

Univerzita Karlova
Přírodovědecká fakulta

Studijní program: Demografie
Studijní obor: Demografie se sociální geografii



Petr Havej

Prostorová statistická analýza bodových dat: možnosti na příkladu výzkumu kriminality
Spatial statistics of point data: possibilities on example of crime research

Bakalářská práce

Vedoucí bakalářské práce: RNDr. Pavlína Netrdová, Ph.D.

Praha, 2019

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze dne 15. 7. 2019

Podpis

Poděkování:

Děkuji především paní RNDr. Pavlíně Netrdové, Ph.D. za probuzení zájmu o problematiku svým nadšeným přístupem, za neutuchající ochotu a cenné rady. Dále můj dík patří rodině a nejbližším za podporu při studiu, což mi dovolilo prožít krásné období.

ABSTRAKT

Bodová lokalizovaná data se ve světě kolem nás nacházejí doslova na každém kroku. Ať už jde o data o umístění různých objektů nebo místa, kde se stal nějaký incident, jako třeba trestný čin či místo kde došlo k výskytu vzácnější choroby. Jedním z důležitých cílů výzkumníka je analyzovat jejich prostorové rozmístění. K takovému cíli existuje množství metod, které se začaly rozvíjet zejména s rozvojem na poli výpočetní techniky a informačních technologií. Cílem této práce je zjistit, jaké jsou současné metody přístupu k prostorové analýze bodových dat. Na skutečných datech za kriminalitu otestovat a aplikovat zkoumané metody prostorové analýzy a zhodnotit jejich přínosy a výsledky. Diskutovány budou problémy, které mohou vyvstat při aplikování metod a také skutečné výzkumy využívající takové metody.

Klíčová slova: prostorová analýza, bodová data, výzkum kriminality

ABSTRACT

Localized point data are literally everywhere in the world around us. Whether it is about the location of different objects or a place where an incident happened, such as crime, or a place where a more rare disease occurred. One of the researcher's important goals is to analyze their spatial distribution. There is a number of methods to achieve such a goal, which have begun to develop in particular along the development of the field of computing and information technology. The aim of this work is to find out what are the current methods of approach to spatial analysis of point data. To test and apply the spatial analysis methods studied on actual data for crime and evaluate their benefits and results. Problems that may arise in applying methods and actual research using such methods will be discussed.

Keywords: spatial analysis, point data, crime research

Počet znaků: 70 923

OBSAH

PŘEHLED POUŽITÝCH ZKRATEK.....	2
SEZNAM OBRÁZKŮ.....	3
1 ÚVOD.....	4
2 PROSTOROVÁ STATISTICKÁ ANALÝZA BODOVÝCH DAT.....	6
3 VÝZKUM PROSTOROVÝCH ASPEKTŮ KRIMINALITY.....	10
3.1. VÝZKUMY ZAMĚŘENÉ NA ČESKO.....	13
4 DATA O KRIMINALITĚ.....	15
5 METODY PROSTOROVÉ STATISTICKÉ ANALÝZY BODOVÝCH DAT.....	17
5.1. ZÁKLADNÍ POPISNÉ STATISTIKY.....	18
5.2. ANALÝZY KONCENTRACÍ.....	21
5.3. PROSTOROVÉ MODELOVÁNÍ, INTERPOLACE.....	29
6 ZÁVĚR.....	33
SEZNAM LITERATURY A POUŽITÝCH ZDROJŮ.....	35

PŘEHLED POUŽITÝCH ZKRATEK

ESDA	Explorační prostorová analýza dat (<i>Exploratory Spatial Data Analysis</i>)
GIS	Geografický informační systém
GPS	<i>Global Positioning System</i>
IKSP	Institut pro kriminologii a sociální prevenci
LISA	<i>Local Indicators of Spatial Association</i>
NNH	Metoda hierarchického prostorového shlukování nejbližších sousedů (<i>Nearest neighbor hierarchical spatial clustering</i>)
NNI	Index nejbližšího souseda (<i>Nearest neighbour index</i>)
RNNH	Metoda hierarchického prostorového shlukování nejbližších sousedů přizpůsobena riziku (<i>Risk-adjusted nearest neighbor hierarchical clustering</i>)
STAC	Prostorová a časová analýza zločinu (<i>Spatial and Temporal Analysis of Crime</i>)

SEZNAM OBRÁZKŮ

Obr. 1: Ukázka centrografických statistik rozmístění vloupání v Norwood Parku v Chicagu v roce 2018	19
Obr. 2: Srovnání tří typů trestných činů (vloupání, krádeže aut, ublížení na zdraví) v Norwood Parku v Chicagu z roku 2018 a jejich prostorové distribuce pomocí geografického středu a směrodatné elipsy odchylek	20
Obr. 3: Modus krádeží vozidel a výsledek metody „fuzzy mode“ s radiusem 500 čtverečních stop pro krádeže vozidel v Chicagu v roce 2018.....	23
Obr. 4: Shluky krádeží motorových vozidel v Chicagu v roce 2018 utvořené metodou hierarchického shlukování nejbližších sousedů (NNH)	25
Obr. 5: Detail jednoho ze shluků druhého řádu vytvořeného pomocí metody hierarchického shlukování nejbližších sousedů.....	26
Obr. 6: Shluky případů krádeží motorových vozidel v Chicagu v roce 2018 identifikované pomocí metody STAC	28
Obr. 7: Výsledek metody jádrových odhadů na krádeže motorových vozidel v Chicagu v roce 2018 v porovnání s ohnisky identifikovanými pomocí metody STAC	31

Kapitola 1

Úvod

Bodová lokalizovaná data se ve světě kolem nás nacházejí doslova na každém kroku. Ať už jde o data o umístění různých objektů (školy, obchody apod.) nebo místa, kde se stala nějaká událost (trestný čin, výskyt vzácnější choroby apod.). Jedny z důležitých otázek, které si kladou nejenom geografové, jsou spojeny s výskytem různých jevů a procesů. Kde přesně a proč zrovna tam se jevy vyskytují? Dochází k jejich koncentraci a mění se tato koncentrace v průběhu času? Lze vypořádat prostorový vzorec v jejich výskytu, který by umožnil predikci? K zodpovězení uvedených otázek je nutné mít k dispozici data o výskytu sledovaných jevů a správně je analyzovat s důrazem na jejich prostorové rozmístění. K takovému účelu lze využít množství metod, které se začaly rozvíjet zejména s rozvojem na poli výpočetní techniky a informačních technologií.

Bakalářská práce je metodologicky zaměřená, když se snaží představit současné metody prostorové statistické analýzy bodových dat a identifikovat aktuální trendy na tomto poli. Vývoj v této oblasti, zejména vývoj softwarů a samotných metod je velmi rychlý. Můžeme za tím hledat především rozvoj informačních technologií v posledních letech a také stále větší dostupnost dat, i těch lokalizovaných. Je tak potřeba sledovat tento vývoj a snažit se nové metody a postupy aplikovat také v Česku. Práce se přitom zaměřuje na využití metod prostorové statistické analýzy bodových dat v sociálně-geografickém výzkumu.

Cílem bakalářské práce je představit současné přístupy a trendy týkající se prostorové statistické analýzy bodových dat a otestovat a zhodnotit nejčastěji používané metody. Provedena bude rešerše současné metodologické literatury i sociálně-geografických výzkumů analyzujících bodová data. Konkrétní vybrané metody analýzy budou aplikovány a testovány na reálných lokalizovaných datech o kriminalitě, což umožní zhodnocení jejich možností a omezení. Diskutovány budou problémy, které mohou vyvstat při aplikaci metod a interpretaci výsledků. K analýze je využíván software ArcGIS a software CrimeStat, který se na tento druh analýz specializuje a který není v českém prostředí běžně používán.

Jedna z důležitých oblastí, kde se využívá prostorová statistická analýza bodových dat je oblast výzkumu kriminality. Výzkum v této oblasti se časem vyvíjel od mapování trestných činů ještě před dobami využívání geografických informačních systémů až po sofistikované využívání nejrůznějších metod k odhalení prostorových vzorců rozmístění trestných činů. Právě na výzkumu kriminality budou testovány vybrané metody prostorové analýzy bodových dat, k čemuž poslouží skutečná, přesně lokalizovaná data o trestných činech.

Mezi výzkumné otázky, na které je v práci hledána odpověď, patří: Jaké jsou možnosti a přínosy prostorové statistické analýzy bodových dat? V jakých oblastech (nejen) sociálně-geografického výzkumu se využívají metody prostorové statistické analýzy bodových dat? Jaké jsou možnosti

a omezení jednotlivých metod a jejich aplikace ve vybraném softwaru? Je software CrimeStat užitečný nástroj k analýze bodových dat (oproti softwaru ArcGIS) a jaké jsou jeho výhody a nevýhody? Zodpovězením těchto otázek by práce měla přispět k většímu povědomí české odborné veřejnosti o aktuálních možnostech a trendech v prostorové statistické analýze bodových dat. Konkrétní aplikace by se také mohly stát jakýmsi návodem pro další odborné práce zabývající se analýzou bodových dat.

Kapitola 2

Prostorová statistická analýza bodových dat

Prostorová statistická analýza se řadí mezi kvantitativní metody výzkumu. Za prostorové analýzy je možné obecně považovat kvantitativní studium jevů, které se projevují v prostoru. Prostorové analýzy je tedy možno chápat jako specifickou podskupinu kvantitativních analýz, ke které jsou potřeba prostorová data - data která se vztahují k určitému místu (Netrdová, 2010). Využití takových metod v geografii jde ruku v ruce s vývojem myšlení a nastupujícími paradigmaty v průběhu vývoje geografie jako vědy. Nastupující paradigmatické směry se lišily charakteristickými filozofickými a metodologickými přístupy, jež se týkaly také názorů na možnosti využití kvantitativních metod. S využitím kvantitativních metod v geografii souvisí i diskuse o monistickém a dualistickém pojetí geografie. Při dualistickém chápání je geografie pojímána rozděleně jako fyzická geografie, patřící mezi vědy přírodní, a sociální geografie, patřící mezi vědy společenské. Jak uvádí Netrdová (2010), rozdílný charakter těchto dvou disciplín se projevuje i v množství odlišných problémů spojených s kvantitativní analýzou. Ve fyzické geografii se geografové často snaží o predikci přírodních jevů a procesů v prostoru a tíhnou k pozitivismu mnohem více než sociální geografové, kde je fenoménem zklamání z pozitivistické filosofie, spojené právě i s odmítavým postojem k využití kvantitativních metod (Netrdová, 2010). To je způsobeno třeba i často obtížnou kvantifikací sociálních jevů spojeného se stochastickým charakterem lidského chování, což dále rozvádí například Pavlínek (1993): „Společnost a jednotlivci nemohli poskytnout v období kvantitativní revoluce ideální terén pro geografický výzkum v důsledku nepředvídatelnosti jejich chování.“ Cíl většiny studií v kvantitativní geografii tak není produkovat bezchybné výzkumy (což je v mnoha případech, zejména právě pokud se pracuje se společensko-vědními daty, nemožné), ale spíše maximalizovat pochopení prostorových procesů s minimální chybou. (Fotheringham a kol, 2000)

Prostorová analýza má své kořeny už v dobách dávných, před geografickými informačními systémy, dokonce i před samotnou výpočetní technikou. K prostorové analýze jsou zapotřebí prostorová data, tedy data, jež v sobě zahrnují pozorování nějakého jevu a zahrnují i prostorovou informaci. Ta může být vyjádřena explicitně ve formě adresy nebo souřadnice, nebo implicitně jako pixel uprostřed satelitního snímku (Fotheringham a kol., 2000). Za ranou formu práce s prostorovými daty by se daly považovat například ty, které uvádí Fotheringham a kol. (2000): raní astronomové, kteří se snažili zmapovat vesmír nebo Ptolemaios a jeho experimenty zmapovat svět. Prostorovou analýzu bychom mohli popsat jako proces využívání technik k zjišťování prostorových vzorců, rozmístění, lepšímu pochopení procesů tvořící hodnoty nějakých atributů. (Fotheringham, Rogerson, 2009). Není správné označovat prostorovou analýzu a analýzu prostorových dat za synonyma, protože za prostorovou analýzu se dá považovat pouze ta analýza, jež využívá prostorový aspekt dat, což ne každá analýza prostorových dat splňuje. (Netrdová, 2010).

Prostorová data obsahují jak informaci o poloze tak informaci o nějakém atributu. Tudíž jsou to data, pro která je nějaký atribut zaznamenán na různých polohách a tyto polohy jsou zakódovány jako

součástí dat. Stále více si v dnešní době uvědomujeme, že většina dat je prostorová. Jako běžné příklady prostorových dat uvádí Fotheringham s Rogersonem (2009) data z cenzu, incidence chorob nebo třeba distribuce rostlinných druhů. Prostorová data se liší od neprostorových dat v několika vlastnostech. Předně je to v tom, že data na sobě typicky nejsou nezávislá, hodnoty atributů tíhnou k tomu být podobnější v k sobě blízkých místech než v místech od sebe více vzdálených. (Fotheringham, Rogerson, 2009)

Prostorová analýza se často spoléhá na data agregovaná do různých zón (polygonů), sledované vzorce se však mohou radikálně měnit v závislosti na změně volby hranic těchto zón. Agregovaná data ve většině případů pocházejí ze souboru individuálních událostí. Agregace však málokdy slouží jako pomoc k prostorové analýze, spíše se jedná o administrativní důvody. Ve skutečnosti je však docela možné analyzovat prostorové vzorce desagregovaných individuálních dat většinou tak, že jsou každé jednotlivé události přiřazeny k jednotlivým bodům v prostoru a poté se posuzuje prostorová distribuce těchto bodů. (Fotheringham a kol., 2000)

Tato práce se tedy zaměřuje právě na možnosti statistické analýzy prostorových bodových dat. To jsou data, jež nám dávají informaci o přesném umístění nějakých jevů, věcí nebo událostí. Jako příklad mohou sloužit stromy v lese, epicentra zemětřesení, hovory z mobilních telefonů, případy vzácné choroby, naleziště zlata nebo třeba nehody na silnici. Hlavním a nejdůležitějším bodem zájmu zkoumání a analýz takovýchto dat, je právě jejich prostorové rozmístění, hledání vzorců a odchylek od náhodného rozmístění. V této práci se z velkého množství oblastí a zdrojů dat zaměřím primárně na příklady výzkumu kriminality a představené metody analýzy budu testovat právě na datech za trestné činy. Prostorová analýza bodových dat se samozřejmě neomezuje pouze na jevy spojené s kriminální činností, ba ani pouze na jevy společenské. Jejich využití lze nalézt napříč vědeckým spektrem od věd jako ekologie, epidemiologie, astronomie, ekonometrie nebo geologie (Baddeley a kol., 2017).

To, co jsou bodová data, nebo co můžeme považovat za bodová data je do velké míry dáno měřítkem, ve kterém jevy sledujeme. Rogerson a kol. (2000) uvádí příklad – zaměstnanec policie zkoumá vloupání do domácností v určitém městě a každé je reprezentováno bodem, jednotkou analýzy jsou tedy domácnosti. Na více lokálním měřítku, pokud budou dvě sousední domácnosti v konfliktu ohledně pozemku, nebudou se považovat za body, ale za zony. Podobně města mohou být v menším měřítku brána jako zony, v případě třeba národního měřítka je vhodnější reprezentovat všechna města jako body. (Rogerson a kol., 2000)

Statistická analýza prostorového rozmístění bodů může výzkumníkovi dopomoci k zjištění důležitých rysů a vlastností zkoumaného souboru dat, jako například těch, které uvádí Baddeley a kol. (2017), tedy třeba tendenci nalezišť zlata se nacházet blízko závažnému geologickému zlomu nebo zjištění, že případy výskytu určité choroby se více nacházejí blízko zdroji nějakého druhu znečištění nebo nalezení faktu, že si hnízda ptáků udržují jistou minimální separaci mezi sebou. Analýzy bodových dat tak poskytují klíčové důkazy u důležitých výzkumů různé škály, od výzkumu přenosu cholery, přes výzkum chování sériových vrahů až třeba k výzkumům makrostruktur vesmíru (Baddeley a kol., 2017).

