

Geokódovanie a jeho aplikácia na obce Slovenska – prípadová štúdia regiónu Liptov

Abstract:

The paper deals with the issue of geocoding and its use based on specific geocoding mechanisms implemented in geoinformation tools, applied in a case study of geocoded spatial units (communes) in Slovakia in a selected region (Liptov). Except the detailed procedure that addresses geocoding using Google Geocoding API, it also provides the analysis of the results of obtained pointed locations through the procedure of defined central reference points (centroids) based on the GIS MapInfo Professional. Pointed localised communes, obtained in geocoding process, are subsequently compared with the official state registry of definitional points of Slovak communes. Different quality of the results is compared using the distance analyses (buffering) applied to selected elements of the road network (highways and roads of first and second class) with subsequent evaluation of the distribution of population in the selected region according these distance zones.

Keywords:

geocoding, definition point, centroid, Slovak communes, address, Google Geocoding API

Úvod

Geokódovanie, ako nástroj prevodu z nepriameho stanovenia polohy geografických objektov a javov (prostredníctvom adresy, geografického názvu alebo identifikátora) na priamu vo forme geografických alebo iných priestorových súradníc, sa čoraz širšie využíva v geografickom výskume a praxi v súvislosti s rozvojom a dostupnosťou geokódovacích služieb, systémov a registrov. Tvorba účelových geokódovacích systémov a databáz, akými sú napr. registre adries a ich definičné body, nadobúda čoraz väčší význam pre prax úmerne s rastom dostupnosti nástrojov internetového mapovania zacielených na ich využitie, alebo tvorbu.

Cieľom príspevku je uviesť do problematiky geokódovania a jeho použitia na báze konkrétnych geokódovacích mechanizmov v prostredí geoinformačných systémov a ich služieb na modelovom príklade geokódovania základných územných jednotiek Slovenska vo vybranom regióne Liptova. Zameriavame sa na detailný opis postupu hromadného, resp. dávkového adresného geokódovania s využitím služby Google Geocoding API. Správnosť a presnosť geokódovania podstatne ovplyvňuje kvalitu výskumu, ktorý využíva výsledné lokalizácie modelovaných geografických objektov, a to najmä v kontexte využívania už vytvorených štandardizovaných štatistických a adresných priestorových systémov, ich registrov a rôznych geoinformačných techník (Bracken a Martin 2001, Ratcliffe 2001,). Príspevok preto poskytuje analýzu výsledkov bodových lokalizácií získaných ďalším spôsobom, a to s postupom získavania referenčných centrálnych bodov (centroidov) z areálovo vymedzených územní obcí na báze funkcie konkrétneho programu GIS (MapInfo Professional). Ako tretí zdroj pre analýzy sme použili oficiálny register definičných bodov obcí Slovenska. Rôznu kvalitu výsledkov komparujeme pomocou vzdialenostnej analýzy (zónovania – buffering) s využitím komunikácií a distribúcie obyvateľstva v regióne Liptova.

1. Základné pojmy

Geokódovanie

V úvode uvádzame definície základných používaných pojmov úzko spojených s medzinárodnými normami pre oblasť geografických informácií, ktoré zastrešuje medzinárodná normalizačná komisia ISO/TC 211 Geographic Information/Geomatics.

Norma ISO 19133 (podľa Geografická informácia, 2003, TeSlo 2012) definuje *geokódovanie* (angl. *geocoding*) ako prevod jednej formy umiestnenia (miesta) do inej, kde v poznámke rozširuje definíciu o tento text: "Geokódovanie sa zvyčajne vzťahuje na prevod adres (alebo iných nepriamych identifikátorov miesta) na priamu polohu (priestorové súradnice). Mnohí poskytovatelia služieb vkladajú do svojho geokódovača rozhranie na spätné geokódovanie, čím rozširujú definíciu služby na všeobecný prevodník miesta (polohy). Pretože služby sledovania (trasy, cesty) využívajú vnútorné kódovanie polohy (vyhľadávanie trasy), ktoré zvyčajne nie je dostupné pre iných (iné služby), geokódovač tvorí nedeliteľnú súčasť takej služby.". Norma ISO 19133 tiež definuje *geokódovač* (angl. *geocoder*) ako službu vykonávajúcu geokódovanie a *geokódovacia služba* (angl. *geocoder service*) ako sieťovo dostupnú službu, ktorá transformuje popis polohy, ako napr. názov miesta, adresa, smerové číslo, do normalizovaného popisu polohy bodového geometrického objektu. Pojem geokódovanie sa často zamieňa s pojmom *polohové priradenie* (angl. *georeferencing*), ktorý podľa normy ISO 19130 označuje „proces určovania vzťahu medzi polohou údajov v snímkovom súradnicovom systéme a ich geografickou polohou alebo polohou na mape.“ (TeSlo 2012).

Definičný bod

Problematika určenia *definičného bodu* modelovaného (geo)objektu je v geodézii a kartografii rozpracovaná už dávnejšie. Definičný, resp. reprezentačný bod objektu (angl. *centroid, label point, point of an object*) definuje technická prax ako „bod objektu (vzťahnutý najčastejšie na jeho ťažisko), ktorého poloha (súradnice) určujú globálnu lokalizáciu objektu v území“. (Terminologický slovník geodézie, kartografie a katastra 1998, s. 28). V súlade s tým sa *definičný bod nehnuteľnosti* definuje ako bod (najčastejšie v ťažisku parcely), ktorého súradnice sú základným globálnym lokalizačným prvkom nehnuteľnosti v katastri nehnuteľností.

V súčasnosti poskytuje Geodetický a kartografický ústav v Bratislave dátovú sadu oficiálnych definičných bodov základných územných jednotiek – obcí a častí obcí za celé územie Slovenska vo vektorovom digitálnom tvare (formát súboru SHP) v národnom súradnicovom systéme Jednotnej trigonometrickej siete katastrálnej (S-JTSK) spolu so základnými charakteristikami, ako sú napr. názov a kód, počet obyvateľov, kategória sídla a časti obce, výmera obce (Cenník, 2011). Ide o produkt – priestorový register, ktorý bol vytvorený v súlade požiadaviek Európskej únie (Nariadenie 2008), bez uvedenia metodiky spôsobu určenia konkrétnej polohy bodu v hraniciach zastavaného územia základnej územnej jednotky.

