

**UNIVERZITA KOMENSKÉHO V BRATISLAVE**  
**PRÍRODOVEDECKÁ FAKULTA**

**PRIESTOROVÉ DÁTA O KRAJINNEJ POKRÝVKE**  
**A HODNOTENIE TEMATICKEJ**  
**PRESNOSTI EURÓPSKEJ VRSTVY HUSTOTY**  
**NEPRIEPUSTNOSTI**

Dizertačná práca

2022

Mgr. Andrea ČORBOVÁ

**UNIVERZITA KOMENSKÉHO V BRATISLAVE**  
**PRÍRODOVEDECKÁ FAKULTA**

**PRIESTOROVÉ DÁTA O KRAJINNEJ POKRÝVKE**  
**A HODNOTENIE TEMATICKEJ**  
**PRESNOSTI EURÓPSKEJ VRSTVY HUSTOTY**  
**NEPRIEPUSTNOSTI**

Dizertačná práca

Študijný program: Regionálna geografia  
Študijný odbor: Vedy o Zemi  
Školiace pracovisko: Geografický ústav SAV  
Školiteľ: Mgr. Pavol Hurbánek, PhD.

**Bratislava 2022**

**Mgr. Andrea ČORBOVÁ**



Univerzita Komenského v Bratislave  
Prírodovedecká fakulta

## ZADANIE ZÁVEREČNEJ PRÁCE

**Meno a priezvisko študenta:** Mgr. Andrea Čorbová  
**Študijný program:** regionálna geografia (Jednoodborové štúdium, doktorandské III. st., externá forma)  
**Študijný odbor:** vedy o Zemi  
**Typ záverečnej práce:** dizertačná  
**Jazyk záverečnej práce:** slovenský  
**Sekundárny jazyk:** anglický

**Názov:** Priestorové dáta o krajinnej pokrývke a hodnotenie tematickej presnosti Európskej vrstvy hustoty nepriepustnosti  
*Spatial data on land cover and thematic accuracy assessment of the European Imperviousness Density Layer*

**Anotácia:** Rozvoj geoinformačných technológií a dátových vrstiev programu Európskej únie pre pozorovanie Zeme nazývaného Copernicus, predovšetkým v rámci služby pre monitorovanie krajiny (Land Monitoring Service), ako aj príbuzných dátových vrstiev, zväčša s rozsahom pre celú Európu alebo jej veľkú časť, ponúka vhodné nástroje pre analýzu stavu a zmien priestorového rozmiestnenia nepriepustných a zastavaných plôch v krajine. Keďže tieto plochy indikujú rozmiestnenie obyvateľstva a jeho aktivít, uvedené dátové vrstvy majú široké využitie, okrem iného aj pre populačné modelovanie alebo dezagregáciu sociálnoekonomických premenných. Úspešné využívanie týchto dátových vrstiev však vyžaduje poznanie rozličných charakteristík ich presnosti. Keďže rôzne aplikácie majú špecifické požiadavky z pohľadu týchto charakteristík, cieľom práce je hodnotenie rozličných aspektov tematickej presnosti uvedených dátových vrstiev o krajinnej pokrývke.

**Školiteľ:** Mgr. Pavol Hurbánek, PhD.  
**Katedra:** PriF.KRGRR - Katedra regionálnej geografie a rozvoja regiónov  
**PriF vedúci katedry:** prof. RNDr. Ladislav Tolmáči, PhD.

**Dátum zadania:** 12.03.2015

**Dátum schválenia:** 12.03.2015

prof. RNDr. Viliam Lauko, CSc.  
garant študijného programu

.....  
študent

.....  
školiteľ

## **Čestné vyhlásenie**

Čestne vyhlasujem, že som predloženú dizertačnú prácu spracovala samostatne, s použitím uvedenej literatúry, vlastných poznatkov a ďalších informačných zdrojov a neporušila som autorský zákon.

Bratislava 2022

.....

Podpis autora práce

## **Pod'akovanie**

Ďakujem vedúcemu dizertačnej práce, Mgr. Pavlovi Hurbánkovi, PhD., za usmerňovanie a cenné rady pri vypracovaní dizertačnej práce, za jeho pomoc, trpezlivosť a pochopenie. Zároveň ďakujem všetkým, ktorí mi pri písaní práce vyšli v ústrety, hlavne môjmu manželovi a mojej mame.

## ABSTRAKT

ČORBOVÁ, Andrea (2022): *Priestorové dáta o krajinnej pokrývke a hodnotenie tematickej presnosti Európskej vrstvy hustoty nepriepustnosti* [Dizertačná práca]. Bratislava: Univerzita Komenského v Bratislave, Prírodovedecká fakulta, 109 strán.

V rámci európskeho programu Copernicus monitoruje Európska únia našu Zem a poskytuje nám priestorové dáta o krajinnej pokrývke, ktoré majú širokospektrálne využitie. Dôležitým aspektom je ich presnosť nakoľko zodpovedajú realite. Jednou z takzvaných vrstiev s vysokým rozlíšením je aj vrstva hustoty nepriepustnosti (Imperviousness Density – IMD), ktorá predstavuje raster s rozlíšením 100 m s celočíselnými hodnotami pixelov od 0 do 100 %, reprezentujúcimi stupeň nepriepustnosti. Vytvorená bola takmer pre celú Európu a je dostupná pre roky 2006, 2009, 2012, 2015 a 2018. Doposiaľ vykonané validácie zo strany producentov boli založené na stratifikovaných náhodných vzorkách, ktorých veľkosť bola však v porovnaní so záujmovým územím pomerne malá, vďaka čomu bola podľa požiadaviek dosiahnutá jej vysoká tematická presnosť. Cieľom tejto dizertačnej práce je čitateľovi priblížiť nielen ponuku a dostupnosť priestorových dát o krajinnej pokrývke, ktoré poskytuje spomínaný program Copernicus, ale ponúknuť aj rôzne štatistické metódy, pomocou ktorých sa následne dajú obdobným spôsobom validovať tieto intenzitné mapy s vysokým rozlíšením presnejšie. Vykonaná bola komplexná validácia, teda zhodnotenie tematickej presnosti IMD, pre referenčné roky 2006 a 2009 na Slovensku a v Česku, pomocou kvantitatívnych štatistických metód založených na veľkej jednoduchej náhodnej vzorke (20 000 pixelov s priestorovým rozlíšením 100 m pre obidva roky v oboch záujmových územiach). Veľká vzorka totiž dokáže presnejšie odhadnúť najmä chybu vynechania v porovnaní s menšími vzorkami. Dôraz bol kladený na to, ako sa táto presnosť vyvíjala. Analýza odhaľuje postupné zlepšovanie tematickej presnosti IMD takmer s každou novou publikovanou verziou. Avšak nepriepustných povrchov, ktoré boli nesprávne klasifikované ako priepustné vo veľkom počte pixelov s mapovými hodnotami nulovej nepriepustnosti a s malými nenulovými referenčnými hodnotami nepriepustnosti, teda výskyt menších (ale z istého pohľadu závažnejších) chýb podhodnotenia, ostáva naďalej veľmi veľký.

**Kľúčové slová:** Copernicus. Priestorové dáta o krajinnej pokrývke. IMD. Nepriepustnosť. Tematická presnosť.

## **ABSTRACT**

ČORBOVÁ, Andrea (2022): *Spatial data on land cover and thematic accuracy assessment of the European Imperviousness Density Layer* [Dissertation work]. Bratislava: Comenius University in Bratislava, Faculty of Science, 109 pages.

Under the European Copernicus program, the European Union monitors our Earth and provides us with spatial data on land cover that has a wide range of uses. An important aspect is their transferability as they correspond to reality. One of the so-called High Resolution Layers is also the Imperviousness Density Layer (Imperviousness Density – IMD), which represents a raster with a resolution of 100 m and pixel integer values from 0 to 100 % representing the imperviousness degree. It has been created for almost the whole of Europe and is available for years 2006, 2009, 2012, 2015 and 2018. The validations performed so far by the producers were based on stratified random samples, the size of which, however, was relatively small compared to the area of interest, thanks to which its high thematic accuracy was achieved as required. The aim of this Dissertation work is not only to elucidate the reader the supply and availability of spatial data on land cover provided by the Copernicus program, but also to offer various statistical methods that can then be used to validate these high-resolution intensity maps more accurately. A comprehensive validation was performed, i.e. an evaluation of the IMD thematic accuracy, for the reference years 2006 and 2009 in Slovakia and Czechia, using quantitative statistical methods based on a large simple random sample (20,000 pixels with a spatial resolution of 100 m for both years in both areas of interest). In particular, a large sample can more accurately estimate the omission error compared to smaller samples. Emphasis was placed on how this accuracy developed. The analysis reveals a gradual improvement in IMD thematic accuracy with almost every new version released. However, impervious areas that have been incorrectly classified as pervious in a large number of pixels with zero imperviousness map values and small non-zero imperviousness reference values, i.e. the occurrence of small (but major from a certain perspective) underestimation errors., remains very high.

**Key words:** Copernicus. Spatial land cover data. IMD. Imperviousness. Thematic accuracy.

## PREDHOVOR

Priestorové dáta o krajinnej pokrývke ponúkajú širokú paletu využitia, ako pre bežného človeka, tak aj pre vedu a výskum, či pre politické účely. Vo všeobecnosti, kým obyčajný človek sa často nezamýšľa nad zdrojom dát a informácií, s ktorými narába, špecializované ústavy a vedecké inštitúcie siahajú po prameni, ktorého neoddeliteľnou súčasťou je práve poznatok o jeho vierohodnosti, presnosti a vedomosti o tom, na koľko sú informácie stotožniteľné so skutočnosťou. A tak je to aj v geografii, resp. konkrétne aj pri výskume a využívaní priestorových dát o krajinnej pokrývke.

Ešte pred vznikom plánovanej vrstvy monitorujúcej vybrané územie sa diskutuje o požiadavkách kladených na daný produkt, súčasťou ktorých je dosiahnutie stanovenej presnosti. Následne sa stanoví spôsob validácie a verifikácie daného produktu, ktorý však nie je mnohokrát dostatočný. Štatistické metódy, využité pri zisťovaní tematickej presnosti skúmaných produktov, sú často vopred nastavené tak, aby výsledky postupov dosahovali požadované hodnoty. Zabúda sa aj na podmienky, ktoré síce nie sú súčasťou samotných požiadaviek na validáciu či verifikáciu konkrétnych produktov, ale sú súčasťou využiteľnosti použitých štatistických metód. Takou podmienkou je napr. stanovenie dostatočnej veľkosti rozsahu vzorky, na ktorej sa má daná presnosť skúmať. Výber príliš malej databázy na preverenie vierohodnosti často až príliš skresľuje výsledky, ktoré sa síce na prvý pohľad zdajú postačujúce, ale po preskúmaní skutočnosti nemajú ani zďaleka takú výpovednú hodnotu, akou sú prezentované. A práve to bolo motiváciou k napísaniu tejto dizertačnej práce.

V tejto oblasti sme sa rozhodli ponúknuť vierohodný postup skúmania tematickej presnosti priestorových dát o krajinnej pokrývke výberom dostatočne veľkej vzorky a navrhnutím vhodných validačných metód. Zamerali sme sa na jednu konkrétnu vrstvu, ktorou je Európska vrstva hustoty nepriepustnosti, zverejňovaná postupne od roku 2006 v trojročných intervaloch. Aby boli naše výsledky porovnateľné s výsledkami iných validácií, sami sme niekoľkokrát vylepšovali naše jednotlivé referenčné databázy, ktoré sme zároveň umiestnili zámerne na územie dvoch štátov, čím sme mohli výsledky navzájom porovnávať. Samozrejme, nastali aj chyby a nepresnosti, ktorým sme sa snažili prísť na koreň tým, že ponúkame možné zdroje chýb a návrhy na ďalšie zlepšenia pri obdobných postupoch validácií priestorových dát o krajinnej pokrývke.



# OBSAH

ZOZNAM GRAFOV, OBRÁZKOV A TABULIEK.....	10
ZOZNAM SKRATIEK A ZNAČIEK .....	13
ÚVOD.....	15
1 TEORETICKÉ VÝCHODISKÁ .....	18
1.1 Pozadie vzniku priestorových dát – európska iniciatíva Copernicus.....	18
1.1.1 Štruktúra fungovania programu .....	19
1.1.2 Správa a financovanie programu .....	20
1.2 Copernicus – program poskytujúci priestorové dáta o krajinnej pokrývke .....	22
1.2.1 Copernicus globálne služby monitorovania Zeme.....	22
1.2.2 Copernicus celoeurópske služby monitorovania Zeme .....	24
1.2.3 Copernicus lokálne služby monitorovania Zeme.....	31
1.2.4 Služby poskytujúce snímky a referenčné údaje na celosvetovej a celoeurópskej úrovni.....	32
1.3 Súčasný stav riešenej problematiky doma a v zahraničí.....	33
1.3.1 Význam a zdôvodnenie potreby validácie .....	35
1.3.2 Definície základných pojmov .....	35
1.3.3 Niektoré prípadové štúdie validácií doma i v zahraničí .....	38
2 CIEĽ PRÁCE.....	59
3 METODIKA PRÁCE A METÓDY SKÚMANIA.....	60
3.1 Záujmové územie a príprava referenčnej databázy.....	62
3.2 Metóda histogramov.....	67
3.3 Korelačná analýza .....	68
3.4 Výpočet totálnej absolútnej chyby normalizovanej a jej štruktúra .....	72
3.5 Ďalšie navrhované metódy .....	75
4 VÝSLEDKY PRÁCE .....	79
4.1 Hodnotenie tematickej presnosti metódou histogramov .....	82
4.2 Hodnotenie tematickej presnosti korelačnou analýzou.....	88
4.3 Hodnotenie tematickej presnosti výpočtom priemernej absolútnej chyby normalizovanej a jej štruktúry.....	90

5	DISKUSIA.....	96
5.1	Možné zdroje chýb v pôvodných databázach .....	96
5.2	Možné zdroje chýb v referenčných databázach .....	97
	ZÁVER .....	99
	ZOZNAM POUŽITEJ LITERATÚRY .....	102

## **ZOZNAM GRAFOV, OBRÁZKOV A TABULIEK**

### **Grafy**

Graf 1 Znázornenie priebehu 95 % intervalu spoľahlivosti pre mriežku  $5 \times 5$  bodov v celom rozsahu hodnôt stupňa nepriepustnosti

Graf 2 Príklad histogramu rozdelenia hodnôt hustoty nepriepustnosti v IMD a IRD databázach

Graf 3 Príklad histogramu porovnania početnosti nadhodnotených a podhodnotených pixelov podľa rozdielu stupňa nepriepustnosti v IMD a IRD databáze

Graf 4 Histogram rozdelenia hodnôt hustoty nepriepustnosti vo všetkých dostupných databázach IMD2006 a v referenčnej databáze IRD2006\_A\_v3 pre Slovensko

Graf 5 Histogram rozdelenia hodnôt hustoty nepriepustnosti vo všetkých dostupných databázach IMD2006 a v referenčnej databáze IRD2006\_A\_v3 pre Česko

Graf 6 Histogram rozdelenia hodnôt hustoty nepriepustnosti vo všetkých dostupných databázach IMD2009 a v referenčnej databáze IRD2009\_A\_v3 pre Slovensko

Graf 7 Histogram rozdelenia hodnôt hustoty nepriepustnosti vo všetkých dostupných databázach IMD2009 a v referenčnej databáze IRD2009\_A\_v3 pre Česko

Graf 8 Histogramy rozdelenia hodnôt hustoty nepriepustnosti v databázach IMD2006\_v5 a IMD2009\_v3 a v referenčných databázach pre Slovensko a Česko

Graf 9 Histogramy početností rôznych hodnôt chýb nadhodnotenia a podhodnotenia získané porovnaním stupňa hustoty nepriepustnosti IMD databáz s príslušnými databázami IRD na Slovensku

Graf 10 Histogramy početností rôznych hodnôt chýb nadhodnotenia a podhodnotenia získané porovnaním stupňa hustoty nepriepustnosti IMD databáz s príslušnými databázami IRD v Česku

Graf 11 Hodnoty korelačných koeficientov pre vzťahy medzi referenčnými databázami IRD\_A/B a príslušnými validovanými databázami IMD pre roky 2006 a 2009 na Slovensku a v Česku

Graf 12 Celková absolútna chyba normalizovaná (TAEn) v percentách a jej štruktúra vypočítaná pre referenčnú databázu IRD\_A

Graf 13 Porovnanie hodnôt všetkých druhov chýb vypočítané pre referenčnú databázu IDR\_A pre rok 2006 na území Slovenska a Česka

Graf 14 Porovnanie hodnôt všetkých druhov chýb vypočítané pre referenčnú databázu IDR\_A pre rok 2006 na území Slovenska a Česka

Graf 15 Percentuálny podiel jednotlivých druhov chýb na celkovej absolútnej chybe normalizovanej vypočítanej pre referenčnú databázu IRD\_A

## **Obrázky**

Obrázok 1 Schéma zapojených subjektov

Obrázok 2 Vizuálna interpretácia bodov vo vzorke  $100 \times 100$  m

Obrázok 3 Rozmiestnenie náhodného výberu 73 pixelov na Slovensku

Obrázok 4 Referenčné lokality s rozmermi  $6 \times 6$  km

Obrázok 5 Rozdelenie krajín alebo skupín krajín EHP39 pri validáciach

Obrázok 6 Príklad jedného pixelu s rozmermi  $100 \times 100$  m s  $5 \times 5$  bodmi s 20 m rozostúpením

Obrázok 7 Chyby vynechania a priradenia pre kvantitatívne kontinuálne hodnoty od 0 po 1

Obrázok 8 Tvorba referenčnej databázy

Obrázok 9 Tvorba referenčnej databázy v prostredí programu ArcMap 10.2.2

Obrázok 10 Pixely z náhodného výberu, interpretované ako priepustné a) momentálne bez vegetácie, b) s vegetáciou, c) vodné plochy, d) dlhodobé bez vplyvu človeka – skaly

Obrázok 11 Pixely z náhodného výberu, interpretované s určitou hodnotou podielu nepriepustných povrchov

Obrázok 12 Záporná korelácia

Obrázok 13 Kladná korelácia

Obrázok 14 Ilustračný obrázok

Obrázok 15 Nadhodnotené pixely z náhodného výberu 20 000 vzoriek IMD2006

Obrázok 16 Podhodnotené pixely z náhodného výberu 20 000 vzoriek IMD2006

## **Tabuľky**

Tabuľka 1 Základné údaje o CLC

Tabuľka 2 Všeobecná podoba chybovej matice pre mapu krajinej pokrývky s q triedami

Tabuľka 3 Výsledky validácie vo forme chybovej matice na Slovensku

Tabuľka 4 Výsledky validácie vo forme chybovej matice v Maďarsku

Tabuľka 5 Chyba prijatia, chyba vynechania a parametre spoľahlivosti

Tabuľka 6 Chybová matica vytvorená producentom

Tabuľka 7 Chybová matica

Tabuľka 8 Základné údaje štátov záujmového územia validácie

Tabuľka 9 Typy pixelov

Tabuľka 10 Porovnanie počtu nadhodnotených a podhodnotených pixelov v jednotlivých verziách IMD

Tabuľka 11 Vypočítané korelačné koeficienty pre vzťahy medzi IMD a IRD\_A/B na území Slovenska

Tabuľka 12 Vypočítané korelačné koeficienty pre vzťahy medzi IMD a IRD\_A/B na území Česka

Tabuľka 13 Celková absolútna chyba normalizovaná (TAEn) pre referenčnú databázu IRD\_A

Tabuľka 14 Chyby nadhodnotenia

Tabuľka 15 Chyby podhodnotenia

## ZOZNAM SKRATIEK A ZNAČIEK

<b>AI</b>	správne klasifikované pixely ako nepriepustné
<b>AP</b>	správne klasifikované pixely ako priepustné
<b>CGLS</b>	Copernicus globálne služby monitorovania Zeme
<b>CLC</b>	CORINE Land Cover
<b>CLMS</b>	Copernicus služby monitorovania Zeme
<b>DPZ</b>	dial'kový prieskum Zeme
<b>ECMWF</b>	Európske centrum pre strednodobé predpovede počasia
<b>EEA</b>	Európska environmentálna agentúra
<b>EGMS</b>	Európska služba pozemného pohybu
<b>EHP39</b>	Európsky hospodársky priestor – 39 krajín
<b>EK</b>	Európska komisia
<b>EMSA</b>	Európska námorná bezpečnostná agentúra
<b>ERDF</b>	Európsky fond regionálneho rozvoja
<b>ES</b>	Európske spoločenstvo
<b>ESA</b>	Európska vesmírna agentúra
<b>ESM</b>	Európska mapa osídlenia
<b>EUMETSAT</b>	Európska organizácia pre využitie meteorologických družíc
<b>FRONTEX</b>	Európska agentúra pre riadenie operačnej spolupráce na vonkajších hraniciach členských štátov Európskej únie
<b>GAC</b>	poradný výbor GMES
<b>GHSL</b>	Globálna vrstva ľudského osídlenia
<b>GIO</b>	Iniciačná operačná fáza programu GMES
<b>GIS</b>	Geografický informačný systém
<b>GMES</b>	Globálny monitoring pre životné prostredie a bezpečnosť
<b>HR IMAGE</b>	rastrový produkt s vysokým rozlíšením, súvisle pokrývajúci štáty Európy
<b>HRL</b>	vrstvy s vysokým rozlíšením
<b>IHRL</b>	vrstva nepriepustnosti s vysokým rozlíšením
<b>IMD</b>	vrstva hustoty nepriepustnosti
<b>IRD_A</b>	referenčná databáza typu A
<b>IRD_B</b>	referenčná databáza typu B

<b>ISRSE33</b>	33. medzinárodná konferencia o diaľkovom snímaní životného prostredia v Strese
<b>JRC</b>	Spoločné výskumné centrum
<b>LC</b>	krajinná pokrývka
<b>LMCS</b>	Jadrové služby monitorovania krajiny
<b>LU</b>	využitie zeme
<b>MaO</b>	závažnejšie nadhodnotenie
<b>MaU</b>	závažnejšie podhodnotenie
<b>MCK</b>	Súprava na porovnávanie máp
<b>MiO</b>	menej závažné nadhodnotenie
<b>MiU</b>	menej závažné podhodnotenie
<b>MMU</b>	minimálna mapovacia jednotka
<b>MSE</b>	priemerná štvorcová odchýlka
<b>NDVI</b>	normalizovaný diferenčný vegetačný index
<b>NRC</b>	Národné referenčné centrum
<b>N2K</b>	Sústava Natura 2000
<b>SatCen</b>	Satelitné stredisko Európskej únie
<b>SAŽP</b>	Slovenská agentúra životného prostredia
<b>TAE</b>	totálna absolútna chyba
<b>TAEn</b>	totálna absolútna chyba normalizovaná

## ÚVOD

V mnohých oblastiach ľudskej činnosti zaujali veľmi dôležité postavenie čo najpresnejšie informácie o priestorovej diferenciácii rozmiestnenia obyvateľstva na našej Zemi. Vedomosť o rozmiestnení obyvateľstva je veľmi dôležitá najmä pri plánovaní využitia krajiny, napr. pri lokalizácii nových bytových jednotiek, nemocníc, nákupných stredísk atď., ale aj pri procesoch tvorby krízových scenárov, či pri procesoch zmierňovania následkov po prírodných katastrofách (GSE Land 2008). Vhodnou metódou mapového zobrazenia rozmiestnenia socioekonomických javov v krajine je využitie jednoduchých a prehľadných kartogramov, ktoré vyjadrujú hodnotu skúmanej premennej zvyčajne farbou, ktorá je priradená administratívnym jednotkám (obec, okres, kraj, ...). Kartogramy však nie sú vhodné pre analytické účely, nakoľko dochádza k domnienke, že hodnoty skúmanej premennej sú v každom bode vytýčenej oblasti rovnaké a menšie územné jednotky sa interpretujú ako vnútorne homogénne, čo však určite nezodpovedá realite.

Veľmi dobrým reprezentantom prítomnosti a priestorového rozmiestnenia ľudí sú práve zastavané plochy, na ktorých sa obyvateľstvo sústreďuje. Zaberanie pôdy je však často nekontrolované. Zástavba pôdy a z toho vyplývajúce negatívne dôsledky na životné prostredie patria k závažným problémom súčasnej doby. Pre určenie stavu zabraných plôch sa za vhodné považuje mapovanie nepriepustných povrchov.

Údaje o výskyte zastavaných plôch a o predpokladanej prítomnosti ľudí na týchto plochách sú v súčasnosti najčastejšie získavané z dát z diaľkového prieskumu Zeme (DPZ). Tieto dáta sa zbierajú v pravidelných časových intervaloch, a preto je možné na sledovaných územiach nielen pozorovať, ale aj porovnávať zmeny, ktoré tam nastali (EEA 2006). Rozsiahle územia sa spracovávajú relatívne rýchlo a zozbierané dáta majú širokospektrálne využitie.

Mapovaním využitia krajiny človekom, či zaberaniu zemského povrchu nepriepustnými plochami sa venuje veľa projektov. V rámci európskeho priestoru ide predovšetkým o aktivity pod záštitou Európskej únie, ktorej program Copernicus (Copernicus Land Monitoring Service – CLMS) je základom úsilia o pozorovanie a monitorovanie Zeme. Tento program sprístupňuje poznatky o našej planéte na úplnom, voľne dostupnom a otvorenom základe. Ide o služby poskytujúce geografické informácie o využití Zeme a iné údaje súvisiace s vegetačným stavom krajiny a stavom vodných tokov, vhodné na aplikáciu



v rôznych oblastiach, ako je územné plánovanie, lesné hospodárstvo a poľnohospodárstvo. Tieto služby sú rozdelené na globálne, celoeurópske a lokálne.

Jednou z celoeurópskych služieb programu Copernicus je vrstva nepriepustnosti s vysokým rozlíšením (Imperviousness High Resolution Layer – IHRL), skrátene tiež vrstva hustoty nepriepustnosti (Imperviousness Density – IMD). Predchádzajúcim poskytovateľom tejto služby bol program Globálne monitorovanie pre životné prostredie a bezpečnosť (Global Monitoring for Environment and Security – GMES) (EEA 2020). IMD je rastrová vrstva s rozlíšením 100 m s hodnotami pixelov 0 – 100 % reprezentujúcimi podiel nepriepustných plôch. Stupeň nepriepustnosti je vypočítaný pomocou automatického algoritmu založeného na kalibrovanom normalizovanom diferenčnom vegetačnom indexe (NDVI) (Pennec a Smith 2020). Nepriepustné plochy charakterizujeme podľa definície zo systému klasifikácie krajinej pokrývky FAO: „Nepriepustné plochy sú charakterizované nahradením pôvodnej prírodnej alebo poloprírodnej pokrývky či vodnej plochy umelým, často nepriepustným povrchom; umelosť je obvykle charakteristická dlhou trvácnosťou“ (EEA 2007, s. 2). Táto vrstva pokrýva 39 krajín Európskeho hospodárskeho priestoru a bola prvýkrát vyrobená pre rok 2006 (vtedy pokrývala 38 krajín) popri aktualizácii už desaťročia známeho projektu CORINE Land Cover (CLC), na základe rovnakých vstupných dát z DPZ ako pre CLC triedu umelých povrchov, presnejšie ako pre jej podtriedy 1.1 a 1.2. IMD vznikla ako riešenie problému s minimálnou mapovacou jednotkou 25 ha, ktorú má CLC. Na Slovensku sa to prejavilo tak, že v CLC z roku 2006 sa až v 455 obciach s malou rozlohou nenachádzali zastavané plochy. Od roku 2006 sa Európska environmentálna agentúra (EEA) v spolupráci s Európskou vesmírnou agentúrou (ESA) a Európskou komisiou (EK) snažia o jej výrobu a aktualizáciu v pravidelných trojročných intervaloch (EEA 2020). V súčasnosti je dostupná pre roky 2006, 2009, 2012, 2015 a 2018 (Pennec a Smith 2020).

IMD od svojho vzniku sľubuje do budúcnosti pozitívny vývoj. Využitie tejto vrstvy možno v súčasnosti nájsť rovnako vo fyzickej geografii ako aj v humánnej geografii. Mapovanie rozmiestnenia nepriepustných plôch sa využíva aj pri ochrane životného prostredia v environmentalistike, či v rámci poľnohospodárskej politiky pri ochrane pôd. Podľa GSE Land (2008): „Tak ako CLC aj IMD môže byť využitá ako podkladová vrstva pre dezagregáciu prakticky ľubovoľných sociálnoekonomických štatistík a pri populačnom modelovaní.“

To čo si pri využívaní IMD stanovili európske inštitúcie za cieľ, nie je až také jednoduché splniť. Vytvorením novej rastrovej vrstvy sa malo začať s presnejšími analýzami pri riešení geografických, ekonomických či sociologických problémov (EK 2015), na ktoré mali mať nepriepustné plochy zaznačené v IMD a získané informácie o nich veľký vplyv.

Viacere štáty totižto poukazujú na to, že IMD nespĺňa požiadavky kladené na tematickú presnosť. Slovenským národným referenčným centrom (NRC) ktoré spolupracuje s EEA je Slovenská agentúra životného prostredia (SAŽP). SAŽP vykonala kvalitatívne hodnotenie IMD s výsledným komentárom: „Zatiaľ na GIS aplikácie nepoužiteľný a málo vierohodný... vyskytujú sa v ňom hrubé nedostatky oproti realite a SAŽP ho neodporúčala použiť v EEA na akékoľvek analýzy, ktoré sú dôležité pre rozhodovacie procesy Európskej komisie a majú legislatívne, reštriktívne alebo finančné dopady na členské štáty,“ (Machková – Nováček 2009, p. 13). Preto sa považuje za potrebné priebežne vykonávať validácie jednotlivých vrstiev. Ich podstatou je overiť, či určitý produkt spĺňa stanovené kritériá.

Hlavným cieľom tejto dizertačnej práce je zhodnotenie tematickej presnosti Európskej vrstvy hustoty nepriepustnosti, konkrétne jej postupne vylepšených verzií pre referenčné roky 2006 a 2009, pomocou náhodnej vzorky na území Slovenka a Česka.

# 1 TEORETICKÉ VÝCHODISKÁ

V snahe rozšírenia teoretických poznatkov a priblíženia skúmanej problematiky považujeme náčrt a deskripciu pozadia vzniku priestorových dát za nevyhnutnú. Rovnako definovanie kľúčových pojmov a ich jednoznačná interpretácia býva zásadným krokom pri štúdiu javov a procesov. V prípade pojmov vzniku a využitia priestorových dát o krajinnej pokrývke sa tento, zdanlivo jednoduchý krok stáva o to zložitejším, že použitie uvedených pojmov, ako aj chápanie ich ideovej podstaty je častokrát závislé od cieľovej skupiny užívateľov. Vzhľadom na túto skutočnosť a zámer, že na základoch identity týchto pojmov bude postavená aj celá dizertačná práca, považujeme za nutné ich presné vymedzenie.

## 1.1 Pozadie vzniku priestorových dát – európska iniciatíva Copernicus

Program Copernicus je základom úsilia Európskej únie o monitorovanie Zeme a jej mnohých ekosystémov a zároveň sa ním zaisťuje, že jej občania budú pripravení na krízy a prírodné a človekom spôsobené katastrofy a chránení pred nimi. Program Copernicus, vychádzajúci zo základov hlboko zakorenených vedeckých poznatkov a desaťročí investícií EÚ do výskumu a technologického vývoja, predstavuje príklad európskej strategickej spolupráce v oblasti vesmírneho výskumu a priemyselného rozvoja (EK 2015).

Copernicus, vo svojej podstate multidisciplinárny a skutočne európsky projekt, spája komunity z geoinformačného a environmentálneho vedeckého spektra a jeho výsledkom sú prevádzkové služby od monitorovania arktického morského ľadu, cez núdzovú reakciu, až po odhaľovanie únikov ropy a monitorovanie rozširovania miest. V rámci služieb programu Copernicus sa podporuje široká škála environmentálnych a bezpečnostných aplikácií vrátane monitorovania zmeny klímy, trvalo udržateľného rozvoja, dopravy a mobility, regionálneho a miestneho plánovania, námorného dozoru, poľnohospodárstva a zdravia (EK 2015).

Služby programu Copernicus sú založené na informáciách zo špecializovanej sústavy družíc známej ako „Sentinel“, ako aj z desiatok družíc tretích strán známych ako „prispievajúce vesmírne misie“, ktoré dopĺňajú údaje z meraní „in situ“ (teda miestne alebo na mieste).

Program Copernicus na základe voľnej dostupnosti a sprístupnenia veľkej väčšiny svojich údajov, analýz, prognóz a máp prispieva k rozvoju nových inovatívnych aplikácií a služieb

prispôsobených potrebám špecifických skupín používateľov, ktoré sa dotýkajú rôznych hospodárskych a kultúrnych alebo rekreačných činností, a to od mestského plánovania, plavby a poistenia až po archeológiu.

Copernicus predstavuje začiatok novej éry v pozorovaní Zeme. Vypúšťajú sa družice Sentinel, k dispozícii sú prevádzkové služby programu Copernicus a vo forme aplikácií s pridanou hodnotou (využívajúcich výstupy programu) sa začínajú objavovať prvé náznaky „hospodárstva spojeného s programom Copernicus“ (EK 2015).

Program Európskej únie Copernicus je revolučný program pozorovania a monitorovania Zeme skúmajúci našu planétu a jej životné prostredie v konečný prospech všetkých európskych občanov. V rámci programu Copernicus sa vďaka rozmanitosti technológií – od družíc vo vesmíre po meracie systémy na zemi, v mori a v ovzduší – poskytujú prevádzkové údaje a informačné služby otvoreným a voľne dostupným spôsobom v širokom rozsahu oblastí využitia.

### **1.1.1 Štruktúra fungovania programu**

#### **A. Údaje z vesmíru**

Program Copernicus podporuje skupina špecializovaných družíc vo vlastníctve EÚ – družice Sentinel, ktoré sú osobitne navrhnuté tak, aby spĺňali potreby služieb programu Copernicus a ich používateľov. Únia vypustením prvej z nich – Sentinel-1A – v roku 2014 rozbehla proces umiestnenia sústavy viac ako dvanástich družíc na obežnú dráhu v priebehu nasledujúcich desiatich rokov. Družice Sentinel plnia potrebu konzistentného a nezávislého zdroja vysokokvalitných údajov pre služby programu Copernicus.

Program Copernicus vychádza aj z existujúcej vesmírnej infraštruktúry: družice, ktoré prevádzkuje Európska vesmírna agentúra (ESA), Európska organizácia pre využitie meteorologických družíc (EUMETSAT), členské štáty EÚ a ďalšie tretie krajiny a komerční poskytovatelia. Tieto sú známe ako prispievajúce misie a údaje z družíc pre program poskytovali už od jeho počiatku. Túto činnosť budú vykonávať aj naďalej, najmä v prípade potreby dát s veľmi vysokým rozlíšením, napríklad pri núdzových alebo bezpečnostných aplikáciách.

## **B. Údaje na zemi, v mori a vo vzduchu**

V rámci programu Copernicus sa vychádza aj z veľkého počtu meracích systémov in situ (teda miestnych alebo na mieste), ktoré na účely programu poskytujú členské štáty. Patria sem senzory umiestnené na brehoch riek, prenášané vzduchom meteorologickými balónmi, ťahané na mori loďami alebo plávajúce v oceáne. Údaje in situ sa používajú na kalibráciu, overovanie a dopĺňanie informácií z družíc, čo je dôležité na poskytovanie spoľahlivých a konzistentných údajov v priebehu času (EK 2015).

## **C. Služby v rámci programu Copernicus: premena údajov na informácie**

Službami v rámci programu Copernicus sa transformuje toto veľké množstvo údajov z družíc a údajov in situ na informácie s pridanou hodnotou, a to spracovaním a analýzou údajov, ich integráciou do iných zdrojov a potvrdením výsledkov. Dátové súbory za celé roky a desaťročia sa upravujú tak, aby boli porovnateľné a dalo sa v nich vyhľadávať, čím sa zaisťuje monitorovanie zmien; skúmajú sa modely, ktoré sa používajú na vytvorenie lepších prognóz, napríklad v súvislosti s oceánom a atmosférou. Zo snímok sa vytvárajú mapy, identifikujú sa znaky a anomálie a získavajú sa štatistické informácie.

Tieto činnosti, ktorými sa vytvára pridaná hodnota, sa zefektívňujú prostredníctvom šiestich tematických skupín služieb programu Copernicus: monitorovanie atmosféry, monitorovanie morského prostredia, monitorovanie pôdy, zmena klímy, riadenie núdzových situácií, bezpečnosť. V týchto oblastiach sa vyvinuli jasne vymedzené komunity používateľov, no služby programu Copernicus priamo aj nepriamo využíva i mnoho ďalších sektorov.

### **1.1.2 Správa a financovanie programu**

Program Copernicus predstavuje skutočne európske spoločné úsilie. Copernicus, program Európskej únie financovaný z verejných zdrojov, je predovšetkým majetkom všetkých občanov Únie, ktorí zostávajú jeho konečnými vlastníckmi a prijímateľmi.