Pokud analyzujeme bodová data, hledáme v nich často nějaké vzorce. Proto se něčemu takovému často říká analýza bodových vzorců (point pattern analysis). A jako ve většině prostorových analýz, i při hledání takových vzorců v bodech se snažíme o to, aby nám nalezené vzorce rozmístění řekly něco o procesech, které stojí za takovým rozmístěním bodů. Například jestli se dají nalézt místa s vyšší koncentrací vloupání do domácností, jak taková místa vypadají, co jsou jejich charakteristiky a jestli místa s podobnými charakteristikami mají také vyšší koncentrace vloupání. Nebo, jak zmiňuje Rogerson a kol. (2000), nás právě může zajímat jestli se hnízda některých druhů teritoriálních ptáků nacházejí dále od ostatních hnízd. V takovém případě by výzkumník testoval rozptýlenost souboru bodů, a pokud by byla nalezena, bylo by možné zjistit minimální vzdálenost, která se vyskytuje mezi dvěma hnízdy. Tyto dva příklady ukazují, že při prostorové analýze bodových dat se snažíme obecně nalézt 2 hlavní fenomény, a sice shlukování, kde se body nacházejí u sebe blíže než kdyby byly rozdělené

náhodně, nebo rozptýlení, kde se body nacházejí od sebe dále než by se dalo očekávat při náhodném rozdělení. (Rogerson a kol., 2010)

Ke statistické analýze bodových dat ovšem neexistuje žádné jednoduché univerzální řešení, je důležité mít na paměti, že korektní způsob analyzování dat nezávisí pouze na formátu dat, ale i na tom jak byla data získána a samozřejmě na cílech analýzy (Baddeley a kol., 2017).

Další využití analýz můžeme spatřit ve srovnávání a to jak ve srovnání souborů z jiných oblastí, kde hledáme jestli jsou například vzorce rozmístění podobné na dvou různých místech, nebo dvou a více typů souborů bodů ve stejné oblasti. Při takovém srovnání se můžeme ptát na otázky typu: Mají krádeže aut tendenci nacházet se na podobných místech jako vloupání do domácnosti? Avšak i pokud zkoumáme pouze jeden soubor bodů, často tak činíme v komparativním rámci, typicky porovnáváme distribuci souboru bodů s nějakou referenční distribucí. Pro příklad uvádí Rogerson a kol. (2000) zkoumání incidence choroby, kdy je důležité zjistit, zda jsou případy distribuovány v souladu s rozložením populace. Pokud jsou případy výskytu choroby více prostorově seskupené než rozložení populace, může to být důkaz shlukování kolem nějakých míst. V tomto případě není na místě porovnávat výskyt chorob s náhodným rozmístěním ve studované oblasti, protože se dají očekávat shluky výskytu právě ve vztahu k prostorovému rozložení populace. To, že náhodné rozmístění je nevhodné k využití jako referenční distribuce, je v sociální geografii častý jev. (Rogerson a kol., 2000)

Prostorové analýzy v geografii začaly opět nabírat na důležitosti na přelomu tisíciletí, po období, kdy byly v útlumu v souvislosti s obecnou kritikou konceptu geografie jako prostorové vědy, se jejich narůstající popularita a důležitost spojuje především s rozvojem na poli informačních a výpočetních technologií, zejména s rozvojem geografických informačních systémů (GIS), který sehrál nejdůležitější roli právě ve vývoji a aplikaci prostorových analýz. Díky těmto GIS softwarům byl umožněn vývoj sofistikovanějších nástrojů a metod pro analyzování prostorových dat a jejich prostorového aspektu. To vedlo k usnadnění provádění územně podrobných analýz, náročných zejména početně. GIS je prostředí pro manipulaci, analýzu, modelování a zobrazování prostorových dat které umožňuje použití množství nástrojů a metod prostorové analýzy. (Netrdová, 2010)

Metody prostorové analýzy jsou rozšířené v řadě oblastí, data v desagregované formě jako body jsou dostupná v široké škále oborů. Analýza prostorových vzorců bodů je populární zejména v biologii, epidemiologii a kriminologii (Rogerson a kol., 2000). Z oboru biologie stojí za zmínku například studie Buluška a kol. (2016) o prostorových vzorcích lesů reliktního buku v Česku a Polsku nebo studie Zemana (2011) o rozmístění mravenišť mravenců *Formica polyctena* a faktorech ovlivňující toto rozmístění. V epidemiologii je prostorová analýza považovaná za klíčový nástroj pro vytváření a testování modelů předpovídajících efekt změny klimatu na prostorové rozmístění chorob (Graham a kol., 2004). Studiu věnovanému prostorovým aspektům kriminality je věnována pozornost v kapitole 3, kde je dále rozvedeno.

Pro analýzu prostorových dat existuje v dnešní době již množství softwarů, kde jsou pro takovou analýzu dostupné nástroje. Mezi takové patří i ArcMap vyvinutý společností ESRI a CrimeStat, který vyvinul Ned Levine. V této práci bude užito právě těchto dvou softwarů. Druhý jmenovaný, CrimeStat ve verzi 4.02, bude podroben větší examaci, jelikož se specializuje právě na prostorovou analýzu bodových dat. Bude zhodnocen přínos tohoto softwaru a jeho potenciální využití v tuzemském výzkumu.

CrimeStat je prostorově statistický program využívaný původně hlavně k mapování kriminality, nalezl ovšem využití i v jiných oborech využívající zejména bodová data. Tento program je vyvíjen od 90. let díky grantu od National Institute of Justice. Je to následovník programu Hawaii Pointstat. CrimeStat je program pro Windows, napsaný v C++ a jeho první verze byla vydána v roce 1999. Následovala verze druhá z roku 2002, verze 3 z roku 2004 a poslední verze 4, jejíž nejnovější aktualizace, verze 4.02 vyšla

v roce 2015. (Levine, 2008). CrimeStat není GIS, jeho součástí není možnost vytváření map, vizualizace dat a výsledků výpočtů, je tak vhodné s ním pracovat v kombinaci s GIS softwarem. Vstupními daty jsou přesně lokalizovaná (X a Y souřadnice) místa činu, incidenty (body) a výstupem jsou statistiky, které mohou být uloženy ve formátu shapefile, které jsou kompatibilní s většinou GIS softwarů, včetně ArcMapu, a mohou tak být dále graficky zobrazeny v jiných programech geografických informačních systémů. V rozsáhlé uživatelské příručce je software CrimeStat popsán jako: „balík pro prostorovou statistiku, který může analyzovat rozmístění trestných činů“ (Levine, 2013). Mezi hlavní funkce softwaru CrimeStat 4.0 patří (1) prostorová deskripce (spatial description), která umožňuje uživateli provést analýzy, týkající se prostorové distribuce dat, (2) analýzy koncentrací (hot spot analysis), pomocí kterých uživatel identifikuje skupiny případů spolu se shlukujícími, (3) prostorové modelování (spatial modeling), interpolace, umožňující uživateli analyzovat prostorové chování dat, (4) Crime Travel Demand Modeling, umožňující analyzovat chování potenciálních sériových zločinců. (Levine, 2006; Levine, 2013, Heraux, 2007)

Kapitola 3

Výzkum prostorových aspektů kriminality

Kriminální činnost je něco, co se v nejrůznějších podobách vyskytuje po celém světě a po celou dobu existence lidstva. Kriminalita je komplexní fenomén, zločin nastává konvergencí trestněprávní normy, která označí posuzované jednání jako trestné, pachatele, oběti, proti které jednání směřuje a místa trestného činu, které je definováno jak místem tak časem, ve kterém se předchozí tři složky setkají a vytvoří tzv. kriminální událost (crime event) (Barilik, 2014). Ta nabývá většinou formu krádeží, přepadení, vražd atd. a přináší s sebou újmu peněžní (ztráta majetku, pojištění, vymáhání práva, soudnictví apod.) i psychologickou (bezpečí, viktimizace). Různé typy trestných činů mají odlišný dopad na peněžní nebo psychologickou újmu společnosti, je však možné konstatovat, že redukce zločinu má obecně pozitivní vliv na společnost (Andresen, 2006). Aby byla společnost schopna redukovat zločin, musí pochopit aspekty a podmíněnosti jeho výskytu. Kriminalita patří mezi interdisciplinární témata, jíž se věnuje řada oborů zejména z oblasti sociálních věd, např. kriminologie, sociologie, antropologie, psychologie, dotýká se ale rovněž pedagogiky, sociální politiky či genderových studií. Neméně důležitou roli v pochopení kriminality hraje geografie, která přispívá k jejímu hlubšímu poznání zejména tím, že klade důraz na její prostorové aspekty, čímž tak přidává důležitou dimenzi, která je jinými obory opomíjena. (Jíchová, 2014) Geografie se mimo jiné zaměřuje i na mapování kriminality. To lze spolu s analýzou jejich ohnisek považovat za základní metodu a postup při analýze kriminality. Metod a různých postupů, jak odhalit ohniska zločinnosti a objasnit prostorové vzorce trestných činů, existuje velké množství, které se liší v míře generalizace a mají různý aplikační potenciál. Avšak smyslem všech těchto metod je identifikovat problémová místa a tím přispět k porozumění výskytu trestných činů. V ideálním případě by na základě těchto zjištění měly být zavedeny aktivní i preventivní opatření vedoucí ke snižování míry kriminality. (Jíchová, Netrdová, 2017)

Ve vývoji geografického (prostorového) přístupu ke studiu kriminality lze identifikovat tři až čtyři etapy. Za počátek tohoto přístupu se dá považovat první polovina 19. století, například práce evropských (francouzských) statistiků Adolpha Queteleta (1833, 1842) a André-Michela Guerryho (1833), kteří položili základ zohledňování prostorového aspektu při studiu kriminality. Jíchová (2014) uvádí, že v 19. století se výzkum kriminality zaměřuje jak na diferenciaci územního rozložení kriminality, tak rovněž na hledání souvislostí s dalšími proměnnými. V jejich pracích zkoumali vliv sociálních ukazatelů (věk, pohlaví, vzdělání, profese) na míru násilné a majetkové kriminality na úrovni francouzských departementů. Poukázali tak na fakt, že kriminalita není v prostředí distribuována náhodně, nýbrž různá místa vykazují různou míru kriminality (Jíchová, Netrdová, 2017; Barilik, 2014). Přestože mnoho analýz bylo zaměřeno na velké územní celky, kde bylo upozorňováno na problematiku vysoké vnitřní variability takových území, již ke konci 30. let se začínají objevovat studie zkoumající kriminalitu v detailním měřítku (typu náměstí, ulice) (Jíchová, 2014). Ze studií 19. století mimo jiné plyne, že na mikro úrovni se kriminalita koncentruje do míst, kde se vhodné příležitosti překrývají se známým prostorem největších skupin pachatelů, tedy například do pomezí chudinských čtvrtí a bohatších míst, do míst komerce a zábavy a podobně (Barilik, 2014). Vliv pozitivismu na studium kriminality v tomto

období je zřetelný ve volbě metod a v přístupu k tématu. Dominovaly totiž studie založené na analýze dat sbíraných ve vznikajících specializovaných statistických institucích. Vedle kvantitativních statistických analýz dat byly využívány i vlastní terénní šetření či analýzy dopisů, graffiti nebo tetování. Protože byly výsledky analýz autorů 19. století především mapy, později se začali označovat jako tzv. kartografičtí kriminalisté (Jíchová, 2014).

Za druhou etapu ve vývoji prostorového přístupu ke studiu kriminality se dá považovat počátek 20. století a zejména autoři spojovaní s proudem urbánní ekologie známého podle svého působiště jako tzv. Chicagská škola (Barilik, 2014). Mezi nové přístupy ve výzkumu kriminality patřilo studium zločinnosti v rámci města a důraz na vliv prostředí na páchání trestné činnosti, spojené se studiem indikátorů prostředí, sociálně prostorových odlišností různých částí měst a analýzami lokalit přesných míst trestných činů nebo bydlišť pachatelů. Oblastem, ve kterých dochází k vyššímu počtu trestných činů byla v tomto období věnována zvýšená pozornost, stejně jako charakteristikám pachatelů. Studie v tomto období využívaly zejména kvantitativní data a na ně navázané statistické analýzy. Ty zaznamenaly velký rozmach hlavně po druhé světové válce v souvislosti s kvantitativní revolucí. Díky novým technologickým možnostem vzrostl zájem o různé komplexnější statistické modely, využívaly se tak nové metody jako například regresní analýzy či faktorové analýzy. Hlavními typy výzkumu byly v tomto období areálové a ekologické studie. Podle Jíchové (2014) se areálové studie zaměřují na identifikaci vzorců v prostoru, např. distribuce různých druhů trestných činů v prostoru a využívají různé kartografické techniky a častou pomoc počítačů. Ekologické studie hledají souvislosti mezi kriminalitou a dalšími socioekonomickými a jinými ukazateli obyvatelstva či území. Tím navazují na studie z 19. století a rozšiřují je díky lepším možnostem zpracování dat. Důležitými výstupy z tohoto období, ovlivňující pozdější výzkumy kriminality, jsou model zonálního členění města, teorie sociální dezorganizace (více např. Barilik, 2014) a analýzy bydlišť pachatelů (Jíchová, 2014).

Třetí etapa, odehrávající se na konci 20. století, se dá již označit za období moderní teorie environmentální kriminologie. Na rozdíl od první a druhé etapy, ve kterých převažovalo zaměření na sociologické aspekty prostředí, resp. sociální problémy způsobující rozdíly v míře kriminality mezi různými místy, tuto etapu (environmentální kriminologie) zavedla dle Jíchové (2014) tehdejší kritika prostorové vědy a její odlidštěnosti a falešné objektivitě k hledání alternativních přístupů a pozornost se přesunula na jedince v kontextu prostředí. Podle Barilika (2014) tak nastává přesun ze sociologického přístupu k přístupu geografickému. Argumentuje tvrzeními Andersena a kol. (2010), že se začíná zkoumat, jak se jedinci pohybují v rámci určitého prostředí, namísto posuzování toho, jak je jednotlivec ovlivněn sociologickými podmínkami svého okolí (Barilik, 2014). Ve studiích se tedy objevuje zaměření na odhalování souvislostí mezi zločinem, obětí zločinu a místem, kde se čin odehrál. Podle environmentální kriminologie je velmi důležité rozhodování jedince v daném čase a na daném místě. Pro uskutečnění zločinu je nezbytné kromě motivovaného pachatele, přítomnosti oběti a nepřítomnosti hlídače nebo ochránce právě také vhodné místo a vhodný čas. Výzkum se tedy zaměřuje na chování pachatelů v prostředí, jež vytváří příležitost k trestnému činu (Barilik, 2014; Jíchová, 2014). Upozadil se tedy názor, že kriminalitu lze potlačovat posilováním materiálních standardů zejména chudých a znevýhodněných skupin obyvatel, i z důvodu toho, že rehabilitační a preventivní programy pro mládež, svobodné matky, drogově závislé a pro trestané osoby či investice do sociálního bydlení a pracovních příležitostí pro dlouhodobě nezaměstnané za cílem snížit kriminalitu až na výjimky selhávaly. Environmentální kriminologické teorie lze označit za oportunistické teorie, důležité je v nich jak se jedinec chová v daný čas na daném místě (Barilik, 2014). V souladu s těmito tvrzeními jsou i ústřední koncepty tohoto období, jmenovitě například koncept prevence kriminality prostřednictvím designu prostředí, teorie rutinních aktivit, teorie racionální volby (o ní více třeba Zapletal (2003)) nebo teorie struktury trestné činnosti či výzkum strachu ze zločinu, věnovaný diferencovanému vnímání strachu různými skupinami obyvatel (rozdíly mezi muži a ženami, mezi různými věkovými skupinami a podobně) (Barilik, 2014; Jíchová, 2014). Dochází i k většímu rozpracování konceptu tzv. hot spots, tedy míst s vyšší než průměrnou koncentrací kriminálních aktivit. Vychází z něj, že rozložení těchto aktivit není náhodné a lze taková místa nebo typy míst (parky, podchody, náměstí) vymezit a dále se snažit o odhalení příčin vyšší koncentrace kriminálních aktivit v těchto místech. Metody výzkumu se v tomto období obohatily především o kvalitativní metody.