Centroid

Extrakcia lokalizačných dát vo forme reprezentatívnych bodov centier, stredov alebo ťažísk najmä graficky zložitých objektov, akým je napr. areál obce, alebo inej územnej jednotky, je v prostredí geoinformačných programov založená na rôznych matematických, resp. geometrických postupoch (Richeson 2005). Takto získané body, budeme ich označovať ako *centroidy*, môžu byť vo veľkom rozpore s obligátne stanovenými lokalizáciami na základe účelových metodík, ako je napr. pravidelná revízia definičných bodov základných sídelných jednotiek, kde je metodika jednoznačne určená právnym predpisom a viaže sa jednoznačne na osídlenie, resp. na obytné časti zastavanej plochy územnej jednotky (Metodika revízie základných sídelných jednotiek 1999). Geometricky určené body nevyhovujú tiež geograficky orientovaným modelačným technikám a analýzám (Madajová, 2010).

Tvoba centroidov predstavuje v prostredí geografických informačných systémov najmä manipulačný nástroj spracovania priestorových dát, resp. účelovú „manipuláciu s objektmi

a ich vlastnosťami, ktorý sa však dá aplikovať aj ako analytický nástroj skúmania základných geografických princípov“ (Tuček, 1998, s. 235-4).

Postupy extrakcie centroidov nie sú obvykle prezentované v konkrétnych geoinformačných programoch a ostávajú pre používateľa čiernou skrínkou. Taktiež implementačné štandardy medzinárodnej organizácie Open Geospatial Consortium (<http://www.opengeospatial.org/>), snažiacej sa o otvorenú a dobrovoľnú zhodu technických štandardov pre geopriestorové dáta a služby, nevyžadujú špecifikáciu použitých algoritmov, dôraz je kladený na ich funkčnosť, čo vytvára rôzne výsledky v závislosti od ich implementácie (Ryden, 2005, s. 10).

Adresa

Pre oblasť zjednocovania postupov tvorby adresných systémov, ktoré sú ťažiskové pre použitie geokódovacích služieb, sú významné aktivity súvisiace s budovaním európskej infraštruktúry priestorových informácií prostredníctvom projektu INSPIRE (Infrastructure for Spatial Information) v súlade so Smernicou Európskeho parlamentu a Rady 2007/2/ES a príslušnou dátovou špecifikáciou (Data Specification on Addresses – Guidelines D2.8.I.5 INSPIRE, 2011)

Dátová špecifikácia pre adresy je potrebná aj na uľahčenie interoperability, čiže prepojenia adresných informácií medzi členskými štátmi Európskej únie (EÚ), najmä z dôvodu rozdielnych koncepcií a všeobecných vlastností národných, miestnych alebo lokálnych adresných systémov. Existujú rozdiely v štandardoch, pravidlách, schémach, modelových dátach, registroch a referenčných systémoch adries. Dátová špecifikácia definuje adresu ako „identifikáciu pevne umiestneného objektu (napr.: pozemku, budovy) prostredníctvom štruktúrovaného zloženia geografických názvov a identifikátorov. Základný koncept špecifikácie spočíva v určení hierarchického adresného „lokátora“ (číslo domu, budovy). Adresa je definovaná ako priestorový objekt, ktorý identifikuje pevnú polohu nehnuteľnosti. Pre tento účel má adresa identifikátor, napr. číslo adresy alebo názov budovy, ktoré umožňujú používateľovi odlišovať sa od adries susedných nehnuteľností a určiť presnú geografickú polohu. Geografická poloha adresy (nehnuteľnosti) je reprezentovaná ako geografický bod. Na zistenie presnej adresy treba použiť „adresné komponenty“, ktoré definujú jej umiestnenie v určitej geografickej oblasti v súlade s ďalšími dátovými špecifikáciami smerníc INSPIRE bližšie uvedené v citovanej špecifikácii. Aplikačnú schému pre výmenu adries tvorí 5 častí: 1. *názov administratívnej jednotky* (napr. okres, municipalita), 2. *názov adresnej oblasti* (napr. názov mesta, obce), 3. *názov dopravnej komunikácie* (napr. názov ulice), 4. *adresný lokátor* (napr. číslo domu) a 5. *poštový „deskriptor“* (napr. poštové smerové číslo).

Dátová špecifikácia pre adresy je dôležitým krokom k vybudovaniu komplexnej európskej infraštruktúry priestorových informácií, ktorý ohrozuje najmä nedodržiavanie stanovených časových limitov pre tvorbu registrov adries a s nimi spojených jednotlivých členských krajín EÚ. Bohužiaľ ku krajinám, ktoré doposiaľ tento základný register adries nemajú patrí aj Slovensko. Problematiku adries pre Slovenskú republiku by mal riešiť projekt *Operačného plánu informatizácie spoločnosti* (OPIS) zameraný na vytvorenie informačného systému registra adries a sprístupnenie jeho služieb, ako jedného zo základných priestorových registrov integrovaného informačného systému verejnej správy (IS registra adries, 2009). Ak sa tak nestane, budú u nás dostupné len finančne náročné komerčné adresné systémy a služby.

2. Geokódovacia služba od spoločnosti Google

Pri práci s geografickými informačnými systémami (GIS) sú veľmi často využívané priestorové výbery, na základe ktorých možno identifikovať existujúce vzťahy medzi rôznymi jednotkami v priestore. Pri týchto operáciách sa často využíva pozícia centroidu konkrétnej entity reprezentujúcej modelovaný geografický objekt alebo jav. Túto možno využiť aj ako bodovú reprezentáciu danej priestorovej jednotky. Vzhľadom na metodiku vytvárania

centroidov pri tvorbe mapových vrstiev je však pochopiteľné, že takéto bodové vyjadrenie nevystihuje skutočnú polohu obce v rámci hraníc jej územia. Na túto skutočnosť sme chceli poukázať v tejto časti príspevku na modelovom území regiónu Liptov (územie okresov Liptovský Mikuláš a Ružomberok), ktorý predstavuje územie, v ktorom je väčšia časť obyvateľstva obcí koncentrovaná v sídelnej osi kopírujúcej údolie Váhu, kým hranice jednotlivých obcí prebiehajú po svahoch okolitých geomorfologických jednotiek. Práve v takýchto horských oblastiach s výraznou vertikálnou členitosťou terénu, sú rozdiely medzi pozíciou centroidu a reálnou pozíciou obce (v zmysle obytných domov) pomerne výrazné a môže teda dôjsť k pomerne značným skresleniam pri rôznych priestorových analýzach.