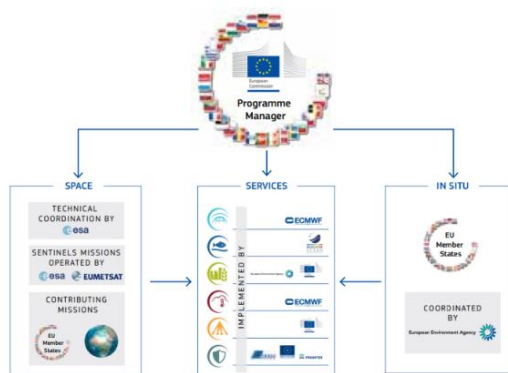
Európska komisia (EK) v partnerstve s členskými štátmi EÚ dohliada na program, koordinuje ho a zaisťuje, aby zostal zameraný na používateľov. Je zodpovedná za stanovenie a vypracovanie politickej vízie programu Copernicus a za zavedenie prvkov, ktorými sa umožní bezproblémové a riadne fungovanie systému. Spolu s Európskym parlamentom a Radou zodpovedá za dlhodobý finančný záväzok, ktorým sa zaisťuje základ pre udržateľnosť programu (EK 2015).

Členské štáty EÚ a ESA prispievajú do programu Copernicus niekoľkými spôsobmi: vývojom družíc v ESA, poskytovaním údajov z vesmírnych infraštruktúr vo vlastníctve jednotlivých štátov, poskytovaním údajov z dátových zdrojov mimo vesmíru (in situ) alebo účasťou na „pozemných segmentoch pre spoluprácu“ koordinovaných Európskou komisiou, v rámci ktorých môžu mať členské štáty priamy prístup k údajom z družíc Sentinel tým, že financujú a vyvíjajú vlastné zariadenia pozemných segmentov (prijímacie stanice, centrá spracovania a archivácie, aplikácie špecifické pre danú krajinu atď.).

Vývoj vesmírnej zložky vrátane vypustenia špecializovaných družíc Sentinel bol zverený ESA, ktorá zároveň pôsobí ako celkový architekt systémov vesmírnej zložky a zaisťuje jej technickú koordináciu. Operácie družíc Sentinel boli zverené agentúram ESA a EUMETSAT na základe ich osobitných odborných znalostí. Družice Sentinel vlastní Európska únia.

Operatívne vykonávanie programu, na ktorý dohliada Generálne riaditeľstvo Európskej komisie pre vnútorný trh, priemysel, podnikanie a malé a stredné podnikanie, vykonávajú buď priamo príslušné útvary Európskej komisie (najmä Spoločné výskumné centrum – JRC), alebo je zverené európskym subjektom a agentúram s príslušnými odbornými znalosťami.

K zainteresovaným európskym agentúram patria Európska environmentálna agentúra (EEA), Európska agentúra pre riadenie operačnej spolupráce na vonkajších hraniciach členských štátov Európskej únie (FRONTEX), Európska námorná bezpečnostná agentúra (EMSA) alebo Satelitné stredisko Európskej únie (SatCen). K ostatným zainteresovaným príslušným organizáciám patrí Európske centrum pre strednodobé predpovede počasia (ECMWF) a výskumný inštitút Mercator Ocean. Zabezpečovanie údajov z infraštruktúry in situ koordinuje Európska environmentálna agentúra (EEA).



Obrázok 1 Schéma zapojených subjektov  
(Zdroj: EK 2015)

## **1.2 Copernicus – program poskytujúci priestorové dáta o krajinnej pokrývke**

Program Copernicus pozostáva z troch komponentov: zo satelitného monitorovania Zeme, zo zberu in situ údajov a zo služieb, ktorých cieľom je poskytnúť informácie nevyhnutné na monitorovanie prostredia zemského povrchu.

Copernicus služby monitorovania Zeme poskytujú geografické informácie o využití Zeme a iné údaje súvisiace s vegetačným stavom krajiny a stavom vodných tokov, vhodné na aplikáciu v rôznych oblastiach, ako je územné plánovanie, lesné hospodárstvo a poľnohospodárstvo. Tieto služby rozdeľujeme do štyroch hlavných nasledujúcich úrovní:

- Globálne služby
- Celoeurópske služby
- Lokálne služby
- Služby poskytujúce snímky a referenčné údaje na celosvetovej a celoeurópskej úrovni

### **1.2.1 Copernicus globálne služby monitorovania Zeme**

Európsky program pre monitorovanie Zeme Copernicus, začal v roku 2013 v nadväznosti na niekoľko rokov úsilia v oblasti výskumu Zeme a v nadväznosti na program GMES – Global Monitoring for Environment and Security, ktorý vstúpil v rokoch 2011 – 2013 do iniciačnej operačnej fázy, známej pod názvom GMES Initial Opera (GIO).

Copernicus globálne služby monitorovania Zeme (CGLS) sú súčasťou Jadrových služieb monitorovania krajiny (LMCS), ktoré sú európskym vlajkovým programom pre výskum Zeme. Globálne služby systematicky poskytujú sériu kvalitatívnych biofyzikálnych produktov (každých 10 dní nové snímky) o stave a vývoji povrchu krajiny na globálnej úrovni v strednom až nízkom priestorovom rozlíšení, doplnenú o tvorbu dlhodobých časových sérií. Produkty sa používajú na monitorovanie vegetácie (napr. index plochy listov, podiel pokrytia zelenej vegetácie, index vegetačného stavu), vodných tokov (napr. aj index pôdnej vody), energetického potenciálu a pozemnej kryosféry.

Jadrové služby monitorovania krajiny (LMCS) spracovávajú dáta z výskumu Zeme s cieľom poskytnúť odborné vysoko hodnotné produkty a informácie o krajinnej pokrývke a v čo najkratšom aktuálnom čase tieto dáta poskytuje širokej skupine používateľov. Systematické pokrytie celej planéty a špecifických oblastí má vo svojej kompetencii Generálne

riadiťstvo Európskej komisie Spoločného výskumného centra (JRC). Dôkladnejším priestorovým vzorkovaním s nižšou frekvenciou aktualizácie na lokálnej a celoeurópskej úrovni sa zaoberá Európska agentúra pre životné prostredie (EEA).

Táto globálna zložka Jadrových služieb monitorovania krajiny sa označuje ako Copernicus globálne služby monitorovania Zeme a zahŕňa nasledovné:

- V prvom rade ide o systematické monitorovanie na globálnej úrovni, ktoré bežne a spoľahlivo v reálnom čase poskytuje na tejto úrovni rôzne biogeofyzikálne parametre na základe senzorov nízkeho až stredného rozlíšenia a následné zostavenie dlhodobých konzistentných časových sérií.
- Mapovanie a validáciu tzv. hot spotov v určitých regiónoch sveta, ktoré sa monitorujú na vyžiadanie, kvôli ich obmedzenému geografickému prístupu a ktoré majú nízku frekvenciu návštevnosti, ale satelitné dáta sú poskytované s vysokým rozlíšením.
- Pozemný výskum, podporujúci hlavne validáciu produktov založených na satelitných pozorovaniach, ale aj lokalizačné aplikácie a služby.

Celkový rozsah, kľúčové priority a výstavba globálnej zložky Copernicus služby monitorovania pôdy, vrátane hlavných otázok implementácie a podmienok jej trvalej udržateľnosti sú opísané v konsenzuálnom pracovnom dokumente diskutovanom počas mimoriadneho zasadnutia 33. medzinárodnej konferencie o diaľkovom snímaní životného prostredia (ISRSE33) v Strese (Taliansko) v roku 2009 a revidované delegátmi členských štátov EÚ z Poradného výboru GMES č. 14 (GMES Advisory Committee – GAC) 12. júna 2009.

Podľa tohto konsenzuálneho dokumentu by táto krajinná služba mala odpovedať na potreby aspoň týchto oblastí politiky EÚ: poľnohospodárstva a potravinovej bezpečnosti, biodiverzita, zmena podnebia, lesné a vodné zdroje, degradácia pôdy a dezertifikácia, rozvoj vidieka.

Po prijatí nariadenia Copernicus v roku 2014 sa začala plne funkčná fáza programu Copernicus, ktorá je základom pre zabezpečenie kontinuity operačnej globálnej služby monitorovania pôdy do roku 2020. Druhá výzva na implementáciu v roku 2015 zároveň predpokladá podstatné rozšírenie aktivít: väčšie produktové portfólio a spustenie ďalších komponentov služieb (napr. hot spot, in situ).



## 1.2.2 Copernicus celoeurópske služby monitorovania Zeme

### A. CORINE Land Cover

Projekt CORINE Land Cover (CLC), ktorý mapuje krajinnú pokrývku s využitím Zeme a funguje už niekoľko desaťročí, je súčasťou programu CORINE. Program „Coordination of Information on the Environment“ (CORINE) je realizovaný Európskou komisiou od roku 1985. Jeho cieľom sú výsledky troch typov (Tazberik 1994):

- „Informačný Systém stavu životného prostredia Európskeho spoločenstva (ES), ktorý pozostáva z viacerých databáz. Tento systém by mal byť následne používaný pri orientácii a aplikácii politiky ES v živom prostredí.“
- „Vývoj štandardizácie a metodík v rámci štátov ES, ale aj v nečlenských štátoch zapojených do niektorých projektov.“
- „Koncentrácia aktivít v oblasti vývoja optimálnych metodík a spôsobov zberu údajov na vzájomnom porozumení a v medzinárodnom rozsahu.“

V rámci programu CORINE bolo „jednou z hlavných úloh spomínané vytvorenie počítačovej databázy krajinného krytu CLC“ (Tazberik 1994). CORINE Land Cover je projekt mapovania povrchu krajiny Európy. Na Slovensku pripravila mapovú službu pre CORINE Slovenská agentúra životného prostredia (SAŽP). Získané údaje sú dôležité pre politické záujmy v oblasti životného prostredia, regionálneho rozvoja, poľnohospodárstva, dopravy či územného plánovania. „Program v súčasnej dobe predstavuje tiež jeden zo základných zdrojov pre tvorbu zložitejších informácií o ďalších témach, ako je erózia pôdy, emisie znečisťujúce ovzdušie a vegetáciu, atď.“ (Machková, Nováček 2009).

Inventár CORINE Land Cover sa začal v roku 1985 (referenčný rok 1990). Aktualizácie boli vytvorené v rokoch 2000, 2006, 2012 a 2018. Pozostáva zo súpisu pokrytia pôdy v 44 triedach. CLC používa minimálnu mapovaciu jednotku (MMU) 25 ha pre plošné javy a minimálnu šírku 100 metrov pre lineárne javy. Vrstvy z jednotlivých časových období sú doplnené vrstvami zmien, ktoré zdôrazňujú zmeny v pokrytí pôdy s MMU o rozlohe 5 ha. Rôzne MMU znamenajú, že zmena vrstvy má vyššie rozlíšenie ako stavová vrstva.

CLC je produkovaná vo väčšine krajín vizuálnou interpretáciou satelitných snímok s vysokým rozlíšením. V niekoľkých krajinách sa uplatňujú poloautomatické riešenia s použitím národných údajov in situ, spracovania satelitných snímok, integrácie GIS a generalizácie.

Tabuľka 1 Základné údaje o CLC  
(Zdroj: upravené podľa <https://land.copernicus.eu/pan-european/corine-land-cover/view>)

	CLC1990	CLC2000	CLC2006	CLC2012	CLC2018
<b>Satelitné dáta</b>	Landsat-5 MSS/TM jednotlivé dáta	Landsat-7 ETM jednotlivé dáta	SPOT-4/5 and IRS P6 LISS III dvojité dáta	IRS P6 LISS III and RapidEye dvojité dáta	Sentinel-2 a Landsat-8 na vyplnenie medzier
<b>Obdobie mapovania</b>	1986 – 1998	2000 +/- 1 year	2006 +/- 1 year	2011 – 2012	2017 – 2018
<b>Geometrická presnosť</b>	≤ 50 m	≤ 25 m	≤ 25 m	≤ 25 m	≤ 10 m (Sentinel-2)
<b>Min. mapovacia jednotka / šírka</b>	25 ha / 100m	25 ha / 100m	25 ha / 100m	25 ha / 100m	25 ha / 100m
<b>Geometrická presnosť, CLC</b>	100 m	lepšia ako 100 m	lepšia ako 100 m	lepšia ako 100 m	lepšia ako 100 m
<b>Tematická presnosť, CLC</b>	≥ 85 % (pravdepodobne nedosiahnutá)	≥ 85 % (dosiahnutá)	≥ 85 %	≥ 85 % (pravdepodobne dosiahnutá)	≥ 85 %
<b>Čas výroby</b>	10 rokov	4 roky	3 roky	2 roky	1,5 roka
<b>Počet zmapovaných krajín</b>	26	30	38	39	39

## B. CLC+

Európska vesmírna agentúra (EEA) začala rozvíjať a implementovať sériu produktov, tzv. sadu produktov CLC+, ktoré rozširujú a dopĺňajú v súčasnosti existujúce produkty služieb Copernicus pre monitorovanie zeme (CLMS). Dôvodom tejto produkcie je lepšie zosúladenie sa so zvyšujúcimi sa požiadavkami na monitorovanie a podávanie správ o využití pôd v Európe (Land Cover – LC, Land Use – LU). Tento proces vývoja CLC+ sa nazýva tiež „CORINE Land Cover 2. generácie“.

Koncepcia CLC+ pozostáva z troch komponentov – z kostry „CLC+ Backbone“ (detailná geometrická vektorová vrstva s 10 metrovým rozlíšením), z jadra „CLC+ Core“ (konzistentná viacúčelová sieť dát o využití zeme) a z tzv. inštancií „CLC Instances“ (na mieru šité produkty pokrytia a využitia zeme na úrovni 100 metrového rozlíšenia).

Verejná súťaž na vývoj CLC+ Backbone bola vyhlásená v roku 2019 a jej výroba začala vo februári 2020. Po verejnom obstarávaní bol v novembri 2020 odštartovaný aj vývoj a implementácia CLC+ Core, ktorý mala byť funkčný od polovice roku 2021, avšak verejne dostupný stále nie je.

### **C. Vrstvy s vysokým rozlíšením**

Vrstvy s vysokým rozlíšením (High resolution layers – HRL), ktoré majú celoeurópske pokrytie, poskytujú informácie o špecifických charakteristikách pôdneho pokrytia a sú doplnkom k mapovaniu využívania pôdy, ako napríklad v súboroch údajov CORINE Land Cover (CLC). HRL sa vyrábajú zo satelitných snímok s rozlíšením 20 m prostredníctvom kombinácie automatického spracovania a klasifikácie založenej na interaktívnych pravidlách.

Doposiaľ bolo identifikovaných 5 tried, ktoré zodpovedajú hlavným triedam z CLC, t. j. vrstva nepriepustnosti, lesy, (prírodné) trávne porasty, mokrade a trvalé vodné plochy. Tieto celoeurópske produkty budú postupne pokrývať všetky krajiny EHP39. HRL sa potom môžu použiť napríklad ako atribúty pre rôzne druhy agregovaných priestorových jednotiek, ako sú polygóny NUTS3, CLC, pravidelné siete alebo presne stanovené oblasti.

#### **1.2.2.1 Imperviousness Density – IMD**

Vrstva hustoty nepriepustnosti s vysokým rozlíšením IMD2006 bola prvou, ktorá bola vyrobená v rokoch 2006 – 2008 z multi-senzorových, bitemporálnych a orto-rectifikovaných satelitných snímok, a to popri aktualizácii CORINE Land Cover 2006. Jej výroba sa týkala 38 európskych krajín (32 členských štátov EHP a 6 krajín západného Balkánu). Na zhotovenie prvej verzie IMD bol v roku 2007 vypísaný tender EEA/IDS/07/001 v hodnote 1,2 – 2 mil. €. Predmetom kontraktu bola „produkcia konzistentnej celoeurópskej databázy zastavaných plôch s vysokým rozlíšením, obsahujúcej stupeň nepriepustnosti pre referenčný rok 2006. Vstupné dáta pre túto GMES službu – orto-rectifikované satelitné snímky budú poskytnuté z ESA“ (EEA 2007, p. 2). Účelom kontraktu bola „podpora širokej škály požiadaviek na informácie o životnom prostredí pre viaceré

oblasti politik na európskej, národnej a regionálnej úrovni, vrátane poľnohospodárstva, kvality pôd, vody, ovzdušia a mestského prostredia, ochrany prírody...“ (EEA 2007, p. 1).

Od produkcie v roku 2006 sa pre referenčné roky 2009, 2012, 2015 a nedávno 2018 vytvoril časový rad stavových vrstiev hustoty nepriepustností v celej oblasti pokrytej 39 členskými a spolupracujúcimi krajinami Európskej environmentálnej agentúry (EEA), ktoré predstavujú celkovo 6 miliónov km<sup>2</sup> (EEA 2020).

Prvé dva referenčné roky (IMD2006 a IMD2009) boli pôvodne vyrobené v rámci aktivít európskeho projektu Geoland2 a GMES. Počas výroby IMD2012, kedy sa začala súčasne vyrábať aj prvá zmenová vrstva (vrstva zachytávajúca len prírastok alebo úbytok nepriepustných plôch medzi dvoma referenčnými rokmi) medzi rokmi 2009 a 2012, sa zistilo, že je veľmi pravdepodobné, že došlo najmä k určitému nadhodnoteniu absolútnych hodnôt podielu nepriepustných plôch a taktiež k zvýšeniu nepriepustnosti v už nepriepustných oblastiach podľa rokov 2006 a 2009. Prišlo sa totižto na to, že absolútne hodnoty nepriepustnosti v roku 2009 boli v mnohých oblastiach vyššie, ako v roku 2012, hoci zmenová vrstva popisovala práve zvýšenie nepriepustnosti v týchto oblastiach (EEA 2017). To bol dôvod, prečo služba Copernicus následne prepracovávala postupne tieto referenčné roky. V rámci produkcie vrstvy hustoty nepriepustnosti k referenčnému roku 2015 (IMD2015) boli preto opätovne spracované a navzájom harmonizované stavové vrstvy a vrstvy zmien vzťahujúce sa ku všetkým predchádzajúcim referenčným rokom 2006, 2009, 2012, a teda vznikli novšie, vylepšené dostupné verzie pre jednotlivé referenčné roky (EEA 2017).

Opisu metodológie aktualizovanej verzie referenčného roka 2009 sa venoval Gangkofner a kol. (2010) a aktualizácii referenčného roka 2012 Lefebvre a kol. (2013). Podobné metódy boli aplikované aj v USA (2011) na vývoj databázy National Land Cover (Xian a kol. 2011).

Najnovšie špecifikácie opravených verzií rádu vrstiev hustoty nepriepustnosti ponúka Tobias Langanke a kol. (2016), ktorý popisuje zlepšenia berúc do úvahy skúsenosti z predchádzajúcich výsledkov overenia produktov, týkajúce sa zabezpečenia správnej klasifikácie priepustných a nepriepustných povrchov.

Zlepšenia sa týkajú hlavne opravy týchto často vyskytujúcich sa chýb (Langanke a kol. 2016, s. 7):

- Povrchy zodpovedajúce triedam CLC 1.3.x (bane, lomy, skládky a staveniská) sa nesprávne považovali za nepriepustné povrchy.
- Skleníky a povrchy pokryté skleníkovými komplexami sa majú klasifikovať ako nepriepustný povrch.
- Zelené strechy by sa nemali klasifikovať ako nepriepustné povrchy napriek tomu, že ide o veľmi utesnený povrch.
- Pláže, piesočnaté duny (CLC 3.3.1), holé skaly (CLC 3.3.2) a územia s riedkou vegetáciou (CLC 3.3.3) sú veľmi často nesprávne klasifikované. Nepatria k nepriepustným povrchom.
- Letiská a prístavy sú ďalším častým zdrojom nesprávnej klasifikácie a mali by sa opraviť, nakoľko dochádza k častým chybám priradenia alebo naopak vynechania.

Pri úprave vrstiev hustoty nepriepustnosti bola navyše použitá aj metóda vyplňania medzier, ktorou sa „malo zabrániť medzerám vo vrstvách spôsobených chýbajúcimi vstupnými údajmi pri pozorovaní Zeme, napr. prítomnosťou oblakov nad daným terénom v čase snímkovania“ (EEA, 2007). Metóda spočíva v tom, že medzery, ktoré boli zachytené vo vrstve z referenčného roku 2009 sa doplnili údajmi z referenčného roku 2006 a medzery z referenčného roku 2012 sa vyplnili údajmi z referenčného roku 2009, atď. Predpokladá sa teda, že práve v týchto oblastiach môže dôjsť k určitému podhodnoteniu nárastu nepriepustných plôch v danom referenčnom roku, keďže sa využívajú údaje z predchádzajúcich referenčných rokov.

V tomto prípade je vhodné presnejšie definovať pojem referenčný rok/obdobie podľa EEA (2006): „Za referenčný rok/obdobie sa pokladá časové obdobie najviac 3 roky (rok pred referenčným rokom, referenčný rok a rok po referenčnom roku). To znamená, že údaje pozorovania Zeme použité na vytvorenie napr. referenčného roku IMD2009 sú v skutočnosti kombináciou snímkovania z rokov 2008, 2009, 2010.“

„Nepriepustné plochy sú charakterizované nahradením pôvodného (polo-prírodného) zemského povrchu alebo vodného povrchu umelým, často nepriepustným materiálom. Tieto umelé povrchy sa zvyčajne na týchto miestach nachádzajú po dlhší čas,“ (EEA 2007). Vrstva hustoty nepriepustnosti s vysokým rozlíšením „zachytáva priestorové rozloženie umelo nepriepustných oblastí vrátane úrovne nepriepustnosti na plošnú jednotku“ (EEA 2007).

Stupeň nepriepustnosti (1 – 100 %) sa vypočítava pomocou automatického algoritmu založeného na kalibrovanom normalizovanom diferenčnom vegetačnom indexe (NDVI).

Údaje o nepriepustnosti sú dostupné pre referenčné roky 2006, 2009, 2012, 2015 a 2018 obsahujú dva typy produktov (PROJECT MANAGER COPERNICUS LAND SERVICES 2020):

#### 1. Stavové vrstvy:

- Vrstvy hustoty nepriepustnosti IMD „Imperviousness Density“: Tieto vrstvy sú dostupné pre každý z 5 referenčných rokov (napr. vrstva hustoty nepriepustnosti 2009 – IMD2009). Sú k dispozícii v pôvodnom 10 m rozlíšení (IMD 2018) a v 20 m rozlíšení (IMD2006 – IMD2015).
- Vrstvy zastavanej nepriepustnosti IBU „Impervious Built-up“: Tento produkt zobrazuje zastavané plochy, časť nepriepustných plôch, kde sa nachádzajú budovy. Zastavané plochy sú teda podskupinou nepriepustných plôch. Týka sa to oblastí, kde sa nachádzajú nadzemné stavby. Na rozdiel od nepriepustnosti charakterizovanej dlhou trvácnosťou je IBU2018 binárnym produktom vyjadrujúcim zastavané alebo nezastavané oblasti. Tento produkt je k dispozícii v rozlíšení 10 metrov, rovnako ako 100 metrov agregovaná verzia s názvom „Share of Built-up“ (SBU).

#### 2. Vrstvy zmien nepriepustnosti:

Pre každé trojročné obdobie medzi 5 referenčnými rokmi (2006 – 2009, 2009 – 2012, 2012 – 2015, 2015 – 2018) sú k dispozícii dva typy vrstiev zmien nepriepustnosti:

- Jednoduché vrstvy mapujúce percentuálne zvýšenie alebo zníženie nepriepustnosti pre tie pixely, ktoré vykazujú skutočnú zmenu nepriepustnosti v časovom období. Tieto produkty sú k dispozícii v 20 m a 100 m rozlíšení.
- Vrstvy s klasifikovanou zmenou, ktoré mapujú najrelevantnejšie kategórie zmien nepriepustnosti (nezmenená priepustnosť, nové pokrytie nepriepustnosti, strata pokrytia nepriepustnosti, nezmenená nepriepustnosť, zvýšený podiel nepriepustnosti, znížený podiel nepriepustnosti). Tento produkt je k dispozícii iba v 20 m rozlíšení.

### **D. Služby monitorujúce biofyzikálne parametre**

Systematické sledovanie biofyzikálnych parametrov na zemskom povrchu vyúsťuje do tvorby biogeofyzikálnych produktov o stave a vývoji zemského povrchu. Tieto produkty

s vysokým priestorovým rozlíšením sa používajú na monitorovanie dynamiky vegetácie, vodného cyklu, energetického potenciálu a premenných zemskej kryosféry. Vyrábajú sa v celoeurópskom rozsahu. V súčasnosti sú implementované dva produkty, „Fenológia a produktivita vegetácie s vysokým rozlíšením“ a „Monitorovanie snehu a ľadu s vysokým rozlíšením“.

## **E. Európska služba pozemného pohybu**

Európska služba pozemného pohybu (European Ground Motion Service – EGMS) poskytuje konzistentné, pravidelné a spoľahlivé informácie o prirodzenom a antropogénnom pohybe na zemi. Služba monitoruje všetky štáty patriace do programu Copernicus. EGMS predstavuje základ pre aplikácie pozemného pohybu na kontinentálnej, národnej a miestnej úrovni. Služba sa považuje za bezprecedentnú príležitosť na štúdium dopadov ľudskej činnosti na zemský povrch, ako sú pomaly sa pohybujúce zosuvy pôdy, poklesy pôdy v dôsledku využívania podzemnej vody alebo podzemnej ťažby, sopečné nepokoje a mnohé ďalšie. EGMS je tiež základom pre skúmanie pohybu zeme ovplyvňujúceho budovy a lineárne infraštruktúry.

V roku 2016 bola vytvorená pracovná skupina, ktorá na základe potrieb a požiadaviek budúcich používateľov určila kľúčové špecifikácie pre túto európsku službu pozemného pohybu.

V roku 2019 konzorcium EEA definovalo parametre použité na vymedzenie technickej implementácie EGMS v dokumente: „Európska služba pozemného pohybu: Plán implementácie a špecifikácie produktu“. Ďalší dokument „Interakcia a spolupráca medzi EGMS a inými operačnými globálnymi monitorovacími službami“ popisuje charakter už existujúcich vnútroštátnych služieb a poskytuje predbežný a počiatkový plán možnej spolupráce medzi európskymi a vnútroštátnymi službami (Alessandro Ferretti a kol. 2021).

Produkty EGMS budú prístupné pomocou špeciálneho prehliadača. Výroba produktov za referenčné obdobie 2015 – 2020 je stále v procese. Validáciu by malo vykonať nezávislé konzorcium do konca októbra 2022. Následne sa plánuje prvá aktualizácia produktu a opäť následná validácia. Prístup k údajom na stiahnutie by mal byť dostupný koncom roka 2022 (Capes a Passera 2022).

## **F. Príbuzné celoeurópske produkty**

Do tejto poslednej skupiny celoeurópskych produktov patrí Európska mapa osídlenia (European Settlement Map – ESM) a pilotné projekty CORINE Land Cover v krajinách Východného partnerstva.

Európska mapa osídlenia je rastrová vrstva, ktorá mapuje ľudské sídla na základe satelitných snímok SPOT5 a SPOT6. Bola vytvorená pre referenčné obdobia rokov 2012 a 2015. ESM2016 (označovaná tiež ako Global Human Settlement Layer – GHSL) predstavuje percentuálny podiel pokrytia zastavanej plochy na priestorovú jednotku. Dostupná je v 2,5 m a 10 m rozlíšení ako poloprodukt, v 100 m rozlíšení ako výsledný produkt (Kemper a kol. 2021).

### **1.2.3 Copernicus lokálne služby monitorovania Zeme**

Lokálne služby koordinuje Európska agentúra pre životné prostredie a ich cieľom je poskytnúť špecifické a podrobnejšie informácie, ktoré dopĺňajú informácie získané prostredníctvom celoeurópskych služieb. Tieto služby sa zameriavajú na rôzne hot spoty, t. j. oblasti, ktorých sa týkajú špecifické environmentálne výzvy a problémy, ktorých riešenie bude založené na snímkach s veľmi vysokým rozlíšením (2,5 × 2,5 m pixely) v kombinácii s inými dostupnými dátovými súbormi (obrázky s vysokým a stredným rozlíšením) v celoeurópskom priestore. Medzi tieto lokálne služby patrí: Urban Atlas, Pobrežné zóny, Natura 2000.

Urban Atlas obsahuje komplexné informácie o skutočnom „land cover“ a „land use“ (teda o krajinej pokrývke a využití krajiny) na území veľkých európskych miest a ich okolia. Tieto dáta predstavujú cenný materiál nielen pre zhodnotenie a kontrolu aktuálneho rozvoja mestských aglomerácií podľa územného plánu, ale môžu slúžiť ako podklad pre zhodnotenie rizík a príležitostí na území, od hrozieb záplav až po identifikáciu nových infraštruktúrnych potrieb. Značnou výhodou je tiež homogenita dát, s akou sú záujmové územia pokryté. Pri porovnaní dát z niekoľkých časových horizontov je možné vyhodnotiť dynamiku rastu aglomerácií a v závislosti od nej tiež lepšie skúmať faktory, ktoré rast ovplyvňujú. Z hľadiska dopravy má Urban Atlas význam v monitoringu výstavby budov, logistických centier alebo nových priemyselných areálov, ktoré sú na dopravnú infraštruktúru úzko naviazané a z ich vzniku možno odhadovať nielen zaťaženie existujúcej dopravnej infraštruktúry, ale zároveň plánovať jej optimalizáciu, či prípadné rozšírenie (Schmitt a kol. 2020).



Projekt Urban Atlas je financovaný Európskou komisiou predovšetkým z Európskeho fondu regionálneho rozvoja (ERDF). V roku 2010 boli v atlase spracované územia 117 európskych miest s minimálnym počtom 100 000 obyvateľov. Do roku 2011 bol však zoznam rozšírený na viac ako 300 miest, pričom aktualizácia je plánovaná každé 1 – 3 roky. Jednou z podmienok pre zaradenie mesta do atlasu bola tiež účasť v európskom projekte Urban Audit, v rámci ktorého sa porovnávali vybrané európske mestá na základe výskytu rôznych indikátorov a spracovávaní dát. Atlas bol vyhotovený v mierke 1:10 000. Základom pre vytvorenie atlasu sa stali snímky z družice Spot 5, s priestorovým rozlíšením 2,5 m. Minimálna výmera zobrazenej plochy je v závislosti na mierke 0,25 ha (EK 2015).

Ďalšia lokálna služba sa zaoberá pôdnym pokrytím a využívaním pôdy v oblastiach pozdĺž riek, t. j. pobrežných zón. Dôvod pre vznik tejto lokálnej služby bola potreba monitorovať biodiverzitu na európskej úrovni, okrem iného v rámci zlepšenia "zelenej" a "modrej" infraštruktúry v Európskej únii.

Oblasti Natura 2000 (N2K) sú tiež dôležitými bodmi, ktoré treba brať do úvahy. Cieľom prvého projektu N2K bolo posúdiť, či sa lokality Natura 2000 účinne zachovali a či je zastavený pokles určitých typov biotopov trávnatých plôch.

#### **1.2.4 Služby poskytujúce snímky a referenčné údaje na celosvetovej a celoeurópskej úrovni**

K službám poskytujúcim snímky a referenčné údaje na celoeurópskej úrovni patrí služba IMAGE (High Resolution Image Mosaics – HR IMAGE) je rastrový produkt súvisle pokrývajúci štáty Európy. Sú to satelitné snímky s vysokým rozlíšením, ktoré vytvárajú vstupy pre tvorbu mnohých informačných produktov a služieb, ako sú mapy využitia krajiny alebo vrstvy s vysokým rozlíšením podľa charakteru krajiny. Keďže všetky satelitné snímky sú k dispozícii pre 39 krajín Európskeho hospodárskeho priestoru (EHP39) vznikla celoeurópska orto-rektifikovaná mozaika.

Od roku 2000 Európska komisia (EK) poverila Európsku vesmírnu agentúru (ESA) v pravidelných intervaloch systematicky poskytovať aktuálne satelitné obrazy na celoeurópskej úrovni. Táto časová séria obrazových mozaik s vysokým rozlíšením poskytuje základ pre podporu a monitorovanie politik Európskeho spoločenstva (ES).

Vytváranie takýchto mozaiek zahŕňa rôzne kroky spracovania, ako napr. radiometrické vyhladzovanie snímok z rôznych dátumov a snímačov, geometrická korekcia na ortorektifikované produkty a minimalizácia oblačnosti. Podľa Müller a kol. (2007) „kombináciou uvedených satelitných snímok zo satelitov majú byť vytvorené dve celoeurópske konzistentné pokrytia k danému referenčnému roku (+/- jeden rok). Obidve pokrytia majú byť nasnímané vo vegetačnom období, mať maximálne 5 % pokrytie oblakmi na celoeurópskej úrovni a časový rozdiel dátumov snímkovania by na žiadnom mieste nemal byť menší ako 6 týždňov. Výsledná európska mozaika je dodávaná v štandardnej európskej projekcii (Lambertovo azimutálne rovnakoplošné zobrazenie). Podmienka na geometrickú presnosť danej mozaiky je, aby RMSE (odmocnina priemeru štvorcov chýb) chyba bola menšia ako 20 m v každom smere v porovnaní s mozaikou predošlého obdobia, prípadne s iným referenčným podkladom“.

V súčasnosti sú vyrobené vrstvy s vysokým 20 m rozlíšením (referenčné roky 1990, 2000, 2006, 2009, 2012, 2015) a vrstvy s veľmi vysokým 2,5 – 2 m rozlíšením (referenčné obdobia rokov 2012 a 2018). Na vytvorenie týchto IMAGE produktov boli používané snímky zo satelitov. Ako prvý, na tvorbu IMAGE1990, bol použitý satelit Landsat MSS a TM. Snímky zo satelitu Landsat 7 ETM+ 12.5/25m sa používali pri tvorbe IMAGE2000. Pri tvorbe rastrového produktu IMAGE2006 boli použité „snímky s vysokým rozlíšením z francúzskeho satelitu SPOT4 (rozlíšenie 20m), SPOT5 (rozlíšenie 10m) a indického satelitu IRS-P6 (rozlíšenie 23,5 m)“ (Müller a kol. 2007). Verzia IMAGE2009 vznikla zo snímok satelitov SPOT4-5, IRS-P6 LISS-3. Na výrobu IMAGE2012, boli použité snímky satelitu IRS-RapidEye (rozlíšenie viac ako 6,5m), SPOT-5 (2,5m rozlíšenie), SPOT-6 (1,5 m rozlíšenie), FORMOSAT-2 (2 m rozlíšenie). IMAGE2015 a IMAGE2018 vznikli na základe snímok zo satelitov Resourcesat-1/2, SPOT-5, Sentinel-2 MSI.

### **1.3 Súčasný stav riešenej problematiky doma a v zahraničí**

Európska vrstva hustoty nepriepustnosti (Imperviousness Density – IMD) bola pôvodne na začiatku tvorby označovaná ako Európska mapa stupňa nepriepustnosti, celým názvom „Fast Track Service Precursor on Land Monitoring – Degree of soil sealing“. Je to prvá dátová vrstva o krajinskej pokrývke (land cover – LC) s vysokým rozlíšením, ktorá pokrýva celú Európu. Veľkosť pixelov vo finálnej verzii je 100 × 100 m a v každom pixeli je vyjadrený stupeň nepriepustnosti povrchu hodnotou 0 – 100 %.

V roku 2006 EEA vypísala prvý tender, v ktorom boli špecifikované požiadavky na výrobu IMD. Použitá bola definícia pojmu zastavaných plôch zo systému FAO klasifikácie krajinej pokrývky: „Zastavané plochy sú definované ako nahradenie pôvodnej prírodnej alebo poloprírodnej pokrývky zemskeho či vodného povrchu umelou pokrývkou, ktorá je často nepriepustná. Táto zmena na umelé nepriepustné plochy je obvyčajne dlhodobá,“ (FAO 2005). Aj napriek tejto definícii sa v tomto dokumente tiež uvádza, že pre účely tendra, resp. pre účely validácie, možno za zastavanú plochu považovať územie s podielom nepriepustných plôch 80 – 100 % (stupeň nepriepustnosti), a tým sa zavádza pojem „threshold“ (prahová hodnota), prostredníctvom ktorého sa metrické (pomerové) hodnoty nepriepustných plôch 0 – 100 % transformujú na hodnoty dichotomické – zastavané a nezastavané. Maucha a kol. (2010) neskôr navrhujú zmeniť tento „threshold“ z 80 % na 30 %. Vysvetľujú, že „validácia s „thresholdom“ 30 % (na rozdiel od validácie s pôvodne navrhovaným „thresholdom“ 80 %) vedie ku chybám vynechania a priradenia menším ako 15 %“, čo bola požiadavka definovaná v tendri.