Může za to právě změna přístupu a zaměření výzkumů, ale také vliv feministicky orientovaných geografek (otázky na rozdílnosti vnímání prostředí mezi pohlavími a podobně). Začínají se tedy více využívat dotazníková šetření, rozhovory, ankety, využívá se ale i kreslení mentálních map nebo psaní deníků. Vliv environmentální kriminologie zaměřil výzkum kriminality na ty nejmenší měřítka a výjimkou není ani studium na úrovni ulic, míst či zaměření na rezidenty nebo uživatele konkrétních míst. Společně s tím je stále hodnocena diferenciací úrovně a struktury kriminality nejen v rozličných regionech (např. urbánní a rurální prostředí, suburbie), ale také ve vztahu k sociálním a ekonomickým charakteristikám vázaných na rozdílná prostředí a jejich obyvatele. (Jíchová, 2014)

Od této etapy do současnosti nedošlo k výrazným změnám v přístupu k prostorovým aspektům kriminality. Rozvíjející se technologické možnosti, zejména geografické informační systémy, výpočetní technika a možnosti sběru dat nicméně umožňují kvalitnější zpracování map, vizualizaci, zpracování a správu dat. Výzkum se rozšiřuje do všech oblastí světa, původní koncepty z vyspělých zemí se aplikují na odlišná prostředí, kromě tradičních témat (prostorová distribuce vybraných trestných činů, výzkum pachatelů) začínají dostávat prostor i nová aktuální témata, například problematika světové bezpečnosti, terorismus. (Jíchová, 2014)

S pokrokem na poli informačních technologií pokračuje zájem o užívání a vývoj technik, které mohou pomoci vysvětlit výskyt kriminálních aktivit. Mezi nejvýznamnější počítačové nástroje patří doposud zejména geografické informační systémy (GIS) a (crime mapping) software k mapování kriminality. Tyto dvě příbuzné technologie usnadňují explorační prostorového rozmístění trestných činů (Murray, 2001). GIS přinesly hlavně velké zefektivnění práce s lokalizovanými daty, tedy daty obsahující GPS souřadnice s informací o přesném místě spáchání trestného činu. To umožnilo nejen zrychlení výzkumu, ale i zjednodušení úprav dat a předávání informací (Jíchová, Netrdová, 2017). Zobrazování umístění trestných činů na mapě, obecně známo jako „crime mapping“ (mapování kriminality), je jedním z užitečných aspektů GIS při snahách o prevenci zločinnosti. V kriminologické teorii se však podobné techniky využívají již téměř po dvě století, tedy dávno před GIS i výpočetní technikou (Murray, 2001). Za běžnou metodu dříve používanou policejními stanicemi k analýze a mapování kriminality se dá považovat zapichování špendlíků do mapy, jehož modernější verzí je v podstatě tzv. bodová (tečková) metoda. Tato metoda je jednoduchá a z části určitě účinná, protože je schopna poskytnout základní přehled o tom, kde ve městě k trestným činům dochází. Navzdory jednoduchosti takové metody, ji však zdaleka nemůžeme považovat za ideální, třeba kvůli časové náročnosti na její přípravu, nebo proto, že základní přehled, který tato analýza poskytuje je pouze vizuální a tudíž interpretovat prostorové vzorce nebo identifikovat vzorce jiných informací může být obtížné, zejména pokud pracujeme s velkým počtem trestných činů. Z takové mapy nemusí být ani dobře čitelné zda se na jednom místě událo více trestných činů, nebo pouze jeden a identifikaci ohnisek to může znepříjemnit. Také je samozřejmě takový způsob analýzy fyzicky omezen, mapa se špendlíky se nedá jednoduše přesouvat, data z terénu je nemožné rychle doplnit atd. (Bates, 1987; Jíchová, Netrdová, 2017)

Mapování kriminality je však nyní již důležitá technická součást moderních policejních postupů. Policejní analytici opakovaně mapují incidenty trestných činů, jednak proto aby detekovali obecné vzorce rozmístění kriminality k soustředění policejních sil a snah o prevenci, tak také proto aby identifikovali a dopadli specifického pachatele, který provádí trestné činy. Informace získané z takových analýz jsou poté aplikovány různě, od soustředění hlídek více do ohnisek trestných činů za účelem cílené prevence, přes sledování chování sériových pachatelů, které policie chce dopadnout, až třeba k mapování dopravních nehod (Levine, 2006). Analýza trestných činů, je observační a ne experimentální, což znamená, že to nejsou vědečtí pracovníci, kteří specifikují geografickou lokaci jejich pozorování. Proto je lokace trestných činů velmi důležitá informace pro výzkumníka, zejména když charakter místa nebo prostoru, ve kterém k trestnému činu došlo je brán jako faktor, který ovlivňuje procesy generující trestné činy (Huebner, Bynum, 2016).

Z odborné literatury zabývající se prostorovou analýzou kriminálních činů se můžeme dozvědět o široké škále analytických postupů, které je možné použít například pro identifikaci lokalit se zvýšenou

kriminalitou. Za zmínku stojí například práce Anselina a kol. (2000, 2005), ti se ve svých pracích zaměřují zejména na otázky související s metodologií v prostorových statistických analýzách dat za kriminalitu, zvláštní pozornost věnují popisným analýzám (exploratory spatial data analysis – ESDA), které považují za základ všech empirických analýz kriminality ve vztahu k místu. Mezi další důležitou literaturu patří práce, kterou napsal Eck a kol. (2005). Ta se zaměřuje především na techniky a software k analýze koncentrací (hotspot analysis). Při výčtu základní literatury nesmím opomenout ani Neda Levine, který stojí za softwarem CrimeStat. V jeho pracích (např. 2006, 2008, 2013) se věnuje rozličným technikám a metodám k analyzování prostorových aspektů a distribuce kriminality. Jako užitečné hodnotím i další publikace jako Murray (2001), Wortley, Mazzerolle (2008) nebo Hirschfield, Bowers (2001), které přináší poměrně ucelené pohledy na analýzu kriminality v kontextu prostředí.

3.1. Výzkumy zaměřené na Česko

V české geografii téma kriminality nepatří ke stěžejním oblastem zájmu. Výzkumů zabývajících se analýzou rozmístění zločinnosti, které se snaží identifikovat základní prostorové vzorce rozmístění kriminality nenalezneme v českém prostředí mnoho. Studiu percepce, pocitu bezpečí, vlivu prostředí na páčání kriminality se věnuje několik prací. V takových výzkumech se využívají především kvalitativní metody jako dotazníková šetření nebo rozhovory, kvantitativní metody se spíše neobjevují. Mimo to se řada děl zaměřuje svým tématem spíše na charakteristiky typů trestných činů, výkon práva, či obecně kritizuje vývoj společnosti. Prostorové aspekty kriminality se v takových pracích ovšem tolik nerozebírají. (Jíchová, 2013; Jíchová, 2014)

Pro současný výzkum kriminality v Česku znamenala zlom 90. léta. Změny v politické, socioekonomické a technologické oblasti s sebou přinesly i změny úrovně a struktury kriminality. Na změny v oblasti kriminality (hospodářská kriminalita), nové typy trestných činů (počítačová kriminalita) reagují odborníci řadou výzkumů, které se zpracovávají především v rámci Institutu pro kriminologii a sociální prevenci (IKSP) (Jíchová, 2014). Tyto výzkumy se zaměřují na organizovaný zločin, oběti a pachatele zločinů, drogovou problematiku a podobně. IKSP poskytuje podrobný popis jím provedených výzkumů v pravidelných publikacích, viz poslední taková publikace Zeman (2017). Zároveň IKSP vydává každoroční zprávy o trendech v oblasti kriminality. Dle Jíchové (2014) se nadále rozšiřují nejen zkoumaná témata v oblasti kriminality, ale diverzifikují se i přístupy, metodiky a územní měřítko výzkumu. Výzkumy v Česku začínají intenzivněji reflektovat zahraniční studie. Analýzy úrovně a struktury kriminality na různých měřítcích (typicky na úrovni krajů, nižší měřítko než okresy jsou raritou) jsou jedním ze dvou základních přístupů ke studiu kriminality v českém prostředí, tím druhým jsou analýzy jednotlivých témat z oblasti kriminality, často doplněné případovými studiemi. (Jíchová, 2014) Ve výzkumech IKSP se využívá často kombinace metod, většinou jsou to však metody kvalitativní (dotazníky, rozhovory, obsahové analýzy či terénní šetření), prostorové analýzy jsou využívány velmi zřídka, respektive vůbec ve výzkumech IKSP poslední dekády. Ačkoliv jedním z výzkumných úkolů střednědobého plánu v letech 2012–2015 byl i úkol s názvem „Regionální rozložení a podmíněnost kriminality, problematika a výskyt tzv. hot-spots“, výsledná publikace od Holase a kol. (2016) byla zaměřena spíše na míru obav z kriminality, pocit bezpečí, spokojenost se životem v lokalitě, a v neposlední řadě na viktimizaci respondentů a byly využity kvalitativní metody výzkumu.

V posledních letech dochází k využívání geografických informačních systému v některých českých výzkumech zaměřených na kriminalitu. Jíchová (2014) se domnívá, že využití GIS se v posledních letech rozvíjí ve větší míře a jako příklady udává Karbana (2008), či studie ve sborníku Bezpečnostní situace v České republice. Konkrétní místa trestných činů, tedy přesně lokalizovaná data, kterými se zabývá tato práce, se však stále téměř nevyužívají, navzdory tomu, že by pro řadu analýz byla právě tato data nejhodnější, a která by mj. umožnila aplikovat poznatky a koncepty environmentální kriminologie. (Jíchová, 2014) Je tomu tak zejména kvůli problémům s dostupností dat a jejich kvalitou (Ivan, 2018) o čemž se více zmiňuji v kapitole 4. Za zmínku stojí také internetový mapový portál mapakriminality.cz, který poskytuje pohled na aktuální úroveň kriminality a její vývoj a strukturu až do úrovně obvodních

(místních) oddělení policie. Jako základní nástroj pro srovnání mezi územními jednotkami používá tento portál index kriminality (počet zjištěných skutků za zvolené období na 10 000 obyvatel). Data zveřejněná na tomto webu jsou oficiální policejní statistiky, které poskytuje Policie ČR každý měsíc. (mapakriminality.cz, 2019)

Využíváním přesně lokalizovaných dat za trestné činy se mimo jiné zabývají autoři Horák a Ivan. Ti publikovali několik prací na toto téma, kde se zaměřují jak na metodiku analýz těchto dat, tak na samotné analýzy skutečných dat za trestné činy na mikro-úrovni, konkrétně ve městě Ostrava. V těchto pracích Horák a Ivan (např. 2012, 2013, a v navazujícím díle 2018, které dále rozšiřuje předchozí zjištění) analyzují prostorové rozmístění trestných činů ve městě Ostrava, k tomu využívají prostorově statistické metody hot-spot analýzy a interpolační metodu jádrových odhadů. Pro potřeby Magistrátu města Ostravy byla v rámci těchto studií připravena sada map, dokumentující vývoj jednotlivých kategorií a vybraných tříd trestných činů a přestupků (např. mapy výskytu a meziročního vývoje násilných trestných činů, loupeží, výtržnictví, krádeží, vloupání a dalších) na území tohoto města. (Horák, Ivan, 2012) Kromě těchto analýz, vypracovali stejní autoři také metodické dílo Metodika identifikace anomálních lokalit kriminality pomocí jádrových odhadů. To má dle slov autorů za cíl „doporučit standardizovaný postup využívání metody plošných jádrových odhadů pro identifikaci anomálních lokalit kriminality“. Jako hlavní metodu pro identifikaci anomálních lokalit, které bývají často nazývány jako hot spots, v tomto díle preferují metodu jádrových odhadů (kernel density estimation) či jádrového vyhlazení. Výsledky takových analýz bývají autory označovány také jako „heat maps“. Tato metoda je populární pro svou vizuální názornost právě hlavně v oblasti analýz kriminality. Obecně metoda jádrových odhadů přiřazuje každému bodu v mapě odhad intenzity na základě vzdálenosti k ostatním událostem. (Ivan, Horák, 2016)

Jiným autorem využívajícím přesně lokalizovaná data o trestných činech je Kostka. Ve své diplomové práci (2016) se zabýval analýzou rozmístění krádeží automobilů na území města Liberce a možností aplikace technik situační prevence. Data na tento výzkum byla poskytnuta Policií ČR, obsahovala více než tisíc přesně lokalizovaných (GPS souřadnicemi) záznamů o krádežích automobilů ta za roky 2011 až 2014. Oblasti s větším, než průměrným výskytem krádeží automobilů identifikuje Kostka pomocí metody jádrového odhadu. Práce měla za cíl pomocí map znázorňující hot spoty ilustrovat změny v prostorové distribuci hot spotů ve sledovaných letech, a rovněž navrhnou preventivních opatření, které by bylo vhodné ve vybraných hot spotech přijmout. (Kostka, 2016)

Kapitola 4

Data o kriminalitě

K analyzování kriminality jsou potřeba data o kriminalitě. Ta jsou ve světě různě dostupná a pojí se s nimi také několik otázek, problémů a omezení, které je třeba si při práci s takovými daty uvědomit. Důležitým omezením je tzv. latentní kriminalita, což je termín pro kriminalitu nehlášenou a nevidovanou ve statistikách. Taková kriminalita může dle odhadů u některých typů trestných činů dokonce převyšovat činy ve statistikách evidované. Pokud bychom chtěli při analýzách odhadnout nevidované trestné činy, nastane problém s jejich přesným umístěním. S lokalizací se pojí několik problémů i u evidovaných dat, které se pojí například s přesností zápisu. Ta je u jednotlivých záznamů rozdílná, někdy může být problém v nejednoznačném určení místa trestného činu. To může nastat hned z několika důvodů. Přesné místo činu totiž nemusí být vůbec známo ani přibližně, oběť si nemusí vybavit, kde k činu došlo nebo to ani přesně neví – významnou část takových případů tvoří činy (kapesní krádeže) při cestování, zejména městskou hromadnou dopravou. Dalším důvodem, proč může docházet k chybám a nejednoznačné lokalizaci trestných činů je, že popis místa činu je v evidenci příliš obecný vzhledem k velikosti objektu (parkoviště u obchodu, náměstí, nákupní centrum apod.) a to vede ke koncentraci míst činu do jistých bodů. Také mohou být některé evidované skutky spojeny s páčáním trestné činnosti na více místech, např. ve více obchodech nebo ve více garážích. Důležité je také znát podrobněji systém evidence kriminality, se kterým pracujeme, protože v některých systémech jsou lokalizovány trestné činy podle míst, kde k daným trestným činům došlo, v jiných zase podle místa bydliště pachatele nebo podle místa bydliště oběti. Problém může nastat i u procesu geokódování, kdy se záznamům přiřazují přesné souřadnice podle zadané adresy, nebo také u vzájemné nekompatibility systémů, projevující se například v jinak používaných zkratkách, v jiných názvech míst (ulic, měst) či neaktuálnosti map. (Ivan, Horák, 2012; Jíchová, Netrdová, 2017)

