativního jevu relativními hodnotami tak, aby byly dílčí územní celky srovnatelné. Při tvorbě kartogramu musí být kvantitativní data přepočtena na jednotku plochy dílčího územního celku. Dle autorů existuje mnoho různých druhů kartogramů, nejčastěji bývá použit homogenní jednoduchý kartogram, kdy je celek pokryt odstíny barvy podle předem sestavené objektivní stupnice. Pokud nemá smysl přepočítávat sledovaný kvantitativní jev na jednotku plochy, lze pro vizualizaci využít metody pseudokartogramu. Pseudokartogramy lze používat pouze v případě, že dílčí územní celky mají přibližně stejné velikosti ploch, na které se vyjadřovaný kvantitativní jev přepočítává (Voženílek a Kaňok, 2011).

#### **Metoda gridu**

Metoda gridu je dalším speciálním typem areálové metody sloužící k vizualizaci kvantitativních i kvalitativních dat. Pod pojmem grid se rozumí pravidelná polygonová síť složená ze stejně velkých pravidelných útvarů, nejčastěji čtverců, hexagonů nebo trojúhelníků. Území je rozděleno pomocí mřížky na stejně velké územní celky – buňky, které mají jednoznačný identifikátor a k nimž jsou přiřazeny statistické údaje (Klauda, 2016). Podle Huanga a kol. (2007) je možné statistické údaje do odlišných územních jednotek přiřazovat metodou agregace, disagregace nebo interpolací.

Podle Klaudy (2016) patří mezi hlavní výhody gridové sítě nezávislost na územně a časově proměnlivé struktuře administrativních jednotek. Tato nezávislost umožňuje snadno porovnávat statistické údaje v různých časových obdobích, nezávisle na změnách administrativního členění. Další zmíněnou výhodou je vzájemná srovnatelnost všech buněk v gridu, což je zajištěno pravidelností polygonové sítě. Veškeré údaje jsou v gridu anonymizovány a generalizovány.

### **Metoda disagregace**

Analytická metoda disagregace je metoda přepočtu statistických informací z vyšších územních jednotek do nižších, popřípadě jiných územních celků. Pokud je potřeba zjistit neznámou hodnotu statistického ukazatele pro dílčí část územní jednotky, lze využít metody prostorové disagregace k odhadu neznámé hodnoty na základě známé hodnoty z vyššího územního celku (Starý, 2012). Přenos dat v procesu disagregace je složitý z důvodu neshodných hranic zdrojových a cílových oblastí a z důvodu heterogenity hustoty výskytu daného jevu (Li a kol., 2007). Protože jev ve sledovaném území není konstantní, je vhodné použít pro disagregaci pomocná data znázorňující rozložení jevu v daném území. Výsledkem disagregace je pak polygonová, rastrová nebo gridová vrstva s nově odvozenými hodnotami.

### **Metoda agregace**

Analytická metoda agregace je metoda převádějící statistické informace z nižších územních jednotek do vyšších. Na vstupu do agregace se nachází detailnější bodová nebo polygonová data, při agregaci pak dochází k sečtení statistických údajů ve všech dílčích částech územního celku. Metoda agregace se používá při převodu informací z bodů na polygony, při anonymizaci, generalizaci nebo vizualizaci statistických dat.

Podle Klaudy (2016) lze metodu agregace použít samostatně nebo v kombinaci s metodou disagregace při přepočtu statistických dat do gridové sítě. Agregace může být také závěrečnou fází metody disagregace, kdy jsou statistická data z vyšších územních jednotek nejprve disagregována do detailní gridové sítě, z níž jsou pak zpětně agregována do územních jednotek.

## **2.2 Použitý software a data**

Příprava a zpracování dat, samotná disagregace statistických dat a následná vizualizace přepočtených údajů proběhla v programu ArcGIS for Desktop ve verzi 10.4.1. Pro tvorbu hexagonální sítě byl využit skript Generate Tesselation. Tento skript je součástí nástrojů programu ArcMap od verze 10.4 a nachází se v Data Management toolboxu. V prostředí ArcGIS Pro byly do tohoto nástroje přidány navíc dvě nové možnosti tvorby pravidelné polygonové sítě. Disagregace statistických údajů proběhla v prostředí ArcGIS for Desktop 10.4.1 jako posloupnost několika nástrojů. Z důvodu zachování posloupnosti nástrojů, zjednodušení a zautomatizování procesu disagregace s využitím pomocných dat byl v rozhraní ModelBuilder vytvořen nástroj pro automatický přepočet statistických dat. Při disagregaci byl také využit nástroj Areal Interpolation. Tento nástroj je součástí balíčku Geostatistical Wizard a je dostupný pod licenci Geostatistical Analyst. Pro zpracování dat, tvorbu tabulek a grafů byl použit tabulkový software Microsoft Excel 2016. Pro stahování dat z RÚIANu byl použit nástroj QGIS RUIAN Plugin 1.0.0 z laboratoře GeoForAll na Fakultě stavební ČVUT v Praze. Nástroj funguje jako plugin do programu QGIS verze 2.18 a umožňuje stahovat aktuální data z RÚIANu pro větší množství obcí najednou. Pro statistické zpracování byl použit program GeoDa verze 1.12 umožňující přes uživatelské rozhraní statisticky analyzovat prostorová data. Pro grafické zpracování posteru byl využit program Adobe Illustrator CS6.

Jako referenční data pro disagregaci počtu obyvatel byla použita zdůvěrněná statistická data ze SLDB 2011 za ZSJ České republiky poskytnutá Českým statistickým úřadem. Pro testování přesnosti disagregačních technik byla převzata data o počtu obyvatel s obvyklým pobytem ze SLDB 2011 z Veřejné databáze ČSÚ. Jako pomocná data pro disagregaci byla vzata data z OpenStreetMap za Českou republiku, platná k 5. 11. 2018, dále vektorová data CORINE Land Cover 2018, data Open Land-Use Map za Olomoucký kraj ze dne 28. 3. 2019. Z databáze RÚIAN za Olomoucký kraj byly využity vrstvy adresních bodů a budov, platné ke dni 1. 4. 2019. V práci byla dále využita vrstva adresních bodů „Statistické budovy (vchody k bytům)“, poskytnutý Katedře geoinformatiky od ČSÚ. Data jsou platná k 1. dubnu 2016 a obsahují informace například o počtu podlaží, počtu bytů v budově nebo způsobu využití budovy. Jako administrativní hranice byla použita datová sada ArcČR® 500 verze 3.3. Pro disagregaci průměrné mzdy a podílu nezaměstnaných osob byla použita data z Veřejné databáze ČSÚ.

## 2.3 Postup zpracování

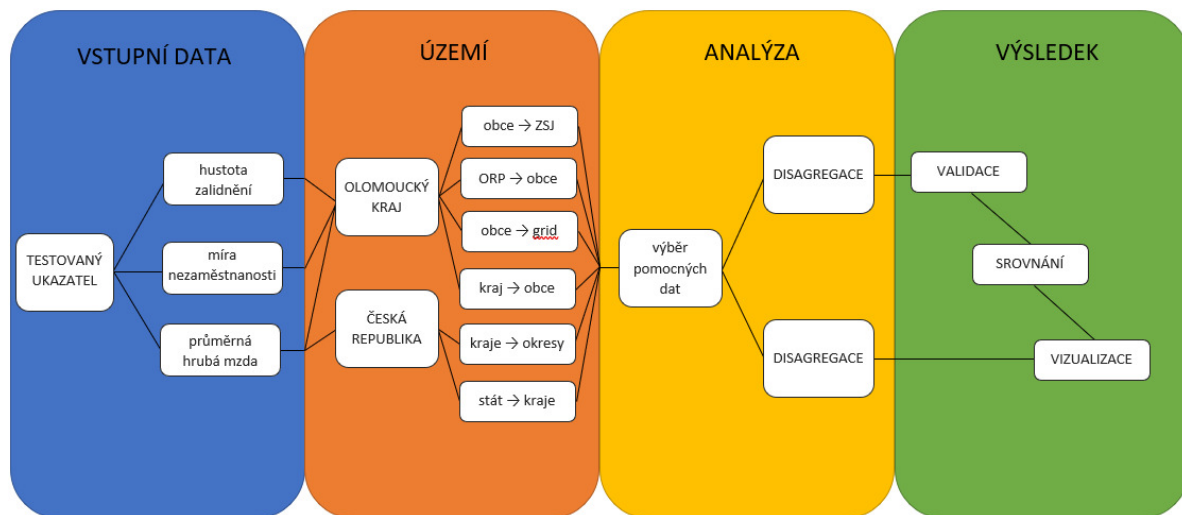
Samotnému vlastnímu zpracování předcházelo studium řešené problematiky. Byla nastudována teorie disagregačních metod a konkrétní příklady jejich využití. Během rešerše byla nastudována některá dosavadní zpracování tématu disagregací v České republice i v zahraničí. Dále byla nastudována tematika pomocných dat pro disagregaci, jejich vhodné varianty a jejich dostupnost. Také byly nastudovány možnosti vizualizace přepočtených statistických dat do gridové sítě. Schéma postupu práce (Obr. 2.3.1) znázorňuje samotný průběh práce.

Vzhledem k zadání práce nebylo testování disagregačních metod prováděno pouze na úrovni státního území České republiky, ale i zvláště pro Olomoucký kraj a jeho dílčí územní jednotky. Nejprve byly vybrány tři statistické ukazatele, na nichž testování proběhne. Pro každý statistický ukazatel byla pak na základě konzultace s vedoucím práce vybrána pomocná data použitá při přepočtu. Při disagregaci statistických ukazatelů byl nejprve otestován již existující nástroj v programu ArcMap a poté byl vytvořen vlastní postup přepočtu statistických údajů. Pro každý statistický údaj bylo vybráno testovací území s dostupnými daty a referenční území sloužící pro validaci. Samotný výběr ukazatelů a postup práce je popsán v kapitole 4.

Byly vybrány tři různé statistické ukazatele, pro něž byl proveden přepočet do nižších nebo odlišných územních jednotek prostřednictvím pomocných dat. Statistické údaje pro přepočet a validaci výsledků pochází ze SLDB 2011 a byly autorovi poskytnuty Katedrou geoinformatiky v podobě tabulek ve formátu List Microsoft Excelu (.xlsx). Pomocná data pro přepočet byla získána z volně dostupných zdrojů nebo byla poskytnuta Katedrou geoinformatiky ve formátu ESRI Shapefile (.shp).

Výsledkem práce je 15 tištěných mapových výstupů, obsahujících konkrétní vizualizace přepočtených statistických údajů. Dále pak 156 vizualizací přepočtených dat, 8 tabulek obsahujících provedené přepočty a hodnocení výsledků jednotlivých disagregačních metod. Nedílnou součástí jsou také tři nástroje pro disagregaci vytvořené v rozhraní ModelBuilder. Netištěné výstupy jsou dostupné v mapovém projektu přílohy číslo 18.

Na závěr byla sepsána textová část práce, byl zhotoven informační poster a byly vytvořeny webové stránky prezentující výsledky práce. Součástí práce jsou také tabulky, obrázky a grafy.



Obr. 2.3.1 Grafické schéma postupu práce

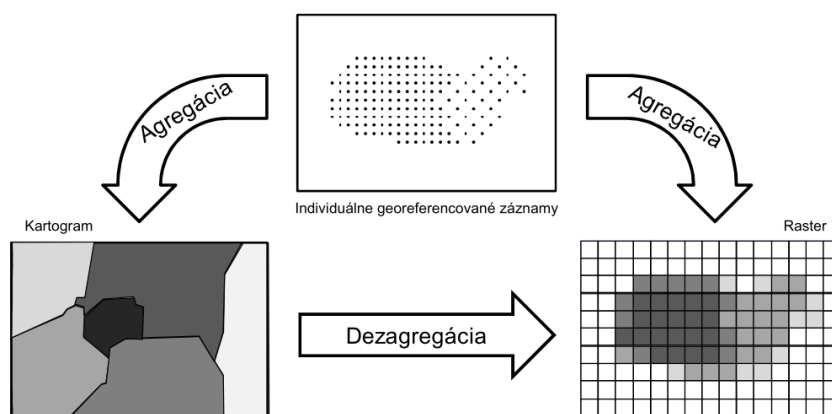
### 3 REŠERŠE STUDOVANÉ LITERATURY

Celkovému řešení práce předcházelo podrobné studium dostupné literatury, vyhledávání použitelných metod, nástrojů, postupů a softwarových řešení aplikovatelných na problém disagregace. Pomocí důkladné rešerše byly vybrány vhodné nástroje a metody použité při přepočtu statistických dat.

#### 3.1 Metoda disagregace

V současné době je k dispozici velké množství bodových či polygonových sekundárních dat použitelných pro různé populační analýzy v daném území. Občas jsou však již existující data buď v nevyhovujících územních jednotkách, v nevyhovujícím formátu, příliš drahá nebo nedostupná. V tomhle případě lze potřebná data získat jinými způsoby. Prvním možným způsobem je vytvoření vlastních primárních dat za námi zvolené území. Tato možnost je však časově a finančně velmi náročná. Druhou, dostupnější metodou je získání potřebných dat z již existujících sekundárních dat pomocí přepočtu statistických informací do námi zvolených územních jednotek.

Podle Klauudy (2016) lze přiřazovat statistické informace do námi zvolených územních jednotek dvěma způsoby. Prvním z nich je agregace detailnějších statistických dat do námi zvoleného území (bottom-up přístup). Na vstupu se nachází detailní bodová nebo polygonová data, při agregaci pak dochází k sečtení statistických údajů ve všech dílčích částech územního celku. Metoda agregace se používá při převodu informací z bodů na polygony, při anonymizaci a generalizaci statistických dat. Druhým možným způsobem je metoda disagregace (top-down přístup). Tato metoda převádí statistické informace z vyšších územních celků do nižších celků, do gridu nebo do odlišného typu územních jednotek. Metoda agregace však, podobně jako generalizace nebo převzorkování, není reversibilní, nemůžeme tedy z agregovaných dat získat zpět původní vstupní hodnoty. Pro disagregaci je proto vhodné použít pomocná data umožňující geoprocesoru odhadnout rozložení jevu v daném území. Přesnost výsledků disagregace závisí na vhodnosti použitých pomocných dat a nastavení parametrů statistického přepočtu. Výsledkem disagregace je pak polygonová nebo gridová vrstva s novými odvozenými hodnotami. Obecně platí, že součet statistických informací ve všech dílčích celcích by měl odpovídat hodnotě statistického údaje ve vyšším celku. Jako pomocná data pro disagregaci statistických ukazatelů lze použít například data o krajinném pokryvu, rozlohu zastavěného území, hustotu silniční sítě nebo přesně lokalizované adresní body. Metoda disagregace se používá pro odhad chybějících statistických údajů v dílčích územních celcích.



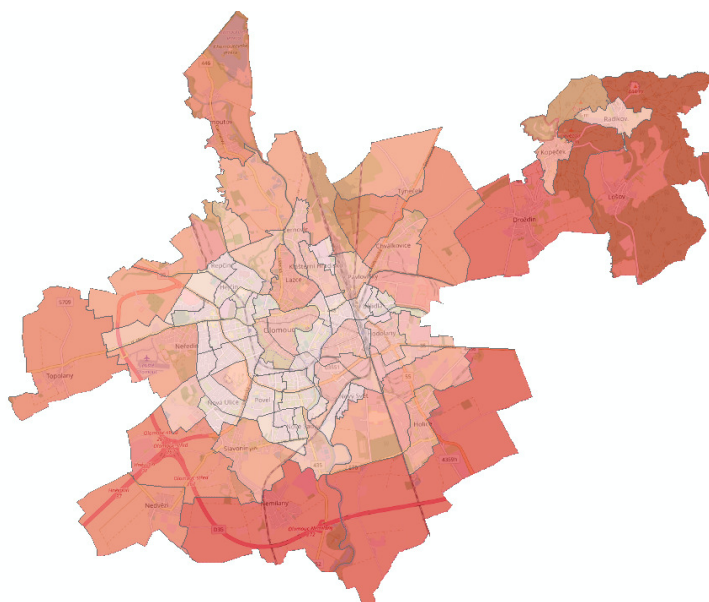
Obr. 3.1.1 Schéma metody agregace a disagregace (autor: Rosina a Hurbánek, 2012)

### Disagregace bez využití pomocných dat

Přenos statistických informací z vyšších územních jednotek do nižších nebo z jednoho typu územních jednotek (zdrojová oblast) do druhého (cílová oblast) je složitý proces. Protože většina zdrojových dat je zgeneralizovaná na určité územní jednotky a neznáme přesné rozložení statistického jevu v námi zvoleném území, nemůžeme přesně určit hodnoty statistického jevu pro námi zvolené územní jednotky.

Areálová interpolace je proces přenosu prostorových dat z jednoho typu územních jednotek na druhý a obvykle bývá použita pro porovnávání více datasetů vytvořených v různých územních jednotkách (Eicher a Brewer, 2001). Areálová interpolace zahrnuje různé metody, které většinou používají vážení statistických údajů, založené na oblasti průniku mezi zdrojovou a cílovou oblastí, sloužící k rozložení charakteristik v území (Zandbergen, 2011). Na vstupu do areálové interpolace se nachází územní jednotky se statistickými informacemi a pomocná data pro výpočet. Pomocná data slouží k přesnějšímu určení rozložení statistického jevu v daném území a určují složitost samotného výpočtu areálové interpolace.

Nejjednodušší metodou areálové interpolace je areálové vážení (areal weighting). Tato metoda pro svůj výpočet nepotřebuje žádná externí statistická data, jako pomocná data používá rozlohu jednotlivých dílčích částí celku. Podle Flowerdewa a kol. (1991) v případě, že statistickým údajem je počet, areálové vážení předpokládá, že daný jev je v prostoru rozložen rovnoměrně. Poměr hodnoty pro dílčí část a hodnoty v celku se rovná poměru rozlohy dílčí části a celku. V případě, že statistickým údajem je poměr (procentuální část), areálové vážení předpokládá, že hodnota pro dílčí část je stejná jako hodnota v celém území. Pokud cílová oblast zasahuje do více zdrojových oblastí, pak je hodnota jevu v cílové zóně rovna váženému průměru hodnot jevů těchto zdrojových oblastí, kde váhou je procentuální část rozlohy cílové zóny zabrané danou zdrojovou oblastí (Flowerdew a kol., 1991). Tato metoda je vhodná, pokud nemáme k dispozici žádné doplňující informace o cílových oblastech nebo je statistický jev relativně rovnoměrně rozložen v celém území. Pokud však jev není v území rovnoměrně rozložen, dochází při disagregaci k výrazným odchylkám a nepřesnostem. Na obr. 3.1.2 je znázorněný počet obyvatel v ZSJ Olomouce. Lze vidět, že areálové vážení přiřazuje největší počet obyvatel (tmavě červená) do největších ZSJ, ve kterých reálně bydlí výrazně méně lidí.



Obr. 3.1.2 Nedostatky areálového vážení znázorněné na počtu obyvatel Olomouce

Všechny metody používající externí pomocná data by měly přesností svého výsledku překonat areálové vážení (Zandbergen, 2011).

Dalším způsobem, jakým lze disagregovat statistické informace je využití interpolačních metod. Krivoruchko a kol. (2011) ve své práci testovali využití různých variant interpolační metody Kriging pro disagregaci statistických dat. Pro tři různé typy úloh byly využity tři různé druhy rozdělení – Gaussovo, binomické a Poissonovo. V rámci práce byl vytvořen nástroj *Areal Interpolation*, který byl zařazen do toolboxu Geostatistical Analyst v programu ArcMap. Tento nástroj je dostupný pouze s licencí Geostatistical Analyst. Náповěda nástroje navrhuje možné využití jednotlivých typů rozdělení. Nástroj pak umožňuje mnoho možností úprav nastavení variogramu. Výsledkem nástroje je zejména rastrová vrstva interpolovaných hodnot pokrývající celé území a vrstva zobrazující chyby odhadu. Po použití nástroje lze interpolovanou vrstvu převést do polygonů pomocí nástroje *Areal Interpolation Layer To Polygons*.

### **Disagregace s využitím pomocných dat**

Metoda disagregace má na svém vstupu hodnoty statistického jevu agregované za určité území, proto je při disagregování potřeba modelovat způsob, jakým rozložit danou statistickou veličinu v námi sledovaném území. Pro tyto účely se využívá tzv. pomocných dat (ancillary data). Pomocnými daty se rozumí datové vrstvy s informacemi o rozložení jiného než disagregovaného jevu v území. Vybraný jev jako pomocná informace pro disagregaci by měl mít souvislost s disagregovaným jevem (např. hustota zástavby pro disagregaci počtu obyvatel). Pomocná data mají za úkol zpřesňovat informace o rozložení jevu v daném území, proto je nutné, aby byla k dispozici na detailnější úrovni než disagregovaný jev. Přesnost výsledku disagregace pak bude záviset zejména na vhodnosti použitých pomocných dat, úrovně jejich přesnosti a detailu.

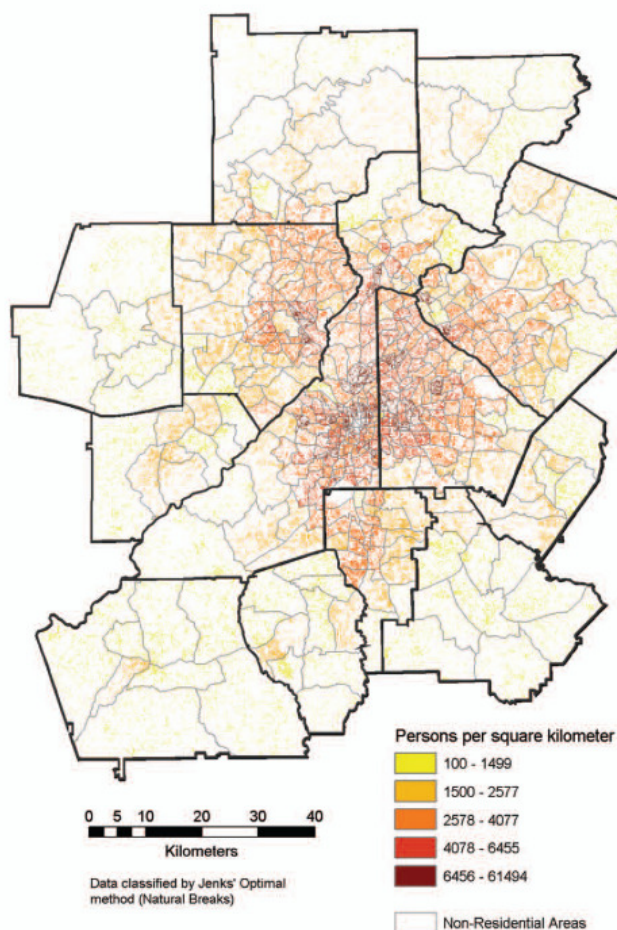
Prvním a nejčastěji využívaným typem pomocných dat pro zpřesňování výsledku disagregace jsou data o krajinném pokryvu nebo využití půdy. Tato data rozčleňují sledované území do jednotlivých kategorií krajinného pokryvu pomocí manuálně nebo automatickou klasifikací vytvořených polygonů využití území. Velké využití tato data nacházejí zejména při disagregaci počtu obyvatel, kdy je předpokládáno, že se obyvatelstvo koncentruje převážně do určitého typu využití území (zastavěná oblast). Výhodou použití dat o krajinném pokryvu jako pomocných dat je jejich dostupnost, pokrytí rozsáhlého území a jednoduchost jejich použití.

Starý (2012) ve své práci popisuje využití regresní analýzy pro disagregaci socioekonomických údajů pomocí dat krajinného pokryvu. Pro přepočítání statistických údajů byla využita data CORINE Land Cover a Urban Atlas Land Use, nesoucí informace o využití území v Evropě. Před samotnou analýzou bylo vypočítáno procentuální zastoupení jednotlivých typů krajinného pokryvu v obcích. Analýza pak zahrnovala srovnání statistického jevu s jednotlivými třídami krajinného pokryvu a vytvoření X-Y bodového grafu jejich závislosti. Pro zjištění míry závislosti statistických ukazatelů na vrstvách krajinného pokryvu byl vypočítán korelační koeficient. Bylo zjištěno, že významná korelace se vyskytuje pouze u hustoty zalidnění, pro tu pak byla vytvořena rovnice regresního modelu a každé třídě krajinného pokryvu byla přiřazena váha pro výpočet hustoty zalidnění.

Holt a kol. (2004) se ve své práci zabývali srovnáním hustoty zalidnění v okolí Atlanty v letech 1980, 1990 a 2000. Mezi těmito roky docházelo ke změnám hranic sčítacích okrsků, proto bylo nutné využít areálovou interpolaci k přepočtu dat na stejné sčítací okrsky. Jako pomocná data pro práci byly využity družicové snímky z družice Landsat 5 pro roky 1984, 1990 a 2000. Tyto snímky byly pomocí nekontrolované klasifikační metody



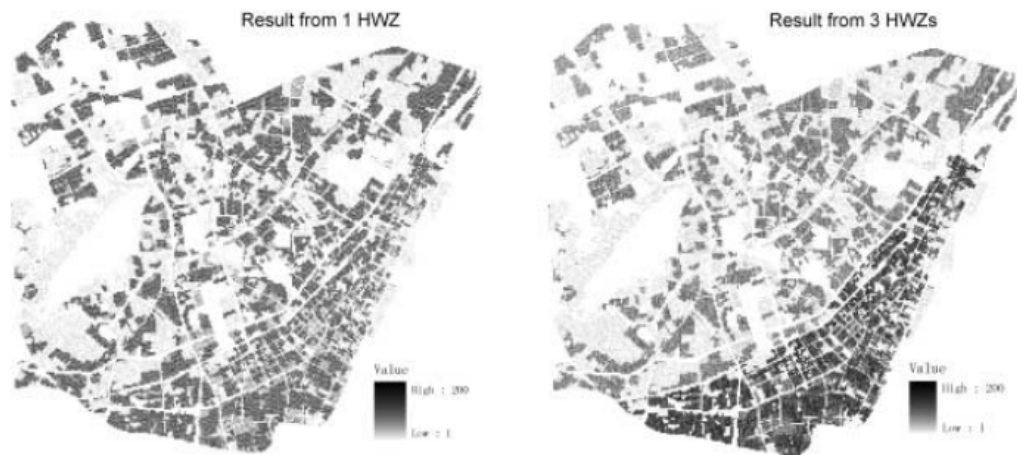
ISODATA převedeny na jednotlivé kategorie využití území. Pro samotnou disagregaci byla použita binární disagregační metoda. Ta spočívá v rozdělení dat o využití území na místa, v nichž se jev nachází (obytné plochy), a místa, v nichž se jev nenachází (ostatní plochy). Pro každou oblast ovlivněnou změnou hranic sčítacího okrsku bylo vypočítáno procentuální zastoupení obytných ploch v dané oblasti, stanovení vah založených na procentuálním zastoupení obytných zón a aplikování vah na data odvozená ze sčítání obyvatel v letech 1980 a 2000. Následně byly zjištěny změny v počtu obyvatel ve sčítacích okrscích. Závěrem byla vytvořena mapa (obr. 3.1.3) znázorňující dasymetrickou metodu hustotu zalidnění v okolí města Atlanta.