Používatelia, takto podľa požiadaviek tendra vyrábanej vrstvy IMD, sa však nestotožňovali s na začiatku použitým termínom „soil sealing“, nakoľko už dávno pred vypísaním tendra definovala EEA tento pojem širšie: „Pod termínom „soil sealing“ (znepriepustnenie, zapečatenie pôdy) sa rozumie zmenenie vlastností pôdy, ktorá sa tým mení na nepriepustné médium, napr. zhutnením poľnohospodárskou mechanizáciou. Zároveň sa však tento istý pojem používa na popísanie prekrytia pôdneho povrchu nepriepustnými materiálmi ako je betón, kov, sklo, asphalt, makadam,“ (EEA 2015). Pri tvorbe IMD sa pod pojmom „soil sealing“ chápe iba prekrytie pôdneho povrchu nepriepustnými materiálmi, tak ako to popisuje druhá časť vyššie uvedenej definície. Zo strany samotných používateľov teda vzišiel návrh nahradiť tento pojem „soil sealing“ pojmom „imperviousness“ (nepriepustnosť). Následne bola používateľmi jednotlivých členských štátov EEA vyjadrená aj potreba presnejšej definície toho, čo IMD presne mapuje. Žiaľ ďalší tender vypísaný v roku 2009 z dôvodu vylepšenia pôvodnej IMD z roku 2006 reagoval len na prvú z týchto dvoch požiadaviek a nahrádza pojem „soil sealing“ pojmom „impreviuousness“ (EEA 2009). Nejasnosť obsahu nového používaného pojmu „impreviuousness“, teda konkrétne toho, aké plochy tento pojem zahŕňa, však odstránená nebola.

### 1.3.1 Význam a zdôvodnenie potreby validácie

Podstatou validácie je overenie, či daný produkt spĺňa určité kritériá. Validácia by sa mala na rozdiel od verifikácie vykonávať až po ukončení procesu produkcie a viesť k prijatiu alebo zamietnutiu databázy, ale nie k jej modifikácii. V našom prípade si pod validáciou môžeme predstaviť overenie, či tematická mapa odvodená z IMD spĺňa alebo nespĺňa kritérium 85 % presnosti (tematickej), ktorá bola stanovená v podmienkach tendra na zhotovene produktu. Maucha a Büttner (2008) konštatujú, „že iba skontrolovaním všetkých pixelov mapy a zistením, že presne 15 % z nich bolo nesprávne klasifikovaných by bolo možné preukázať, že tematická presnosť mapy je presne 85 %“. Skontrolovať však takto každý pixel by bolo finančne a aj časovo veľmi nákladné, ak nie nákladnejšie ako samotná produkcia hodnoteného produktu, a preto je nevyhnutné využívať pri validáciách štatistické metódy výberu. Ako typickú metódu hodnotenia vrstiev krajinnej pokrývky uvádza štatistický výber aj Stehman (1999): „Hodnotenie presnosti je typicky založené na výbere referenčných lokalít, v ktorých je skutočná (referenčná) klasifikácia krajinnej pokrývky porovnaná s klasifikáciou v hodnotenej mape“.

Posúdiť tematickú presnosť danej vrstvy je dôležitý krok hlavne pre budúcich užívateľov, ktorí ju budú aplikovať v rôznych oblastiach, budú prostredníctvom nej riešiť rozličné problémy, preto je vedomosť o tom, do akej miery je presná, resp. spoľahlivá, veľmi dôležitá.

Podľa Koláča (2003) „kvalita hodnôt v databáze je spravidla hodnotená na základe štatistického porovnania jednotlivých hodnôt s referenčnými údajmi, ktoré majú vyššiu kvalitu a pochádzajú z nezávislého zdroja“. Zistiť tematickú presnosť skúmaného objektu podľa Koláča (2003) „znamená uviesť pravdepodobnosť, s akou má objekt skutočne vlastnosť, ktorá je udávaná príslušným atribútom. Základné požiadavky na uskutočnenie validácie sú jasne definované kritériá“.

### 1.3.2 Definície základných pojmov

Pred overovaním tematickej presnosti dátovej vrstvy IMD a lepšie pochopenie interpretácie výsledkov jednotlivých validácií je potrebné zadefinovať nasledujúce pojmy:

**Nepriepustnosť** alebo **stupeň hustoty nepriepustnosti** je stupeň zakrytia pôdy. Odhaduje sa k ploche pixelu a je vyjadrený hodnotou od 0 % po 100 %, na základe čoho rozlišujeme kategórie nepriepustnosti. Podľa Maucha a kol. (2010) „za nepriepustné povrchy

sa považujú najmä umelé štruktúry, ako sú napr. chodníky (cesty, chodníky, dvory a parkoviská), ktoré sú pokryté nepriepustnými materiálmi, ako sú napr. asfalt, betón, tehla, kameň a strechy budov“.

Podľa FAO (2005) sú priemyselné a mestské oblasti klasifikované nasledovne na základe výskytu nepriepustných povrchov v porovnaní s priepustnými povrchmi:

- Vysoká hustota: viac ako 75 % celkovej plochy pozostáva z nepriepustného povrchu (povrchov).
- Stredná hustota: 50 % až 75 % celkovej plochy pozostáva z nepriepustného povrchu (povrchov).
- Nízka hustota: 50 % až 30 % celkovej plochy pozostáva z nepriepustného povrchu (povrchov).
- Rozptýlené: 30 % až 15 % celkovej plochy pozostáva z nepriepustného povrchu (povrchov).

**Zastavaná plocha** je „plocha charakterizovaná nahradením pôvodnej prírodnej alebo polo-prírodnej pokrývky, či vodnej plochy, umelým, často nepriepustným povrchom; nepriepustnosť je obvykle charakteristická dlhou trvácnosťou“ (EEA 2007).

**IMD hodnoty**<sup>1</sup> o percentuálnom zastúpení nepriepustnosti v jednotlivých pixeloch IMD sú dáta dostupné ku stiahnutiu (stupeň nepriepustnosti v jednotlivých pixeloch pre jednotlivé verzie IMD).

**IRD hodnoty**<sup>2</sup> o percentuálnom zastúpení nepriepustnosti v jednotlivých pixeloch sú referenčné dáta. Možno ich považovať za skutočný stupeň nepriepustnosti v príslušných pixeloch. Sú získavané vizuálnou interpretáciou leteckých ortofotosnímkov vytvorených pre stav krajinej pokrývky pre jednotlivé referenčné roky.

**Správne klasifikovaný pixel** je pixel, ktorého hodnota percentuálneho zastúpenia nepriepustnosti v IMD databáze sa rovná hodnote o percentuálnom zastúpení nepriepustnosti v IRD databáze.

---

<sup>1</sup> IMD hodnoty – spojitá metrická premenná 0 – 100 %

<sup>2</sup> IRD hodnoty – spojitá metrická premenná 0 – 100 %

**Nadhodnotený pixel** je pixel, ktorého hodnota nepriepustnosti v IRD je v skutočnosti menšia ako hodnota nepriepustnosti v IMD.

**Podhodnotený pixel** je pixel, ktorého hodnota nepriepustnosti v IRD je v skutočnosti väčšia ako hodnota nepriepustnosti v IMD.

**Celková presnosť** predstavuje „pomer sumy správne klasifikovaných pixelov k celkovému počtu pixelov“ (Rosina 2010).

Pre účely validácie sa IMD hodnoty percentuálneho vyjadrenia podielu rozlohy nepriepustných plôch v určitom pixeli modifikujú na tematickú mapu s dvoma triedami. Prvou triedou podľa EEA (2006) sú „zastavané plochy, pri ktorých sa uvádza miera nepriepustnosti  $\geq 80$  % (resp. 30 %) v danom pixeli s rozlíšením  $100 \times 100$  m“. Druhou triedou sú „nezastavané plochy s mierou nepriepustnosti  $< 80$  % (resp. 30 %). EEA (2006) považuje túto hodnotu 80 % (resp. 30 %) za odporúčaný „threshold“ (prahová hodnota).

Po transformácii spojitej metrickej premennej (IMD hodnoty, IRD hodnoty) na dichotomickú premennú (priepustné, nepriepustné<sup>3</sup>) na základe stanovenej prahovej hodnoty „thresholdu“, je možné hovoriť tiež o chybách vynechania a priradenia a o presnosti producenta a presnosti užívateľa.

**Chyba vynechania** „omission error“ je „pomer počtu pixelov, ktoré boli nesprávne vynechané z triedy, do ktorej skutočne patria“ (Congalton 1991).

**Producentova presnosť** „producer’s accuracy“ je „pomer počtu správne klasifikovaných pixelov určitej kategórie s počtom pixelov danej kategórie v referenčnej databáze, t. j. presnosť s akou bola zmapovaná daná trieda. Doplnkom k presnosti producenta je chyba vynechania“ (Congalton 1991).

**Chyba priradenia** „comission error“ je „pomer počtu pixelov, ktoré boli nesprávne zaradené do určitej triedy klasifikácie“ (Congalton 1991).

**Presnosť používateľa** „user’s accuracy“ je „pomer počtu správne klasifikovaných pixelov určitej kategórie s počtom pixelov danej kategórie v hodnotenej mape, t. j. pravdepodobnosť,

---

<sup>3</sup> V tejto dizertačnej práci navrhujeme metrickú premennú nepriepustnosť po transformácii na dichotomickú premennú, už ďalej nesprávne nenazývať podľa EEA (2006) ako zastavané a nezastavané plochy, ale ako priepustné a nepriepustné plochy (EEA 2009).

že pixel klasifikovaný v mape do určitej kategórie patrí do tejto kategórie aj v skutočnosti. Doplňkom k presnosti užívateľa je chyba prijatia“ (Congalton 1991).

Nasledujúci príklad vysvetľuje význam presnosti producenta a používateľa: „Producent mapy môže deklarovať, že 90 % plôch, ktoré sú v skutočnosti lesmi, bolo v mape správne klasifikovaných ako lesy a iba 10 % skutočných lesov bolo z triedy lesov vynechaných a nesprávne klasifikovaných do iných tried. Presnosť tejto kategórie je dostatočná. Avšak používateľ mapy môže zistiť, že z tých plôch, ktoré boli v mape klasifikované ako lesy tvoria v skutočnosti lesy iba 50 %, zvyšných 50 % tvoria iné kategórie chybné zaradené do triedy lesov. Presnosť triedy lesov je nedostatočná. Iný príklad popisuje, že ak by bolo celé skúmané územie klasifikované ako les, je isté, že všetky reálne lesy boli klasifikované správne, teda presnosť producenta je 100 %. Avšak z pohľadu užívateľa, len malá časť z území klasifikovaných ako lesy, skutočne lesy sú – preto je presnosť používateľa mizivá,“ (Rosina 2010).

Keď v roku 2006 EEA stanovila mieru požadovanej presnosti 85 % na výrobu IMD, bol použitý zrejme len ukazovateľ celkovej presnosti, čo nie je evidentne vhodné. Podľa Rosinu (2010), je „dôvodom fakt, že nepriepustné plochy, obzvlášť ak sú definované podielom nepriepustných plôch 80 % a viac, tvoria len veľmi malý podiel rozlohy štátov“. Rosina (2010) poukazuje na to, že „je potrebné pre každú triedu zvlášť vypočítať presnosť producenta a presnosť užívateľa, ktorá vypovedá o tom, do akej miery sa môže používateľ na danú klasifikáciu spoliehať“.

### **1.3.3 Niektoré prípadové štúdiá validácií doma i v zahraničí**

#### **A. Kvalitatívne hodnotenie IMD na Slovensku**

Jednou z úloh Národných referenčných centier (NRC) bolo po doručení finálneho produktu spracovať kvalitatívne hodnotenie. Pre Slovensko bolo kvalitatívne hodnotenie vykonané v réžii Slovenskej agentúry životného prostredia (SAŽP). Vykonalo sa na prvej verzii produktu. Kvalita dát bola posudzovaná v rámci 18 vybraných tried CLC, hodnotila sa na päťstupňovej verbálnej stupnici – veľmi zlá, nedostatočná, prijateľná, dobrá, excelentná. Z 18 tried bolo na Slovensku posudzovaných 13 (zvyšných 5 sa na našom území nevyskytuje). Iba v jednej z tried, aj to málo významnej (prístavy), bolo hodnotenie dobré. Pri štyroch druhoch pokrývky (krovinaté a bylinné vegetačné spoločenstvá, plochy s ojedinelou vegetáciou, holé skaly a vnútrozemské močiare) bolo hodnotenie prijateľné

a v ôsmich prípadoch bolo hodnotenie nedostatočné (urbanizované štruktúry, priemyselné a obchodné jednotky, cestná a železničná sieť, letiská, ťažba nerastov, skládky a lokality vo výstavbe, orná pôda, lesy, rôznorodé poľnohospodárske plochy). Faktom, že presnosť bola vo väčšine z kontrolovaných tried zhodnotená ako nedostatočná, zodpovedá aj celkové kvalitatívne hodnotenie. Produkt „nebol odporúčaný pre použitie na spoľahlivé modelovanie, GIS aplikácie, strategické plánovanie či reguláciu“ (Machková a kol. 2008). Autori, okrem iného, videli problém aj v spomínanej nejasnej definícii nepriepustnosti a zastavanosti, celkovej nespoľahlivosti klasifikácie v rámci hlavných tried. „Problémom môže byť prekrytie umelých objektov tenkým porastom zelených rastlín“ (zrejme mali na mysli machy, lišajníky a pod), keďže sú identifikované ako vegetácia (Machková a kol. 2008).

## **B. Validácia Maďarsku a na Slovensku podľa odporúčaní EEA**

V roku 2008 vydáva EEA odporúčania Maucha a Büttnera (2008) pre kvantitatívne hodnotenie tejto vrstvy s vysokým rozlíšením. Podľa nich má byť validovaná tematická mapa odvodená z IMD s dvoma triedami – priepustné a nepriepustné. Maucha a Büttner (2008) navrhujú pomocou stratifikovaného náhodného výberu vybrať z prvej triedy aspoň 2000 vzoriek (priepustné pixely) a 500 vzoriek (nepriepustné pixely) z druhej triedy. Takéto rozdelenie vysvetľujú tým, že „je potrebné zabezpečiť, aby bola početnosť výberu dostatočná aj v triedach s veľmi malou rozlohou“ (Maucha a Büttner 2008). Keďže trieda nepriepustných plôch tvorí asi iba mizivé percento rozlohy štátov (zväčša okolo 0,81 % z rozlohy štátu), je takýto prístup potrebný. Ako jednotka výberu má byť použitý jeden pixel  $100 \times 100$  m, čiže areálová jednotka. Ďalej odporúčajú:

- „porovnať hodnoty získané validáciou s hodnotami v pôvodnej databáze a vypočítať chybovú maticu,
- použiť skript pre open-source štatistický software „R“ na vypočítanie chyby priradenia, chyby vynechania a hodnoty ich spoľahlivosti,
- ak je viac ako 87 nesprávnych pixelov v rámci 500 pixelov v prvej triede klasifikácie (takže chyba prijatia je väčšia ako 15 % so spoľahlivosťou vyššou ako 95 %), presnosť hodnotenej mapy nedosiahla cieľových 85 %,
- obyčajne, chybu vynechania nie je možné vypočítať s dostatočnou spoľahlivosťou (jej miera závisí od podielu rozlohy triedy nepriepustných plôch, ktorý je spravidla veľmi nízky),



- ak sa celý interval spoľahlivosti nachádza nad 15 % hranicou (takže so spoľahlivosťou väčšou ako 95 % je chyba vynechania väčšia ako 15 %), potom presnosť hodnotenej mapy nedosiahla 85 %“ (Maucha a Büttner 2008).

Všetky tieto odporúčania je však možné aplikovať až po určení referenčných hodnôt vo vybraných vzorkách. Podľa odporúčaní EEA (2008) má byť referenčná hodnota odvodená pomocou systematického výberu 100 bodov v pravidelnej matici 10 × 10 bodov s rozstupom 10 m v rámci areálu každého z 2500 náhodne vybraných pixelov. Hodnota stupňa nepriepustnosti potom môže byť následne určená spočítaním bodov nachádzajúcich sa na nepriepustnom povrchu (na báze vizuálnej interpretácie leteckej snímky). Ako podklad pre odhadovanie referenčných hodnôt mali byť použité prednostne letecké ortofotosnímky s veľmi vysokým rozlíšením zosnímané čo najbližšie k danému referenčnému roku. Interpretácia má prebiehať „na slepo“, t. j. hodnotiaci osoba nemá poznať hodnotu pixela vo validovanej databáze, aby úsudok nebol ovplyvnený. Maucha a Büttner (2008) odporúčajú najmä v prípade nepriepustných plôch porovnávať podkladové snímky z dvoch rôznych zdrojov, príp. vyhotovené v dvoch rôznych časových obdobiach a pri prípadných zmenách použiť hodnotu z novšej snímky. Ak pri porovnaní dvoch dátumov nebudú evidentné zmeny v nepriepustnosti, avšak hodnoty sa mierne líšia (napr. v dôsledku rozdielneho zatienenia budovami či vegetáciou, alebo iným vnímaním obrazu pri rôznych rozlíšeniach), výsledná hodnota má byť určená ako ich priemer. Interpretácia má prebiehať „z pohľadu stroja“, t. j. ak sú na snímke koruny stromov, hoci povrch pod nimi je pravdepodobne nepriepustný, bod má byť posúdený ako priepustný.



Obrázok 2 Vizuálna interpretácia bodov vo vzorke 100 x 100 m  
(Zdroj: Maucha a Büttner 2008)

Výsledky validácie sa podľa Maucha a Büttnera (2008) majú následne zobrazit' v chybovej matici. „Chybová matica je zvyčajne používaná validačná schéma pri overovaní máp využitia zeme alebo pri klasifikácii satelitného obrazu,“ (Congalton a Green 1999). „Chybová matica je kontingenčná tabuľka, v ktorej bunky na diagonále obsahujú správne klasifikácie (referenčné dáta súhlasia s mapou) a bunky mimo diagonály obsahujú nesúhlasné klasifikácie,“ (Stehman – Czaplewski, 1998). Podľa Foodyho (2002) je matica (ne)zhody v súčasnosti stredobodom literatúry, ktorá sa týka hodnotenia presnosti (v tomto prípade sa uprednostňuje termín confusion matrix – matica (ne)zhody pred error matrix – chybová matica). Tematická presnosť vrstvy nepriepustnosti je teda prezentovaná vo forme chybovej matice vytvorenej z výsledkov interpretácie vzoriek a ich skutočných hodnôt v nepriepustnej vrstve. Foody (2002) ďalej konštatuje, že „z matice (ne)zhody možno odvodiť mnoho mier presnosti klasifikácie. Jednou z najčastejšie používaných mier je podiel správne klasifikovaných prípadov z celkového počtu prípadov – celková presnosť. Ak sa pozornosť venuje presnosti jednotlivých tried klasifikácie, z matice možno odvodiť podiel správne priradených jednotiek do danej triedy z celkového počtu jednotiek v danej triede“. Toto možno vykonať dvomi odlišnými spôsobmi, ktoré sme už spomínali a vysvetľovali ako presnosť používateľa a presnosť producenta, v závislosti od toho, či sú výpočty založené na riadkoch alebo stĺpcoch matice.

Tabuľka 2 Všeobecná podoba chybovej matice pre mapu krajinnej pokrývky s q triedami  
(Zdroj: upravené podľa Stehman – Czaplewski 1998)

		Referenčné dáta				
		1	2	...	q	
Hodnotená mapa	1	$p_{11}$	$p_{12}$	...	$p_{1q}$	$\sum_{j=1}^q p_{1j}$
	2	$p_{21}$	$p_{22}$	...	$p_{2q}$	$\sum_{j=1}^q p_{2j}$
	...	...	...	...	...	...
	q	$p_{q1}$	$p_{q2}$	...	$p_{qq}$	$\sum_{j=1}^q p_{qj}$
		$\sum_{i=1}^q p_{i1}$	$\sum_{i=1}^q p_{i2}$	...	$\sum_{i=1}^q p_{iq}$	

Sériu opisných techník vychádzajúcich z chybovej matice podrobne opisuje aj Congalton (1991). Už zmienená celková presnosť sa podľa neho javí ako najjednoduchšia metóda. Je možné ju vypočítať ako pomer sumy hodnôt na hlavnej diagonále k celkovému počtu jednotiek. Obdobným spôsobom je možné vypočítať aj presnosť za jednotlivé kategórie tak, že počet správnych pixelov v určitej kategórii možno vydeliť buď sumou príslušného riadka

alebo stĺpca. Ak vydelíme počet správne klasifikovaných pixelov v určitej kategórii počtom pixelov tej istej kategórii v referenčnej databáze (teda sumou stĺpca), výsledná miera je spomínaná presnosť producenta a dá sa interpretovať ako presnosť, s ktorou bola daná trieda zmapovaná (je to pravdepodobnosť, že bod bol zaradený do správnej triedy). Nezabudnime, že doplnkom ku presnosti producenta je chyba vynechania. Tá je pomerom počtu pixelov, ktoré boli nesprávne vynechané z triedy, do ktorej v skutočnosti patria. Naopak, ak počet správne klasifikovaných pixelov v danej triede vydelíme počtom pixelov v príslušnej kategórii v hodnotenej mape (teda sumou riadka), výsledkom bude presnosť používateľa. Tá predstavuje pravdepodobnosť, že pixel, ktorý bol v mape klasifikovaný do určitej triedy, bude aj v skutočnosti patriť do tejto triedy. Pripomíname, že doplnkom k presnosti používateľa je už skôr popísaná chyba prijatia, t. j. podiel pixelov, ktoré sú nesprávne zaradené do určitej triedy klasifikácie.

Na základe týchto odporúčaní prebehli následne dve obdobné validácie. Prvá bola vykonaná v Maďarsku v roku 2008. Výsledky tejto validácie interpretoval Maucha a Petrik (2008). Druhá validácia na rovnakej báze náhodného výberu 2 500 vzoriek prebehla v roku 2010 a vykonal ju Rosina (2010). V oboch prípadoch bol na transformáciu vrstvy IMD2006 na tematickú mapu s dvoma triedami – priepustné a nepriepustné, použitý pre potreby týchto validácií odporúčaný „threshold“ (prahová hodnota) 80 %.

Výsledky oboch validácií sú prehľadne zobrazené v nasledujúcich chybových maticiach. Je vhodné si všimnúť, že sa pri porovnaní pôvodnej a referenčnej databázy podľa Rosinu (2010) na Slovensku nenašla ani jedna chyba vynechania, teda všetky pixely klasifikované ako priepustné boli priepustné aj v referenčnej databáze. Na druhej strane, počet pixelov chybné prijatých do triedy nepriepustných plôch bol dokonca vyšší ako počet tých, ktoré boli klasifikované správne. Z toho vyplýva „veľmi nízka hodnota presnosti používateľa pre triedu nepriepustných plôch – 22,40 %. Presnosť producenta pre túto triedu je naopak 100 %, pretože všetky reálne nepriepustné pixely boli vo vrstve klasifikované ako nepriepustné. Počet chybné prijatých pixelov do triedy nepriepustných plôch je tak vysoký (388), že ovplyvnil natol'ko celkovú presnosť i presnosť producenta pre triedu nepriepustných plôch, že sú tesne pod vyžadovaným 85 % limitom presnosti,“ (Rosina 2010). Čo sa týka validácie v Maďarsku (Maucha a Petrik 2008), počas overovacieho procesu sa ukázalo, že v prípade 11 vzoriek z celkového počtu 2 500 neboli k dispozícii referenčné údaje. Takže presne 498 vzoriek bolo vybratých pre triedu nepriepustných, 1991 bolo vybratých pre triedu priepustných povrchov. Pri porovnaní

výsledkov oboch matíc je zjavne jasné, že hoci výsledky validácie sú na prvý pohľad veľmi podobné, predsa len validácia v Maďarsku bola vykonaná s vyššou celkovou presnosťou (o približne 4 % lepši výsledok) a vo veľkej miere je na tom lepšie aj presnosť používateľa, ktorá je vyššia takmer dvojnásobne.

Tabuľka 3 Výsledky validácie vo forme chybovej matice na Slovensku  
(Zdroj: upravené podľa Rosina 2010)

	Triedy	Referenčné dáta		Suma	Presnosť používateľa
		Nepriepustné plochy	Priepustné plochy		
<b>Hodnotená vrstva</b>	Nepriepustné plochy	112	388	500	22,40 %
	Priepustné plochy	0	2 000	2 000	100 %
	Suma	112	2 388	Celkový počet vzoriek: 2500	
	Presnosť producenta	100%	83,75 %	Celková presnosť 84,48 %	

Tabuľka 4 Výsledky validácie vo forme chybovej matice v Maďarsku  
(Zdroj: upravené podľa Maucha – Petřík 2008)

	Triedy	Referenčné dáta		Suma	Presnosť používateľa
		Nepriepustné plochy	Priepustné plochy		
<b>Hodnotená vrstva</b>	Nepriepustné plochy	222	278	498	44,58 %
	Priepustné plochy	6	1 985	1 991	99,69 %
	Suma	228	2 269	Celkový počet vzoriek: 2489	
	Presnosť producenta	97,37 %	87,48 %	Celková presnosť 88,67 %	

Treba však poznamenať, že kým v Maďarsku prebehla validácia úplne prvej vrstvy IMD z referenčného roku 2006, môžeme ju označiť aj IMD2006\_v1 (verzia jedna), tak Rosina (2010) validoval tú istú vrstvu z roku 2006, ale konkrétne, už jej vylepšenú verziu, môžeme ju označiť ako IMD2006\_v2 (verzia dva) .

Tender EEA/SES/09/003 za účelom zdokonalenia pôvodnej IMD2006\_v1 vyžadoval (EEA 2009):

- Odstrániť oblasti ťažby, skládok a výstavby – triedy CLC 1.3. Tieto triedy boli v prvej verzii priradené k vrstve nepriepustných plôch. V druhej verzii mali byť považované za priepustné plochy. Odstránil sa ich funkčný pôdorys, no budovy, ktoré pripadali na tieto oblasti sa zachovali.
- Opraviť/vyplniť medzery vyskytujúce sa osídlení. V prvej verzii IMD2006 sa často vyskytovali problémy s vynechaním sídelnej zástavby. Hlavne rozptýlené oblasti, ktoré sa nachádzajú aj na území Slovenska, sa javili ako najproblematickejšie oblasti. Stalo sa, že nie jedna obec na mape neobsahovala žiadnu nepriepustnú plochu.
- Opraviť nesprávne klasifikované oblasti pláží, piesočnatých dún, holých skál, ktoré boli klasifikované ako nepriepustné plochy.
- Opraviť nesprávne klasifikované letiská a prístavy.

Produkcia tzv. druhej, vylepšenej, či aktualizovanej verzie IMD2006\_v2 bola ukončená v decembri 2009. Výsledný produkt bol zverejnený 26. januára 2010 cez dátovú službu EEA.

Podľa odporúčaní Mauchu a Bütnera (2008) posledným krokom odmietnutia alebo prijatia IMD pre vybrané územie je výpočet parametrov spoľahlivosti (hoci zamietnutie databázy bolo takmer isté, vzhľadom na vysoký podiel chybných klasifikácií, ktoré boli zistené už v samotnom výbere). Na tento účel sa použil skript napísaný v jazyku pre open-source – štatistický softvér „R“, ktorého kód je prílohou odporúčaní Mauchu a Bütnera (2008) pre kvantitatívne hodnotenie vrstiev s vysokým rozlíšením. Za vstupné parametre sa mali použiť 4 hodnoty z chybovej matice, podiel pixelov IMD v triede nepriepustných plôch (pre Slovensko 0,81 %), podiel územia vylúčeného z výberu (pre Slovensko 0 %) a hladina štatistickej významnosti ( $\alpha=0,05$ ). Výstupné parametre boli dolná ( $p_0$ ) a horná ( $p_1$ ) hranica intervalu spoľahlivosti, ich priemerná chyba (stredná hodnota), polovica šírky intervalu spoľahlivosti (spoľahlivosť priemernej chyby) a pravdepodobnosť, že IMD mala reálne chybu viac ako 15 %. Výstupné parametre mali byť vypočítané osobitne pre chybu vynechania a pre chybu prijatia. Ako názorný príklad uvádzame v nasledujúcej tabuľke výsledky pre územie Slovenska podľa validácie Rosinu (2010) aj s verbálnou interpretáciou.

Tabuľka 5 Chyba prijatia, chyba vynechania a parametre spoľahlivosti  
(Zdroj: Rosina 2010)

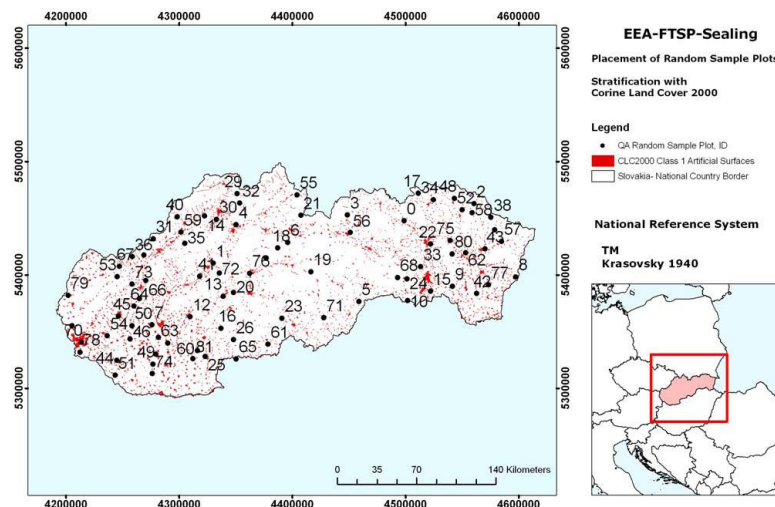
Druh chyby	Početnosť výberu	Počet chybných jednotiek	$p_0$ – dolná hranica intervalu spoľahlivosti	$p_1$ – horná hranica intervalu spoľahlivosti	Priemerná chyba	Interval spoľahlivosti	Pravdepodobnosť skutočnej chyby väčšej ako 15%
Chyba prijatia	500	388	74,52 %	80,64 %	77,58 %	± 3,06 %	100,00 %
Chyba vynechania	2000	0	0,00 %	18,31 %	9,15 %	± 9,15 %	8,55 %

- Trieda nepriepustných plôch mala s 95 % pravdepodobnosťou chybu prijatia väčšiu ako 74,52 % a menšiu ako 80,64 %. Priemerná chyba prijatia bola 77,58 %. Chyba prijatia bola väčšia ako 15 % s pravdepodobnosťou 100 %.
- Trieda nepriepustných plôch mala s 95 % pravdepodobnosťou chybu vynechania väčšiu ako 0 % a menšiu ako 18,31 %. Priemerná chyba prijatia bola 9,15 %. Chyba vynechania bola väčšia ako 15 % s pravdepodobnosťou 8,55 %.

Nakoľko stačilo, aby aspoň jedna z vyššie uvedených chýb presahovala hranicu 15 % s dostatočnou spoľahlivosťou, IMD2006\_v2 na území Slovenska nespĺňa požadované kritériá a nemôže byť akceptovaná (Rosina 2010).

### C. Hodnotenie presnosti od dodávateľa produktu

Spolu s produkciou IMD sú vyhotovované aj správy o doručení produktu pre každý štát jednotlivo. Tieto správy zhotovuje vždy firma, ktorá spracováva územie daného štátu. Územie Slovenska (IMD2006) monitorovala v tom čase firma GISAT, s.r.o., ktorá následne podala aj správu o doručení produktu. Tá obsahovala technickú špecifikáciu dodaných dát, komentovala metodiku tvorby a hodnotila presnosť. Podľa veľkosti územia štátu bolo spracované hodnotenie na základe výberu pixelov, ktorý bol náhodný. Na Slovensku bolo vybraných len 82 pixelov, z čoho bola polovica zámerne vybratá z triedy umelých povrchov CLC. Pomocou Google Earth boli zistené referenčné hodnoty. Následne sa vytvorila chybová matica priradením pôvodných a referenčných hodnôt do tried priepustných a nepriepustných povrchov. Podľa odporúčania Maucha a Büttnera bol použitý „threshold“ 80 %. Vypočítala sa presnosť producenta a používateľa, a celková presnosť. V konečnom dôsledku bolo na Slovensku zohľadnených len 73 pixelov, nakoľko sa vynechali tie, ktoré boli na území nespĺňajúcom podmienky satelitných snímok použitých pri tvorbe IMAGE2006.



Obrázok 3 Rozmiestnenie náhodného výberu 73 pixelov na Slovensku  
(Zdroj: Kopecký – Kahabka 2008)

Tabuľka 6 Chybová matica vytvorená producentom  
(Zdroj: Kopecký – Kahabka 2008)

		Classification			Producer's Accuracy	Omission Error
		>80%	<80%	Σ		
Validation	>80%	2	0	2	100,0%	0,0%
	<80%	3	68	71	95,8%	4,2%
	Σ	5	68	73		
User's Accuracy		40,0%	100,0%			
Commission Error		60,0%	0,0%			
Overall Accuracy		95,9%				

Na základe vyššie uvedených údajov môžeme konštatovať, že celkový počet vzoriek použitých pri validácii producentom je len 73, čo je príliš malý počet (z čoho vyplýva nízka spoľahlivosť tohto výsledku).

Hoci bola do triedy umelých povrchov umiestnená polovica vzoriek, do triedy nepriepustných povrchov sa dostalo len 5 z nich, z ktorých boli 3 vzorky klasifikované dokonca nesprávne. Presnosť používateľa nadobudla hodnotu 40 %, no nemožno ju považovať za relevantný výsledok, pretože bola vypočítaná z malých absolútnych čísel. Už jednotkový rozdiel zmenil výsledok až o 20 %. Celkovú presnosť potom výrazne ovplyvnila druhá trieda, trieda priepustných plôch s výrazne vyšším podielom v zastúpení počtu pixelov. To sa odzrkadlilo vo výslednej hodnote až 96 %. Takéto percento samozrejme spĺňalo tendrom požadovanú podmienku celkovej presnosti vyššej ako 85 %. Produkt bol

producentom hodnotený ako prijateľný, čo možno považovať za veľmi odvážne tvrdenie, vzhľadom na veľmi malú vzorku. V celej správe o doručení nebolo dokonca ani popísané, akým spôsobom bola určená referenčná hodnota jednotlivých pixelov. Vytknúť je možné aj absenciu výpočtov parametrov spoľahlivosti.

#### **D. Validácie referenčných lokalít s rozmermi 6 × 6 km**

V roku 2010 sa Hurbánek a kol. zaoberali validáciou IMD2006\_v2 na základe cieľného výberu referenčných lokalít s rozmermi 6 × 6 km. Šlo o prípadové štúdie od najmenších roztratených sídel a malých kompaktných vidieckych obcí, cez dlhé uličné obce a veľké vidiecke obce, až po menšie a väčšie mestské sídla, ktoré reprezentovali vybrané typy štruktúr osídlenia Slovenska. Validované bolo rozptýlené osídlenie pri meste Myjava, okolie obce Havranec, mesto Podolínec, Dubnica nad Váhom a jej okolie a mesto Poprad.

Tvorba referenčnej databázy bola uskutočnená v programe Q-GIS. Pomocou ortofotosnímkov bol interpretovaný skutočný podiel zastavaných plôch vo vopred nadefinovaných štvorcoch vybraných lokalít nasledovným spôsobom:

- V každom modelovom území bola vytvorená pravidelná sieť 1 440 000 bodov 5 metrovými rozstupmi nad leteckými ortofotosnímkami.
- Každému bodu bola priradená hodnota 0, 1 alebo 2. Hodnota 0<sup>4</sup> reprezentovala priepustné plochy. Hodnota 1<sup>5</sup> reprezentovala nepriepustné plochy s alochtónnym materiálom premiestneným na dané miesto aktivitou človeka a hodnota 2 reprezentovala nepriepustné plochy s autochtónnym materiálom.

---

<sup>4</sup> 0 reprezentuje priepustnú plochu s vegetáciou, poprípade v danom časovom úseku priepustnú plochu bez vegetácie. Jedná sa o poľnohospodárske plochy (polia, záhrady), na ktorých bola vegetácia odstránená vplyvom človeka. Druhou skupinou sú priepustné plochy dlhodobé bez vegetácie, ktorá bola odstránená prirodzene, bez vplyvu človeka – skaly, štrkové a piesočné pláže alebo sítě, vodné plochy atď. (Lajčáková 2013).

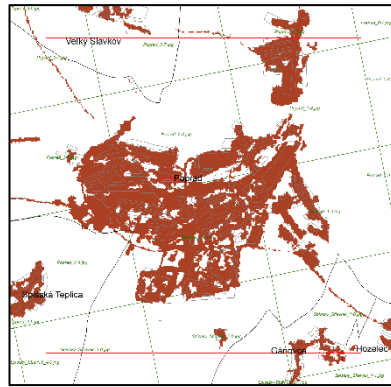
<sup>5</sup> 1 reprezentuje nepriepustnú plochu s alochtónnym materiálom, ktorý bol premiestnený na dané miesto činnosťou človeka. Patria tu územia bez vegetácie, napr. zastavané plochy – budovy, spevnené komunikácie (asfaltové, betónové, štrkové a vydlážené cesty, chodníky, železničné trate) a dvory v blízkosti hospodárskych a obytných budov, od podlažia odizolované skládky komunálneho alebo priemyselného odpadu (Lajčáková 2013).