Co se týče dostupnosti dat v Česku, také zde není situace úplně bezproblémová. Přesně lokalizovaná data o kriminalitě sbírají v Česku policejní složky, a to od roku 2013. Využívají je však hlavně v rámci vlastních projektů a bezpečnostní praxe. Pro veřejnost tato data nejsou k dispozici a pro odborníky k základnímu i aplikovanému výzkumu je dostupnost těchto dat velmi omezená (Šimon, Jíchová, 2018a). Autoři, kteří získali lokalizovaná data za trestné činy, tak učinili v rámci žádosti či smlouvy o spolupráci s příslušným orgánem policie. Například data o krádežích automobilů v Liberci, se kterými pracuje Kostka (2016) nebo Jíchová, Netrdová (2017), byla poskytnuta Krajským ředitelstvím policie Libereckého kraje. Tato data o krádežích osobních automobilů v letech 2011 až 2014 dle Jíchové a Netrdové (2017) vykazují 90% úplnost a správnost, což považují za standardní a přijatelnou chybovost u mapování trestné činnosti. Po očištění od nežádoucích záznamů obsahoval tento soubor údaje o 1299 krádežích. Jako druhý příklad získaných dat o kriminálních činech v českém prostředí slouží několik studií od autorů z Vysoké školy báňské v Ostravě Ivana a Horáka. Ti ve svých pracích (např. 2012, 2013, 2018) analyzují data, která jim byla poskytnuta v rámci smlouvy o spolupráci v oblasti geoinformačních technologií mezi Vysokou školou báňskou – Technickou univerzitou Ostrava a Policií ČR. Pro studii z roku 2012 (která byla poté aktualizována a rozšířena v dalších pracích) byl zdrojem dat

anonymizovaný export z evidence trestné činnosti, poskytnutý Krajským ředitelstvím Policie ČR v součinnosti s Policejním prezídiem ČR a dokumentovala výskyt trestných činů i přestupků na území města Ostrava pro rok 2009. Data obsahovala informace o adresních atributech (ulice, číslo popisné, atd.) spolu se subjektivním popisem lokality tvořený policistou. Čas skutku byl vymezen datem a časem zahájení a ukončení páčání skutku, kódové označení klasifikace trestného činu pak sloužilo k věcnému vymezení (Ivan, Horák 2012). Existuje ještě několik výzkumů využívající lokalizovaná data omezená časově či typem kriminality, dle Šimona a Jíchové (2018a) je však obecně výzkum geolokalizovaných dat o kriminalitě a jeho využití v experimentální kriminologii v Česku v počáteční fázi (Šimon, Jíchová, 2018a). I když má detailní výzkum kriminality s lokalizovanými daty (např. výzkum míry škodlivosti kriminality) velký potenciál (kromě oblasti bezpečnosti obyvatelstva či prevence kriminality) také třeba v pojišťovnictví či realitním sektoru, není v Česku ani v jiných post-komunistických zemích v současnosti výzkum na toto téma velmi rozvinutý (Šimon, Jíchová, 2018b).

Prezentace dat veřejnosti ve srozumitelné podobě a představení detailních dat o kriminalitě je to, co i přes volání o potřebě takového nástroje ze stran zástupců samospráv, bezpečnostních složek či zástupců vzdělávání v bezpečnostní praxi, dle Šimona a Jíchové (2018b), v Česku chybí. V některých městech v Česku se začínají využívat k prezentaci dat online mapy. Ty se obecně mohou lišit v podrobnosti uvedených informací. Příkladem takových map jsou například ty z Kolína nebo Uherského Hradiště, využití online map je u nás však spíše výjimkou. V oblasti online map jsou rozvinutější například v USA. New York nebo Chicago dávají dobrý příklad online map v USA, na mapě si lze zvolit typ trestného činu, časové určení, na online mapě Chicaga je dokonce k vidění jméno oběti včetně odkazu na články u vražd. (Jíchová, Netrdová, 2017)

V některých jiných státech (např. USA) je i dostupnost dat za kriminalitu na lepší úrovni. Již zmíněné město Chicago dává na svém webovém portálu <https://data.cityofchicago.org/> volný přístup k datovým souborům z mnoha oblastí včetně kriminality. Tento datový portál umožňuje dle vlastních slov „najít městská data, najít fakta o sousedství, vytvořit mapy, grafy o městě a stáhnout data zdarma pro vlastní analýzy“ (City of Chicago, 2019). Podobný zdroj skýtá např. i americká metropole Los Angeles, data za kriminalitu jsou zde poskytována tamním policejním oddělením. Data jsou volně přístupná z webu <https://data.lacity.org/>, sám starosta Los Angeles Eric Garcetti říká: „Sdíleme tyto data s veřejností, abychom zvýšili transparentnost, zodpovědnost a zákaznické služby, a abychom poskytli společnostem, jednotlivcům a neziskovým organizacím schopnost použít toto množství informací ke zlepšení života v našem městě.“ (City of Los Angeles, 2019) Z podobně dostupných dat z jiných států stojí za zmínku například data z Velké Británie, které na svém webu (<http://data.police.uk/>) zdarma zveřejňuje britská policie a získání je velmi jednoduché. (Veronesi, 2015)

V této práci jsou k testování metod prostorové statistické analýzy bodových dat využita bodová data za trestné činy z amerického města Chicago. Je tak učiněno zejména z důvodu snadné dostupnosti těchto dat. Využívaný datový soubor obsahuje podrobná lokalizovaná data za trestné činy od roku 2001 do současnosti a soubor je pravidelně aktualizován. Data konkrétně obsahují prostorovou informaci ve formě X a Y souřadnic. Obsahují také informaci o datu a přibližném času, ve kterém ke zločinu došlo. Velmi cenná je informace o typu trestného činu. Každému zločinu je přiřazen jeden ze 32 typů. V této analýze budou nejvíce využity data za krádeže motorových vozidel (n=5 253), vloupání (n=6 308) a ublížení na zdraví (n=25 585) (vše za rok 2018), soubor však dále obsahuje například data za vraždy, sexuální kriminalitu a další. Dále soubor obsahuje popis skutku a detailnější popis místa, kde ke skutku došlo (ulice, budova atd.). V neposlední řadě také informaci o tom, zda došlo k zadržení či v jakém policejním obvodu či jiné jednotce k činu došlo. Pro práci jsou také použity vrstvy (soubory shapefile) budov, ulic a parků ze stejného datového portálu.

Kapitola 5

Metody prostorové statistické analýzy bodových dat

Tato kapitola obsahuje představení vybraných metod prostorové statistické analýzy bodových dat. Metody jsou aplikovány na bodová data o trestných činech (detailněji popsána v kapitole 4) z amerického města Chicago. Jednotlivé metody jsou v textu představeny, diskutovány jsou jejich výhody či nevýhody a omezení, dále jsou také u některých metod zmíněny studie z různých oblastí výzkumu, které tyto metody využívají. Metody jsou rozděleny do tří základních skupin – 1. skupinu tvoří metody pro zjištění prostorové distribuce, sem patří základní metody explorační prostorové analýzy jako geografický střed (mean center), elipsa směrodatné odchylky (standard deviation ellipse) nebo geografický medián (median center) a další, zároveň do této skupiny řadím metodu nejbližšího souseda (nearest neighbor analysis) ke zjištění zda se body shlukují či nikoliv; 2. skupina obsahuje metody pro analýzu koncentrací a hledání tzv. ohnisek (hot spots). Sem zahrnuji metody jako modus, fuzzy mode, metoda hierarchického prostorového shlukování nejbližších sousedů nebo analýzu STAC; 3. skupina zahrnuje metody prostorového modelování a interpolace.

Všechny tyto metody mají za cíl povědět nám něco o prostorovém rozmístění zkoumaného jevu. O tom jestli jsou studované body rozdělené náhodně, či tvoří shluky nebo naopak tvoří rozptýlenou distribuci. V případě zkoumání kriminality představené v této práci, jevy s ní spojené nejsou v prostoru rozmístěné náhodně a existuje významné rozdíly mezi zeměmi, regiony, městy, ale právě i ulicemi jednoho města, na které se konkrétně v této práci zaměřuji. Znalost rozmístění kriminality ve městě je velmi důležitá pro samotné kriminální vyšetřování, avšak i pro tvorbu znalostní báze kriminologie. Jak uvádí Šimon a Jíchová (2018a): „poznání prostorově vázaných kriminogenních faktorů umožňuje soustředit omezené zdroje do míst, kde lze docílit nejvyššího dopadu preventivních či punitivních opatření. Prostorové faktory mají výraznou určující schopnost pro koncentraci kriminality, řada typů kriminality má jasné lokalizační předpoklady. Polovina kriminality je ve velkých zahraničních městech realizovaná ve dvou až čtyřech procentech uličních segmentů města, proto je důležité znát vzorce prostorové distribuce kriminálních jevů. (Šimon, Jíchová, 2018a)

Při výzkumu kriminality, na který se tato práce zaměřuje, je základní cíl prostorových analýz nalézt vzorce prostorového rozmístění a tuto znalost pak využít v hledání příčin vzniku zločinu. Například pokud se velký počet krádeží stává v okolí určitého místa, kriminology bude zajímat prostředí tohoto místa nebo jeho okolí. Vlastnosti prostředí mohou poté kriminologovi umožnit pátrat po podobných vlastnostech na jiných místech a zkoumat, zda se větší počet zločinů vyskytuje i na takových místech (Leung, 2015). Tato práce se však omezuje pouze na hledání prostorových vzorců, ambice zkoumat prostředí nemá.

V odborné literatuře je představena široká škála analytických postupů, které je možné využít pro identifikaci lokalit se zvýšenou koncentrací bodových dat. Mnoho z nich jsou metody, jež pracují s daty agregovanými do polygonů. Nad těmito agregovanými daty je pak možné posuzovat, nakolik se shlukují

jednotlivé polygony s vysokými, resp. nízkými hodnotami. K tomuto slouží metody pro hodnocení globální a lokální prostorové autokorelace (LISA – Moranovo I, metody G_i a G_i^* apod.) (Ivan, Horák, 2012). Ty však považuji nad rámec této práce, jelikož jsou zde představovány výhradně metody jež pracují přímo s bodovými daty jako takovými.

5.1. Základní popisné statistiky

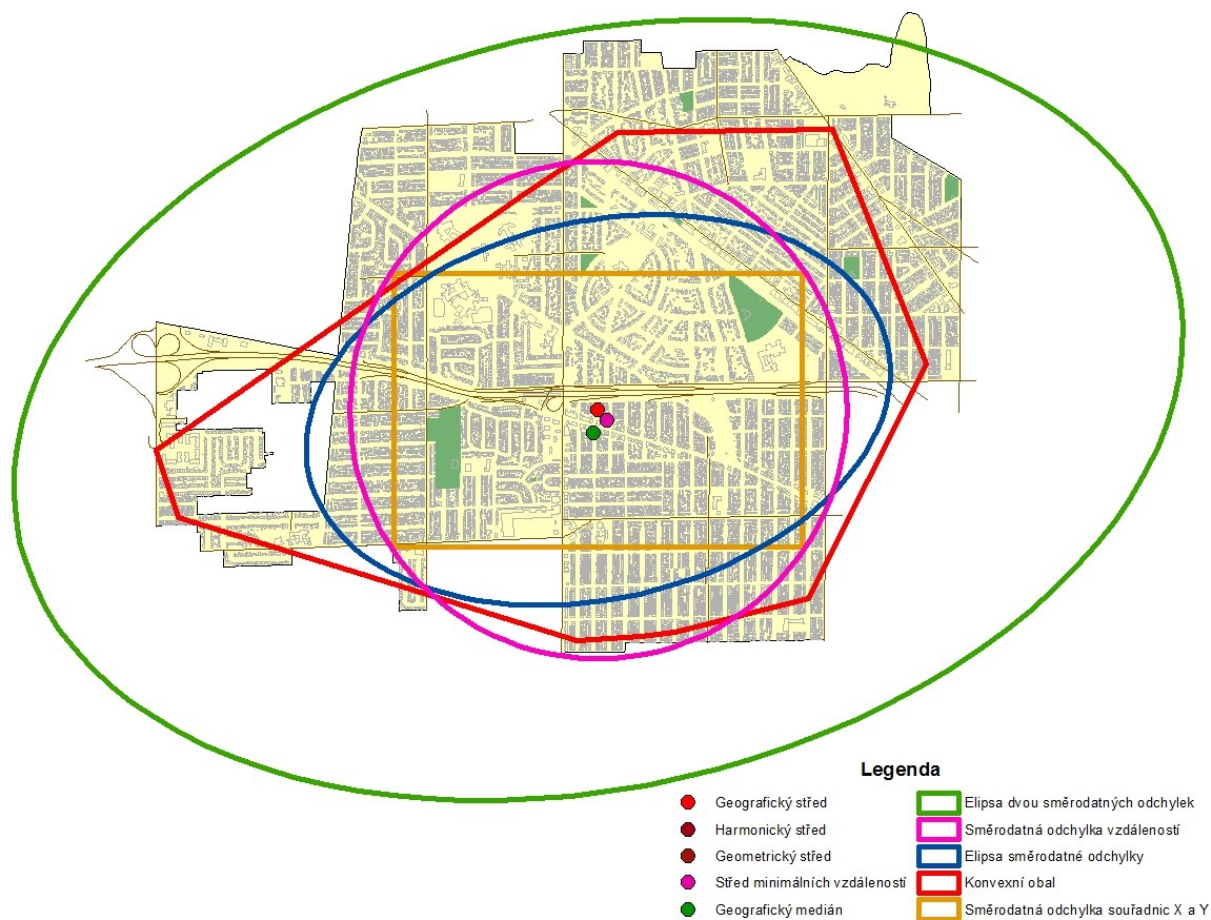
V této části jsou představeny základní prostorově popisné statistické metody. Tyto základní metody prostorové analýzy bodových dat popisují obecné vlastnosti prostorového rozmístění a představují první krok prostorové analýze dat. Zahrnují v sobě globální charakteristiky, popis regionálních rozdílů a popis malých koncentrovaných shluků (clusterů). Obecně tyto metody slouží k hledání středních hodnot a celkové prostorové distribuce v souboru a mohou být využity k srovnání různých typů bodových distribucí (např. různých typů trestných činů) nebo ke srovnání stejného typu distribuce v různých časech, například mezi jednotlivými měsíci. (Eck a kol., 2005; Levine, 2006).

Do této skupiny řadím metody, jejichž výstupem jsou body středových hodnot. Prvním z nich je **geografický střed (mean center)**. To je nejjednodušší statistika, souřadnice geografického středu jsou průměrem souřadnic X a průměrem souřadnic Y všech bodů. Těm se může být přiřazena i váha (Smith, Bruce, 2008). Této metody může být využito právě například pro srovnání prostorové distribuce jednotlivých typů trestných činů (Eck a kol., 2005). Další středovou hodnotou je **střed minimálních vzdáleností (mean center of minimum distance)**. To je bod, ve kterém je součet vzdáleností ke všem ostatním bodům nejmenší. **Geografický medián (median center)** je bod se souřadnicemi mediánu X souřadnic a mediánu Y souřadnic a jelikož je více odolný vůči odlehlým hodnotám než geogr. střed, může se využívat v souborech ovlivněných právě odlehlými hodnotami. Oproti ArcGISu nabízí software CrimeStat navíc výpočet **harmonického (harmonic mean) a geometrického (geometric mean) středu**. V analýze kriminality jsou v praxi tyto hodnoty téměř totožné s geografickým středem, jejich použití tedy není nutně obohacující (Smith, Bruce, 2008). Avšak větší využití našly třeba v ekologii při studiu oblastí zvířecí aktivity, například Dixon, Chapman (1980).