Obr. 3.1.3 Hustota zalidnění v oblasti Atlanty (Holt a kol., 2004)

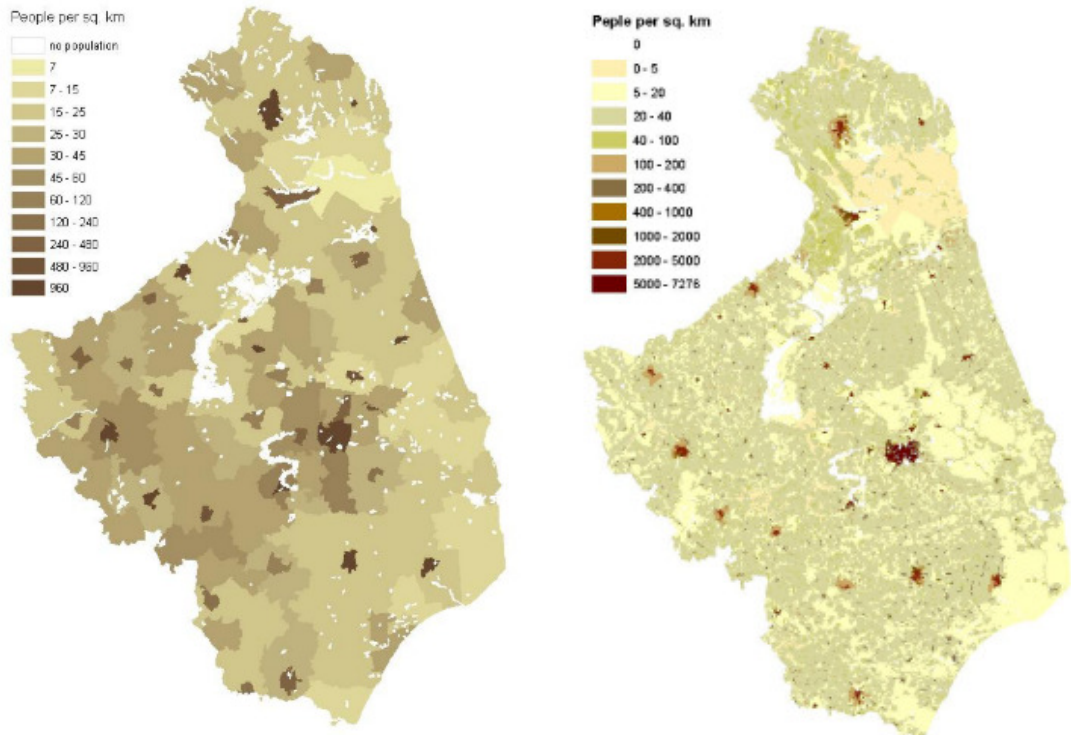
Huang a kol. (2007) ve své práci disagregovali počet obyvatel v části čínského města Wuhan pomocí dat o krajinném pokryvu. Při použití dat o využití území však měla každá kategorie krajinného pokryvu v celé řešené oblasti stejnou váhu. Je však rozdíl, jestli se jedná o zastavěnou oblast na okraji nebo v centru města. Proto byly vytvořeny ve studovaném území rozsáhlé oblasti pokrývající rozdílné regiony města, tzv. homogenní váhové zóny (HWZ). Hustota sledovaného statistického jevu v jedné HWZ je poté přibližně stejná. Poté byl na řešenou oblast aplikován princip dvojitěho vážení, kdy každé rastrové buňce ve studované oblasti byly přiřazeny váhy pro krajinný pokryv a pro homogenní zónu. Nejprve byla hustota zalidnění disagregována bez použití homogenní váhové zóny, poté s použitím tří vytvořených váhových zón (Obr. 3.1.4). Oba dva přístupy byly srovnány a bylo zjištěno, že použití tří homogenních váhových zón výrazně snížilo rozdíly mezi odhadnutou a skutečnou hodnotou hustoty zalidnění v řešené oblasti.





Obr. 3.1.4 Rozdíl mezi disagregací bez použití a s použitím HWZ (Huang a kol., 2004)

Bielecka (2005) se zabývala disagregací hustoty zalidnění v Podleském vojvodství ve východním Polsku. Jako pomocná data pro disagregaci byla použita data o krajinném pokryvu z volně dostupné datové sady CORINE Land Cover z roku 2000. Data o počtu obyvatel pak byla převzata z polského sčítání lidu v roce 2002. Na data byl nejprve aplikován binární přístup, poté byla využita metoda upraveného areálového vážení (Modifying areal weighting method), kdy byl použit regresní model pro zjištění vztahu mezi počtem obyvatel a typem krajinného pokryvu. Výsledkem bylo rozčlenění vrstvy CORINE Land Cover do šesti kategorií a rozčlenění komun do šesti tříd. Pro každou kombinaci komuny a krajinného pokryvu byl vypočítán váhový koeficient, ten byl pak dosazen do výsledné rovnice pro disagregaci. Bylo zjištěno, že rozčlenění komun a krajinného pokryvu do více kategorií přináší přesnější výsledky disagregace než binární metoda (Obr. 3.1.5).



Obr. 3.1.5 Rozdíl mezi binární metodou a upraveným areálovým vážením (Bielecka, 2005)

Data o krajinném pokryvu jsou sice snadno dostupná a jednoduchá na použití, nemusí však dokonale reprezentovat rozložení statistického jevu v území. Například zastavěná oblast v centru města, na sídlišti a okraji města se bude v hustotě zalidnění výrazně lišit. Pokud je potřeba dosáhnout ještě přesnějších výsledků a jsou k dispozici podrobná data o území, lze jako pomocná data využít jiné datové sady korespondující s disagregovaným jevem. Mezi použitelné datové sady patří zejména hustota silniční sítě, adresní body, hustota zástavby nebo intenzita nočního osvětlení. Problémem těchto datových sad však může být jejich obtížnější dostupnost.

Zandbergen a Ignizio (2010) porovnávali areálové vážení a využití krajinného pokryvu s dalšími datovými sadami ve čtyřech státech USA pro disagregaci počtu obyvatel. První datovou sadou je nepropustnost podloží odvozená z dat leteckého snímkování. Výhodou nepropustnosti je její větší detail a absence potřeby kalibrace vrstvy. Nevýhodou však je, že vysokou nepropustnost mají i silnice a průmyslové oblasti. Z původní datové sady nepropustnosti byly odvozeny tři další, aby došlo k odstranění nežádoucího zkreslení vlivem silnic a průmyslových zón. Další datovou sadou byla hustota silniční sítě. V hustě obydlených zónách se nachází velké množství silnic, hustota silniční sítě je zde větší. Problémem jsou zde však dálnice, na nichž nikdo nebydlí. Pro účely práce byly vytvořeny dvě kategorie – celková hustota silniční sítě a hustota místních silnic. Poslední datovou sadou byla intenzita nočního osvětlení. Po disagregaci byly nově vytvořené údaje porovnány se skutečnými údaji a vypočítáno procento špatně umístěných obyvatel. Pro každý stát byly pak techniky seřazeny od nejlepší po nejhorší. Bylo zjištěno, že pro každý stát je nejvhodnější jiná metoda. V součtu hodnocení (méně je lépe) za všechny státy však nejlépe vychází upravená nepropustnost podloží, po ní následuje krajinný pokryv a hustota místních silnic. Mezi nejhorší techniky pak patří areálové vážení a intenzita nočního osvětlení (Obr. 3.1.6).

Technique	Sum of Banks
Areal weighting	107
Land cover	39
Total imperviousness	61
Imperviousness < 60%	35
Imperviousness < 75%	31
Cleaned imperviousness	72
Total road density	61
Minor road density	43
Nighttime lights	85

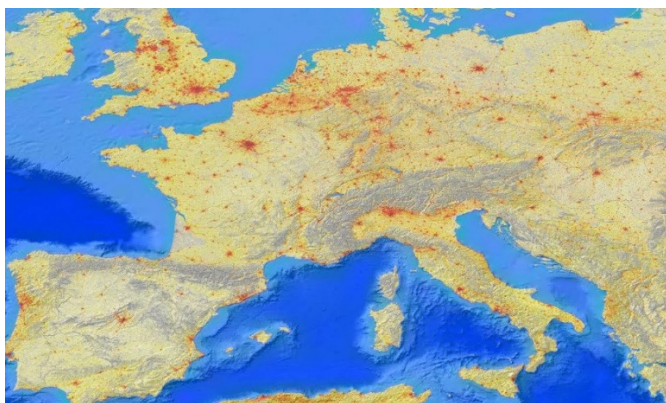
Obr. 3.1.6 Srovnání různých disagregačních technik (Zandbergen a Ignizio, 2010)

Zandbergen (2011) testoval přesnost disagregace za použití přesně lokalizovaných adresních bodů v porovnání s areálovým vážením, nepropustností povrchu, krajinným pokryvem, hustotou silniční sítě a nočním osvětlením. Přesnost každé metody byla určena porovnáním vypočítaných hodnot pro dané území se skutečnými hodnotami. Adresní body byly pro větší přesnost rozděleny na obytné a neobytné. Následně byl vypočítán průměrný počet obyvatel na jeden adresní bod a ten byl vynásoben počtem adresních bodů v každé lokalitě. Výsledkem výpočtu je odhad počtu obyvatel pro danou lokalitu. Na ostatní datové sady byl aplikován stejný postup, jako aplikovali Zandbergen a Ignizio (2010). Pro každou lokalitu pak byl pomocí známých hodnot vypočítán počet nepřesně umístěných lidí a jednotlivé metody byly navzájem srovnány. Bylo zjištěno, že využití adresních bodů představuje významné zpřesnění odhadu oproti ostatním technikám (Obr. 3.1.7).

Technique	People Placed Incorrectly		
	Count	Percentage	Rank
Areal weighting	182,363	21.2	12
Land cover (agriculture zero)	100,063	11.6	5
Land cover (agriculture not zero)	121,692	14.2	9
Total imperviousness	113,422	13.2	6
Imperviousness < 60%	92,410	10.8	3
Imperviousness < 75%	94,407	11.0	4
Cleaned imperviousness	148,900	17.3	10
Nighttime lights	160,151	18.6	11
Total road density	116,027	13.5	8
Minor road density	114,077	13.3	7
Address Points	41,992	4.9	2
Residential Address Points	36,523	4.2	1

Obr. 3.1.7 Srovnání různých disagregačních technik (Zandbergen, 2011)

V roce 1998 Národní laboratoře Oak Ridge (ORNL) poprvé představily globální populační datovou sadu LandScan™. LandScan™ je každoročně aktualizovaný celosvětový grid nesoucí informace o počtu obyvatel v rozlišení 30"×30". Pokrývá celou planetu od 84° s.z.š. do 90° j.z.š. Je dostupný ve formátu ESRI grid a ESRI binary raster, pro stažení je nutná registrace. LandScan™ byl vytvořen kombinací dat ze sčítacích okrsků v každém státě a pomocných dat, včetně dat o krajinném pokryvu, silničních sítích, sklonu svahu, zastavěných oblastech a analýzy leteckých snímků s vysokým rozlišením. V každém státě algoritmus spočítá pro každou buňku pravděpodobnostní koeficient a tyto koeficienty poté aplikuje na výsledky sčítání lidu v daném státě. Celková populace v daném státě je pak rozložena do buněk podle vah pravděpodobnostních koeficientů. Celá metodika zahrnuje závěrečnou manuální kontrolu. Po analýze leteckých snímků byly vyhodnoceny očividné anomálie a byl vytvořen nový rastr s opravenými koeficienty pro výpočet. Na obr. 3.1.8 lze vidět vizualizaci datové sady LandScan™ pro střední a jižní Evropu.



Obr. 3.1.8 Vizualizace datové sady LandScan™ 2013 (zdroj: <http://www.gisuser.com>)

### 3.2 Možnosti software

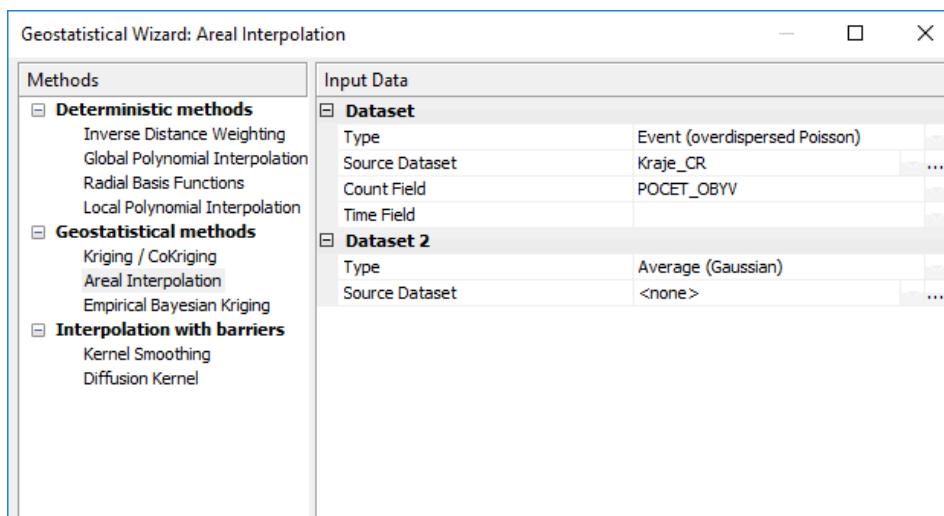
Před samotným řešením práce byly vyhledány již existující nástroje vhodné pro řešení problematiky disagregace a pro stažení pomocných dat z databáze RÚIAN. Byly vybrány některé nástroje, které byly testovány a následně případně použity pro potřeby práce.

## 3.2.1 Možnosti přepočtu dat

### Areal Interpolation

Nástroj Areal Interpolation je dostupný pod extenzí Geostatistical Analyst v programu ArcMap verze 10.3 a vyšší. Nástroj umožňuje provádět interpolaci geostatistickou metodou kriging pro data zprůměrovaná nebo agregovaná do polygonů (ESRI, 2016). Nástroj je dostupný pouze s licencí Geostatistical Analyst.

Pro použití nástroje je nejprve potřeba aktivovat extenzi Geostatistical Analyst a poté zapnout nástrojovou lištu Geostatistical Analyst. Po spuštění rozhraní Geostatistical Wizard a zvolení nástroje Areal interpolation se zobrazí okno s výběrem vstupních datasetů (obr. 3.2.1.1). Areal Interpolation umí pracovat s jedním datasetem, umí ale přidat i druhý dataset jako proměnnou pro cokriging. Oba datasety mohou mít shodné či úplně odlišné hranice. Po volbě datasetů je nutné vybrat vhodný typ dat. Nástroj nabízí k výběru tři druhy nastavení – zprůměrovaná data (Gaussovo rozdělení), data vyjadřující míru (binomické rozdělení) a data vyjadřující počet (Poissonovo rozdělení). Gaussovo rozdělení je vhodné pro spojitá data zprůměrovaná ve sledovaném území, například úroveň znečištění ovzduší nebo průměrná teplota. Binomické rozdělení je použitelné, pokud byla zdrojová data v daném území náhodně vybrána a byl spočítán počet případů splňujících určitá kritéria. Tento přístup lze použít například pro zjištění počtu lidí volících jednu stranu nebo pro zjištění míry nezaměstnanosti. Poissonovo rozdělení je použitelné pro interpolaci počtu jevů na daném území za určitý čas, například pozorování počtu velryb v oceánu. Pokud je údaj o čase v nástroji vynechán, předpokládá se, že doba pozorování je ve všech polygonech stejná. Toto nastavení je pak použitelné například pro interpolaci údaje o počtu obyvatel (ESRI, 2016).



Obr. 3.2.1.1 Rozhraní nástroje Areal Interpolation

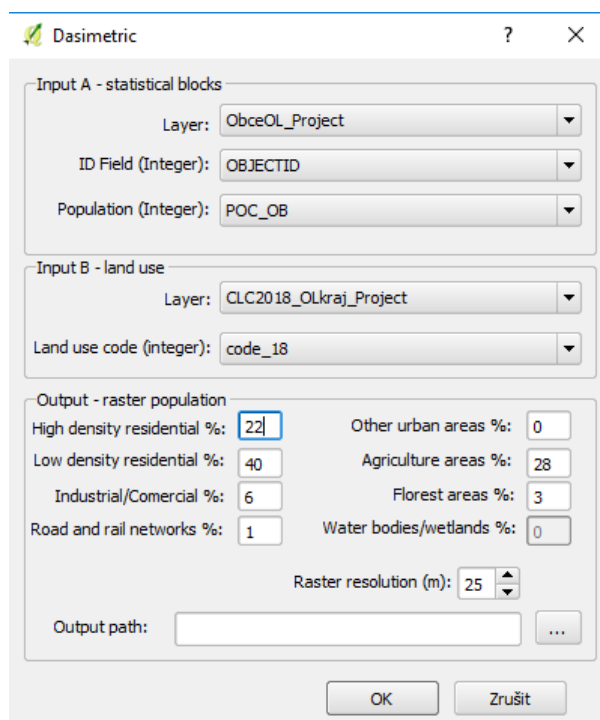
Po volbě vhodného rozdělení následuje úprava kovariance nebo semivariogramu tak, aby model přinášel co nejlepší výsledky. Na závěr dojde ke křížové validaci, kdy nástroj predikuje nové hodnoty a vyhodnocuje jejich chyby. Po dokončení procesu se vytvoří rastr s interpolovanými hodnotami. Tento rastr je pak pomocí nástroje Areal Interpolation Layer To Polygons převeden do cílových zón. Nástroj Areal Interpolation byl v rámci práce testován při disagregaci počtu obyvatel.

### QGIS Dasymetric Plugin

QGIS Dasymetric Plugin je volně dostupný plugin do programu QGIS 2.18 vytvořený na univerzitě v Lisabonu. Tento plugin, naprogramovaný v jazyce Python 2.7, disagreguje



počet obyvatel v libovolných územních jednotkách do rastrové mřížky využitím datové sady CORINE Land Cover (CLC) jako pomocných dat. Na vstupu jsou zdrojové územní jednotky s informací o počtu obyvatel, dále vektorová vrstva CLC a relativní hustoty obyvatel pro osm hlavních typů krajinného pokryvu (viz obr. 3.2.1.2). Nástroj převede vektorovou vrstvu CLC na rastr o stanovené velikosti pixelu, reklasifikuje rastr podle kódů CLC do osmi hlavních kategorií krajinného pokryvu. Těm pak přiřadí stanovené relativní hustoty. Nástroj převede informaci o počtu obyvatel na rastr a na základě relativních hustot pro každou buňku vypočítá předpokládaný počet obyvatel v každé buňce (Mileu a Queirós, 2018). Nástroj byl testován pro přepočítání dat o počtu obyvatel pomocí krajinného pokryvu, i přes různá nastavení nástroje se však nepodařilo získat použitelné výsledky, proto nebyl nástroj při práci použit.

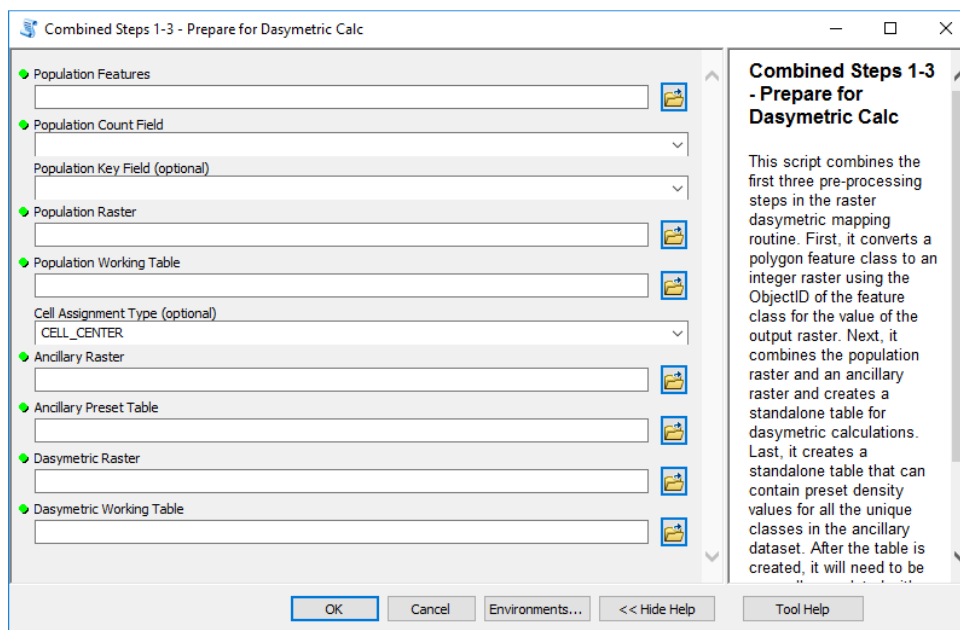


Obr. 3.2.1.2 Rozhraní nástroje QGIS Dasymetric Plugin

## IDM Toolbox

Intelligent dasymetric mapping (IDM) Toolbox je sada nástrojů pro ArcMap provádějící disagregaci počtu obyvatel podle váhového rastru z pomocných dat. Nástroj vytvořil Torrin Hultgren z United States Environmental Protection Agency (EPA, 2017). Nástroj je dostupný ve dvou variantách, první pro ArcGIS verze 10.2, druhá pro verzi 10.3 a vyšší. Pro instalaci je nutné stáhnout variantu nástroje pro odpovídající verzi softwaru ArcGIS a rozbalit obsah komprimované složky. Po rozbalení stačí jen v ArcToolboxu přidat IDM Toolbox jako nový Toolbox. Po přidání se nahraje Toolbox s osmi Python skripty.

Pro disagregaci počtu obyvatel je třeba do nástroje vložit zdrojové zóny s informací o počtu obyvatel a váhový rastr nesoucí podrobné informace o daném území (Obr. 3.2.1.3). Toolbox obsahuje osm Python skriptů. Přepočítání probíhá v sekvenci pěti kroků, během svého průběhu vytvoří nástroj tři tabulky a tři rastrové vrstvy. Výsledkem je rastrová vrstva o velikosti buňky váhového rastru, obsahující předpokládaný počet obyvatel v každé buňce. Přesnost odhadu poté záleží na zvoleném váhovém rastru. Nástroj je funkční a poskytuje použitelné výsledky, pro jeho složitost a tvorbu velkého množství nadbytečných dat nakonec nebyl v práci použit.



Obr. 3.2.1.3 Rozhraní nástroje IDM Toolbox

## 3.2.2 Možnosti importu dat z RÚIANu

### VFR Import

Nástroj VFR Import od společnosti ARCDATA PRAHA, s.r.o. slouží pro převod dat stažených z databáze RÚIAN do geodatabáze softwaru ArcGIS (ARCDATA PRAHA, 2019). Pro použití nástroje je nejprve potřeba stáhnout a nainstalovat aktuální verzi nástroje odpovídající nainstalované verzi programu ArcGIS. Poté je nutné v ArcToolboxu vytvořit nový uživatelský Toolbox, do něhož je třeba nainportovat nástroj VFR Import. Po úspěšném importu je možné začít nástroj používat.

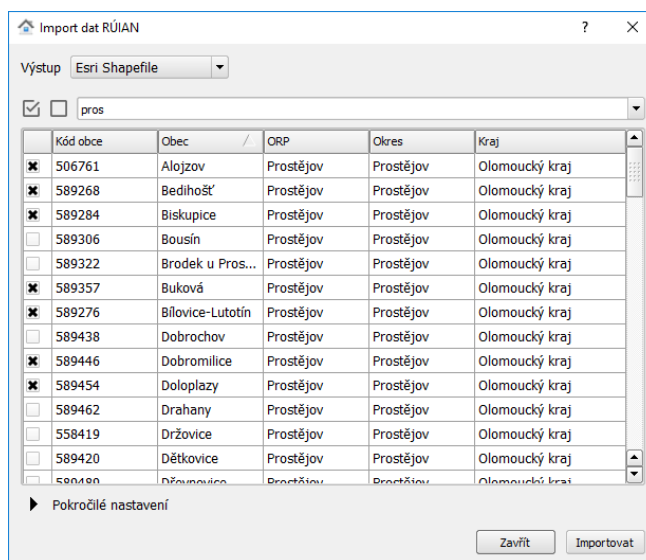
Pro použití nástroje je třeba nejprve stáhnout XML soubor ve formátu VFR (Výměnný formát RÚIAN) z webových stránek Veřejného dálkového přístupu do Katastru nemovitostí od ČÚZK. Stahovat je možné po jednotlivých obcích podle kódu obce, výběr obcí lze filtrovat podle kraje nebo ORP. Při stahování je možné vybrat stažení aktuálních nebo historických dat, kompletní datovou sadu nebo jen přírůstky od určitého data. Dále je také možné vybrat si mezi základní a kompletní datovou sadou, základní datová sada není nástrojem VFR Import podporována. Poté je v nástroji potřeba vybrat vstupní (stažený) XML soubor a cílovou geodatabázi, do níž nástroj data z RÚIANu převede. Nástroj umožňuje v základní verzi importovat pouze jednu obec současně, proto není pro stahování dat ve větším rozsahu příliš efektivní a nebyl v práci využit.

### QGIS RUIAN Plugin

QGIS RUIAN Plugin 1.0.0 je nástroj sloužící k převodu dat z databáze RÚIAN do formátu ESRI Shapefile (SHP). Nástroj funguje jako plugin do programu QGIS verze 2.18 a vyšší. Tento plugin byl vytvořen laboratoří GeoForAll na Fakultě stavební ČVUT v Praze a umožňuje stahovat aktuální data z RÚIANu přes uživatelské rozhraní v QGISu (Obr. 3.2.2.2). Pro použití nástroje je nejprve potřeba v nastavení zásuvných modulů QGISu připojit repozitář GeoForAll ČVUT. Poté lze do QGISu nainstalovat RÚIAN Plugin. Po úspěšné instalaci se do nástrojové lišty QGISu přidá nová ikona, přes kterou se nástroj spouští (Laborať GeoForAll, 2016).

Po spuštění nástroje se zobrazí uživatelské rozhraní s tabulkou všech dostupných obcí ke stažení. Pro filtrování obcí je možné vybrat obce podle názvu, ORP, okresu nebo

kraje. Nástroj umožňuje vybrat více obcí najednou, podle autorů se však doporučuje zvolit menší objem dat, maximálně ve velikosti okresu. V pokročilém nastavení nástroje lze vybrat stažení kompletní nebo základní datové sady. Na závěr je třeba zvolit složku, do níž budou data uložena a výstup dat. Lze zvolit mezi ESRI Shapefile, OGC GeoPackage nebo SQLite databázi. Po doběhnutí importu lze data přímo zobrazit v QGISu. Pro rychlé a snadné stahování kompletních dat z RÚIANu za větší územní celky byl nástroj pro práci využit.



Obr. 3.2.2.2 Rozhraní nástroje QGIS RUIAN Plugin

## 4 PŘÍPRAVA A ZPRACOVÁNÍ DAT

Vlastnímu testování disgregačních metod předcházela výběr disgregovaných statistických ukazatelů, příprava dostupných pomocných dat pro disgregaci a výběr vhodných atributů použitých pro přepočítání. Vzhledem k zadání práce byla definována testovací území, na nichž bude proveden přepočítání a výsledná srovnání. Bylo třeba také zvolit vhodnou velikost buňky váhového rastru pro co největší přesnost přepočítání.

### 4.1 Výběr statistických ukazatelů

Samotnému řešení práce předcházela výběr statistických ukazatelů, pro něž bude přepočítání provedeno. Po konzultaci s vedoucím práce byly vybrány tři základní statistické ukazatele – počet obyvatel (hustota zalidnění), podíl nezaměstnaných osob a průměrná hrubá mzda. Pro každý z těchto ukazatelů bylo pro přepočítání potřeba najít vhodná data za vyšší územní jednotky a následně pro validaci data za nižší územní jednotky.

#### **Počet obyvatel (hustota zalidnění)**

Nejprve byla provedena analýza dostupných primárních dat poskytnutých Katedrou geoinformatiky, tedy zdůvěřovaná data ze SLDB 2011 za ZSJ-díl České republiky, tabulky 111 až 123. Data byla Katedře geoinformatiky poskytnuta Českým statistickým úřadem. Tabulky obsahují kompletní výsledky SLDB 2011 za celou Českou republiku, agregované do územních jednotek ZSJ-díl. Pro disgregaci počtu obyvatel byla použita tabulka 111 „Obyvatelstvo podle pohlaví a podle druhu pobytu, státního občanství, způsobu bydlení, národnosti a náboženské víry“. Tato tabulka obsahuje celkem 23 303 řádků a 167 sloupců. Zdůvěřování znamená, že pro území ZSJ-díl, v nichž bydleli maximálně dva obyvatelé, byly z důvodu ochrany osobních údajů hodnoty přepočteny do největší ZSJ v obci a hodnoty v dané ZSJ-díl nahrazeny hodnotou i.d. (důvěrný údaj). Vzhledem k tomu, že zdůvěřování se týká řádově jednotek obyvatel, byla tato chyba při práci zanedbána. Jako data pro validaci disgregace byl vybrán sloupec „Obyvatelstvo celkem“, obsahující údaje o počtu obyvatel s obvyklým pobytem.