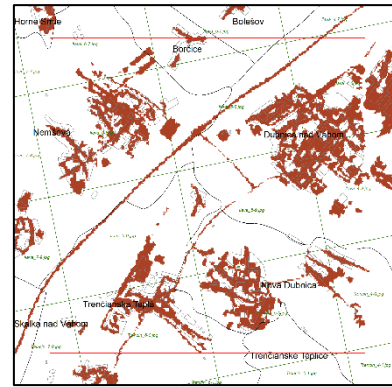
- Body s hodnotami 2<sup>6</sup> boli pripočítané k bodom s hodnotami 1 (boli považované za nepriepustné s alochtónnym materiálom). Autori plánovali neskôr body s hodnotami 2 pripočítať k bodom s hodnotami 0 (mali byť považované za priepustné) a porovnávať skúmané zmeny v tematickej presnosti.  
Takýto postup bol zvolený z dôvodu nejasnej definície pojmu nepriepustnosti, nakoľko už dávno pred vypísaním tendra, v ktorom boli špecifikované požiadavky na výrobu IMD, definovala EEA tento pojem širšie: „Pod termínom znepriepustnenie, zapečatenie pôdy sa rozumie zmenenie vlastností pôdy, ktorá sa tým mení na nepriepustné médium, napr. zhutnením poľnohospodárskou mechanizáciou. Zároveň sa však tento istý pojem používa na popísanie prekrytia pôdneho povrchu nepriepustnými materiálmi ako je napr. betón, kov, sklo, asfalt, makadam a podobne“ (EEA 2015).
- Na základe binárnej interpretácie bodov (priepustné a nepriepustné povrchy) bol vyrátaný najpresnejší podiel nepriepustných povrchov vzhľadom k celému územiu pre každé modelové územie.
- Klasifikované body boli postupne agregované do pixelov s hranou 20 m, 40 m, 60 m, 80 m, 100 m, 120 m, 200 m, 240 m, 300 m, 400 m, 500 m, 600 m, 1000 m, 1200 m, 1500 m, 2000 m, 3000 m. Pre každú agregovanú úroveň bol vyrátaný podiel zastavaných plôch v jednotlivých pixeloch.

---

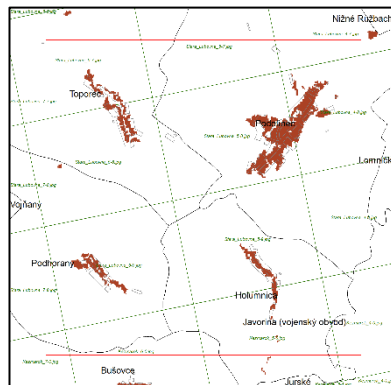
<sup>6</sup> 2 reprezentuje nepriepustnú plochu s autochtónnym materiálom, t. j. bez alochtónneho materiálu premiestneného na dané miesto činnosťou človeka. Patria tu územia bez vegetácie, odkryté opakujúcou sa činnosťou človeka (poľné a lesné cesty ujazdené a ušliapané ľuďmi alebo hospodárskymi zvieratami; dvory v blízkosti budov, plochy na skladovanie dreva pri drevospracujúcich prevádzkach). Takýto stav zemskej pokrývky sa v danej lokalite opakuje aspoň na dvoch snímkach z rôznych období (Lajčáková 2013).



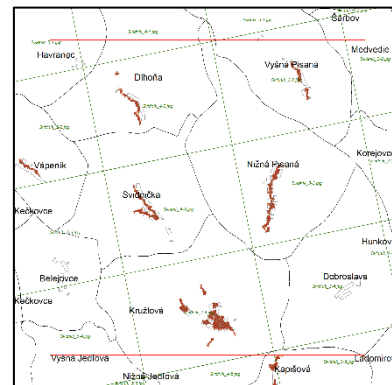
Poprad



Dubnica nad Váhom



Podolínec



Havranec

Obrázok 4 Referenčné lokality s rozmermi 6 × 6 km

(Zdroj: Lajčáková 2013)

V programe MS Excel boli následne takto klasifikované a agregované referenčné údaje o podiele zastavaných plôch spolu s údajmi o dostupných hodnotách nepriepustnosti v daných pixeloch (dostupné pre všetky agregované úrovne) spracované. Tematická presnosť produktu IMD2006\_v2 bola porovnávaná s referenčnou databázou prostredníctvom nasledujúcich kvantitatívnych metód: korelačnou analýzou dát (Pearsonov, Kendallov, Spearmanov korelačný koeficient), vypočítaním chýb (priemernej chyby, priemernej chyby nadhodnotenia vypočítanej z nadhodnotených pixelov a podhodnotenia vypočítanej z podhodnotených pixelov) a rozsiahlymi kontingenčnými tabuľkami s podmieneným formátovaním.

Týmito validáciami podľa Hurbánka a kol. (2010) a podľa Lajčákovvej (2013) sa dokázalo, že tematická presnosť mapy IMD2006\_v2 je tým nižšia, čím je osídlenie menej kompaktné a viac rozptýlené.

Hurbánek a kol. (2010) sa tiež venoval problému s určením hodnoty „thresholdu“ tak, aby vrstva IMD2006\_v2 čo najviac súhlasila s referenčnou databázou vytvorenej tematickej

mapy s dvoma triedami (priepustné, nepriepustné), vytvorenou použitím práve pomocou určitej prahovej hodnoty. Pomocou referenčných lokalít Popradu a Dubnice poukázal, že vrstva IMD v niektorých prípadoch extrémne nadhodnocuje stredné a väčšie hodnoty nepriepustnosti a naopak extrémne podhodnocuje stredné a menšie hodnoty nepriepustnosti. To podľa Hurbánka (2010) „môže viesť k nadhodnoteniu podielu nepriepustných plôch v oblastiach s kompaktnjším osídlením (zvyčajne v mestských oblastiach) a nedostatočným zastúpením alebo úplným vynechaním malých nepriepustných plôch rozptýlených vidieckych sídel“.

#### **E. Odporúčania na validáciu vrstiev IMD stupňa nepriepustnosti s vysokým rozlíšením z roku 2006, 2009, 2012 a 2015**

Validácia všetkých vrstiev Copernicus IMD s vysokým rozlíšením si vyžaduje značné úsilie. Gallego a kol. (2016) navrhovali na jej realizáciu systematickú tvorbu vzoriek vyvinutú prostredníctvom vzorkovacieho štvorca LUCAS na základe dvojstupňovej stratifikácie.

Pre každý referenčný rok bola k dispozícii rastrová vrstva s rozlíšením 20 m, avšak za referenčný rok 2012 bolo v čase realizácie týchto validácií dodaných len približne 90 % z celkovej plochy (chýbalo Španielsko, Grécko, Cyprus a francúzske zámorské regióny).

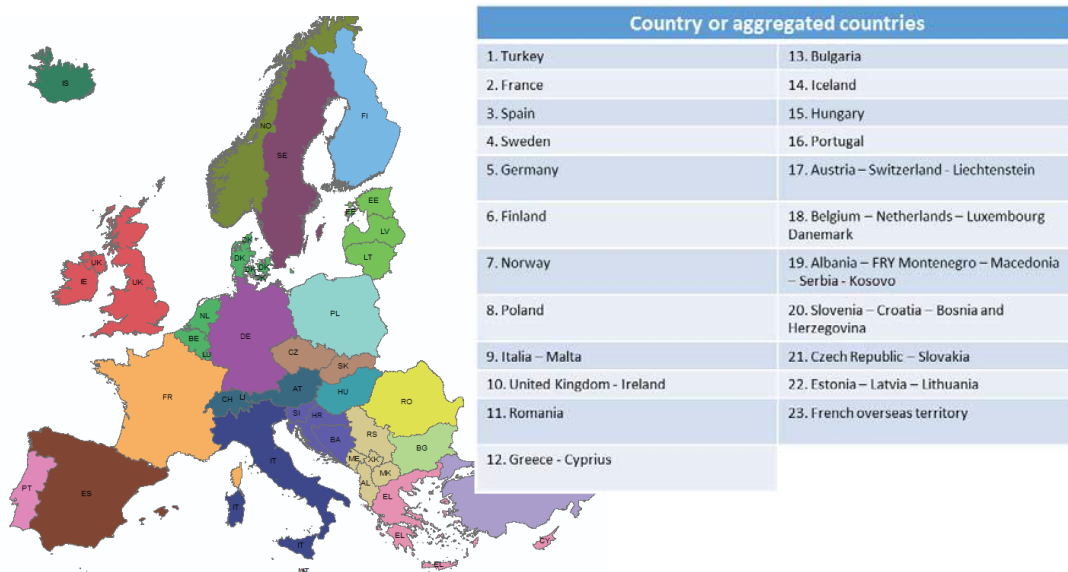
Cieľom pilotnej realizácie validácií bolo (Gallego a kol. 2016) bolo vyvinúť a implementovať postup hodnotenia presnosti, ktorý je schopný potvrdiť výsledky získané zúčastnenými krajinami počas overovacej fázy na základe vytvorenej vrstvy nepriepustnosti (Büttner 2012), ďalej je vhodný na posúdenie presnosti vrstvy nepriepustnosti na EHP39 a zároveň dokáže zabezpečiť, aby sa výsledky mohli analyzovať v biogeografickom regióne alebo na rozsiahlej úrovni krajín, aby sa zabezpečilo, že neexistujú veľké regionálne rozdiely.

Je potrebné poukázať na to, že validácie takéhoto typu dát nad takou rozsiahlou oblasťou si vyžadujú naozaj značné a väčšinou bezprecedentné úsilie. Väčšina literatúry o hodnotení presnosti vymedzenia nepriepustného povrchu pomocou diaľkového prieskumu Zeme má tendenciu zamerať sa na relatívne malé oblasti (Ji a Jensen 1999, Yang a kol. 2003, Chabaeva a kol. 2007). Jediným porovnateľným príkladom je národná databáza krajinej pokrývky Spojených štátov amerických (Wickham a kol. 2013) a nedávna štúdia od Hansen a kol. (2014). Táto štúdia sa však primárne zameriava na zmeny nepriepustnosti a preto sa hodnotenie presnosti kombinuje s hodnotením iného využitia zeme. V skutočnosti

sa zaznamenáva relatívne nízka presnosť pre triedu zmien nepriepustnosti, ktorá je čiastočne spojená s tým, že nepriepustné oblasti, aj keď stále rastú, stále zaberajú veľmi malú časť celkovej rozlohy. V Európe sa v súčasnosti odhaduje, že umelé povrchy predstavujú menej ako 5 % celkového EHP39 (Büttner a kol. 2012) a malo by sa očakávať, že nepriepustné povrchy predstavujú len podmnožinu tejto oblasti, čo z nej robí veľmi zriedkavý druh a naďalej tak vzniká osobitná výzva na posúdenie jeho presnosti s vysokou mierou presnosti.

Wickham a kol. (2013) použil stratifikovaný náhodný výber vzoriek, pričom použil ako vrstvy triedy využitia zeme. To zaručuje, že intenzita odberu vzoriek je prispôbená výskytu každej triedy, aby sa dosiahla dostatočná úroveň presnosti používateľa aj v prípade zriedkavých tried (Wickham a kol., 2013). Pokiaľ ide o európsku vrstvu nepriepustnosti, už pred týmito validáciami sa uplatnil podobný prístup, ale založený na stratifikovanom systematickom výbere vzoriek s použitím štatistického zisťovacieho štvorca EUROSTAT pre využitie/pokrytie územia (LUCAS) (Gallego a Delincé 2010). Hlavnou výhodou použitia prístupu založeného na systéme LUCAS je, že systematický prístup zaručuje úplnú sledovateľnosť a je tiež možné zdieľanie vzoriek na posúdenie viacerých produktov, čím sa poskytnú potenciálne úspory z rozsahu.

Prvým krokom pri validácii vrstiev IMD z roku 2006, 2009, 2012 a 2015 bolo určenie krajín alebo skupín krajín s plochou väčšou ako 90 000 km<sup>2</sup> (Obrázok 5). Vytvorilo sa tak 23 zón. Tento krok umožnil analýzu výsledkov pre rôzne krajiny a biogeografické regióny, aby sa posúdilo, či sa objavila akákoľvek rôznorodosť v kvalite údajov v rôznych regiónoch.



Obrázok 5 Rozdelenie krajín alebo skupín krajín EHP39 pri validáciách  
(Zdroj: Sannier a kol. 2016)

V druhom kroku bolo podľa Sanniera a kol. (2016) pre každú krajinu alebo skupinu krajín vytvorených 6 vrstiev nasledovne:

- Vrstva nepriepustnosti 2006 – 2009 – 2012 – 2015: Stupeň nepriepustnosti 30 – 100 % v období 2006 – 2009 – 2012 – 2015. Táto kategória bola rozdelená na dve vrstvy definované prienikom s umelými a neumelými triedami CLC.
- Vrstva priepustnosti s vysokou pravdepodobnosťou 2006 – 2009 – 2012 – 2015: Stupeň nepriepustnosti 0 – 29 % a CLC nepriepustné triedy 2006 – 2009 – 2012 – 2015.
- Vrstva priepustnosti s nízkou pravdepodobnosťou 2006 – 2009 – 2012 – 2015: zvyšok oblasti 2006 – 2009 – 2012 – 2015.
- Zmeny nepriepustnosti 2006 – 2009: všetky zmeny (zvýšenie a zníženie).
- Zmeny nepriepustnosti 2009 – 2012: všetky zmeny (zvýšenie a zníženie).
- Zmeny nepriepustnosti 2012 – 2015: všetky zmeny (zvýšenie a zníženie).

Následne boli vypočítané chyby priradenia na zvýšenie presnosti používateľa a zaradenie chyby vynechania bolo kvôli zvýšeniu presnosti producenta.

Minimálne 50 až 100 jehnohektárových pixelov (100 × 100 m) bolo vybratých na vrstvu v každej z 23 zón. Toto množstvo bolo považované za dostatočné na dosiahnutie požadovanej presnosti a celkovo bolo vybraných 20 164 vzoriek nad krajinami Európskeho hospodárskeho priestoru (EHP39). Podmienkou bolo (Gallego a kol. 2016), aby krajina

alebo skupina krajín, nad ktorými bola vytvorená vzorka, mali plochu väčšiu ako 90 000 km<sup>2</sup>. V každom štvorci sa použila mriežka 5 × 5 bodov s rozostúpením 20 m. Tieto body boli interpretované na ortofotosnímках s rozlíšením lepším než 2,5 m, aby sa zistilo, či boli znepriepustnené. Ak bod padol na hranicu nepriepustného prvku, uplatňovalo sa pravidlo posunu tak, že zhruba polovica bodov v tejto situácii bola klasifikovaná ako nepriepustná.



Obrázok 6 Príklad jedného pixelu s rozmermi 100 × 100 m s 5 × 5 bodmi s 20 m rozostúpením  
(Zdroj: Sannier a kol. 2016)

Aj v tomto prípade bol podľa novších odporúčaní Büttnera (2012) použitý „threshold“ – prah 30 %, čím sa tento produkt s kontinuálnou hustotou premenil na binárnu masku, tematickú mapu s dvoma triedami (Büttner 2012). Iné alternatívy možno nájsť v literatúre na riešenie kvantitatívnych produktov podľa Wilmotta (1981), Legatesa a Mc Caba (1999) a Duveiller a kol. (2016). V literatúre o diaľkovom prieskume Zeme sa často používajú indikátory dohody pre kvantitatívne parametre (Ji a Gallo 2006, Silván Cárdenas a Wang 2008, Meroni a kol. 2013).

Tematická presnosť vrstvy nepriepustnosti bola následne prezentovaná vo forme chybovej matice vytvorenej z výsledkov interpretácie vzoriek a ich skutočných hodnôt v nepriepustnej vrstve. Ako vysvetľujú Selkowitz a Stehman (2011), nerovnaká intenzita odberu vzoriek vyplývajúca zo stratifikovaného prístupu systematického odberu vzoriek by sa mala zohľadniť použitím váhového faktora ( $p$ ) na každú jednotku vzorky založeného na pomere medzi počtom vzoriek a veľkosťou zvažovanej vrstvy

$$p_{ij} = \frac{1}{N} \sum_{x \in (ij)} \frac{1}{\pi}$$

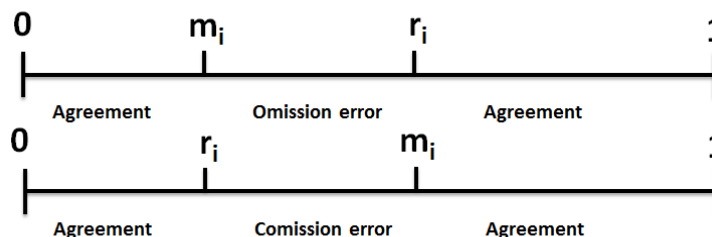
kde  $i$  a  $j$  sú stĺpce a riadky v matici,  $N$  je celkový počet možných jednotiek a  $\pi$  je intenzita odberu vzorky pre danú vrstvu.

Dôvodom je, že vzorky z menších vrstiev vykazujú vyššiu intenzitu odberu vzoriek ako tie z väčších vrstiev. Korekcia intenzity odberu sa preto uplatňovala na vytvorené chybové matice podľa spomínaného postupu, ktorý opisujú Selkowitz a Stehman (2011), a ktorý aplikoval Olofson a kol. (2013). To vedie k váhovému faktoru nepriamo úmernému pravdepodobnosti zaradenia vzoriek z danej vrstvy. Neuplatnenie tejto opravy by mohlo viesť k podhodnoteniu alebo nadhodnoteniu presnosti mapy.

Počiatkové výsledky založené na binárnej konverzii mapy použitím prahu 30 % indikujú úroveň chýb priradenia a chýb vynechania podstatne vyššiu ako je požadovaná maximálna úroveň 15 % stanovená v špecifikáciách produktu. Existuje tiež značná variabilita v rôznych biogeografických regiónoch a skupinách krajín a treba poznamenať, že vrstva 2012 vykazuje vyššiu presnosť ako vrstva v rokoch 2009 a 2006, čo možno vysvetliť postupným zlepšovaním pri produkcii novších verzii.

Chyby vynechania sa zdajú byť výrazne ovplyvnené zastúpením vynechaní vo vrstvách s nízkou pravdepodobnosťou výskytu chýb vynechania, čo má značnú váhu. To môže naznačovať, že stratifikačný prístup vybraný na priepustnú vrstvu sa pravdepodobne musí prehodnotiť.

Hodnotenie založené na binárnej maske predpokladá, že pre každý pixel sú k dispozícii úplné informácie, čo však nie je v tomto prípade, pretože odhad úrovne nepriepustnosti pixela je založený na interpretácii vybraných bodov. Z toho dôvodu bola medzi navrhnutými ukazovateľmi tematickej presnosti aj kvantifikácia nezhody na úrovni pixelov ako rozdiel medzi hodnotou vrstvy  $m_i$  a referenčnou hodnotou  $r_i$ . Ak bola hodnota vrstvy väčšia ako referenčná hodnota pre daný pixel, tak šlo o chybu priradenia, ale ak nastal opak, šlo o chybu vynechania. Tento postup mal vykazovať realistickejší odhad množstva chýb priradenia a vynechania.



Obrázok 7 Chyby vynechania a priradenia pre kvantitatívne kontinuálne hodnoty od 0 po 1  
(Zdroj: upravené podľa Sannier a kol. 2016)

Ak by autori mali k dispozícii kompletné informácie o bunke, tak pre referenčné údaje by primeraná miera chyby priradenia a vynechania bola:

$$\varphi = \frac{\sum_i \text{pos}(m_i - r_i)}{\sum_i m_i} \quad \vartheta = \frac{\sum_i \text{pos}(r_i - m_i)}{\sum_i r_i}$$

kde  $\text{pos}(x)$  je pozitívna časť, t. j.  $\text{pos}(x) = x$ , ak  $x > 0$  a  $\text{pos}(x) = 0$  ak  $x \leq 0$ ,  $m_i$  udáva podiel nepriepustnosti podľa vrstvy IMD,  $r_i$  je referenčná hodnota podielu nepriepustnosti. Pre každú vzorkovaciu jednotku s rozlíšením  $100 \times 100$  m bola použitá kvantitatívna hodnota IMD (odhadovaná v % podľa klasifikácie satelitného obrazu) a referenčná hodnota, ktorá je odhadom nepriepustnosti získaným interpretáciou 25 bodov.

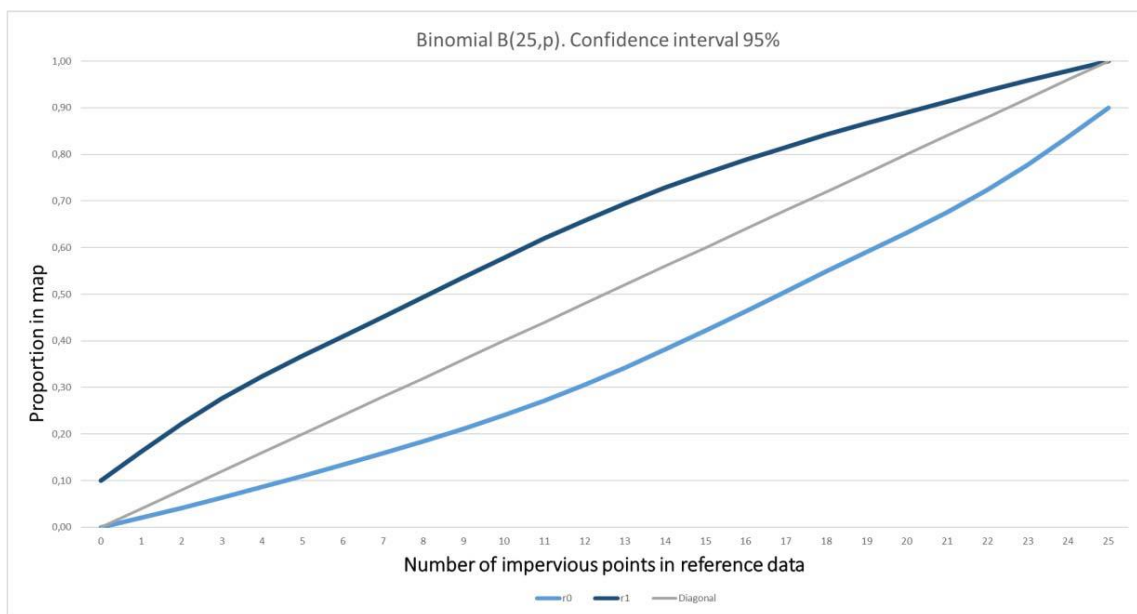
Takýmito výpočtami vznikla podľa Gallega a kol. (2016) pri IMD2015 chyba priradenia vo výške 21,9 % a chyba vynechania 40,2 %, aj keď celková presnosť bola 98,4 %. Časť nesúhlasu medzi referenčnými dátami a údajmi z vrstvy IMD pochádza z neúplnej informácie o bunke pre referenčné údaje: aj keď predpokladáme, že vrstva IMD a referenčné údaje sú v dokonalom súhlase, stále by existoval rozdiel medzi  $m_i$  a  $r_i$ , pretože priestorová podpora je iná:  $m_i$  sa vzťahuje na celý pixel, zatiaľ čo  $r_i$  sa vzťahuje na vzorku bodov vo vnútri. Počet nepriepustných bodov, ktoré používame ako referenčnú hodnotu, má pravdepodobnostné rozloženie v dôsledku vzorkovania v rámci buniek. Ak je vzorkovanie vnútri náhodné, počet bodov kopíruje binomické rozdelenie  $B(25, p)$  s o niečo nižším rozptylom v dôsledku systematického vzorkovania. To má vplyv na očakávané chyby vynechania a pridania.

Jeden možný spôsob, ako to vziať do úvahy, bolo vymedzenie intervalu spoľahlivosti pod nulovou hypotézou  $m_i = p_i$ . Pre akýkoľvek pár  $(r_i, m_i)$  však neexistuje úplne presvedčenie, že  $m_i \neq p_i$ . Možným spôsobom riešenia tohto zdroja neistoty bola kvantifikácia chýb pridania a chýb vynechania:



$$\varphi = \frac{\sum_i \text{pos}(m_i - r_{1i})}{\sum_i m_i} \quad \vartheta = \frac{\sum_i \text{pos}(r_{0i} - m_i)}{\sum_i m r_i}$$

kde  $r_{0i}$  a  $r_{1i}$  sú najbližšie body v intervale spoľahlivosti. Získané výsledky ukazujú, že chyba priradenia je menej ako 10 % a chyba vynechania je menej ako 20 %, čo presahuje požadovanú úroveň presnosti tejto vrstvy vzhľadom na chyby vynechania. Ale opäť je to silne ovplyvnené vynechaniami v rozsiahlych oblastiach vo vrstvách s nízkou pravdepodobnosťou vynechania. Pri odstraňovaní iba 39 bodov zo vzorky zodpovedajúcej týmto rozsiahlym vrstvám (z viac ako 20 000) sa úroveň chyby vynechania znížila na 13 %, čo je v prijateľnom rozmedzí presnosti očakávanej pre vrstvu nepriepustnosti. Znovu to naznačuje, že stratifikačný prístup pri posudzovaní chýb musí byť revidovaný tak, aby lepšie zacielil oblasti potenciálneho vynechania (Gallego a kol. 2016).



Graf 1 Znáročnenie priebehu 95 % intervalu spoľahlivosti pre mriežku 5 x 5 bodov v celom rozsahu hodnôt stupňa nepriepustnosti (Zdroj: Gallego a kol. 2016)

Graf 1 predstavuje priebeh 95 % intervalu spoľahlivosti pre binomické rozdelenie  $B(25, r_i)$ . Všimnite si, že len pre hodnoty blízke 0,5 je možné použiť bežnú Gaussovú aproximáciu, ktorá vedie k intervalu približne  $(r_i \pm 2s_i)$ , zatiaľ čo pre hodnoty blízke 0 alebo 1 sú intervaly silne asymetrické.

Posledným krokom autorov (Gallego a kol. 2016) bol odhad veľkosti územia pokrytej nepriepustnými plochami. Prirodzený prístup k odhadovaniu územia pomocou

klasifikovaných satelitných snímok je počítanie pixelov alebo iná, podobná klasifikácia, ktorá sa zameriava na podiel určitého parametra, tak ako sa to bolo aj v tomto prípade:

$$A = \sum_i m_i$$

Výsledok tohto odhadu pre 23 vytvorených zón (okrem Španielska, Grécka, Cypru a francúzskych zámorských území) je 2,06 %, približne 107 000 km<sup>2</sup>. Je však jasné, že tento odhad má skreslenie, ktoré môže byť veľmi veľké (Bauer a kol. 1978, Houston a Hall 1984, Czaplewski 1992) a zhruba zodpovedá rozdielu medzi chybou priradenia a chybou vynechania (Carfagna a Gallego 2005). Jednoduchým rozumným nápadom môže byť náprava skreslenia s rozdielom medzi chybou priradenia a chybou vynechania vypočítanými pomocou chybovej matice. Bohužiaľ odborníci pracujúci v oblasti diaľkového prieskumu Zeme vypočítavajú chybovú maticu stále často na základe účelového súboru jednotiek namiesto skutočnej pravdepodobnej vzorky. Vo vedeckej literatúre účelové vzorky skoro zmizli vďaka úsiliu viacerých autorov (Congalton a Green 1999, Stehman 2009, Olofsson a kol. 2014), ale sú stále časté v správach o projektoch uskutočnených pre rôzne inštitúcie a zostávajú nepublikované. Lepší odhad nepriepustného územia bol dosiahnutý jednoduchým pridaním rozdielu mimodiagonálnych hodnôt: 107 000 + 57 000 – 23 200 ~ 140 000 km<sup>2</sup>, t. j. 2,74 % územia.

Tabuľka 7 Chybová matica  
(Zdroj: Gallego a kol. 2016)

		Map			Omission error
		Impervious	Other	Total	
Reference	Impervious	82859	56981	139840	40.7%
	Other	23232	4961592	4984824	
	Total	106091	5018573	5124664	
Commission error		21.9%			

## F. Hodnotenie presnosti IMD2018 od dodávateľa produktu

Správa o hodnotení kvality IMD2018 bola zverejnená spoločnosťou EEA v júni 2020, ako jeden zo základných dokumentov, podľa požiadaviek tendra.

Do roku 2015 bola produkcia IMD založená najmä na údajoch z Landsat-8 s 30 metrovým priestorovým rozlíšením, na výrobu IMD2018 boli už použité údaje zo satelitu Sentinel-2 s priestorovým rozlíšením 10 m. Výsledkom je zlepšenie priestorového rozlíšenia vrstiev, čo viedlo k presnejšiemu zachyteniu nepriepustných povrchov. Kým vrstvy IMD boli

tak doteraz vyrábané s priestorových rozlíšením 20 m, od roku 2018 je priestorové rozlíšenie 10 m. No napriek tomu boli pre účely validácie jednotlivé 10 m pixely agregované do štvorcov s rozlíšením 100 m (Smith a kol. 2020).

Validácia IMD2018 prebehla na 2 úrovniach:

- Bola vytvorená regresná krivka, ktorá vyjadrovala koreláciu medzi referenčnými a mapovými hodnotami s cieľom identifikovať akékoľvek systematické odchýlky medzi posudzovanými dátami.
- Na základe „thresholdu“ 30 % sa vytvorila tematická mapa s dvoma triedami – priepustné plochy (0 – 29 % podiel nepriepustných plôch na územnú jednotku) a nepriepustné povrchy (30 – 100 % podiel nepriepustných plôch na územnú jednotku).

Validácia prebehla rovnako ako aj v predošlých rokoch na základe náhodnej stratifikovanej vzorky, ktorá zabezpečila, že pri výbere bola zastúpená každá trieda krajinej pokrývky. Takto bola pomocou vzorkovacieho štvorca LUCAS na základe dvojstupňovej stratifikácie vytvorená sieť štvorcov naprieč celou Európou (EHP39). Z predchádzajúcich validácií sa zachoval aj postup, že v každom štvorci so 100 m rozlíšením bolo rozmiestnených 5 × 5 bodov s 20 m rozostúpením, hoci rozlíšenie IMD2019 je už 10 m a nie 20 m. Takto bolo vybraných 25 777 vzoriek (Smith a kol. 2020).

Kompletná analýza dát prebehla na rovnakom princípe (popísaný v predchádzajúcej podkapitole E) ako to bolo doposiaľ pri vrstvách z roku 2012 či 2015.

Podľa nej IMD2018 spĺňa väčšinu špecifikácií, ale požiadavka na minimálnu tematickú presnosť 90 %, založená na 30 % „thresholde“, nebola vo viacerých oblastiach splnená. Aj táto vrstva má podľa výsledkov hodnotenia podceňovať výskyt nepriepustných plôch, ukazuje sa však vyváženejšie rozdelenie medzi chybami vynechania a chybami priradenia v porovnaní s ostatnými rokmi. Celkovo došlo k výraznému zlepšeniu kvality vrstvy IMD v porovnaní s rokom 2015. Hlavný dôvod zlepšenia presnosti používateľa a presnosti producenta súvisí s lepším rozlíšením pri monitorovaní krajinej pokrývky (Smith a kol. 2020).

## 2 CIEĽ PRÁCE

Hlavným cieľom tejto dizertačnej práce je komplexná validácia Európskej vrstvy hustoty nepriepustnosti, konkrétne jej postupne vylepšených verzií pre referenčné roky 2006 (IMD2006) a 2009 (IMD2009).

Čiastkové ciele, ktoré podmieňujú dosiahnutie hlavného cieľa sú:

- priblížiť čitateľovi ponuku a dostupnosť priestorových dát o krajinnej pokrývke, ktoré poskytuje európsky program Copernicus,
- popísať Európsku vrstvu hustoty nepriepustnosti, jej vznik, priebeh vývoja a vylepšovania jednotlivých dostupných verzií a spôsoby ich validácii zo strany producenta či iných nestranných používateľov,
- vytvoriť dostatočne veľkú referenčnú databázu IRD2006 a IRD2009 nad vybraným záujmovým územím Slovenska a Česka, popísať spôsob tvorby a presný postup vzniku postupne vylepšených vysoko kvalitných referenčných verzií,
- navrhnúť štatistické kvantitatívne metódy vhodné na čo najpresnejšiu komplexnú validáciu intenzitných máp, porovnať ich navzájom, prípadne uviesť výhody a nevýhody ich použitia,
- pomocou štatistických metód kvantitatívne zhodnotiť presnosť jednotlivých postupne vylepšovaných verzií IMD2006 a IMD2009, navzájom porovnať vylepšené verzie a dosiahnuté výsledky na dvoch záujmových územiach,
- diskutovať o možných zdrojoch chýb a ponúknuť návrhy na zlepšenie.

### 3 METODIKA PRÁCE A METÓDY SKÚMANIA

Maucha, Bütner a Kosztra (2010) sa zmieňujú, že pri každej validácii je nutné položiť si dve otázky: „Na koľko výrobok zodpovedá špecifikáciám? Na koľko výrobok zodpovedá realite?“ Odpoveď na prvú otázku podľa autorov informuje o kvalite hotového diela, tvorí základ pre prijatie, odmietnutie alebo revíziu produktu. Druhá odpoveď vypovedá o použiteľnosti výrobku. Zdôrazňujú, že ak musí byť rozhodnuté o výrobku, musia tiež byť hotové kvantitatívne výsledky. Kladú dôraz na to, že rôzne použitia vyžadujú rôzne druhy akostných parametrov. Autori zdôrazňujú, že obe otázky sa týkajú prekrývajúcich sa problémov, obe sú dôležité a zvyčajne majú inú odpoveď, teda špecifikácie produktu a kontrola kritérií hodnotenia by mala viesť k správne vyhodnoteniu hodnoty produktu.

V teoretickej časti tejto práce, v kapitole 1.3.3 Niektoré prípadové štúdie validácií doma i v zahraničí, sme sa snažili priblížiť čitateľovi niekoľko doposiaľ vykonaných štúdií, ktoré sa venujú validácií vrstiev IMD jednotlivých referenčných rokov.

Nakoľko sa používané metódy interných validácií zo strany poskytovateľa (chybová matica, presnosť používateľa a presnosť producenta, celková presnosť) a ich výsledky nezdajú byť dôveryhodne interpretované vzhľadom na relatívne malú vzorku referenčných dát v pomere ku rozlohe skúmaného územia (EHP39), v nasledujúcej časti našej dizertačnej práce pristúpime k nami navrhovaným metódam možných validácií takýchto vrstiev s vysokým rozlíšením a k ich odporúčaniam na používanie do budúcnosti pri obdobných validáciách. Použitie metód načrtujeme na konkrétnej validácii IMD2006 (validácia postupne vylepšovaných dostupných verzií IMD2006\_v1\_v2\_v3\_v5) a IMD2009 (validácia dvoch dostupných verzií IMD2009\_v1\_v3), ku ktorej sme si vytvorili vlastnú referenčnú databázu nad územím Slovenska a Česka (IRD2006\_v3 a IRD2009\_v3):

IMD2006:

- IMD2006\_v1 (sprístupnená 12. 6. 2009, tzv. „precursor“)
- IMD2006\_v2 (sprístupnená 26.1.2010, tzv. „enhancement“)
- IMD2006\_v3 (sprístupnená 22. 1. 2013, publikovaná pri príležitosti vytvorenia IMD2009)
- IMD2006\_v4 (sprístupnená v roku 2014, publikovaná pri príležitosti vytvorenia IMD2012) – v čase dostupnosti sa nám nepodarilo ju získať a EEA rozhodla neposkytovať staršie verzie verejnosti po zverejnení novších verzií.

- IMD2006\_v5 (sprístupnená 15. 1. 2019, publikovaná pri príležitosti vytvorenia IMD2015)

#### IMD2009:

- IMD2009\_v1 (sprístupnená 22. 1. 2013, publikovaná pri príležitosti vytvorenia IMD2009)
- IMD2009\_v2 (sprístupnená v roku 2014, publikovaná pri príležitosti vytvorenia IMD2012) – v čase dostupnosti sa nám nepodarilo ju získať
- IMD2009\_v3 (sprístupnená 15. 1. 2019, publikovaná pri príležitosti vytvorenia IMD2015)

#### IRD2006:

- IRD2006\_v3 – naša referenčná databáza, ktorú sme 3-krát upravovali a vylepšovali

#### IRD2009:

- IRD2009\_v3 – naša referenčná databáza, ktorú sme 3-krát upravovali a vylepšovali

Najviac využívaný postup validácie v prípade, že produkt, ktorý sa validuje, má podobu rastra, je nasledovný (Hurbánek, Rosina 2013):

- Zo všetkých pixelov validovanej vrstvy vyberieme jednoduchý náhodný výber pixelov pomerne veľkej vzorky.
- V pixeloch danej vytvorenej vzorky zistíme referenčné hodnoty prislúchajúce premennej, ktorú skúmame.
- Referenčné hodnoty a hodnoty validovanej vrstvy porovnáme.
- Zhody, resp. nezhody medzi porovnávanými hodnotami kvantifikujeme.

