

Univerzita Karlova
Přírodovědecká fakulta
Katedra aplikované geoinformatiky a kartografie

Studijní program: Geografie
Studijní obor: Geoinformatika, kartografie a dálkový průzkum Země



Bc. Markéta Růžičková

**POKROČILÉ METODY KLASIFIKACE DRUŽICOVÝCH DAT
PRO ÚČELY LULUCF**

**ADVANCED SATELLITE DATA CLASSIFICATION TECHNIQUES
FOR LULUCF APPLICATIONS**

Diplomová práce

Vedoucí práce: doc. RNDr. Přemysl Štych, Ph. D.

Praha 2025

Zadání diplomové práce

Jméno a příjmení: Markéta Růžičková

Studijní program: Geoinformatika, kartografie a dálkový průzkum Země

Název: Pokročilé metody klasifikace družicových dat pro účely LULUCF

Zásady pro vypracování

Téma DP

Diplomová práce se bude zaměřena na sektor LULUCF, zahrnující využití krajiny, změny využití krajiny a lesnictví. Jde o jeden ze sektorů lidské činnosti, za který jsou v rámci mezinárodních dohod každoročně reportovány emise či propady skleníkových plynů.

Přístupy jednotlivých států k LULUCF reportingu se od sebe více či méně odlišují. Tato práce se bude věnovat LULUCF reportingu zejména v podmínkách Česka. Teoretická část práce bude spočívat v důkladném popisu jednotlivých tříd používaných při vykazování emisí skleníkových plynů ze sektoru LULUCF. V praktické části bude zkoumána možnost klasifikace těchto tříd pomocí moderních metod založených na dálkovém průzkumu Země, s využitím klasifikátorů založených na strojovém učení.

Cíle DP

- rešerše reportingu LULUCF dat v Česku
- tvorba referenčních dat – třídy LULUCF v českých k. ú., provedená na základě terénního výzkumu
- tvorba klasifikačního modelu pro účely LULUCF za využití klasifikátorů založených na strojovém učení
- představení vybraných existujících land cover/land use datasetů

Výzkumné otázky

Jak jsou vymezeny třídy sektoru LULUCF v Česku?

Jak vytvořit trénovací data pro automatickou metodu klasifikace tříd LULUCF pomocí strojového učení? Jak tuto metodu nastavit, aby klasifikace dosahovala co nejlepších výsledků?

Jaké jsou přínosy a limity využití dat DPZ pro účely LULUCF reportingu?

Základní popis použitých dat a metod s uvedením výzkumného úkolu v oblasti geoinformatiky

Datové zdroje: Corine Land Cover 2018, LPIS, katastr nemovitostí ČR, ortofoto a ZM10 ČÚZK, CLC+, ESA World Cover, Dynamic World, ESRI Land Cover

SW: ArcGIS Pro, QGIS, aplikace QField (v tabletu), platforma Google Earth Engine

Předpokládané výstupy práce

- literární rešerše věnující se LULUCF reportingu v Česku
- trénovací a validační dataset pro klasifikaci krajiny do tříd LULUCF vzešlý z terénního mapování
- klasifikace dat DPZ do tříd LULUCF s využitím klasifikátorů založených na strojovém učení
- srovnání již existujících land cover/land use datasetů

Seznam odborné literatury:

Blujdea, V., Abad Vinas, R., Federici, S., Grassi, G. (2016): The EU greenhouse gas inventory for LULUCF sector: I. Overview and comparative analysis of methods used by EU member states. Carbon management 6 (5); str. 247-259.

Svoboda, J.; Štych, P.; Laštovička, J.; Paluba, D.; Kobliuk, N. Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia. Remote Sens. 2022, 14, 1189.

Nařízení Evropského parlamentu a Rady (EU) 2018/841 ze dne 30. května 2018 o zahrnutí emisí skleníkových plynů a jejich pohlcování v důsledku využívání půdy, změn ve využívání půdy a lesnictví do rámce politiky v oblasti klimatu a energetiky do roku 2030 a o změně nařízení (EU) č. 525/2013 a rozhodnutí č. 529/2013/EU.

ČHMÚ (2023): National greenhouse gas inventory report of the Czech Republic. Submission under UNFCCC.

Vedoucí diplomové práce: doc. RNDr. Přemysl Štych, Ph. D.

Konzultant diplomové práce: Mgr. Jan Svoboda

Datum zadání diplomové práce: 18. 5. 2024

.....

Vedoucí diplomové práce

doc. RNDr. Přemysl Štych, Ph.D.

.....
Garant studijního programu

V Praze dne 18. 5. 2024

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a uvedla jsem všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla použita k získání jiného nebo stejného akademického titulu.

V Praze dne 22. 4. 2025

Bc. Markéta Růžičková

Poděkování

Ráda bych na tomto místě poděkovala svému školiteli doc. RNDr. Přemyslu Štychovi, Ph. D., za jeho odborný dohled a vstřícné jednání. Velké poděkování patří mému konzultantovi Mgr. Janu Svobodovi za mnoho cenných rad i podnětů, a nadšení pro věc, které mi v průběhu práce předal. Děkuji i svému příteli za všestrannou podporu, kterou mi poskytoval během celého studia.

Abstrakt

Předkládaná diplomová práce se zabývá možnostmi využití dat dálkového průzkumu Země pro účely reportingu v sektoru LULUCF (Land Use, Land Use Change and Forestry) v Česku. V současnosti jsou data pro LULUCF reporting čerpána z katastru nemovitostí. Cílem práce bylo vytvořit klasifikační model umožňující určovat LULUCF třídy na základě snímků Sentinel-2. Tvorba klasifikačního modelu probíhala za využití vlastního trénovacího datasetu vzešlého z podrobného terénního mapování. V prostředí Google Earth Engine byly testovány algoritmy Random Forest (RF) a Gradient Tree Boosting (GTB). Hodnocen byl vliv vstupních pásem Sentinel-2, způsobu tvorby kompozitu, čistoty dat pro trénování modelu a nastavení parametrů obou klasifikátorů. Metodou RF se podařilo dosáhnout celkové přesnosti 87,4 %, metodou GTB pak 86,7 %. Nejúspěšnější model RF byl následně aplikován na celé území Česka. Práce kriticky hodnotí potenciál i limity využití dat DPZ pro LULUCF reporting, zejména ve vztahu k rozdílům mezi land cover a land use. Součástí práce bylo také představení a vzájemné porovnání vybraných volně dostupných globálních land cover/land use datasetů (CLC+, ESA World Cover, Dynamic World a ESRI Land Cover).

Klíčová slova: LULUCF, GEE, Random Forest, Gradient Tree Boosting, Sentinel-2

Abstract

Presented thesis explores the potential of remote sensing data for reporting in the LULUCF sector (Land Use, Land Use Change and Forestry) in Czechia. Currently, LULUCF data are derived from the national cadastral records. The aim of this thesis was to develop a classification model capable of determining LULUCF categories based on Sentinel-2 imagery. The model was trained using a custom dataset obtained through detailed field mapping. The processing was carried out on the Google Earth Engine platform, using two classification algorithms: Random Forest (RF) and Gradient Tree Boosting (GTB). The evaluation focused on the influence of spectral bands from Sentinel-2, the method of composite creation, the quality of training data, and algorithm parameter settings. The RF algorithm achieved an overall accuracy of 87.4 %, while the GTB algorithm reached 86.7 %. The best-performing RF model was subsequently applied to the entire territory of Czechia. The thesis provides a critical assessment of the potential and limitations of using remote sensing data for LULUCF reporting, particularly in relation to the differences between land cover and land use concepts. It also includes an overview and mutual comparison of selected freely available global land cover/land use datasets (CLC+, ESA World Cover, Dynamic World, ESRI Land Cover).

Keywords: LULUCF, GEE, Random Forest, Gradient Tree Boosting, Sentinel-2

Obsah

1 Úvod	9
2 Úvod do problematiky a literární rešerše	11
2.1 Legislativní rámec.....	11
2.2 Sektor LULUCF	12
2.3 LULUCF reporting v podmínkách Česka	14
2.4 Další metody klasifikace krajinných tříd (podobných) LULUCF	19
2.5 Existující zdroje land cover/land use dat.....	21
3 Řešená katastrální území.....	26
4 Data a metody.....	31
4.1 Definice klasifikační nomenklatury a tvorba referenčních vrstev LULUCF zájmových k. ú.	31
4.2 Porovnání vybraných datových zdrojů na řešených k. ú.	34
4.3 Použitá data a tvorba kompozitů pro klasifikace	35
4.4 Tvorba trénovacích a validačních dat	37
4.5 Použité klasifikátory	40
4.6 Klasifikace	42
5 Výsledky	46
5.1 Porovnání vybraných datových zdrojů na řešených k. ú.	46
5.2 Výsledky klasifikace.....	48
5.2.1 Výsledky klasifikace Random Forest.....	48
5.2.2 Výsledky klasifikace Gradient Tree Boosting.....	55
5.2.3 Srovnání klasifikátorů.....	56
6 Diskuse	59
7 Závěr	65
Zdroje	67
Seznam příloh.....	73

Seznam zkratk

ČHMÚ Český hydrometeorologický ústav

ČSÚ Český statistický úřad

ČÚZK Český zeměměřický úřad katastrální

DPZ dálkový průzkum Země

GEE Google Earth Engine

GTB Gradient Tree Boosting

IPCC International Panel on Climate Change (Mezivládní panel pro změnu klimatu)

k. ú. katastrální území

KN katastr nemovitostí

LULUCF Land Use, Land Use Change and Forestry

LPIS Land Parcel Identification System (Veřejný registr půdy)

MMU minimální mapovací jednotka

MŽP Ministerstvo životního prostředí

POP polygony odpovídající pixelům

NIZ Národní inventarizační zpráva

RF Random Forest

ÚHÚL Ústav pro hospodářskou úpravu lesů

UNFCCC United Nations Framework Convention on Climate Change (Rámcová úmluva OSN
o změně klimatu)

ZM10 Základní mapa 1 : 10 000

1 ÚVOD

Změna klimatu patří mezi hojně diskutovaná témata současného světa. Hlavní příčinou jejího urychlení je zesílení přirozeného skleníkového efektu atmosféry Země vlivem antropogenních emisí skleníkových plynů. Skleníkové plyny jsou emitovány do atmosféry například spalováním fosilních paliv, z různých průmyslových a zemědělských provozů, z likvidace odpadů, ale také například vlivem odlesňování či dalších změn v krajině. Legislativa Evropské Unie nařizuje členským státům každoročně vykazovat emise či propady skleníkových plynů za jednotlivé sektory lidské činnosti. Odlesňování a změny využití krajiny jsou zahrnuty v sektoru označovaném jako LULUCF (Land use, Land use Change and Forestry, česky tedy využití krajiny, změny ve využití krajiny a lesnictví).

Tato diplomová práce bude zaměřena právě na sektor LULUCF. Je zřejmé, že přesné informace o změnách využití krajiny jsou jedním z důležitých zdrojů dat o emisích či pohlcování skleníkových plynů. Přístupy jednotlivých členských států k LULUCF reportingu se od sebe více či méně odlišují. Tato práce se bude věnovat LULUCF reportingu zejména v podmínkách Česka. Teoretická část práce bude spočívat v představení definic jednotlivých tříd tak, jak jsou využívány k vykazování emisí skleníkových plynů ze sektoru LULUCF v Česku. V praktické části bude testována možnost klasifikace těchto tříd z dat dálkového průzkumu Země (DPZ). Pro tyto účely budou využívána multispektrální data Sentinel-2 a testovány klasifikační algoritmy Random Forest a také Gradient Tree Boosting. Hlavní pozornost při klasifikacích bude věnována testování různých způsobů, jak vytvořit trénovací dataset a natrénovat klasifikační model. Ke trénování klasifikátorů budou využita data získaná z důkladného terénního mapování zvolených katastrálních území. Hlavním cílem práce je vytvoření robustního klasifikačního modelu, schopného pokrýt co největší heterogenitu krajiny Česka a využitelného pro klasifikaci rozsáhlejších územních celků. Takto vytvořený klasifikační model a tato DP celkově by se mohly stát významným přínosem pro účely LULUCF reportingu v Česku s využitím dat DPZ.