Pro důkladnější analýze získaných dat bylo zjištěno, že se některé ZSJ skládají z více dílů. Nejnižší územní jednotky dostupné v rámci databáze ArcČR® 500 jsou však ZSJ, proto musela být data z ZSJ-díl agregována do ZSJ. Agregace proběhla v programu Microsoft Excel, kdy byly pomocí jazyka VBA do sloupce vypsány pouze unikátní kódy ZSJ a poté pomocí funkce SUMIF sečteny hodnoty pro každý unikátní kód ZSJ.

Protože bylo potřeba provádět disgregace počtu obyvatel na více úrovních podrobnosti, jako zdrojová data byla vybrána data o počtu obyvatel s obvyklým pobytem z SLDB 2011 z Veřejné databáze ČSÚ za obce a ORP. Tato data byla uložena ve formátu Microsoft Excel, každé územní jednotce byl přiřazen její kód a podle něho byla data pomocí nástroje Join nahrána do prostředí ArcMap.

#### **Průměrná hrubá mzda**

Pro disgregaci průměrné hrubé mzdy byla použita data z Veřejné databáze ČSÚ. Data jsou dostupná za Českou republiku a kraje jako absolutní hodnoty nebo meziroční index. Data lze pro roky 2000–2018 získat za jednotlivá čtvrtletí nebo celý rok. Pro potřeby práce byla stažena data za Českou republiku a jednotlivé kraje z roku 2018. V rámci práce byl využit atribut „Průměrná měsíční mzda (na fyzické osoby) v Kč“. Data byla získána ve formátu Microsoft Excel (XLSX), každému kraji byl přiřazen jeho kód a podle něj byla data pomocí nástroje Join nahrána do prostředí ArcMap.



## **Podíl nezaměstnaných osob**

Podíl nezaměstnaných osob vyjadřuje podíl dosažitelných uchazečů o zaměstnání ve věku 15-64 let ze všech obyvatel ve stejném věku (ČSÚ, 2019). Pro disagregaci podílu nezaměstnaných osob byla použita data „Vývoj indikátorů v Olomouckém kraji“ poskytnutá Katedrou geoinformatiky ve formátu Microsoft Excel (XLSX). Data obsahují 12 tabulek popisujících vývoj 23 indikátorů mezi lety 2007 a 2016 za obce Olomouckého kraje. Pro validaci disagregace byl použit atribut NEZAM\_16 obsahující údaj o podílu nezaměstnaných osob v procentech. Jako data zdrojových zón byla stažena data o podílu nezaměstnaných osob z ČSÚ za okresy Olomouckého kraje. Každému okresu byl v programu Microsoft Excel přiřazen jeho příslušný kód a podle něj byla data pomocí nástroje Join nahrána do prostředí ArcMap.

## **4.2 Testovací území**

Vzhledem k charakteru práce bylo nejprve nutné zvolit několik úrovní podrobnosti, pro něž bude přepočten statistických dat probíhat. V rámci získání co nejvyšší přesnosti výsledků disagregace byla snaha provádět přepočten statistických dat na co největší úrovni podrobnosti. Provádět disagregaci statistických údajů pro celé území České republiky by bylo velmi časově a datově náročné, proto byl vybrán jako testovací území pouze Olomoucký kraj. Volba úrovní podrobnosti korespondovala s dostupností zdrojových a referenčních dat.

Olomoucký kraj se rozkládá ve střední části Moravy. Celková výměra kraje 5 267 km<sup>2</sup> tvoří 6,7 % z celkové rozlohy České republiky. Svou rozlohou se řadí k menším regionům, zaujímá osmé místo mezi kraji ČR. K 31.12.2018 zde žilo 632 492 obyvatel (6,1 % ČR) a s průměrnou hustotou zalidnění 121,1 obyvatel/km<sup>2</sup> se kraj řadí k průměrně lidnatým krajům České republiky. Zemědělská půda zaujímá v Olomouckém kraji 53,0 z celkové plochy kultur v kraji (Krajský úřad Olomouckého kraje, 2019).

Hustota zalidnění byla řešena pouze v rámci Olomouckého kraje a byly pro ni zvoleny čtyři úrovně podrobnosti:

- z ORP do obcí,
- z obcí do ZSJ,
- z obcí do gridu,
- ze ZSJ do gridu.

Podíl nezaměstnaných osob byl řešen pouze v rámci Olomouckého kraje a byl disagregován ve dvou úrovních podrobnosti – z kraje do obcí a z ORP do obcí.

Průměrná mzda byla řešena nejprve v rámci celé České republiky ve dvou úrovních podrobnosti – ze státu na kraje a z krajů na okresy. Poté byl proveden odhad i pro Olomoucký kraj – z kraje do obcí.

## **4.3 Výběr pomocných dat**

Metoda disagregace má na svém vstupu hodnoty statistického jevu agregované za určité území, proto je při disagregování potřeba modelovat způsob, jakým rozložit danou statistickou veličinu v námi sledovaném území. Pro tyto účely se využívá tzv. pomocných dat. Pomocnými daty se rozumí datové vrstvy s informacemi o rozložení jiného než disagregovaného jevu v území. Pomocná data mají za úkol zpřesňovat informace o rozložení jevu v daném území, proto je nutné, aby byla k dispozici na detailnější úrovni než disagregovaný jev. Přesnost výsledku disagregace pak bude záviset zejména na vhodnosti použitých pomocných dat, úrovně jejich přesnosti a detailu. Pro každou

testovanou veličinu bylo vybráno několik druhů pomocných dat, pro ně byl následně proveden přepočítání a výsledky pak byly zhodnoceny na referenčních datech.

### **4.3.1 Počet obyvatel**

Na základě rešerše dosavadních zpracování tématu disagregace a konzultace s vedoucím práce byla jako pomocná data pro disagregaci počtu obyvatel vybrána data o krajinném pokryvu, hustotě silniční sítě, adresních bodech a budovách. Pro každý typ pomocných dat pak bylo vybráno více různých datových zdrojů nebo více variant využití jednoho datového zdroje.

#### **Data o krajinném pokryvu**

Prvním typem pomocných dat byla data o krajinném pokryvu. Pro krajinný pokryv byla zjišťována vhodnost použití dat pro dva různé datové zdroje – CORINE Land Cover, Open Land-Use Map.

CORINE Land Cover (CLC) je evropská databáze o krajinném pokryvu vytvářená službou Copernicus Land Monitoring Service. Tvorba datové sady byla zahájena v roce 1985 a první kompletní verze vyšla v roce 1990. Od roku 2000 vychází aktualizované verze datové sady pravidelně každých šest let. CLC používá jako minimální mapovací jednotku území o rozloze 25 ha pro plošné jevy a minimální šířku 100 m pro liniové jevy. Tvorba probíhá ve většině států vizuální interpretací satelitních snímků, v některých státech je část tvorby automatizována pomocí klasifikace satelitních snímků nebo dostupných národních dat (Copernicus, 2019). Pro potřeby práce byla stažena nejnovější vektorová datová sada z roku 2018.

Open Land-Use Map (OLM) je evropská databáze o využití území zpracovávaná v projektu SDI4Apps koordinovaném Západočeskou univerzitou v Plzni. OLM je bezesvá datová sada sloužící k tvorbě detailních map o krajinném pokryvu různých oblastí, založená na celoevropských datových sadách jako CORINE Land Cover nebo UrbanAtlas, doplněná o dostupná detailní regionální data (SDI4Apps, 2015). Data jsou do sebe spojena tak, že podrobnější datová vrstva má přednost před méně podrobnou. Pro Českou republiku jsou data brána nejprve z Katastru nemovitostí, dále pak databáze LPIS, Urban-Atlas a CORINE Land Cover. Datová sada je k volně dispozici ke stažení ve formátu ESRI Shapefile (SHP) za administrativní jednotky NUTS-3 pro celou Evropskou unii. Kódy využití území odpovídají evropské hierarchické klasifikaci HILUCS (Hierarchical INSPIRE Land Use Classification System). Data mají po stažení definovaný špatný souřadnicový systém, pro opravu této chyby byl datům pomocí nástroje Define Projection nastaven souřadnicový systém WGS 1984 Web Mercator (Auxiliary Sphere).

#### **Hustota silniční sítě**

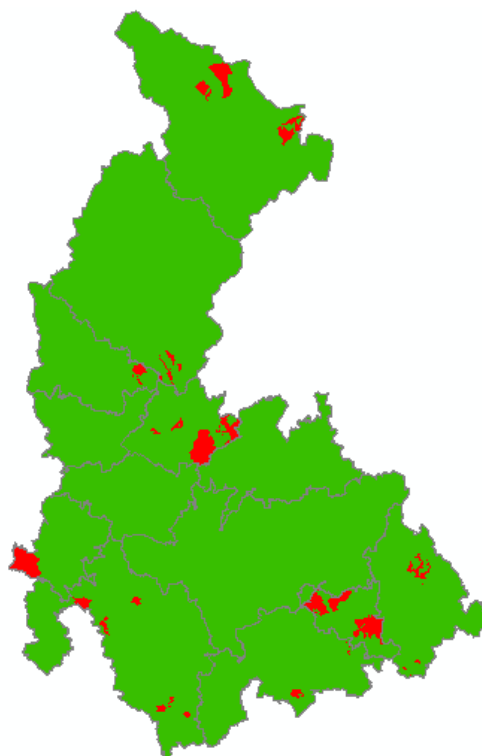
Pro získání informací o hustotě silniční sítě v daném území byla využita datová sada OpenStreetMap (OSM) dostupná pod licencí Open Database License (ODbL). OSM jsou otevřená data vytvářená komunitou přispěvatelů OpenStreetMap. Pro potřeby práce byla stažena data z německého serveru GEOFABRIK za Českou republiku ve formátu ESRI Shapefile (SHP). Kompletní balíček obsahuje 18 vrstev, pro přepočítání byla využita pouze liniová vrstva `gis_osm_roads_free_1` obsahující všechny zakreslené cesty, informace o jejich typu, označení a maximální povolené rychlosti. Datová sada za celou Českou republiku obsahuje 1 128 360 záznamů a 12 atributů. Pro potřeby práce byly využity pouze atributy `fclass` a `name`.

#### **Adresní body a budovy**

Prvním využitým datovým zdrojem byl dataset adresních bodů „Statistické budovy (vchody k bytům)“ poskytnutý autorovi Katedrou geoinformatiky ve formátu ESRI

Shapefile. Bodová vrstva obsahovala celkem 2 809 711 adresních bodů a 63 atributů. Bylo zjištěno, že celkem 1 988 563 záznamů (71 %) obsahovalo v atributu NAZ\_OBEC město Choceň a většina atributů obsahovala nulové hodnoty. Po důkladnější analýze však bylo zjištěno, že chybou částí vrstvy je pouze atributová část, prostorová lokalizace adresních bodů byla v pořádku. Pro potřeby práce byla tedy využita pouze geometrie vrstvy.

Jako zdroj dat o rozmístění budov v území a adresních bodech byly použita datové sady z databáze RÚIAN. Pomocí nástroje QGIS RUIAN Plugin (viz kapitola 3.2.2) byla stažena ke dni 27. 3. 2019 nejnovější data za obce Olomouckého kraje a uložena do formátu ESRI Shapefile. I přes doporučení autorů stahovat data za menší územní jednotky byl nástroj dostatečně robustní na to, aby zvládl stáhnout najednou data za celý kraj. Do adresáře bylo uloženo 11 souborů, z nich byly při práci použity pouze datasety „adresnimista“ a „stavebniobjekty“. Po důkladném zkoumání dat bylo zjištěno, že 138,59 km<sup>2</sup> (2,36 %) rozlohy Olomouckého kraje databáze RÚIAN nepokrývá (viz Obr. 4.3.1.1). V těchto místech (na Obr. 4.3.1.1 červeně) je v databázi k dispozici pouze bodová vrstva adresních bodů.



Obr. 4.3.1.1 Pokrytí Olomouckého kraje daty RÚIAN

Datová sada stavebních objektů obsahuje 43 atributů a 245 427 záznamů. Pro potřeby práce byly využity tři atributy:

- ZpusobVyuz – způsob využití budovy,
- PocetBytu – počet bytů v budově,
- PocetPodlaz – počet podlaží budovy.

Před použitím dat pro práci bylo nejprve potřeba opravit alespoň evidentní chyby. Stavební objekty byly seřazeny sestupně nejprve podle počtu podlaží, poté podle počtu bytů. Nejextrémnější hodnoty byly poté opraveny podle údajů podobných stavebních objektů v okolí nebo vizuální interpretací panoramatických snímků od Mapy.cz. Pro přepočítání pomocí kombinace adresních bodů a obytných budov pak bylo potřeba, aby

stavební objekty sloužící k bydlení měly alespoň jedno podlaží. Pomocí nástroje Calculate Field byl tedy u všech obytných budov s nula podlažími stanoven počet podlaží na jedno.

Datová sada adresních bodů za Olomoucký kraj obsahuje 176 852 záznamů a 16 atributů, mezi nimi číslo domovní, číslo orientační a PSČ. Adresní body nenesou informace o využití budovy, ale po prozkoumání jejich prostorové lokalizace bylo zjištěno, že se vždy nachází uvnitř budovy. Pro získání obytných adresních bodů byly nejprve vybrány všechny budovy s využitím jiným než pro bydlení a poté byly pomocí nástroje Erase odstraněny všechny adresní body v těchto budovách. Tím bylo docíleno zachování adresních bodů i v místech bez pokrytí daty RÚIAN. Takto připravené datové sady byly poté využity v práci.

### **4.3.2 Průměrná hrubá mzda**

Pro disagregaci průměrné hrubé mzdy byly testovány čtyři atributy – index vzdělanosti, podíl obyvatel zaměstnaných v terciéru, míra zaměstnanosti a počet dokončených bytů.

Index vzdělanosti byl pro potřeby práce stanoven jako podíl obyvatel s ukončeným minimálně úplným středním vzděláním na počtu obyvatel s obvyklým pobytem starších 15 let. Tento jednoduchý index pro potřeby práce dostatečně popisuje vzdělanostní charakteristiku obyvatelstva. Pro výpočet indexu vzdělanosti byla použita data o nejvyšším ukončeném vzdělání z Veřejné databáze ČSÚ platná ke dni 26. 3. 2011. Data jsou k dispozici ve formátu XLSX za obce, ORP, okresy, kraje a stát. Díky tomu je možné index vzdělanosti využít pro disagregaci průměrné hrubé mzdy v různých úrovních podrobnosti. Index byl vypočítán v prostředí Microsoft Excel.

Podíl obyvatel zaměstnaných v terciéru byl stanoven jako podíl obyvatel zaměstnaných v jiném sektoru než priméru a sekundéru na počtu zaměstnaných osob včetně pracujících studentů a učňů. Pro výpočet byla použita data o odvětví ekonomické činnosti z Veřejné databáze ČSÚ platná ke dni 26. 3. 2011. Data jsou k dispozici ve formátu XLSX za obce, ORP, okresy, kraje a stát. Pro výpočet počtu obyvatel zaměstnaných v terciéru byly sečteny sloupce:

- stavebnictví,
- velkoobchod a maloobchod; opravy a údržba motorových vozidel,
- doprava a skladování,
- ubytování, stravování a pohostinství,
- informační a komunikační činnosti,
- peněžnictví a pojišťovnictví,
- činnosti v oblasti nemovitostí, profesní, vědecké a technické činnosti a administrativní a podpůrné činnosti,
- veřejná správa a obrana; povinné sociální zabezpečení,
- vzdělávání,
- zdravotní a sociální péče.

Tento součet byl poté podělen počtem zaměstnaných včetně pracujících studentů a učňů. Index byl vypočítán v prostředí Microsoft Excel.

Míra zaměstnanosti vyjadřuje podíl počtu zaměstnaných na počtu všech osob 15letých a starších (ČSÚ, 2019). Data byla stažena z Veřejné databáze ČSÚ platná ke dni 26. 3. 2011. Data jsou dostupná pouze za stát a kraj. Pro získání podrobnějších dat bylo potřeba využít datové sady z výpočtu indexu vzdělanosti a podílu obyvatel zaměstnaných

v terciéru. Pro nižší územní jednotky byla míra zaměstnanosti vypočtena jako podíl zaměstnaných včetně pracujících studentů a učňů na počtu obyvatel starších 15 let.

Posledním testovaným atributem byl počet dokončených bytů. Data byla stažena ve formátu XLSX z Veřejné databáze ČSÚ platná ke dni 31. 12. 2018. Data jsou dostupná za obce, ORP, okresy, kraje a stát. Při výpočtu pak byl použit počet dokončených bytů ve stovkách.

Pomocí funkce CORREL byla zjišťována korelace mezi jednotlivými atributy a průměrnou hrubou mzdou na území krajů. Bylo zjištěno (viz Obr. 4.3.2.1), že nejbližší vztah k průměrné hrubé mzdě mají atributy IND\_VZD a TERCIER, proto byly tyto atributy využity. Jako třetí atribut pro disagregaci pak byla využita míra zaměstnanosti (MIRA\_ZAM).

KRAJ	IND_VZD	TERCIER	MIRA_ZAM	DOK_BYTY	MZDA_2018
Hlavní město Praha	58,87	74,38	59,2	52,9	38 221 Kč
Jihočeský kraj	42,43	54,05	55,4	17,13	27 887 Kč
Jihomoravský kraj	45,56	59,67	53,4	44,5	29 627 Kč
Karlovarský kraj	34,72	54,29	54,8	4,88	27 036 Kč
Kraj Vysočina	40,60	49,27	53,5	14,41	28 444 Kč
Královéhradecký kraj	41,83	52,58	53,3	14,74	28 439 Kč
Liberecký kraj	39,47	50,25	53,3	10,44	28 772 Kč
Moravskoslezský kraj	40,67	54,29	51,4	25,49	27 907 Kč
Olomoucký kraj	42,09	55,95	51,7	16,31	27 680 Kč
Pardubický kraj	41,03	52,46	54,3	16,71	27 798 Kč
Plzeňský kraj	41,34	52,68	56,3	24,27	29 711 Kč
Středočeský kraj	43,91	59,04	56,9	72,46	31 460 Kč
Ústecký kraj	35,74	54,48	51,9	11,38	28 582 Kč
Zlínský kraj	41,65	52,01	53	13,06	27 531 Kč
KORELACE	0,89	0,90	0,76	0,70	-

Obr. 4.3.2.1 Tabulka vypočtených atributů a míry korelace

### 4.3.3 Podíl nezaměstnaných osob

Jako pomocná data pro disagregaci podílu nezaměstnaných osob byla použita data poskytnutá Katedrou geoinformatiky popisovaná v kapitole 4.1, aktuální k roku 2016. Byly testovány atributy:

- MIGR\_SA\_16 – hrubá míra migračního salda na 100 obyvatel,
- DOK\_BYT\_16 – počet dokončených bytů na 100 obyvatel,
- IND\_VZD\_11 – index vzdělanosti,
- DOSTUP\_17 – index dostupnosti dálniční sítě,
- DOJIZDK\_11 – hrubá míra salda dojížděky na 100 ekonomicky aktivních obyvatel,
- OVZD\_16 – kvalita ovzduší (koncentrace prachových částic PM10),
- LES\_17 – podíl lesa na rozloze území.

## 5 PŘEPOČET DAT

Po přípravě a zpracování dat následovala hlavní část práce. Hlavní náplní práce bylo otestovat vybrané disagregační metody za účelem přepočtu vybraných statistických ukazatelů do nižších nebo odlišných územních jednotek v několika úrovních podrobnosti a zhodnocení přesnosti použitých metod na referenčních datech. Po konzultaci s vedoucím práce byly vybrány tři statistické ukazatele – počet obyvatel (hustota zalidnění), podíl nezaměstnaných osob a průměrná hrubá mzda.

### 5.1 Průběh zpracování

Za účelem přepočtu dat z vyšších do nižších nebo odlišných územních jednotek byla využita analytická metoda disagregace. Metoda disagregace má na svém vstupu hodnoty statistických dat agregované do vyšších územních jednotek, z nich je pak proveden odhad hodnoty jevu v nižších územních jednotkách. Protože jev ve sledovaném území není konstantní, je vhodné použít pro disagregaci pomocná data znázorňující rozložení jevu v daném území. Výsledkem disagregace je pak polygonová, rastrová nebo gridová vrstva s nově odvozenými hodnotami.

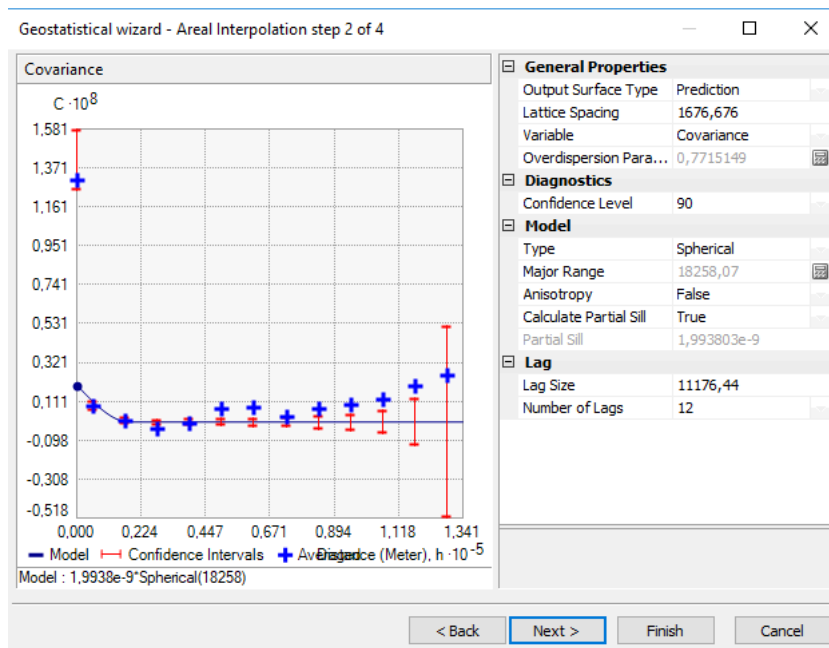
Pro každý z vybraných statistických ukazatelů bylo otestováno několik různých typů pomocných dat, byl vytvořen nástroj pro automatický přepočet v rozhraní ModelBuilder. Dále byla sledována časová náročnost jednotlivých přepočtů. Přepočty byly prováděny na osobním počítači s konfigurací AMD Ryzen 7 1700 Eight-Core CPU (3,70 GHz), 16 GB RAM, NVIDIA GeForce GTX 1060 6GB. Metody s nejlepšími výsledky byly následně vizualizovány.

#### 5.1.1 Počet obyvatel

V rámci testování disagregačních metod na počtu obyvatel proběhl přepočet ve čtyřech různých úrovních, s čtyřmi typy pomocných dat. Pro každý typ pomocných dat bylo zvoleno více různých nastavení tak, aby bylo dosaženo co nejlepších výsledků.

##### **Areálová interpolace**

Prvním použitým nástrojem pro přepočet počtu obyvatel byl nástroj Areal Interpolation, popisovaný v kapitole 3.2.1. Tímto nástrojem byla prováděna disagregace pouze z úrovně obcí do úrovně ZSJ. Jako vstupní data byla použita vrstva ObceOL obsahující hranice obcí a informaci o počtu obyvatel v nich. Vzhledem k povaze dat bylo jako typ dat zvoleno Poissonovo rozdělení. Jako pole s počtem obyvatel byl zvolen sloupec POC\_OB. Protože data byla sbírána za jeden konkrétní čas, byl údaj o čase vynechán. Bylo ponecháno původní nastavení vlastností modelu (viz Obr. 5.1.1.1). Po proběhnutí modelu byla vytvořena rastrová vrstva areálové interpolace, ta byla pak pomocí nástroje Areal Interpolation Layer To Polygons převedena do nových územních jednotek (ZSJ) a uložena do vrstvy ZSJ\_OLkraj\_ArealInt do atributu Predicted. Přepočet nebyl nijak časově náročný, tvorba rastrové vrstvy trvala 36 vteřin a následné převedení do polygonů poté 37 vteřin.



Obr. 5.1.1.1 Nastavení modelu

### Disagregace s využitím pomocných dat

Pro disagregaci počtu obyvatel bylo dále využito pět různých typů pomocných dat:

- rozloha areálu (areálové vážení),
- krajinný pokryv,
- silniční síť,
- adresní body,
- budovy.

Pro každý typ pomocných dat byly nejprve stanoveny váhy, poté byla zvolena vhodná velikost buňky budoucího rastru. Následně byla pomocná data převedena do rastru. Tento rastr sloužil jako váhový rastr nesoucí detailní informace o území a společně s hodnotami statistického ukazatele vstupoval do samotného přepočtu. Po dobytí přepočtu byla data zpětně agregována z rastrové mřížky do nových územních jednotek.

První testovanou metodou disagregace bylo areálové vážení. Tato metoda pro svůj výpočet nepotřebuje žádná externí statistická data, jako pomocná data používá rozlohu jednotlivých dílčích částí celku. Podle Flowerdewa a spol. (1991) v případě, že statistickým údajem je počet, areálové vážení předpokládá, že daný jev je v prostoru rozložen rovnoměrně. Na základě tohoto předpokladu byl vytvořen váhový rastr, v němž mají všechny buňky stejnou hodnotu. Tento rastr poté společně se zdrojovou a cílovou zónou sloužil jako vstupní data do modelu.

Dalším testovaným typem dat byla data o krajinném pokryvu. Pro přepočet byl využit binární přístup. Tento přístup ve své práci aplikovali např. Holt a kol. (2004) nebo Bielecka (2005) a spočívá v rozdělení území na místa, kde se jev vyskytovat může a místa, kde se vyskytovat nemůže. Nejprve byla vzata datová sada CORINE Land Cover. Na základě rešerše významu kódů byly označeny jako obytné kódy 111 (Continuous urban fabric) a 112 (Discontinuous urban fabric). Ve vrstvě byl vytvořen nový atribut WEIGHT, pro tyto kódy mu byla stanovena hodnota 1. Ostatní kódy pak nabývaly hodnoty 0. Datová sada byla převedena na rastr o stanovených velikostech pixelu. Dále byla vzata data Open Land-Use Map. Tato data obsahují informace o využití území ve sloupci

hilucs\_lan. Po zjištění významu kódů byly jako obytné plochy stanovené kódy 500 (bydlení), 510 (čistě obytné území) a 520 (smíšené obytné území). Do sloupce WEIGHT pak byly podobně jako u vrstvy CORINE Land Cover stanoveny váhy a pomocí nástroje Feature To Raster byla data převedena na rastr o stanovené velikosti buňky.

Jako další typ pomocných dat byla využita hustota silniční sítě. Využití hustoty silniční sítě ve své práci testoval Zandbergen (2011), podle něj vykazují tato data výsledky podobné jako data o krajinném pokryvu. Jako zdrojová data byla použita volně dostupná datová sada OpenStreetMap. Informace o typu silnice se nachází v atributu fclass. Po zjištění významu jednotlivých hodnot pak bylo možné hodnotit přesnost několika různých variant těchto pomocných dat. Pro disagregaci byly využity tři varianty dat z OSM:

- kompletní – obsahují všechny liniové prvky sloužící jako cesty,
- silnice – obsahují pouze silniční a uliční síť,
- lokální silnice – silniční a uliční síť bez dálnic, silnic I. a II. třídy mimo obec.

Kompletní varianta je kompletní staženou verzí dat OSM. Pro její následné použití bylo potřeba pomocí nástroje Project převést tuto vrstvu do souřadnicového systému S-JTSK Krovak East North. Varianta silnice se nachází ve vrstvě OSM\_OLkraj\_roads, pro její tvorbu byly z kompletní verze podle atributu fclass vybrány pouze hodnoty „living\_street“, „motorway“, „pedestrian“, „primary“, „residential“, „secondary“, „tertiary“, „trunk“ a „unclassified“. Tím byly odstraněny zejména lesní a polní cesty, které nijak nekorespondují s rozmístěním obyvatelstva v území. Varianta lokální silnice pak vycházela z varianty silnice, pouze z ní byly odstraněny silnice I. a II. třídy neobsahující v atributu name název ulice, dále pak dálnice a silnice pro motorová vozidla. Hustota silniční sítě byla poté pro každou variantu vypočítána nástrojem Line Density a použita jako váhový rastr pro přepočítání.