Validácia sme realizovali prostredníctvom kvantitatívnych štatistických metód s využitím dostupných dát o percentuálnom zastúpení nepriepustnosti v jednotlivých pixeloch pre všetky dostupné verzie IMD2006 a IMD2009 a referenčných dát o percentuálnom zastúpení nepriepustnosti v jednotlivých pixeloch získaných vizuálnou interpretáciou leteckých ortofotosnímkov vytvorených z diaľkového prieskumu Zeme, príp. s reálnym stavom krajinnej pokrývky v teréne.

### 3.1 Zájmové územie a príprava referenčnej databázy

Na to, aby bolo možné venovať sa jednotlivým metódam validácie, pokladáme za dôležité priblížiť čitateľovi základnú sídelnú štruktúru zájmového územia Slovenska a Česka a postup tvorby referenčnej databázy na vybranom území.

Slovensko a Česko sú štáty strednej Európy so spoločnými, ale aj odlišnými znakmi. V oboch krajinách je nerovnomerný výskyt obyvateľstva. Najvyššiu hustotu obyvateľov na území Slovenska má oblasť Považia a okolie hlavného mesta, v Česku je to severovýchod krajiny a severná Morava.

Rozloha Slovenska je 49 036 km<sup>2</sup>. Počet obyvateľov dosiahol v roku 2021 číslo 5 449 270 obyvateľov (ŠÚSR 2022) a hustota obyvateľstva je 111 obyvateľov na km<sup>2</sup> (ŠÚSR 2022). Najväčší vplyv na rozloženie sídelnej štruktúry má povrch Slovenska. Najviac sa ľudia koncentrujú v blízkosti veľkých riek, pozdĺž kotlín a v nížinách na východe a západe krajiny. V strede Slovenska, kde sa nachádzajú rozľahlé pohoria, je koncentrácia ľudí a osídlenia podstatne nižšia. Prevažne kompaktné osídlenie je sústredené v mestách a obciach. Na území Slovenska sa vyskytujú aj roztratené sídla. Je to špecifické osídlenie, ktoré má v rôznych oblastiach krajiny rôzne pomenovanie. V oblasti Myjavy sa nachádzajú kopanice, okolie Hriňovej a Detvy je typický výskytom lazov, stále nájdeme pri Novej Bani a Žarnovici a na Orave sú rale. Zastavané plochy pokrývajú cca 2,4 % z celkovej rozlohy Slovenska. Hlavné mesto Bratislava leží na juhozápade krajiny.

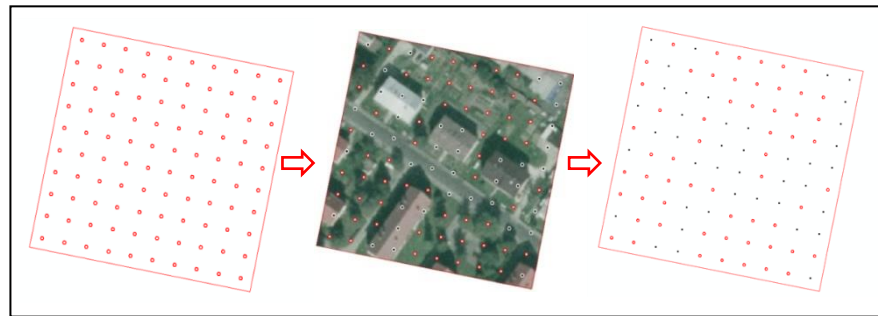
Naším druhým zájmovým územím je Česko, ktorého rozloha je 78 866 km<sup>2</sup>. Podľa posledného sčítania obyvateľov v roku 2021 má krajina 10 515 669 obyvateľov (ČSÚ 2022), čo je 2-krát viac ako má Slovensko. Hustota obyvateľstva dosahuje hodnotu 133 obyvateľov na km<sup>2</sup> (ČSÚ 2022). Českej republike s pravidelnou sídelnou štruktúrou dominuje jej hlavné mesto Praha. Ak porovnáme osídlenie Česka s osídlením na Slovensku, je určite menej kompaktné. Zastavané plochy zaberajú o niečo vyššie percento, približne 3,3 %. Tento podiel postupne narastá rovnako ako aj na území Slovenska.

Tabuľka 8 Základné údaje štátov zájmového územia validácie  
(Zdroj dát: ČSÚ, ŠÚSR 2022)

Názov štátu	Rozloha (km <sup>2</sup> )	Odhad podielu zastavaných plôch z celkovej rozlohy krajiny (%)	Počet obyvateľov	Hustota zaľudnenia (obyv./km <sup>2</sup> )
Slovensko	49 036	2,6	5 415 459	111
Česko	78 866	3,3	10 515 669	133

Na našom záujmovom území sme si vytvorili referenčnú databázu. Pomocou metódy jednoduchého náhodného výberu sme vybrali 40 000 pixelov s rozlíšením  $100 \times 100$  m, z čoho bolo 20 000 pixelov umiestnených na Slovensku a 20 000 pixelov na území Česka. V každom pixeli sme vybrali 100 bodov, ktoré boli rozložené v pravidelnej matici ( $10 \times 10$  bodov). Body boli od seba vzdialené 10 m.

V každej krajine sme na základe ortofotosnímkov z roku 2006 a z roku 2009 vizuálne interpretovali spolu 4 000 000 bodov (2 000 000 za rok 2006, 2 000 000 za rok 2009). Dokopy bolo interpretovaných 8 000 000 bodov (spolu 80 000 pixelov v oboch krajinách). Interpretácia prebehla na základe toho, či sa daný bod nachádzal alebo nenachádzal na nepriepustnom, resp. priepustnom povrchu.

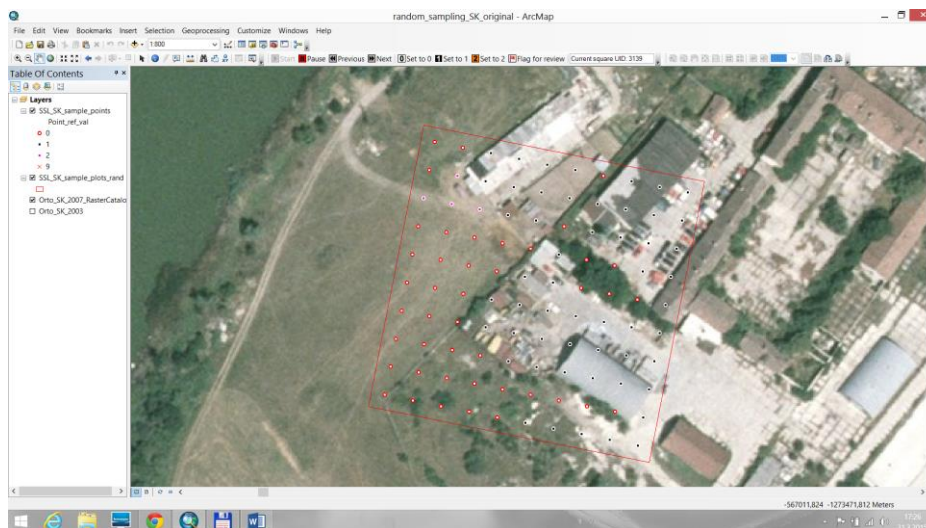


Obrázok 8 Tvorba referenčnej databázy  
(Zdroj: vlastné spracovanie)

Každému bodu bola priradená jedna z nasledujúcich hodnôt (viď podkapitolu 1.3.3 D. Validácie referenčných lokalít s rozmermi  $6 \times 6$  km, klasifikácia podľa Lajčákovej (2013)):

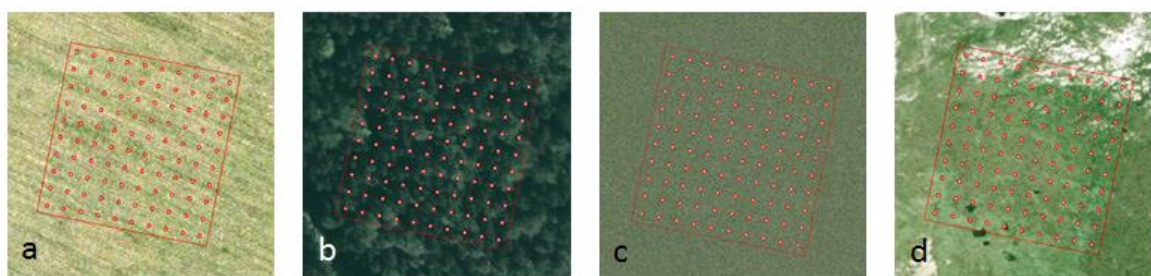
- 0 – priepustná plocha
- 1 – nepriepustná plocha s alochtónnym materiálom premiestneným na dané miesto aktivitou človeka – napr. budovy, cesty, parkoviská, ...
- 2 – nepriepustná plocha s autochtónnym materiálom znepriepustnená vďaka aktivite človeka – napr. poľné cesty, ...





Obrázok 9 Tvorba referenčnej databázy v prostredí programu ArcMap 10.2.2  
(Zdroj dát: EUROSENSE a GEODIS SLOVAKIA 2005 – 2007)

Hodnotu každého bodu automaticky zapísalo do stĺpca s názvom „ref value“. Ten bol súčasťou atribútovej tabuľky. Na rýchle posúvanie z jedného pixela na druhý sme vytvorili skript zabezpečujúci aj zväčšenie práve zobrazenej časti mapy. Pri interpretácii bol na obrazovke počítača zobrazený vždy len jeden vybraný pixel, ktorý interpretujúca osoba hodnotila. Spočítaním takto hodnotených bodov bola vyjadrená referenčná hodnota podielu nepriepustných povrchov daného pixelu (od 0 do 100 %), keďže každý pixel obsahoval vyššie spomínaných presne 100 bodov. Čím vyššiu percentuálnu hodnotu mal pixel, tým viac sa správal ako nepriepustný.



Obrázok 10 Pixely z náhodného výberu, interpretované ako priepustné a) momentálne bez vegetácie, b) s vegetáciou, c) vodné plochy, d) dlhodobu bez vplyvu človeka – skaly  
(Zdroj dát: EUROSENSE a GEODIS SLOVAKIA 2005 – 2007)



Obrázok 11 Pixely z náhodného výberu, interpretované s určitou hodnotou podielu nepriepustných povrchov

(Zdroj dát: EUROSENSE a GEODIS SLOVAKIA 2005 – 2007)

Pri interpretácii nastali aj prípady, kedy interpretujúca osoba musela vyhodnotiť snímku s prihliadnutím na spôsob, akým bola nasnímaná. Ak sa na snímke nasnímanej kamerou pod zjavne viditeľným uhlom vyskytol bod umiestnený na vertikálnej ploche, napr. na múre výškovej budovy, bol tento bod interpretovaný ako bod nepriepustný, s hodnotou 1. Opačný prípad nastal v prípade, že sa bod nachádzal na korune stromu, pod ktorým sa zjavne nachádzal priepustný povrch. Vtedy mu bola priradená hodnota 0, bod bol považovaný za priepustný. Takého kontextuálne informácie musela interpretujúca osoba vyhodnocovať aj v prípadoch, kedy na skúmané objekty dopadal tieň, čo spôsobovalo sťaženie viditeľnosti interpretácie konkrétnych povrchov. Medzi ďalšie zaujímavé príklady patria prípady ťažko identifikovateľných objektov, ktoré boli umiestnené na dvoroch, ako napr. parkujúce autá na tráve, kopy štrku, a pod., kedy bolo zjavné, že sa pod nimi nachádza priepustný povrch. Preto bolo potrebné skúmať tieto objekty z hľadiska toho, ako dlho sa na danej priepustnej ploche nachádzajú, a to porovnaním danej oblasti na snímkach z aspoň 2 časových období. Ak bol objekt považovaný za krátkodobo lokalizovaný na danom priepustnom povrchu, bola mu priradená hodnota 0. Naopak, ak bol objekt považovaný za dlhodobo lokalizovaný na priepustnom povrchu, dostal hodnotu 1.

Boli vytvorené dve verzie referenčných databáz:

- IRD\_A – body s hodnotou 2 boli považované za priepustné povrchy (rovnako ako body s hodnotou 0)
- IRD\_B – body s hodnotou 2 boli považované za nepriepustné povrchy (rovnako ako body s hodnotou 1)

Takýto postup sme zvolili opäť z dôvodu nejasnej definície pojmu nepriepustnosti (podobne ako v práci Lajčákovej (2013), v ktorej boli vytvorené dva typy databáz pre referenčné štvorce s rozmermi  $6 \times 6$  km), nakoľko doposiaľ nebolo nikde zadefinované presne a jasne, či nepriepustné povrchy s autochtóнным materiálom patria alebo nepatria do skupiny

nepriepustných. Preto sme sa rozhodli s nimi v prvom prípade pracovať ako s priepustnými a v druhom prípade ako s nepriepustnými povrchmi.

Vizuálnu interpretáciu snímok sme vykonali v programe ArcMap 10.2.2 v jednotlivých pixeloch na slepo (EEA 2008). Osoba, ktorá hodnotila body, nepoznala hodnoty podielu nepriepustnosti jednotlivých pixelov vo vrstve, ktorú validovala, aby sme neovplyvnili úsudok. Odhadovanie referenčných hodnôt prebehlo za rok 2006 a rok 2009. Spolu tak bolo interpretovaných dvakrát 80 000 pixelov (pre verziu A aj pre verziu B). Každá vzorka (SK\_IRD2006\_A, SK\_IRD2006\_B, SK\_IRD2009\_A, SK\_IRD2009\_B, CR\_IRD2006\_A, CR\_IRD2006\_B, CR\_IRD2009\_A, CR\_IRD2009\_B) bola následne opäť dvakrát skontrolovaná a v prípade menších nedostatkov opravená. V prípade Česka bola náhodne vytvorená vzorka za rok 2009 – CR\_IRD2009\_A a CR\_IRD2009\_B interpretovaná najprv Radkou Pohlovou (2013) a následne však upravená nami interpretujúcou osobou, aby sa zachovala jednotnosť interpretácie, resp. prípadná chybovosť interpretujúcej osoby. Dvojnásobným opakovaným kontrolovaním každej jednej vzorky vznikli vylepšené, v poradí tretie verzie referenčných databáz s nasledovným označením IR2006\_v3, resp. IR2009\_v3 pre obe krajiny, na území ktorých sme vykonávali validáciu.

Na Slovensku boli použité ako podkladové mapy ortofotosnímky firiem © GEODIS Slovakia s.r.o. a EUROSENSE s.r.o. nasnímané okolo roku 2006, ktoré boli dostupné cez Google Earth. Referenčné databázy vytvorené podľa nich nám poslúžili na validáciu vrstvy IMD2006. Ako podklad pre tvorbu referenčných databáz na validáciu vrstvy IMD2009 boli použité ortofotosnímky nasnímané z roku 2009, ktoré sme si zakúpili (EUROSENSE a GEODIS SLOVAKIA 2008 – 2010). Ich cena bola však veľmi vysoká, a preto sme boli schopní zakúpiť len 5000 vzoriek. Na to, aby sme zakúpili čo najlepší výber vzoriek, museli sme najprv identifikovať zvyšných 15 000 vzoriek, pri ktorých sme predpokladali, že sa v nich podiel nepriepustných povrchov za dané obdobie rokov 2006 až 2009 nezmenil. Šlo prevažne o pixely so 100 % podielom priepustných povrchov (trávnaté oblasti, lesy a vodné plochy).

V prípade Česka boli podkladové snímky nasnímané v období rokov 2006, 2008 a 2009. Využité boli ortofotosnímky WMS služby národného geoportálu ČÚZK.

Hodnoty „IMD a IRD\_A“ a hodnoty „IMD a IRD\_B“ sme porovnávali na úrovni jednotlivých 100 m pixelov použitím vybraných kvantitatívnych metód. Štatistické hodnotenie prebehlo v prostredí programu MS Excel.

### 3.2 Metóda histogramov

Metóda histogramov je metóda vhodná na porovnávanie početnosti jednotlivých hodnôt stupňa nepriepustnosti v pôvodnej databáze (v IMD databáze) s početnosťou hodnôt stupňa nepriepustnosti v referenčnej databáze (v IRD databáze).

Porovnávaním rozdielu IMD a IRD hodnôt miery nepriepustnosti sa získava miera nadhodnotenia a podhodnotenia.

Popri porovnaní počtov nadhodnotených a podhodnotených pixelov možno dodatočnú informáciu o kvalite dát získať aj porovnaním rozdelenia hodnôt stupňa nepriepustnosti v pôvodných a referenčných databázach. Takýmto porovnaním sa vykreslia ďalšie nedostatky dát (nahodnotenie a podhodnotenie niektorých častí škály hodnôt). V ideálnom prípade by boli početnosti hodnôt stupňa nepriepustnosti v pôvodnej IMD databáze zhodné s referenčnými IRD hodnotami.

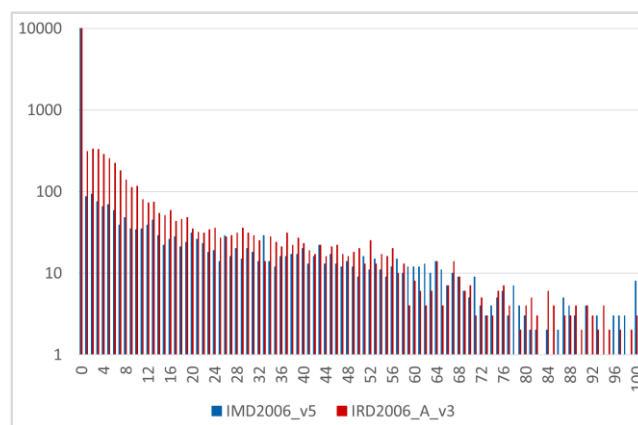
Validácia sa uskutočnila pomocou:

- histogramov rozdelenia hodnôt hustoty nepriepustnosti v IMD a IRD databázach
- histogramov porovnania početnosti nadhodnotených a podhodnotených pixelov podľa rozdielu stupňa nepriepustnosti v IMD a IRD databáze:

$$\text{nahodnotenie (podhodnotenie)} = IMD_i - IRD_i$$

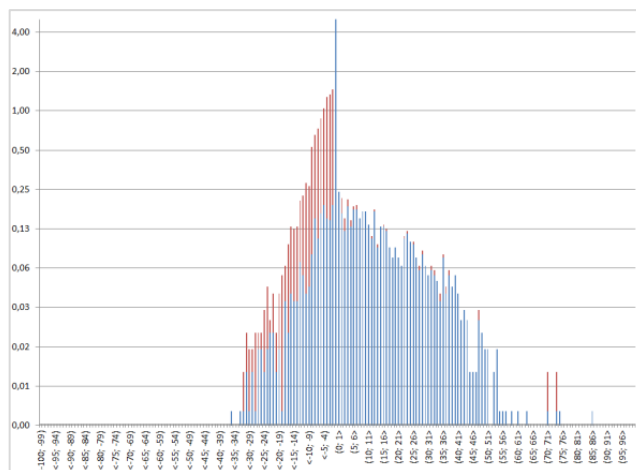
$$IMD_i = IMD \text{ hodnota v } i - \text{tom pixeli; } IMD_i \in (0, 1, 2, 3, \dots, 100)$$

$$IRD_i = IRD \text{ hodnota v } i - \text{tom pixeli; } IRD_i \in (0, 1, 2, 3, \dots, 100)$$



Graf 2 Príklad histogramu<sup>7</sup> rozdelenia hodnôt hustoty nepriepustnosti v IMD a IRD databázach

<sup>7</sup> Výška prvého stĺpca (modrého aj červeného) v Graf 2 je väčšia ako 10 000. Pre lepšiu čitateľnosť grafu je zobrazené maximum na osi y len 10 000.



Graf 3 Príklad histogramu porovnania početnosti nadhodnotených a podhodnotených pixelov podľa rozdielu stupňa nepriepustnosti v IMD a IRD databáze

Os x je v Graf 3 rozdelená:

- na záporné hodnoty ( $IMD_i - IRD_i < 0$ ) – podhodnotené pixely
- kladné hodnoty ( $IMD_i - IRD_i > 0$ ) – nadhodnotené pixely

Vzniknuté chyby nadhodnotenia a podhodnotenia v jednotlivých pixeloch sme rozdelili na dve skupiny (podľa odporúčaní Gallego a kol. 2016):

- závažne podhodnotené pixely a závažne nadhodnotené pixely (v Graf 3 sú vyznačené červenou farbou)
- menej závažne podhodnotené pixely a menej závažne nadhodnotené pixely (v Graf 3 sú vyznačené modrou farbou)

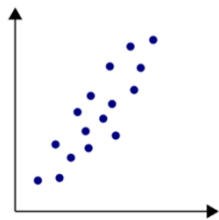
### 3.3 Korelačná analýza

Ďalšou metódou, ktorá bola v tejto dizertačnej práci využitá a ktorú navrhujeme využiť pri ďalších obdobných validáciách vrstiev s vysokým rozlíšením je korelačná analýza pomocou Pearsonovho, Kendallovho a Spearmanovho korelačného koeficientu.

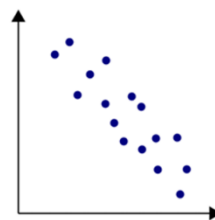
Korelačná analýza skúma tesnosť závislosti medzi kvantitatívnymi premennými, v našom prípade ukazuje smer a silu vzťahu medzi pôvodnými IMD a referenčnými IRD dátami. Korelačná analýza na rozdiel od regresie nevyjadruje príčinnno-následné vzťahy. To znamená, že premenná Y (v našom prípade IRD) nezávisí od premennej X (v našom prípade IMD), ale dve náhodné premenné Y a X sa spoločne menia. Regresná analýza predpokladá, že premenná Y je náhodná a premenná X fixná. Korelačné koeficienty

sú mierou lineárnej závislosti týchto dvoch premenných, teda dokazujú tematickú presnosť medzi našimi hodnotami IMD a IRD.

Korelačný koeficient je bezrozmerné číslo, ktoré nezávisí od toho, v akých jednotkách boli namerané hodnoty  $X$  a  $Y$ . Hodnoty všetkých korelačných koeficientov sa nachádzajú v intervale  $(-1, 1)$ . Kladné hodnoty koeficientu znamenajú kladnú koreláciu – premenné  $X$  a  $Y$  majú tendenciu meniť sa rovnakým smerom; záporné hodnoty znamenajú zápornú koreláciu – premenné  $X$  a  $Y$  majú tendenciu meniť sa rôznym smerom (Zvára 2003). Čím viac sa hodnota korelačného koeficientu blíži k hodnote 1 alebo  $-1$ , tým silnejší je lineárny vzťah medzi skúmanými veličinami, a teda tým je tematická presnosť IMD väčšia. Hodnoty blízko k 0 vyjadrujú slabnutie tohto lineárneho vzťahu, resp. žiadny vzťah, a teda tým menšia je presnosť IMD. Nemožno však usúdiť, že ak by korelačný koeficient nadobudol hodnotu nula, tak veličiny  $X$  a  $Y$  by boli nezávislé.



Obrázok 13 Kladná korelácia



Obrázok 12 Záporná korelácia

### A. Pearsonov výberový korelačný koeficient ( $r$ )

Pearsonov korelačný koeficient je najviac používaným a najdôležitejším korelačným koeficientom v súčasnosti. Používa sa v prípade, že premenné  $X$  a  $Y$  sú merané prinajmenšom v intervalovej škále. Je možné korelovať dve úplne rozličné veličiny, ako je napr. výška a hmotnosť. Nezávisí od mierky, v ktorej boli premenné namerané. Ak budeme najprv merať výšku v centimetroch a hmotnosť v kilogramoch a potom v palcoch (výška) a pondoch (hmotnosť), tak korelačný koeficient bude mať rovnakú hodnotu. Korelácia bude vysoká, ak bude dať nameranými bodmi v rovine preložiť regresnú priamku, pričom využijeme metódu najmenších štvorcov<sup>8</sup>.

<sup>8</sup> Metóda najmenších štvorcov znamená, že súčet obsahov štvorcov = súčet umocnených vzdialeností meraných bodov, je od regresnej priamky najmenší možný.

Jeho výpočet je realizovaný z  $n$  párových hodnôt  $(x_i, y_i)$ . Výberový Pearsonov korelačný koeficient  $r_{xy}$  definovaný vzťahom (Zvára 2003):

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

kde  $n$  je rozsah výberu;

$x_i$  = IMD hodnota v  $i$  – tom pixeli,  $x_i \in (0, 1, 2, 3, \dots, 100)$ ;

$y_i$  = IRD hodnota v  $i$  – tom pixeli,  $y_i \in (0, 1, 2, 3, \dots, 100)$ .

Pearsonov korelačný koeficient je parametrický korelačný koeficient, to znamená, že ak má byť správne interpretovaný, mali by sme predpokladať normálne rozdelenie skúmaných dát. V našom prípade je potrebné v tomto prípade zohľadniť príliš vysoký podiel dát, ktoré v oboch databázach (IMD aj IRD) predstavovali nulovú hodnotu, teda priepustné plochy bezo zmeny, kde nevznikla žiadna chyba (nadhodnotenia/podhodnotenia).

## **B. Kendallov korelačný koeficient (*tau*)**

Ďalším typom korelačného koeficientu je Kendallov korelačný koeficient, ktorý patrí do skupiny neparametrických koeficientov. Kendallov korelačný koeficient, ktorý vyjadruje silu závislosti medzi dvoma poradovými premennými, je založený na inverziách v poradí. Je možné ho interpretovať ako rozdiel medzi pravdepodobnosťou, že hodnoty dvoch premenných sú v tom istom poradí, oproti pravdepodobnosti, že tieto hodnoty nie sú v tom istom poradí.

Postup pri vypočítavaní Kendallovho korelačného koeficientu je nasledovný: Majme súbor s rozsahom  $n$  pozostávajúci z hodnôt  $x_i$  a  $y_i$ . Najprv je potrebné jednotlivé dvojice usporiadať tak, aby hodnoty  $x_i$  tvorili rastúcu postupnosť a zistiť, či rastú aj hodnoty  $y_i$ . Ak áno, jedná sa o kladnú asociáciu medzi veličinami X a Y, tzv. konkordanciu. Ak nie a hodnoty  $y_i$  majú klesajúcu (opačnú) postupnosť, ide naopak o zápornú asociáciu, tzv. diskordanciu.

- Ak  $x_i > x_j$  a  $y_i > y_j$  alebo  $x_i < x_j$  a  $y_i < y_j$ , tak ide o súhlasnú asociáciu – konkordanciu.
- Ak  $x_i > x_j$  a  $y_i < y_j$  alebo  $x_i < x_j$  a  $y_i > y_j$ , tak ide o nesúhlasnú asociáciu – diskordanciu.

Ak počet všetkých kladných asociácií označíme ako  $P$  a počet všetkých záporných asociácií označíme ako  $Q$ , tak rozdiel medzi nimi  $S$  ( $S = P - Q$ ) interpretuje jednoduchú mieru závislostí. To či je  $S$  kladné alebo záporné, o tom rozhoduje prevaha konkordancie, resp. diskordancie. Hodnoty, ktoré  $S$  nadobudne, budú závisieť od rozsahu  $n$ .

Na to, aby sme vypočítali konečnú hodnotu Kendallovho korelačného koeficientu, musíme ešte odhadnúť teoretickú hodnotu koeficientu  $tau$ . Následne bude Kendallov korelačný koeficient definovaný vzťahom (Hendl 2015):

$$tau_{xy} = \frac{S}{D} = \frac{P - Q}{D}$$

kde  $D$  je maximálny možný počet konkordancií, resp. diskordancií;  $D = \frac{n(n-1)}{2}$ ;  $n$  je rozsah výberu.

Hoci je výpočet Kendallovho korelačného koeficientu náročný, jeho interpretácia je priamočiarejšia ako interpretácia Spearmanovho korelačného koeficientu. Poznamenajme ešte, že Kendallov korelačný koeficient nadobúda vo všeobecnosti menšie absolútne hodnoty než Spearmanov korelačný koeficient (Markechová a kol. 2011).

### C. Spearmanov korelačný koeficient ( $rho$ )

Spearmanov korelačný koeficient patrí do skupiny neparametrických poradových korelačných koeficientov, rovnako ako Kendallov korelačný koeficient. Využíva princíp poradovej korelácie, ktorej podstata spočíva v tom, že sa nahradí realizácia náhodných veličín  $x_i$  a  $y_i$  ich poradím  $R_{x_i}$  a  $R_{y_i}$  vzhľadom k ostatným hodnotám výberu zoradeného podľa veľkosti. Ak nastane situácia, že sa hodnoty  $x_i$  alebo  $y_i$  opakujú viackrát, vtedy im je priradená priemerná hodnota v poradí, ktorú by mali mať, tak aby šli za sebou a boli rôzne.

Odhadom teoretickej hodnoty Spearmanovho korelačného koeficientu ( $rho$ ) je výberový korelačný koeficient  $r_{xy}$  definovaný vzťahom (Hendl 2015):

$$rho_{xy} = 1 - \frac{6 \sum_i^n D_i^2}{n(n^2 - 1)}$$

kde  $D_i$  je rozdiel poradí  $R_{x_i}$  a  $R_{y_i}$  a  $n$  je rozsah výberu;  $R_{x_1}, \dots, R_{x_n}$  sú poradia náhodných veličín  $x_1, \dots, x_n$ ;  $R_{y_1}, \dots, R_{y_n}$  sú poradia náhodných veličín  $y_1, \dots, y_n$ .



Spearmanov korelačný koeficient, zachytáva monotónne vzťahy, na rozdiel od Pearsonovho korelačného koeficientu, ktorý popisuje iba lineárne vzťahy. Medzi monotónne vzťahy patria vo všeobecnosti rastúce vzťahy, prípadne klesajúce. Koeficient je vďaka princípu poradovej korelácie odolný voči odľahlým hodnotám vo vzorke.

### 3.4 Výpočet totálnej absolútnej chyby normalizovanej a jej štruktúra

Absolútna a relatívna chyba patria medzi chyby experimentálne. Pri skúmaní presnosti nameraných rôznych vedeckých dát a informácií je každom prípade vhodné vypočítavať obidva typy chýb, takže je dobré rozdiel medzi nimi správne pochopiť a poznať spôsob ich výpočtu.

Absolútna chyba je údajom o neistote merania. Vyjadruje mieru toho, ako ďaleko sa meranie nachádza od skutočnej hodnoty. Absolútna chyba sa vypočítava vždy v rovnakých jednotkách ako bolo realizované meranie. Alternatívne, ak máme známu a vypočítanú hodnotu, pomocou absolútnej chyby môžeme vyjadriť, ako blízko je naše meranie k ideálnej hodnote. V tomto prípade sa chyba vyjadruje ako rozdiel medzi očakávanou a skutočnou hodnotou.

Na to, aby sme vypočítali relatívnu chybu, musíme najprv určiť absolútnu chybu. „Relatívna chyba vyjadruje, aká je veľká absolútna chyba v porovnaní s celkovou veľkosťou objektu, ktorý sme merali“ (Hazelwinkel 2001). Relatívnu chybu vyjadríme v podobe pomeru (zlomku) absolútnej chyby ku známej hodnote, alebo ju ešte vynásobíme hodnotou 100 a vyjadríme ju v percentách.

Pri validácii intenzitných máp sa obdobným spôsobom môžeme zamerať na výpočet hodnôt absolútnych chýb jednotlivých pixelov v IMD a IRD databázach, na základe ktorej vieme vypočítať totálnu absolútnu chybu (TAE):

$$TAE = \sum_i^n |IMD_i - IRD_i|$$

kde  $IMD_i$  je IMD hodnota podielu nepriepustných plôch v  $i$  – tom pixeli,

$i \in (0, 1, 2, 3, \dots, 20\ 000)$ ,

$IRD_i$  je IRD hodnota podielu nepriepustných plôch v  $i$  – tom pixeli,

$i \in (0, 1, 2, 3, \dots, 20\ 000)$ .

Nevýhodou takto vypočítanej chyby sa však ukazuje skutočnosť, že podľa nej by nebolo možné následne porovnať tematickú presnosť vrstiev dvoch území, ktoré by nemali rovnaký reálny podiel nepriepustných povrchov (Hurbánek, Rosina 2013). Preto je vhodným riešením práve výpočet vyššie spomínanej relatívnej chyby. Relatívnu chybu môžeme v tomto prípade interpretovať ako totálnu absolútnu chybu, ktorú normalizujeme, a to čo najpresnejšou hodnotou podielu nepriepustných plôch vo vybranej vzorke. Ak máme dostatočne veľkú vzorku, čo sme v našom prípade mali, získame tento údaj z referenčných dát tak, že všetky hodnoty pixelov vo vzorke spočítame. Výsledkom takto vypočítanej relatívnej chyby je totálna absolútna chyba normalizovaná (TAEn) vyjadrená v percentách:

$$TAEn = \frac{\sum_i^n |IMD_i - IRD_i|}{\sum_i^n IRD_i} \times 100 \%$$

kde  $IMD_i$  je IMD hodnota podielu nepriepustných plôch v  $i$  – tom pixeli (%)

$IRD_i$  je IRD hodnota podielu nepriepustných plôch v  $i$  – tom pixeli (%)

Na to, aby sme vedeli skúmať totálnu absolútnu chybu normalizovanú do hĺbky, rozhodli sme sa zhodnotiť aj jej štruktúru a vyjadriť pomocou nej vzniknuté chyby nadhodnotenia a podhodnotenia, ktoré sme následne rozdelili na menej závažné a závažnejšie.

Tabuľka 9 Typy pixelov

IMD hodnoty	IRD hodnoty	Rozdiel hodnôt	Klasifikácia pixelov
$IMD_i$	$IRD_i$	$IMD_i - IRD_i$	Typy pixelov
0	0	0	<b>AP</b> – správne klasifikované pixely ako priepustné „pervious agreement“
> 0	> 0	0	<b>AI</b> – správne klasifikované pixely ako nepriepustné „impervious agreement“
> 0	> 0	> 0	<b>MiO</b> – menej závažné nadhodnotenie „minor overestimation“
> 0	> 0	< 0	<b>MiU</b> – menej závažné podhodnotenie „minor underestimation“
> 0	0	> 0	<b>MaO</b> – závažnejšie nadhodnotenie „major overestimation“
0	> 0	< 0	<b>MaU</b> – závažnejšie podhodnotenie „major underestimation“

Keď chceme vrstvu IMD využívať, považujeme za vhodné poznať nie len celkovú chybu, ako to navrhuje vo svojich požiadavkách EEA, ale aj vzniknuté štyri chyby nezhody, resp. dva typy zhody, ktoré sú bližšie popísané v Tabuľka 9.

Ak chceme vedieť, ako často sa daná chyba, resp. zhoda, nachádza v rámci celej vzorky, tak je potrebné skúmať frekvenciu ich výskytu vzhľadom na celkový rozsah vzorky  $n$ .

Pre frekvenciu výskytu platí:

$$n_{AP} + n_{AI} + n_{MiO} + n_{MiU} + n_{MaO} + n_{MaU} = n$$

$$\frac{n_{AP}}{n} + \frac{n_{AI}}{n} + \frac{n_{MiO}}{n} + \frac{n_{MiU}}{n} + \frac{n_{MaO}}{n} + \frac{n_{MaU}}{n} = 100 \%$$

Okrem frekvencie výskytu (ne)zhôd sme sa rozhodli skúmať aj to, aký podiel majú jednotlivé typy (ne)zhôd v rámci celej totálnej absolútnej chyby, nakoľko sme chceli vedieť, akou mierou sa podieľajú na celkovej presnosti IMD.

Pre podiel na celkovej chybe  $TAEn$  platí:

$$TAEn = TAEn_{MiO} + TAEn_{MaO} + TAEn_{MiU} + TAEn_{MaU} \quad (TAEn_{AP} = 0 ; TAEn_{AI} = 0)$$

$$\text{pre MiO: } TAEn_{MiO} = \frac{\sum_i^n n_{MiO} |IMD_i - IRD_i|}{B} \cdot 100 (\%)$$

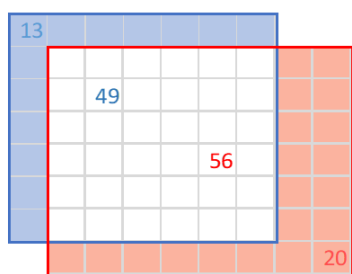
$$\text{pre MaO: } TAEn_{MaO} = \frac{\sum_i^n n_{MaO} |IMD_i - IRD_i|}{B} \cdot 100 (\%)$$

$$\text{pre MiU: } TAEn_{MiU} = \frac{\sum_i^n n_{MiU} |IMD_i - IRD_i|}{B} \cdot 100 (\%)$$

$$\text{pre MaU: } TAEn_{MaU} = \frac{\sum_i^n n_{MaU} |IMD_i - IRD_i|}{B} \cdot 100 (\%)$$

kde  $B = \frac{\sum_i^n IR_i}{n}$  je priemer referenčných hodnôt všetkých pixelov vo vzorke.