Z výše uvedeného lze definovat hlavní cíle práce:

- rešerše problematiky LULUCF se zaměřením na Česko
- definování klasifikační nomenklatury LULUCF v souladu s požadavky LULUCF a využitelností pro klasifikaci dat DPZ
- terénní mapování zájmového území za využití definované nomenklatury LULUCF
- přehled a vzájemné porovnání vybraných existujících land cover/land use databází
- tvorba trénovacího datasetu a testování různých způsobů natrénování klasifikačního modelu

- klasifikace krajiny z dat Sentinel-2 pomocí klasifikátorů Random Forest a Gradient Tree Boosting za využití platformy Google Earth Engine
- vyhodnocení přesnosti, interpretace výsledků a diskuse použité metodiky se zaměřením na potenciál využití výstupu DP k dalšímu výzkumu a využití v praxi

Výzkumné hypotézy a otázky

Jak jsou vymezeny třídy sektoru LULUCF v Česku?

Jak sestavit nomenklaturu LULUCF tak, aby byla v souladu s LULUCF reportingem a zároveň využitelná pro terénní mapování a klasifikaci dat DPZ?

Jak vytvořit trénovací data a algoritmus pro klasifikaci tříd LULUCF z dat DPZ pomocí strojového učení? Jak tuto metodu nastavit, aby klasifikace dosahovala co nejlepších výsledků?

Jaký je potenciál využití dat Sentinel-2 pro klasifikaci krajiny pro účely LULUCF reportingu?

2 ÚVOD DO PROBLEMATIKY A LITERÁRNÍ REŠERŠE

Přítomnost skleníkových plynů v atmosféře je nezbytná pro život na Zemi, jaký známe. Tyto plyny způsobují skleníkový efekt, tedy umožňují zadržování tepla v atmosféře Země. Sluneční záření prochází atmosférou a ohřívá povrch Země. Povrch Země následně vyzařuje teplo zpět do atmosféry ve formě dlouhovlnného záření. Skleníkové plyny toto záření pohlcují a způsobují ohřev atmosféry. Bez přítomnosti atmosféry by průměrná teplota na Zemi byla 18 °C, s atmosférou a za působení skleníkového efektu je to 15 °C. Nadměrné emise skleníkových plynů v důsledku lidské činnosti mohou nicméně způsobit zesílení tohoto přirozeného efektu.

Hlavními skleníkovými plyny jsou CO₂, CH₄, N₂O, tzv. F-plyny (HFC), přízemní ozon a vodní pára. Z těchto plynů má oxid uhličitý v atmosféře nejvyšší koncentraci, avšak vykazuje nejnižší skleníkový efekt na jednotku hmotnosti. Jako nepřímé skleníkové plyny jsou považovány NO_x, CO, NMVOC (nemethanové těkavé organické sloučeniny) a SO₂. Nepřímé znamená, že se přímo nepodílí na skleníkovém efektu, ale jejich přítomnost v atmosféře může klima také ovlivňovat (ČHMÚ 2023).

Většina antropogenních emisí CO₂ pochází ze spalování fosilních paliv a stavebnictví (výroby cementu), významný podíl mají i změny ve využívání půdy, zejména odlesňování. Antropogenní emise CH₄ mají původ zejména v těžbě uhlí, transportu zemního plynu, chovu hospodářských zvířat, odpadového hospodářství a hospodaření se živočišnými odpady. Emise N₂O způsobené lidskou činností pocházejí především ze zemědělství, spalování biomasy a některých průmyslových činností (ČHMÚ 2024).

2.1 Legislativní rámec

Evropská unie se jako smluvní strana Kjótského protokolu z roku 1997 a Pařížské dohody z roku 2015 zavázala podílet na celosvětové snaze snižovat emise skleníkových plynů. Pařížská dohoda formuluje dlouhodobý cíl ochrany klimatu, jímž je udržet globální nárůst teploty výrazně pod úrovní 2 °C ve srovnání s úrovní před průmyslovou revolucí a usilovat o to, aby tento nárůst nepřesáhl 1,5 °C ve srovnání s úrovní před průmyslovou revolucí (MŽP 2023). Pařížská dohoda tímto navazuje na Kjótský protokol, jehož platnost skončila v roce 2020.

K dosažení cíle vytyčeného Pařížskou dohodou je zapotřebí, aby EU do roku 2050 dosáhla uhlíkové neutrality (Savaresi a kol. 2020). Podle nařízení Evropského parlamentu přijatého v roce 2018 (EP 2018) má každý členský stát EU povinnost zajistit, aby byly emise skleníkových plynů v odvětví LULUCF v období 2021–2030 kompenzovány alespoň

rovnocenným množstvím pohlcení uhlíku. V rámci LULUCF nařízení mají členské státy povinnost vést záznamy o emisích skleníkových plynů v tomto sektoru. Bilance skleníkových plynů členské státy publikují v dokumentu s názvem Greenhouse Gas Inventory (GHGI), v českém jazyce Národní inventarizační zpráva (NIZ). Tyto zprávy jsou odhadem ročních emisí skleníkových plynů vyprodukovaných lidskou činností na území dané země. Cílem GHGI je reflektovat, jak se ekonomické aktivity státu a jeho uhlíková stopa odrážejí na produkci či pohlcování skleníkových plynů. Souhrnné GHGI za celou EU vydává Evropská komise, v rámci plnění závazku Rámcové úmluvy OSN o změně klimatu (UNFCCC, United Nations Framework Convention on Climate Change). Skleníkové plyny, které jsou v rámci nařízení o LULUCF sledovány, jsou CO₂, NH₄ a N₂O.

2.2 Sektor LULUCF

Sektor LULUCF zahrnuje využívání půdy, změny ve využívání půdy a lesnictví. LULUCF může fungovat jako carbon sink (ukládání uhlíku v biomase, půdě a vytěženém dřevě), ale i jako zdroj emisí skleníkových plynů (deforestace, požáry vegetace). Ve světovém měřítku je sektor LULUCF zdrojem cca 11 % emisí CO₂ (IPCC 2019), na úrovni EU převládá propad uhlíku nad jeho emisemi (EEA 2023). Kromě emisí ze sektoru LULUCF jsou v rámci UNFCCC reportovány také emise ze sektorů energetika, průmyslové procesy a využití produktů, zemědělství a odpad. LULUCF je nicméně jediným sektorem, který skleníkové plyny jak emituje, tak pohlcuje.

Třídy klasifikace krajiny, ve kterých je LULUCF reporting realizován, jsou definované Mezivládním panelem pro změnu klimatu (IPCC) a jde o: Forest land, Cropland, Grassland, Wetlands, Settlements a Other land. Názvy tříd jsou z důvodu eliminace zavádějícího překladu v textu ponechány v původním anglickém znění.

Definice jednotlivých tříd dle Penman a kol. (2003) jsou následující:

Forest land = všechno území s dřevinnou vegetací, která splňuje prahové hodnoty danými jednotlivými státy. Zahrnuje také biotopy, kde je vegetace zatím nízká či řídká a nesplňuje definované parametry lesa, ale předpokládá se, že v budoucnu vyroste a prahových hodnot dosáhne.

Cropland = orná půda, obdělávaná půda a zemědělsko-lesnické systémy, kde vegetace nesplňuje prahové hodnoty potřebné pro třídu Forest land.

Grassland = louky, pastviny a plochy s vegetací, které nesplňují prahové hodnoty třídy Forest land, a bez lidského zásahu se nepředpokládá, že by těchto hodnot dosáhly.

Settlements = veškeré zastavěné území.

Wetlands = území, která jsou pokrytá nebo satureovaná vodou celý rok nebo alespoň po část roku, a nespádají do jiné třídy. Zahrnuje vodní nádrže, vodní toky i jezera.

Other land = holá půda, skály, zaledněné plochy a území, která nespádají do žádné jiné třídy.

Pro účely efektivního naplňování společného programu EU je klíčové, aby výstupy členských států byly plně porovnatelné mezi sebou. Mezi přístupy k LULUCF reportingu napříč jednotlivými státy ovšem existují značné rozdíly. Blujdea a kol. (2016) publikovali souhrnnou analýzu LULUCF reportingu ve členských státech EU. Zabývají se používanými metodami sběru a analýzy dat, definicemi jednotlivých tříd a konzistencí mezi více různými zdroji dat použitými pro stejný účel.

Odlišnosti mezi údaji jednotlivých států začínají již ve stadiu definic klasifikačních tříd. Například pro definici lesní plochy jsou využívány tři hlavní parametry: minimální rozloha, minimální plocha pokrytá korunami, a minimální výška vegetace. Prahové hodnoty pro tyto parametry se ale u členských států různí, z historických i ekonomických důvodů. Minimální kritérium pro plochu lesa je 0,05 ha, maximální pak 1 ha, minimální kritérium pro plochu pokrytou korunami je 10 %, maximální pak 30 %. V procesu rozhodování o prahové hodnotě je v některých případech více uvažována fragmentace krajiny, v jiných méně. Autoři článku nicméně konstatují, že pro splnění účelu reportování mají rozdílné definice lesa zanedbatelný efekt. Rozdíly v definicích jsou také u tříd Cropland a Grassland. Jejich definice jsou většinou vyjádřeny kvalitativními popisy a někdy se od sebe rozeznávají pomocí doby, po kterou jsou pokryty trávou. Zemědělské plochy toho času ležící ladem jsou v některých státech zahrnovány do třídy Cropland, v jiných do Grassland.

Co se sběru dat týče, nejčastějšími zdroji dat o Forest land jsou státní statistické databáze (například katastrální data), pro Cropland a Grassland jsou to letecké snímky a provozní evidence. Jemnější dělení tříd pak většinou vychází ze statistických údajů ze zemědělského sektoru, například celkové plochy osázené kterými plodinami.

Závěrem autoři konstatují, že ačkoli je v celém procesu napříč členskými státy stále velká heterogenita co se týče sběru dat i definic tříd, pro účely věci je dle autorů přesnější a spolehlivější vycházet z národních dat uzpůsobených prostředím jednotlivých států, než se opírat o celková, pan-evropská data, která, přestože by měla jednotné definice a metody, by byla zobecňující na úkor přesnosti. Navrhovaným možným vylepšením by mohl být strukturovaný systém sdílení dat, metod a „know how“, kterými by se jednotlivé státy mohly vzájemně inspirovat a zkvalitňovat tak celý projekt. Mohl by například pomoci vyřešit rozdílnosti v odhadech emisí v hraničních oblastech, kde dva státy monitorují různými způsoby stejné prostředí.

Z hlediska metodiky výpočtu a celkové bilance skleníkových plynů je LULUCF nejkomplicovanějším sektorem. Nejdůležitější složkou krajiny z hlediska uhlíkové bilance jsou lesy, které pohlcují CO₂ z atmosféry a ukládají uhlík v biomase a v půdě. Do jaké míry a na jak dlouho jsou schopny lesy, vegetace a půda vázat uhlík je značně závislé na způsobu využívání krajiny. Výsadba nových lesů a správné hospodaření s půdou zvyšuje sekvestraci atmosférického CO₂ v uhlíkových rezervoárech. Na druhou stranu odlesňování, intenzivní využívání půdy a odvodňování rašelinných půd zvyšuje emise CO₂ z uhlíku uloženého v půdě a biomase (Hendriks a kol. 2021). Další negativní vlivy na lesní ekosystémy, a tudíž emise skleníkových plynů mají lesní požáry, bouře či kůrovcové kalamity. Zmíněné vlivy stojí rovněž za poměrně vysokou meziroční variabilitou v uhlíkové bilanci.