Další vstupní vrstvou byly adresní body. Byly využity dva typy adresních bodů popisované v kapitole 4.3.1. Jako kompletní adresní body byla využita datová sada „Statistické budovy (vchody k bytům)“, jako obytné adresní body potom upravená bodová vrstva z databáze RÚIAN. V rámci práce a z důvodu nedostatku dat bylo předpokládáno, že každý adresní bod má stejnou váhu. Oba druhy dat poté byly pomocí nástroje Point Density převedeny na váhový rastr o stanovené velikosti buňky.

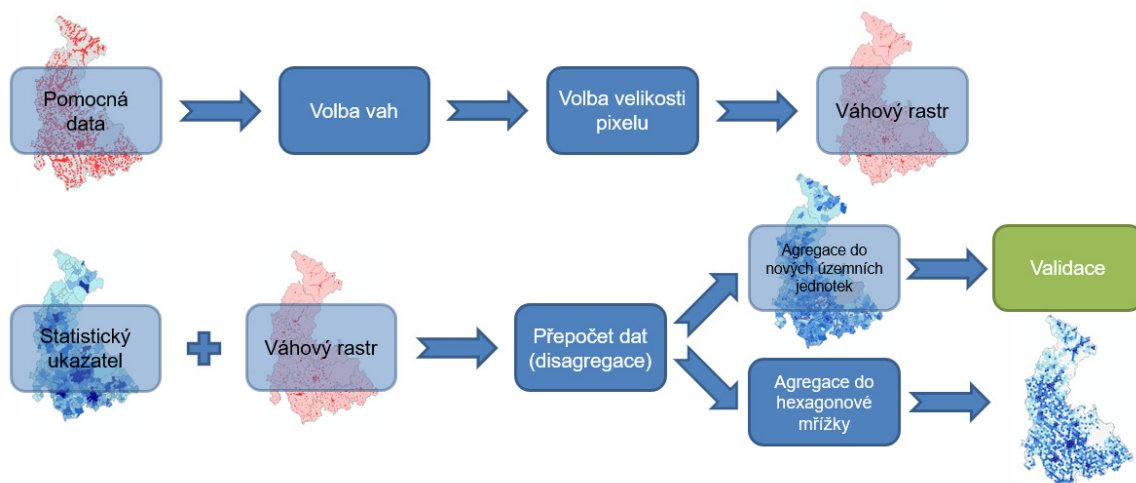
Posledním typem pomocných dat byla polygonová vrstva budov z databáze RÚIAN. Data však nepokrývají kompletní Olomoucký kraj (Obr. 4.3.1.1) a obsahují některé chyby, které bylo nutné opravit (viz kapitola 4.3.1). Data o budovách byla využita v pěti variantách – všechny budovy, obytné budovy, počet podlaží obytných budov, počet bytů obytných budov a kombinace obytných adresních bodů s obytnými budovami. Varianty byly vytvořeny na základě zjištění významů jednotlivých atributů. Varianta všechny budovy byla bez úprav převedena na rastr o stanovené velikosti pixelu tak, že všechny budovy měly stejnou váhu. Jako obytné budovy byly vybrány podle atributu ZpusobVyuz budovy s kódem 3 (objekt k bydlení), 6 (bytový dům) a 7 (rodinný dům). Obytné budovy byly převedeny na rastr stejným způsobem jako všechny budovy. Při testování počtu podlaží a počtu bytů obytných budov byl jako váha do váhového rastru stanoven počet podlaží, resp. bytů.

Kombinace obytných adresních bodů s počtem podlaží obytných budov byla navržena tak, aby eliminovala chyby přepočtu vzniklé nekompletním pokrytím území daty RÚIAN. Nejprve byly obytné adresní body převedeny do rastru o velikosti buňky 5 metrů. Místa s adresním bodem dostala hodnotu 1, ostatní místa hodnotu 0. Poté byl na rastr převeden i počet podlaží obytných budov tak, že každému typu budov byla přiřazena váha a ta byla



poté vynásobena počtem podlaží. Pro potřeby práce byly testovány tři různé váhy. Oba rastry byly následně sečteny a vzniklý rastr sloužil jako váhový rastr pro přepočítání.

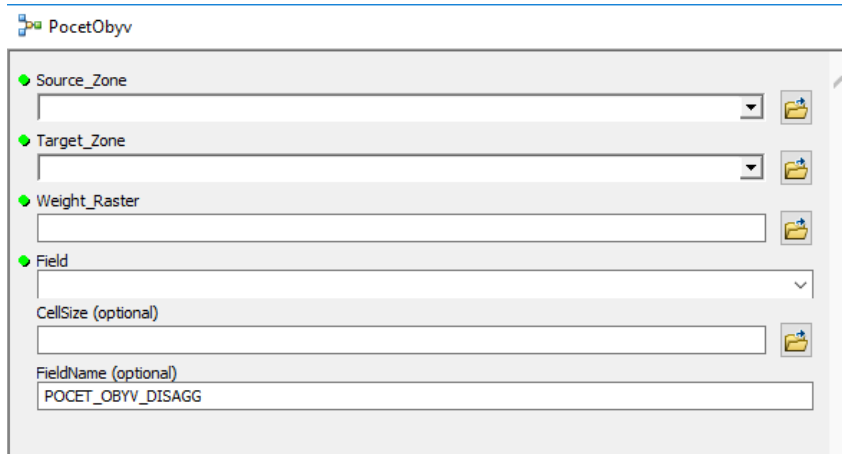
Samotný přepočítání byl navržen jako sekvence několika kroků. Nejprve byl počet obyvatel ve zdrojových zónách převeden na rastr o stanovené velikosti buňky (viz obr. 5.1.1.2). Poté byla z váhového rastru vypočítána suma vah pomocných dat v každé zdrojové zóně. Váha buňky byla poté vydělena sumou vah v příslušné zdrojové zóně a vynásobena počtem obyvatel v tomto území. Tím vznikl odhad počtu obyvatel v jedné buňce. Pro zjištění počtu obyvatel v cílové zóně pak stačilo sečíst všechny dílčí odhady v rámci cílové zóny. Protože by bylo komplikované a časově náročné tento přepočítání realizovat pro všechny varianty pomocných dat manuálně, byl pro tento účel vytvořen nástroj v rozhraní ModelBuilder (Příloha 16).



Obr. 5.1.1.2 Schéma postupu práce

Model funguje jako posloupnost osmi nástrojů. Na vstupu se nachází šest parametrů:

- Source\_Zone – zdrojová zóna,
- Target\_Zone – cílová zóna,
- Weight\_Raster – váhový rastr vytvořený rasterizací pomocných dat,
- Field – atribut počtu obyvatel v atributové tabulce zdrojové zóny,
- CellSize – velikost buňky rastrů v přepočtu v metrech,
- FieldName – název nově vytvořeného sloupce s přepočtenými daty.



Obr. 5.1.1.3 Rozhraní nástroje

Nejprve byl proveden převod statistických dat do rastru, poté v rastru proběhl přepočítání z vyšších do nižších nebo odlišných územních jednotek podle návrhu. Výsledek byl poté převeden do atributové tabulky cílové zóny. Všechny průběžně vytvořené vrstvy jsou po úspěšném doběhnutí modelu odstraněny. Model tedy nevytváří žádné nové vrstvy, pouze jeden sloupec v atributové tabulce s hodnotami. Model také umožňuje zvolit název nově vytvořeného sloupce. Názvy jednotlivých atributů byly definovány jako kombinace testovaného statistického jevu, zdroje pomocných dat, způsobu použití dat a velikosti buňky.

Tento princip byl poté využit pro testování různých typů pomocných dat a jejich vzájemného srovnání. Popsané vstupní vrstvy byly použity pro disagregaci počtu obyvatel z úrovně ORP do obcí a z úrovně obcí do ZSJ. Pro přepočítání z obcí do ZSJ byly testovány tři velikosti buněk – 5, 10 a 25 metrů. Pro přepočítání z ORP do obcí pak byla jako testovací velikost buňky stanovena hodnota 5 metrů.

### **5.1.2 Průměrná hrubá mzda**

Při testování disagregace průměrné hrubé mzdy byla použita pomocná data popisovaná v kapitole 4.3.2. Před samotným přepočítáním byl navržen postup v několika krocích. Nejprve byly stanoveny váhy pro jednotlivá pomocná data. Poté byla pomocná data těmito váhami vynásobena a vznikl index pro disagregaci. Tento index byl ve zdrojové zóně zprůměrován. Hodnota indexu v cílové zóně byla následně podělena zprůměrovanou hodnotou indexu ve zdrojové zóně. Vznikl koeficient, kterým byla vynásobena hodnota průměrné hrubé mzdy ve zdrojové zóně. Výsledkem je odhadnutá průměrná hrubá mzda v cílové zóně. Přesnost disagregace pak závisí na vhodnosti volby parametrů a jejich vah.

Vzhledem k náročnosti navrženého postupu byl pro usnadnění tvorby odhadů průměrné hrubé mzdy vytvořen nástroj v rozhraní ModelBuilder (Příloha 17). Nástroj umožňuje provádět disagregaci pomocí až tří různých vstupních parametrů. Nástroj vyžaduje na vstupu jako pomocná data tři atributy, pro použití méně atributů stačí jako pomocná data vybrat libovolný sloupec s číselným datovým typem a nastavit hodnotu jeho váhy na 0. Na vstupu do nástroje se nachází 11 parametrů (Obr. 5.1.2.1):

- SourceZone – zdrojová zóna,
- TargetZone – cílová zóna,
- Output cell size – velikost buňky rastrů v přepočtu v metrech,
- StatValue – jméno sloupce s průměrnou hrubou mzdou ve zdrojové zóně,
- New Field Name – jméno nově vytvořeného sloupce s přepočtenou mzdou,
- Variable1, Variable2, Variable3 – atributy sloužící jako pomocná data,
- Variable1\_Weight, Variable2\_Weight, Variable3\_Weight – váhy pomocných dat.

PrumMzda

SourceZone  
CeskaRepublika

Output cell size (optional)  
25000

TargetZone  
Kraje\_CR\_mzda

Variable1  
MIRA\_ZAM

Variable2  
TERCIER

Variable3  
IND\_VZD

StatValue  
MZDA\_FYZ\_2018

New Field Name (optional)  
Mzda\_Disagg

Variable1\_Weight  
1

Variable2\_Weight  
1

Variable3\_Weight  
1

Obr. 5.1.2.1 Rozhraní nástroje

Nástroj předpokládá, že zprůměrovaná hodnota statistického jevu v území odpovídá zprůměrovanému disagregačnímu indexu. Podíl indexu v cílové zóně a zprůměrovaného indexu pak odpovídá podílu průměrné hrubé mzdy v cílové zóně a podílu průměrné hrubé mzdy ve zdrojové zóně. Nástroj umožňuje volbu až tří pomocných dat a jejich vah. Je také použitelný i pro disagregaci jiných statistických ukazatelů, u nichž bude platit, že čím vyšší bude hodnota disagregačního indexu pomocných dat, tím vyšší bude hodnota statistického jevu.

Pro disagregaci byly vybrány atributy IND\_VZD, MIRA\_ZAM a TERCIER (viz kapitola 4.3.2). Bylo testováno několik různých nastavení vah těchto atributů. Dále byla testována vhodná velikost buňky pro přepočítání a doba trvání přepočtu pro jednotlivé metody. Průměrná hrubá mzda byla testována na několika úrovních podrobnosti. Nejprve byla přepočítávána v rámci celé České republiky – ze státu do krajů a z krajů do okresů. Následně byl proveden přepočítání v rámci Olomouckého kraje z kraje do obcí.

### 5.1.3 Podíl nezaměstnaných osob

Při testování disagregace podílu nezaměstnaných osob byla použita pomocná data popisovaná v kapitole 4.3.3. Postup disagregace pak byl velmi podobný postupu u průměrné hrubé mzdy (kapitola 5.1.2). Byly testovány atributy index dojížděky (IND\_DOJIZDK), míra znečištění ovzduší (OVZD), index vzdělanosti (IND\_VZD) a dostupnost dálniční sítě (DOSTUP). Index dojížděky byl vytvořen z atributu DOJIZDK\_11 posunutím hodnoty 0 do hodnoty 50 a vydělením původní hodnoty hrubé míry salda dojížděky dvěma. Hodnoty vytvořeného indexu dojížděky mohou nabývat rozmezí 0 až 100. Index dojížděky vyšší než 50 může indikovat přítomnost většího zaměstnavatele a tím i nižší nezaměstnanost v území. Hodnoty nižší než 50 mohou značit nedostatek pracovních míst v regionu a tím i vyšší nezaměstnanost. Vyšší míra znečištění ovzduší může indikovat přítomnost velkého zaměstnavatele a tím i nižší nezaměstnanost. Index vzdělanosti nese vzdělanostní charakteristiku obyvatelstva. Byl stanoven předpoklad, že více vzdělaní lidé budou méně častěji nezaměstnaní. Index dostupnosti dálniční sítě byl pro potřeby práce upraven tak, aby nejlépe dostupná místa měla hodnotu 100 a nejhůře dostupná pak hodnotu 0. Lepší dostupnost dálniční sítě může značit vyšší koncentraci

zaměstnavatelů v regionu i snadnější dostupnost pro zaměstnance, což může mít za následek nižší nezaměstnanost.

Pro disagregaci podílu nezaměstnaných osob byl využit stejný postup jako u průměrné hrubé mzdy (kapitola 5.1.2). Protože se u všech typů pomocných dat předpokládá, že jsou s podílem nezaměstnaných osob nepřímo úměrné, byl vytvořen nový nástroj v rozhraní ModelBuilder. Model je stejný jako model pro průměrnou hrubou mzdu (Obr. 5.1.2.2), jen při výpočtu koeficientu pro násobení průměrného podílu nezaměstnaných osob počítá s nepřímou úměrností. V rámci práce bylo testováno několik různých nastavení vah pomocných dat. Podíl nezaměstnaných osob byl testován v rámci Olomouckého kraje pro přepočítání z kraje do obcí a z ORP do obcí.

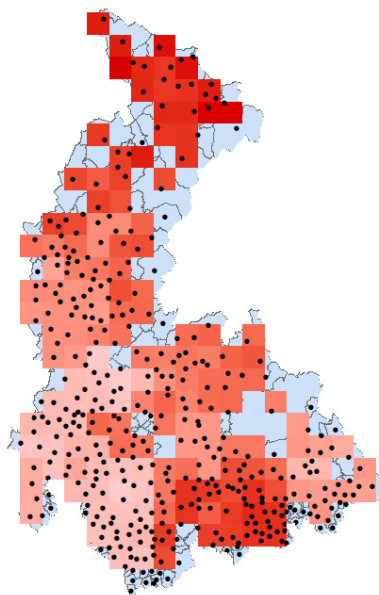
## 5.2 Zhodnocení metod

Důležitou součástí práce bylo zhodnotit vhodnost použití jednotlivých metod a typů pomocných dat pro přepočítání. Dále byla hodnocena přesnost přepočítání na referenčních datech dostupných na lokální úrovni. Klíčovými faktory hodnocení byl výběr vhodné velikosti buňky pro přepočítání a její vliv na dobu a přesnost přepočítání, dále pak přesnost přepočítání pomocí různých metod nebo pomocných dat.

### 5.2.1 Výběr velikosti buňky

Jedním z dílčích cílů práce bylo vybrat vhodnou velikost buňky použitou při přepočítání a zjistit dobu trvání samotného přepočítání. Podle Zajícové (2017) mohou příliš velké buňky způsobit zkreslené výsledky a příliš malé buňky mohou naopak být velmi náročné na výpočet. Přepočítání probíhalo na stolním počítači popisovaném v kapitole 5.1. Pro průměrnou hrubou mzdu a podíl nezaměstnaných osob byl také zjišťován počet nulových hodnot vygenerovaných během přepočítání.

Nulové hodnoty jsou při přepočítání generovány z důvodu příliš hrubého rastru, který není schopný vytvořit jednu buňku pro každou územní jednotku. Centroidům územních jednotek pak při přepočítání nejsou zpětně přiřazeny hodnoty tohoto rastru. Na Obr. 5.2.1.1 lze také pozorovat, že příliš hrubý rastr vlivem nedostatečného detailu přiřazuje více obcím stejnou hodnotu přepočteného jevu. Nulové hodnoty tedy mohou sloužit jako indikátor nepřesnosti přepočítání.



Obr. 5.2.1.1 Vznik nulových hodnot při přepočítání

### Podíl nezaměstnaných osob

Při disagregaci podílu nezaměstnaných osob bylo testováno šest různých velikostí buněk ve dvou úrovních podrobnosti. Byly zvoleny velikosti strany buňky 0,1; 0,25; 0,5; 1; 2,5 a 5 kilometrů. Jako testovací území bylo zvoleno území Olomouckého kraje z důvodu dostupnosti dat. Byly vybrány úrovně podrobnosti z ORP do obcí a z kraje do obcí. Při analýze byla zjišťována doba přepočtu v sekundách, počet vytvořených rastrových buněk a počet nulových hodnot ve výsledné tabulce.

Tab. 5.2.1.1 Kritéria hodnocení velikosti buňky pro přepočet podílu nezaměstnaných osob

Olomoucký kraj				
Úroveň podrobnosti	Velikost hrany buňky [km]	Doba přepočtu [s]	Počet vytvořených buněk	Počet nulových hodnot
z ORP do obcí	0,1	40,36	399	0
	0,25	23,92	399	0
	0,5	23,45	399	0
	1	24,54	398	1
	2,5	22,93	349	9
	5	22,77	163	38
z kraje do obcí	0,1	40,38	399	0
	0,25	38,88	399	0
	0,5	39,11	399	0
	1	38,85	398	1
	2,5	38,76	349	9
	5	39,50	163	38

Na základě Tab. 5.2.1.1 bylo zjištěno, že velikost vybraných buněk výrazně neovlivňuje rychlost přepočtu. Výsledek však byl zkreslen rychlostí posledního nástroje Alter Field, který v některých případech proběhl během chvilky a jindy trval až 20 vteřin.

Důležitějším faktorem hodnocení byl počet vytvořených buněk. Pro snížení zkreslení přepočtu je vhodné, aby každá cílová územní jednotka měla svoji vlastní vytvořenou buňku. Cílovými územními jednotkami byly v tomto případě obce Olomouckého kraje, kterých je 402. Obce Kozlov, Luboměř pod Strážnou a Město Libavá však mají společně data přiřazena do VVP Libavá. Maximální množství buněk je tedy 399. Dalším faktorem byl počet nulových hodnot. V rámci úplnosti přepočtených informací je třeba, aby každá cílová územní jednotka měla po přepočtu svoji vlastní hodnotu. Tato kritéria pro zvolené testovací území splňují buňky o straně 0,5 km a menší.

### Průměrná hrubá mzda

Při disagregaci průměrné hrubé mzdy bylo testováno šest různých velikostí buněk ve třech úrovních podrobnosti. Pro přepočet z úrovně státu do krajů byly vybrány buňky o hraně 1, 5, 10, 25 a 50 kilometrů. Pro přepočet z kraje do okresů byly vybrány velikosti buněk 1; 2,5; 5; 10 a 25 kilometrů. Při analýze byla zjišťována doba přepočtu v sekundách, počet vytvořených rastrových buněk a počet nulových hodnot ve výsledné tabulce.

Tab. 5.2.1.2 Kritéria hodnocení velikosti buňky pro přepočítání průměrné hrubé mzdy.

Česká republika				
Úroveň podrobnosti	Velikost hrany buňky [km]	Doba přepočtu [s]	Počet vytvořených buněk	Počet nulových hodnot
ze státu do krajů	1	15,80	14	0
	5	18,39	14	0
	10	18,29	14	0
	25	17,53	14	0
	50	17,96	13	1
z kraje do okresů	1	19,70	77	0
	2,5	19,96	77	0
	5	19,87	76	1
	10	19,78	77	0
	25	20,32	67	4

Výsledky byly hodnoceny podobným stylem, jako u podílu nezaměstnaných osob. Na základě Tab. 5.2.1.2 bylo zjištěno, že velikost buněk v těchto rozměrech opět nehraje při rychlosti přepočtu žádnou roli. Pro testované velikosti buněk bylo zjištěno, že ve většině případů byla pro každou územní jednotku vytvořena jedna buňka. Jako vhodné velikosti buněk byly vybrány buňky o hraně 25 kilometrů a menší pro převod ze státu do krajů a buňky o hraně 2,5 kilometru a menší pro převod z krajů do okresů. Pro přepočítání z kraje do obcí byly použity výsledky popisované v Tab. 5.2.1.1.

### Počet obyvatel

Při disagregaci počtu obyvatel byly v testovacím území Olomouckého kraje v úrovni z obcí do ZSJ testovány velikosti buňky 5, 10 a 25 metrů. Vhodnost velikosti buňky byla zjišťována na základě přesnosti přepočtu použitím různých typů pomocných dat. Vhodná velikost buňky byla vybrána v kapitole 5.2.2. Bylo zjištěno, že doba přepočtu je nepřímo úměrná velikosti buňky. Pro velikost buňky 5 metrů trval přepočítání 3 minuty a 19 sekund, pro buňku 10 metrů 1 minutu a 37 sekund a pro buňku o straně 25 metrů pak 1 minutu a 8 sekund. Pro další úrovně podrobnosti pak byla vybrána nejpřesnější metoda.

## 5.2.2 Zhodnocení přesnosti

### Počet obyvatel

Pro přepočítání počtu obyvatel do nižších územních jednotek bylo použito více různých typů pomocných dat a jejich nastavení. Výsledné hodnoty pak bylo potřeba porovnat na referenčních datech a výsledná zjištění interpretovat. Pro zhodnocení přesnosti byl pro každou metodu vypočítán počet špatně umístěných lidí. Tento způsob hodnocení aplikoval při řešení své práce Zandbergen (2011) na 12 různých typů pomocných dat. Podle autora je počet špatně umístěných lidí definován jako suma absolutních hodnot rozdílu mezi odhadnutou a známou hodnotou jevu za každou územní jednotku, vydělená dvěma. Špatně umístěná osoba má za následek dvojnásobnou chybu: o jednu osobu více v jednom

území a o jednu méně v druhém území (Zandbergen, 2011). Výsledné metody poté seřadil od nejpřesnější po méně přesnou pomocí procenta špatně umístěných obyvatel.

Při disagregaci byly použity tyto metody a pomocná data:

- POCET\_OBYV\_AW – areálové vážení,
- POCET\_OBYV\_CLC – CORINE Land Cover (binární přístup),
- POCET\_OBYV\_OLM – Open Land-Use Map (binární přístup),
- POCET\_OBYV\_RUIAN\_BUD – kompletní adresní body z RÚIAN,
- POCET\_OBYV\_BUDVCHODY\_POIDEN – statistické budovy (vchody k bytům),
- POCET\_OBYV\_OBYTNE\_POIDEN – obytné adresní body z RÚIAN,
- POCET\_OBYV\_OSM\_FULL – kompletní síť cest z OSM,
- POCET\_OBYV\_OSM\_ROADS – silniční síť z OSM,
- POCET\_OBYV\_OSM\_MINOR – lokální silnice z OSM,
- POCET\_OBYV\_RUIAN\_BUDVSE – budovy z RÚIAN,
- POCET\_OBYV\_RUIAN\_BUD\_BYDLENI – obytné budovy z RÚIAN,
- POCET\_OBYV\_RUIAN\_BUD\_PODLAZI – počet podlaží obytných budov,
- POCET\_OBYV\_RUIAN\_BUD\_BYTY – počet bytů obytných budov,
- AREAL\_INTERPOLATION – areálová interpolace,
- POCET\_OBYV\_RUIAN\_ADV – kombinace obytných adresních bodů a váženého počtu podlaží.

Bodové a polygonové vrstvy byly převedeny do rastrů pomocí nástrojů Point Density a Line Density. Číslo za názvem vrstvy pak značí nastavení rádiusu v nástroji.

Toto hodnocení bylo vyhotoveno pro úroveň podrobnosti z ORP do obcí a z obcí do ZSJ. Přepočtené hodnoty byly z prostředí ArcMap exportovány do formátu Microsoft Excel. V tomto prostředí byl za každý sloupeček s výslednými hodnotami přidán sloupec s absolutními chybami. Absolutní chyby byly na závěr sečteny a podle stanoveného způsobu hodnocení vyděleny dvěma. Tento součet byl poté vydělen celkovým počtem obyvatel a převeden na procenta. Výsledné hodnoty se nachází v XLS souboru dostupném na DVD (příloha číslo 18).

Nejprve byla zhodnocena vybraná velikost buňky. Pro výběr vhodné velikosti buňky bylo stanoveno procento špatně umístěných lidí pro každý typ pomocných dat. Poté byla vytvořena pořadí přesnosti metod pro každý typ pomocných dat. Součet těchto pořadí je pak kritériem pro výběr vhodné velikosti buňky (méně je lépe).

Typ pomocných dat	5m	10m	25m
POCET_OBYV_AW	1	2	3
POCET_OBYV_CLC	1	2	3
POCET_OBYV_OLM	2	1	3
POCET_OBYV_RUIAN_BUD1	1	2	3
POCET_OBYV_RUIAN_BUD5	1	2	3
POCET_OBYV_RUIAN_BUD10	1	2	3
POCET_OBYV_BUDVCHODY_POIDEN1	1	3	2
POCET_OBYV_BUDVCHODY_POIDEN5	1	2	3
POCET_OBYV_BUDVCHOD_POIDEN10	1	2	3
POCET_OBYV_OBYTNE_POIDEN1	2	3	1
POCET_OBYV_OBYTNE_POIDEN5	1	2	3
POCET_OBYV_OBYTNE_POIDEN10	1	2	3
POCET_OBYV_OSM_FULL1	2	1	3
POCET_OBYV_OSM_FULL5	1	2	3
POCET_OBYV_OSM_FULL10	1	2	3
POCET_OBYV_OSM_ROADS5	2	3	1
POCET_OBYV_OSM_ROADS10	1	2	3
POCET_OBYV_OSM_ROADS25	1	3	2
POCET_OBYV_OSM_MINOR5	3	2	1
POCET_OBYV_OSM_MINOR10	3	1	2
POCET_OBYV_OSM_MINOR25	3	2	1
POCET_OBYV_RUIAN_BUDVSE	1	2	3
POCET_OBYV_RUIAN_BUD_BYDLENI	2	1	3
POCET_OBYV_RUIAN_BUD_PODLAZI	1	2	3
POCET_OBYV_RUIAN_BUD_BYTY	1	2	3
<b>CELKEM</b>	<b>36</b>	<b>50</b>	<b>64</b>

Obr. 5.2.2.1 Hodnocení velikosti buňky při disagregaci počtu obyvatel

Na základě tohoto srovnání hodnocení byla vybrána jako vhodná velikost buňky buňka o straně 5 metrů. Na Obr. 5.2.2.1 je však patrné, že pro různé metody jsou vhodné různé velikosti buňky. Po stanovení vhodné velikosti buňky pro nejdetailnější úroveň byla poté tato velikost použita pro přepočítání z ORP do obcí a poté i pro přepočítání z územních jednotek do gridu pro následné vizualizace.

Dalším dílčím cílem bylo vybrat vhodnou metodu přepočtu. Pro tento účel byla převzata data za velikost buňky 5 metrů. Data byla v prostředí Microsoft Excel filtrována postupně podle jednotlivých ORP a pomocí funkce SUBTOTAL bylo zjištěno procento špatně umístěných obyvatel v každé z nich. Poté bylo zjištěno pořadí jednotlivých metod v každém z ORP a tato pořadí byla sečtena. Metoda s nejnižším součtem pořadí (viz Obr. 5.2.2.2.) pak byla vybrána jako nejvhodnější pro dané testovací území.