Metody jejichž výstupem jsou polygony, které nám vypovídají o úrovni rozptylu v souboru, tvoří **směrodatná odchylka souřadnic X a Y (standard deviation of the X and Y coordinates)**, tato metoda vytvoří polygon, který uzavírá oblast kde se protínají čtyři úsečky: jednu směrodatnou odchylku nad průměrem na ose X, jednu směrodatnou odchylku pod průměrem na ose Y, jednu směrodatnou odchylku nad průměrem na ose Y, jednu směrodatnou odchylku pod průměrem na ose X; **směrodatná odchylka vzdáleností (standard distance deviation)** je dvoudimenzionální ekvivalent směrodatné odchylky, která vytvoří kruh kolem průměru, jehož poloměr tvoří směrodatná odchylka. Ač dává užitečnou informaci prostorovém rozptýlení bodů, nebere v úvahu směr, který může být charakterizován mnohem širším rozšířením v určitém směru ve srovnání s jiným. Proto soubor bodů lépe charakterizuje **směrodatná elipsa odchylek (standard deviation ellipse)**, která je určena směrodatnou odchylkou vzdáleností všech bodů ke geografickému středu a její velikost a tvar pomáhá vysvětlit stupeň a směr rozptylu; dále **konvexní obal (convex hull)**, což je polygon, který obsahuje 100% bodů a slouží k zobrazení tvaru distribuce. (Eck a kol., 2005; Lisitsin 2015; Smith, Bruce, 2008; Levine, 2013)

Tyto centrografické metody zjišťování prostorové distribuce mohou při studiu kriminality odpovědět na několik základních otázek. Na obrázku č. 1 jsou zobrazeny vypočítané zmíněné body a polygony. Z výsledků si můžeme odpovědět například na otázku, kam by bylo nejhodnější zaparkovat policejní auto, aby mělo nejbližší k incidentům. Bylo by to na jeden ze středových bodů, nejlépe na střed minimálních vzdáleností. Díky elipsám směrodatných odchylek můžeme predikovat, že k dalšímu trestnému činu dojde pravděpodobně právě uvnitř elipsy, téměř jistě pak uvnitř konvexního obalu. (Smith, Bruce, 2008)

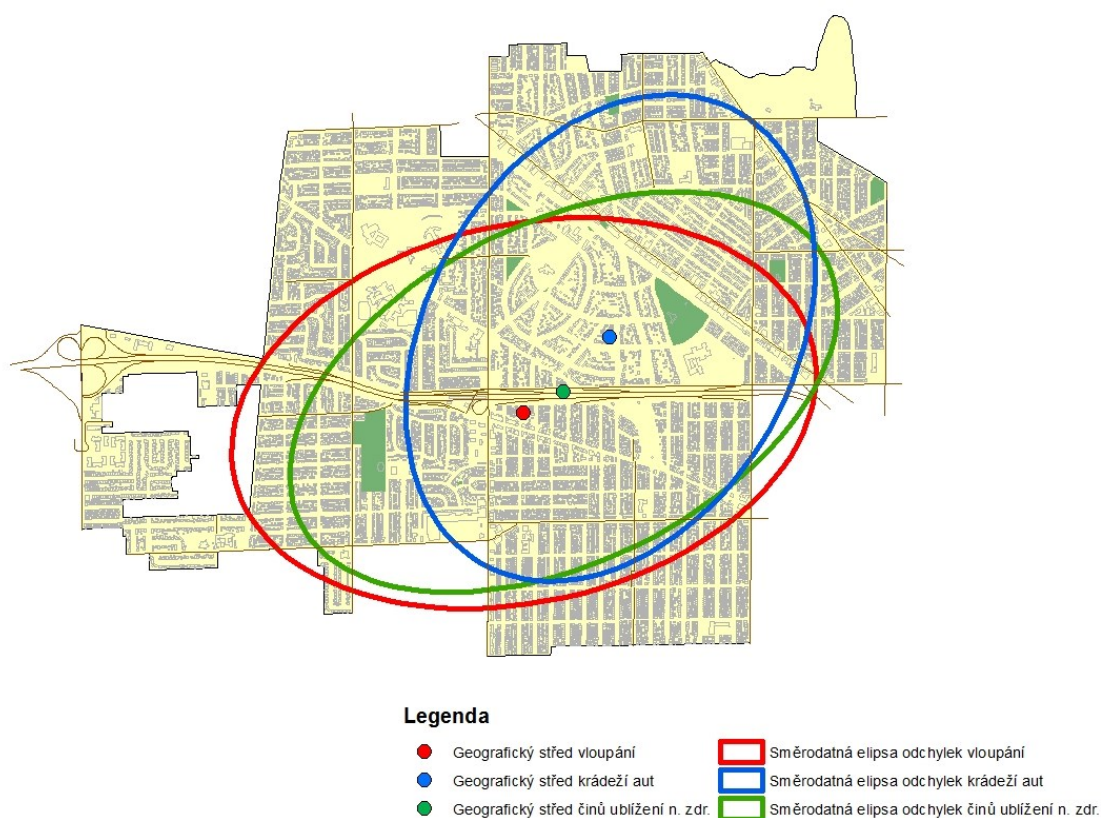
Obr. 1: Ukázka centrografických statistik rozmístění vloupání v Norwood Parku v Chicagu v roce 2018



Zdroj dat: City of Chicago (2019); vlastní vypracování; pozn.: geografický střed překrývá v tomto případě harmonický i geometrický střed

Obrázek č. 2 zobrazuje rozmístění tří různých typů trestných činů pomocí geografického středu a směrodatné elipsy odchylek. Pomocí takového srovnání jsme schopni identifikovat, že krádeže aut se odehrávají více na severovýchodě než zbylé dva typy trestných činů a směrodatná elipsa odchylek krádeží aut má lehce odlišný směr než ostatní dva typy.

Obr. 2: Srovnání tří typů trestných činů (vloupání, krádeže aut, ublížení na zdraví) v Norwood Parku v Chicagu z roku 2018 a jejich prostorové distribuce pomocí geografického středu a směrodatné elipsy odchylek



Zdroj dat: City of Chicago (2019); vlastní vypracování

Všechny tyto centrografické a prostorově popisné statistiky mají široké využití v celé škále výzkumů, jelikož jejich výsledky mohou poskytnout základní přehled o prostorové distribuci zkoumaných jevů. Mimo kriminologii a obecně sociální geografii se podobných technik může využít např. v již zmíněné ekologii či v geologii. Jako příklad uvedu práci Vladimira Lisitsina (2015), který, kromě jiných metod, využívá právě exploračních prostorově popisných statistik a popisuje využití těchto technik při výzkumu minerálních ložisek.

Jsou však případy, kdy nám mohou tyto metody dát „falešné“ povědomí o souboru dat. Pokud se body shlukují například do dvou shluků, výpočty nebudou reprezentovat skutečné rozmístění bodů, jelikož budou počítány z obou shluků, které mají uvnitř sebe určitou distribuci bodů, ale dohromady se tato distribuce bude jevit jinak. V takových případech je vhodné utvořit dva separátní soubory a prostorovou distribuci zkoumat odděleně (Smith, Bruce, 2008).

Možná nejdůležitější metodou z této skupiny je **metoda nejbližšího souseda (nearest neighbor analysis)** k odhalování typu prostorové distribuce, tedy jestli data pochází z rozmístění náhodného, či tvoří shluky, nebo jsou více rozptýlená než při náhodném rozmístění. Metoda je založena na vzdálenostech mezi sousedy, data se v prostoru shlukují pokud jsou vzdálenosti mezi sousedy menší než očekávané vzdálenosti za náhodného rozdělení. Pro každý bod je vypočítána vzdálenost k nejbližšímu sousedovi, součet všech těchto vzdáleností je vydělen počtem bodů v souboru – to je hodnota skutečné průměrné vzdálenosti mezi nejbližšími sousedy, poté se vytvoří náhodné rozmístění stejného počtu bodů pokrývající stejné území a vypočítá se stejná hodnota jako pro skutečné body – to je hodnota očekávané průměrné vzdálenosti mezi nejbližšími sousedy. Index nejbližšího souseda

(NNI – nearest neighbor index) je poměr skutečné hodnoty průměrné vzdálenosti mezi nejbližšími sousedy ku hodnotě očekávané průměrné vzdálenosti mezi nejbližšími sousedy. Pokud se tedy NNI rovná 1, body pochází z náhodného rozmístění, v případě že hodnota NNI je menší než 1, data vykazují shlukování v prostoru. Hodnota NNI větší než 1 ukazuje na rovnoměrné rozmístění bodů v prostoru. (Eck a kol., 2005) Tato metoda nám říká něco o tom, zda v souboru dochází ke shlukování, či nikoliv, avšak neříká nic o tom, kde v prostoru ke shlukům dochází, což může být velmi užitečná informace. K takovému cíli však vedou metody jiné, dále diskutované níže. (Anselin a kol. 2000) Pokud je známa oblast zachycení incidentů, k výpočtu NNI by měla být použita rozloha této oblasti. Pokud však není známa, měla by ke zjištění rozlohy zkoumané oblasti výskytu incidentů využít rozloha polygonu vypočítaného pomocí metody konvexního obalu (popsané výše), která dává relativně přesné výsledky, které realitu příliš nezkrusují. (Eck a kol., 2005)

Pokud bereme za oblast výskytu bodů určité administrativní území, mohou být ve skutečnosti nejbližší sousedé incidentů za hranicemi tohoto území, které v analýze neuvažujeme. V takovém případě metoda nejbližšího souseda nadhodnocuje pozorované vzdálenosti. Software CrimeStat nabízí korekci tohoto problému možnostmi obdélníkové nebo kruhové korekce hranic (ArcGis 10.6.1. tuto možnost u metody nejbližšího souseda nenabízí). Pokud je bod blíže k hranicím (obdélníku nebo kruhu) než k nejbližšímu sousedovi, je pak vzdálenost k hranicím brána jako upravená vzdálenost k nejbližšímu sousedovi a je s ní dále počítáno. Pokud např. analyzujeme zločin v území uvnitř velké metropolitní oblasti obklopené podobnými zločiny, dává smysl využít této korekce. (Smith, Bruce, 2008; Levine 2013)

Pro otestování této metody jsem zvolil krádeže motorových vozidel v americkém městě Chicago. Výpočet indexu nejbližšího souseda pomocí softwaru CrimeStat vyšel 0,769, což poukazuje na prostorové shlukování v souboru.

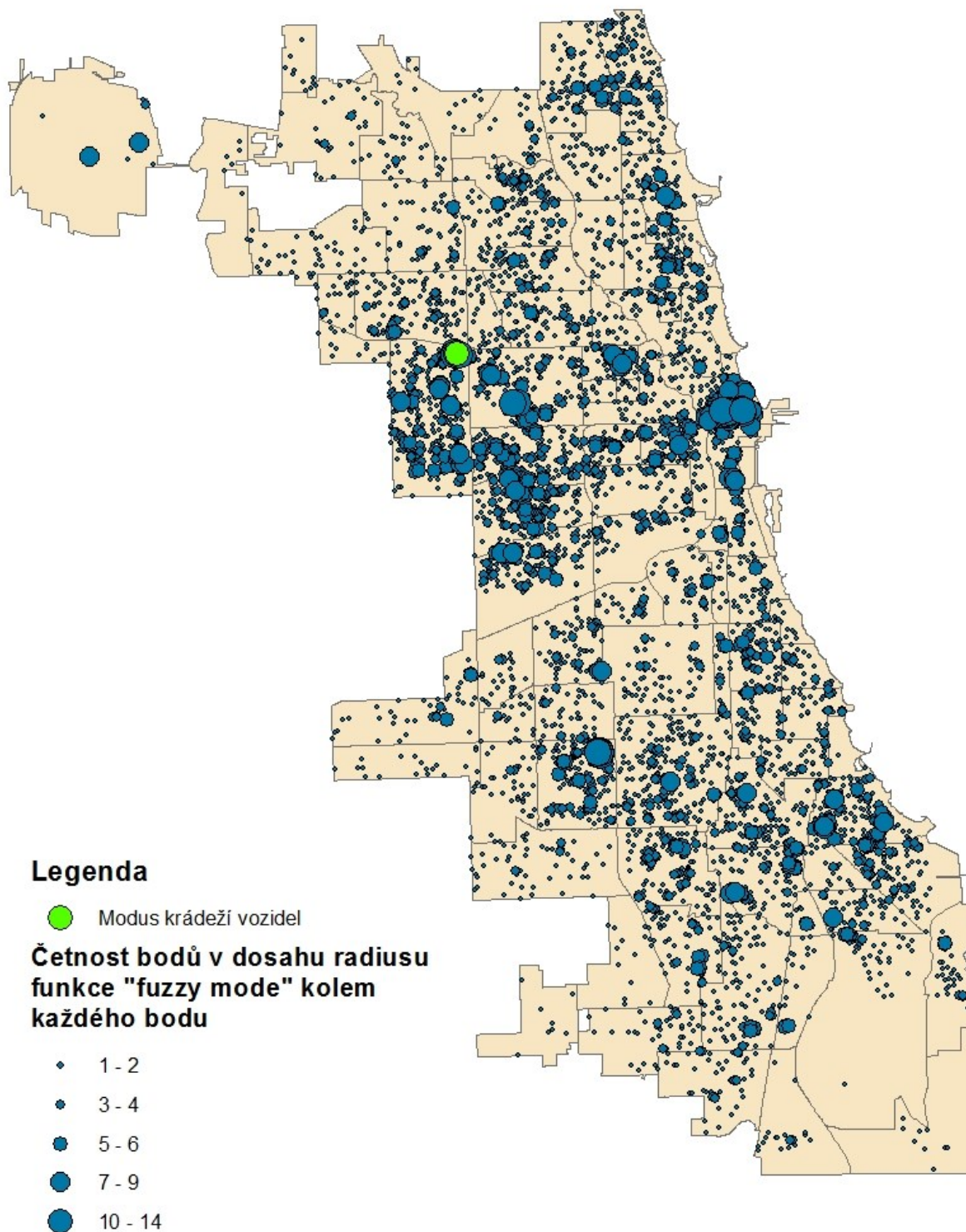
Podobná metoda je **Ripleyho K funkce**. Porovnává počet bodů (sousedů) v jakékoliv vzdálenosti s očekávaným počtem bodů pro náhodné rozmístění. Poskytuje tak test náhodnosti pro všechny vzdálenosti, od nejmenších po velikost celé zkoumané oblasti. Pokud je počet nalezených bodů vyšší než se očekávalo při náhodném rozmístění, ukazuje to na shlukování (Eck a kol., 2005). V analýze kriminality má tato metoda dle Smith a Bruce (2008) pouze malé využití, jelikož je složitější na interpretaci než metoda nejbližšího souseda a ta je v oblasti kriminality po většinu času dostačující. Každopádně využití Ripleyho K funkce našla například v biologii, Lagache a kol. (2013) využívají Ripleyho K funkci ke zkoumání prostorových distribucí v buněčných procesech.

5.2. Analýzy koncentrací

Pokud existují oblasti s vyšší koncentrací bodů než průměrnou, hovoříme o těchto oblastech jako o tzv. ohniskách či oblastech koncentrace (hot spots). Existuje několik metod jak tyto oblasti identifikovat. Obecně nás zajímá, kde se nachází místa se zvýšeným výskytem námi studovaných jevů, ohniska tak nelze jednoznačně definovat. Nejjednodušším ukazatelem tak může být **modus (mode)**, pomocí kterého identifikujeme bod s největším výskytem incidentů. Takovou funkci nabízí CrimeStat, identifikuje bod s největším výskytem incidentů, který si můžeme snadno zobrazit např. v ArcGISu (obr. 3). V realitě se však zřídkakdy dva incidenty vyskytnou přesně na stejných souřadnicích, zejména v oblasti kriminality. Pro příklad si lze představit parkoviště, na kterém se stalo 20 krádeží vozidel, těžko však byla všechna ukradena přesně ze stejného místa na stejných souřadnicích. Avšak v systému policejní evidence bude mít všech 20 případů stejnou adresu, která se geokóduje do stejného bodu na parkovišti a všech 20 případů krádeže vozidel bude mít stejné souřadnice. Tato metoda tak využívá nepřesnosti v policejních záznamech (Smith, Bruce, 2008). A skutečně, testování této metody odhalilo pravdivost přechozího tvrzení, jelikož jako modus, tedy bod s největším výskytem incidentů byl identifikován bod na souřadnicích parkoviště u obchodního domu Walmart (obr.3). Druhou variantou je metoda tzv. **fuzzy mode**. Tato metoda staví na modu, jejím principem je, že kolem každého bodu utvoří radius, jehož velikost zvolí sám výzkumník, a zachytí všechny body uvnitř tohoto radiusu.