2.3 LULUCF reporting v podmínkách Česka

V Česku je vytvářením a publikací NIZ pověřen Český hydrometeorologický ústav (ČHMÚ). Ten koordinuje další organizace, které připravují data o emisích skleníkových plynů v rámci jednotlivých sektorů. Za inventuru skleníkových plynů ze sektoru LULUCF je zodpovědný Ústav pro výzkum lesních ekosystémů (Institute of Forest Ecosystem Research, IFER), a částečně i Ústav výzkumu globální změny Akademie věd (Global Change Research Institute of the Czech Academy of Sciences, GCRI). IFER využívá data Českého statistického úřadu (ČSÚ), Českého úřadu zeměměřického a katastrálního (ČÚZK) a Ústavu pro hospodářskou úpravu lesů (ÚHÚL). ČÚZK poskytuje každoročně aktualizovaná data o land use. ÚHÚL poskytuje data o lesním managementu (těžba, přírůstek lesa atd.).

Nejnovější Národní inventarizační zpráva za Česko byla publikována v roce 2024 a vztahuje se k období 1990 až 2022. V roce 2022 byl nejdůležitějším skleníkovým plynem CO₂, který tvořil 81,7 % všech emisí skleníkových plynů. Následoval CH₄ s 10,9 % a N₂O s 4,3 %. Za období 1990 až 2022 klesly emise CO₂ ze sektoru LULUCF o 36,6 %, emise CH₄ o 51,2 % (ČHMÚ 2024). Tabulka 1 uvádí přehled množství emisí skleníkových plynů dle jednotlivých sektorů.

Tabulka 1: Emise skleníkových plynů v roce 2022 dle sektorů

Sektor	Množství emisí (kt CO ₂ eq.)	Podíl na celkových emisích (%)
Energetika	87907,24	72,98
Průmyslové procesy	15045,20	12,49

Sektor	Množství emisí (kt CO ₂ eq.)	Podíl na celkových emisích (%)
Zemědělství	8422,28	6,99
LULUCF	3378,06	2,80
Odpady	5702,11	4,73

Zdroj: ČHMÚ 2024

Tabulka 2 dává přehled o emisích skleníkových plynů jednotlivých složek LULUCF v letech 1990 a 2022. Zatímco v roce 1990 byla složka Forest land charakterizována jako nejdůležitější sink skleníkových plynů, v roce 2022 jde o jeden z významných zdrojů. Na druhou stranu, emise skleníkových plynů ze třídy Cropland se snížily na méně než polovinu, a u třídy Grassland, která je stabilně sinkem, se toto pohlcení významně zvýšilo. Při porovnávání jednotlivých roků mezi sebou je nicméně zapotřebí mít na paměti významnou meziroční variabilitu v rámci LULUCF sektoru.

Tabulka 2: Emise skleníkových plynů LULUCF v letech 1990 a 2022 dle složek

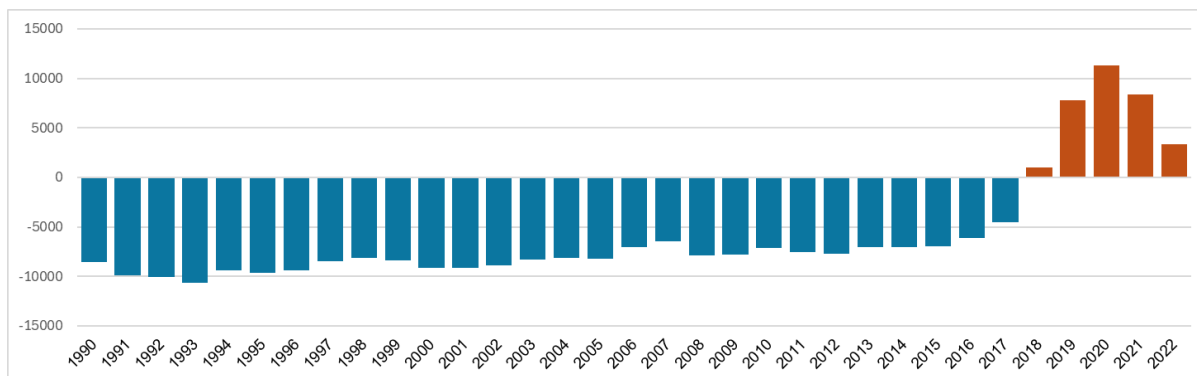
Složka	Množství emisí v roce 1990 (kt CO ₂ eq.*)	Množství emisí v roce 2022 (kt CO ₂ eq.*)
Forest land	-7222	5528,01
Cropland	116	45,28
Grassland	-144	-500,87
Wetlands	24	56,53
Settlements (vč. Other land)	319	194,87
Harvested wood products	-1680	-1946,26

* jako CO₂ eq. jsou přepočítány i ostatní skleníkové plyny

Zdroj: ČHMÚ 2024

V souhrnech emisí je zahrnuta i místně nespécifická třída Harvested wood products. Jedná se o uhlík uložený v produktech vyrobených z vytěženého dřeva.

Obrázek 1 znázorňuje vývoj emisí ze sektoru LULUCF od roku 1990. Od roku 2015 započalo postupné snižování pohlcení skleníkových plynů sektorem LULUCF, ale od roku 2018 je sektor naopak zdrojem jejich emisí. Důvodem je kůrovcová kalamita způsobená extrémními suchy, která české lesy v posledních letech postihla. Jarní oteplení přichází již v dubnu, ne až v květnu, což broukům umožňuje dřívější počátek jarního rojení. Vysoké teploty na jaře i v létě dále přispívají k urychlení jeho vývoje a zvýšení počtu generací. Sucho a vysoké teploty navíc oslabují samotné stromy, které se pak napadení obtížně brání (Biologické centrum AV ČR, 2024).



Obrázek 1: Vývoj emisí skleníkových plynů ze sektoru LULUCF od roku 1990

Zdroj: ČHMÚ 2024

LULUCF reporting je v Česku prováděn na podkladě katastrálních dat. Tabulka pro převod třídy LULUCF a druhu využití pozemku dle KN, publikovaná v Národní inventarizační zprávě 2022 (ČHMÚ 2024), je uvedena v Tabulce 3.

Tabulka 3: Převod druhu pozemku dle KN na třídu LULUCF dle NIZ 2022

LULUCF	KN
Forest land	10 Lesní pozemek
Cropland	2 Orná půda 3 Chmelnice 4 Vinice 5 Zahrada 6 Ovocný sad
Grassland	7 Trvalý travní porost
Settlements	13 Zastavěná plocha a nádvoří 14 Ostatní plocha

Wetlands	11 Vodní plocha
Other land	od roku 2018 pod Settlements

Forest land

Třída Forest land je definována jako dřevinný porost s rozlohou minimálně 0,05 ha, pokrytí plochy korunami je alespoň 30 % a stromy mají v dospělosti minimálně 2 m do výšky. Liniové lesní porosty mají šířku minimálně 20 m.

Co se týče výpočtu uhlíkové bilance třídy forest land, informace o lesích jsou získávány z více zdrojů. Primárním zdrojem jsou data ÚHÚL, který vytváří tzv. Lesní hospodářské plány. Každoročně je v této databázi aktualizována 1/10 celkové lesní plochy rozptýlená po celé zemi. Doplňkovým zdrojem informací o lesích jsou data Národní inventarizace lesů (NIL), které rovněž realizuje ÚHÚL. Metodou NIL je systematický sběr dat založený na statistické teorii. Informace jsou získávány pravidelným terénním průzkumem na trvale umístěných inventarizačních plochách, které pokrývají celé území České republiky. ÚHÚL realizoval tři inventarizační cykly NIL v letech 2001–2004, 2011–2015 a 2016–2020. Dalším pomocným zdrojem jsou výstupy z celorepublikového statistického šetření Inventarizace krajiny CzechTerra, kombinující analýzu leteckých snímků a terénní šetření (CzechTerra 2015).

Specifickou podtřídou Forest land jsou tzv. clear-cut areas. Clear-cut area je definována jako dočasně nezalesněná oblast po definitivním nebo záchranném vykácení lesa. Přestane existovat, jakmile je opět zalesněna, což musí podle českého lesního zákona nastat do dvou let. Pro tuto třídu není zjištěna žádná změna uhlíkových zásob a je zavedena výhradně pro účely konzistentní zprávy o lesních pozemcích (ČHMÚ 2024).

Další důležitou informací představuje roční těžba lesů. Tyto informace zajišuje ČSÚ, který sbírá informace od více než 600 vlastníků lesů. Inventarizace emisí zahrnuje také informace o ztrátách při sklizni (harvest loss), což představuje dodatečné kácení zejména z důvodu přírodních kalamiť a disturbancí, ale také jako dodatečné kácení během plánované sklizně. Nejvýznamnější kalamitou vyžadující záchranné kácení byla kůrovcová kalamita, která dosáhla vrcholu v roce 2020. Dalšími disturbancemi jsou pak například silný vítr či sníh. Poškození způsobené znečištěním je oproti 80. a počátku 90. let zanedbatelné.

Cropland

Třída Cropland je v Česku reprezentována druhy pozemku dle KN orná půda, chmelnice, vinice, ovocné sady a zahrady. Cropland je nejzastoupenější třídou v Česku, avšak od 70. let dochází k trvalému snižování jejího podílu. Zatímco v roce 1990 pokrývala třída Cropland téměř 44 % plochy Česka, do roku 2022 se tento podíl snížil na cca 40 %. Orná

půda mizí ve prospěch travních porostů (zejména v méně produktivních horských regionech), a sídelní a dopravní infrastruktury (ČHMÚ 2024).

Grassland

Jako třída Grassland je v Česku považován druh pozemku dle KN trvalý travní porost. Od 90. let se rozloha třídy Grassland postupně zvyšuje, od roku 1990 do roku 2022 se zvýšila o 24 %. Nejzastoupenější jsou travní porosty v podhorských a horských málo úrodných regionech, a zatravnění je navíc státem dotačně podporováno.

Z hlediska emisí skleníkových plynů je třeba rozlišovat travní porosty na extenzivně či intenzivně obhospodařované, nebo travní porosty bez produkční funkce. Tyto informace jsou získávány pomocí statistických průzkumů Farm Production Surveys, které provádí všechny státy EU (Eurostat 2023). V Česku jsou výsledky těchto průzkumů publikovány ČSÚ (ČSÚ 2014). Po roce 2013 se zvýšil podíl extenzivně obhospodařovaných travních porostů, což mělo za následek zvýšený carbon sink ve třídě Grassland (ČHMÚ 2024).

Settlements

Od roku 2018 se do třídy Settlements zahrnují i plochy, které byly předtím zahrnovány do třídy Other land. V současnosti tedy třída Settlements zahrnuje dva druhy pozemků dle KN, zastavěná plocha a nádvoří (13), a ostatní plocha (14). Třída Settlements zahrnuje mimo sídelní, dopravní a technické infrastruktury, sportovišť či dobývacích prostor také vojenské prostory, které byly v minulosti zahrnovány do třídy Grassland (ČHMÚ 2024).

Wetlands

V Česku je třída Wetlands většinou odpovídá druhu pozemku KN vodní plocha, dále jsou do třídy Wetlands zahrnuty podmáčené půdy, které v KN spadají pod druh pozemku ostatní plocha (14) a způsob využití pozemku zamokřená plocha (11). V této třídě jsou zahrnuty vodní toky, rybníky, jezera, vodní nádrže a mokřady. V roce 2022 plocha třídy Wetlands pokrývala cca 2 % Česka, a tato plocha průběžně narůstá. To je dáno zejména programy na podporu retence vody v krajině, které podněcují budování větších či menších vodních nádrží a dalších vodních prvků (ČHMÚ 2024).