Pomocná data	Hranice	Jeseník	Konice	Lipník na	Litovel	Mohelnic	Olomouc	Prostějov	Přerov	Šumperk	Stemberk	Uničov	Záběh	SUM	POŘADÍ
Areálová interpolace	46,8 %	75,5 %	34,5 %	37,6 %	38,6 %	55,7 %	47,6 %	37,2 %	36,5 %	71,1 %	60,8 %	44,0 %	44,2 %	178	14
Areálové vážení	42,6 %	46,3 %	21,3 %	42,2 %	35,7 %	50,4 %	50,4 %	36,8 %	44,8 %	55,2 %	44,2 %	51,0 %	36,4 %	173	13
Sílniční síť - kompletní	29,2 %	35,2 %	18,8 %	35,4 %	26,7 %	39,6 %	33,0 %	21,1 %	27,8 %	40,7 %	31,0 %	35,2 %	26,4 %	151	12
Sílniční síť - očištěná	28,4 %	23,7 %	13,3 %	32,8 %	22,8 %	32,1 %	29,3 %	17,6 %	27,7 %	26,2 %	23,6 %	32,5 %	22,7 %	132	11
Open Land-Use Map	31,9 %	21,8 %	15,5 %	26,8 %	19,3 %	33,9 %	33,8 %	18,3 %	34,2 %	24,8 %	20,4 %	31,1 %	17,8 %	130	10
Sílniční síť - malé silnice	25,7 %	21,9 %	12,4 %	25,7 %	18,8 %	31,8 %	26,6 %	16,0 %	26,9 %	23,4 %	23,6 %	29,1 %	21,0 %	111	9
Budovy	25,9 %	19,6 %	11,0 %	23,0 %	16,9 %	25,8 %	28,1 %	18,8 %	26,3 %	15,5 %	22,6 %	37,0 %	17,2 %	108	8
CORINE LandCover	28,3 %	23,9 %	18,8 %	15,1 %	22,3 %	24,7 %	23,7 %	16,4 %	23,1 %	20,3 %	15,6 %	27,0 %	14,8 %	99	7
Adresní body	21,1 %	14,9 %	10,0 %	11,7 %	11,6 %	19,3 %	25,1 %	12,0 %	24,9 %	17,5 %	16,1 %	20,4 %	14,0 %	72	5
Obytné budovy	21,0 %	18,3 %	11,0 %	22,8 %	11,4 %	20,3 %	18,9 %	12,7 %	19,1 %	14,1 %	17,8 %	29,8 %	12,0 %	72	5
Obytné adresní body	20,8 %	14,1 %	6,8 %	10,3 %	11,4 %	18,5 %	23,3 %	11,7 %	22,1 %	15,2 %	14,0 %	19,3 %	12,6 %	51	4
Obytné budovy - byty	10,8 %	12,1 %	8,0 %	16,2 %	9,3 %	9,1 %	13,2 %	9,9 %	9,2 %	9,7 %	8,6 %	25,1 %	12,0 %	40	3
Obytné budovy - podlaží	9,5 %	11,7 %	9,1 %	18,5 %	5,9 %	11,3 %	14,1 %	5,4 %	6,1 %	6,5 %	10,9 %	17,0 %	6,5 %	33	2
Adresní body + podlaží ob. budov	5,3 %	9,4 %	4,3 %	10,0 %	5,6 %	6,0 %	11,3 %	4,2 %	5,4 %	7,1 %	8,4 %	12,5 %	6,9 %	15	1

Obr. 5.2.2.2 Hodnocení přesnosti pomocných dat

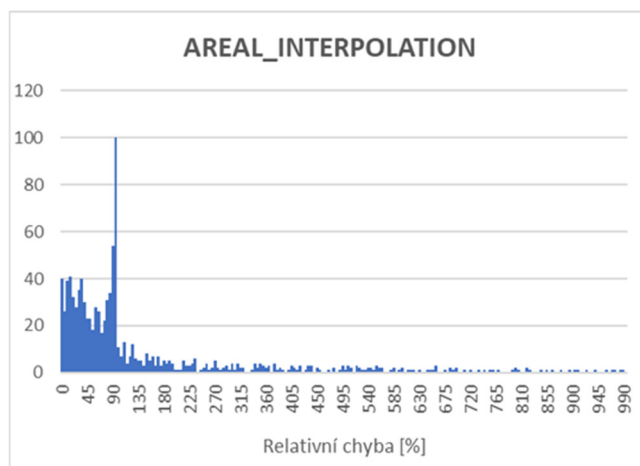
Tato tabulka byla vyhodnocena pro obě úrovně podrobnosti a v rámci ZSJ i pro všechny tři velikosti buněk. Výsledná data se nachází na DVD. Pro přepočítání z obcí do ZSJ i z ORP do obcí vyšla jednoznačně nejpřesnější kombinace adresních bodů a počtu podlaží obytných budov. Tato metoda byla poté využita pro následný přepočítání informací o počtu obyvatel z územních jednotek do gridu za účelem vizualizace statistických dat.

Dalším způsobem hodnocení bylo zobrazení četností relativních chyb do histogramů. Toto hodnocení bylo provedeno pouze pro úroveň podrobnosti z obcí do ZSJ. Pro toto hodnocení byla nejprve absolutní chyba podělena skutečným počtem obyvatel v ZSJ. Tato relativní četnost pak byla převedena na procenta a pro každý testovaný typ pomocných dat znázorněna v histogramu. Histogramy jsou dostupné na DVD jako příloha číslo 18.



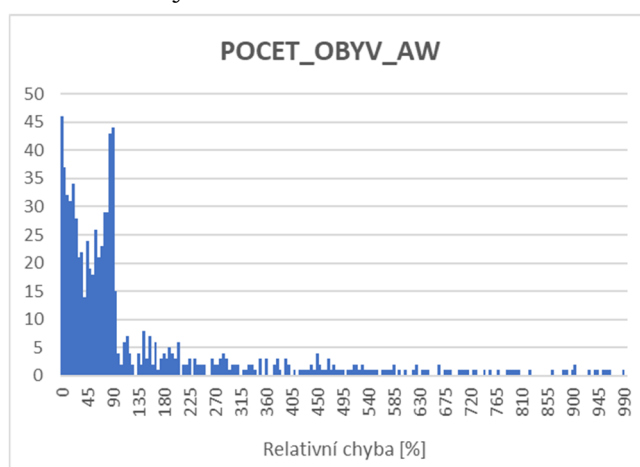
V histogramech byly znázorněny pouze chyby od 0 do 1000 %, ve všech vrstvách se však nachází ZSJ s vyššími chybami, než 1000 %.

Na Obr. 5.2.2.3 lze vidět, že dostupný nástroj Areal Interpolation vykazuje nejčastěji hodnotu chyby v okolí 100 %. Takovýto výsledek znamená chybu stejně velkou jako je počet obyvatel obce. Z výsledků je patrné, že tento nástroj nedokáže dobře modelovat jev v daném regionu i přes to, že zdrojové zóny byly dostupné na dostatečně detailní úrovni. Dále také bylo zjištěno, že po přepočtu pomocí tohoto nástroje 158 hodnot (13,4 % všech ZSJ) vykazuje relativní chybu vyšší, než 1 000 %. Na základě interpretace výsledků bylo zjištěno že se jedná o velmi nepřesnou metodu.



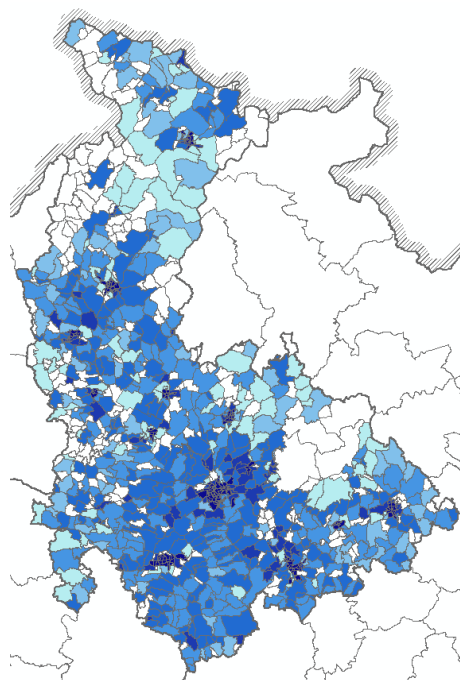
Obr. 5.2.2.3 Histogram relativních chyb areálové interpolace

Na Obr. 5.2.2.4 se nachází histogram relativních chyb pro areálové vážení. Areálové vážení předpokládá rovnoměrné rozložení jevu v prostoru. Obyvatelstvo v prostoru však není rozmístěno rovnoměrně, existují malé ZSJ s velkým množstvím obyvatel a pak velké ZSJ s nízkým počtem obyvatel. Z histogramu je patrné, že velké množství ZSJ vykazuje relativní chybu do 5 %, v tomto případě se však jedná o ZSJ územně shodné se zdrojovou obcí. Dále lze vidět, že velké množství chyb se pohybuje v okolí 100 %, stejně jako u areálové interpolace. Tyto chyby se vyskytují zejména v ZSJ s malou rozlohou a vysokým počtem obyvatel, popřípadě u větších ZSJ s menší hustotou zalidnění. Celkem 136 hodnot (11,6 %) vykazuje relativní chybu vyšší, než 1 000 %, což je způsobeno zejména malými a hustě osídlenými ZSJ.



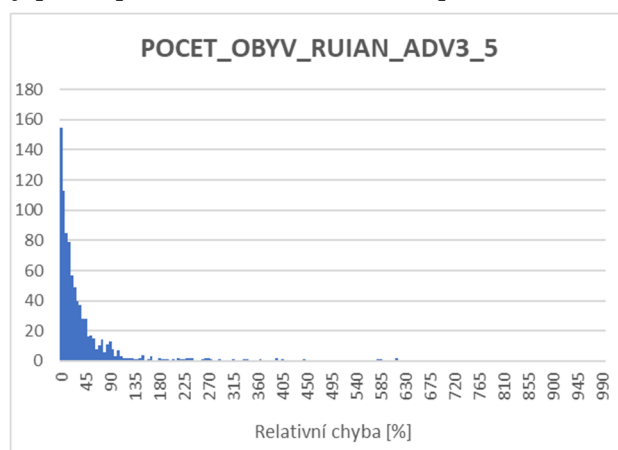
Obr. 5.2.2.4 Histogram relativních chyb areálového vážení

Jedním z typu hodnocených pomocných dat při přepočtu byla i datová sada CORINE Land Cover. Na Obr. 5.2.2.6 je vidět 260 ZSJ (22,2 %), které mají přepočtenou hodnotu obyvatel rovnu nule a přitom v nich žijí lidé. Na histogramu relativních chyb (viz příloha 18) lze vidět, že největší množství relativních chyb je okolo 100 %. Důvodem těchto chyb je absence obytného typu krajinného pokryvu v těchto lokalitách. Při minimální mapovací jednotce 25 ha (Copernicus, 2019) není možné zachytit veškeré zastavěné plochy, obzvláště pak v malých vesnicích, proto těmto územím nebyli přiřazeni žádní obyvatelé.



Obr. 5.2.2.5 Disagregace počtu obyvatel pomocí CORINE Land Cover

Na Obr. 5.2.2.6 se nachází histogram relativních chyb pro kombinaci adresních bodů a počtu podlaží obytných budov. Po důkladném prozkoumání dat bylo zjištěno, že relativní chybu do 50 % vykazuje 948 ZSJ (80,1 %), relativní chybu nad 1 000 % pouze devět ZSJ (0,77 %). Vysoké relativní chyby u těchto ZSJ mohou vzniknout v areálech s jednotkami obyvatel nebo vlivem kombinace datových sad z různých období – statistických dat ze SLDB 2011 a pomocných dat platných k březnu 2019. Bylo zjištěno, že do relativní chyby 5 % spadá 441 ZSJ, což je více než jedna třetina všech ZSJ. Takovéto výsledky lze pro účely práce považovat za dostatečně přesné.



Obr. 5.2.2.6 Histogram relativních chyb kombinace adresních bodů a budov

### Podíl nezaměstnaných osob

Pro přepočtení podílu nezaměstnaných osob byl použit postup popisovaný v kapitole 5.1.3. Vstupními parametry byla váha indexu dojížděky, znečištění ovzduší a indexu vzdělanosti. Data byla poté pomocí nástroje v rozhraní ModelBuilder (příloha 17) přepočtena do nižších územních jednotek. Přepočet probíhal v rámci Olomouckého kraje ve dvou úrovních podrobnosti – z kraje do obcí a z ORP do obcí. Druhou metodou bylo vytvoření regresní rovnice v statistickém softwaru GeoDa a tuto rovnici poté aplikovat na přepočet. Za účelem zjištění přesnosti byly testovány dva faktory – počet obcí s chybou přepočtené nezaměstnanosti do 1 % a absolutní chyba počtu nezaměstnaných v rámci celého kraje. Výsledné hodnoty jsou znázorněny v tabulce 5.2.2.1.

Tab. 5.2.2.1 Hodnocení přesnosti přepočtu podílu nezaměstnaných osob

Olomoucký kraj					
Úroveň podrobnosti	Váha indexu dojížděky	Váha znečištění ovzduší	Váha indexu vzdělanosti	Počet obcí s chybou < 1 %	Absolutní chyba počtu nezaměstnaných
z ORP do obcí	1	1	6	158	6073
	1	1	10	161	5774
	1	1	20	168	5494
	1	1	40	164	5325
	1	1	50	164	5287
	1	1	100	164	5207
	regrese				128
z kraje do obcí	1	1	10	120	7043
	1	1	20	124	6892
	1	1	30	124	6846
	1	1	40	124	6824
	1	1	50	125	6809
	1	1	100	121	6779
	regrese				128

Bylo zjištěno, že se zvyšováním váhy indexu vzdělanosti roste počet obcí s chybou nižší než 1 %. Dále se snižuje absolutní chyba počtu nezaměstnaných, což je způsobeno vyšším podílem nezaměstnaných osob ve velkých městech. Proto bylo testováno zejména navýšování váhy tohoto atributu. Bylo zjištěno, že nejvíce obcí s chybou přepočtené nezaměstnanosti nižší než 1 % vychází pro přepočet z ORP při váze 20 a z kraje při váze 50. Na základě získaných výsledků byla jako nejvhodnější váha indexu vzdělanosti pro přepočet v rámci testovaného území stanovena na hodnotu 50. Dále bylo zjištěno, že vhodná volba vah dosahuje lepších výsledků než vytvořená regresní rovnice. Kompletní výsledky jsou součástí přílohy 18.

### Průměrná hrubá mzda

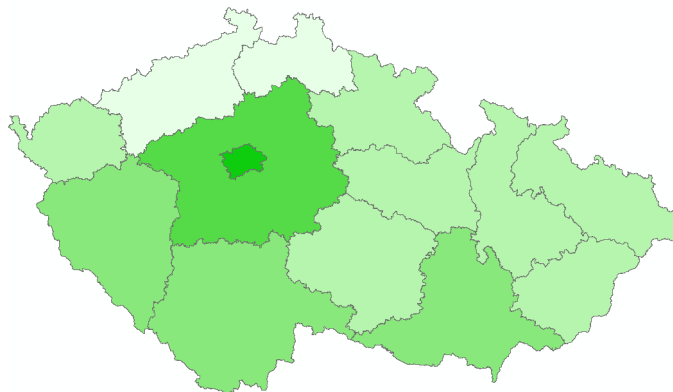
Pro přepočtení podílu nezaměstnaných osob byl použit postup popisovaný v kapitole 5.1.2. Vstupními parametry byla váha indexu vzdělanosti, podílu obyvatel zaměstnaných

v terciéru a míry zaměstnanosti. Data byla poté pomocí nástroje v rozhraní ModelBuilder (příloha 17) přepočtena do nižších územních jednotek. Přepočet probíhal v rámci České republiky a Olomouckého kraje ve třech úrovních podrobnosti. Na základě výpočtů v programu Microsoft Excel byly stanoveny váhy pro jednotlivé parametry (Obr. 5.2.2.7).

ATRIBUT	IND_VZD	TERCIER	MIRA_ZAM
VÁHA	1,00	1,00	0,5

Obr. 5.2.2.7 Váhy atributů

Použitím těchto vah byl poté proveden přepočet průměrné hrubé mzdy z krajů do okresů a poté z Olomouckého kraje do obcí. Z důvodu absence referenčních dat nebylo možné tyto výsledky validovat. Pro výpočet průměrné hrubé mzdy z úrovně státu do krajů byla také vytvořena v statistickém softwaru GeoDa regresní rovnice, pomocí níž byl poté přepočet proveden (viz Obr. 5.2.2.8). Bylo zjištěno, že průměrná chyba při přepočtu ze státu do krajů je u regresní rovnice menší než u přepočtených hodnot. Kompletní výsledky jsou k dispozici ve zdrojovém souboru Microsoft Excel v příloze 18.

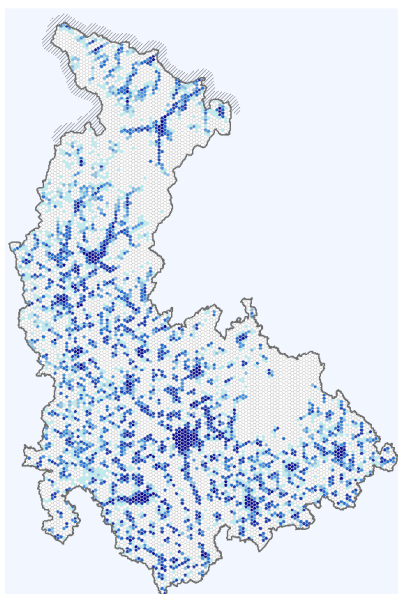


Obr. 5.2.2.8 Disagregace průměrné hrubé mzdy pomocí regrese

### 5.2.3 Vizualizace

V rámci práce bylo otestováno velké množství metod disagregace a různých pomocných dat, vytištěné však byly pouze nejzásadnější výsledky přepočtů. V případě detailnějšího studia je na DVD (příloha 18) k dispozici mapový projekt ve formátu MXD obsahující vizualizace všech otestovaných metod přepočtu. V rámci práce vzniklo 15 tištěných mapových výstupů, které jsou součástí příloh 1 až 15. V přílohách 1 až 7 jsou vizualizovány výsledky disagregací počtu obyvatel z úrovně obcí do ZSJ pomocí sedmi hlavních typů pomocných dat. Pro každá z pomocných dat byla vybrána nejpřesnější metoda přepočtu a ta byla zvizualizována. V příloze 8 se nachází přepočtené hodnoty počtu obyvatel z ORP do obcí pomocí nejlepší metody vybrané v kapitole 5.2.2.

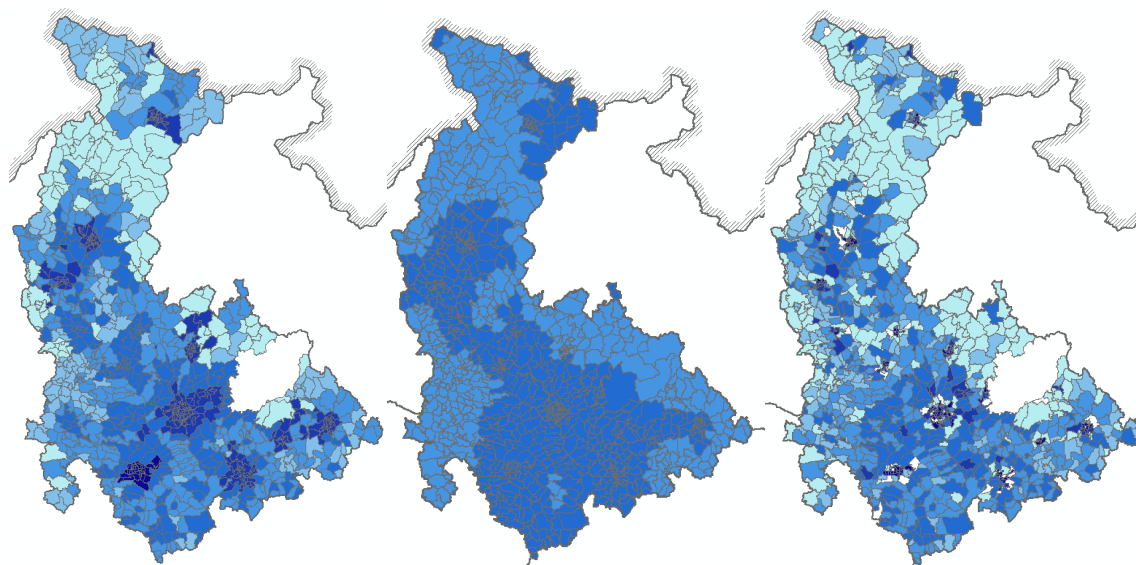
V přílohách 9 a 10 je zvizualizován přepočet počtu obyvatel pomocí stejné metody a stejné velikosti buňky v hexagonové síti (Obr. 5.2.3.1). Podle Zajícové (2017) hexagonová struktura svým obléjším tvarem výstižněji kopíruje hranici sledovaného území a zdá se být flexibilnější než čtvercová síť. Velikost buňky byla vybrána na základě doporučení autorky. Pro Olomoucký kraj v měřítku 1 : 500 000 byla vzata velikost buňky 0,5 km<sup>2</sup>; pro ORP Prostějov v měřítku 1 : 200 000 pak velikost buňky 0,1 km<sup>2</sup>. Obě velikosti buněk jsou dostatečně dobře čitelné při daném měřítku při vizualizaci na formátu A4. Hexagonová struktura o stanovené velikosti buňky byla vytvořena pomocí skriptu Generate Tesselation, dostupného v programu ArcMap.



Obr. 5.2.3.1 Vizualizace hustoty zaldnění v hexagonové síti

V přílohách 11 a 12 je zvizualizován přepočten podílu nezaměstnaných osob v rámci Olomouckého kraje ve dvou úrovních podrobnosti. Nejprve z úrovně kraje do obcí a poté z ORP do obcí. V příloze 13 a 14 se nachází výsledné mapy přepočtu průměrné hrubé mzdy na státní úrovni, v příloze 15 pak přepočten průměrné hrubé mzdy z kraje do obcí pro testovací území Olomouckého kraje.

Na Obr. 5.2.3.2 jsou znázorněny dvě v rámci testovacího území nejméně přesné metody disagregace (viz kapitola 5.2.2) a jejich vzájemné srovnání se skutečným stavem hustoty zaldnění. Na obrázku vlevo lze vidět, že areálové vážení rozpočítalo obyvatele v rámci jedné obce podle rozlohy dílčích ZSJ, proto mají všechna dílčí ZSJ stejnou hustotu zaldnění. Na obrázku uprostřed lze vidět disagregaci provedenou areálovou interpolací. Výsledné hustoty obyvatel se nachází mezi 50 a 250 obyv./km<sup>2</sup>. Lze pozorovat, že areálová interpolace metodou kriging výrazně shlazuje vstupní data. V tmavším pásu se sice skutečně více lidí nachází, ale jejich prostorová distribuce není zdaleka takto rovnoměrná. Obrázek vpravo poté znázorňuje skutečný stav počtu obyvatel.



Obr. 5.2.3.2 Srovnání přesnosti disagregačních metod (areálové vážení, areálová interpolace, skutečný stav)

## 6 VÝSLEDKY

Cílem práce bylo otestovat vybrané disagregační metody v prostředí GIS za účelem přepočtu vybraných statistických ukazatelů do nižších územních jednotek. Pro zpracování bakalářské práce byly nejprve vybrány statistické ukazatele, které byly předmětem testování – počet obyvatel, průměrná hrubá mzda a hustota zalidnění. Pro každý statistický ukazatel byla poté na základě rešerše a konzultace s vedoucím práce vybrána vhodná pomocná data. Pro každý ukazatel byla pak na základě dostupnosti zdrojových a referenčních dat definována testovací území a úrovně podrobnosti. Jako testovací území byl vybrán Olomoucký kraj, případně Česká republika.

Pro samotný přepočet byla nejprve otestována již existující softwarová řešení v prostředí GIS, poté byl pro každý statistický ukazatel vytvořen vlastní postup řešení, sestávající se z několika nástrojů. Pro automatizaci přepočtu byl vytvořen nástroj ve vizuálním programovacím rozhraní ModelBuilder. V rámci přepočtu byla testována přesnost výsledků na dostupných referenčních datech a doba trvání přepočtu. Celý proces výběru všech proměnných vstupujících do přepočtu je popsán v kapitole 4.

### 6.1 Srovnání metod

Na základě testování přesnosti přepočtu počtu obyvatel pomocí různých pomocných dat v úrovni podrobnosti ZSJ byla vybrána vhodná velikost strany buňky. Byly testovány buňky velikosti 5, 10 a 25 metrů. Jako vhodnější byla zvolena buňka, která v součtu za všechna použitá pomocná data vykazovala nejmenších procentuálních odchylek u špatně umístěných lidí. Ve většině případů vykazovala nejlepší hodnoty buňka velikosti strany 5 metrů (viz Obr. 5.2.2.1). Tato velikost buňky vykazovala nejhorší výsledky pouze pro přepočet pomocí lokálních silnic z OSM. Na základě těchto zjištění bylo 5 metrů vybráno jako vhodná velikost strany buňky.

Pro přepočet ostatních dvou ukazatelů byla vybírána vhodná velikost buňky na základě počtu nulových hodnot a počtu vytvořených buněk. Pro dosažení správné funkčnosti nástroje bylo potřeba, aby pro každou územní jednotku v cílové zóně byla vytvořena jedna rastrová buňka. Pokud toto pravidlo neplatí, může být dvěma obcím přiřazena stejná hodnota vlivem příliš hrubého rastru, popřípadě nemusí být obci přiřazena žádná hodnota. Toto pravidlo splňovaly buňky o straně:

- pro úroveň obcí - 0,5 km a menší,
- pro úroveň okresů - 2,5 km a menší,
- pro úroveň krajů - 25 km a menší.

Při disagregaci počtu obyvatel byly výsledky všech postupů srovnány v prostředí Microsoft Excel. Nejprve byl pro každý typ pomocných dat vytvořen sloupec absolutních hodnot chyb špatně umístěných lidí. Tyto chyby pak byly v rámci testovacího území sečteny a na základě metodiky hodnocení přesnosti disagregačních metod od Zandbergen (2011) vyděleny dvěma. Výsledná hodnota pak byla podělena počtem obyvatel v testovacím území. V rámci zhodnocení metod bylo vypočteno procento špatně umístěných lidí pro každé ORP testovacího území zvlášť a poté bylo pro každé ORP vytvořeno pořadí jednotlivých metod. Výsledné pořadí metod pak vzniklo součtem dílčích pořadí tak, že metoda s nejmenším součtem dílčích pořadí byla stanovena jako nejpřesnější v rámci testovaného území. Na Obr. 5.2.2.2 je stanoven žebříček přesnosti jednotlivých pomocných dat v rámci testovaných území. Nejlépe vyšla pro úroveň obcí i pro ZSJ pomocná data vzniklá kombinací obytných adresních bodů a počtu podlaží obytných budov vážených způsobem využití budovy. Nejhorší výsledky naopak zaznamenal nástroj Areal Interpolation střídavě s areálovým vážením.

Pro zhodnocení přesnosti přepočtu podílu nezaměstnaných osob v obcích Olomouckého kraje ve dvou různých úrovních podrobnosti byla data z programu ArcMap převedena do formátu XLS. Pro každou kombinaci vah pomocných dat byla vypočtena relativní odchylka vypočtené nezaměstnanosti od skutečné. Pro každou obec byl zjištěn počet obyvatel s trvalým pobytem v produkční skupině 15–64 let. Tímto počtem pak byl vynásoben skutečný i vypočtený podíl nezaměstnaných osob. Po odečtení těchto dvou hodnot od sebe a převedení rozdílu do absolutní hodnoty byla zjištěna absolutní chyba počtu nezaměstnaných obyvatel. V rámci hodnocení vah pomocných dat byla sledována dvě kritéria – počet obcí s relativní odchylkou nižší než 1 % a součet absolutních chyb počtu nezaměstnaných ve všech obcích. Na základě těchto kritérií byly stanoveny nejhodnější váhy jednotlivých atributů pro testovací území:

- Index dojížděky – váha 1,
- Znečištění ovzduší – váha 1,
- Index vzdělanosti – váha 50.

Pro přepočet průměrné hrubé mzdy byly testovány váhy pomocných dat pouze pro přepočet ze státu do krajů, v nižších úrovních nejsou referenční data dostupná. Volba atributů použitých při přepočtu je popsána v kapitole 5.2.2. Na základě výpočtů v programu Microsoft Excel byly stanoveny váhy pro jednotlivé parametry (viz Obr. 5.2.2.7). Pro řešení byla dále vytvořena v statistickém softwaru GeoDa regresní rovnice, pomocí níž byl poté proveden přepočet z úrovně státu do krajů.