Na reálnejšie bodové vyjadrenie pozície obce v rámci jej územia sme použili jednu zo služieb spoločnosti Google, konkrétne geokódovaciu službu fungujúcu na štandardnom princípe konverzie textového adresného reťazca do podoby geografických súradníc daného bodu. Podobných geokódovacích mechanizmov existuje väčšie množstvo, pri našom výbere zohralo hlavnú úlohu predovšetkým primárne využitie získaných súradníc za účelom zobrazenia jednotlivých obcí SR v Google Maps. Nespornou výhodou je však aj množstvo návodov a dostupných informačných zdrojov o tomto nástroji. K zaujímavým prácam opisujúcim použitý algoritmus patria Pejic et al. (2009), Zhang a Shi (2007) a Bačík (2012).

Technický princíp využitej služby je detailne opísaný priamo na webových stránkach spoločnosti Google (The Google Geocoding API¹). V princípe možno definovať tri základné spôsoby geokódovania objektov, a to pomocou:

1. požiadavky odoslanej cez protokol HTTP,
2. vytvoreného mapového rozhrania,
3. dávkového geokódovania, ktoré je vhodné pri práci s veľkým objemom dát a bolo využité aj v našom prípade.

1. Výsledkom geokódovacieho procesu adresného určenia pomocou požiadavky cez protokol HTTP je štruktúrovaný dokument vo formáte XML obsahujúci okrem potrebných geografických súradníc aj viaceré informácie o začlenení danej priestorovej jednotky do vyšších administratívnych jednotiek (štát, VÚC, poštové smerovacie číslo, atď.).

Príklad takéhoto postupu ukážeme na hľadanej adrese: M. R. Štefánika, 154/39, Považská Bystrica, kde do prehľadávača stránok zadáme adresu v tvare:

```
„http://maps.googleapis.com/maps/api/geocode/xml?address=M.R.Štefánika 154/39 Považská Bystrica&sensor=true“.
```

Adresa sa následne upraví do potrebného tvaru:

```
„http://maps.googleapis.com/maps/api/geocode/xml?address=M.R.Štefánika%20154/39+Považská%20Bystrica+Slovensko&sensor=true“.
```

Výsledkom celého procesu je výpis súradníc v štruktúrovanom dokumente XML (obr. 1). Generovaná poloha adresného bodu je súčasťou „know-how“ poskytovateľa geokódovacej služby, v tomto prípade spoločnosti Google a jej zmluvných partnerov pre dodávku databázových systémov adries.

2. Jednoduchší a častejšie používaný spôsob predstavuje geokódovanie adresy na základe jej zadania do textového poľa na internetovej stránke (samotné spracovanie prebehne pomocou JavaScriptu). Výhodou je, že okrem potrebných súradníc sa daný bod okamžite objaví priamo v mape, na druhej strane však nezískame ďalšie informácie o danom objekte, tak ako pri použití protokolu HTTP. Porovnanie výsledkov procesov geokódovania je na obr. 1.

Obr. 1 Výsledok geokódovania adresy „M. R. Štefánika, 154/39, Považská Bystrica“

¹ <https://developers.google.com/maps/documentation/geocoding/>

3. Pre naše potreby (2927 obcí) sa nedali predchádzajúce spôsoby geokódovania z pochopiteľných dôvodov použiť. Pri spracovaní veľkého množstva adresných určení je najvhodnejšie použiť dávkové geokódovanie adres. Tento spôsob využíva skriptovací jazyk PHP, pomocou ktorého sa jednotlivé súradnice (v súradnicovom systéme World Geodetic System 84) vložia do dočasného súboru formátu CVS a následne konvertujú do vytvorenej databázovej tabuľky uloženej na serveri. Základným vstupom je teda adresné určenie jednotlivých miest – obcí. Z dôvodu duplicitných názvov slovenských obcí (na Slovensku je spolu 95 názvov obcí, ktoré sa vyskytujú viacnásobne, čo predstavuje spolu 205 obcí), bolo potrebné uviesť adresu v takom tvare, ktorá by eliminovala uvedenú zhodu názvov. Z tohto dôvodu bola adresa zapísaná v tvare: *nazov_obce*, *nazov_okresu*, *nazov_statu*. Táto informácia slúžila ako vstup do geokódovacieho procesu, výsledkom bolo získanie geografických súradníc takto definovaných adresných určení (obr. 2). Uvedený výstup bol primárne využitý v internetovej aplikácii venovanej slovenským obciam, ktorý detailne opisuje práca Bačík (2012).

Obr. 2: Databázová tabuľka so získanými geografickými súradnicami (lat, lng)

3. Rozdiely medzi pozíciou centroidu a geokódovanou adresou obce

Hlavným cieľom bolo zistiť, ako sa bude líšiť distribúcia obyvateľstva obcí regiónu Liptov (ako aj počtu obcí) vo vzťahu k vybraným prvkom cestnej siete, pri použití centroidov ako bodových prezentácií obcí na jednej strane a ich geokódovaných adres na strane druhej. Pri príprave potrebných mapových podkladov sme využili zdrojové vektorové údajové vrstvy areálov obcí, zástavby a líniovú vrstvu cestnej siete Slovenska poskytnuté firmou ESPRIT s.r.o. Banská Štiavnica. Súradnice centroidov jednotlivých obcí v národnom systéme (S-JTSK) sme získali pomocou funkcií *CentroidX(obj)* a *CentroidY(obj)* z areálovej vrstvy hraníc obcí. v programe desktop GIS MapInfo Professional (ver. 10), v ktorom boli spracované ďalšie prezentované výsledky. Ide o štandardnú funkcionality programu typu „plocha do bodu“, ktorá je obvyklou súčasťou podobných desktopových programov GIS (ArcGIS, TopoL a ďalšie) na báze algoritmu čiernej skrinky obvykle založenom na geometrickom výpočte ťažiska (stred) plochy. Hodnoty týchto, ako aj geografických súradníc jednotlivých obcí Slovenska generované geokódovacou adresnou službou Google, boli transformované do vektorových bodových vrstiev mapového formátu TAB v programe MapInfo Professional (funkcia *Create points...*). Polohu oboch bodových vrstiev ilustruje výsledná mapová kompozícia (obr. 3), kde je viditeľné prepojenie bodovej vrstvy geokódovaných adres s vrstvou sídiel (zastavané plochy obcí). Dá sa konštatovať, že takéto bodové vyjadrenie je oveľa bližšie k reálnej pozícii ako použitie centroidov.