Other land

Od roku 2018 není třída Other land samostatně vykazována a její emise jsou připočítány emisím třídy Settlements. Před rokem 2018 byla třída Other land v Česku reprezentována neobhospodařovanými (neobhospodařovatelnými) plochami. Šlo o druh pozemku dle KN ostatní plocha, a způsob využití pozemku neplodná půda. Důvodem uvedeným pro tuto změnu je to, že v podmínkách Česka plochy většinou nejsou nedotčené, do jejich využívání je

nějakým způsobem člověkem upravováno, a tudíž nesplňují kritérium, že do nich není zasahováno. Dle Penmana a kol. (2003) do třídy Other land patří také holá půda, skály a zaledněné oblasti. První dvě zmíněné se v Česku vyskytují pouze okrajově a třetí se nevyskytuje vůbec, proto není nutné sloučení třídy Other land s třídou Settlements považovat za zdroj výraznějších nepřesností v reportování emisí (ČHMÚ 2024).

2.4 Další metody klasifikace krajinných tříd (podobných)

LULUCF

V Česku je primárním zdrojem dat pro LULUCF katastr nemovitostí. Tento zdroj má nicméně své rezervy a v některých případech obsahuje zastaralá či záměrně neaktualizovaná data (Pazúr a kol. 2015, Míček a kol. 2020). Jednou z výzev na poli geoinformatiky je klasifikace tříd LULUCF za pomoci dat dálkového průzkumu Země a využití moderních klasifikačních algoritmů.

Svoboda a kol. (2022) ve své studii testují potenciál dat Sentinel-2 pro účely klasifikaci tříd LULUCF. Ze snímků byla vytvořena mozaika, která obsahovala také SRTM model a informaci o ročním rozptylu NDVI. Tvorba testovacího datasetu proběhla částečně automatizovaně na základě databáze CLC, jejíž třídy byly převedeny na své ekvivalenty v LULUCF nomenklatuře, ale pro dostatečné zastoupení i méně hojných či specifických typů povrchů bylo přikročeno také k manuálnímu výběru trénovacích polygonů. Ke klasifikaci autoři využívají metodu Random Forest. Výhoda této metody, jak autoři konstatují, spočívá mimo jiné v tom, že si dokáže poradit i s rozdílnými typy povrchů v rámci jedné třídy LULUCF, tudíž tyto povrchy nemusí být klasifikovány zvlášť a poté agregovány. V rámci studie je testováno celkem 450 kombinací parametrů Number of Trees, Variables Per Split a Bag Fraction, nejúspěšnější kombinace dosáhla celkové přesnosti 89,01 %. Co se týče jednotlivých tříd, třída Settlements byla nejčastěji chybně klasifikována jako Cropland, třída Cropland byla nejčastěji chybně klasifikována jako Grassland a naopak. Svoboda a kol. (2022) na základě svých výsledků shledávají, že navržená metoda a využití dat Sentinel-2 programu Copernicus se pro systematickou implementaci dat DPZ pro účely LULUCF jeví velmi perspektivně.

Další důležitou třídou LULUCF z hlediska pohlcování uhlíku je třída Grassland. V Evropě tohoto biomu stále ubývá (Schils a kol. 2022), což sebou mimo výše zmíněných rizik nese také snížení biodiverzity (Sperry a kol. 2019). Relevantním výzkumným tématem je detekce travních porostů, obdělávané půdy a jejich přeměn pomocí dálkového průzkumu Země, a to zejména pomocí časových řad či snímků velmi vysokého rozlišení.

Dle Esch a kol. (2014) nelze provádět klasifikaci travních porostů a obdělávané půdy založenou pouze na spektrálních charakteristikách, protože tyto se v průběhu vegetační sezóny výrazně proměňují, kromě toho jsou si spektrální charakteristiky některých zemědělských plodin a travin v některých obdobích roku velmi podobné. Autoři se ve své studii zabývají kombinací high resolution a medium resolution dat a trénovacích dat z existující databáze LUCAS. Postup spočívá v segmentaci krajiny na homogenní parcely pomocí high resolution dat, a dále za pomoci medium resolution dat získání charakteristiky sezónních proměn vegetace. Sezónní proměnlivost vegetace v jednotlivých segmentech je vyjádřena průměrným NDVI a údaji, v jakých částech roku nabývá extrémních hodnot. Ke klasifikaci je využit klasifikátor C5.0. Autorům se tímto postupem podařilo dosáhnout celkové přesnosti 86 %.

Šandera a Štych (2020) se zabývali mapováním přeměn třídy trvalých travních porostů na obdělávanou půdu. Cílem autorů bylo identifikovat nejvhodnější biologické proměnné pro nejefektivnější zachycení těchto změn. Zkoumané biologické proměnné byly FAPAR (fraction of absorbed photosynthetically active radiation), FCOVER (fraction of vegetation cover), LAI (leaf area index), CAB (chlorophyll content in the leaf), CWC (canopy water content) a NDVI (normalized difference vegetation index). K tomuto účelu autoři využívají klasifikátor Random Forest a jeho parametr mean decrease accuracy (MDA). Vhodnost prediktorů je posuzována s ohledem na čas pořízení scény, celkovou přesnost zachycení změn a také čas potřebný k výpočtu. Je vyvinuta snaha o minimalizaci počtu vstupních prediktorů při stejné dosažené přesnosti. Autoři využívali Sentinel-2 time-series data, jako trénovací data byl využit LPIS. Jako nejvhodnější prediktory byly vyhodnoceny FAPAR, CWC a LAI. Na druhou stranu, nejnižších přesností klasifikace bylo dosaženo s FCOVER a NDVI. Detekce změn celkově dosáhla vysoké přesnosti, přes 95 %.

Výběrem vhodných proměnných pro detekci opačných změn, tedy z travních porostů na obdělávanou půdu, se zabývali Klouček a kol. (2018). Autoři pracovali se snímky vysokého rozlišení Landsat 8, jako trénovací data byl využit LPIS. Bylo testováno téměř šest desítek proměnných indikujících změnu (vegetační indexy, texturní proměnné, komponenty PCA, komponenty transformace Tasseled Cap ad.) a s těmito proměnnými byly vytvářeny modely pomocí GLM (Generalised Linear Models). Pro klasifikaci byl použit algoritmus SVM. Základní otázka spočívala v tom, kolik proměnných a jaké jsou zapotřebí ke spolehlivé klasifikaci sledovaných změn. Jako nejefektivnější model byl vyhodnocen model pracující se třemi proměnnými: NDVI, Wetness, a pátá komponenta Tasseled Cap. Zvyšování počtu proměnných vstupujících do výpočtu nevedlo k výraznému zvýšení přesnosti klasifikací. Nejúspěšnější klasifikace změn dosáhla celkové přesnosti 89 %.

Sledování, analýzy a predikce vývoje emisí ze sektoru LULUCF nejsou záležitostí jen Evropské unie. Latta a kol. (2018) vyvinuli model LURA (Land Use and Resource Allocation) určený k prostorově lokalizovaným odhadům emisí či pohlcení uhlíku v lesích ve Spojených státech amerických. Autoři vyjadřují obavu, že faktory jako stárnutí lesů, zvyšující se poptávka po produktech ze dřeva, urbanizace či vliv klimatické změny mohou zapříčinit zmenšení tohoto uhlíkového sinku. Záměrem autorů je modelovat vývoj emisí a sekvestrace CO₂ na základě vybraných makroekonomických ukazatelů a ilustrovat, jak moc se mohou odhady emisí či pohlcení uhlíku souvisejících s lesními ekosystémy lišit právě v závislosti na ekonomickém vývoji státu. Hlavními ekonomickými indikátory, se kterými LURA pracuje, je HDP, zahájená výstavba, cena ropy, spotřeba dřeva a využití bioenergie. I když je LURA ekonomickým modelem, jeho důležitým přínosem je zahrnutí prostorových dat. Model pracuje s detailními prostorovými daty o zalesňování, růstu lesů, odlesnění i regeneraci lesů. Predikce modelu LURA pro období 2014–2035 ukazují, že lesy v USA pravděpodobně zůstanou významným carbon sinkem, avšak roční pohlcení uhlíku se bude snižovat v důsledku poptávky po produktech ze dřeva a stárnutí lesů. Vyšší ekonomický růst bude znamenat rychlejší tempo snižování sekvestrace uhlíku lesy.

2.5 Existující zdroje land cover/land use dat

Tato kapitola si klade za cíl představit vybrané globální, volně dostupné a často aktualizované land cover/land use databáze vycházející z dat DPZ.

Corine Land Cover+

V současné době je ve vývoji nová, rozšířená a modernizovaná generace hojně využívaných vrstev Corine Land Cover (CLC), tzv. Corine Land Cover+ (CLC+). Struktura CLC+ sestává ze 3 komponent: CLC+ Backbone, CLC+ Core a CLC+ Instances. Komponentu CLC+ Backbone tvoří prostorová data – rastrové a vektorové vrstvy land cover dat. K únoru 2025 jsou k dispozici ke stažení rastrové vrstvy z referenčních let 2018 a 2021 a vektorová vrstva 2018. Vektorový produkt pracuje s kombinací liniových vrstev tvořících „kostru krajiny“, typicky hydrologické a dopravní sítě (hardbone) a segmentovaných družicových snímků (softbone). V hardbone fázi je vytvořena kostra krajiny pomocí liniových vrstev stabilních linií v krajině. Tato data jsou převzata převážně z open source databází, např. OSM, CLMS-EU Hydro a WISE Water Framework Directive. MMW je 20 m, v místech řidší sítě i méně. V softbone fázi se tato kostra stává podkladem pro segmentaci krajiny na snímcích Sentinel-2, tedy rozdělení krajiny do plošek o podobných spektrálních či texturálních příznacích. Ploškám ve vektorové vrstvě jsou pak přiřazeny třídy na základě klasifikované rasterové vrstvy CLC+ Backbone. Počet tříd u vektorové vrstvy je vyšší než u vrstvy rastrové. Rastrové vrstvy

disponují rozlišením 10 m a 11 třídami krajinného pokryvu, vektorové vrstvy mají MMU 0,5 ha a 18 krajinných tříd. Třídy si navzájem odpovídají, pouze ve vektorovém produktu byly některé původní třídy rastrového produktu pro získání detailnější informace rozděleny (EEA 2020 User manual).

Databázová komponenta CLC+ se nazývá CLC+ Core. CLC+ Core poskytne nástroj pro vytváření „na míru šitých“ produktů pomocí různých kombinací dostupných dat. CLC+ Core bude integrovat existující a budoucí produkty CLMS a národní produkty o land use a land cover pomocí standardizovaného integračního přístupu v souladu s datovým modelem EAGLE. To znamená, že jednotlivé třídy těchto zdrojů budou zachyceny jako prvky modelu EAGLE – tedy rozděleny na komponenty pokrytí půdy (LCC), atributy využití půdy (LUA) a další charakteristiky (CH). Tento způsob umožní kombinace dříve navzájem neharmonizovaných dat, zejména informace o land cover z produktů CLMS se specifickými informacemi o využití půdy z jednotlivých zemí. Takto vytvořené produkty se budou nazývat CLC+ Instances. Půjde například právě o podklady pro LULUCF reporting nebo pro „tradiční“ CLC, aby bylo možné pokračovat v časové řadě.

Vrstvy jsou zatím poskytovány v zobrazení ETRS89 LAEA, produkty pro jednotlivé státy v jejich národních zobrazeních budou postupem času dodány rovněž.