Výpočet snadno vykonáme v softwaru CrimeStat (ArcGIS tuto metodu nenabízí). Výsledek můžeme zobrazit na mapě jako body, které se liší velikostí podle toho, kolik bodů spadá do rádiusu daného bodu, názorně je to předvedeno na obr. 3. Pomocí této metody již můžeme zpozorovat několik oblastí se zvýšenou koncentrací incidentů, v tomto případě krádeží vozidel v Chicagu v roce 2018. Největší koncentrace bychom na základě těchto výsledků mohli hledat opět ve stejné oblasti, kde leží modus a nachází se tam 2 body s největší koncentrací, dále třeba v oblasti spojení Grand Ave. a State street. Tato metoda je vhodná pro detekování koncentrací incidentů, které mají mezi sebou krátké vzdálenosti, jako např. několik parkovišť kolem stadionu nebo parkoviště sdílené několika bytovými domy (Levine, 2017; Smith, Bruce, 2008). Její výhodou je snadný výpočet i její nenáročnost na pochopení. Nevýhoda této metody spočívá v tom, že body blízko sebe spadnou sami sobě do radiusu. V oblastech s vysokou koncentrací to pak vyústí ve více „hot spotů“, které se navzájem započítávají. Avšak i přes tento nedostatek funguje tato metoda dobře k identifikaci jednotlivých nejžhavějších ohnisek. (Smith, Bruce, 2008)

Obr. 3: Modus krádeží vozidel a výsledek metody „fuzzy mode“ s radiusem 500 čtverečních stop pro krádeže vozidel v Chicagu v roce 2018



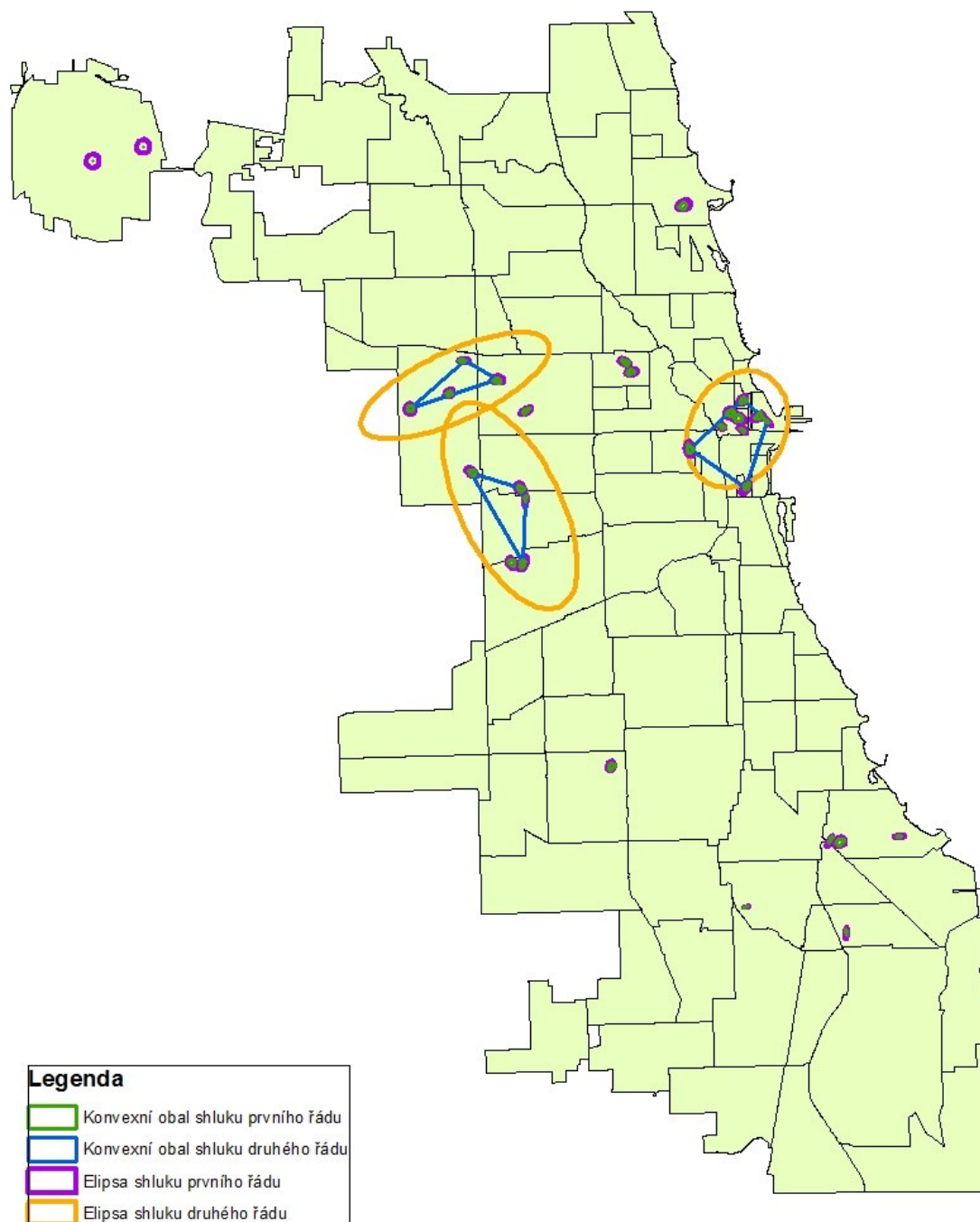
Zdroj dat: City of Chicago (2019); vlastní vypracování

Následující metoda analýzy koncentrací navazuje na metodu nejbližšího souseda, která odpovídá na otázku zda-li body v souboru mají blíže ke svým sousedům než by se očekávalo při náhodném rozmístění a dochází tedy ke shlukování (či nikoliv). Metoda **hierarchického prostorového shlukování nejbližších sousedů (nearest neighbor hierarchical spatial clustering - NNH)** nám může odpovědět na logickou navazující otázku, a sice kde v prostoru k takovým shlukům dochází. Incidenty jsou seskupeny do shluků nejbližších sousedů, které obsahují určitý minimální počet bodů, který zvolí výzkumník. Tyto shluky prvního řádu, jsou stejným principem seskupeny do shluků druhého řádu a tak tento proces pokračuje do dalších úrovní shlukování (odtud termín „hierarchické“ shlukování v názvu), dokud je shlukování možné. V každém kroku jsou seskupeny pouze body (shluky), mezi kterými je vzdálenost menší než očekávaná za náhodného rozmístění. (Eck a kol., 2005; Smith, Bruce, 2008)) Ve většině případů skončí shlukování na druhém nebo třetím řádu. Incidenty mohou patřit pouze do jednoho shluku, ale ne všechny incidenty patří do nějakého shluku (Levine, 2008). Tato metoda je součástí softwaru CrimeStat. Lze v něm nastavit několik parametrů, které mění průběh analýzy. Výzkumník by měl zkoušet měnit nastavení, aby docílil výsledku, který mu vyhovuje. Může měnit například minimální počet bodů k identifikaci shluku. To slouží k redukci počtu shluků, budou identifikovány pouze ty shluky, které obsahují daný minimální počet bodů (většinou nás moc nezajímají shluky dvou nebo tří bodů). Větší minimální počet bodů redukuje šanci, že budou identifikovány „falešné shluky“ (takové, vyskytující se i v náhodném rozmístění). Výstup této analýzy můžeme vytvořit ve formě elips (definovaných směrodatnou odchylkou či jejím násobkem, kterou zvolí výzkumník) nebo konvexních obalů. Konvexní obaly jsou přesnější, avšak elipsy představují elegantnější grafické zobrazení (Smith, Bruce, 2008).

Curtis a kol. (2003) využili metodu hierarchického prostorového shlukování nejbližších sousedů ke zkoumání prostorového rozmístění výskytu vzteklinou nakažených mývalů v Západní Virginii v letech 1999 a 2000. Tato metoda je využita k identifikování správného měřítka a lokace k provádění dalších úkonů. (Curtis a kol., 2003)

Obrázek číslo 4 zobrazuje grafické zpracování výstupu metody hierarchického prostorového shlukování nejbližších sousedů na skutečných datech o krádežích motorových vozidel ve městě Chicago v roce 2018 (srovnej s obrázkem číslo 3). Minimální počet bodů ve shluku byl nastaven na 7 a výsledné shluky jsou zobrazeny pomocí konvexních obalů a elips 1,5 směrodatné odchylky. Celkem bylo takto identifikováno 30 shluků prvního řádu a 3 shluky druhého řádu.

Obr. 4: Shluky krádeží motorových vozidel v Chicagu v roce 2018 utvořené metodou hierarchického shlukování nejbližších sousedů (NNH)

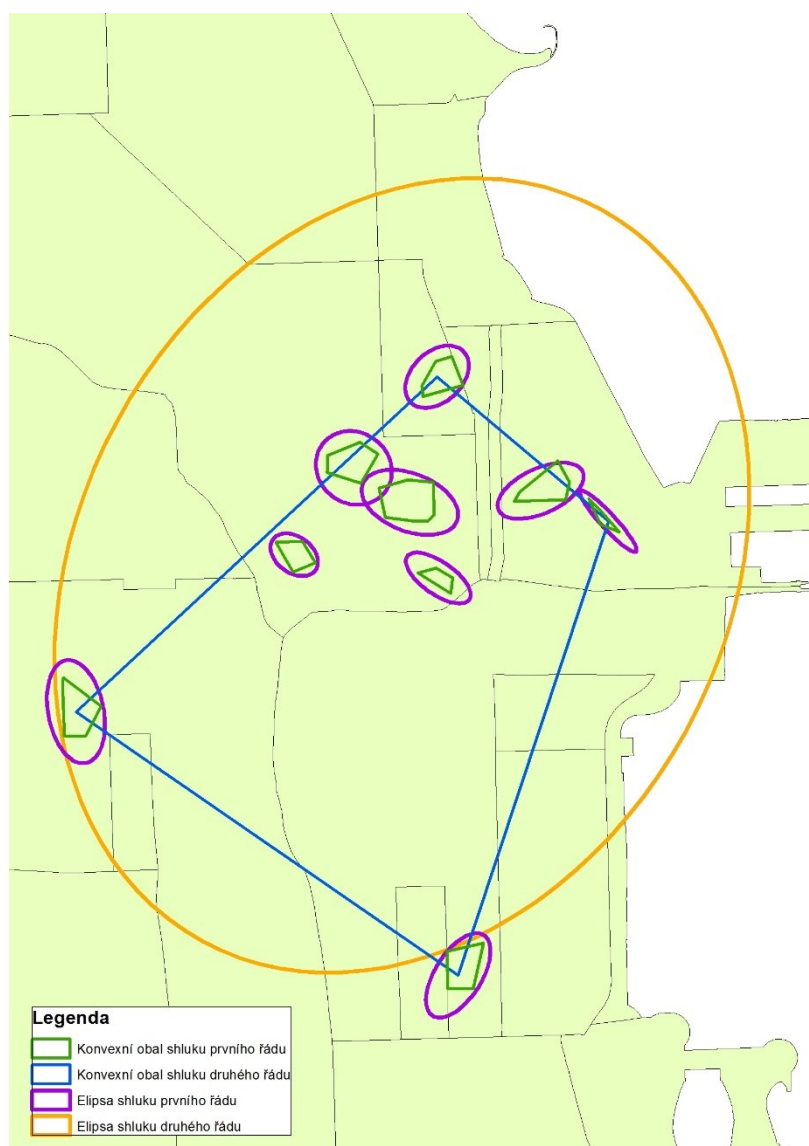


Zdroj dat: City of Chicago (2019); vlastní vypracování

Obrázek číslo 5 zobrazuje detailní pohled na jeden ze vzniklých shluků druhého řádu, konkrétně západní shluk, lokalita rohu ulic Grand Ave. a State street a širšího okolí, které jsme již identifikovali

pomocí metody fuzzy mode. Shluky prvního řádu nám poskytují detailnější pohled na lokalitu a umožní pohled na oblasti koncentrací kriminality v menším měřítku.

Obr. 5: Detail jednoho ze shluků druhého řádu vytvořeného pomocí metody hierarchického shlukování nejbližších sousedů (NNH)



Zdroj dat: City of Chicago (2019); vlastní vypracování

Mezi výhody této metody obecně patří její největší objektivita ze všech analýz koncentrací. Porovnání skutečných vzdáleností mezi sousedy s očekávaným náhodným rozmístěním, založeným na statistické pravděpodobnosti je neobjektivnější možností. Další výhodou je možnost identifikace ohnisek jak malých tak rozsáhlých, což je výhodné pro strategické účely. Shluky identifikované pomocí této metody jsou ty nejvíce koncentrované. (Smith, Bruce, 2008)