Obr. 3. Región Liptov – porovnanie pozície centroidov obcí a geokódovaných adres

Na pozíciu bodu obce z geokódovanej adresy, na rozdiel od pozície centroidu, nemá žiadny vplyv tvar a veľkosť územia. Práve z tohto dôvodu bola ako modelové územie vybraná oblasť s výraznou vertikálnou členitosťou, ktorá do značnej miery ovplyvňuje výsledky ďalších analýz. Obce v horských oblastiach majú odlišný tvar a veľkosť spôsobený práve geomorfologickým profilom tohto územia. Obyvateľstvo je sústredené v dolinách, vo významných dopravných koridoroch a v blízkosti hlavných centier. Preto sme v ďalšej fáze zisťovali rozdiely v pozícii jednotlivých bodov. Rozdiely sme graficky znázornili pomocou línií vytvorených na základe použitia funkcie *SpiderGraph* v programe MapInfo Professional. Presnosť výslednej mapovej vrstvy vytvorenej geokódovaním sme porovnali aj s vrstvou

oficiálnych definičných bodov slovenských obcí distribuovaných Geodetickým a kartografickým ústavom SR v Bratislave (obr. 4).

Obr. 4. Porovnanie rozdielov medzi (a) centroidmi a geokódovanými adresami, (b) geokódovanými adresami a definičnými bodmi

Na obr. 4 vidieť výrazné rozdiely medzi danými pozíciami bodov. Niektoré štatistické hodnoty, vyplývajúce so vzájomného porovnania, sú pomerne zaujímavé. Celková suma rozdielov predstavuje dĺžku 195,8 km, a priemerná hodnota rozdielu je 2,41 km, čo možno považovať za pomerne vysoké číslo. Príčinou takejto vysokej hodnoty je už spomínaná skutočnosť tvaru georeliéfu v tejto oblasti a možno predpokladať, že v rovinných oblastiach Slovenska je tento rozdiel oveľa menší. Maximálne hodnoty rozdielov boli zaznamenané v obciach Kráľova Lehota (10,45 km) a Lubochňa (10,04 km). Ide o obce s najväčšou rozlohou spomedzi sledovaného súboru obcí. Výslednú vrstvu geokódovaných obcí sme následne porovnali s definičnými bodmi, čím sme si chceli overiť presnosť takto získanej bodovej vrstvy obcí. V tomto prípade sa suma rozdielov dostala na hodnotu 16,8 km, čo predstavuje priemer 0,2 km, a maximálna hodnota nepresiahla úroveň 0,9 km. Tieto hodnoty sú veľmi zaujímavé z hľadiska ďalšieho možného postupu pri tvorbe bodovej témy obcí pre rôzne územia, práve na základe možností geokódovacieho procesu využívajúceho uvedenú službu.

4. Distribúcia obyvateľstva regiónu Liptova podľa bodových lokácií obcí a vzdialeností od ciest

Rozdiely medzi pozíciou centroidu a polohou obce, stotožnenú s jej geokódovanou adresou, sú pomerne výrazné, a preto je predpoklad, že aj distribúcia obyvateľov vo vzťahu k vybraným prvkom cestnej siete bude vykazovať rozdielne výsledky. Získanie základnej predstavy o rozdieloch medzi oboma vrstvami bolo primárnym vstupom k ďalšej analýze, v ktorej bolo hlavným cieľom zistenie, ako táto skutočnosť vplýva na štatistické vyhodnotenie distribúcie obyvateľstva vo vzťahu k vybraným prvkom cestnej siete (obr. 5).

Obr. 5. Porovnanie polohy (a počtu) centroidov a geokódovaných adries v rôznych vzdialenostných zónach od (a) diaľnice (b) ciest 1. a 2. triedy

Na základe vytvorených vzdialenostných máp operáciami vektorovej mapovej algebry boli štatisticky vyhodnotené počty jednotlivých obcí, ako aj počet ich obyvateľov v jednotlivých vzdialenostných zónach. Základné získané hodnoty sú prezentované v obr. 6.

Obr. 6: Porovnanie počtu obcí a počtu obyvateľov v jednotlivých vzdialenostných zónach od diaľnic a ciest 1. a 2. triedy (pri použití troch rozdielnych bodových interpretácií polohy obce)

Z uvedenej tabuľky možno identifikovať niekoľko základných rozdielov, vyplývajúcich z rozdielnej interpretácie polohy obcí vybraného regiónu, vzťahnutej k trom odlišným bodovým reprezentáciám. Zo vstupných máp bolo zrejmé, že relatívne malé polohové rozdiely medzi geokódovanými adresami a definičnými bodmi sa premietnu do podobných štatistických výsledkov. Použitie oboch vyjadrení je veľmi podobné z hľadiska početnosti jednotlivých obcí, ako aj v počte obyvateľov zaradených do jednotlivých vzdialenostných zón. Existujú pochopiteľne menšie rozdiely, ktoré sú výsledkom rozdielov medzi pozíciami týchto objektov. Pri globálnom pohľade je vidieť, že sumárne hodnoty sú v oboch prípadoch zhodné (počet obcí aj počet obyvateľov). Táto skutočnosť je dôležitá z hľadiska ďalšieho možného využitia geokódovacieho procesu bodovo definovaných objektov. Keď však tieto hodnoty porovnáme s pozíciou centroidov, tu sú už jednotlivé rozdiely oveľa výraznejšie. Pri

hodnotení napr. rozmiestnenia obyvateľstva podľa ciest 1. a 2. triedy do 1 km od ciest je zaradených len 20 obcí, čo predstavuje 48755 obyvateľov. Pri použití predchádzajúcich dvoch vrstiev je to 37 obcí, čo tvorí v oboch prípadoch viac ako 100000 obyvateľov (viac ako 76% obyvateľov sledovaného regiónu). Z toho logicky vyplýva, že vo vzdialenosti väčšej ako 1 km býva necelých 25% obyvateľov. Naopak, podľa pozície centroidu je v tejto zóne (nad 1 km) viac ako 83000 obyvateľov a väčšia časť obcí. Podobné rozdiely, aj keď z nášho pohľadu menej významné, sú viditeľné aj pri sledovaní distribúcie obyvateľstva vo vybraných zónach od diaľnic (tab. 1).