K nejvýraznějším přínosům tohoto produktu patří vyšší prostorové rozlišení oproti původní vrstvě CLC (10 m namísto 100 m pro raster, 0,5 ha namísto 25 ha pro vektor), častější aktualizace dat (3 roky namísto 6 let), a usnadnění integrace získaných dat o krajině s dalšími daty, například hydrologickými či biogeografickými. Novinkou je zcela odlišná klasifikační nomenklatura. Vrstvy CLC+ se řídí pravidly datového modelu EAGLE, který je představen v textu níže. Jedním z nejdůležitějších rozdílů je to, že neobsahuje klasifikační třídy kombinující land cover a land use, což u vrstev CLC znemožňovalo některé druhy využití. Databáze bude také obsahovat informace o druhé či třetí nejzastoupenější třídě v daném polygonu, která je uchována v jeho attributech. V neposlední řadě bude obsahovat také atributy, které reflektují vývoj polygonu v čase (například jakými změnami či disturbancemi prošel).

CLC+ Backbone je v současné době využíván k podpoře implementace nařízení o zahrnutí emisí skleníkových plynů a sekvestrace z LULUCF. Vývoj produktu CLC+ LULUCF Instance i jeho praktické využití probíhá v rámci pracovní skupiny LULUCF při EEA. Již existující datové sady (od r. 2021) jsou v současnosti dostupné pouze na vyžádání. Rezervou oproti vrstvám CLC+ Backbone je prostorové rozlišení 100 m (Copernicus.eu 2025).

ESA World Cover

ESA World Cover představuje globální land cover produkt Evropské kosmické agentury (ESA), vytvořený pomocí kombinace dat ze satelitů Sentinel-1 a Sentinel-2, tedy za pomoci

optických i radarových dat. Vrstvy ESA World Cover je vztaženy k referenčním rokům 2020 a 2021. Kromě globálního pokrytí a volného přístupu k datům patří k devizám tohoto produktu jeho vysoké prostorové rozlišení, které činí 10 m. Pro rok 2020 dosahovala celková přesnost produktu 75 %, pro rok 2021 pak 76,7 %. Krajinový pokryv je členěn do 11 tříd dle nomenklatury FAO Land Cover Classification System (Gregorio, Jansen 2000).

Nejsnazší cestou k prohlížení či stahování dat je jednoduchá webová mapová aplikace WorldCover viewer. Data je možné stáhnout za administrativní celky či za jakoukoli oblast definovanou uživatelem. V aplikaci jsou k dispozici také mediánové kompozity Sentinel-1 a Sentinel-2, ze kterých byl produkt vytvářen. Ke stažení jsou data k dispozici ve formátu Cloud Optimised GeoTIFF. Další možností je například využití aplikace World Cover vytvořené pomocí platformy GEE, která umožňuje porovnání vrstvy ESA World Cover a kompozitů Sentinel-1 a Sentinel-2 pomocí nástroje Swipe.

Dynamic World

Dalším zajímavým zdrojem land cover dat je Dynamic World, který byl vyvinut ve spolupráci mezi Google a World Resources Institute. Jde o globální dataset s prostorovým rozlišením 10 m, vycházející ze snímků Sentinel-2. Ke klasifikaci a zpracování dat je využíváno moderních deep learning metod, díky čemuž je jsou data Dynamic World k dispozici v téměř reálném čase – jsou aktualizována každých 2 až 5 dní v závislosti na zeměpisné šířce dané lokality. Časová řada dat Dynamic World sahá až do roku 2015.

Data mají deset pásem. Devět pásem udává každému pixelu pravděpodobnost, s jakou náleží do níže popsaných tříd land cover, a jedno pásmo udává třídu s nejvyšší pravděpodobností. Krajinový pokryv je dělen do 9 tříd.

Autoři deklarují, že nomenklatura krajinových tříd byla vytvořena v souladu se třídami LULUCF dle IPCC Good Practice Guidance, aby bylo usnadněno využití dat Dynamic World k výpočtům emisí skleníkových plynů (Brown a kol. 2022).

Dataset je dostupný v prostředí GEE, kde je možné využívat ho k dalším pokročilým analýzám. Mimo GEE jsou data volně k prohlížení k dispozici také na ResourceWatch.

ESRI Land Cover

ESRI Land Cover je další z volně dostupných globálních datových zdrojů o krajinovém pokryvu. Dataset byl vytvořen společností ESRI, vychází ze snímků Sentinel-2, nabízí prostorové rozlišení 10 m a je dostupný pro roky 2017 až 2023. Od roku 2018 je vytvářen z celoroční sady snímků Sentinel-2. Byl vytvořen pomocí klasifikačních modelů založených na deep learning. ESRI Land Cover rozděluje krajinu do 9 tříd.

Dataset ESRI Land Cover je součástí dynamické a často aktualizované sbírky geografických dat společnosti ESRI, Living Atlas. Living Atlas shromažďuje a zpřístupňuje

data z celého světa ve formě map, datových sad a vrstev. ArcGIS Living Atlas of the World je přístupný v prostředí ArcGIS Online, což uživatelům nabízí možnost jednoduše data ESRI Land Cover kombinovat či integrovat s dalšími produkty Living Atlas.

Samotná vrstva ESRI Land Cover je však zpřístupněna i v samostatné jednoduché webové mapové aplikaci Sentinel-2 Land Cover Explorer, ve které je možné prohlížet data za každý referenční rok, prohlížet zdrojové satelitní snímky či stahovat data ve formátu GeoTIFF nebo rastrového obrázku.

Nomenklatury všech představených datasetů jsou v Příloze 1.

Model EAGLE

Klasifikačních systémů a nomenklatur land cover/land use databází již existuje celá řada. Každý klasifikační systém je navrhován s jiným účelem, vychází z jiných zdrojů dat (vedle dat DPZ také například z leteckých snímků, katastrální evidence či terénního mapování), a také z rozdílných prostředí. Z toho pak vyplývají jejich různá měřítka (resp. MMU), tematické zaměření, soustředění na konkrétní krajinné jevy a upozadování jiných, a tak dále.

Rozdílnost klasifikačních systémů komplikuje jejich kombinace či konverze, a tím pádem například vytváření delších časových řad. Častou nevýhodou klasifikačních systémů je také kombinování land cover a land use přístupu. Většinou jsou pomocí land use popisovány plochy v intenzivně využívaných oblastech, a pomocí land cover plochy v „přírodnějších“, extenzivně využívaných oblastech.

Novým evropským standardem pro klasifikační systémy se má stát datový model EAGLE. Snaha o eliminaci výše zmiňovaných komplikací vedla k vytvoření integrovaného datového rámce, který je schopen pokrýt potřeby nejen celoevropské, ale i národní, a zároveň zajistit sémantickou konzistenci pomocí harmonizace toku vstupních dat z různých zdrojů a přístupů – to znamená od dat Sentinel-2 po národní databáze jednotlivých států. Koncept prvně představil Land Information System Austria (LISA), od té doby ho dále vyvíjí Environment Information and Observation Network (EIONET) Action Group on Land monitoring in Europe (EAGLE). Zásadní otázka řešena v rámci projektu zněla: Jak popsat povrch Země koncepčně a jak tyto informace o land use a land cover efektivně a konzistentně uchovávat? Navazující otázky pak zněly: Jaké druhy land use (LU) a land cover (LC) tříd jsou schopna zachytit data DPZ? Jak popisovat krajinu odděleně pro LU a LC? Jak datový model udělat víceúčelový a nezávislý na měřítku a zdrojích dat? Je možné zajistit zpětnou kompatibilitu s existujícími zdroji dat (např. CLC) i po změně sémantického přístupu, bez ztráty informačního obsahu?

EAGLE model se tedy snaží krajinu odděleně popsat z hlediska land cover (Land cover components), land use (Land use attributes), a dalších doplňujících informací, které se nedají jednoduše kategorizovat, tzv. Landscape characteristics. Tyto tři složky představují tři základní hlediska, pomocí kterých je krajina v EAGLE modelu popsána. Takovýto popis krajiny zaručuje, že o krajinné plošce je uchována informace odděleně pro LU i LC, ale zároveň je uchována i jejich kombinace, což společně s Landscape characteristics podává ucelený a komplexní popis oné krajinné plošky.

Model EAGLE je možné reprezentovat 2 způsoby: pomocí EAGLE matrix a EAGLE data modelu.

EAGLE matrix je vyjádřena ve formě tabulky. Jde v zásadě o tři víceúrovňové tabulky, přičemž každá z tabulek představuje jeden z pohledů: LC components, LU attributes a Landscape characteristics. Tyto víceúrovňové tabulky postupují od nejjobecnějšího k nejkonkrétnějšímu. První úroveň LC components zahrnuje povrchy nepokryté vegetací, povrchy pokryté vegetací a vodní plochy, a tyto se dále dělí. LU attributes se v první úrovni dělí na primární, sekundární a terciární hospodářský sektor. Landscape characteristics je nejkompaktnější tabulkou, příklady prvních úrovní jsou způsob obhospodařování, typ plodiny či sezónnost změn. Druhá možná reprezentace, EAGLE data model, je ve formě UML diagramu.

Základem pro popis krajiny je v EAGLE modelu land cover. Ploška krajiny se dá popsat jedním LC komponentem, více LC komponenty, nebo více LC komponenty s procentuálním zastoupením každého z nich. Land cover je pak specifikován dále pomocí LU attributes a Landscape characteristics. Tím vznikne kombinace těchto tří složek, která může být přiřazena ke třídám, které známe z klasických klasifikačních systémů, což nám umožní porovnávat jednotlivé klasifikační systémy mezi sebou. Rozdělení do těchto tří bloků navíc nechává prostor EAGLE konceptu pro modifikaci (např. přidání či ubrání některých položek matice).

Pro porovnávání tříd s klasickými klasifikačními nomenklaturami vznikl systém kódování. Ten funguje tak, že při porovnávání EAGLE matrix s třídou klasické nomenklatury je každému prvku z EAGLE matice přidán kód xx-x-0-1-2-3-4, podle toho, jak moc je pro danou třídu určující. (xx znamená, že je atribut pro třídu irelevantní, 4 znamená přítomnost více než jednoho klíčového a určujícího atributu). Například pro třídu „ledovce a trvalý sníh“ (z CLC) bude irelevantní složka EAGLE matice „skupinky stromů“. Naopak třída smíšené hospodářské kultury (z CLC) bude muset obsahovat alespoň dvě složky z EAGLE matice „obdělávané plodiny“, „trvalé plodiny“ či „trvalé travní porosty“. Pomocí kódů složek matice, které byly přiřazeny každé třídě na základě jejich definic v klasifikační nomenklatuře, je možné mezi sebou porovnávat definice tříd různých klasifikačních systémů, či odhalit různé překryvy či mezery mezi definicemi tříd v rámci jednoho klasifikačního systému (Arnold a kol. 2013).

3 ŘEŠENÁ KATASTRÁLNÍ ÚZEMÍ

V rámci této diplomové práce bylo podrobně zmapováno šest katastrálních území (k. ú.) v Česku, jejichž hlavní charakteristiky shrnuje následující kapitola. Tato území dále posloužila zaprvé pro porovnání vybraných existujících zdrojů land use dat s reálným stavem, a zadruhé jako ground truth data pro klasifikaci. Byla zvolena k. ú. České Vrbné, Chvaletice, Klec, Hojovice, Prachovice a Škrdlovice.

Výběr katastrálních území nebyl proveden náhodně, ale cíleně tak, aby každé území vykazovalo co největší vnitřní heterogenitu a zahrnovalo široké spektrum typů využití krajiny. Proces výběru takovýchto k. ú. probíhal následujícím způsobem. Do polygonové vrstvy Corine Land Cover 2018 pro území celého Česka byl přidán atribut určující příslušnost polygonů k třídám LULUCF. Převod mezi třídami CLC (atribut Code_2018) a třídami LULUCF byl proveden dle převodní tabulky z publikace Svoboda a kol. (2022).