## 6.2 Vizualizace

Pro testované metody byly vytvořeny vizualizace všech provedených přepočtů. Celkem bylo zhotoveno 156 vizualizací, které jsou uloženy v mapovém projektu ArcGIS Desktop dostupném na DVD v příloze 18. Výsledné vizualizace jsou přiloženy v jedné geodatabázi včetně všech zdrojových dat. Výsledné vrstvy jsou v datovém rámci pro přehlednost seskupeny podle typu statistického ukazatele, testovaného území a velikosti strany buňky. Vizualizace jsou jedním z výstupů práce, postup a jejich nastavení je popsán v kapitole 5.

Výsledné vrstvy byly vizualizovány tak, aby byly po doplnění zbylých základních kompozičních prvků mapy ihned k použití. Všem vrstvám disagregovaného počtu obyvatel byly stanoveny stupnice o šesti intervalech, hodnota nula byla z vizualizace vynechána. Intervaly jsou v rámci testovaného ukazatele nastaveny jednotně pro všechny úrovně podrobnosti, typy pomocných dat a velikost strany buňky. Pro průměrnou hrubou mzdu byla v úrovni krajů stanovena stupnice o šesti intervalech, v úrovni okresů o sedmi intervalech a v úrovni obcí o osmi intervalech. Menší množství intervalů v úrovni krajů bylo voleno z důvodu malého množství hodnot, větší množství intervalů pro obce pak bylo voleno z důvodu dobrého znázornění většího rozpětí hodnot míru. Pro průměrnou hrubou mzdu byla barevná stupnice nastavena pro různé úrovně podrobnosti různě. Pro podíl nezaměstnaných osob byla volena stupnice o sedmi intervalech. Jako hranice intervalů jsou zvoleny pouze celá čísla, která jsou pro uživatele mapy představitelná. Stupnice jsou zhotoveny sekvenčně ve třech barevných provedeních. K vrstvám byly do výsledného projektu přidány administrativní hranice a velká města ve studované oblasti. Příklady některých vizualizací jsou přiloženy jako tištěné vázané přílohy.

## 7 DISKUZE

Cílem práce bylo otestovat vybrané disgregační metody na několika statistických ukazatelích a výsledky zhodnotit na referenčních datech. Úkolem bylo navrhnout a otestovat různé typy pomocných dat pro vybraná testovací území a stanovené velikosti buněk.

První problém při práci nastal při analýze datového zdroje adresních bodů „Statistické budovy (vchody k bytům)“ poskytnutém autorovi Katedrou geoinformatiky ve formátu ESRI Shapefile. Bodová vrstva obsahovala celkem 2 809 711 adresních bodů a 63 atributů. Bylo zjištěno, že celkem 1 988 563 záznamů (71 %) obsahovalo v atributu NAZ\_OBEC město Choceň a většina atributů obsahovala nulové hodnoty. Po důkladnější analýze však bylo zjištěno, že chybnou částí vrstvy je pouze atributová část, prostorová lokalizace adresních bodů byla v pořádku. Data nebyla nijak upravována, pro potřeby práce byla využita pouze geometrie vrstvy.

Další dva problémy nastaly při využití databáze RÚIAN. Po důkladném zkoumání dat bylo zjištěno, že 138,59 km<sup>2</sup> (2,36 %) rozlohy Olomouckého kraje databáze RÚIAN nepokrývá (viz Obr. 4.3.1.1). V těchto místech je v databázi k dispozici pouze bodová vrstva adresních bodů. Dále bylo zjištěno, že požadované atributy ve vrstvě stavebních objektů nemají ve všech případech správně uvedené hodnoty. Před použitím dat pro práci bylo nejprve potřeba opravit alespoň evidentní chyby. Nejextrémnější hodnoty počtu podlaží a bytů byly poté opraveny podle údajů podobných stavebních objektů v okolí nebo vizuální interpretací panoramatických snímků od Mapy.cz. Pro přepočty pomocí kombinace adresních bodů a obytných budov pak bylo potřeba, aby stavební objekty sloužící k bydlení měly alespoň jedno podlaží. Počet podlaží byl tedy u všech obytných budov s nula podlažími stanoven na jedno.

Největší časovou a zejména datovou náročnost představoval převod pomocných dat do rastru o velikosti strany buňky 5 metrů. Některá pomocná data byla převáděna i více než 10 minut a výsledná testovací geodatabáze se všemi vstupními rastry měla po dokončení všech analýz 44,9 GB. Z tohoto důvodu také nemohly být vstupní rastry začleněny do příloh práce. Řešením těchto problémů by mohlo být začlenění tvorby vstupního rastru do nástroje. Toto sice časovou náročnost přepočtu nesníží, celý přepočty však ve výsledku zanechá jen nový řádek v atributové tabulce, což by ušetřilo místo na disku. Pro potřeby práce však bylo potřeba některé rastry využít vícekrát, proto nebyla tato myšlenka zrealizována.

Při důkladném zkoumání výsledků jednotlivých přepočtů počtu obyvatel bylo zjištěno, že některé ZSJ, které jsou součástí jedné obce, obsahují přičtené hodnoty z jiné obce. Toto je způsobeno přiřazováním některých buněk rastru z jedné do druhé obce. Závěrečná zonální statistika přičte do jedné obce i hraniční buňky s přiřazeným počtem obyvatel z jiné obce. Možností řešení tohoto problému by mohlo být snížení velikosti buňky rastru. Tato varianta je ale výpočetně, časově i datově náročná. Druhou možností je vytvoření nástroje na disgregaci statistických ukazatelů ve vektorovém prostředí. Tento přístup by umožnil zachovat pomocná data ve vektorové podobě, to by kromě ušetření času a místa na disku mohlo přinést přesnější výsledky přepočtů. Vektorová disgregace je přístup, který není prozatím příliš zdokumentovaný a může být předmětem dalšího studia.

V práci bylo v rámci snahy o dosažení co největší přesnosti testováno velké množství pomocných dat a různých metod jejich využití. Některým atributům byly také přiřazovány váhy pro přepočty. Z časových důvodů bylo možné otestovat jen omezené množství zvolených pomocných dat a jejich vah. Předmětem dalšího zkoumání může být nejen testování disgregace pro další statistické ukazatele, ale i využití jiných dostupných



pomocných dat pro přepočtení. Bylo by možné se zaměřit například na jeden konkrétní statistický ukazatel, pro něj otestovat větší množství pomocných dat a jejich kombinací, pro každý typ pomocných dat pak vybírat různé váhy tak, aby bylo dosaženo co největší přesnosti. Tyto výsledky by poté bylo vhodné validovat na větším testovacím území pro dosažení co největší přesnosti odhadu a vytvoření univerzálního nástroje pro disagregaci. V rámci hodnocení výsledků přesnosti by také bylo možné zjistit místa, kde nástroj hodnoty statistického jevu přičítá a kde naopak odečítá. Z těchto informací by šlo vyvodit závěr, že například v městských oblastech nástroj výrazně podhodnocuje přepočtené hodnoty statistických jevů.

Nástroj Areal Interpolation dostupný jako součást Geostatistical Wizard programu ArcMap během práce nevykazoval příliš přesné výsledky pro přepočtení počtu obyvatel z obcí do ZSJ. Nástroj byl testován i pro přepočtení míry nezaměstnanosti, v rámci přepočtu z úrovně ORP do obcí měl však nástroj nedostatek zdrojových dat pro výpočet. Nadstavbou této práce může být otestování nástroje Areal Interpolation pro více různých typů statistických dat. Další možností je přidání druhé datové sady do nástroje jako druhou proměnnou statistické metody cokriging. Pokud by do nástroje byla dodána vhodná pomocná data, mohla by výrazně zvýšit přesnost výsledků areálové interpolace.

V rámci návaznosti na práci by bylo také možné navázat spolupráci s ČSÚ a otestovat jejich dostupné přesnější a složitější nástroje pro disagregaci různých statistických ukazatelů.

V práci byly splněny všechny dílčí cíle a výsledky mohou sloužit jako doporučení při disagregaci testovaných nebo případně i dalších statistických dat. Nedostatky a návrhy zmíněné v diskuzi mohou být předmětem dalšího studia.

## 8 ZÁVĚR

Hlavním cílem bakalářské práce bylo provést testování vybraných disgregačních metod za účelem přepočtu vybraných statistických ukazatelů do nižších územních jednotek. Testovanými statistickými ukazateli byl počet obyvatel, míra nezaměstnanosti a průměrná hrubá mzda. Úkolem bylo navrhnout a otestovat různé metody přepočtu informací a různá pomocná data do něj vstupující. Výsledky bylo poté třeba patričně vizualizovat a porovnat.

Teoretickou částí bylo nastudovat příklady v minulosti již použitých disgregací, různé typy pomocných dat použité při přepočtu a kartografické možnosti vizualizace výstupů. Dalším cílem bylo studium existujících softwarových možností pro disgregaci a export dat z databáze RÚIAN. Možnosti a vybrané příklady jsou popsány v kapitolách 2 a 3.

Na základě konzultace s vedoucím práce byly vybrány statistické ukazatele pro přepočet do nižších územních jednotek. Pomocí rešerše byla vybrána vhodná pomocná data pro přepočet a nástroje pro jejich stažení. Dále byl navržen postup pro přepočet všech statistických ukazatelů a byly vytvořeny automatizované nástroje na jejich disgregaci. Na základě dostupnosti primárních dat byla definována vhodná testovací území, pro které byl přepočet proveden.

Dílčím výsledkem práce byl návrh nejvhodnějších pomocných dat pro každý statistický ukazatel a nejvhodnější velikost strany buňky použité při přepočtu. Pro přepočet počtu obyvatel byla vybrána jako nejvhodnější kombinace adresních budov a počtu podlaží bytů. Jako vhodná velikost buňky byla na základě srovnání vybrána délka strany 5 metrů. Tato varianta přináší pro testovaná území nejpřesnější výsledky. Pro přepočet ostatních statistických ukazatelů byla v rámci testovacích území stanovena jako doporučená velikost strany buňky 500 metrů. Jako pomocná data pro přepočet průměrné hrubé mzdy byl vybrán index vzdělanosti, podíl obyvatel zaměstnaných v terciéru a míra zaměstnanosti. Pro disgregaci podílu nezaměstnaných osob byly vybrány atributy index dojíždky, míra znečištění ovzduší a index vzdělanosti. Zvolené váhy všech atributů jsou popsány v kapitole 5.2.2. Výsledné přepočty byly zhodnoceny na dostupných referenčních datech a poté vizualizovány.

Závěrem byla sepsána textová část práce. Výsledky práce jsou prezentovány na webových stránkách bakalářské práce a na informačním posteru. Celá práce (textová část, přílohy, výstupy, zdrojová a vytvořená data) je dostupná v příloze 18 na DVD.

## POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

CORINE Land Cover — Copernicus Land Monitoring Service [online]. 2019 [cit. 2019-04-27]. Dostupné z: <https://land.copernicus.eu/pan-european/corine-land-cover>

ČSÚ. *Zaměstnanost a nezaměstnanost podle výsledků VŠPS - Metodika* [online]. 2019 [cit. 2019-04-29]. Dostupné z: [https://www.czso.cz/csu/czso/zam\\_vsps](https://www.czso.cz/csu/czso/zam_vsps)

*Dasymetric Toolbox | EnviroAtlas | US EPA* [online]. 2016 [cit. 2019-04-25]. Dostupné z: <https://www.epa.gov/enviroatlas/dasymetric-toolbox>

EICHER, Cory L. a Cynthia A. BREWER. Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation. *Cartography and Geographic Information Science* [online]. 2001, 28(2), 125-138 [cit. 2019-05-09]. DOI: 10.1559/152304001782173727. ISSN 1523-0406. Dostupné z: <http://www.tandfonline.com/doi/abs/10.1559/152304001782173727>

FLOWERDEW, Robin, Mick GREEN a Evangelos KEHRIS. Using areal interpolation methods in geographic information systems. *Papers in Regional Science* [online]. 1991, 70(3), 303-315 [cit. 2019-05-09]. DOI: 10.1007/BF01434424. ISSN 1056-8190. Dostupné z: <http://doi.wiley.com/10.1007/BF01434424>

HOLT, James B., C.P. LO a Thomas W. HODLER. Dasymetric Estimation of Population Density and Areal Interpolation of Census Data. *Cartography and Geographic Information Science* [online]. 2004, 31(2), 103-121 [cit. 2019-05-09]. DOI: 10.1559/1523040041649407. ISSN 1523-0406. Dostupné z: <http://www.tandfonline.com/doi/abs/10.1559/1523040041649407>

HUANG, Zhengdong, Henk OTTENS a Ian MASSER. A Doubly Weighted Approach to Urban Data Disaggregation in GIS: A Case Study of Wuhan, China. *Transactions in GIS* [online]. 2007, 11(2), 197-211 [cit. 2019-05-09]. DOI: 10.1111/j.1467-9671.2007.01041.x. ISSN 1361-1682. Dostupné z: <http://doi.wiley.com/10.1111/j.1467-9671.2007.01041.x>

KLAUDA, Petr. Prostorově určená statistická data. *Statistika&My*. Praha: ČSÚ, 2016, 6(5), 18 - 19. ISSN 1804-7149.

KRIVORUCHKO, Konstantin, Alexander GRIBOV a Eric KRAUSE. Multivariate Areal Interpolation for Continuous and Count Data. *Procedia Environmental Sciences* [online]. 2011, 3, 14-19 [cit. 2019-05-09]. DOI: 10.1016/j.proenv.2011.02.004. ISSN 18780296. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1878029611000053>.

LI, Tiebei, David PULLAR, Jonathan CORCORAN a Robert STIMSON. *A comparison of spatial disaggregation techniques as applied to population estimation for South East Queensland (SEQ), Australia* [online]. 2007 [cit. 2019-02-12]. DOI: 10.4225/03/57E9AECEBA789. Dostupné z:

[https://figshare.com/articles/A\\_comparison\\_of\\_spatial\\_disaggregation\\_techniques\\_as\\_applied\\_to\\_population\\_estimation\\_for\\_South\\_East\\_Queensland\\_SEQ\\_Australia/385980](https://figshare.com/articles/A_comparison_of_spatial_disaggregation_techniques_as_applied_to_population_estimation_for_South_East_Queensland_SEQ_Australia/385980)

MILEU, Nelson a Margarida QUEIRÓS. Development of a QGIS Plugin to Dasymetric Mapping. 2018. Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings: Vol. 18 , Article 9. DOI: <https://doi.org/10.7275/3628-0a51> Available at: <https://scholarworks.umass.edu/foss4g/vol18/iss1/9>

*O Olomouckém kraji | Olomoucký kraj* [online]. 2019 [cit. 2019-04-25]. Dostupné z: <https://www.olkraj.cz/o-olomouckem-kraji-cl-1362.html>

*Open Land-Use Map* [online]. 2015 [cit. 2019-04-27]. Dostupné z: [https://sdi4apps.eu/open\\_land\\_use/](https://sdi4apps.eu/open_land_use/)

ROSINA, Konstantin a Pavol HURBÁNEK. Na báze európskych služieb pre monitoring krajiny [online]. *Kartografické listy*, 2012, 20(2), 56–65. Dostupné z: <https://gis.fns.uniba.sk/kartografickelisty/archiv/KL20/14.pdf>

STARÝ, Vlastimil. Prostorová disagregace s využitím dat LU/LC [online]. VŠB - TU Ostrava, 2012 [cit. 2019-02-08]. Dostupné z WWW: [http://gisak.vsb.cz/GISacek/GISacek\\_2012/papers/stary.pdf](http://gisak.vsb.cz/GISacek/GISacek_2012/papers/stary.pdf)

*VFR Import - Geografické informační systémy (GIS) - ARCDATA PRAHA* [online]. 2019 [cit. 2019-04-24]. Dostupné z: <https://www.arcdata.cz/produkty/software-arcdata/vfr-import>

VOŽENÍLEK, Vít a Jaromír KAŇOK. *Metody tematické kartografie: vizualizace prostorových jevů*. Olomouc: Univerzita Palackého v Olomouci pro katedru geoinformatiky, 2011. ISBN 9788024427904.

*What is areal interpolation?* [online]. 2019 [cit. 2019-04-24]. Dostupné z: <http://desktop.arcgis.com/en/arcmap/latest/extensions/geostatistical-analyst/what-is-areal-interpolation.htm>

ZAJÍCOVÁ, Hana. *Srovnání Prostorové Reprezentace Adresních Bodů – Grid/areály*. Olomouc, 2017. Bakalářská práce. Univerzita Palackého v Olomouci.

ZANDBERGEN, Paul A. Dasymetric Mapping Using High Resolution Address Point Datasets. *Transactions in GIS* [online]. 2011, 15, 5-27 [cit. 2019-02-15]. DOI: 10.1111/j.1467-9671.2011.01270.x. ISSN 13611682. Dostupné z: <http://doi.wiley.com/10.1111/j.1467-9671.2011.01270.x>

BIELECKA, Elzbieta. (2005). *A Dasymetric Population Density Map of Poland*. Proceedings of the 22nd International Cartographic Conference, no. 48 22: 9–15.

ZANDBERGEN, Paul A. a Drew A. IGNIZIO. Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates. *Cartography and Geographic Information Science* [online]. 2010, 37(3), 199-214 [cit. 2019-05-09]. DOI: 10.1559/152304010792194985. ISSN 1523-0406. Dostupné z: <http://www.tandfonline.com/doi/abs/10.1559/152304010792194985>

*Zásuvný modul QGIS pro stahování datových vrstev RÚIAN - Dokumentace pro QGIS RUIAN Plugin 1.0.0* [online]. 2016 [cit. 2019-04-24]. Dostupné z: <https://ctu-geoforall-lab.github.io/qgis-ruian-plugin/>

## **PŘÍLOHY**

# SEZNAM PŘÍLOH

## Vevázané přílohy

Příloha 1	Hustota zalidnění v ZSJ Olomouckého kraje v roce 2011 (areálová interpolace)
Příloha 2	Hustota zalidnění v ZSJ Olomouckého kraje v roce 2011 (areálové vážení)
Příloha 3	Hustota zalidnění v ZSJ Olomouckého kraje v roce 2011 (hustota silniční sítě)
Příloha 4	Hustota zalidnění v ZSJ Olomouckého kraje v roce 2011 (krajinný pokryv)
Příloha 5	Hustota zalidnění v ZSJ Olomouckého kraje v roce 2011 (obytné adresní body)
Příloha 6	Hustota zalidnění v ZSJ Olomouckého kraje v roce 2011 (podlaží obytných budov)
Příloha 7	Hustota zalidnění v ZSJ Olomouckého kraje v roce 2011 (podlaží obytných budov a obytné adresní body)
Příloha 8	Hustota zalidnění v obcích Olomouckého kraje v roce 2011
Příloha 9	Hustota zalidnění v Olomouckém kraji v roce 2011 (z obcí do hexagonové sítě)
Příloha 10	Hustota zalidnění v Olomouckém kraji v roce 2011 (ze ZSJ do hexagonové sítě)
Příloha 11	Podíl nezaměstnaných osob v obcích Olomouckého kraje v roce 2016 (z kraje do obcí)
Příloha 12	Podíl nezaměstnaných osob v obcích Olomouckého kraje v roce 2016 (z ORP do obcí)
Příloha 13	Průměrná hrubá mzda v krajích České republiky v roce 2018
Příloha 14	Průměrná hrubá mzda v okresech České republiky v roce 2018
Příloha 15	Průměrná hrubá mzda v obcích Olomouckého kraje v roce 2018
Příloha 16	Schéma nástroje na disagregaci počtu obyvatel
Příloha 17	Schéma nástroje na disagregaci průměrné hrubé mzdy

## Volné přílohy

Příloha 18	DVD
Příloha 19	Poster
Příloha 20	Histogramy četností relativních chyb

## Popis struktury DVD

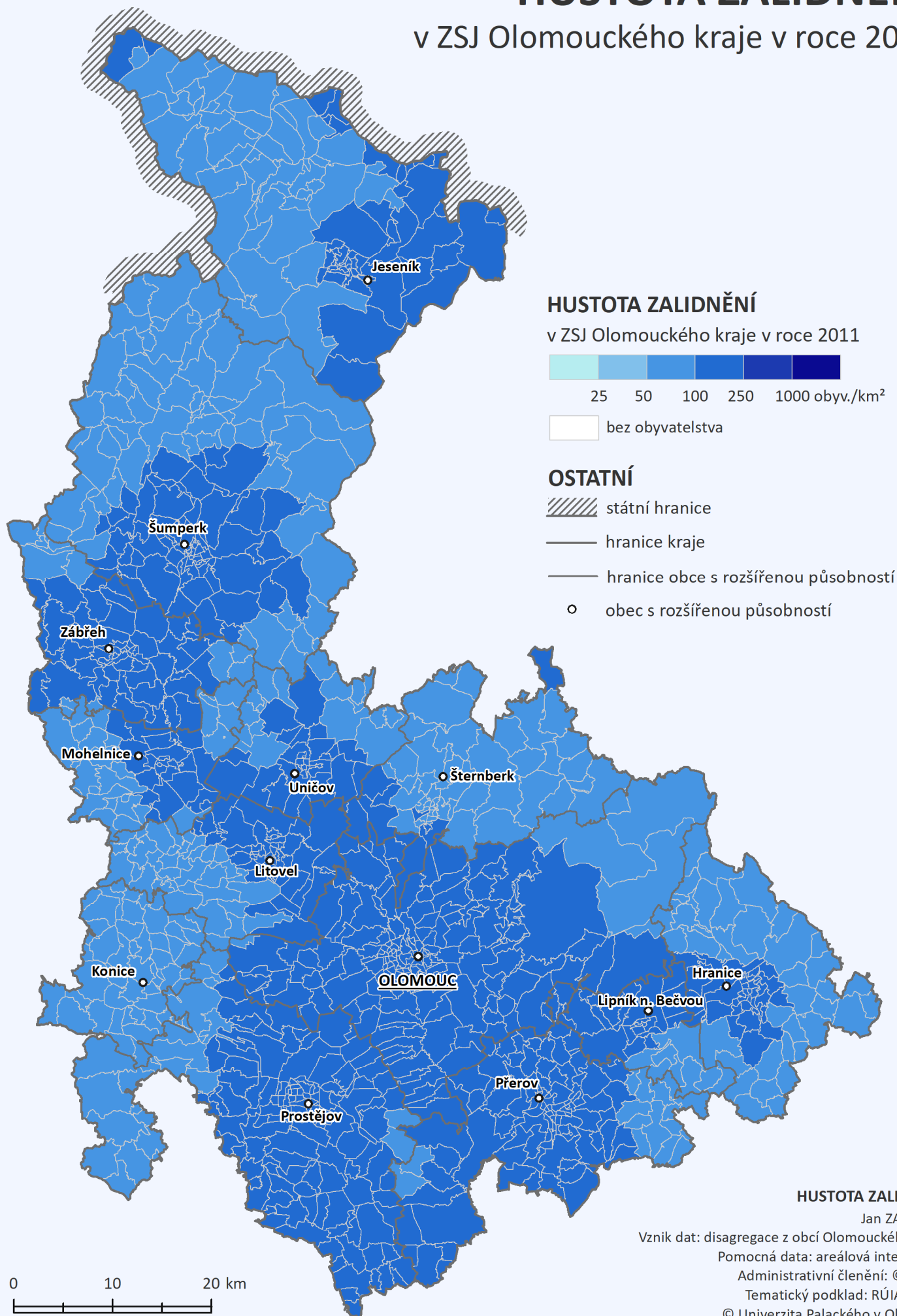
### Adresáře:

Data  
data.gdb  
Disagregace.tbx  
vizualizace.mxd

Přilohy  
Text\_prace  
Web

# HUSTOTA ZALIDNĚNÍ

v ZSJ Olomouckého kraje v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace z obcí Olomouckého kraje

Pomocná data: areálová interpolace

Administrativní členění: © ArcČR

Tematický podklad: RÚIAN, ČSÚ

© Univerzita Palackého v Olomouci

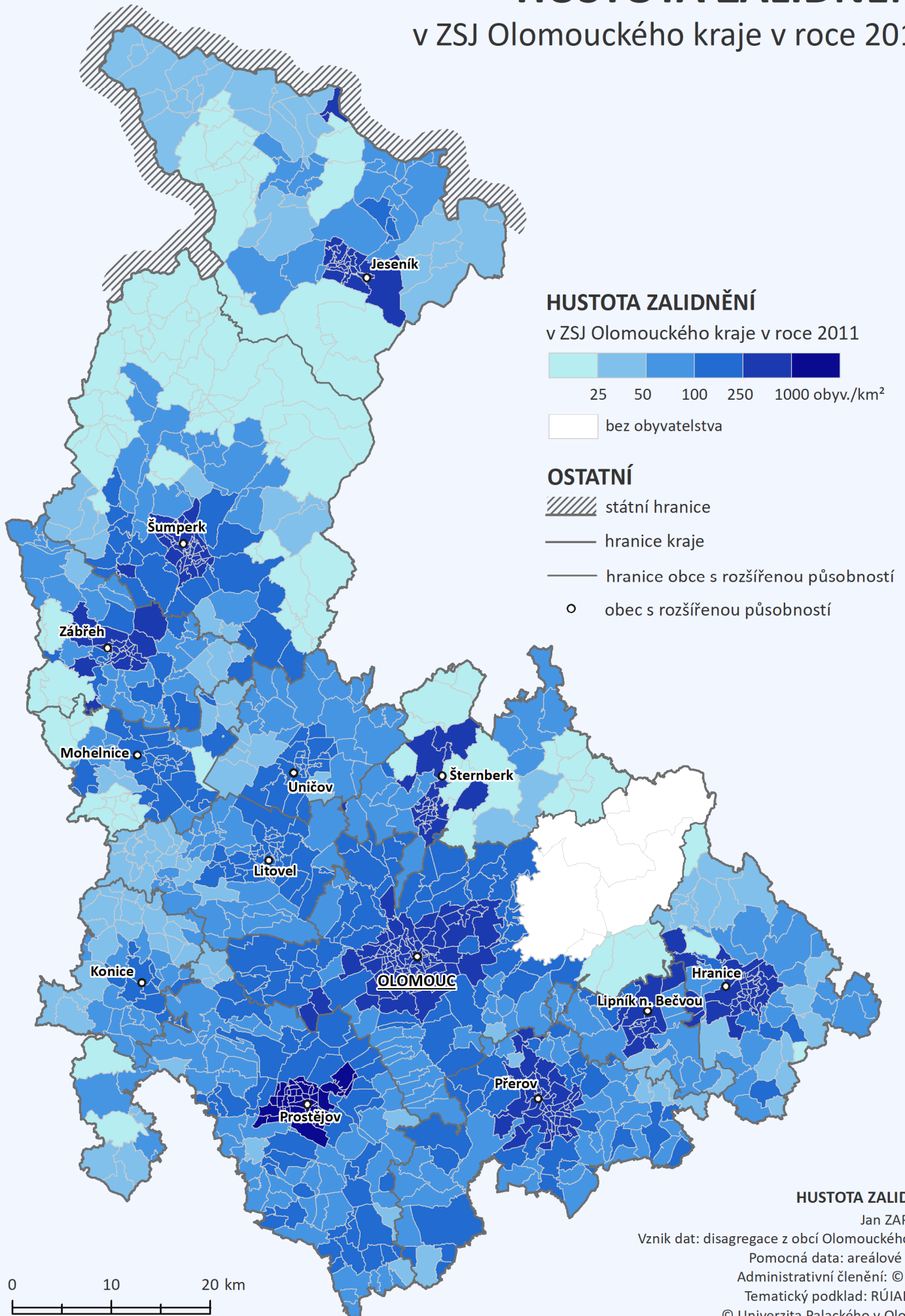
Coordinate system: WGS 1984 UTM Zone 33N

Příloha 1



# HUSTOTA ZALIDNĚNÍ

v ZSJ Olomouckého kraje v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace z obcí Olomouckého kraje