Záver

Potreba čo najpresnejšieho polohového určenia je dôležitým faktorom ovplyvňujúcim výslednú presnosť prezentovaných analýz. V bodovo prezentovaných modelovaných geografických objektoch a javoch (napr. autobusové zastávky, bankomaty), resp. ďalších bodov záujmu sa dajú procesom geokódovania získať presné geografické súradnice pomerne jednoducho. Existuje viacero spôsobov, jedným z nich je použitie adresnej geokódovacej služby od spoločnosti Google, ktorý sme prezentovali. Môže však byť použitý ktorýkoľvek iný geokódovací algoritmus, v ktorom sa dá transformovať adresná identifikácia geografického objektu do podoby geografických súradníc lokalizačného bodu. V prípade priestorových jednotiek areálového typu je však použitie bodového vyjadrenia pomerne zložité. Ideálnym stavom by bola dostupnosť štatistických dát na úrovni čo najmenších priestorových jednotiek (napr. budova, dom), na základe ktorých by sme vedeli s pomerne veľkou presnosťou vyjadriť rôzne ukazovatele v priestore. Pri práci s rôznymi desktopovými programami GIS sa pomerne často používajú centroidy týchto jednotiek, ktorých poloha je ovplyvnená celkovou rozlohou a tvarom daného územia. Ich poloha však často nereflektuje skutočnú pozíciu sídla (v zmysle obývaných plôch tohto územia). Túto skutočnosť sme demonštrovali na príklade regiónu s výraznejšou vertikálnou členitosťou, pretože práve sídla v takýchto oblastiach sú lokalizované v nižšej nadmorskej výške, pri hlavných cestných ťahoch, pričom ich územie pokrýva celé oblasti až po hrebeň príslušných geomorfologických celkov. Pri lokalizovaní jednotlivých obcí sme použili jednoduchý adresný zápis „nazov_obce, nazov_okresu, nazov_statu“, ktorého výsledkom je však priame lokalizovanie daného bodu do zastavaného územia obce (na rozdiel od centroidu). Pri podobných analýzach môžeme výslednú presnosť ovplyvniť zadaním presnej adresy (napr. obecného úradu), čím by sme získali lokalizáciu obce vyjadrenú práve centrom samosprávy zo samosprávneho pohľadu a podobne. Výslednú vrstvu sme porovnali s cieľom overiť presnosť zvoleného postupu s centroidmi obcí a štátnym registrom definičných bodov základných územných jednotiek, ktorý bol vytvorený pre technickú prax. Z hľadiska výslednej presnosti považujeme použitý geokódovací algoritmus za postačujúci pre veľkú časť geografických analýz na sledovanej mierkovej úrovni územných jednotiek. Výhodou je možnosť generovania ľubovoľných bodových vrstiev pre rozličné územia a objekty a vykonávanie ďalších analýz podobného zamerania s možnosťou variabilného stanovenia nepriameho identifikátora (adresy) územnej jednotky pre proces geokódovania v oboch smeroch, teda od názvu adresy k adresnému bodu a naopak. Okrem toho sa takto získané dáta dajú využiť v rozličných aplikáciách typu web mapping v prostredí internetu.

PodĎakovanie

Príspevok vznikol s podporou grantovej výskumnej úlohy VEGA č. 1/0562/12.

Geocoding and its application on the Slovak communes – the case study of region Liptov.

Summary

The requirement for the most accurate determination of a position is an important factor that affects the accuracy of the results analysed in the paper. In the pointed geographical objects and phenomena (e.g. bus stops, ATMs), respectively other points of interest, precise geographic coordinates can be obtained using geocoding process relatively easily. There are several ways of doing so, one of them is the use of Google address geocoding service, presented in this paper. Certainly, any other geocoding algorithm can be used, which can transform an object's address identification into a form of geographical coordinates of localised point. In the case of spatial units of polygon type the point interpretation however is quite difficult. In ideal situation, with available statistical data of the smallest spatial units (e.g. buildings, houses), we are able to express various indicators in space with relatively high accuracy. In various desktop GIS programs the centroids of these units, which position is affected by the overall size and shape of the area, are used quite often. However, their position rarely reflects the actual point of a real central position (populated areas of these units). This fact is shown on the example of the region with high vertical segmentation, where the centres of communes are located at lower altitudes, commonly near main roads, whereas their area sometimes covers the territory up to the ridges of relevant geomorphologic units. Locating individual communes, we have applied the simple addressed record in the form of a commune name, district name and state name (to avoid duplicity in commune's names), resulted however in the direct localisation of such defined point in the built-up area of a commune (in contrary to centroids). Similarly, it is possible to affect the final accuracy by entering the exact address (e.g. the address of a local authority). To verify the accuracy of geocoding procedure, the resulting layer was compared with the communes centroids and with the state registers of defined points of basic territorial units created for the purposes of technical practice. We have demonstrated the differences in pointed localised commune's expressions on the example of population distribution of individual communes in relation to their position in various distance zones of selected elements of the road network in a selected region (Liptov). The results identify several fundamental differences which originate from the different interpretation of a commune position. The use of geocoded commune addresses and defined points brings very similar results in terms of the number of individual communes as well as the number of people enrolled in each distance zones. This fact is important for other possible use of geocoding processes of pointed localised objects in other analyses. However, when these values are compared with the centroid position, there can be observed some significant differences. For example, the evaluation of population distribution by road of first and second class in 1 km zone from these routes includes only 20 communes, representing 48 755 residents (using centroids). Geocoded address and definitional points show totally 37 communes with more than 100 000 residents (more than 76% of all residents in the selected region). From the point of view of final accuracy the geocoded algorithm is considered to be sufficient for the large part of geographical analyses in territorial units of the same scale. The advantage is in the possibility to generate pointed layers for different areas and objects of your own choice, as well as the performance of further similarly oriented analyses with the option of variable determination of indirect identifier (address) of a territorial unit for the geocoding process in both directions, from the address name to the address point and vice versa. In addition, the data acquired in such a way can be used in various web mapping applications on the internet.

Literatúra

BAČÍK, V. 2012: Google API a ich využitie pri tvorbe internetovej stránky obcí Slovenskej republiky. *GEOGRAPHIA CASSOVIENSIS*, VI. roč. (číslo 1), Košice: Prírodovedecká fakulta Univerzity P. J. Šafárika v Košiciach, 2012, s. 5-10.