Na polygonovou vrstvu byla následně použita funkce *Dissolve*, čímž vznikl pro každou třídu LULUCF jeden multipolygon. Dále byl proveden průnik (*Intersect*) této vrstvy s vrstvou katastrálních území ČR (ArcČR 500). Výsledkem byla vrstva polygonů LULUCF s přiřazenou informací o k. ú. Každé k. ú. bylo tedy rozděleno do maximálně 6 plošek dle tříd LULUCF. Pro každý polygon byla vypočítána rozloha, též se získala celková plocha jednotlivých tříd LULUCF v každém k. ú. Z těchto ploch byl vypočítán Shannonův index diverzity dle vzorce

$$H = \sum P_i \times \ln(P_i) ,$$

kde P_i podíl rozlohy každé třídy LULUCF ku celkové rozloze příslušného katastru. Index byl vypočítán samostatně pro každé k. ú. Výsledkem byla tabulka hodnot Shannonova indexu diverzity pro všechna k. ú. v zájmových krajích. Tento index dále fungoval jako jedno z kritérií pro výběr k. ú.

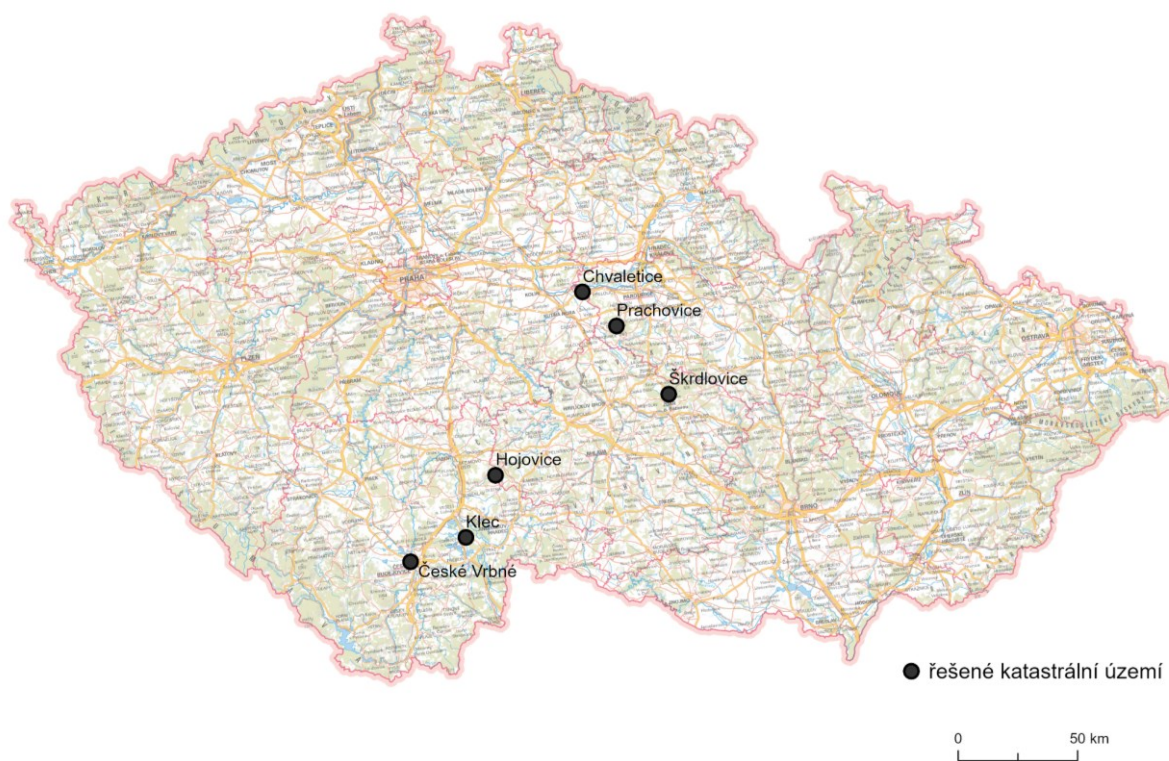
V každém kraji byla vybrána dvě k. ú., která měla v rámci kraje jedno z nejvyšších H a zároveň rozlohu minimálně 5 km². Do rozhodování o výběru k. ú. však vstupovaly i subjektivní faktory, například snaha o navzájem co nejrozmanitější k. ú. (například z hlediska způsobů hospodaření) či dopravní dostupnost a rozloha (aby bylo realizovatelné terénní mapování provést maximálně během dvou dní.).

Tabulka 4 uvádí základní informace o zvolených k. ú., Obrázek 2 je pak přehledovou mapou jejich polohy.

Tabulka 4: Vybrané základní informace o řešených katastrálních územích

Katastrální území	ORP	Počet obyvatel (k 31. 12. 2023)	Minimální nadmořská výška	Maximální nadmořská výška
České Vrbné	České Budějovice	529	371,9	387,5
Klec	Třeboň	162	412,9	431,9
Chvaletice	Přelouč	2927	201,2	306,8
Prachovice	Chrudim	1422	373,1	578,9
Hojovice	Pelhřimov	85	507,1	636,8
Škrdlovice	Žďár nad Sázavou	714	600,2	719,7

Zdroj: Veřejná databáze ČSÚ, Geoprohlížeč ČÚZK – Analýzy výškopisu



Obrázek 2: Přehledová mapa řešených k. ú.

Zdroj: ArcČR 500, podkladová mapa: MČR1M (ČÚZK)

České Vrbné

České Vrbné je katastrální území na severozápadě Českých Budějovic. Osada České Vrbné původně plnila funkci hospodářského zázemí Českých Budějovic, v roce 1980 byla k tomuto městu připojena. Ještě na počátku 21. století byla původní zástavba Českého Vrbného od sídliště Vltava v severozápadní části Českých Budějovic oddělena bloky orné půdy. Postupem času se na těchto blocích velmi výrazně uplatnila komerční suburbanizace a množství obchodních center propojilo zástavbu Českého Vrbného a Českých Budějovic v jeden zastavěný celek (internetový zdroj Encyklopedie Českých Budějovic).

Zástavba v k. ú. České Vrbné je zastoupena pouze v jeho jihovýchodní části, a tvoří ji původní obec a výše zmíněná obchodní centra. Rezidenční zástavbu tvoří rodinné domy. Nachází se zde i poměrně rozlehlá zahradní osada. Severovýchodní část území je pokryta bloky orné půdy, s enklávou přírodní památky Vrbenská tůň. Severozápadní hranici k. ú. tvoří tok Vltavy, na němž je vybudován přístav, MVE a vodácký areál se slalomovým kanálem.

Západní část je převážně přírodního charakteru. Na většině rozlohy se rozkládá přírodní rezervace Vrbenské rybníky. Ta zahrnuje rybníky Černiš, Domin, Bažina a Starý a Nový Vrbenský rybník, a okolní louky, mokřady a lesy. Reliéf k. ú. je převážně rovinný, mírně se svažuje z jihozápadu směrem k severovýchodu.

Klec

Katastrální území Klec leží 2 km severovýchodně od města Lomnice nad Lužnicí a 9 km severně od města Třeboň. Území je součástí CHKO Třeboňsko.

Zástavba v k. ú. Klec je soustředěna v jeho střední části. Jde o vesnici typu ulicovka, tedy podlouhlou vesnici s domy řazenými podél místní komunikace. Zástavba sestává zejména z rodinných domů, zemědělských budov a menších rekreačních zařízení. Za posledních 20 let se zástavba obce prakticky nerozrostla. Jižně od obce Klec katastrální území protíná silnice 2. třídy a rozděluje tak toto protáhlé k. ú. na severní a jižní část. Obec je obklopena bloky orné půdy a roztroušenými lesními porosty menší rozlohy.

Značnou část rozlohy k. ú. pokrývají rybníky. V severní části jde o Nadějskou rybníční soustavu s rybníky Naděje, Víra, Láska, Skutek, Dobrá vůle, Blaník, Měkký rybník a Strakatý rybník. Západní hranici k. ú. zde tvoří řeka Lužnice, východní okraj k. ú. v severní části lemují lesy Kolenecké obory. Jižní část k. ú. je charakteru ryze přírodního. Největší část rozlohy zde zaujímají rybníky Klec a Potěšil, dále pak menší rybník Stehlík, lesy a orná půda. Při západním břehu rybníka Klec se nacházejí pozůstatky hrádku Flughaus ze 14. století. Nejstarší rybníky v lokalitě, založeny Jakubem Krčínem, jsou Naděje, Skutek a Potěšil. Nadějská soustava byla doplněna o řadu menších rybníků v poslední čtvrtině 19. století (webové stránky obce Klec).

Chvaletice

Katastrální území Chvaletice leží 4 km východně od města Týnec nad Labem a 9 km západně od města Přelouč.

Chvaletice jsou známé především díky jedné z největších tepelných elektráren na našem území. Elektrárna Chvaletice byla vybudována v 70. letech 20. století, na místě bývalého pyritového lomu, jehož prostor dnes využívá jako odkaliště. Tento lom byl otevřen v 50. letech 20. století, ze značné části na úkor původní obce Chvaletice. Zástavba, která z původních Chvaletic zůstala, dnes nese název Hornická čtvrť. Dnešní obec Chvaletice byla vytvořena jakožto nové sídlo pro nuceně přestěhované obyvatele (Wikipedie – Elektrárna Chvaletice). Tato obec nicméně neleží v k. ú. Chvaletice, ale Telčice.

Elektrárna se nachází v severovýchodní části k. ú. Jižně od ní se rozkládá odkaliště, v jehož východní části postupně probíhá rekultivace, zejména lesnická. K tomuto areálu přiléhá Hornická čtvrť, jejíž zástavba je tvořena převážně staršími obytnými budovami. V jihovýchodní části k. ú. se nachází dva kamenolomy s probíhající těžbou granitu. Tyto lomy jsou obklopeny lesními porosty, lesní porosty se nachází také mezi elektrárnou a obcí Chvaletice. Zbývající část k. ú. zabírá zemědělská půda, s enklávami menších lesních porostů.

Úsek severní hranice k. ú. je tvořen tokem Labe, na jehož protějším břehu je vybudován přístav Chvaletice. V severní části k. ú. protíná také silnice 2. třídy a nevytíženější železniční trať v Čechách, Praha – Česká Třebová.

Prachovice

Katastrální území Prachovice leží přibližně 13 km jihozápadně od Chrudimi a 16 km východně od Čáslavi. Jižně od k. ú. se rozkládá CHKO Železné hory.

Obec Prachovice leží v centrální části k. ú. Zástavbu tvoří převážně rodinné domy, v centrální části několik bytových domů, a na jihu obce nízké panelové sídliště. V obci se také nacházejí zahradní osady a sportovní areál. Tok Habřinka, v jehož údolí je obec situována, protéká k. ú. v západovýchodním směrem, nachází se na něm několik menších vodních nádrží, a ze značné části je zatrubněn.

Osada vznikla ve 14. století z obydlí skalníků, kteří zde lámali kámen na stavbu nedalekého hradu Lichnice. V okolí Prachovic a nedalekého Vápenného Podolu se totiž nachází velice kvalitní ložiska silursko-devonského vápence, rovněž bohatá na drobné krasové útvary a nalezištěm zkamenělin např. trilobitů či graptolitů. Vápencový lom, který se rozkládá východně od zástavby obce Prachovice, na kopci Boukalka, patří s 6 etážemi a hloubkou 60 metrů k největším v Česku (webové stránky obce Prachovice). Severně od lomu se nachází cementárna.

Vápencový lom s přílehlou cementárnou jsou ze severu i jihu obklopeny lesním porostem, les se také nachází jihozápadně od obce. Zbylá plocha k. ú. je pokryta bloky orné půdy, místy i travními porosty.

Hojovice

Katastrální území Hojovice se nachází při západní hranici kraje Vysočina a sousedí s krajem Jihočeským. Většími městy v okolí jsou Pelhřimov přibližně 25 km severovýchodně, a Pacov přibližně 16 km severně.