Výber tejto metódy pre validáciu IMD ovplyvnili aj nasledujúce predpokladané skutočnosti, na ktoré chceme čitateľa upozorniť. Predstavme si (Obrázok 14), že rozloha nesprávne klasifikovaných plôch podľa niektorej z vrstiev IMD bude na Slovensku 753,59 km<sup>2</sup> ( $TAEn = 1,54 \%$  z 49 036 km<sup>2</sup>), čo ale tvorí až 66 % z celkovej rozlohy nepriepustných plôch na Slovensku (2,33 % z 49 036 km<sup>2</sup>).



754 km<sup>2</sup> – rozloha nesprávne klasifikovaných povrchov na Slovensku podľa IMD2006\_v1 a IRD2006\_A\_v3, čo je 66 % z celkovej rozlohy nepriepustných povrchov (1143 km<sup>2</sup>)

1258 km<sup>2</sup> – rozloha nepriepustných povrchov na SK (IMD2006\_v1)

1143 km<sup>2</sup> – rozloha nepriepustných povrchov na SK (IRD2006\_A)

49 vs. 56

13 + 20 = 33

33 štvorcov = 66 % nesprávne klasifikovaných z nepriepustných povrchov (49 štvorcov)

Obrázok 14 Ilustračný obrázok

### 3.5 Ďalšie navrhované metódy

#### A. Fuzzy numerická metóda

Fuzzy numerická metóda je metóda, ktorá nevychádza z literatúry, ale je jedna z metód, ktorú popisuje a navrhuje Holandská agentúra pre hodnotenie životného prostredia na porovnanie práve rastrových máp. Hlavnou úlohou tejto agentúry je na vedeckej úrovni poskytovať holandskej vláde odpovede na širokú škálu otázok ohľadom životného prostredia. Samozrejme, poradenstvo je často založené na priestorovej analýze a modelovaní. Porovnanie máp je opakujúci sa proces, ktorý je nevyhnutný na kvantifikáciu, vizualizáciu a porozumenie výsledkov analýzy a predchádza procesom modelovania, overovania, validácie a kalibrácie. V spolupráci s Výskumným inštitútom pre vedomostné systémy bol vyvinutý softvérový nástroj, ktorý podporuje tento typ analýzy: Súprava na porovnanie máp (The Map Comparison Kit – MCK) (Visser a kol. 2004), v rámci ktorého bola táto metóda implementovaná. Metóda sa riadi logikou Fuzzy Kappa (Hagen 2003, Hagen-Zanker a kol. 2005), ale je upravená tak, aby fungovala so spojitými a nie kategórickými údajmi. Metóda je dokumentovaná v používateľskej príručke MCK (Hagen-Zanker a kol. 2006).

Hlavnou podstatou fuzzy numerickej štatistiky je, že podobnosť na jednom mieste je daná mierou, do akej je bunka (pixel) v jednej mape podobná jej protipólu v druhej mape, alebo podobná bunkám v jej priamom susedstve berúc ohľad na vzdialenosť susedov. Čím je väčšia vzdialenosť tým susedná bunka menej zaváži na danej podobnosti.

Rovnice 1 – 4 opisujú postup výpočtov. Treba poznamenať, že rovnice 1 – 3 sú všeobecné, pretože v nich vystupuje neznáma funkcia podobnosti  $f(a, b)$ , ktorú si môžeme zvoliť.

Rovnica 4 je konkrétna funkcia podobnosti, ktorú sme aplikovali. Prirodzene, výsledok porovnania bude závisieť od konkrétnej váhovej funkcie, ktorá sa aplikuje a od jej príslušných parametrov. To zavádza stupeň subjektivity.

$$s_i(A, B) = \max_j^N (f(A_i, B_j) * w(d_{i,j})) \quad 1.$$

$$S_i(A, B) = \min(s_i(A_i, B_j), s_i(B, A)) \quad 2.$$

$$S(A, B) = \frac{1}{n} \sum_{i=1}^n S_i(A, B) \quad 3.$$

$$f(a, b) = 1 - \frac{|a-b|}{\max(|a|, |b|)} \quad 4.$$

Touto metódou navrhujeme porovnávať tematickú presnosť IMD v referenčných štvorcoch s rozmermi  $6 \times 6$  km podľa Hurbánka (2010) a Lajčákovej (2013).

Treba poznamenať, že hoci je táto (a sedem ďalších metód) k dispozícii na porovnanie máp v spomínanej príručke MCK, v praxi sa postupy stále spoliehajú na hodnotenie bunky po bunke. Je to uvedené nielen v práci Wealands a kol. (2005), ale potvrdené aj v mnohých prácach o modelových aplikáciách, ako sú: Ahrens et al (1998), Bishop a kol. (2005), Garen a Marks (2005), Liu a kol. (1997), Strasser a Mauser (2001), Viscarra Rossel a Walter (2004). V dôsledku toho boli inovatívne metódy sotva stanovené a je ťažké kvalifikovať ich mieru využitia v praktických aplikáciách. Pravdepodobným príčinným faktorom je, že priestorové explicitné porovnania sú komplikované; nielen z koncepčného hľadiska, ale aj kvôli technickej implementácii v geografickom informačnom systéme alebo inom softvéri.

Dostupnosť nástrojov, ako je MCK, môže v súčasnej výskumnej praxi viesť k rozsiahlejšiemu šíreniu a nasýteniu sľubných metód. Napr. po preskúmaní metódy Fuzzy Kappa (Hagen 2003) nasledovalo vytvorenie softvéru na jej použitie (Visser a de Nijs 2005) a nakoniec výskumníci uplatňujú túto metódu ako súčasť validačného procesu svojich modelov (Prasad a kol. 2006, Ménard a Marceau 2006). Ďalším príkladom šírenia prostredníctvom implementácie softvéru je Idrisi GIS balík, ktorý okrem iného ponúka metodológiu validácie od Pontius a kol. (2004).

## B. Porovnanie intenzít „Intensity-Scale Approach“

Ďalšou metódou z príručky MCK je metóda založená na skúmaní rozdielností máp vzhľadom na ich priemernú štvorcovú („na druhú“) odchýlku, tzv. MSE (Mean Squared Error). Porovnávajú sa pritom intenzitné mapy tvorené len jednotkami (v našom prípade nepriepustné povrchy) a nulami (priepustné povrchy) pre každý pixel, prislúchajúce konkrétnej prahovej hranici, tzv. „thresholdu“. Aj túto metódu považujeme za hodnú preskúmania v súvislosti s porovnávaním tematickej presnosti IMD v referenčných štvorcoch s rozmermi  $6 \times 6$  km podľa Hurbánka (2010) a Lajčákovej (2013) pri použití rôznych hodnôt „thresholdov“.

Výsledkom metódy je funkcia od „thresholdu“, ktorá zobrazuje MSE odchýlku prislúchajúcu konkrétnej prahovej hranici. My by sme chceli nájsť taký „threshold“, kedy by boli pôvodná vrstva a nami odvodená referenčná vrstva najpodobnejšie.

Porovnanie intenzít prebieha v dvoch fázach. Prvá z nich je tzv. prípravná. V nej si obidve vrstvy trochu upravíme. V prvom kroku do obidvoch vrstiev pridáme trochu šumu. Ku každej hodnote pripočítame náhodnú realizáciu z rovnomerného rozdelenia s amplitúdou rovnou polovici najmenšieho nenulového rozdielu v mapách. Táto malá zmena obmení informácie obsiahnuté v pixeloch len nepatrne, zanedbateľne. Dosiahneme tým však to, že každá hodnota sa bude vyskytovať len raz.

Druhú časť prvej prípravnej fázy nazývame re-kalibrácia. Všetky hodnoty v pôvodnej vrstve nahradíme referenčnou hodnotou. Urobíme to pritom tak, že každú hodnotu jednej vrstvy vymeníme za takú hodnotu druhej, ktorá má rovnakú empirickú kumulatívnu pravdepodobnosť výskytu. Inak povedané, ak si hodnoty zoradíme od najmenej po najväčšiu (v každej mape osobitne), hodnoty s rovnakým poradovým číslom tvoria pár použitý vo výmene. Tieto páry sú jedinečné a jednoznačne dané vzhľadom na to, že všetky hodnoty sú navzájom rozdielne. To sme docielili pridávaním šumu.

K ďalšej časti úprav vrstiev je potrebné zvoliť „threshold“. Následne všetky hodnoty, ktoré ho presahujú prepíšeme na hodnotu 1, všetky ostatné na 0, a to v pôvodnej, ako aj referenčnej vrstve. V druhej fáze vypočítame MSE prislúchajúce danému „thresholdu“. Vytvoríme si mapu odchýlok, ako rozdiel predpripravených vrstiev. Hodnoty takejto novej vrstvy môžu nadobúdať len hodnoty 1, 0 a -1. To závisí od toho, či bola nepriepustnosť pozorovaná v oboch mapách súčasne, alebo len v jednej z nich, poprípade žiadnej.

MSE je priemer odchýlok na druhú. Všetky hodnoty rozdielovej vrstvy umocníme na druhú a následne z týchto hodnôt vypočítame aritmetický priemer. Výsledok je číslo medzi 0 a 1. Čím je bližšie k nule, tým je rozdiel vrstiev menší.

## 4 VÝSLEDKY PRÁCE

Hodnotenie tematickej presnosti Európskej vrstvy hustoty nepriepustnosti sme vykonali pomocou štatistických metód, ktoré navrhujeme využívať pri obdobných validáciách intenzitných máp. Tieto metódy sú bližšie samostatne popísané v kapitole 3 METODIKA PRÁCE A METÓDY SKÚMANIA. Na porovnanie dát jednotlivých vrstiev IMD sme v tejto dizertačnej práci využili referenčné databázy typu A, nakoľko sa táto voľba vo všetkých prípadoch javila ako presnejšia. Uvažovali sme teda s možnosťou, že nepriepustné plochy s autochtóнным materiálom sú podľa nejasnej interpretácie EEA (2009) považované za priepustné plochy a priradili sme im hodnotu 0. Navzájom sme porovnali obidva referenčné roky (2006, 2009) obidvoch záujmových území (Slovensko, Česko). Postupne sme sa zamerali na hodnotenie najstarších dostupných verzií IMD2006\_v1 a IMD2006\_v2, potom na hodnotenie starších vylepšených dostupných verzií IMD2006\_v3 a IMD2009\_v1, aj hodnotenie posledných dostupných vylepšených verzií IMD2006\_v5 a IMD2009\_v3 (viac o vývoji vrstiev úvode kapitoly 3 METODIKA PRÁCE A METÓDY SKÚMANIA), ktoré sme tiež navzájom v tomto poradí porovnávali.

Hodnoty IMD a IRD vyjadrujúce mieru nepriepustnosti sme porovnávali na úrovni jednotlivých pixelov. Ak by nastal ideálny prípad, tak by sa tieto hodnoty rovnali. V prípade však, že hodnota IMD bola väčšia ako hodnota IRD, bolo odhalené nadhodnotenie. V opačnom prípade sme objavili mieru podhodnotenia.

Nasledujúce Obrázky 15 a 16 zobrazujú príklady nadhodnotených a podhodnotených pixelov.



Obrázok 15 Nadhodnotené pixely z náhodného výberu 20 000 vzoriek IMD2006  
(Zdroj dát: EUROSENSE a GEODIS SLOVAKIA 2005 – 2007)

V pixeli na Obrázku 15a bolo zistené nadhodnotenie o 40 bodov (IMD hodnota 86, IRD hodnota 46), na Obrázku 15b o 36 bodov (IMD hodnota 98, IRD hodnota 62),



na Obrázku 15c o 34 bodov (IMD hodnota 93, IRD hodnota 59). Na Obrázku 15d je znázornený špecifický prípad – kameňolom, ktorý mal byť podľa EEA (2009), vylúčený z posudzovania ako nepriepustná plocha. Je však isté, že sa to nepodarilo vo všetkých prípadoch, keďže IMD hodnota bola 100 a IRD hodnota 0.



Obrázok 16 Podhodnotené pixely z náhodného výberu 20 000 vzoriek IMD2006  
(Zdroj dát: EUROSENSE a GEODIS SLOVAKIA 2005 – 2007)

Na Obrázku 16 sa nachádzajú príklady pixelov, kde sme objavili chyby podhodnotenia, kedy hodnota IMD bola menšia ako hodnota IRD. V pixeli na Obrázku 16a bolo zistené podhodnotenie o 10 bodov (IMD hodnota 0, IRD hodnota 10), na Obrázku 16b o 12 bodov (IMD hodnota 0, IRD hodnota 12), na Obrázku 16c o 30 bodov (IMD hodnota 6, IRD hodnota 36), na Obrázku 16d o 43 bodov (IMD hodnota 0, IRD hodnota 43).

Nasledujúca Tabuľka 10 zobrazuje reálny stav počtu nadhodnotených a podhodnotených pixelov v jednotlivých verziách IMD. Už na prvý pohľad je zjavné, že vo všetkých vrstvách dominujú chyby podhodnotenia oproti chybám nadhodnotenia (pravý stĺpec tabuľky má vždy vyššiu hodnotu v porovnaní s pravým stĺpcom).

Ak porovnáme vývoj počtu podhodnotených pixelov (1. stĺpec) na Slovenku aj v Česku vyskytujúcich sa v jednotlivých databázach z roku 2006, vidíme, že v oboch krajinách v najnovšej dostupnej vylepšenej verzii IMD2006\_v5 (riadok 4 v porovnaní s 1., 2. a 3. riadkom a riadok 10 v porovnaní so 7., 8. a 9. riadkom) výrazne narástol počet podhodnotených pixelov. A naopak, ak porovnáme vývoj počtu nadhodnotených pixelov (2. stĺpec), tak v oboch krajinách tento počet v poslednej dostupnej verzii pre referenčný rok 2006 výrazne poklesol. Rovnaký prípad nastal v oboch krajinách aj v referenčnom roku 2009 (porovnanie 5. a 6. riadka, porovnanie 11. a 12. riadka). Aj tu sa vo vylepšených verziách potvrdila zvýšená miera podhodnotených pixelov a výrazne znížená miera nadhodnotených pixelov v porovnaní s predchádzajúcou dostupnou verzou pre daný referenčný rok.

Tabuľka 10 Porovnanie počtu nadhodnotených a podhodnotených pixelov v jednotlivých verziách IMD  
(Zdroj dát: Autorka)

Poradové číslo	Štát_IMDdatabáza_IRDdatabáza	Počet podhodnotených pixelov	Počet nadhodnotených pixelov
1.	SK_IMD2006_v1_IRD2006_v3	2087	882
2.	SK_IMD2006_v2_IRD2006_v3	2099	857
3.	SK_IMD2006_v3_IRD2006_v3	2058	897
4.	SK_IMD2006_v5_IRD2006_v3	2661	290
5.	SK_IMD2009_v1_IRD2009_v3	1971	954
6.	SK_IMD2009_v3_IRD2009_v3	2655	264
7.	CZ_IMD2006_v1_IRD2006_v3	2836	1183
8.	CZ_IMD2006_v2_IRD2006_v3	2826	1102
9.	CZ_IMD2006_v3_IRD2006_v3	2821	1095
10.	CZ_IMD2006_v5_IRD2006_v3	3337	594
11.	CZ_IMD2009_v1_IRD2009_v3	3050	1136
12.	CZ_IMD2009_v3_IRD2009_v3	3590	568

Pri prvotnom pohľade na tak jednoduchú a výstižnú Tabuľku 10 sa nám môže zdať, že každá jedna verzia spĺňa 85 % podľa tendrov požadovanú presnosť, a teda že počet chybné klasifikovaných pixelov (súčet v každom riadku) tvorí menej ako 15 % z 20 000 klasifikovaných pixelov. Ak však skutočne spočítame počet podhodnotených a nadhodnotených pixelov v každom aj jednom riadku a vyjadríme ich ako podiel zo všetkých 20 000 hodnotených pixelov, tak túto podmienku spĺňajú všetky databázy len na Slovensku, kde takto chybné klasifikované pixely tvoria približne 14,7 %. V Česku takto vyjadrený počet chybné klasifikovaných pixelov predstavuje približne 20,1 %. Takéto vyjadrenie celkovej presnosti by bolo však veľmi jednoduché a nepresné, ak by sa vôbec dalo hovoriť o nejakej výpovednej hodnote týchto dvoch percentuálne vyjadrených chýb. Spomeňme si na príklad s lesmi vysvetľujúci význam presnosti používateľa a producenta, ktorý sme demonštrovali v kapitole 1.3.2 Definície základných pojmov, kde Rosina (2010) poukazuje, že len vypočítaný ukazovateľ celkovej presnosti nie je vhodnou metódou hodnotenia presnosti intenzitnej mapy. Jedno je však jasné. Databáza IMD na Slovensku sa javí ako presnejšia v každom jednom prípade v porovnaní s databázami v Česku pre jednotlivé roky. Na základe uvedených skutočností sme hodnotili tematickú presnosť Európskej vrstvy hustoty nepriepustnosti nami navrhovanými štatistickými metódami a ich výsledky sme interpretovali v nasledujúcich podkapitolách.

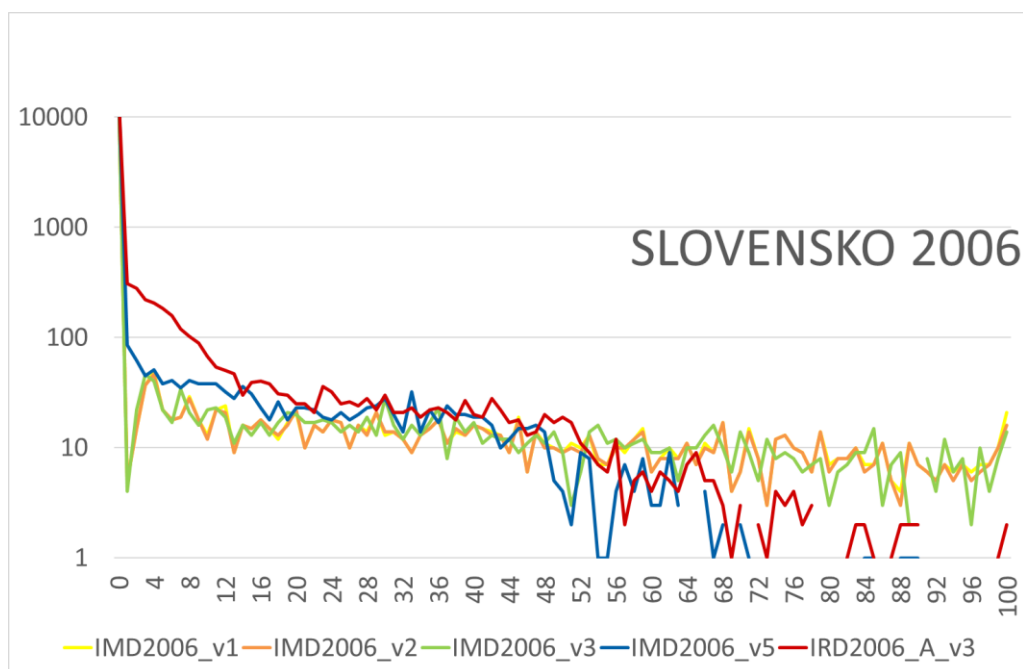
## 4.1 Hodnotenie tematickej presnosti metódou histogramov

Porovnaním počtu nadhodnotených a podhodnotených pixelov vieme s istotou povedať, ktorá z chýb (nadhodnotenie/podhodnotenie) sa v určitej vrstve nachádzala viac. Takýmto porovnaním (Tabuľka 10) však nedokážeme interpretovať, či sa daná miera nadhodnotenia, resp. podhodnotenia vyskytovala viac v pixeloch s nižšou hodnotou hustoty nepriepustnosti, alebo naopak, s vyššou hodnotou hustoty nepriepustnosti. Preto sme túto dodatočnú informáciu získali práve vytvorením histogramov rozdelenia hodnôt hustoty nepriepustnosti v pôvodných a referenčných databázach a ich porovnaním.

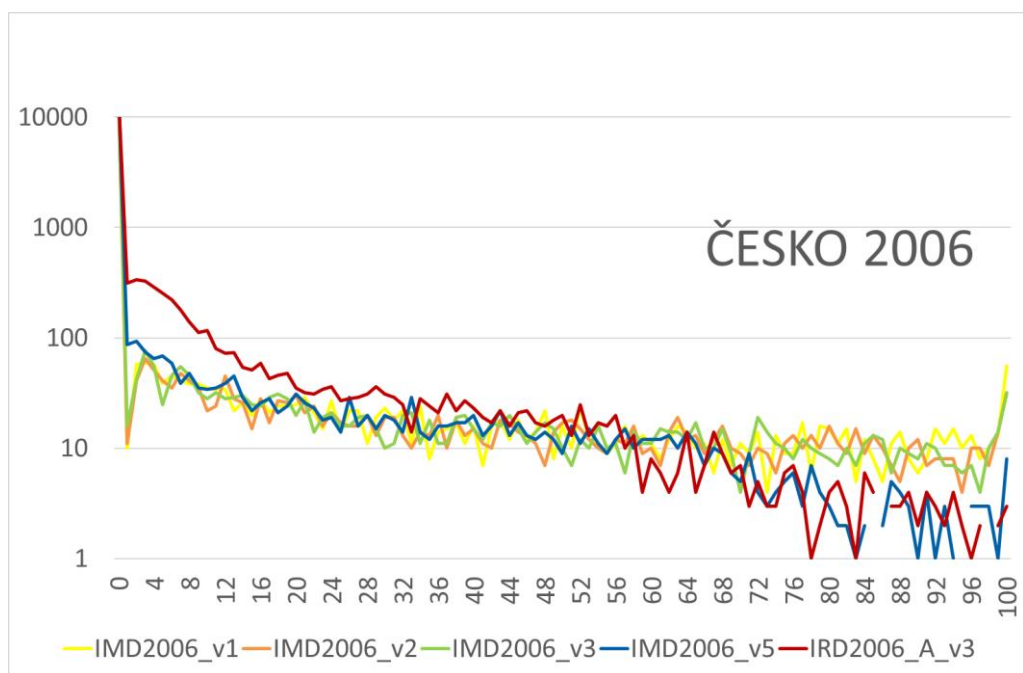
Nasledujúce štyri grafy popisujú najprv stav za rok 2006 (Graf 4 a 5) a následne za rok 2009 pre obidve krajiny (Graf 6 a 7). Pri interpretácii je dôležité upriamiť pozornosť hlavne na červenú krivku, ktorá v každom z grafov reprezentuje skutočný stav rozdelenia hodnôt hustoty nepriepustnosti. Podľa nej je najväčší jednoznačne výskyt nulových hodnôt. Veľký výskyt nízkych hodnôt postupne klesá v oboch rokoch aj krajinách.

Pri starších verziách IMD2006 (v1, v2, v3) a IMD2009 (v1) dominuje relatívne rovnomerné rozloženie početnosti všetkých hodnôt hustoty nepriepustnosti. Tým dochádza k výraznému podhodnoteniu výskytu nízkych hodnôt a nadhodnoteniu vyšších hodnôt v porovnaní s červenou krivkou. Tento prelom medzi nadhodnotením a podhodnotením nastal v starších verziách v oboch krajinách približne pri hodnotách hustoty nepriepustnosti 50 – 60 % (červená krivka v Grafoch 4 a 5 pretína žltú, oranžovú a zelenú krivku; v Grafoch 6 a 7 pretína žltú krivku). Z toho vyplýva, že v pixeloch s nižšími hodnotami hustoty nepriepustnosti sa zaokrúhľovalo smerom nadol a tým došlo ku zjavnému podhodnocovaniu a v pixeloch s vyššími hodnotami hustoty nepriepustnosti sa zaokrúhľovalo smerom nahor, čím sa prejavilo naopak výrazné nadhodnotenie.

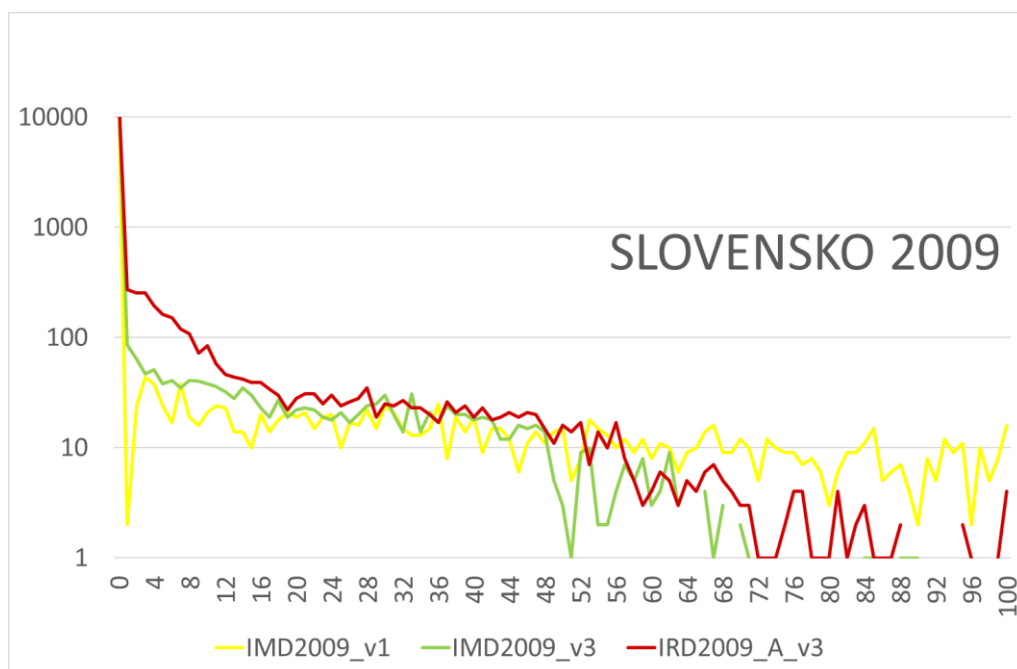
Chýbajúce časti červenej krivky (IRD2006\_A\_v3) vo všetkých štyroch grafoch znamenajú, že sa isté hodnoty hustoty nepriepustnosti v jednotlivých verziách IMD nenachádzali. Takéto medzery nájdeme aj na modrej krivke (IMD2006\_v5) a zelenej krivke (IMD2009\_v3) v oboch krajinách, čím sa tiež vykreslila absencia istých hodnôt hustoty nepriepustnosti v novších verziách pre obidva referenčné roky, na rozdiel od všetkých spomínaných starších verzií, ktoré takéto prerušenie nemajú. V prípade, že krivky starších verzií sú neprerušované v miestach, kde červená krivka referenčnej databázy nie je, nastáva závažné nadhodnotenie, kedy bola pixelu, ktorý neobsahuje žiadne nepriepustné plochy priradená vysoká hodnota hustoty nepriepustnosti.



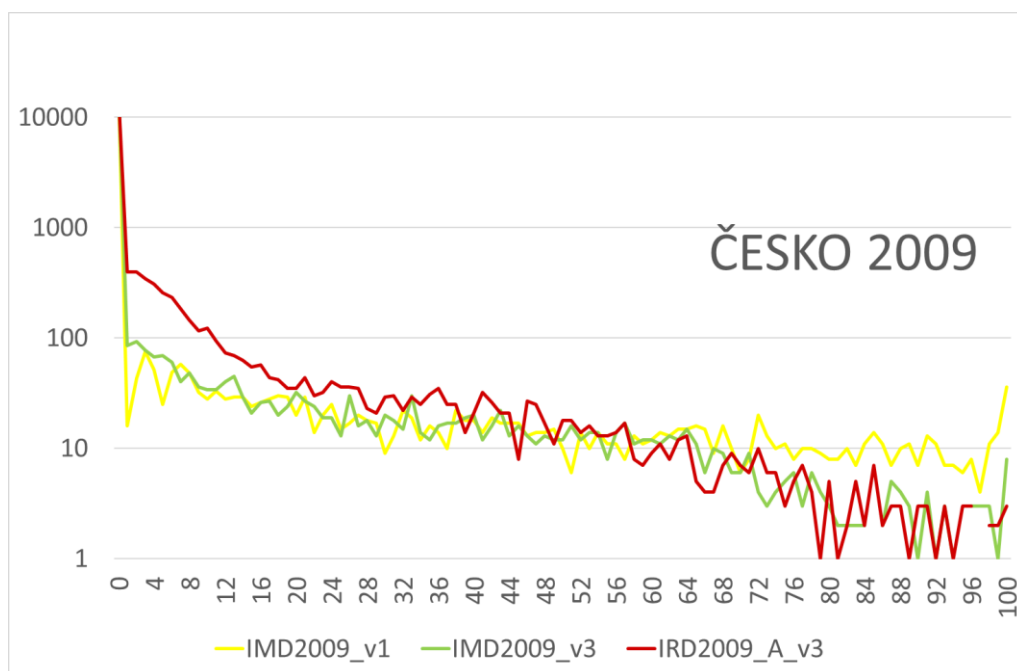
Graf 4 Histogram rozdelenia hodnôt hustoty nepriepustnosti vo všetkých dostupných databázach IMD2006 a v referenčnej databáze IRD2006\_A\_v3 pre Slovensko  
(Zdroj dát: Autorka)



Graf 5 Histogram rozdelenia hodnôt hustoty nepriepustnosti vo všetkých dostupných databázach IMD2006 a v referenčnej databáze IRD2006\_A\_v3 pre Česko  
(Zdroj dát: Autorka)



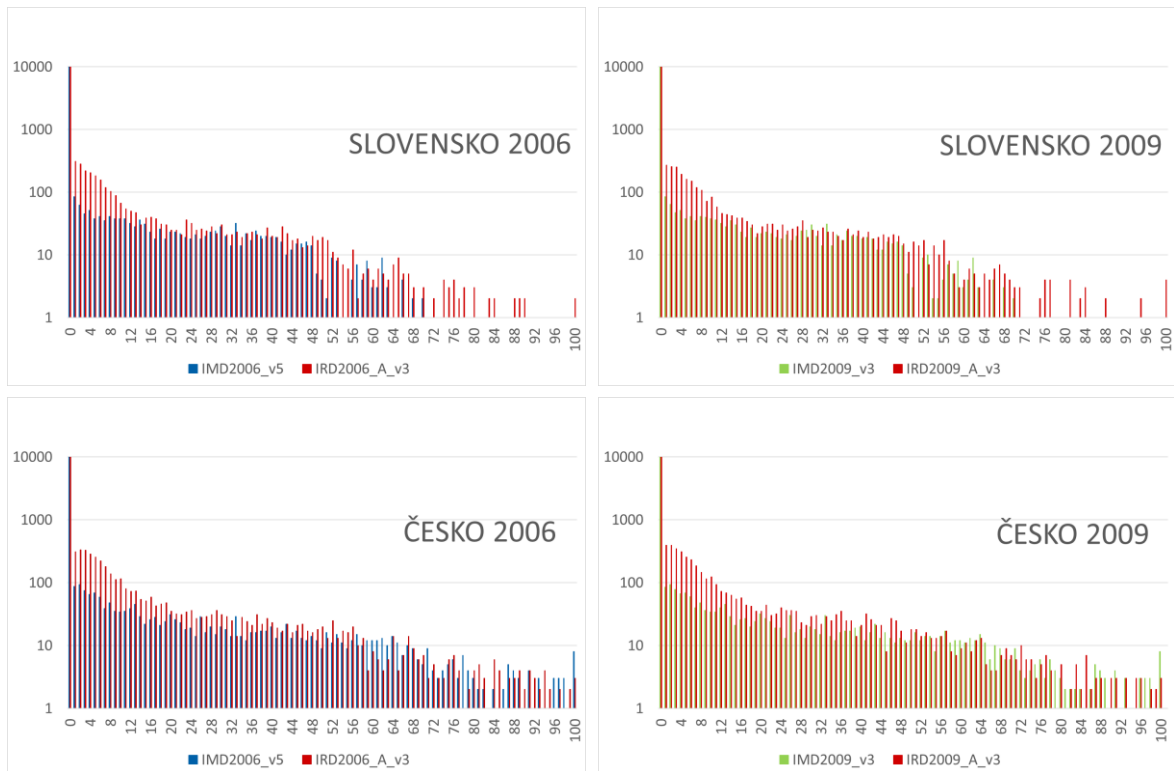
Graf 6 Histogram rozdelenia hodnôt hustoty nepriepustnosti vo všetkých dostupných databázach IMD2009 a v referenčnej databáze IRD2009\_A\_v3 pre Slovensko  
(Zdroj dát: Autorka)



Graf 7 Histogram rozdelenia hodnôt hustoty nepriepustnosti vo všetkých dostupných databázach IMD2009 a v referenčnej databáze IRD2009\_A\_v3 pre Česko  
(Zdroj dát: Autorka)

Podľa Grafov 8 (červené stĺpce prevyšujú modré a zelené) v obidvoch najnovších verziách (IMD2006\_v5, IMD2009\_v3) sa početnosť všetkých nenulových hodnôt nepriepustnosti

(takmer všetkých menších hodnôt v oboch krajinách a väčšiny väčších hodnôt na Slovensku) zdá byť podhodnotená, zatiaľ čo frekvencia nulových hodnôt nepriepustnosti je nadhodnotená.

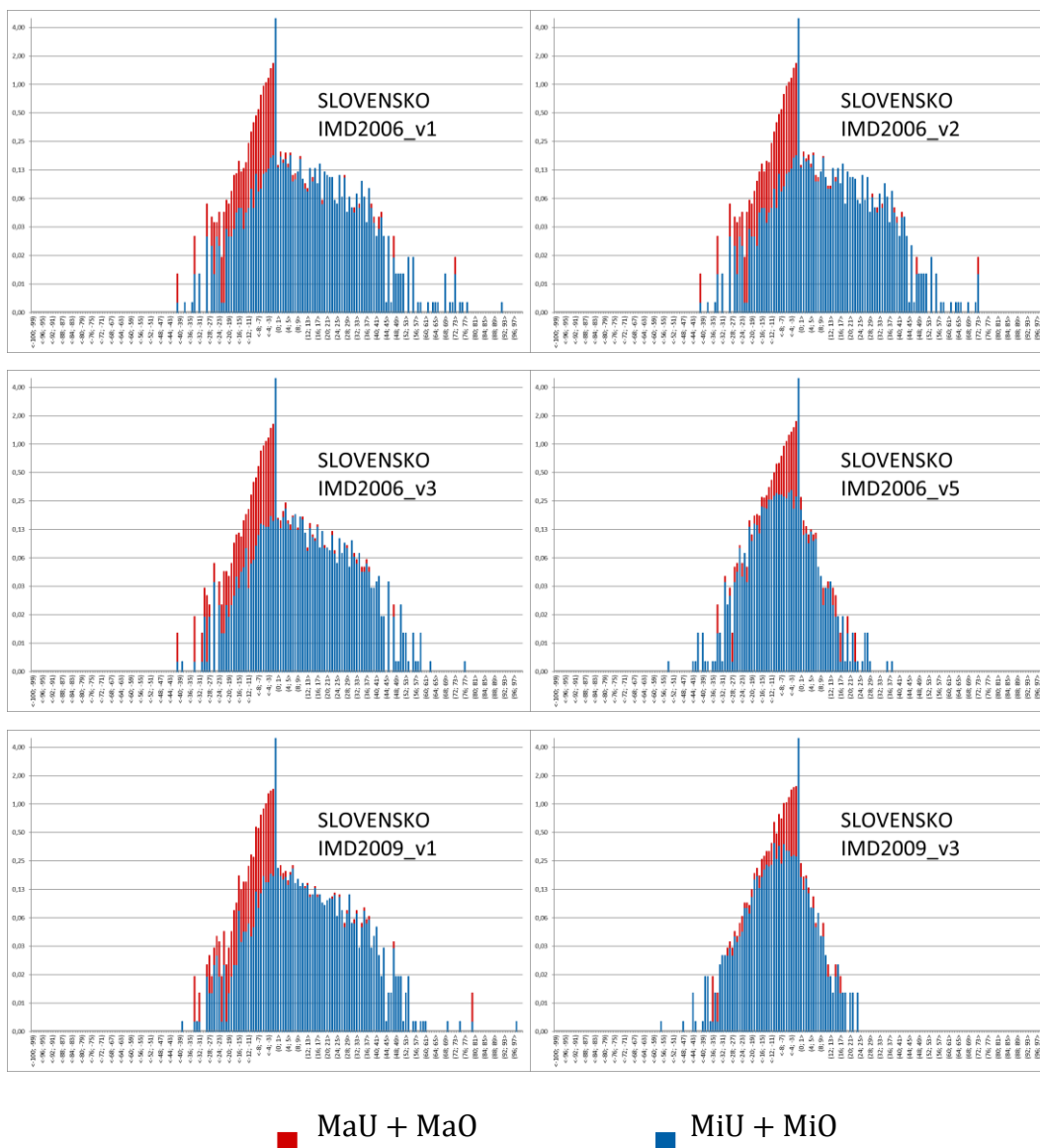


Graf 8 Histogramy rozdelenia hodnôt hustoty nepriepustnosti v databázach IMD2006\_v5 a IMD2009\_v3 a v referenčných databázach pre Slovensko a Česko

Histogramy rozdelenia hodnôt hustoty nepriepustnosti nám vykreslili výskyt nadhodnotených a podhodnotených pixelov v niektorých častiach škál hodnôt v jednotlivých databázach. Nepovedali nám však nič o ich početnosti pri jednotlivých úrovniach hustoty nepriepustnosti. Preto sme zostrojili aj histogramy, ktoré porovnávajú túto početnosť nadhodnotených a podhodnotených pixelov podľa rozdielu stupňa nepriepustnosti v IMD a IRD databázach.