- BRACKEN, I., MARTIN, D. 1989. The generation of spatial population distributions from census centroid data. *Environment and Planning*, 2008, A, 21, pp. 537-543. Dostupné na internete: < <http://www.ij-healthgeographics.com/content/7/1/13> >.
- Cenník, 2011. *Cenník produktov a služieb geodetických a kartografických činností*. Bratislava: Geodetický a kartografický ústav Bratislava, 2011.
- Data Specification on Addresses – Guidelines D2.8.I.5 INSPIRE* 2011. [online], [Citované 24.11.2012]. Dostupné na internete: < http://inspire.gov.sk/Upload/documents/dataspec_3_0_1/INSPIRE_DataSpecification_AD_v3.0.1.pdf >.
- Geografická informácia 2003. *Geografická informácia Terminológia v normách ISO (2002-2003)*. Zborník referátov zo seminára. Pravda, J. ed. Bratislava: Geografický ústav SAV, Bratislava, 2003. 84 s.
- IS registra adries. 2009. *Čiastková štúdia uskutočniteľnosti projektov priority osi I Elektronizácia verejnej správy a rozvoj elektronických služieb OPIS zamerané na celkovú architektúru eGovernmentu, Register adries*. 2009. Bratislava: Ministerstvo financií SR. [online], [Citované 22.12.2012]. Dostupné na internete: < http://www.opis.gov.sk/data/files/2459_5512.pdf >.
- MADAJOVÁ, M. 2010. Metódy transformácie priestorových dát: prehľad, klasifikácia a hodnotenie. *Acta Geographica Universitatis Comenianae*, Vol. 54, 2010, No. 1, pp. 119-135.
- Metodika revízie základných sídelných jednotiek (pre sčítanie obyvateľov, domov a bytov v r. 2001)* 1999. Bratislava: Slovenská agentúra životného prostredia Banská Bystrica, 1991, 31 s. [online], [Citované 30. 12. 2012]. Dostupné na internete: < www.iszp.sk/nastroje/ism/osidlenie/sodb/sodb.doc >.
- Nariadenie 2008. *Nariadenie Európskeho parlamentu a Rady (ES) č. 451/2008 z 23. apríla 2008*, ktorým sa zavádza nová štatistická klasifikácia produktov podľa činností (CPA) a ktorým sa zrušuje nariadenie Rady (EHS) č. 3696/93.
- PEJIC, A., PLETL, S., PEJIC, B. 2009. An expert system for tourists using Google Maps API. In: *Intelligent Systems and Informatics, SISY '09*. 7th International Symposium, 2009, pp. 317-322.
- RATCLIFFE, J., H. 2001. On the accuracy of TIGER-type geocoded address data in relation to cadastral and census areal units. *International Journal of Geographic Information Sciences*, 2001, vol. 15, no. 5, pp. 473-485. Dostupné na internete: < [http://jratcliffe.net/papers/Ratcliffe%20\(2001\)%20On%20the%20accuracy%20of%20TIGER-type%20geocoding.pdf](http://jratcliffe.net/papers/Ratcliffe%20(2001)%20On%20the%20accuracy%20of%20TIGER-type%20geocoding.pdf) >.
- RICHESON, D. 2005. Centers of the United States. *The College Mathematics Journal*, 36 (November 2005) No. 5, pp. 366-373.
- RYDEN, K. (ed.) 2005. *OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 2: SQL option*. Open Geospatial Consortium, Inc. [online], [Citované 24.11.2012]. Dostupné na internete: < http://portal.opengeospatial.org/files/?artifact_id=25354 >.
- Smernica Európskeho parlamentu a Rady 2007/2/ES zo 14. marca 2007*, ktorou sa zriaďuje Infraštruktúra pre priestorové informácie v Európskom spoločenstve (Inspire).
- Terminologický slovník geodézie, kartografie a katastra 1998*. Bratislava: Úrad geodézie, kartografie a katastra Slovenskej republiky, Český úrad zeměměřický a katastrální, 539 s. ISBN 80-88716-36-5.
- TeSlo 2012. *Terminologický slovník Spoločenstva NIPI*. [online], [Citované 30. 12. 2012]. Dostupné na internete: < <http://gis.fns.uniba.sk/TeSlo/index.php/Kateg%C3%B3ria:TeSloSK> >.

TUČEK, J. 1998. *Geografické informační systémy. Principy a praxe*. Praha: Computer Press, 1998, 424 s. ISBN 80-7226-091-X.

ZHANG, J. Y, SHI, H. 2007. Geospatial Visualization using Google Maps: A Case Study on Conference Presenters. In: *Second International Multisymposium on Computer and Computational Sciences*, pp. 472–476.

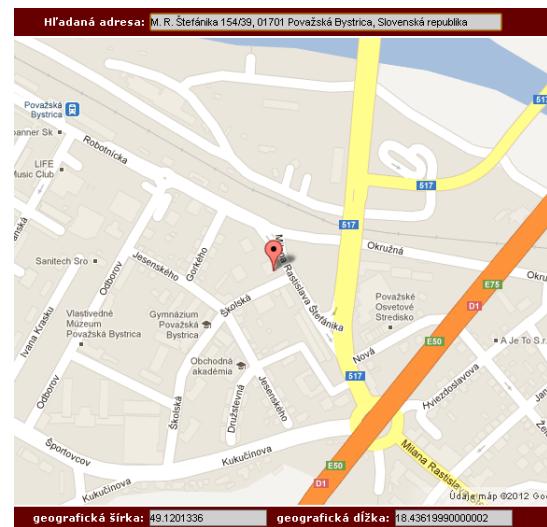
Obrázky

Obr. 1 Výsledok geokódovania adresy „M. R. Štefánika, 154/39, Považská Bystrica“

```

<GeocodeResponse>
  <status>OK</status>
  <result>
    <type>street_address</type>
    <formatted_address>
      M. R. Štefánika 154/39, 01701 Považská Bystrica</formatted_address>
    <address_component>
      <long_name>39</long_name>
      <short_name>39</short_name>
      <type>street_number</type>
    </address_component>
    <address_component>
      <long_name>154</long_name>
      <short_name>154</short_name>
    </address_component>
    <address_component>
      <long_name>M. R. Štefánika</long_name>
      <short_name>M. R. Štefánika</short_name>
      <type>route</type>
    </address_component>
    <address_component>...</address_component>
    <address_component>...</address_component>
    <address_component>
      <long_name>Považská Bystrica</long_name>
      <short_name>Považská Bystrica</short_name>
      <type>administrative_area_level_2</type>
      <type>political</type>
    </address_component>
    <address_component>
      <long_name>Trenčiansky kraj</long_name>
      <short_name>Trenčiansky kraj</short_name>
      <type>administrative_area_level_1</type>
      <type>political</type>
    </address_component>
    <address_component>
      <long_name>Slovenská republika</long_name>
      <short_name>SK</short_name>
      <type>country</type>
      <type>political</type>
    </address_component>
    <address_component>
      <long_name>01701</long_name>
      <short_name>01701</short_name>
      <type>postal_code</type>
    </address_component>
    <geometry>
      <location>
        <lat>49.1201336</lat>
        <lng>18.4361999</lng>
      </location>
      <location_type>ROOFTOP</location_type>
      <viewport>...</viewport>
    </geometry>
  </result>
</GeocodeResponse>