Zástavba obce se nachází v jihozápadní části k. ú. Je tvořena dominantně přestavěnými budovami hospodářských stavení, dnes plnící funkci rodinných domů. Jde o tzv. hromadnou ves, s nepravidelným průběhem cest a ulic. Na návsi se nacházejí dvě vodní nádrže propojené vodotečí, které odvádí vodu do Hojovického potoka. Středem k. ú. probíhá severojižním směrem málo vytížená dopravní komunikace.

Zmíněný Hojovický potok protéká jižně od obce a tvoří jižní hranici k. ú. Vlévá se do Černovického potoka, tvořícího západní hranici k. ú. Obec je obklopena poměrně rozlehlými bloky převážně orné půdy, ale částečně i travních porostů. Tyto bloky jsou od sebe odděleny buď pásy stromů, nebo častěji zpevněnými či nezpevněnými polními cestami. Severně od centra obce se nachází areál farmy a drobná obytná osada Ovčín. Severní a východní část k. ú. pokrývá široký pruh převážně jehličnatého lesa.

Škrdlovice

Katastrální území Škrdlovice je součástí v CHKO Žďárské vrchy, ve vrcholové části Českomoravské vrchoviny. Leží přibližně 6 km severně od Žďáru nad Sázavou. Ze západu k. ú. Škrdlovice sousedí s rybníkem Velké Dářko, největším rybníkem Českomoravské vrchoviny.

Zástavba obce Škrdlovice se nachází v centrální části k. ú. Jižně od obce protíná k. ú. silnice I. třídy, která tvoří obchvat této obce. Jihovýchodně od zástavby obce se nachází Markytský rybník ležící na bezejmenném toku, který obec odvodňuje. Zástavbu Škrdlovic obklopují bloky zemědělské půdy, místy s enklávami skupin stromů. Mezi bloky půdy vede množství zpevněných i nezpevněných cest. Na jihovýchodě a severu území do k. ú. zasahují lesní porosty. Při okrajích těchto porostů se nacházejí osamocené rekreační či obytné stavby. U východní hranice území se nachází myslivecká střelnice.

4 DATA A METODY

Jak bylo výše řečeno, zájmovým územím bylo 6 katastrálních území (po dvou v krajích Vysočina, Jihočeský a Pardubický). Pro účely splnění hlavních cílů práce byla nejprve definována klasifikační nomenklatura pro účely LULUCF. Dle definované klasifikační nomenklatury byl následně vytvořen referenční dataset k roku 2023. Referenční dataset/vrstva vznikla kombinací terénního mapování a dostupných land use/land cover dat (katastr nemovitostí, letecké snímky a topografické mapy). Byla použita jako základ pro vytvoření datových sad pro trénování a validaci klasifikačních algoritmů Random Forest a Gradient Tree Boosting pomocí různých přístupů a nastavení, ale také pro srovnání s existujícími databázemi land cover/land use. Podrobný postup práce je popsán v následujících kapitolách.

4.1 Definice klasifikační nomenklatury a tvorba referenčních vrstev LULUCF zájmových k. ú.

Pro zvolená k. ú. byla získána data z KN a z veřejného registru půdy (LPIS). Kombinací těchto dat vznikly počáteční vrstvy dle klasifikace LULUCF. Tyto vrstvy byly následně porovnány s dalšími zdroji dat a ověřeny přímo v terénu. Podle získaných informací byly upraveny tak, aby odpovídaly skutečnému stavu land use v době terénního průzkumu.

Zpracování dat probíhalo v ArcGIS Pro, pokud není uvedeno jinak. Práce v terénu probíhala zejména v mobilní aplikaci Qfield, která umožňuje zobrazit libovolnou vizualizaci vrstev uloženou v projektu QGIS na počítači a také umožňuje editaci dat v terénu.

V prvním kroku musely být definovány základní LULUCF mapovací parametry, tzn. legenda a minimální mapovací jednotka (MMU) a minimální mapovací šířka (MMW). Ty byly stanoveny s ohledem na následné využití vrstev při klasifikaci území pomocí dat DPZ. MMU byla zvolena 100 m², což odpovídá ploše jednoho pixelu družicového snímku Sentinel-2. Pro lesy byla MMU stanovena na 0,05 ha (což odpovídá regulativu LULUCF a přibližně odpovídá čtverci o straně 22,4 m). MMW byla stanovena 5 m. Definice tříd LULUCF pro tvorbu referenčních vrstev jsou přehledně uvedeny v Tabulce 5. Definice tříd reflektují jak obecné metodické vymezení tříd pro účely reportingu LULUCF, tak i určitou specifičnost krajiny v Česku (Svoboda a kol., 2022).

Tabulka 5: Definice tříd pro tvorbu referenčních dat

Forest land
Lesní pozemky – plochy pokryté dřevinnou vegetací, s rozlohou alespoň 0,05 ha (což přibližně odpovídá čtverci o straně 22,4 m). Stromy v dospělosti dosáhnou minimálně 2 m do výšky. Do této třídy se počítají i menší (užší) nezalesněné plochy, které jsou však při pohledu shora překryty korunami stromů, typicky lesní cesty či cesty podél lesa. Jsou zde zahrnuty lesní oplocenky i v současné době vykácené plochy uvnitř lesa, u kterých se předpokládá, že zde les bude obnoven. Zahrnuty jsou i opuštěné, zarůstající ovocné sady s hustým zápojem dřevin.
Cropland
Plochy orné půdy (v současném roce obdělávané či ležící ladem), chmelnice, vinice i ovocné sady.
Grassland
Travní porosty, tedy plochy s převahou travin a jiných nízkých bylin. Patří sem louky čistě přírodního charakteru, ať už květnaté louky s vysokou biodiverzitou, nebo naopak plochy zarostlé, nesekané, s převahou několika málo dominantních druhů. Patří sem také travní porosty využívané k produkci sena, tedy pravidelně sečené, a jsou zde zahrnuty i pastviny pro zvěř, které jsou v průběhu roku (ne)pravidelně rotovány.
Settlements
Veškeré zastavěné plochy (budovy, dopravní komunikace, sportoviště), a dále prvky v krajině významně přetvořené člověkem. Jde například o lomy, skládky, staveniště, brownfieldy, vojenské prostory. Dále veškeré plochy zahrad kolem domů, a to i když se na nich nachází např. menší lesík, sad, zatravněná plocha či výběh pro domácí zvěř. Je zde zahrnuta i technická zeleň (např. aleje podél komunikací). Jsou zde zahrnuty i udržované městské plochy zeleně, typicky s krátko sečeným trávníkem.
Wetlands
Vodní plochy – tedy vodní toky, rybníky, jezera, vodní nádrže. Jsou zde zahrnuty také podmáčené plochy, typicky mokřady, rašeliniště. Patří sem i různé náplavy oblázků na okrajích vodních toků či nádrží, které jsou při vyšších vodních stavech také pod hladinou vody.

Other land
Zalesněné plochy menší než požadovaných rozměrů pro Forest land, a dále plochy holé půdy a skal.

V následujícím textu je popsán postup přípravy terénního šetření a výsledné referenční vrstvy.

Tvorba vrstev LULUCF pro zájmová k. ú.

Pro zájmová území byly z Geoportálu ČÚZK staženy vrstvy katastru nemovitostí: PARCELY_KN_P.shp (polygony) – vymezení hranic jednotlivých parcel
PARCELY_KN_DEF.shp (body) – definiční body parcel a atributy s kódem druhu pozemku (DRUPOZ_KOD) a způsobu využití pozemku (ZPVYPA_KOD).

Informace o druhu a způsobu využití pozemku byly z bodové vrstvy převedeny do polygonové vrstvy parcel pomocí funkce *Join*. Následně bylo třeba zařadit parcely dle druhu a využití pozemku do tříd LULUCF. Pro tento převod byla využita tabulka dle Langera (2022).

V dalším kroku se vycházelo z předpokladu, že vrstvy LPIS (MZe 2004) obsahují aktuálnější informace stávajícím využití krajiny než vrstvy KN. Proto byly staženy vrstvy půdních bloků LPIS pro zájmová i s nimi sousedící k. ú. (z důvodu, půdní bloky často přesahují hranici jednoho k. ú., avšak evidovány bývají ve vrstvě pouze jednoho z nich). Vrstvy půdních bloků z jednotlivých zájmových a sousedních k. ú. byly spojeny do jedné vrstvy pomocí funkce *Merge* a následně převedeny do tříd LULUCF dle výše citované převodní tabulky.

Výsledné vrstvy LULUCF pro zájmová katastrální území byla vytvořena aktualizací vrstev KN pomocí vrstev LPIS s použitím funkce *Update*. Vrstvy byly následně porovnány s dostupnými zdroji o krajině (ortofoto, ZM10, Google Earth Pro apod.) a dle těchto podkladů upraveny/zpřesněny. Úpravy zahrnovaly jak změny atributů, tak i dělení některých polygonů na menší celky.

Příprava do terénu, práce v terénu a finální úpravy

Připravené vrstvy byly importovány do programu QGIS, kde byla nastavena symbologie vhodná pro terénní mapování. Dále byly nahrány další podkladové vrstvy, například ortofoto nebo Základní topografické mapy (ZM). Výsledný projekt v QGIS byl vytvořen samostatně pro každé katastrální území a poté spolu s daty nahrán do tabletu s nainstalovanou aplikací QField. Terénní mapování probíhalo od července do září 2023, a to pěšky s přejezdy autem nebo na kole. Vrstvy byly přímo v aplikaci QField aktualizovány podle skutečného stavu zjištěného v létě 2023.

Po návratu z terénu byla data finalizována v softwaru ArcGIS Pro, kde byly zapracovány poznámky získané během mapování a provedena kontrola zařazení tříd a topologie vrstev. Po kontrole byla aplikována funkce *Dissolve*. Touto funkcí byly sousedící polygony se stejnou třídou LULUCF sloučeny do jednoho polygonu. Se všemi vrstvami v průběhu celého procesu bylo pracováno v souřadnicovém systému WGS84 UTM33 (EPSG: 32633).

4.2 Porovnání vybraných datových zdrojů na řešených k. ú.

V této části práce jsou navzájem porovnány vybrané zdroje land cover/land use dat na územích katastrů řešených v této práci. Porovnání bylo provedeno mezi vytvořenou referenční vrstvou LULUCF, stavem evidence KN v té podobě, v jaké byl v době mapování využíván pro LULUCF reporting, a vybranými globálními datasety

Zařazení porovnání globálních datasetů (CLC+, ESA WorldCover, Dynamic World a ESRI Land Cover) mělo za cíl poskytnout širší kontext možností využívání volně dostupných dat DPZ pro účely LULUCF reportingu. Tyto datasety představují aktuálně nejednodušší globální produkty land cover/land use, které by v praxi mohly sloužit jako alternativní nebo doplňkový zdroj informací o krajině. Porovnání těchto datasetů s referenčními daty umožnilo rámcově posoudit jejich vhodnost pro využití v LULUCF reportingu. Cílem bylo poukázat na to, že ačkoliv tyto produkty mohou být hodnotným zdrojem informací, při jejich použití je třeba počítat se specifickými chybami v klasifikaci, jako je například přeceňování určitých typů pokryvu, či s odlišným pojetím kategorizace krajiny, které není plně v souladu s požadavky LULUCF.

Vzhledem rozdílným klasifikačním nomenklaturám jednotlivých datasetů bylo zapotřebí jejich legendy sjednotit pro účely porovnání. Jako referenční legenda byla využita legenda LULUCF, na kterou byly nomenklatury ostatních zdrojů na základě podrobného studia definic jejich jednotlivých tříd převedeny. Vrstva KN byla převedena dle Tabulky 3, tedy oficiálního převodního systému dat KN pro účely LULUCF reportingu. Tabulka 6 uvádí definovaný převodní systém pro ostatní datasety. Do převodu nebyla zahrnuta třída Snow and Ice (trvalý sníh a led), která se v českých podmínkách nevyskytuje. Dle celoevropských pravidel LULUCF spadají tyto oblasti do třídy Other land.