Pomocná data: areálové vážení

Administrativní členění: © ArcČR

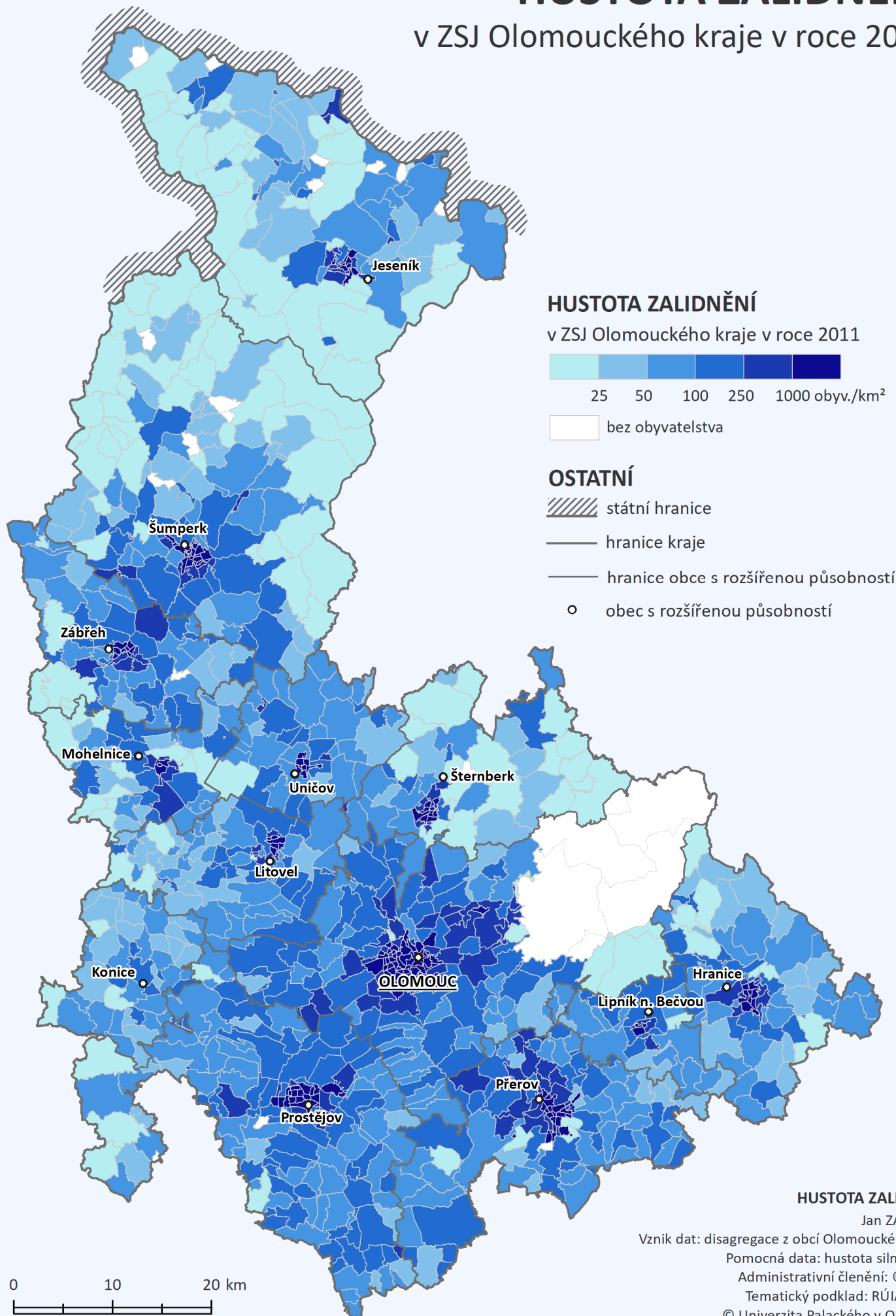
Tematický podklad: RÚJAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

# HUSTOTA ZALIDNĚNÍ

v ZSJ Olomouckého kraje v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace z obcí Olomouckého kraje

Pomocná data: hustota silniční sítě

Administrativní členění: © ArcČR

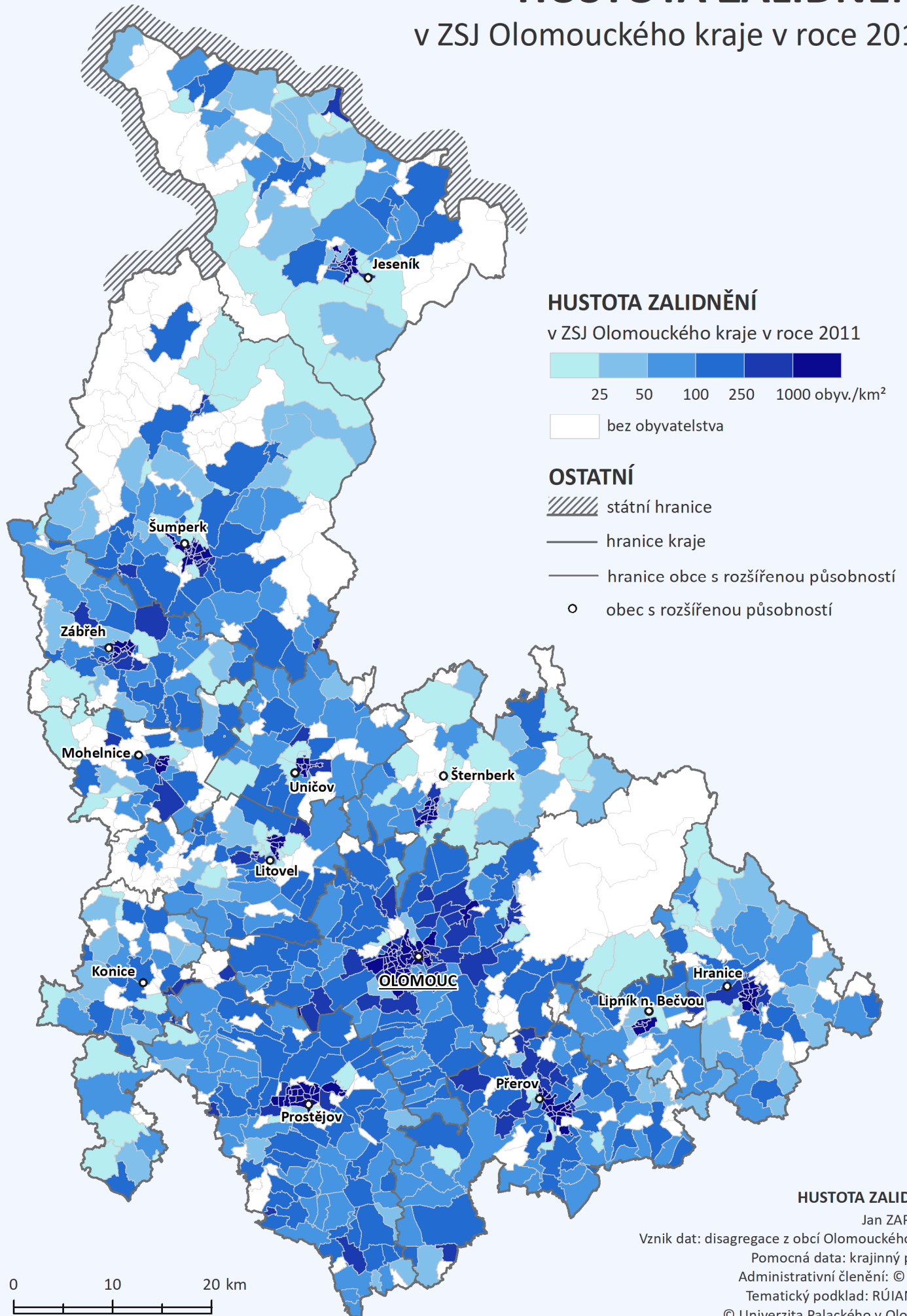
Tematický podklad: RÚIAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

# HUSTOTA ZALIDNĚNÍ

v ZSJ Olomouckého kraje v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace z obcí Olomouckého kraje

Pomocná data: krajinný pokryv

Administrativní členění: © ArcČR

Tematický podklad: RÚIAN, ČSÚ

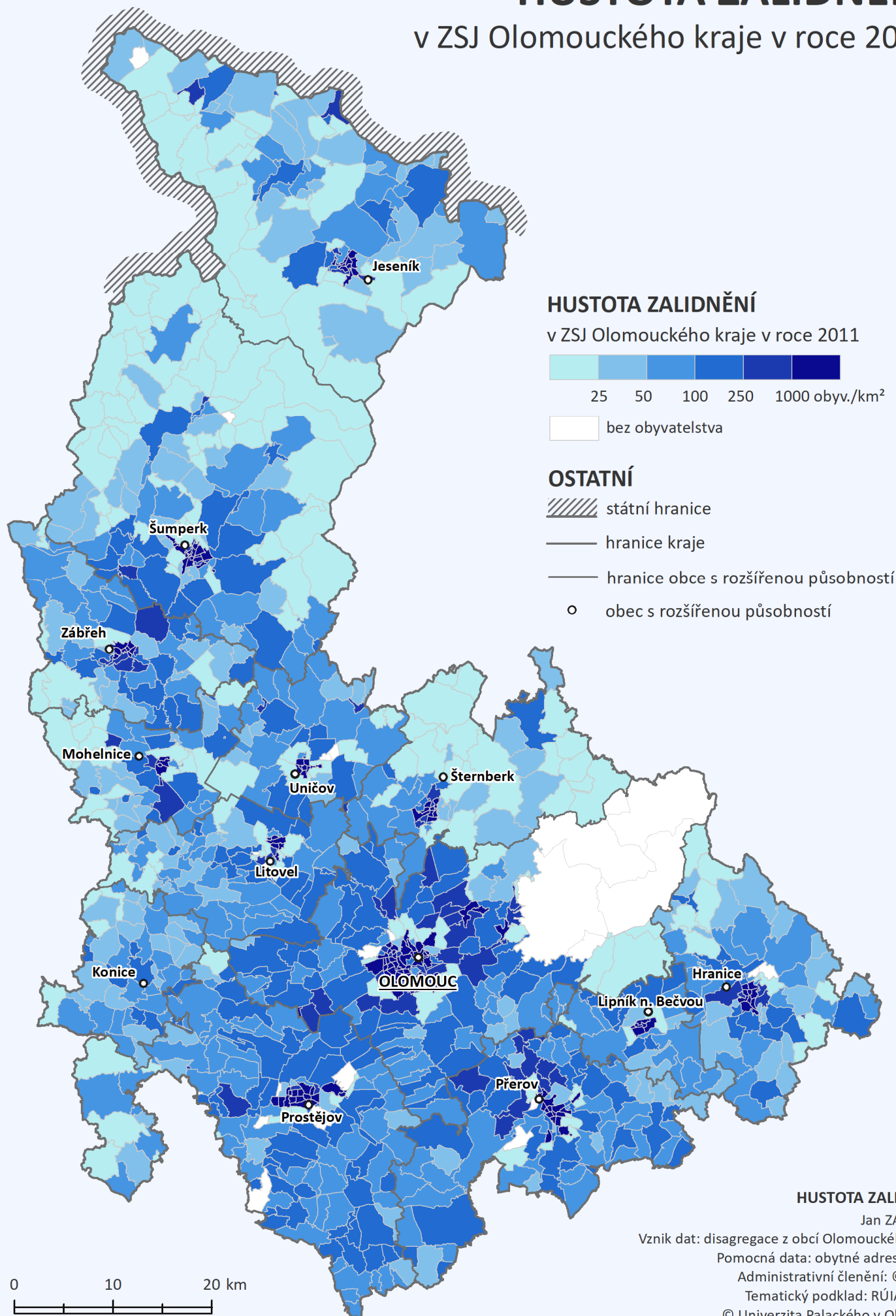
© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N



# HUSTOTA ZALIDNĚNÍ

v ZSJ Olomouckého kraje v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace z obcí Olomouckého kraje

Pomocná data: obytné adresní body

Administrativní členění: © ArcČR

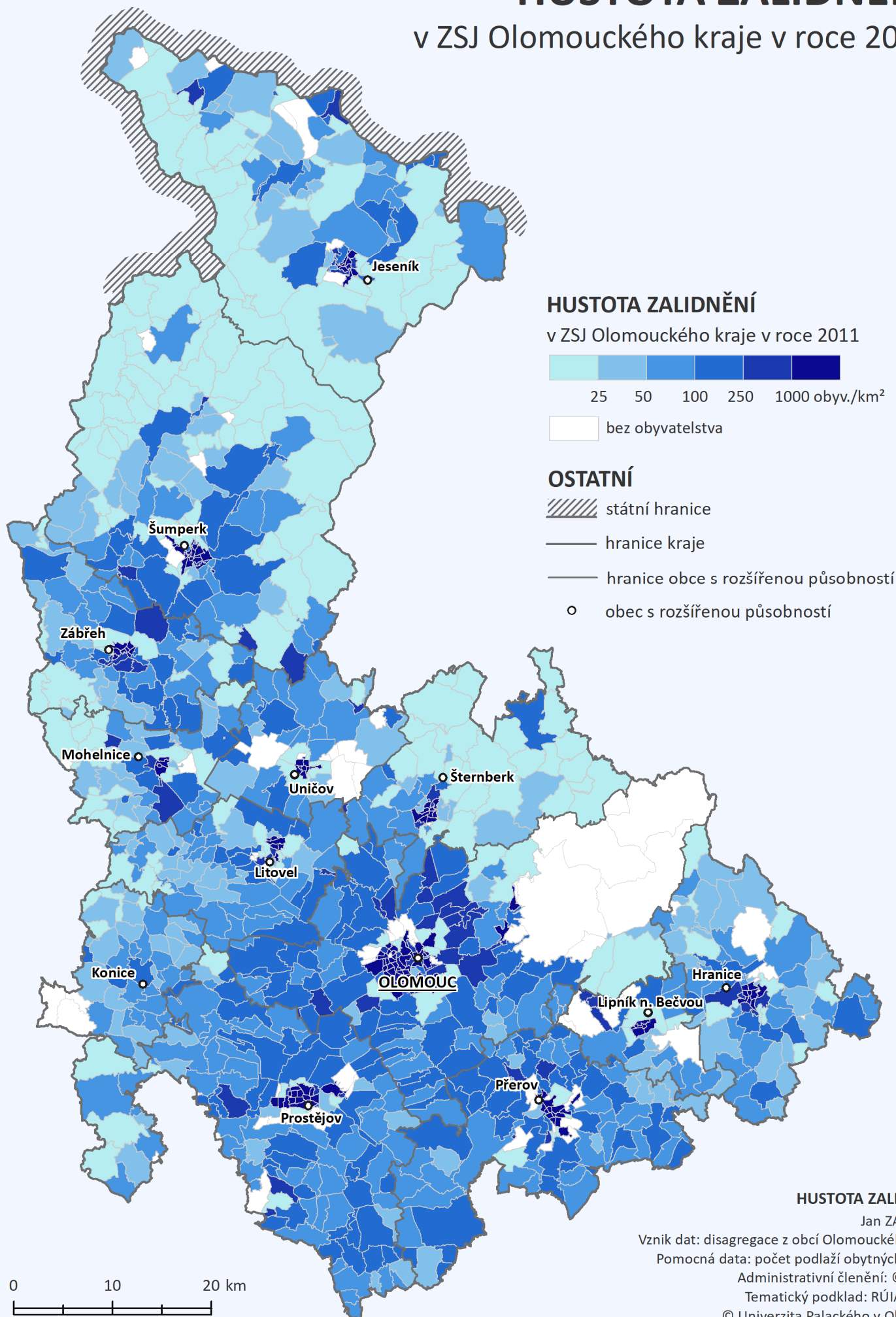
Tematický podklad: RÚIAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

# HUSTOTA ZALIDNĚNÍ

v ZSJ Olomouckého kraje v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace z obcí Olomouckého kraje

Pomocná data: počet podlaží obytných budov

Administrativní členění: © ArcČR

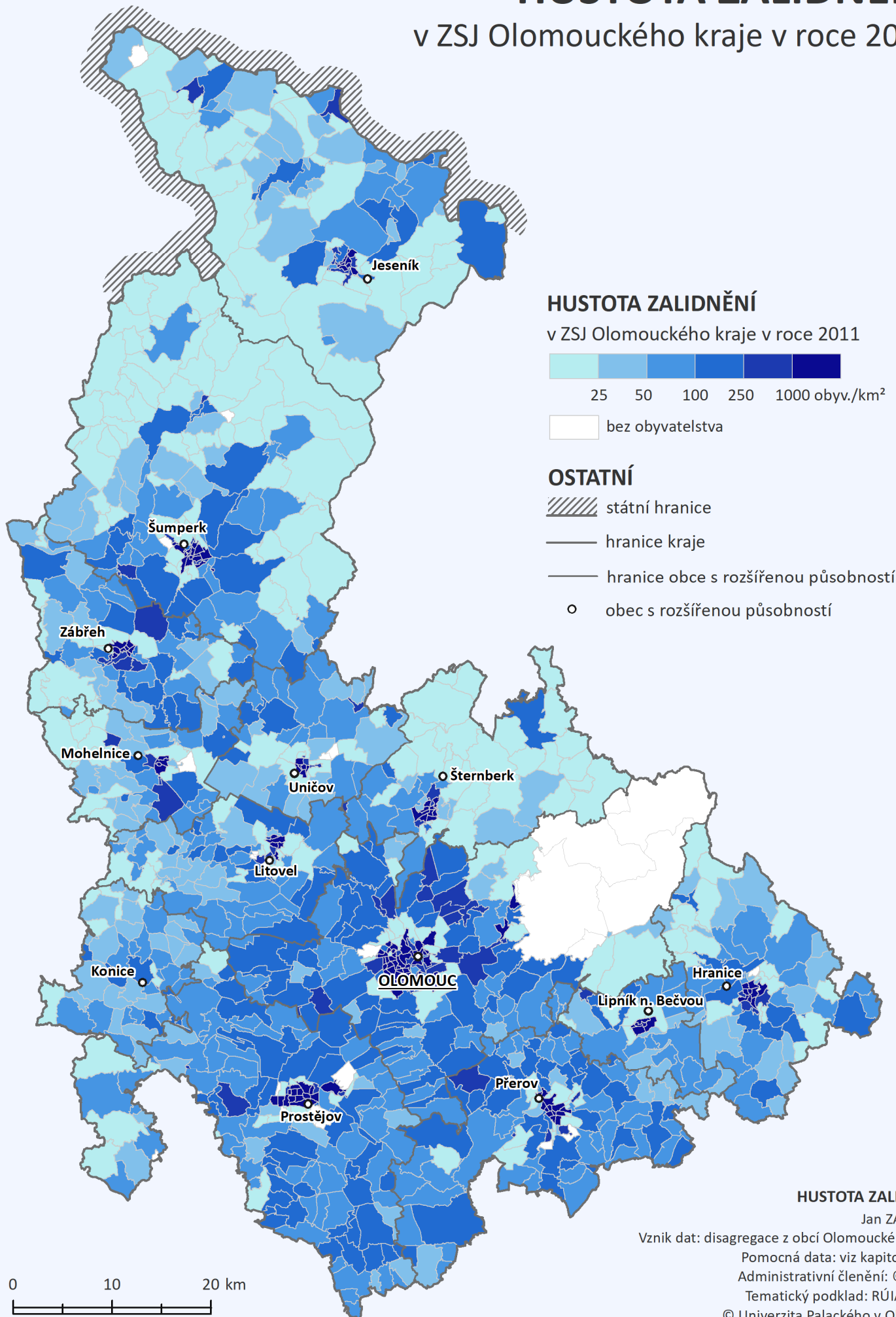
Tematický podklad: RÚIAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

# HUSTOTA ZALIDNĚNÍ

v ZSJ Olomouckého kraje v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace z obcí Olomouckého kraje

Pomocná data: viz kapitola 5.2.2

Administrativní členění: © ArcČR

Tematický podklad: RÚIAN, ČSÚ

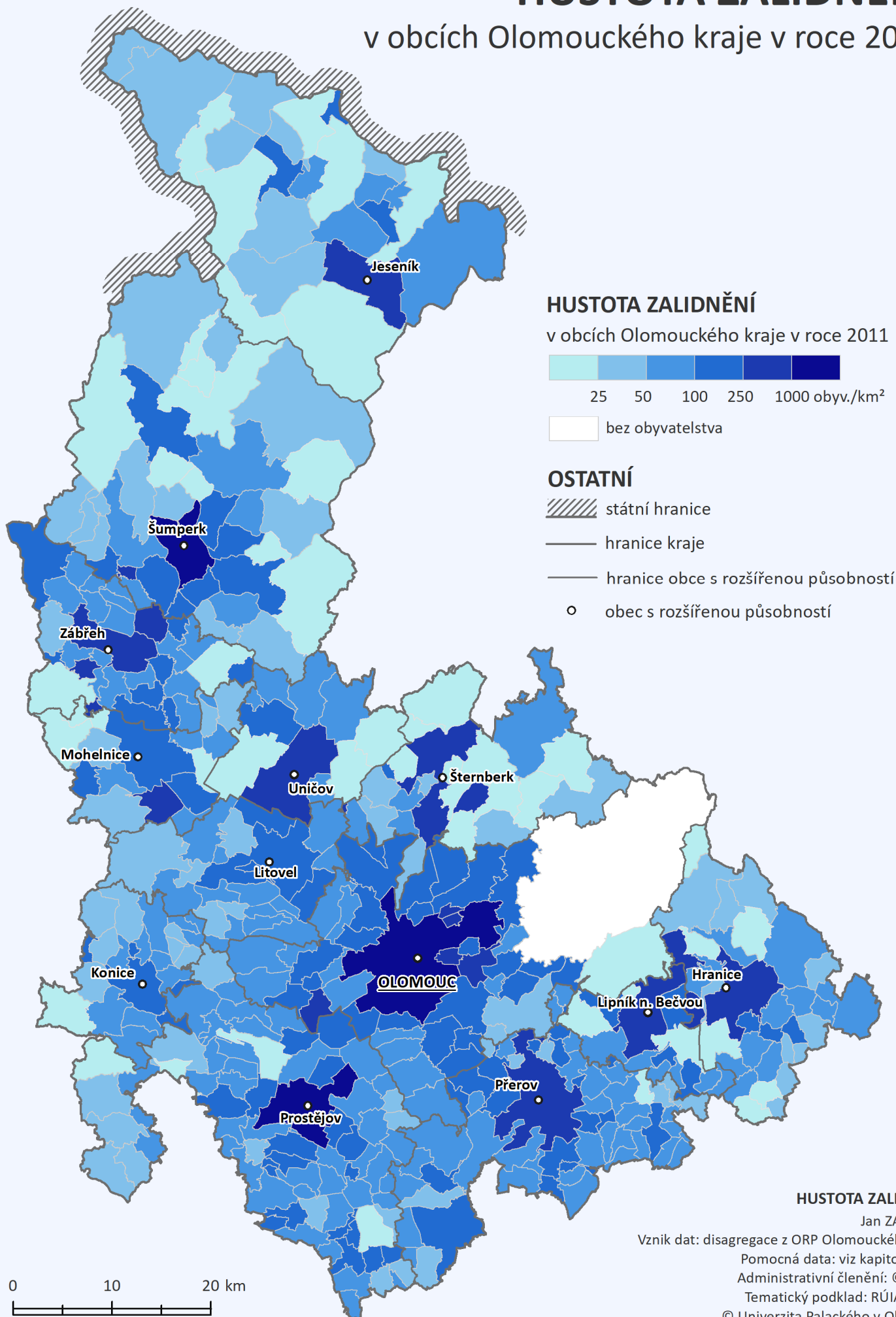
© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

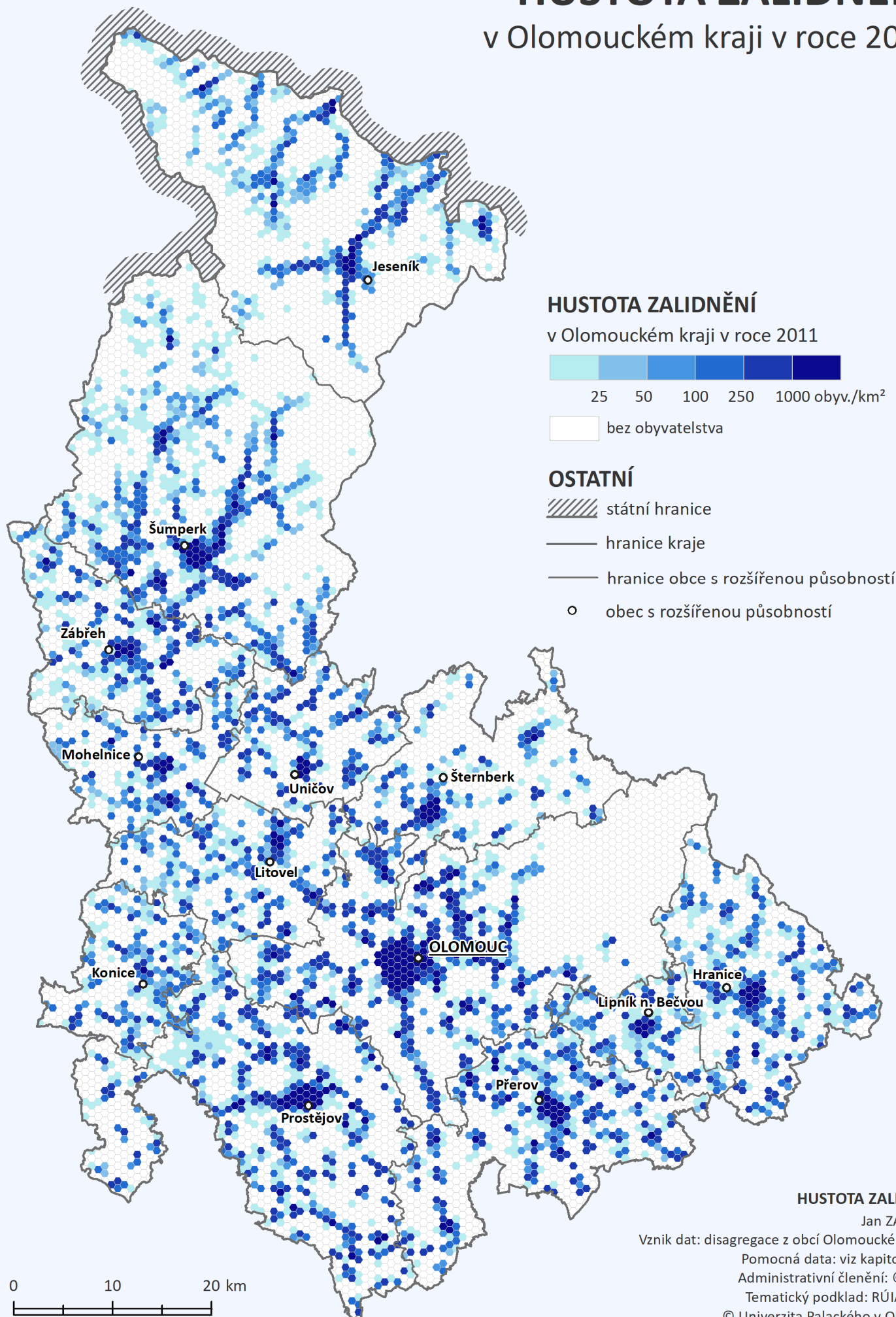


# HUSTOTA ZALIDNĚNÍ

v obcích Olomouckého kraje v roce 2011



# HUSTOTA ZALIDNĚNÍ v Olomouckém kraji v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace z obcí Olomouckého kraje