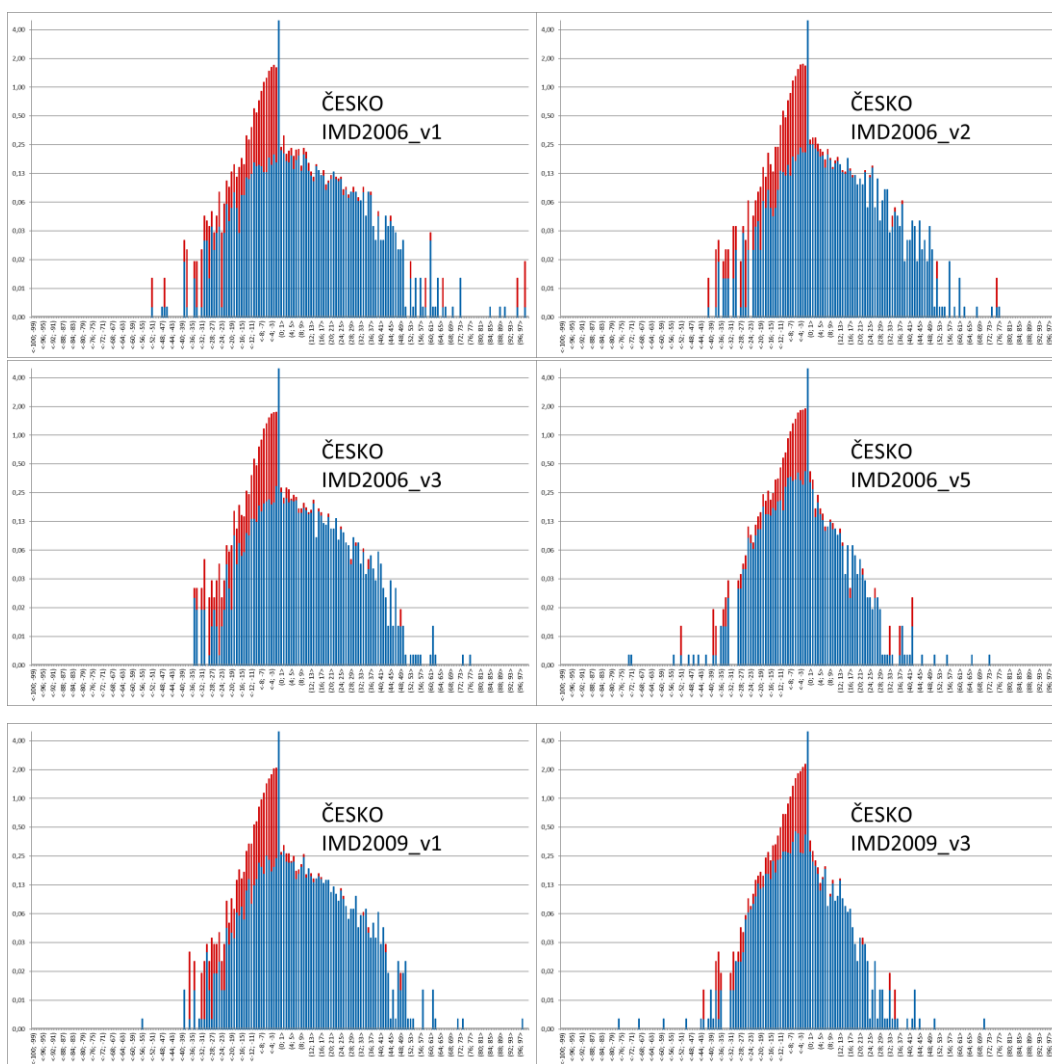
Vypočítané chyby nadhodnotenia a podhodnotenia v jednotlivých vzorkách databáz sú v Grafoch 9 a 10 rozdelené na dve skupiny (podľa Tabuľky 9): chyby závažnejšieho podhodnotenia (MaU) a závažnejšieho nadhodnotenia (MaO) sú vyznačené červenou farbou, a chyby menej závažného podhodnotenia (MiU) a menej závažného nadhodnotenia (MiO) – spolu so vzorkami s nulovým rozdielom medzi IMD a IRD – sú vyznačené modrou farbou. Zatiaľ čo záporné hodnoty na osi x reprezentujú vzorky s chybami podhodnotenia ( $IMD_i - IRD_i < 0$ ), kladné hodnoty na osi x reprezentujú vzorky s chybami nadhodnotenia

( $IMD_i - IRD_i > 0$ ). Čím ďalej od nuly sa v grafoch posúvame oboma smermi (doľava aj doprava), tým veľkosť chýb narastá. Čím sme bližšie k nule, tým je veľkosť chýb menšia.



Graf 9 Histogramy početností rôznych hodnôt chýb nadhodnotenia a podhodnotenia získané porovnaním stupňa hustoty nepriepustnosti IMD databáz s príslušnými databázami IRD na Slovensku (Zdroj dát: Autorka)

Na prvý pohľad sa vývoj Grafov 9 a 10, teda postupný vývoj početností zachytených chýb, v oboch krajinách veľmi podobá. Pomerne široké grafy starších verzií sa postupne zužujú smerom k novším verziám. To signalizuje zjavný úbytok väčších chýb na oboch stranách – smerom k novším upraveným verziám ubúdajú nielen väčšie chyby nadhodnotenia a väčšie chyby podhodnotenia, ale je zjavný aj úbytok závažnejších nadhodnotení aj závažnejších podhodnotení (v grafoch ubúda červená farba).



■ MaU + MaO      ■ MiU + MiO

Graf 10 Histogramy početností rôznych hodnôt chýb nadhodnotenia a podhodnotenia získané porovnaním stupňa hustoty nepriepustnosti IMD databáz s príslušnými databázami IRD v Česku (Zdroj dát: Autorka)

Vo všetkých grafoch je viditeľná výrazná prevaha červenej farby na ľavej strane grafov, čo značí vysoký výskyt závažnejších chýb podhodnotenia, ktoré sa vyskytujú prevažne pri menších chybách, menej pri väčších chybách podhodnotenia (červená farba ubúda v smere od nuly k vyšším hodnotám). Väčšie chyby podhodnotenia v grafoch dokonca úplne absentujú. Oproti tomu dominuje prevaha väčších chýb nadhodnotenia. V niektorých prípadoch dokonca závažnejších (na pravej strane ďalej od nuly miestami vybehla červená farba).

O početnosti chýb nadhodnotenia a podhodnotenia teda platí, že vo všetkých verziách je väčší počet menších chýb podhodnotenia (pričom dominujú závažnejšie chyby) ako počet menších chýb nadhodnotenia. Ďalej platí, že počet väčších chýb nadhodnotenia je výrazne



väčší ako počet väčších chýb podhodnotenia. O výskyte rôznych typov chýb, hlavne vo všetkých starších verziách IMD, teda možno vysloviť úsudok, že obsahujú veľa pixelov s malými chybami podhodnotenia a málo pixelov s veľkými chybami nadhodnotenia.

Čo sa týka postupného vývoja početnosti chýb nadhodnotenia a podhodnotenia od starších vrstiev IMD2006 (v1, v2, v3) a IMD2009 (v1) k novším vrstvám IMD2006\_v5 a IMD2009\_v3 je viditeľný úbytok hlavne väčších chýb nadhodnotenia v obidvoch krajinách. Zjavný je aj úbytok väčších závažnejších chýb podhodnotenia.

## 4.2 Hodnotenie tematickej presnosti korelačnou analýzou

Na ďalšie hodnotenie tematickej presnosti vrstvy hustoty nepriepustnosti sme využili Pearsonov, Kendallov a Spearmanov korelačný koeficient.

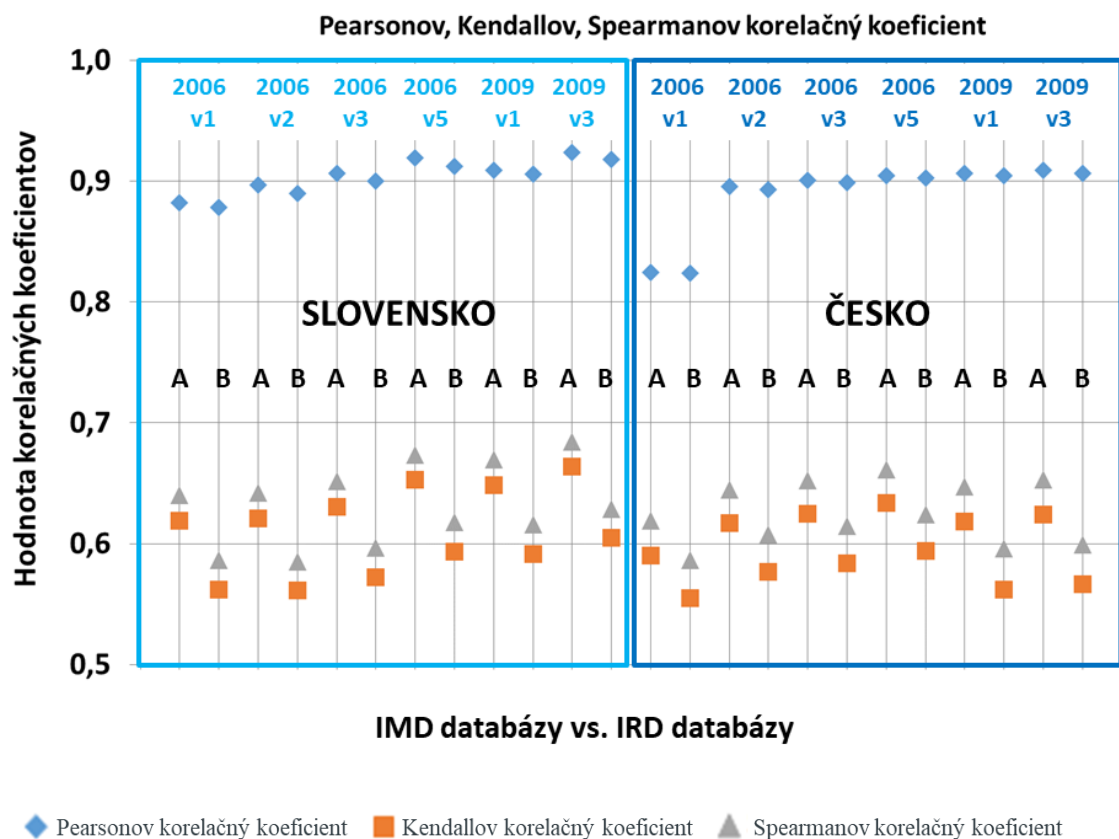
Korelačná analýza kvantifikuje tesnosť a charakter vzťahu medzi referenčnými a validovanými dátami. Čím viac sa hodnota koeficientu približuje k hodnote 1, tým je tematická presnosť IMD väčšia. Vzhľadom na skutočnosť, že vstupné dáta nemajú normálne rozdelenie, získané hodnoty Pearsonovho korelačného koeficientu majú, na rozdiel od zvyšných dvoch koeficientov, len obmedzenú výpovednú hodnotu.

Tabuľka 11 Vypočítané korelačné koeficienty pre vzťahy medzi IMD a IRD\_A/B na území Slovenska (Zdroj dát: Autorka)

SK_IMDdatabáza_IRDdatabáza	Pearson's r	Kendall's tau	Spearman's rho
SK_IMD2006_v1_IRD2006_A_v3	0,88	0,62	0,64
SK_IMD2006_v1_IRD2006_B_v3	0,88	0,56	0,59
SK_IMD2006_v2_IRD2006_A_v3	0,90	0,62	0,64
SK_IMD2006_v2_IRD2006_B_v3	0,89	0,56	0,59
SK_IMD2006_v3_IRD2006_A_v3	0,91	0,63	0,65
SK_IMD2006_v3_IRD2006_B_v3	0,90	0,57	0,60
SK_IMD2006_v5_IRD2006_A_v3	0,92	0,65	0,67
SK_IMD2006_v5_IRD2006_B_v3	0,91	0,59	0,62
SK_IMD2009_v1_IRD2009_A_v3	0,91	0,65	0,67
SK_IMD2009_v1_IRD2009_B_v3	0,91	0,59	0,62
SK_IMD2009_v3_IRD2009_A_v3	0,92	0,66	0,68
SK_IMD2009_v3_IRD2009_B_v3	0,92	0,61	0,63

Tabuľka 12 Vypočítané korelačné koeficienty pre vzťahy medzi IMD a IRD\_A/B na území Česka  
(Zdroj dát: Autorka)

CZ_IMDdatabáza_IRDdatabáza	Pearson's r	Kendall's tau	Spearman's rho
CZ_IMD2006_v1_IRD2006_A_v3	0,82	0,59	0,62
CZ_IMD2006_v1_IRD2006_B_v3	0,82	0,55	0,59
CZ_IMD2006_v2_IRD2006_A_v3	0,90	0,62	0,64
CZ_IMD2006_v2_IRD2006_B_v3	0,89	0,58	0,61
CZ_IMD2006_v3_IRD2006_A_v3	0,90	0,62	0,65
CZ_IMD2006_v3_IRD2006_B_v3	0,90	0,58	0,61
CZ_IMD2006_v5_IRD2006_A_v3	0,90	0,63	0,66
CZ_IMD2006_v5_IRD2006_B_v3	0,90	0,59	0,62
CZ_IMD2009_v1_IRD2009_A_v3	0,91	0,62	0,65
CZ_IMD2009_v1_IRD2009_B_v3	0,90	0,56	0,60
CZ_IMD2009_v3_IRD2009_A_v3	0,91	0,62	0,65
CZ_IMD2009_v3_IRD2009_B_v3	0,91	0,57	0,60



Graf 11 Hodnoty korelačných koeficientov pre vzťahy medzi referenčnými databázami IRD\_A/B a príslušnými validovanými databázami IMD pre roky 2006 a 2009 na Slovensku a v Česku  
(Zdroj dát: Autorka)

V Tabuľkách 11 a 12 sú vypočítané všetky hodnoty navrhovaných korelačných koeficientov pre obidve záujmové územia. Pri tejto metóde sme okrem verzií IRD\_A vypočítali hodnoty aj pre verzie IRD\_B, aby sme poukázali na skutočne vyššiu presnosť verzií IMD v prípade, že nepriepustné plochy a autochtónnym materiálom sú zaradené do skupiny priepustných povrchov. Hodnoty korelačných koeficientov verzie B sú vo všetkých prípadoch nižšie v porovnaní s príslušnými hodnotami korelačných koeficientov verzie A. Údaje z Tabuliek 11 a 12 sme pre lepšiu interpretáciu preniesli do Grafu 11.

Body každého korelačného koeficientu majú tendenciu narastať ku hodnote 1 v smere ku novším verziám IMD (zľava doprava), čo potvrdzuje zvyšovanie tematickej presnosti jednotlivých IMD od starších verzií k novším.

Podľa všetkých vypočítaných korelačných koeficientov sa javia všetky databázy IMD (2006 aj 2009) v porovnaní s IRD\_A presnejšie na Slovensku ako v Česku. Podľa IRD\_B je výnimkou databáza IMD2006\_v2 a IMD2006\_v3, kedy podľa Kendallovho a Spearmanovho korelačného koeficientu, hoci len v priemere len o dve stotiny, je tematická presnosť v Česku vyššia ako na Slovensku.

### **4.3 Hodnotenie tematickej presnosti výpočtom priemernej absolútnej chyby normalizovanej a jej štruktúry**

Doposiaľ sme pomocou vyššie uvedených štatistických metód skúmali vzniknuté chyby nadhodnotenia a podhodnotenia, ich výskyt vzhľadom na stupeň hustoty nepriepustnosti pixelov a porovnávali sme ich početnosť na celej škále nepriepustnosti od 0 do 100 %. Vďaka korelačnej analýze sme analyzovali postupný vývoj zvyšovania tematickej presnosti jednotlivých vrstiev. Na to, aby sme dokázali kvantifikovať aj absolútne chyby jednotlivých pixelov v IMD a IRD databázach, zvolili sme výpočet totálnej absolútnej chyby normalizovanej (TAEn). Okrem hodnotenia celkovej veľkosti TAEn je veľmi dôležité aj hodnotenie jej štruktúry. Vzniknuté chyby nadhodnotenia a podhodnotenia (Tabuľka 14 a Tabuľka 15) v pixeloch boli rozdelené na dva typy – menej závažné a závažnejšie, z pohľadu ich potenciálneho vplyvu na výsledok pri využívaní IMD ako pomocnej vrstvy pre dezagregáciu sociálnoekonomických premenných (Tabuľka 9). Pre tento účel je totiž potrebné poznať nie len celkovú chybu (Tabuľka 13), ale tiež to, ako často sa tieto dva typy zhody a štyri typy nezhody (resp. chýb) vyskytujú v IMD. Pre každý z týchto šiestich typov

(ne)zhody je preto potrebné zistiť frekvenciu jeho výskytu vzhľadom na celkovú veľkosť vzorky.

Tabuľka 13 Celková absolútna chyba normalizovaná (TAEn) pre referenčnú databázu IRD\_A  
(Zdroj dát: Autorka)

IMDdatabáza_IRDdatabáza	TAE/n - Podiel nesprávne klasifikovaných plôch z celkovej rozlohy záujmového územia (%)	Rozloha nesprávne klasifikovaných plôch v záujmovom území (km <sup>2</sup> )	TAEn = (TAE/n)/B (%)	B - Podiel nepriepustných plôch z celkovej rozlohy záujmového územia (%)	Rozloha nepriepustných plôch v záujmovom území (km <sup>2</sup> )
SK_IMD2006_v1_IRD2006_A	1,54	753,59	<b>65,95</b>	2,33	1142,71
SK_IMD2006_v2_IRD2006_A	1,47	720,71	<b>63,07</b>	2,33	1142,71
SK_IMD2006_v3_IRD2006_A	1,41	691,43	<b>60,51</b>	2,33	1142,71
SK_IMD2006_v5_IRD2006_A	1,06	517,57	<b>45,29</b>	2,33	1142,71
SK_IMD2009_v1_IRD2009_A	1,44	705,90	<b>59,89</b>	2,40	1178,70
SK_IMD2009_v3_IRD2009_A	1,07	522,92	<b>44,36</b>	2,40	1178,70
CR_IMD2006_v1_IRD2006_A	2,27	1791,24	<b>70,73</b>	3,21	2532,39
CR_IMD2006_v2_IRD2006_A	1,86	1469,43	<b>58,03</b>	3,21	2532,39
CR_IMD2006_v3_IRD2006_A	1,82	1434,73	<b>56,66</b>	3,21	2532,39
CR_IMD2006_v5_IRD2006_A	1,48	1170,45	<b>46,22</b>	3,21	2532,39
CR_IMD2009_v1_IRD2009_A	1,87	1476,49	<b>56,40</b>	3,32	2618,04
CR_IMD2009_v3_IRD2009_A	1,52	1200,18	<b>45,84</b>	3,32	2618,04

Rovnako dôležité je poznať tiež veľkosť príspevku každého zo štyroch typov nezahody k celkovej hodnote chyby TAEn, a teda do akej miery sa každý z nich podieľa na celkovej tematickej nepresnosti IMD (farebné stĺpce v Tabuľke 14 a v Tabuľke 15).

Tabuľka 14 Chyby nadhodnotenia  
(Zdroj dát: Autorka)

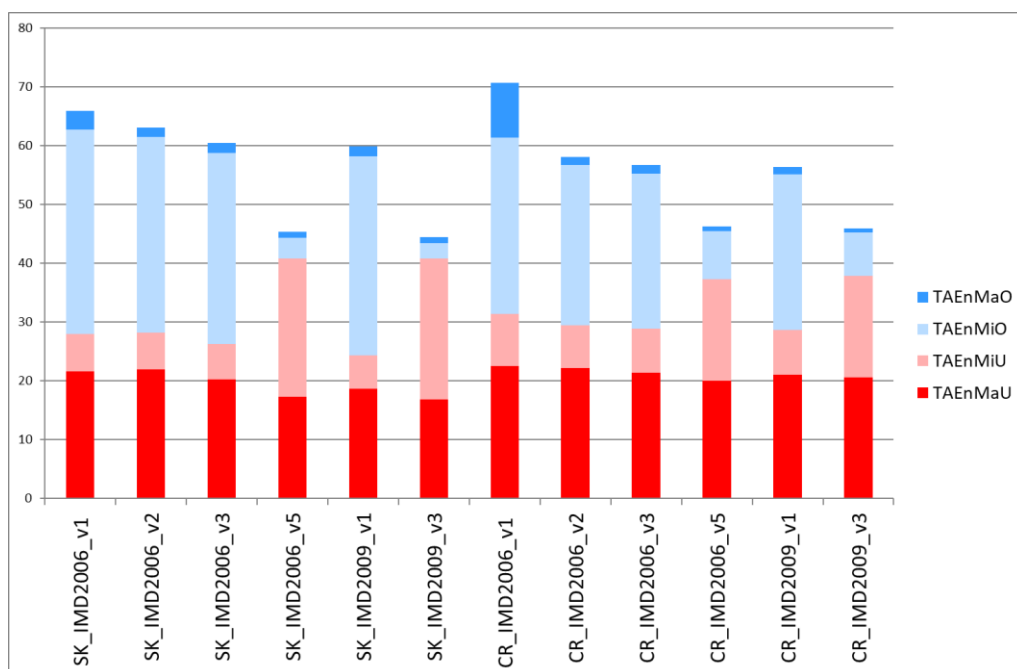
IMDdatabáza_IRDdatabáza	TAEno (%)	TAEo (km <sup>2</sup> )	TAEnMa0 (%)	TAEMa0 (km <sup>2</sup> )	TAEnMio (%)	TAE Mio (km <sup>2</sup> )
SK_IMD2006_v1_IRD2006_A	<b>38,02</b>	434,46	3,17	36,21	34,85	398,25
SK_IMD2006_v2_IRD2006_A	<b>34,92</b>	399,03	1,54	17,58	33,38	381,45
SK_IMD2006_v3_IRD2006_A	<b>34,25</b>	391,33	1,79	20,47	32,45	370,86
SK_IMD2006_v5_IRD2006_A	<b>4,52</b>	51,61	0,97	11,13	3,54	40,48
SK_IMD2009_v1_IRD2009_A	<b>35,55</b>	419,01	1,67	19,69	33,88	399,32
SK_IMD2009_v3_IRD2009_A	<b>3,61</b>	42,54	0,92	10,81	2,69	31,73
CR_IMD2006_v1_IRD2006_A	<b>39,35</b>	996,51	9,37	237,19	29,98	759,32
CR_IMD2006_v2_IRD2006_A	<b>28,57</b>	723,40	1,35	34,19	27,22	689,21
CR_IMD2006_v3_IRD2006_A	<b>27,86</b>	705,50	1,49	37,66	26,37	667,84
CR_IMD2006_v5_IRD2006_A	<b>8,95</b>	226,62	0,83	20,94	8,12	205,68
CR_IMD2009_v1_IRD2009_A	<b>27,75</b>	726,55	1,26	33,08	26,49	693,47
CR_IMD2009_v3_IRD2009_A	<b>8,07</b>	211,24	0,60	15,62	7,47	195,63

Tabuľka 15 Chyby podhodnotenia  
(Zdroj dát: Autorka)

IMDdatabáza_IRDdatabáza	<b>TAEnu (%)</b>	TAEu (km <sup>2</sup> )	<b>TAEnMaU (%)</b>	TAE <sub>MaU</sub> (km <sup>2</sup> )	TAEn <sub>MiU</sub> (%)	TAE <sub>MiU</sub> (km <sup>2</sup> )
SK_IMD2006_v1_IRD2006_A	<b>27,93</b>	319,13	21,59	246,72	6,34	72,40
SK_IMD2006_v2_IRD2006_A	<b>28,15</b>	321,68	21,85	249,64	6,30	72,03
SK_IMD2006_v3_IRD2006_A	<b>26,26</b>	300,10	20,19	230,76	6,07	69,34
SK_IMD2006_v5_IRD2006_A	<b>40,78</b>	465,96	17,19	196,44	23,59	269,53
SK_IMD2009_v1_IRD2009_A	<b>24,34</b>	286,89	18,64	219,75	5,70	67,13
SK_IMD2009_v3_IRD2009_A	<b>40,76</b>	480,38	16,74	197,27	24,02	283,11
CR_IMD2006_v1_IRD2006_A	<b>31,38</b>	794,73	22,43	567,91	8,96	226,82
CR_IMD2006_v2_IRD2006_A	<b>29,46</b>	746,03	22,17	561,53	7,29	184,51
CR_IMD2006_v3_IRD2006_A	<b>28,80</b>	729,23	21,32	540,00	7,47	189,24
CR_IMD2006_v5_IRD2006_A	<b>37,27</b>	943,83	19,94	504,90	17,33	438,93
CR_IMD2009_v1_IRD2009_A	<b>28,65</b>	749,94	20,99	549,62	7,65	200,32
CR_IMD2009_v3_IRD2009_A	<b>37,77</b>	988,94	20,54	988,94	17,24	451,23

Interpretujeme Tabuľku 13: Prvý číselný stĺpec vyjadruje podiel nesprávne klasifikovaných plôch z celkovej rozlohy záujmového prostredia. V tomto stĺpci sa nachádzajú pomerne nízke percentuálne hodnoty, čo by mohlo navádzať ku nesprávnej interpretácii chybovosti jednotlivých vrstiev IMD. Je potrebné si uvedomiť, že toto percento je vyjadrené z celkovej rozlohy jednotlivých štátov, teda aj z rozlohy priepustných aj nepriepustných povrchov. Celková rozloha Slovenska je 49 036 km<sup>2</sup> a Česka 78 866 km<sup>2</sup>. V ďalšom stĺpci je vyjadrená rozloha nesprávne klasifikovaných plôch v km<sup>2</sup>. Najväčšiu výpovednú hodnotu má tretí číselný stĺpec obsahujúci percentuálne údaje celkovej absolútnej chyby normalizovanej vzhľadom na celkovú rozlohu nepriepustných povrchov v danom záujmovom území. Tieto hodnoty presne vyjadrujú, aký podiel zastupujú chybné klasifikované pixely vzhľadom ku všetkým nepriepustným plochám v obidvoch krajinách. Posledné dva stĺpce vyjadrujú skutočnú rozlohu nepriepustných plôch a ich podiel na celkovej rozlohe záujmového územia.

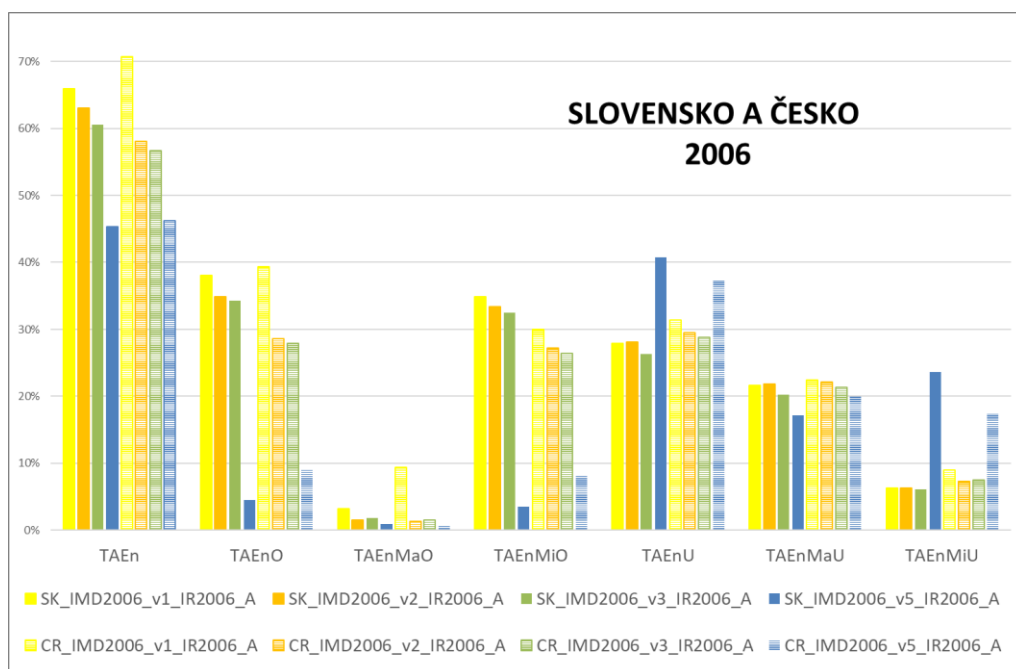
Z hodnôt TAEn vyplýva, že rozloha nesprávne klasifikovaných plôch v pomere k rozlohe nepriepustných plôch v príslušnom záujmovom území tvorí od 44,36 % (na Slovensku v najnovšej verzii IMD2009) až do 70,73 % (v Česku v najstaršej verzii IMD2006). V najstaršej verzii IMD2006 a v najnovších verziách IMD2006 a IMD2009 bola hodnota TAEn väčšia v Česku v porovnaní so Slovenskom, v ostatných verziách to bolo naopak.



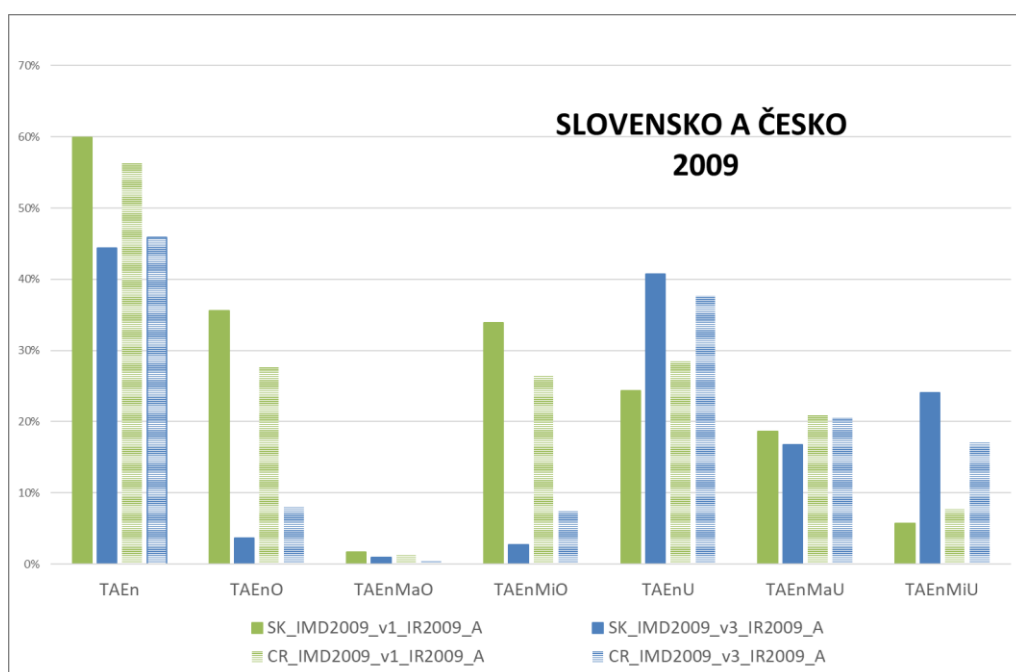
Graf 12 Celková absolútna chyba normalizovaná (TAEn) v percentách a jej štruktúra vypočítaná pre referenčnú databázu IRD\_A  
(Zdroj dát: Autorka)

Z Grafu 12 je zrejmé, že vypočítaná chyba TAEn klesá v oboch krajinách pre obidva skúmané roky od starších verzií po novšie, a teda že tematická presnosť jednotlivých verzií v čase postupne narastá. K najvýraznejšiemu zlepšeniu pritom došlo medzi dvoma najnovšími dostupnými verziami, t. j. medzi IMD2006\_v3 a IMD2006\_v5 pre vrstvu pre rok 2006 (zo 60,51 % na 45,29 % na Slovensku a z 56,66 % na 46,22 % v Česku) a medzi IMD2009\_v1 a IMD2009\_v3 pre vrstvu pre rok 2009 (z 59,89 % na 44,36 % na Slovensku a z 56,40 % na 45,84 % v Česku). Medzi uvedenými dvojicami verzii IMD však zároveň došlo k výraznej zmene štruktúry TAEn, v porovnaní so zmenami štruktúry pri predchádzajúcich vylepšeniach IMD. Hoci v najnovších verziách je podiel menej závažných chýb nadhodnotenia výrazne menší, podiel menej závažných chýb podhodnotenia je výrazne väčší ako v predchádzajúcich verziách.

Hodnoty všetkých chýb podhodnotenia a nadhodnotenia vypočítané pre obidve krajiny pre referenčné databázy IRD\_A sú prehľadne zobrazené v Grafoch 13 a 14.



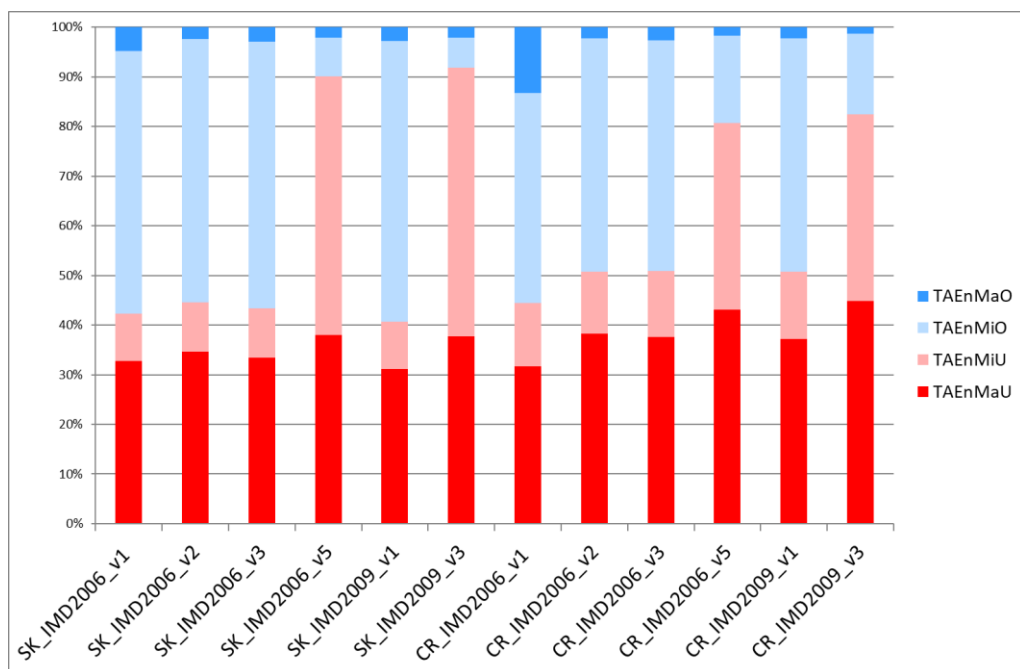
Graf 13 Porovnanie hodnôt všetkých druhov chýb vypočítané pre referenčnú databázu IRD\_A pre rok 2006 na území Slovenska a Česka (Zdroj dát: Autorka)



Graf 14 Porovnanie hodnôt všetkých druhov chýb vypočítané pre referenčnú databázu IRD\_A pre rok 2009 na území Slovenska a Česka (Zdroj dát: Autorka)

Pri niektorých aplikáciách, ako napr. pri dezagregácii sociálnoekonomických premenných, sú za najkritickejšie považované závažnejšie chyby ( $TAEn_{Ma} = TAEn_{MaO} + TAEn_{MaU}$ ): Zatiaľ čo v procese dezagregácie závažnejšie chyby nadhodnotenia ( $TAEn_{MaO}$ ) vytvárajú fiktívne dazymetrické zóny s relatívne veľkou predpokladanou hustotou zaľudnenia, naopak

závažnejšie chyby podhodnotenia (TAEn<sub>MaU</sub>) neodhaľujú skutočné dazymetrické zóny s relatívne veľkou predpokladanou hustotou zaľudnenia. Je pozitívne, že veľkosť týchto závažnejších chýb v smere od starších k novším verziám klesá, avšak kvôli pomalému tempu tohto poklesu (vzhľadom k rýchlejšiemu poklesu celkovej chyby TAEn) ich relatívny podiel na celkovej rozlohe nesprávne klasifikovaných plôch v čase narastá, napr. podiel TAEn<sub>MaU</sub> v najnovších verziách na takmer 38 % na Slovensku (Graf 15 – stĺpec SK\_IMD2006\_v5 a SK\_IMD2009\_v3) a 43 až 45 % v Česku (Graf 15 – stĺpec CZ\_IMD2006\_v5 a CZ\_IMD2009\_v3). Podiel všetkých druhov chýb na celkovej absolútnej chybe normalizovanej, vypočítaný pre referenčnú databázu IRD\_A prehľadne znázorňuje Graf 15.