```



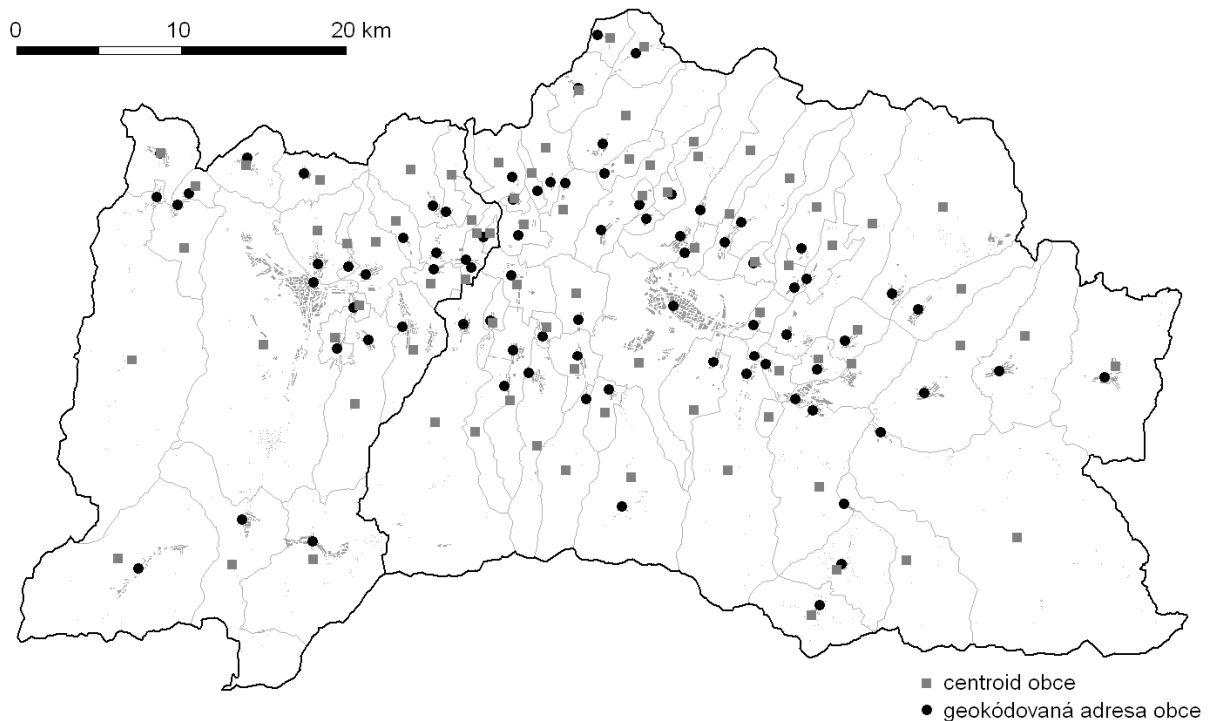
↑
**Pomocou grafického rozhrania
 v podobe internetovej stránky**
 Vidieť presnú polohu bodu v mape

←
**Pomocou http protokolu (zadanie
 konkrétnej adresy v požadovanom
 tvare)**
 Vidieť ďalšie spresňujúce informácie
 o danej adrese

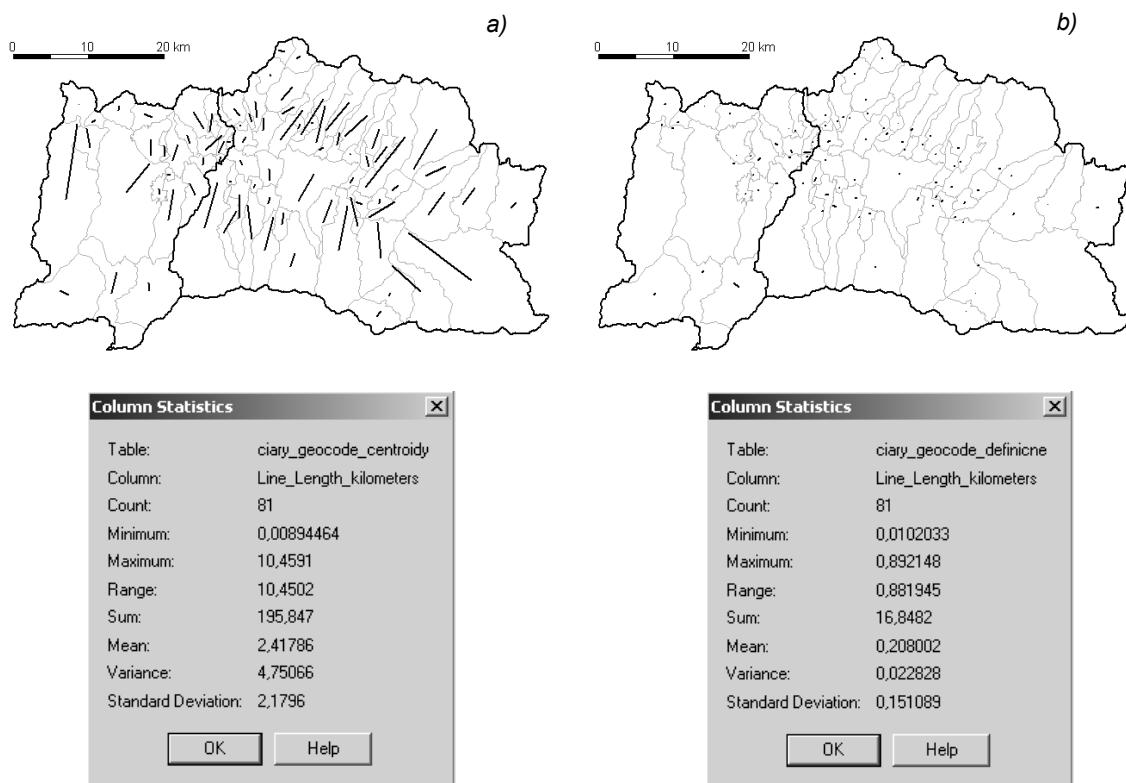
Obr. 2: Databázová tabuľka so získanými geografickými súradnicami (lat, lng) v systéme WGS84