Pomocná data: viz kapitola 5.2.2

Administrativní členění: © ArcČR

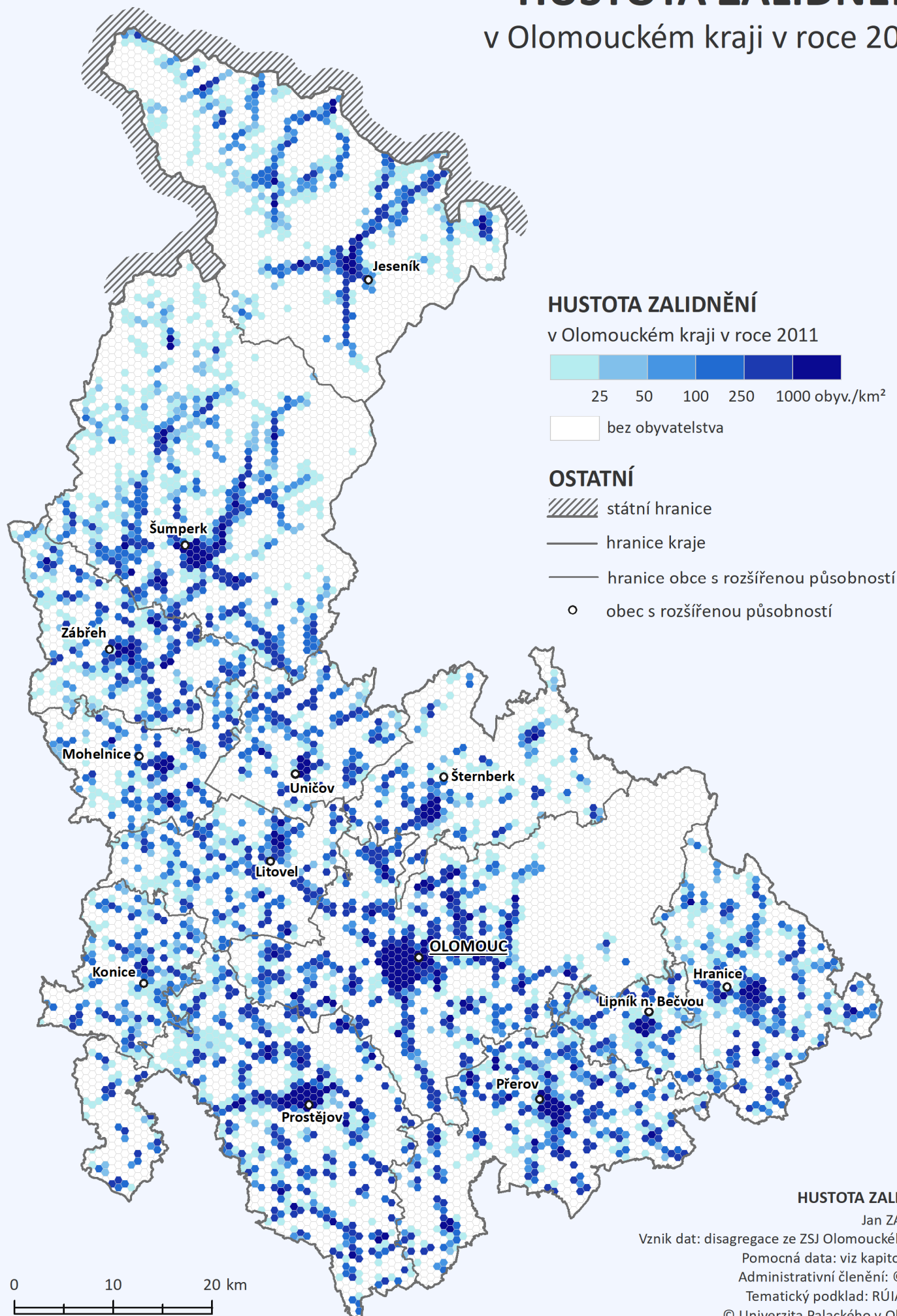
Tematický podklad: RÚIAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N



# HUSTOTA ZALIDNĚNÍ v Olomouckém kraji v roce 2011



## HUSTOTA ZALIDNĚNÍ

Jan ZAPLETAL

Vznik dat: disagregace ze ZSJ Olomouckého kraje

Pomocná data: viz kapitola 5.2.2

Administrativní členění: © ArcČR

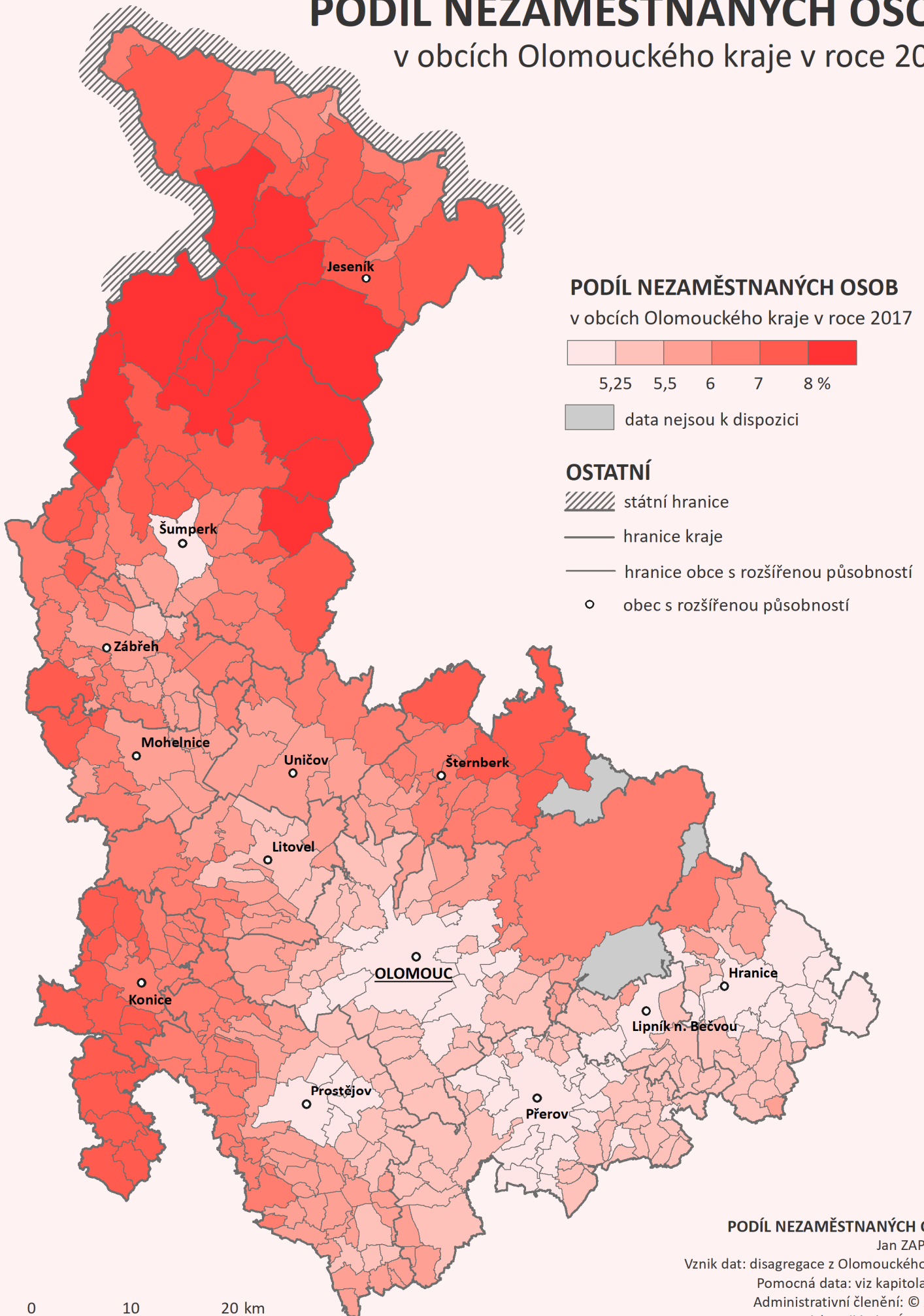
Tematický podklad: RÚIAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

# PODÍL NEZAMĚŠTNANÝCH OSOB

v obcích Olomouckého kraje v roce 2017



## PODÍL NEZAMĚŠTNANÝCH OSOB

Jan ZAPLETAL

Vznik dat: disagregace z Olomouckého kraje

Pomocná data: viz kapitola 5.2.2

Administrativní členění: © ArcČR

Tematický podklad: RÚIAN, ČSÚ

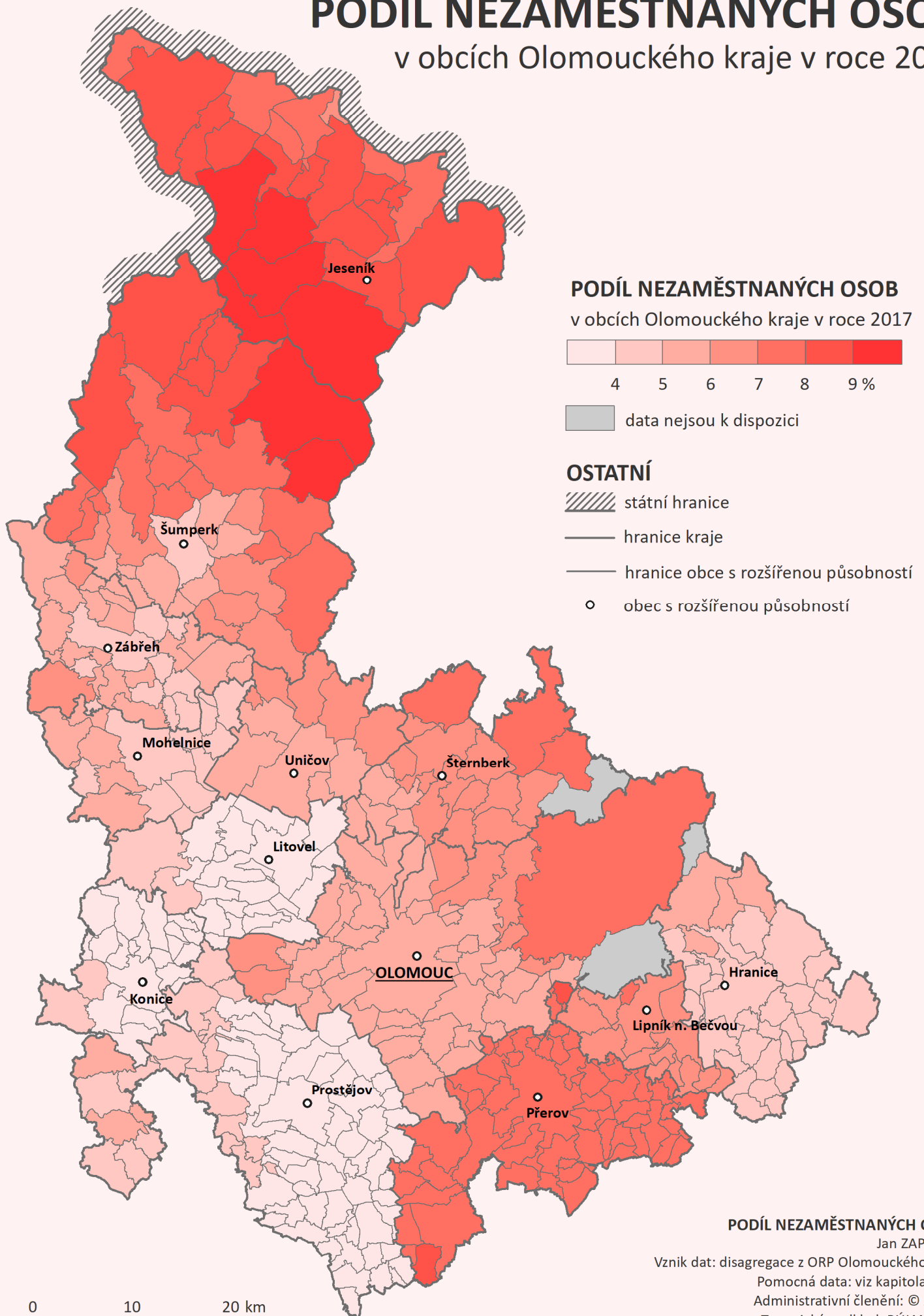
© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

Příloha 11

# PODÍL NEZAMĚŠTNANÝCH OSOB

v obcích Olomouckého kraje v roce 2017



## PODÍL NEZAMĚŠTNANÝCH OSOB

Jan ZAPLETAL

Vznik dat: disagregace z ORP Olomouckého kraje

Pomocná data: viz kapitola 5.2.2

Administrativní členění: © ArcČR

Tematický podklad: RÚIAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

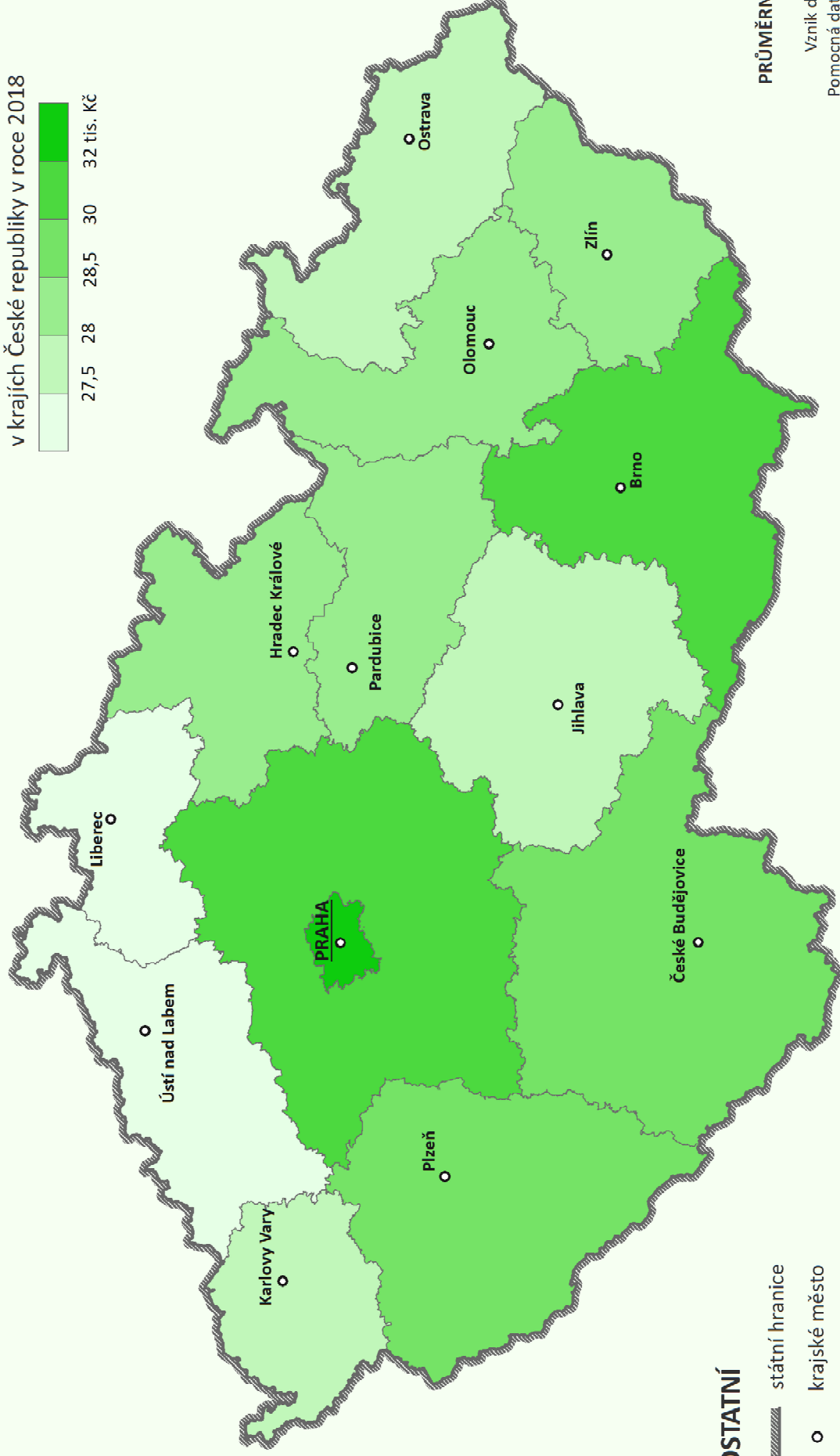
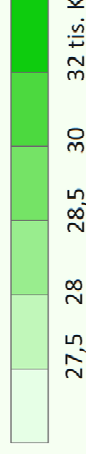
Příloha 12

# PRŮMĚRNÁ HRUBÁ MZDA

v okresech České republiky v roce 2018

## PRŮMĚRNÁ HRUBÁ MZDA

v krajích České republiky v roce 2018



### OSTATNÍ

- státní hranice
- krajské město



### PRŮMĚRNÁ HRUBÁ MZDA

Jan ZAPLETAL

Vznik dat: regresní rovnice

Pomocná data: viz kapitola 5.2.2

Administrativní členění: © ArcČR

Tematický podklad: RÚJAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

Příloha 13

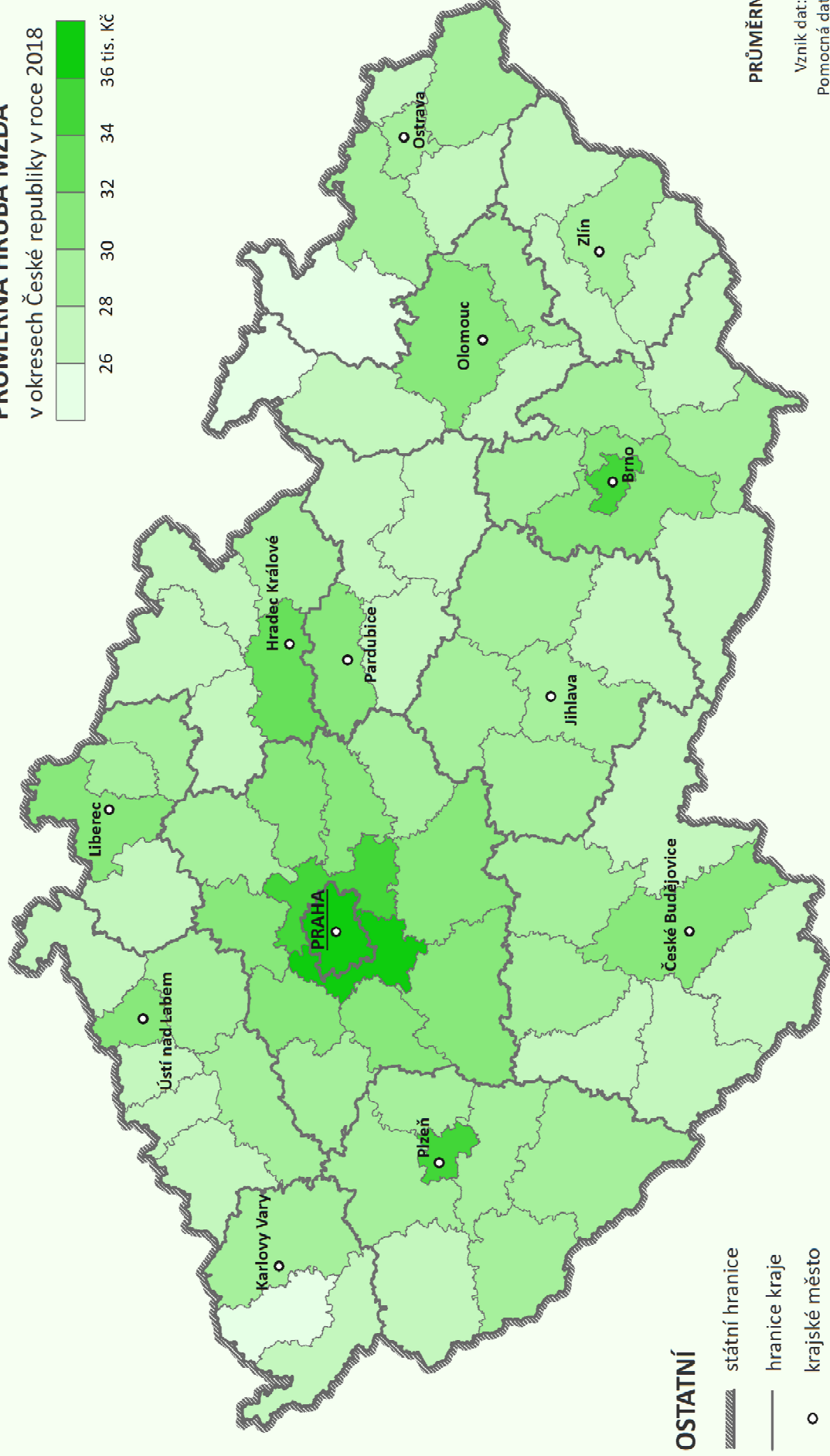


# PRŮMĚRNÁ HRUBÁ MZDA

v okresech České republiky v roce 2018

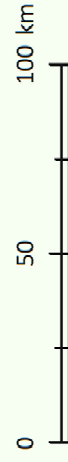
## PRŮMĚRNÁ HRUBÁ MZDA

v okresech České republiky v roce 2018



### OSTATNÍ

- státní hranice
- hranice kraje
- krajské město



1 : 2 000 000

### PRŮMĚRNÁ HRUBÁ MZDA

Jan ZAPLETAL

Vznik dat: disagregace z kraje

Pomocná data: viz kapitola 5.2.2

Administrativní členění: © ArcČR

Tematický podklad: RÚJIAN, ČSÚ

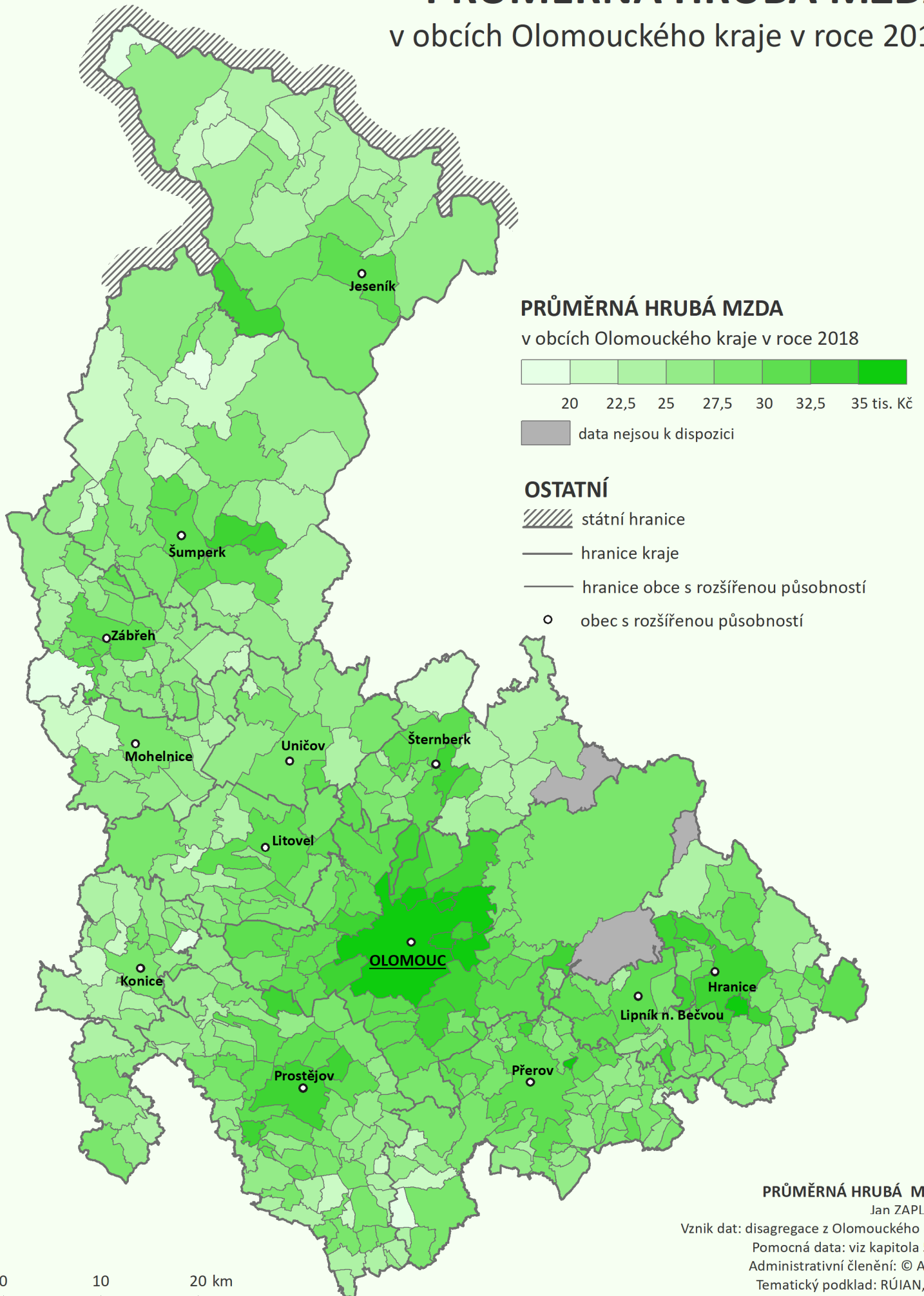
© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

Příloha 14

# PRŮMĚRNÁ HRUBÁ MZDA

v obcích Olomouckého kraje v roce 2018



## PRŮMĚRNÁ HRUBÁ MZDA

Jan ZAPLETAL

Vznik dat: disagregace z Olomouckého kraje

Pomocná data: viz kapitola 5.2.2

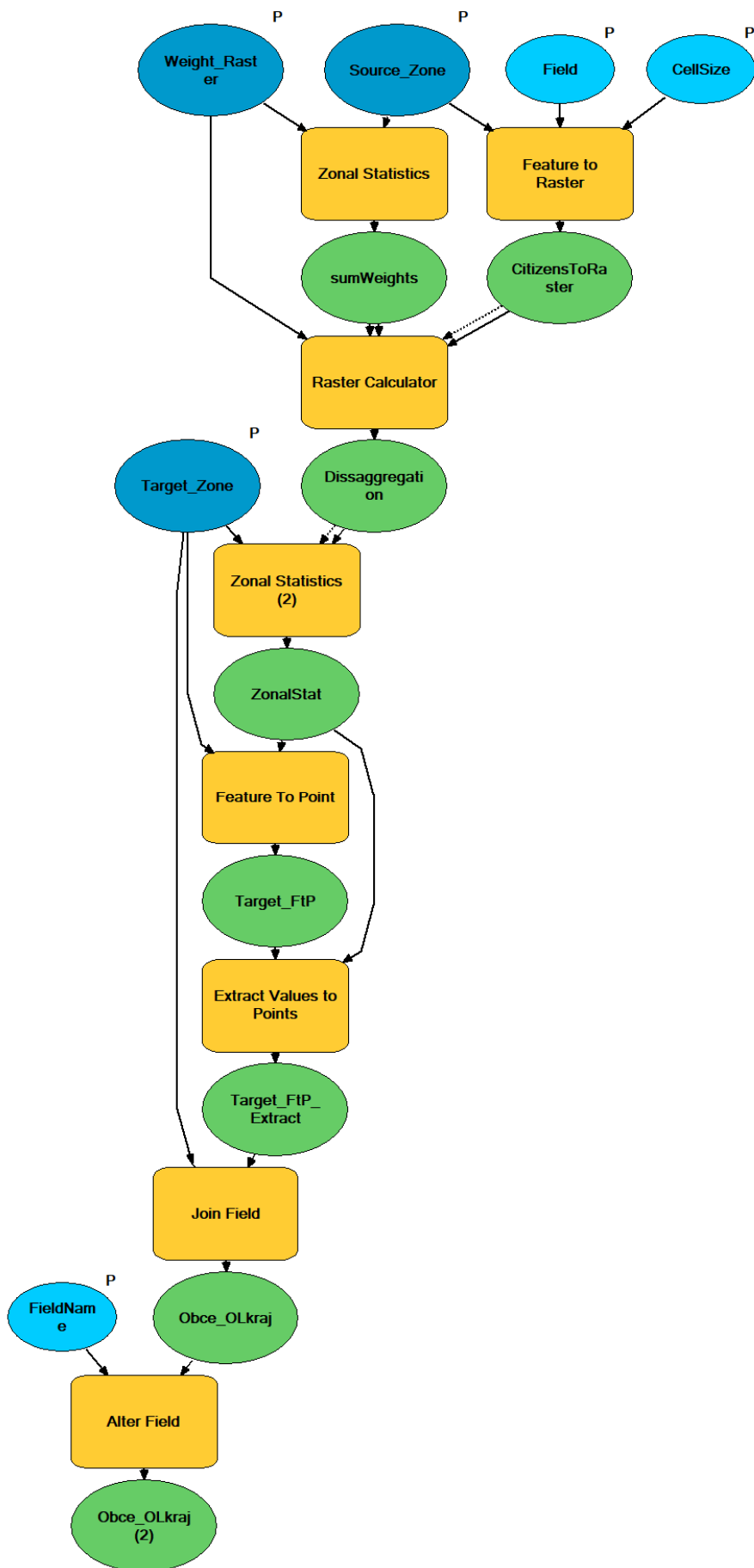
Administrativní členění: © ArcČR

Tematický podklad: RÚIAN, ČSÚ

© Univerzita Palackého v Olomouci

Coordinate system: WGS 1984 UTM Zone 33N

Priloha 16: Schéma nástroje na disagregaci počtu obyvatel



Příloha 17: Schéma nástroje na disagregaci průměrné hrubé mzdy