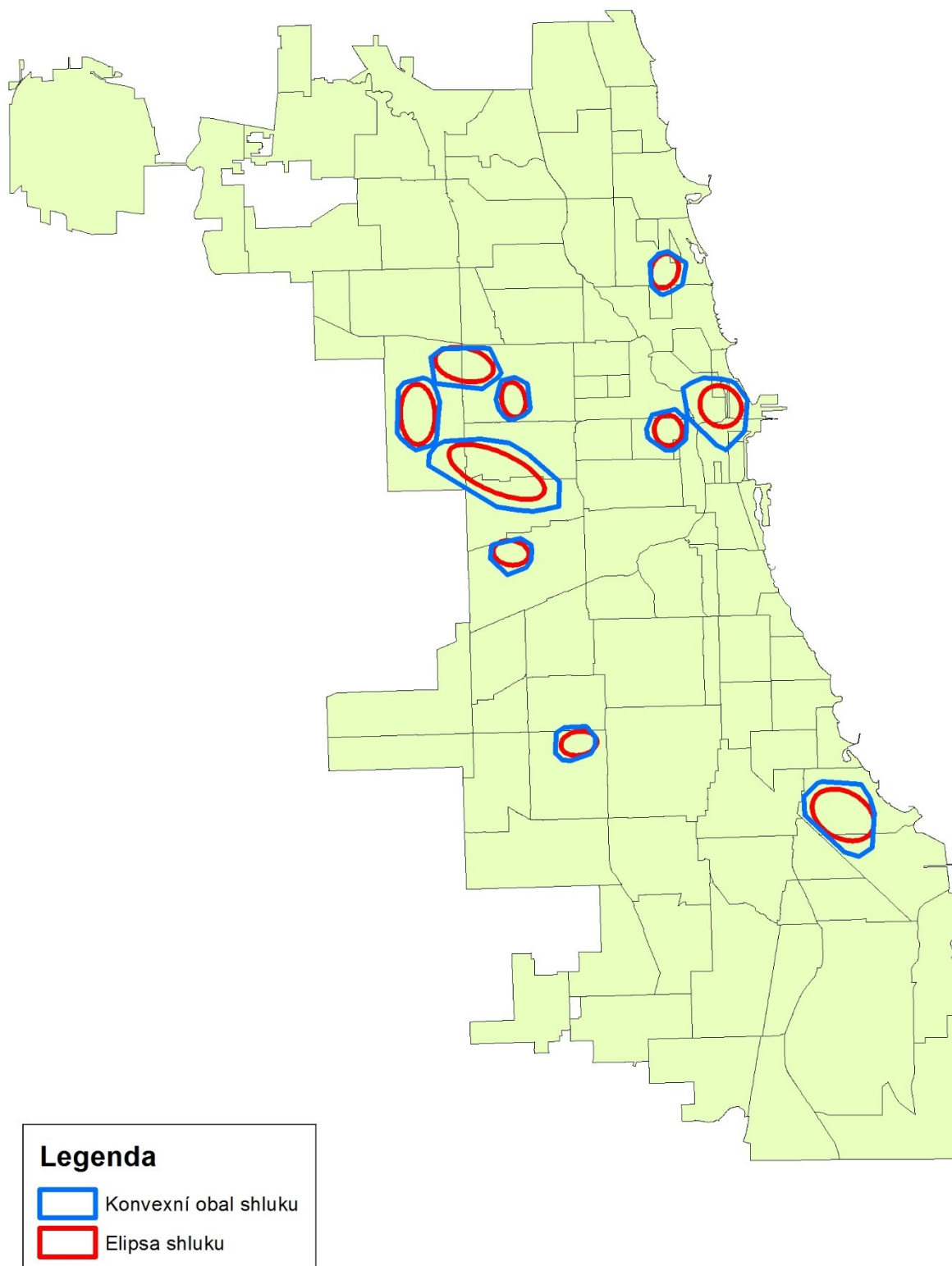
Podle Levine (2013), má tato metoda 4 hlavní využití. Za prvé, může identifikovat malé jednotky, kde jsou incidenty koncentrovány. Toho se dá využít zaměřením preventivních aktivit policií nebo komunitou. Jsou mikroprostředí generující zločinnost, které je tato metoda schopna zachytit. Za druhé, metoda může být aplikována na celé soubory, nemusí se aplikovat na menší jednotky. To usnadňuje využití. Za třetí naznačuje vazby mezi několika menšími shluky, které jsou viditelné skrz shluky vyšších řádů. Za čtvrté, každá z úrovní podněcuje jiné strategie (v oblasti potírání zločinu). Mezi slabiny této

metody patří to že, váha incidentů nehraje žádný vliv. Dále fakt, že je vnesena subjektivita do procesu tím, že si uživatel volí různé parametry a ovlivňuje tak vlastnosti shluků. To ovšem není devíza pouze této metody, ale téměř všech shlukovacích technik. (Levine, 2013)

Tuto metodu lze upravit doplněním sekundárního souboru, kde nějaká proměnná (např. počet obyvatel) slouží jako základna pro shlukování. Taková metoda nese název **hierarchického prostorového shlukování nejbližších sousedů přizpůsobena riziku (risk-adjusted nearest neighbor hierarchical clustering - rnnh)**. Obecně tato metoda funguje stejně jako základní hierarchické shlukování nejbližších sousedů, tato je však založena na poměrech (ne pouze na počtu) a využívá jmenovatele specifikovaného uživatelem. To je specifikum této metody, jiná metoda určena k analýze koncentrací tohoto není schopna. Data do jmenovatele jsou ovšem velmi náročná na získání. Většinou ani nepotřebujeme takové možnosti využít, avšak jsou případy, kdy ano. Shlukování je tedy v tomto případě prováděno s ohledem na tuto základní proměnnou, očekávaná vzdálenost náhodného rozmístění je kalkulována specificky pro jednotlivé buňky (které byly vstupem jako sekundární data se „základovou“ proměnnou ve formě bodů se souřadnicemi a libovolnou proměnnou) a ne jednotnou vzdálenost pro celou studovanou. Více o této metodě píše např. Levine (2008). Pro otestování této metody na krádežích motorových vozidel v Chicagu by byla potřeba sekundární soubor parkovacích míst, který bohužel pro tuto práci není k dispozici. Z čehož vyplývá hlavní nevýhoda této metody pro výzkum kriminality, a sice špatná dostupnost dat využitelných jako sekundární. (Smith, Bruce, 2008)

Jinou metodou analýzy koncentrací je **prostorová a časová analýza zločinu (spatial and temporal analysis of crime – STAC)**, která má název podle jednoho z prvních programů k identifikaci ohnisek (hot spots) vyvinutého při Illinois Criminal Justice Information Authority. Pomocí této metody také získáme elipsy nebo konvexní obaly míst s vyšší koncentrací jevů, funguje však jinak než metoda hierarchického shlukování. Od verze 2 je také implementován v softwaru CrimeStat. (Levine, 2013; Smith, Bruce, 2008). Tato metoda je „skenovací“, na síť pokrývající zkoumanou oblast jsou opakovaně pokládány kružnice a jsou počítány body uvnitř těchto kružnic. Zvolí se kružnice s největším počtem bodů, pokud bod spadá do dvou kružnic, body v těchto kružnicích se zkombinují. Tento proces se opakuje dokud nezbydou žádné překrývající se kružnice. Výsledné shluky jsou ohniska a žádné z nich se nepřekrývá. O přesném algoritmu více v Levine (2013). Výzkumník při práci v softwaru CrimeStat může, podobně jako v metodě NNH, ovlivnit několik parametrů. Velikost kružnic, minimální počet bodů ve shluku nebo velikost výsledných elips. Je na výzkumníkovi zvolit parametry, které vyhovují jeho požadavkům a cílům, avšak je rozumné volit parametry tak, aby nebylo výsledných shluků příliš, či aby nebyly příliš velké nebo malé. Pro ukázkovou analýzu krádeží motorových vozidel v Chicagu jsem zvolil minimální počet bodů ve shluku 7, radius kružnic 0,5 mil, výsledné elipsy jsou o velikosti jedné směrodatné odchylky. Na obrázku 5 jsou k vidění výsledné shluky ve formě elips a konvexních obalů, které je možné porovnat s výsledky předcházejících metod. Metoda STAC při těchto parametrech identifikovala 10 shluků, které obsahují (jejich konvexní obaly) dohromady celkem 1026 případů krádeží motorových vozidel (z 5253 celkem – 19,5%).

Obr. 6: Shluky případů krádeží motorových vozidel v Chicagu v roce 2018 identifikované pomocí metody STAC



Zdroj dat: City of Chicago (2019); vlastní vypracování

Výhodou metody STAC je její rychlost a intuitivnost, také to, že má výzkumník velkou kontrolu nad parametry. Což může být i nevýhoda, protože při nesprávné aplikaci nemusí dávat výsledky smysl. (Smith, Bruce, 2008)

Zmíním ještě další metodu k analýze koncentrací, kterou je metoda zvaná **K-means clustering**. Při ní se vytváří určitý definovaný počet (K) elips. Tato procedura rozdělí data do skupin definovaných uživatelem. To je výhoda a zároveň nevýhoda této metody. K-means clustering nabízí velkou dávku uživatelské kontroly tím, že umožňuje nastavit počet a velikost požadovaných shluků. V případě dat za kriminalitu a využití této metody, K-means se hodí více k velkým geografickým jednotkám a je vhodná spíše k identifikaci oblastí hlídkování než vyloženě ohnisek kriminality (Levine, 2013). Metoda K-means clustering může být vhodná při určitých strategických projektech, ve kterých je počet oblastí na kolik je třeba se zaměřit, dopředu znám. (Smith, Bruce, 2018)

Ve všech těchto metodách je stále množství subjektivity, ačkoliv ne tolik jako v identifikování shluků pouhým okem. Parametry v těchto metodách lze volit libovolně tak, že mohou vzniknout shluky malé i velké velikosti, může jich být malé nebo velké množství. Subjektivita nemusí být brát jako negativní jev, jde o to tyto metody dostatečně pochopit, abychom mohli využít data ke kýženému cíli. K takevému pochopení vede dle Smith a Bruce (2008) zejména experimentování s metodami a praxe. (Smith, Bruce, 2008)

V oblasti výzkumu kriminality se znalost ohnisek výskytu trestných činů využívá ve velké míře k cílení opatření k prevenci. Šimon a Jíhová (2018a) uvádí, že zaměřit kriminologicko-geografický výzkum pouze na místa koncentrace (hotspoty) aniž bychom zohlednili společenskou závažnost a škodlivost konkrétních typů trestných činů je v zahraniční literatuře kritizováno a označováno za suboptimální. Může totiž v praxi docházet k případům, kdy je policejní přítomnost alokována sice do ohnisek kriminality, ale ne tolik společensky škodlivé a paradoxně tak může dojít k nárůstu společenské škodlivosti kriminality. To však neznamená, že by takové strategie ke snížení počtu některých kriminálních činů byly špatné nebo neúčinné, avšak kýžený pokles počtu některých kriminálních činů nemusí automaticky znamenat pokles celkové společenské škodlivosti kriminality. (Šimon, Jíhová, 2018a)

V této subkapitole bylo představeno několik metod, jak mohou výzkumníci analyzovat bodová data, specificky tedy data za kriminalitu, a pomocí nich vizualizovat a blíže pochopit ohniska kriminality. Velice důležité jsou testy shlukování. Výzkumník by mohl ztratit mnoho cenného času a zdrojů, pokud by se snažil vytvořit mapy ohnisek kriminality a test jako metoda nejbližšího souseda by dokázal rychle odhalit, že k žádnému shlukování v souboru vůbec nedochází. (Eck a kol., 2005)

5.3. Prostorové modelování, interpolace

Interpolace je populární metoda založená na agregaci bodů uvnitř určitého radiusu a vytváří vyhlazený, kontinuální povrch, který reprezentuje množství nebo hustotu jevů napříč územím (Eck a kol., 2005). Metody souvislého vyhlazování povrchu neboli interpolační metody jsou velmi oblíbené při mapování a vizualizování distribuce nejen kriminality (hledání ohnisek kriminality), ale i při různých analýzách zaměřených na životní prostředí, např. modelace odtokových poměrů podle zjištěných nadmořských výšek (Křikavová 2009). Ve většina běžných interpolačních technik je používána intenzita určitého jevu, populace nebo hodnota „z“. Pro příklad uvádí Eck a kol. (2005) využití interpolačních technik k vytváření povrchů reprezentující distribuce deště, kde neznáme hodnoty na území mezi srážkoměry jsou odhadnuty. (Eck a kol, 2005) Obecně jsou interpolační techniky velmi cenné pro odhadování hodnot na velkých plochách, ze kterých známe pouze vzorek. Například je nemožné znát přesnou teplotu na každém místě na Zemi, ale izotermální mapy ukazují teploty pro celé zkoumané území. K tomu jsou potřeba vzorky z několika desítek až několika tisíců míst (záleží na měřítku) a zbytek míst je podle teplot z míst vzorků odhadnut. (Smith, Bruce, 2008)

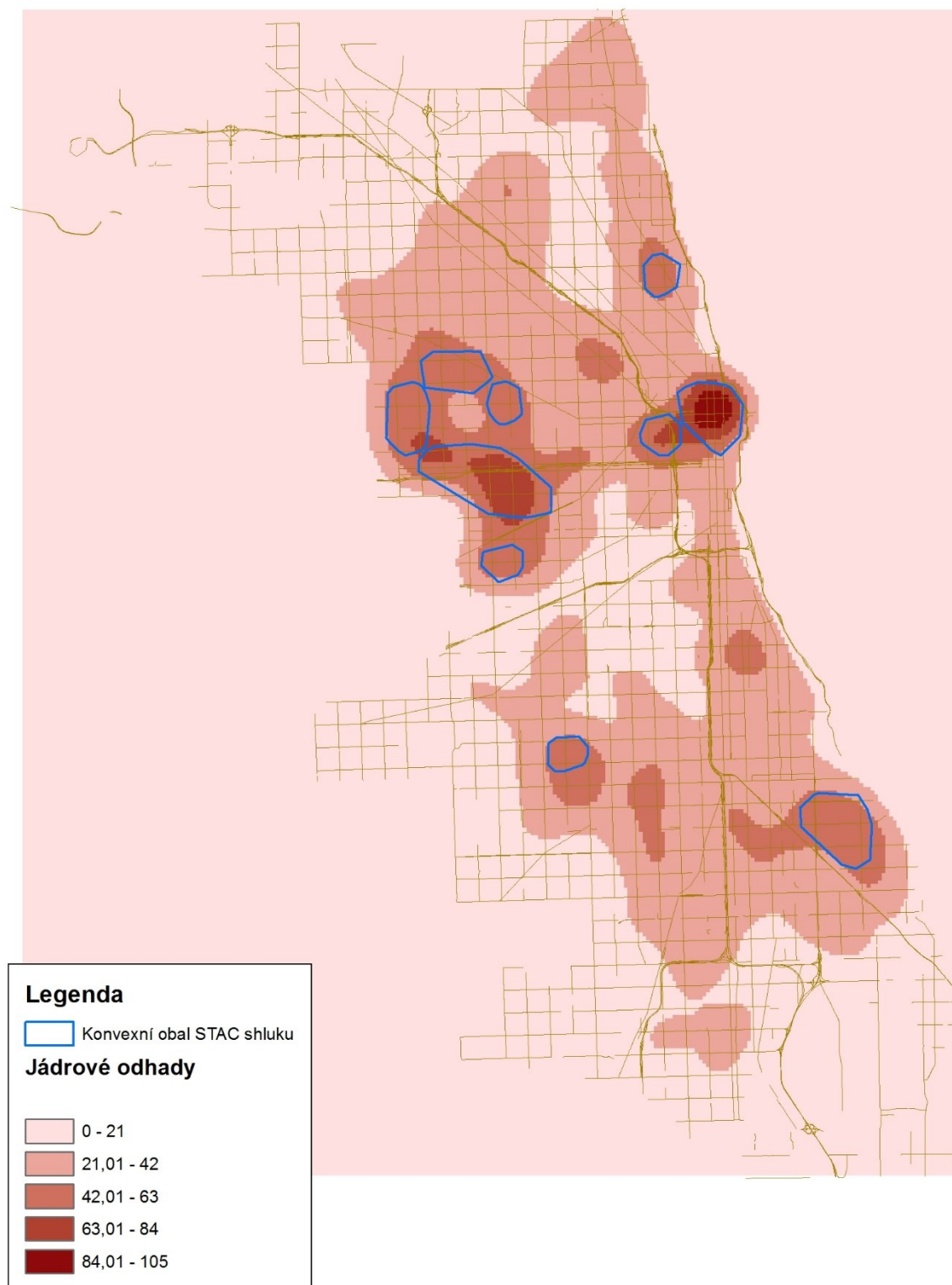
Situace s daty za kriminalitu je však poněkud rozdílná. Taková data se vzorky s různou intenzitou totiž v této oblasti neexistují a ani by podobné techniky odhadování počtu zločinů, které se odehrály mezi existujícími místy činu, nedávalo smysl v oblasti kriminality využívat, žádné zločiny v prostoru mezi zaznamenanými zločiny zaznamenány nebyly, tudíž snaha je odhadovat nedává smysl. Místo toho se výzkumník snaží o vytvoření povrchu, který reprezentuje distribuci, a měl by sloužit jako vizualizace k pochopení rozložení zločinu v prostoru. Mapy vytvořené pomocí interpolace by ve výzkumu kriminality měly nejlépe znázorňovat riziko. Ne každý bod na Zemi má zločin, na každém bodě na Zemi je však šance ke zločinu. Jádrové odhady hustoty odhadují tyto šance na základě míst, kde se zločin odehrával v minulost. (Eck a kol., 2005; Smith, Bruce, 2008)

Metoda jádrových odhadů je tedy ta, která zde bude více examinována a testována. Ivan a Horák (2016) zdůrazňují, že je před použitím metody jádrových odhadů nutné si ujasnit, co je cílem analýzy a pro jaký účel je analýza tvořena. Metoda jádrového odhadu není vhodná pro zobrazení rozsáhlých území, ale je vhodnější pro mapy větších měřítek, které zobrazují obce či jejich části (Ivan, Horák, 2016). Velikost radiusu determinuje, jak generalizovaná se bude mapa jevit. Menší radius ukáže více lokální výkyvy, širší radius ukáže obecnější trendy v datech. Jsou sečteny počty bodů, které spadají do oblasti radiusu pro každou buňku a to je poté vyděleno hustotou každé buňky sítě. Jádrové odhady však dávají větší váhu bodům blíže ke středu radiusu než těm na okraji. Výsledkem toho je hladší distribuce hustot. (Eck a kol., 2005)

Použitá funkce jádrového vyhlazování se zobrazením pouze určitého procentuálního podílu nejvyšších intenzit se ukázala jako velmi výhodná u výzkumů Ivana a Horáka (2012). Uvádí však, že tato metoda není nejvhodnější pro trestné činy či přestupky s nízkým počtem případů rozmístěných po městě, jelikož v daném dosahu se pak často nevyskytuje žádný další případ a nedochází k součtům intenzit více případů. Ve výsledku je tak mapa pouze skružnicemi kolem jednotlivých případů. (Ivan, Horák, 2012)

Mapy jádrových odhadů představují vhodnou metodu pro rychlé nalezení hlavních problémových lokalit. Je však nezbytné každou takto nalezenou lokalitu podrobit další analýze a snažit se odlišit místa, kde je vyšší koncentrace způsobená vyjímečnou anomální událostí od těch způsobených kriminogenním faktorem. (Ivan, Horák, 2016)

Obr. 7: Výsledek metody jádrových odhadů na krádeže motorových vozidel v Chicagu v roce 2018 v porovnání s ohnisky identifikovanými pomocí metody STAC



Zdroj dat: City of Chicago (2019); vlastní vypracování

Na obrázku číslo 7 je znázorněna aplikace metody jádrových odhadů na případy krádeží motorových vozidel v americkém městě Chicago v roce 2018. Jádrové odhady byly zpracovány pomocí softwaru ArcGIS. Můžeme tak identifikovat hlavní problémové lokality. Většina takových oblastí nalezená pomocí této metody se shoduje s výsledky analýzy STAC a zobrazuje stejné lokality ležící zejména ve středo-západě a středo-východě města. Další problémové lokality jsou identifikovatelné na jihu města.