Graf 15 Percentuálny podiel jednotlivých druhov chýb na celkovej absolútnej chybe normalizovanej vypočítaný pre referenčnú databázu IRD\_A  
(Zdroj dát: Autorka)



## 5 DISKUSIA

V prípade, že by aspoň niektoré z verzií poskytovaných vrstiev IMD2006 a IMD2009 mali podľa skúmaných štatistických metód zhodné výsledky s vytvorenými referenčnými databázami IRD2006 a IRD2009, zodpovedali by skutočnému stavu (z roku 2006, resp. z roku 2009). Rozdiely, ktoré sme zistili nastali buď preto, že skutočnosti nezodpovedali verzie pôvodných databáz, alebo boli chybné vytvorené a interpretované referenčné databázy.

### 5.1 Možné zdroje chýb v pôvodných databázach

Jednou z možných príčin, prečo pôvodné databázy skúmaných vrstiev nepriepustnosti nekorešpondujú s realitou je ich geometrická nepresnosť. Ako bolo uvedené v úvodných kapitolách tejto dizertačnej práce, program Copernicus prevádzkuje služby poskytujúce snímky a referenčné údaje, ku ktorým patrí aj rastrový produkt HR IMAGE súvisle pokrývajúci štáty Európy. Práve tieto satelitné snímky s vysokým rozlíšením sa stali základom pre tvorbu vrstiev Európskej vrstvy hustoty nepriepustnosti. Snímky boli počas procesu tvorby jednotlivých databáz prerozdeľované jednotlivým spoločnostiam, ktoré mali produkciu na starosti. Tie ich najprv pretransformovali do vlastných súradnicových systémov, a tak ich spracovali. V závere svoje hotové časti produktov opätovne transformovali späť do európskeho súradnicového systému, aby mohli byť pospájané do jedného výsledného produktu. Predpokladáme, že práve pri týchto procesoch transformácie sa mohli objaviť isté geometrické nepresnosti a posuny. Geometrickú nepresnosť majú však aj samotné vrstvy HR IMAGE, pri ktorých boli zistené posuny v smere rovnobežiek okolo 12 m a v smere poludníkov okolo 8 m priemernej kvadratickej chyby. Posuny sú však v priemere menšie ako 10 m, preto ak máme pixely so 100 m rozlíšením, tak tento zdroj chýb by nemal byť až tak závažný.

Výsledky štatistických metód poukázali na to, že v pixeloch s nižšími hodnotami hustoty nepriepustnosti sa zaokrúhľovalo smerom nadol a tým došlo ku zjavnému podhodnocovaniu a v pixeloch s vyššími hodnotami hustoty nepriepustnosti sa zaokrúhľovalo smerom nahor, čím sa prejavilo naopak výrazné nadhodnotenie. Výsledné produkty môžu mať následne sklon k binarite (vysoký výskyt nulových hodnôt a vysoký výskyt najmä nadhodnotených pixelov). Dôvodom je pravdepodobne samotná metodika produkcie. Metodologická

nepresnosť zahŕňa (Kopecký a Kahabka 2008) hybridnú automatickú klasifikáciu s výsledkom binárnej mapy zastavaných povrchov, následnú korekciu tejto mapy tak, aby bola dosiahnutá požadovaná 85 % tematická presnosť a odvádzanie hustoty nepriepustnosti na základe indexu NDVI popisujúceho výskyt vegetácie podľa spektrálnej odrazivosti.

Ďalšou príčinou nesúladu pôvodných databáz s realitou je už viackrát v tejto dizertačnej práci spomínaná nepresnosť pojmu nepriepustných povrchov, ktorý je súčasťou definície zastavaných plôch. Ak by sme za nepriepustné plochy (ktoré by IMD mala zachytávať) nepovažovali aj nepriepustné plochy s autochtónnym materiálom, teda „z nepriepustnené“ vďaka aktivite človeka, napr. poľné cesty (z tohto predpokladu vychádza referenčná databáza IRD\_A), potom by tematická presnosť IMD bola vyššia v porovnaní s alternatívou, kedy by sme nepriepustné plochy s autochtónnym materiálom za nepriepustné plochy (ktoré by IMD mala zachytávať) považovali (z tohto predpokladu vychádza referenčná databáza IRD\_B).

## **5.2 Možné zdroje chýb v referenčných databázach**

Tak ako v pôvodných databázach, chyby sa objavujú aj v referenčných databázach, kde najdôležitejšiu úlohu pri ich tvorbe zohráva samotný interpretátor. Nie vždy sa totiž dalo určiť jednoznačne, či povrch, ktorý sa nachádzal pod interpretovaným bodom, bol priepustný (hodnota 0) alebo nepriepustný (hodnota 1 a hodnota 2). V tom prípade zohrala značnú úlohu subjektivita.

Veľmi ťažko identifikovateľné boli hlavne veľmi malé objekty pri menšom rozlíšení. Pomerne často sa vyskytujúce boli pixely uložené na rozhraní priepustných a nepriepustných povrchov, napr. cesta a okolité trávnaté porasty. Niektoré body v týchto pixeloch padli presne na rozhranie týchto povrchov a bolo potrebné rozhodnúť o tom, aká hodnota im bude priradená. V neposlednom rade je potrebné spomenúť tieň budov, ktoré boli často tak tmavé, že výrazne sťažovali interpretáciu.

Je zrejmé, že v prípade viacerých interpretátorov by bol výklad odlišný. Práve preto bola už skôr interpretovaná vzorka z územia Česka pre referenčný rok 2009 (Pohlová 2013) nami reinterpretovaná, aby sa zachovala konzistentnosť interpretácie s rokom 2006 v Česku a s obidvoma rokmi na Slovensku. Riešením by mohlo byť tiež vypočítanie hodnôt

jednotlivých pixelov ako aritmetický priemer určených hodnôt, ktoré by im priradili viacerí zapojení interpretátori.

Portugalské NRC (Caetano et al., 2008) navrhlo viacstupňovú validáciu, ktorej podobný princíp bol využitý v tejto dizertačnej práci. Z pôvodných referenčných databáz IRD\_v1 vznikli postupným opätovným hodnotením a opravením hlavne pixelov so závažnými chybami, referenčné databázy IRD\_v2 až IRD\_v3 (s ktorými sme porovnávali tematickú presnosť vrstiev IMD).

Značné rozdiely medzi stavom priepustných a nepriepustných povrchov v referenčnej databáze a skutočným stavom povrchov v danom roku mohli byť spôsobené aj so zmenou krajinnej pokrývky, ktorá nastala v období medzi nasnímaním snímok HR IMAGE, z ktorých IMD vzniká a snímok, ktoré boli použité pri interpretácii a tvorbe referenčných dát. V prípade, že v tomto krátkom časovom úseku (napr. jeden až dva roky staršie snímky IRD pre referenčný rok 2006 ako snímky IMD) vznikla nová výstavba zachytená v IMD, ale nie v IRD, bolo zaznamenané nepravdivé nadhodnotenie pixelov. Ak bola výstavba zlikvidovaná v IMD, ale stále zachytená a interpretovaná v IRD, dochádzalo k nepravdivému podhodnoteniu pixelov, čo je však veľmi málo pravdepodobné. Ako riešenie sme využívali okrem oficiálnych dostupných snímok aj iné zdroje, ktoré ponúkali snímky z rovnakého časového obdobia a jednotlivé oblasti sme navzájom porovnávali.

## ZÁVER

Predložená dizertačná práca poukázala na potrebu validácie Európskej vrstvy hustoty nepriepustnosti (Imperviousness Density – IMD). Hoci očakávania výrobcov a používateľov boli od jej vzniku veľké, ukázalo sa, že tematická presnosť nie je až na takej úrovni, ako sa predpokladalo.

So zreteľom na celkovú presnosť, presnosť používateľa a presnosť producenta predložená práca priblížila niektoré teoretické koncepcie a výsledky doposiaľ vykonaných validácií na Slovensku aj v zahraničí, podľa ktorých IMD vrstva nespĺňa kritérium 85 % presnosti stanovenej v požiadavkách tendra na zhotovenie produktu. Maucha a Büttner (2008) konštatujú, „že iba skontrolovaním všetkých pixelov vrstvy a zistením, že presne 15 % z nich bolo nesprávne klasifikovaných by bolo možné preukázať, že tematická presnosť vrstvy je presne 85 %. Avšak skontrolovať každý pixel by bolo finančne a časovo nákladnejšie ako samotná produkcia hodnotenej vrstvy, preto je nevyhnutné použiť na účel validácie štatistické metódy výberu, na použitie ktorých je ale potrebné vytvorenie referenčnej databázy, ako nástroja na porovnávanie tematickej presnosti validovanej vrstvy“.

Vytvorenie vlastnej referenčnej databázy na území Slovenska a Česka je prvým krokom k naplneniu hlavného cieľa dizertačnej práce. Tento výsledný produkt by mal byť prínosnejší ako doposiaľ existujúce vytvorené referenčné databázy na porovnávanie tematickej presnosti vrstiev nepriepustnosti.

Druhým krokom potrebným na realizáciu validácií, ktoré boli vykonané v dizertačnej práci, bolo navrhnuť štatistické kvantitatívne metódy vhodné na čo najpresnejšiu komplexnú validáciu intenzitných máp, porovnať ich navzájom, prípadne uviesť výhody a nevýhody ich použitia.

V nadväznosti na tieto kroky k naplneniu hlavného cieľa dizertačnej práce (komplexná validácia Európskej vrstvy hustoty nepriepustnosti) boli ako súčasť ďalšieho výskumu uskutočnené tieto konkrétne kroky:

- čitateľovi sme priblížili ponuku a dostupnosť priestorových dát o krajinej pokrývke, ktoré poskytuje európsky program Copernicus,
- popísali sme Európsku vrstvu hustoty nepriepustnosti – jej vznik, priebeh vývoja a vylepšovania jednotlivých dostupných verzií

- pomocou štatistických metód sme kvantitatívne zhodnotili presnosť jednotlivých postupne vylepšovaných verzií IMD2006 a IMD2009, navzájom sme porovnali ich vylepšené verzie a dosiahnuté výsledky v záujmovom území Slovenska a Česka.

Na základe použitých metód hodnotenia sa zdá byť rozdiel v tematickej presnosti IMD v Česku a na Slovensku relatívne malý. V oboch krajinách v predposledných (IMD2006\_v3 a IMD2009\_v1) a starších verziách sa početnosť pixelov malých hodnôt nepriepustnosti zdá byť podhodnotená a naopak početnosť pixelov veľkých hodnôt nepriepustnosti nadhodnotená. Toto nadhodnotenie je následne v najnovších verziách viacmenej primerane (v Česku), resp. až príliš (na Slovensku) redukované. V prípade najnovších verzií IMD2006\_v5 a IMD2009\_v3 (berúc do úvahy už len referenčnú databázu IRD\_A) je tematická presnosť IMD na Slovensku väčšia ako v Česku tak podľa korelačných koeficientov, ako aj podľa veľkostí celkových absolútnych chýb normalizovaných, podľa veľkostí závažnejších chýb podhodnotenia a nadhodnotenia, a tiež podľa ich podielu na celkovej rozlohe nesprávne klasifikovaných plôch. V oboch krajinách tematická presnosť IMD postupne narastá od starších verzií k novším – od IMD2006\_v1 postupne až po IMD2006\_v5 a od IMD2009\_v1 po IMD2009\_v3 – podľa všetkých týchto vymenovaných metód. Presnosť najnovšej verzie IMD pre rok 2006 (IMD2006\_v5) sa zdá byť o málo menšia v porovnaní s presnosťou staršej dostupnej verzie IMD (IMD2009\_v1) pre rok 2009 len v Česku a len pri použití Pearsonovho korelačného koeficientu; podľa všetkých ostatných vyššie vymenovaných metód je tomu v oboch krajinách naopak.

V súvislosti so záverom, že tematická presnosť jednotlivých verzií v čase narastá, dokonca v súvislosti s publikovaním najnovších hodnotených verzií rýchlejšie ako pri zverejnení tých predchádzajúcich, je potrebné upozorniť na skutočnosť, ktorá tento na prvý pohľad jasný a pozitívny záver do istej miery narušuje. Je to zväčšenie rozdielu medzi najpresnejším dostupným odhadom rozlohy nepriepustných plôch (asi 1143 km<sup>2</sup> v roku 2006 a 1179 km<sup>2</sup> v roku 2009 na Slovensku a asi 2532 km<sup>2</sup> v roku 2006 a 2618 km<sup>2</sup> v roku 2009 v Česku) vychádzajúcim z priemeru vzorky príslušných referenčných databáz IRD\_A (2,33 % v roku 2006 a 2,40 % v roku 2009 na Slovensku a 3,21 % v roku 2006 a 3,32 % v roku 2009 v Česku) na jednej strane a odhadom rozlohy nepriepustných plôch z priemeru vzorky príslušných validovaných databáz IMD na strane druhej, práve v súvislosti so zverejnením najnovších hodnotených verzií: Napríklad podľa priemeru vzorky staršej verzie IMD2009\_v1 by mala byť rozloha nepriepustných plôch na Slovensku a v Česku spolu asi 3904 km<sup>2</sup>, podľa priemeru vzorky IMD2009\_v3 by to malo byť len asi 2578 km<sup>2</sup>, čo je

výrazne menej ako najpresnejší dostupný odhad podľa priemeru vzorky IRD2009\_A 3795 km<sup>2</sup> (príp. IRD2006\_B 4143 km<sup>2</sup>). Pre porovnanie Gallego et al. (2016, p. 6) uvádzajú tri rôzne odhady rozlohy nepriepustných plôch pre rok 2012 v týchto dvoch krajinách spolu: 3989 km<sup>2</sup>, 5410 km<sup>2</sup> a 5350 km<sup>2</sup>. V tomto smere sa teda zdá, že najnovšia verzia IMD je od reality vzdialenejšia ako tie predchádzajúce.

Záverom možno dodať, že tematická presnosť vrstvy nepriepustnosti IMD môže byť pre niektoré aplikácie v Česku a na Slovensku dostatočne, no pre iné nedostatočne uspokojivá. V jej najnovších verziách je celková absolútna chyba normalizovaná TAEn, t. j. pomer rozlohy nesprávne klasifikovaných plôch a rozlohy nepriepustných plôch v danej krajine, asi 44 až 46 %. Z toho asi dve pätiny alebo viac tvoria tzv. závažnejšie chyby (z pohľadu potenciálneho využitia IMD pre dezagregáciu sociálnoekonomických premenných), vyskytujúce sa tak v štvorcoch 100 × 100 m bez nepriepustných plôch, v ktorých IMD indikuje nenulovú nepriepustnosť (tvoriacich asi 1 až 2 % nesprávne klasifikovaných plôch), ako aj v štvorcoch 100 × 100 m s istým, zväčša malým nenulovým podielom nepriepustných plôch, v ktorých IMD indikuje nulovú nepriepustnosť (tvoriacich asi 38 až 45 % nesprávne klasifikovaných plôch).

Napriek tomu je možné vrstvu nepriepustnosti IMD považovať za jedinečný zdroj informácií, ktorý si nájde široké uplatnenie, avšak mal by byť používaný s rozvahou a obozrenosťou, s ohľadom na danú aplikáciu a požadovanú presnosť. Kvalita výstupov v minulosti realizovaných aplikácií využívajúcich IMD na Slovensku a v Česku môže byť spätne prehodnotená aj na základe poznania konkrétnej verzie IMD, ktorá do týchto aplikácií vstupovala, keďže jednotlivé verzie – ako sa ukázalo – sa v rôznych charakteristikách tematickej presnosti do značnej miery líšia.

## ZOZNAM POUŽITEJ LITERATÚRY

- AHRENS, B. et. al. 1998. *On the validation of the atmospheric model REMO with ISCCP data and precipitation measurements using simple statistics*. Meteorology and Atmospheric Physics. 68(3-4), 127-142.
- BAUER, M. E. et. al. 1978. *Area estimation of crops by digital analysis of Landsat data*. Photogrammetric Engineering and Remote Sensing, 44: 1033-1043.
- BISHOP, G. D. et. al. 1998. *A comparison of mapped estimates of long-term runoff in the northeast United States*. Journal of Hydrology, 206(3-4), s. 176-190.
- BÜTTNER G. 2012. *Guidelines for verification and enhancement of high resolution layers produced under GMES initial operations (GIO) Land monitoring 2011 – 2013*. EEA, 47 s.
- BÜTTNER G., KOSZTRA B., MAUCHA G. AND PATAKI R. 2012. *Implementation and achievements of CLC2006*. ETC-LUSI, EEA, 65 s.
- CAETANO, M. et al. 2008. *Accuracy Assessment of the High Resolution Built up Map for Continental Portugal*. Portuguese Geographic Institute, Lisbon, 26 s.
- CARFAGNA E., GALLEGO F. J. 2005. The use of remote sensing in agricultural statistics. In: *International Statistical Review*. 73, 3, s. 389-404.
- CAPES, R., PASSERA, E. 2022. Product Description and Format Specification. *End-to-end implementation and operation of the European Ground Motion Service (EEA/DIS/R0/20/011)* [online]. Doc. Version 1.0, Contract ID: SPECIFIC CONTRACT No 3436/R0-COPERNICUS/EEA.58362, s. 30. Dostupné na: <https://land.copernicus.eu/user-corner/technical-library/egms-product-description-document>
- CONGALTON, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, Volume 37, Issue 1, Jul 1991, s. 35-46. ISSN 0034-4257.
- CONGALTON, R.G. AND GREEN, K. 1999. *Assessing the Accuracy of Remotely Sensed Data Principles and Practices*. Lewis Publishers, Boca Raton.
- CZAPLEWSKI R. L. 1992. Misclassification bias in areal estimates. *Photogrammetric Engineering and Remote Sensing*. 58 (2): s. 189-192.
- ČSÚ. 2022. *Obyvatelstvo* [online]. Praha: Český statistický úřad, 2022 [cit. 2022-03-15]. Dostupné na: [https://www.czso.cz/csu/czso/obyvatelstvo\\_lide](https://www.czso.cz/csu/czso/obyvatelstvo_lide)

- ČÚZK. 2015. *Prohlížeč sloužba WMS – Ortofoto* [online]. Praha: Copyright ©2010 ČÚZK, 2015 [cit. 2018-07-15]. Dostupné na: [https://geoportal.cuzk.cz/\(S\(ai bp1i5dithzbq3vt3wk5d0j\)\)/Default.aspx?menu=3121&mode=TextMeta&side=wms.verejne&metadataID=CZ-CUZK-WMS-ORTOFOTO-P&metadataXSL=metadata.sluzba](https://geoportal.cuzk.cz/(S(ai bp1i5dithzbq3vt3wk5d0j))/Default.aspx?menu=3121&mode=TextMeta&side=wms.verejne&metadataID=CZ-CUZK-WMS-ORTOFOTO-P&metadataXSL=metadata.sluzba)
- DUVEILLER, G., FASBENDER, D., MERONI, M. 2016. Revisiting the concept of a symmetric index of agreement for continuous datasets. *Scientific Reports*, 6 doi:10.1038/srep19401.
- EEA. 2006. *GMES Fast Track Service Precursor on Land Monitoring. EEA Proposal 2006 – 2008 – Terms of Reference*. Copenhagen: 2006, 10 s.
- EEA. 2007. Tender Specifications – GMES Fast track service on land monitoring. *High resolution core land cover data for built-up areas, including degree of soil sealing, 2006 (EEA/IDS/07/001)* [online]. Copenhagen: EEA, 2007. 16 s. [cit. 2021-09-22]. Dostupné na: [https://www.eea.europa.eu/about-us/tenders/EEAIDS07001/tender\\_specifications.pdf](https://www.eea.europa.eu/about-us/tenders/EEAIDS07001/tender_specifications.pdf)
- EEA. 2009. Tender Specification – GMES Fast Track Service Precursor on land monitoring. *High resolution core land cover data for built-up areas, including degree of soil sealing, 2006 (EEA/SES/09/003)* [online]. Copenhagen: EEA, 2009 [cit. 2022-02-04]. Dostupné na: <https://www.eea.europa.eu/about-us/tenders/eea-ses-09-003/tender-specifications-eea-ses-09.003/view>
- EEA. 2010. *The European environment – State and Outlook 2010: LAND USE* [online]. Copenhagen: EEA, 2010. 52 s. [cit. 2021-10-15]. Dostupné na: [file:///C:/Users/Asus-SK/Downloads/Land\\_use.pdf](file:///C:/Users/Asus-SK/Downloads/Land_use.pdf)
- EEA. 2015. *European Soil Sealing V2* [online]. Copenhagen: EEA, 2015 [cit. 2017-06-17]. Dostupné na: <http://www.eea.europa.eu/data-and-maps/explore-interactive-maps/european-soil-sealing-v2>
- EK. 2015. *Copernicus Europe's eyes on Earth* [online]. Brussels: 2015 [cit. 2021-11-20]. ISBN 978-92-79-45666-4. Dostupné na: [https://www.videsinstituts.lv/assets/upload/userfiles/files/Copernicus\\_brochure.pdf](https://www.videsinstituts.lv/assets/upload/userfiles/files/Copernicus_brochure.pdf)
- EUROSENSE a GEODIS SLOVAKIA. 2005 – 2007. *Digitálne ortofotomapy dostupné na Google Earth*. EUROSENSE s.r.o. a GEODIS SLOVAKIA s.r.o.
- EUROSENSE a GEODIS SLOVAKIA. 2008 – 2010. *Digitálne ortofotomapy v rozlíšení 50 cm v plošnom rozsahu územia Slovenskej Republiky*. EUROSENSE s.r.o. a GEODIS SLOVAKIA s.r.o.



- FAO. 2005. Land Cover Classification System. *Classification concepts and User Manual* [online]. Rome: Food and Agriculture Organization of United Nations, 2015 [cit. 2022-01-20]. ISBN 92-5-105327-8. Dostupné na:  
<https://www.fao.org/3/y7220e/y7220e00.htm>
- FERRETTI, A., PASSERA, E., CAPES, R. 2021. Algorithm Theoretical Basis Document. *End-to-end implementation and operation of the European Ground Motion Service (EGMS)* [online]. Doc. Version 1.0, Contract ID: SPECIFIC CONTRACT No 3436/R0-COPERNICUS/EEA.58362, s. 81. Dostupné na:  
<https://land.copernicus.eu/user-corner/technical-library/egms-algorithm-theoretical-basis-document>
- FOODY, G. M. 2002. Status of land cover classification accuracy assessment. In: *Remote Sensing of Environment*, Vol. 80, Is. 1, April 2002, s. 185-201. ISSN: 0034-4257.
- GALLEGO J., SANNIER CH., PENNEC A., DUFOURMONT H. 2016. *Validation of Copernicus Land Monitoring Services and Area Estimation* [online]. Rome : ICASVII, 2016 [cit. 2018-05-16]. Dostupné na: <<https://land.copernicus.eu/user-corner/technical-library/validation-of-copernicus-land-monitoring-services-and-area-estimation/view>>
- GALLEGO, F. J., DELINCÉ J. 2010. The European Land Use and Cover Area-frame statistical Survey (LUCAS). In: *Agricultural Survey Methods*, ed: R. Benedetti, M. Bee, G. Espa, F. Piersimoni, Ch. 10. pp. -168, John Wiley & sons.
- GANGKOFNER, U., WEICHSELBAUM, J., KUNTZ, S., BRODSKY, L., LARSSON, K., DE PASQUALE, V. 2010. Update of the European High-resolution Layer of Built up Areas and Soil Sealing 2006 with Image 2009 Data. *30th EARSeL Symposium 2010: Remote Sensing for Science, Education and Natural and Cultural Heritage* [online]. Paris: s. 185-192 [cit. 2021-10-04]. Dostupné na:  
[https://earsel.org/symposia/2010-symposium-Paris/Proceedings/EARSeL-Symposium-2010\\_5-01.pdf](https://earsel.org/symposia/2010-symposium-Paris/Proceedings/EARSeL-Symposium-2010_5-01.pdf)
- GAREN, D. C., MARKS, D. 2005. Spatially distributed energy balance snowmelt modelling in a mountainous river basin: estimation of meteorological inputs and verification of model results. In: *Journal of Hydrology*, 315(1-4), s. 126-153.
- HAGEN, A. 2003. Fuzzy set approach to assessing similarity of categorical maps. In: *International Journal of Geographical Information Science*, 17(3), s. 235-249.

- HAGEN-ZANKER, A., ENGELEN, G., HURKENS, J., VANHOUT, R., ULJEE, I. 2006. *Map Comparison Kit 3: User Manual*. Maastricht: Research Institute for Knowledge Systems.
- HAGEN-ZANKER, A., STRAATMAN, B., ULJEE, I. 2005. Further developments of a fuzzy set map comparison approach. In: *International Journal of Geographical Information Science*, 19(7), s. 769-785.
- HANSEN, M. C., EGOROV, A., POTAPOV, P. V., STEHMAN, S. V., TYUKAVINA, A., TURUBANOVA, S. A., KOMMAREDDY, A. 2014. Monitoring conterminous United States (CONUS) land cover change with web-enabled landsat data (WELD). *Remote sensing of Environment*, 140, s. 466-484.
- HAZEWINKEL, M., 2001. *Teória chýb*. In: Encyklopédia matematiky, Springer Science + Business Media BV / Kluwer Academic Publishers. ISBN 978-1-55608-010-4.
- HENDL, J. 2015. *Přehled statistických metod: analýza a metaanalýza dat*. Páté, rozšířené vydání. Praha: Portál, 2015. ISBN 978-80-262-0981-2.
- HURBÁNEK, P. a kol. 2010. *Accuracy of Built-up Area Mapping in Europe at Varying Scales and Thresholds – Poster* [online]. Joint NASA LCLUC Science Team Meeting and GOFC-GOLD/NERIN, NEESPI Workshop, Monitoring land cover a land use in boreal and temperate Europe. Tartu, Estonia: Atlantis Conference Center, 25.28.August 2010 [cit. 2020-06-11]. Dostupné na: [https://lcluc.umd.edu/sites/default/files/lcluc\\_documents/pazur\\_lcluc\\_8-2010\\_poster\\_0.pdf](https://lcluc.umd.edu/sites/default/files/lcluc_documents/pazur_lcluc_8-2010_poster_0.pdf)
- HURBÁNEK, P., ROSINA, K. 2013. Metódy pre hodnotenie tematickej presnosti Európskej mapy umelých nepriepustných povrchov. In: *Zborník abstraktov – Rekurzívny sprievodca*. Ružomberok: VERBUM – vydavateľstvo KU, 2013. ISBN 978-80-561-0048-6, s. 49 – 52.
- HOUSTON A.G., HALL F.G. 1984. Use of satellite data in agricultural surveys. *Communications in Statistics Theory and Methods*, 23, pp. 2857-2880.
- CHABAEVA, A., HURD, J., & CIVCO, D. 2007. Quantitative assessment of the accuracy of spatial estimation of impervious cover. In: *ASPRS Annual Conference Proceedigns*, 2007. Tampa, Florida.
- JI, M., & JENSEN, J. R. 1999. Effectiveness of subpixel analysis in detecting and quantifying urban imperviousness from Landsat Thematic Mapper imagery. *Geocarto International*, 14(4), s. 33-41.
- JI, L., GALLO, K. 2006. An agreement coefficient for image comparison. *Photogrammetric Engineering and Remote Sensing*, 72(7), 823-833. Retrieved from [www.scopus.com](http://www.scopus.com).

- KEMPER, T., MELCHIORRI, M. and EHRLICH, D. 2021. Global Human Settlement Layer. In: *Publications Office of the European Union*, Luxembourg, 2021, ISBN 978-92-76-42061-3, doi:10.2760/456883, JRC126219.
- KOLÁŘ, J. 2003. *Geografické informační systémy*. Praha: ČVUT, 2003. 161 s. ISBN 8001026876.
- KOPECKÝ, M., KAHABKA, H. 2008. Delivery Report Slovakia. *EEA FTSP Sealing Delivery Report European Mosaic*, Issue 1.0., InfoTerra GmbH, Potsdam, Germany, 13 s.
- LEFEBVRE, A.; BEAUGENDRE, N.; PENNEC, A.; SANNIER, C.; CORPETTI T., 2013 Using data fusion to update built-up areas of the 2012 European High-Resolution Layer Imperviousness. In *Proceedings of the 33rd EARSeL Symposium Conference*, Matera, Italy, 3–6 June 2013. s. 321-328.
- LAJČÁKOVÁ, A. 2013. *Validácia európskej mapy stupňa nepriepustnosti z roku 2006 v lokalitách Podolíneč a Havranec*: bakalárska práca. Ružomberok: KU, 2013. 84 s.
- LAJČÁKOVÁ, A., HURBÁNEK P., ROSINA K. 2016. Validácia európskej mapy nepriepustnosti v Česku použitím náhodnej vzorky. In: *Geografické myšlení jako aktuální společná výzva*. České Budejovice: Jihočeská univerzita, 2016. ISBN 978-80-7394-619-7, p. 195-203.
- LIU, J., CHEN, J. M., CIHLAR, J., PARK, W. M. 1997. A process-based boreal ecosystem productivity simulator using remote sensing inputs. *Remote Sensing of Environment*, 62(2), s. 158-175.
- LEGATES, D. R., MCCABE JR., G. J. 1999. Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1), 233-241. doi:10.1029/1998WR900018.
- MACHKOVÁ, N., NOVÁČEK, J. 2009. CORINE – mapovanie Slovenska. In: *Enviromagazín*, č. 3/2009, s. 12-13.
- MACHKOVÁ, N., NOVÁČEK, J., TOBIK, J. 2008. *Verification of high resolution soil sealing layer Qualitative assessment Prepared by Slovak Republic*. Slovenská agentúra životného prostredia, Banská Bystrica, 26 s.
- MARKECHOVÁ, D., STEHLÍKOVÁ, B., TIRPÁKOVÁ, A. 2011. *Štatistické metódy a ich aplikácie*. Nitra: UKF, 2011. ISBN 978-80-809-4807-8.
- MAUCHA, G., BÜTTNER, G. 2008. *Recommendations for Quantitative assessment of Highresolution soil sealing layer*. Copenhagen: EEA and European Topic Centre for Land Use and Spatial Information, 2008. 31 s.

- MAUCHA, G. – BÜTTNER, G. – KOSZTRA, B. 2010. *European Validation of GMES FTS Soil Sealing Enhancement Data (final draft)*. Budapest: FOMI, 2010. 35 p.
- MAUCHA, G., PETRIK, O. 2008. *Accuracy assesment of the high resolution built up map for Hungary*. Institute of Geodesy, Cartography and Remote Sensing – FÖMI, Budapest, Hungary. 11 s.
- MAUCHA, G. 2010. Validating EUROLAND products. In: *GEOLAND Forum 6*, March 2010. Toulouse: France. 47 s.
- MÉNARD, A., MARCEAU, D. J. 2006. Simulating the impact of forest management scenarios in an agricultural landscape of southern Quebec, Canada, using a geographic cellular automata. *Landscape and Urban Planning*. In: *Press*, Corrected Proof.
- MERONI, M., ATZBERGER, C., VANCUTSEM, C., GOBRON, N., BARET, F., LACAZE, R., LEO, O. 2013. Evaluation of agreement between space remote sensing SPOT-VEGETATION fAPAR time series. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4), 1951-1962. doi:10.1109/TGRS.2012.2212447.
- MÜLLER, R. et. al. 2007. Automatic Production of European Orthoimage Coverage Within The GMES Land Fast Track Service Using Spot 4/5 and Irs-P6 Liss III. Data. In: *ISPRS Conference Proceedings. Volume XXXVI*. Germany: ISPRS Workshop Hannover, 2007. 6 s.
- MÜLLER, R. et. al. 2009. *GMES Fast track land service 2006-2008: IMAGE2006, European coverage, methodology and results*. Wessling: German Aerospace Centre (DLR), 2009.
- OLOFSSON, P., FOODY, G. M., HEROLD, M., STEHMAN, S. V., WOODCOCK, C. E., WULDER, M. A. 2014. Good practices for estimating area and assessing accuracy of land change. In: *Remote Sensing of Environment*, 148, s. 42 – 57.
- POHLOVÁ, R. 2013. *Hodnocení přesnosti databáze zastavěných území v Česku: bakalářská práce*. Praha: Univerzita Karlova: 2013. 52 s.
- PONTIUS, J., ROBERT GILMORE, HUFFAKER, D., DENMAN, K. 2004. Useful techniques of validation for spatially explicit land-change models. In: *Ecological Modelling*, 179(4), s. 445-461.
- PRASAD, A. M., IVERSON, L. R., LIAW, A. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. In: *Ecosystems*, 9(2), s. 181-199.

- ROSINA, K. 2010. *Európska mapa zastavaných plôch a jej validácia na území Slovenska*: diplomová práca. Bratislava: UK, 2010. 60 s.
- SANNIER CH. , GALLEGO J., DAHMER J., SMITH G., DUFOURMONT H., PENNEC A. 2016. Validation of Copernicus High Resolution Layer on Imperviousness degree for 2006, 2009 and 2012. In: *Proceedings of Spatial Accuracy*, 2016, s. 60-66.
- SELKOWITZ, D. J., & STEHMAN, S. V. 2011. Thematic accuracy of the National Land Cover Database (NLCD) 2001 land cover for Alaska. *Remote Sensing of Environment*, June 2011, 115(6), 1401–1407.
- SCHMITT, U. 2020. GMES Initial Operations / Copernicus Land monitoring services – Validation of products. *CLMS UA BBH Validation Report SC04 (EEA/MDI/14/010)* [online]. Copenhagen: EEA, 2020. 49 s. [cit. 2022-02-17]. Dostupné na: <https://land.copernicus.eu/user-corner/technical-library/building-heights-2012-validation-report>
- SILVÁN-CÁRDENAS J. L., WANG L. 2008. Sub-pixel confusion-uncertainty matrix for assessing soft classifications, *Remote Sensing of Environment*, 112, s. 1081-1095.
- SMITH, G. 2020. GMES Initial Operations / Copernicus Land monitoring services – Validation of products. HRL Imperviousness degree 2018 Validation report (EEA/MDI/14/010) [online]. Copenhagen: EEA, 2020. 69 s. [cit. 2022-01-11]. Dostupné na: [https://land.copernicus.eu/user-corner/technical-library/clms\\_hrl\\_imd\\_validation\\_report\\_sc04\\_1\\_3.pdf](https://land.copernicus.eu/user-corner/technical-library/clms_hrl_imd_validation_report_sc04_1_3.pdf)
- STEHMAN, S. V. 1999. Basic probability sampling designs for thematic map accuracy assessment. *International Journal of Remote Sensing*, 20 (12), s. 2423-2441.
- STEHMAN, S. V., CZAPLEWSKI, R. L. 1998. Design and Analysis for Thematic Map Accuracy Assessment - an application of satellite imagery. In: *Remote Sensing of Environment*, Volume 64, Number 3, June 1998 , s. 331-344
- STEHMAN, S. V. 2009. Sampling designs for accuracy assessment of land cover. In: *International Journal of Remote Sensing*, 30(20), s. 5243-5272.
- STRASSER, U., & MAUSER, W. 2001. Modelling the spatial and temporal variations of the water balance for the Weser catchment 1965-1994. *Journal of Hydrology*, 254(1-4), s. 199-214.
- ŠÚSR. 2022. *Počet obyvateľov SR* [online]. Bratislava: Štatistický úrad Slovenskej republiky, 2022. [cit. 2022-03-05]. Dostupné na: <https://www.scitanie.sk/>
- TAZBERIK, A. 1994. *Program CORINE*. Banská Bystrica: Slovenská agentúra životného prostredia, 1997.

- VISSER, H., DE NIJS, T. 2006. The Map Comparison Kit. In: *Environmental Modelling & Software*, 21(3), s. 346-358.
- VISSER, H., HAGEN, A., NIJS (DE), T., KLEIN GOLDEWIJK, C. G. M., BORSBOOM - VAN BEURDEN, J. A. M., NIET (DE), R. 2004. *The Map Comparison Kit: Method, software and applications* (report 550002005). Bilthoven: RIVM.
- VISCARRA ROSSEL, R. A., WALTER, C. 2004. Rapid, quantitative and spatial field measurements of soil pH using an Ion Sensitive Field Effect Transistor. In: *Geoderma*, 119(1-2), s. 9-20.
- WEALANDS, S. R., GRAYSON, R. B., & WALKER, J. P. 2005. Quantitative comparison of spatial fields for hydrological model assessment - some promising approaches. In: *Advances in Water Resources*, 28(1), s.15-32.
- WICKHAM, J. D., STEHMAN, S. V., GASS, L., DEWITZ, J., FRY, J. A., & WADE, T. G. 2013. Accuracy assessment of NLCD 2006 land cover and impervious surface. In: *Remote Sensing of Environment*, 130, s. 294-304.
- WILLMOTT, C. J. 1981. On the validation of models. In: *Physical Geography*, 2(2), s. 184 - 194. doi:10.1080/02723646.1981.10642213.
- XIAN, G., HOMER, C., DEWITZ, J., FRY, J., HOSSAIN, N., & WICKHAM, J. (2011). Change of impervious surface area between 2001 and 2006 in the conterminous United States. In: *Photogrammetric Engineering and Remote Sensing*, 77(8), s. 758 - 762.
- ZVÁRA, K. 2003. *Biostatistika*. Praha: Karolinum, 2003. 213 s. ISBN 8024607395.