			id	name	adresa	lat	lng	typ	kod
<input type="checkbox"/>			1	Nitra	Nitra, Nitra, Slovakia	48.30989075	18.08587646	mesto	500011
<input type="checkbox"/>			2	Alekšince	Alekšince, Nitra, Slovakia	48.36696625	17.94747162	vo	500020
<input type="checkbox"/>			3	Báb	Báb, Nitra, Slovakia	48.31032562	17.86828041	vo	500046
<input type="checkbox"/>			4	Beladice	Beladice, Zlaté Moravce, Slovakia	48.34511948	18.29345512	vo	500062
<input type="checkbox"/>			5	Branč	Branč, Nitra, Slovakia	48.20994568	18.14695358	vo	500071
<input type="checkbox"/>			6	Čakajovce	Čakajovce, Nitra, Slovakia	48.36965942	18.03790474	vo	500101
<input type="checkbox"/>			7	Čaradice	Čaradice, Zlaté Moravce, Slovakia	48.35884476	18.51070786	vo	500127
<input type="checkbox"/>			8	Čeladice	Čeladice, Nitra, Slovakia	48.32523346	18.24174500	vo	500135
<input type="checkbox"/>			9	Čierne Kľačany	Čierne Kľačany, Zlaté Moravce, Slovakia	48.35202026	18.42083549	vo	500151
<input type="checkbox"/>			10	Čířáre	Čířáre, Nitra, Slovakia	48.23239517	18.39216995	vo	500160

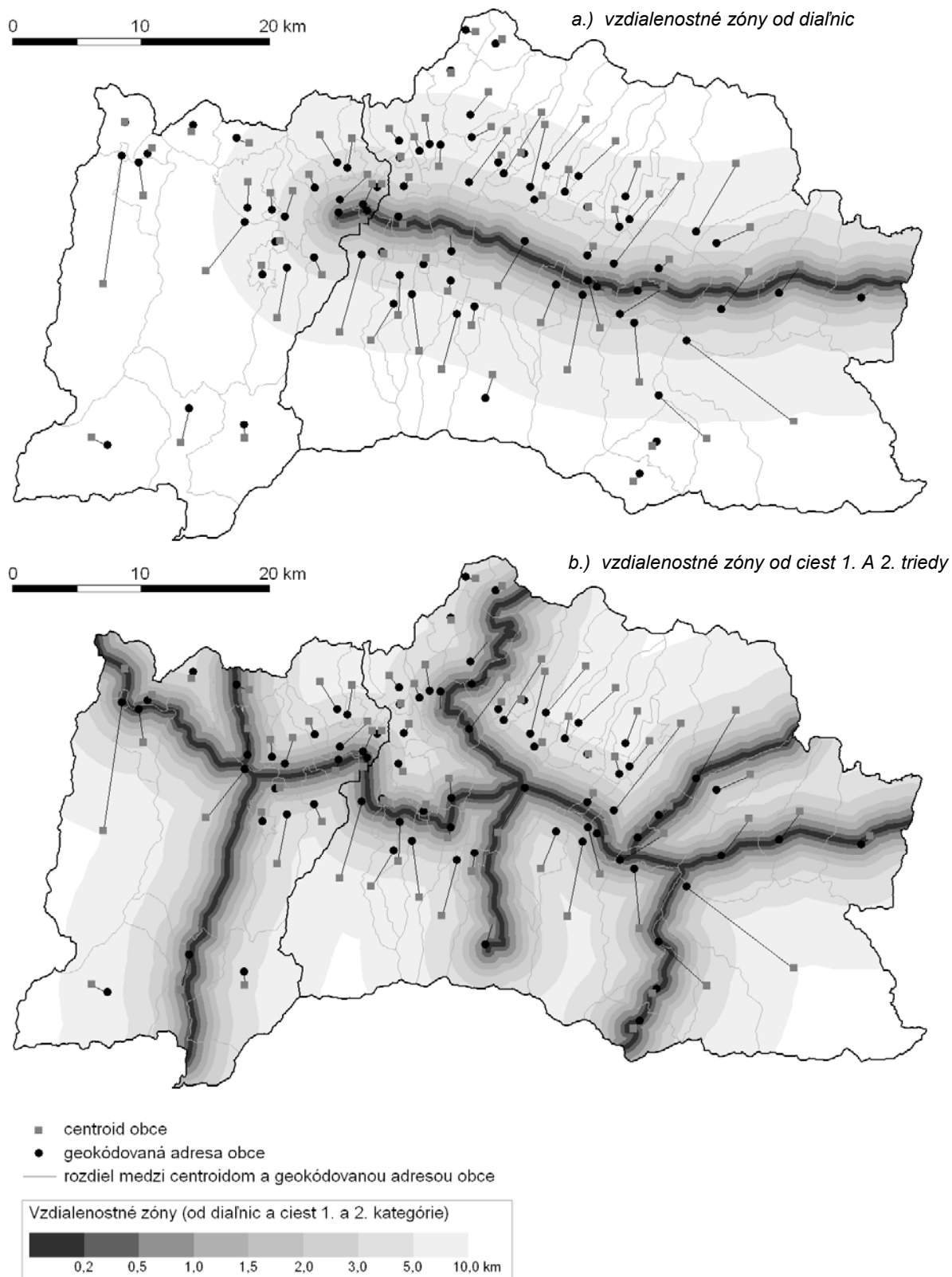
Obr. 3: Región Liptov – porovnanie pozície centroidov obcí a geokódovaných adries



Obr. 4: Porovnanie rozdielov medzi (a) centroidmi a geokódovanými adresami, (b) rozdielov medzi geokódovanými adresami a definičnými bodmi



Obr. 5: Porovnanie polohy (a počtu) centroidov a geokódovaných adries v rôznych vzdialenostných zónach od diaľnic (a), a ciest 1. a 2. triedy (b)



Obr. 6: Porovnanie počtu obcí a počtu obyvateľov v jednotlivých vzdialenostných zónach od diaľnic a ciest 1. a 2. triedy (pri použití 3 rozdielnych bodových interpretácií polohy obce)