Kapitola 6

Závěr

Prostorová analýza, jako specifická podskupina analýz kvantitativních, která se vyznačuje potřebou prostorových dat, tedy dat, která se vztahují k určitému místu na Zemi, je ve společnosti přítomná již dlouhou dobu. Své kořeny má už v dávných dobách, za náznaky prostorových analýz se považují například rané pokusy zmapovat svět nebo vesmír. Její význam v geografii se měnil s převládajícími paradigmaty a jejich postoji ke kvantitativním metodám výzkumu. Stále více si však uvědomujeme, že většina dat je prostorových a je třeba tímto faktem umět pracovat a využít jej. Prostorová data jsou rozdílná od neprostorových dat především v tom, že na sobě nejsou nezávislá a to znemožňuje použití většiny klasických statistických metod. Prostorová bodová data, data o přesném umístění jevů, objektů nebo incidentů, jsou data, která lze najít zejména v oblastech kriminologie, geologie, biologie, epidemiologie či ekologie. U bodových dat nás často zajímá vzorec jejich prostorového rozmístění, zda-li tvoří v prostoru nějaké shluky či nikoliv. Nalezené vzorce prostorového rozmístění se pak dají využít k hledání procesů, které za rozmístěním stojí. Toho je využíváno mimo jiné v kriminologii s cílem zefektivnit preventivní činnost a cílit ji na nejvíce zločinem postižená místa.

Na důležitosti začaly prostorové analýzy nabývat především s rozvojem výpočetní techniky a geografických informačních systémů. V dnešní době existuje již množství softwarů, jež se zaměřují na prostorovou analýzu nebo dokonce na prostorovou analýzu bodových dat. To je případ softwaru CrimeStat, který je v provozu od roku 1999 a nyní je ve své čtvrté verzi. Přínos tohoto freeware softwaru je znatelný, nabízí množství metod a možností, které v tuzemsku více rozšířený software ArcGIS nenabízí. Mezi metody, jež CrimeStat oproti ArcGISu nabízí, patří některé metody k identifikaci ohnisek událostí jako metoda hierarchického shlukování nejbližších sousedů, STAC analýza či Fuzzy mode nebo některé pokročilé metody určené například ke sledování chování seriových pachatelů. Je jednoduchý na ovládání a spolu s uživatelským manuálem tvoří dobrý nástroj k analýze bodových dat. Jeho součástí však není možnost vizualizace, v tom je tak třeba spoléhat na jiný software, například zmíněný ArcGIS.

Široké využití nalézá prostorová analýza bodových dat v oblasti kriminologie. Kriminalita je komplexní fenomén, který se dotýká mnoha oborů včetně geografie, která se zabývá jejími prostorovými aspekty. Ty je třeba pochopit a snažit se objasnit podmíněnosti výskytu kriminality, to může vést k redukci kriminality ve společnosti. Mapování zločinnosti a hledání jejích ohnisek patří k nejdůležitějším činnostem v prostorové analýze kriminality. K tomu slouží velké množství metod. V Česku se výzkumu kriminality jejím prostorovým aspektům věnuje Jíchová (např. 2013, 2014, 2017). V Ostravě se

výzkumem prostorových aspektů zločinnosti zabývají Ivan a Horák (např. 2012). Více se zatím však v Česku provádí kriminologické studie zaměřené na percepci prostředí, výklad práva a podobně, využití přesně lokalizovaných dat je značně omezené, což je dáno velkou měrou i špatnou dostupností dat. Ačkoliv potenciál využití podobných dat je velký nejen pro kriminologii, ale i pro jiné obory, v Česku je jejich využití na počáteční fázi. Diskuse o větší dostupnosti lokalizovaných dat za kriminalitu existuje (například Šimon, Jíchová, 2018a), situaci je tak třeba do budoucna sledovat. Ve světě jsou data o přesně lokalizovaných činech více dostupná, příkladem může být například USA (Chicago, Los Angeles) nebo Velká Británie.

Jak již bylo zmíněno, existuje množství metod pro analyzování bodových dat. Mimo metody využívající agregaci bodů do polygonů, které se dále využívají k posuzování zda dochází ke shlukování či nikoliv, jsou tu metody, které analyzují samotné body a jejich rozmístění. Představené metody jsem rozdělil do tří skupin na základní popisné statistiky, analýzy koncentrací a prostorové modelování a interpolace. V první skupině se nachází metody k popisu rozmístění jako geografický střed, medián, elipsy směrodatných odchylek a jiné centrografické metody. Této skupině dominuje metoda nejbližšího souseda. Ta slouží k odhalování zda dochází v souboru k prostorovému shlukování (nebo naopak disperzi) či nikoliv. Neříká nám nic o tom, kde v prostoru ke shlukům dochází, avšak je důležitým krokem před samotnou identifikací umístění shluků. Pomocí metod ze skupiny analýz koncentrací, lze právě nalézt, kde v prostoru ke zvýšené koncentraci dochází. Analýza STAC je z těchto metod nejjednodušší na pochopení i na provedení, metoda hierarchického prostorového shlukování nejbližších sousedů (NNH) je z těchto analýz nejméně objektivní, jelikož pracuje s náhodným rozmístěním bodů, které porovnává se skutečným. Míra subjektivity však není v této skupině nízká, jelikož u všech těchto metod můžeme měnit parametry a tím tak upravovat výsledek. Při dobrém pochopení metod to však nemusíme považovat za vysloveně negativní jev, navíc míra subjektivity je pořád nižší než identifikování ohnisek pouhým zrakem. Třetí skupinu reprezentuje metoda jádrových odhadů. Výstupem této metody jsou atraktivní mapy a tato metoda je vhodná pro rychlé nalezení problémových lokalit.

S pokračujícím rozvojem na poli prostorových dat vidím v prostorových analýzách velkou budoucnost. Bodová data tvoří velkou část těchto dat a jejich analyzování nám může přinést velmi cenné informace. V oblasti kriminality však snaha nesmí zůstat u znalosti prostorového rozmístění trestných činů a přestupků, je třeba tuto znalost přetavit v efektivní preventivní opatření, aby byla redukce zločinnosti účinná. Nicméně prostorová analýza je užitečný nástroj ke zvýšení povědomí o prostoru a o procesech, které se v prostoru generují a jsou jím ovlivněny.

Seznam literatury a použitých zdrojů

ANDRESEN, M. (2006): A spatial analysis of crime in Vancouver, British Columbia: a synthesis of social disorganization and routine activity theory. *Canadian Geographer*, 50, 4, 487–502.

ANSELIN, L., COHEN, J., COOK, D., GORR, W., TITA, G. (2000): Spatial Analyses of Crime, in: Duffe, D. (ed): *Measurement and Analysis of Crime and Justice*, s. 213–262.

ANSELIN, L., SYABRI I., KHO Y. (2005): GeoDa: An Introduction to Spatial Data Analysis. *Geographical Analysis*, 38, 1, 5-22.

BADDELEY, A., RUBAK, E., TURNER, R. (2017): Spatial Point Patterns: Methodology and Applications with R. *Mathematical Geosciences*, 49, 6, 815–817.

BATES, S. (1987): *Spatial and temporal analysis of crime*. Illinois Criminal Justice Information Authority.

ECK, J. E., CHAINEY, S., CAMERON, J. G., LEITNER, M., WILSON, R. E. (2005): *Mapping crime: Understanding Hot Spots*. National Institute of Justice, Washington DC.

FOTHERINGHAM, S., BRUNSDON, CH., CHARLTON, M. (2000): *Quantitative Geography: Perspectives on Spatial Data Analysis*.

FOTHERINGHAM, S., ROGERSON, P. A., EDS. (2009): *The SAGE Handbook of Spatial Analysis*. SAGE, London, 511 str.

GRAHAM, A.J., ATKINSON, P.M., DANSONC, F.M. (2004): Spatial analysis for epidemiology. *Acta Tropica*, 91, 3, 219–225.

GUERRY, A.-M. (1833): *Essai sur la Statistique Morale de la France*.

HUEBNER, B., BYNUM, T. (2016): *The Handbook of Measurement Issues in Criminology and Criminal Justice*.

HERAUX, C. (2007): *Software Review: Spatial Data Analysis of Crime: A Review of CrimeStat III*.

IVAN, I., HORÁK, J. (2012): Analýza kriminality v Ostravě, in: *Symposium GIS Ostrava 2012: Současné výzvy geoinformatiky*, red. Jan Růžička, Ostrava 2012, 1–7.

IVAN, I., HORÁK, J., NÁVRATOVÁ M., INSPEKTOR, T. (2013): *Prostorová analýza kriminality v Ostravě 2009-2011, Konference GIS Ostrava 2013*.

IVAN, I., HORÁK, J. (2016): Metodika identifikace anomálních lokalit kriminality pomocí jádrových odhadů; Konference GIS Ostrava 2016.

IVAN, I., HORÁK, J. (2018): Registered crime hot spots in Ostrava – location, development and influencing factors. In: Geografický časopis, 70, 1, 3–21.

JÍCHOVÁ, J. (2013): Kriminalita a její percepce v městském prostředí. Disertační práce. Katedra sociální geografie a regionálního rozvoje PŘF UK, Praha, 69 s.

JÍCHOVÁ, J. (2014): Výzkum zločinu v Česku z pohledu geografie. Historická geografie, 40(1), 73–93.

JÍCHOVÁ, J. (2017): Kriminalita v Praze: měřítková diferenciacie a možnosti měření. In: Ouředníček, M., Jíchová, J. (eds.): Sociální prostředí Prahy: město na prahu 21. století. Academia, Praha, s. 63-91.

JÍCHOVÁ, J., NETRDOVÁ, P. (2017): Mapování kriminality jako téma i praktická činnost ve výuce. Geografické rozhledy, 27, 2, 20–23.

KOSTKA, J. (2016): Hot spot analýza krádeží automobilů v Liberci. Diplomová práce. Katedra sociální geografie a regionálního rozvoje PŘF UK, Praha.

LEUNG A review of spatio-temporal pattern analysis approaches on crime analysis

LEVINE, N. (2006): Crime mapping and the CrimeStat program. Geographical Analysis. 38, 1, 41-55.

LEVINE, N. (2008): CrimeStat: a spatial statistical program for the analysis of crime incidents. Shekhar, S. and Xiong, H. (eds): Encyclopedia of Geographic Information Science. Springer. 187-193.

LEVINE, N. (2013): CrimeStat IV: A Spatial Statistics Program for the Analysis of Crime Incident Locations, Version 4.0.

LISITSIN, V. (2015): Spatial data analysis of mineral deposit point patterns: Applications to exploration targeting. Ore Geology Reviews, 71.

MURRAY, A. (2001): Exploratory Spatial Data Analysis Techniques for Examining Urban Crime.

PAVLÍNEK, P. (1993): Anglo-americká geografie ve dvacátém století In: SÝKORA, L. (ed.): Teoretické přístupy a vybrané problémy v současné geografii.

QUETELET, A. (1833): Recherches sur le Penchant au Crime aux Différens Ages. Hayes, Brusel.

QUETELET, A. (1842): A Treatise on Man.

SMITH, S. C., BRUCE, C. W. (2008): CrimeStat III User Workbook. The National Institute of Justice, Washington, DC.

ŠIMON, M., JÍCHOVÁ, J. (2018a): Jsou místa koncentrace kriminality společensky nejškodlivější? Nový analytický přístup pro prevenci kriminality a policejní praxi. In: Lubelcová, G. (ed.): Paralely a divergencie. Zborník z medzinárodnej vedeckej konferencie. Univerzita Komenského v Bratislave, s. 85-95.

ŠIMON, M., JÍCHOVÁ, J. (2018): Zveřejnění dat o kriminalitě na úrovni ulic - návrh řešení pro města v Česku. In: Ščerba, F. (ed.): Kriminologické dny 2018. Sborník z VI. ročníku mezinárodní konference Kriminologické dny. Iuridicum Olomoucense, Olomouc, s. 186-204.

VERONESI, F. (2015): Introductory Point Pattern Analysis of Open Crime Data in London. R tutorial for Spatial Statistics.

WORTLEY, R., MAZEROLLE, L. (2008): Environmental criminology and crime analysis: Situating the Theory, Analytic Approach and Application. *Crime Prevention and Safety*, 11, 2.

ZAPLETAL, J. (2003): Zločin jako volba (K neoklasickým kriminologickým teoriím). *Kriminalistika*, 36, 3, 161–177.

ZEMAN, J. (2011): Faktory ovlivňující prostorové rozmístění mravenišť mravenců druhu *Formica polyctena*. Diplomová práce. Katedra zoologie PŘF UK, Praha.

ZEMAN, P. (ed.) (2017): Research on Crime and Criminal Justice in the Czech Republic (selected results of research activities of IKSP in the years 2012–2015). ISBN 978-80-7338-165-3.

DIXON, K. R., CHAPMAN, J. A. (1980): Harmonic Mean Measure of Animal Activity Areas. *Ecology*, 61, 5, 1040–1044.

Datové a internetové zdroje

CITY OF CHICAGO: About the Open Data Portal [online]. [cit. 9. 4. 2019]. Dostupné z: <https://data.cityofchicago.org/>

CITY OF LOS ANGELES: Los Angeles Open Data. [online]. [cit. 9. 4. 2019]. Dostupné z: <https://data.lacity.org/>

DATA – CITY OF CHICAGO: Crimes - 2001 to present [online]. [cit. 7. 4. 2019]. Dostupné z: <https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2>

MAPA KRIMINALITY: Projekt Otevřené společnosti, o.p.s. [online]. [cit. 21. 6. 2019]. Dostupné z: www.mapakriminality.cz