V prvním kroku byly staženy vrstvy databází ESA World Cover, Dynamic World, ESRI Land Cover, CLC+ a vrstva KN. S výjimkou KN jde o rastrové vrstvy, za účelem srovnávání rozloh byly v SW ArcGIS Pro převedeny do vektorového formátu a byla jim nastaveno zobrazení WGS84 UTM33.

Tabulka 6: Převodní tabulka pro porovnání land cover/land use databází

LULUCF	ESA	Dynamic	ESRI	CLC+
Forest land	Tree cover	Trees	Trees	Woody – needle leaved trees
				Woody – Broadleaved deciduous trees
				Woody – Broadleaved evergreen trees
Cropland	Cropland	Crops	Crops	Periodically herbaceous
Grassland	Grassland	Grass	Grass	Permanent herbaceous
Settlements	Built-up	Built Area	Built Area	Sealed
Wetlands	Open water	Water	Water	Water
	Herbaceous wetland	Flooded vegetation	Flooded vegetation	
Other land	Barren/sparse vegetation	Bare ground	Bare Ground	Non- and sparsely- vegetated
	Shrubland	Shrub and scrub	Shrub and scrub	Low-growing woody plants
	Moss and lichen			

Za využití funkce *Summary statistics* byly v každém k. ú. spočteny rozlohy krajinných tříd dle jednotlivých datových zdrojů.

4.3 Použitá data a tvorba kompozitů pro klasifikace

V další části práce bylo přistoupeno k vlastní klasifikaci dat DPZ do kategorií LULUCF. Pro klasifikaci byly využívány multispektrální snímky Sentinel-2, kolekce Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A, dostupná v prostředí GEE.

Nejprve byl pro každé zájmové k. ú. vytvořen kompozit snímků. Ke tvorbě kompozitů byly využívány snímky z časového období duben až září 2023 a též z užšího časového intervalu květen až červenec téhož roku. Kompozit byl vytvořen jako mediánový i průměrový.

Pro odstranění oblačnosti byl použit algoritmus CloudScore+. Byl využit dataset Cloud Score+ S2_HARMONIZED odvozený z harmonizované kolekce Sentinel-2 L1C s již aplikovanou atmosférickou korekcí (Sen2Cor). Výstupy CloudScore+ lze použít k identifikaci relativně čistých pixelů a efektivnímu odstranění oblačnosti a jejích stínů z obrazů L1C (Top-of-Atmosphere) nebo L2A (Surface Reflectance).

Dataset Cloud Score+ S2_HARMONIZED obsahuje dvě pásma, cs a cs_cdf, která hodnotí kvalitu pixelů z hlediska viditelnosti povrchu na škále od 0 do 1, kde 0 značí zakrytý a 1 zcela čistý pixel.

cs – hodnotí kvalitu pixelu na základě okamžité spektrální podobnosti s teoretickým čistým referenčním stavem.

cs_cdf – představuje pravděpodobnost, že je pixel čistý, na základě historického rozdělení hodnot pro danou lokalitu.

Hlavním důvodem pro výběr tohoto algoritmu byla jeho snadná implementace v prostředí GEE. Pro účely klasifikace bylo nastaveno kritérium $cs \geq 0,8$, tedy pokrytí oblačností nesmělo přesáhnout 20 %.

Na všech snímcích byly dále spočítány indexy NDVI (Normalized Difference Vegetation Index) (Tucker 1979), NDBI (Normalized Difference Built-up Index) (Zha a kol. 2003) a NDWI (Normalized Difference Water Index) (Gao 1996) použitím následujících vztahů:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}$$
$$NDBI = \frac{\rho_{SWIR2} - \rho_{NIR}}{\rho_{SWIR2} + \rho_{NIR}}$$
$$NDWI = \frac{\rho_{NIR} - \rho_{SWIR1}}{\rho_{NIR} + \rho_{SWIR1}}$$

kde ρ_{RED} , ρ_{NIR} , ρ_{SWIR1} a ρ_{SWIR2} odpovídají pásmům 4, 8, 11 a 12 snímku Sentinel-2. Přidání těchto indexů bylo inspirováno studií Sun a Ongsomwang (2023) Indexy byly do klasifikace zařazeny pro svou schopnost zvýraznit zmíněné typy povrchů a tím pomoci jednotlivé třídy lépe rozlišit. Do klasifikace také vstupoval rozptyl hodnot NDVI a NDBI v období od dubna do září 2023. Jejich začlenění mělo za cíl zachytit sezónní proměnlivost sledovaných ploch během vegetačního období. Vysokou užitečnost rozptylu NDVI pro klasifikaci dokládají např. Svoboda a kol. (2022) či Phalke a kol. (2020).

Na základě výsledků terénního výzkumu byl upraven metodický postup klasifikace a modifikována klasifikační nomenklatura pro účely klasifikace multispektrálních družicových dat. Původně samostatné třídy Forest land a Other land byly pro účely klasifikace sloučeny do jedné společné třídy nazvané Woodland (dřevinný porost). V mapovaných územích do třídy Other land byly v rámci terénního mapování zařazovány de facto pouze dřevinné porosty nesplňující prahové hodnoty třídy Forest land (tj. rozloha menší než 0,05 ha, výška stromů v dospělosti nedosahuje 2 m), avšak se shodnými spektrálními charakteristikami. Vzhledem k tomu, že dále budou v práci využívány metody pixelové klasifikace, došlo se k závěru, že nemá opodstatnění oddělovat třídy, které se liší pouze rozměrovým kritériem, ale nikoli spektrálními vlastnostmi. Takové rozlišování by naopak vedlo ke zhoršení výsledků

klasifikace. Opětovné rozlišení třídy Woodland na Forest land a Other land je provedeno až v klasifikovaném obrazu na základě rozměrových charakteristik. Tento postup byl inspirován metodickým přístupem dle Svobody a kol. (2022).

4.4 Tvorba trénovacích a validačních dat

V této diplomové práci byl při tvorbě trénovacích a validačních datasetů využit potenciál, že v rámci této práce bylo detailně zmapováno celé zájmové území. Rozvíjený přístup je snahou o vytvoření robustního klasifikačního modelu, který bude zachycovat co nejširší spektrum typů využití krajiny vyskytujících se v Česku. K vytvoření trénovacího a validačního datasetu bylo využito pět ze šesti zpracovávaných katastrálních území, zatímco jedno území bylo ponecháno jako zcela nezávislé. Data z tohoto k. ú. nebyla využita pro trénování modelu. Z výsledné přesnosti klasifikace nezávislého území je možné do jisté míry usuzovat, jak je model schopen klasifikovat oblasti, které nejsou součástí trénovacích dat a testuje se tak přenositelnost modelu. Klasifikován je i rozsáhlejší územní celek jako ukázka a podklad pro diskusi dalšího využití tohoto modelu v následujícím výzkumu LULUCF s využitím dat DPZ.

Příprava trénovacích a validačních dat

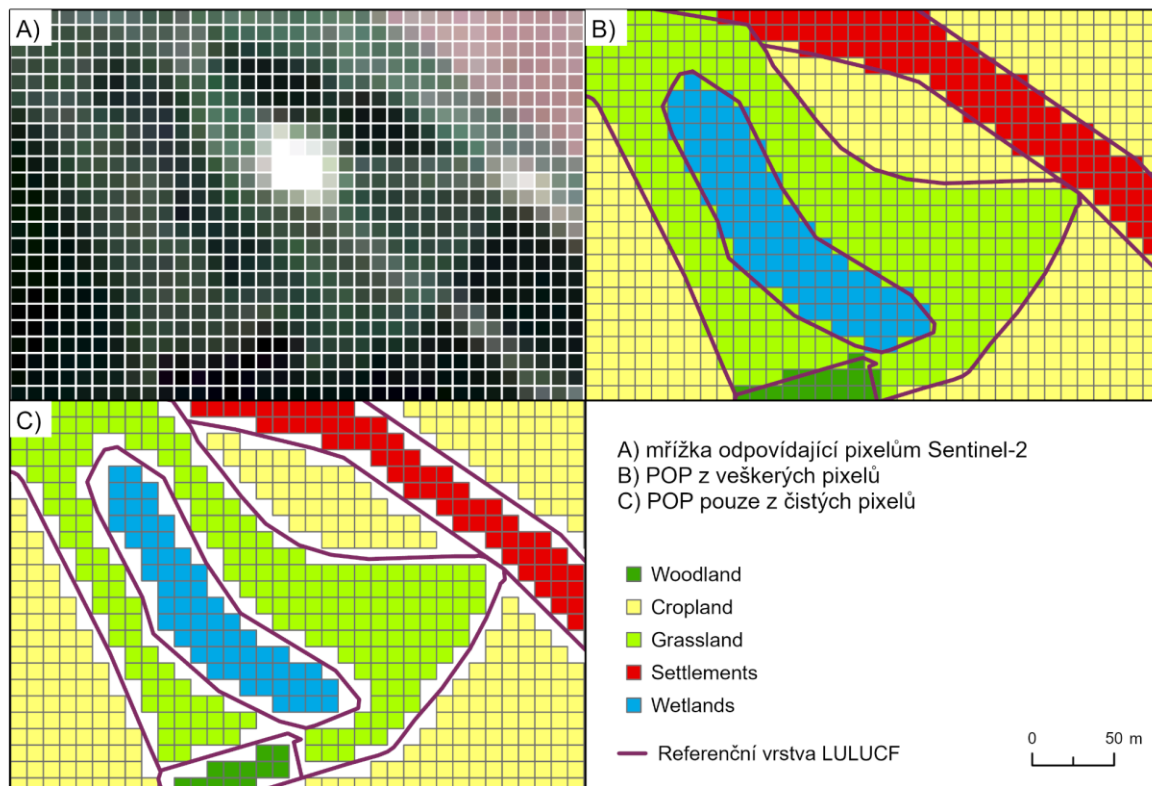
Tvorba datasetu vhodného pro trénování a validaci klasifikačních algoritmů z referenční vrstvy probíhala na cloudové platformě Google Earth Engine (GEE) pomocí programovacího jazyka JavaScript v prostředí JavaScript Code Editor, a částečně i v ArcGIS Pro.

Ze snímku Sentinel-2 (S-2) byla pro každé katastrální území vytvořena pravidelná vektorová mřížka složená z polygonů (čtverců) o rozměrech 10×10 metrů, které odpovídají umístění a velikosti jednotlivých pixelů tohoto snímku. Referenční vektorová vrstva s třídami LULUCF byla nejprve převedena na rastrový formát s rozlišením 2×2 metry. Každému pixelu v tomto rastru byla přiřazena třída LULUCF podle toho, která se v daném místě nacházela v referenční vrstvě. Následně byly tyto dva produkty překryty, tedy každému čtverci vektorové mřížky (10×10 m) byla přiřazena ta třída LULUCF, která se nejčastěji vyskytovala v jemnější mřížce (2×2 m) v rámci jeho plochy. Jinými slovy, modus tříd z 25 buněk 2×2 m. Bylo též ověřeno, že si hranice pixelů obou rastrů vzájemně odpovídají (nejsou vůči sobě posunuty).

K použití jemnějšího rastru před přiřazením hodnot do desetimetrové mřížky bylo přistoupeno z důvodu přesnějšího zachycení prostorového rozložení tříd v oblastech na rozhraní více krajinných tříd. Pokud by byla referenční vektorová vrstva přímo rasterizována na desetimetrovou mřížku, v hraničních oblastech by nebylo možné jednoznačně určit, která třída má v pixelu většinu. Zvolený přístup, tedy jemnější mřížka a následný downsampling na základě nejčastější hodnoty, výrazně zvyšuje pravděpodobnost,

že pixelu na hranici více krajinných tříd bude skutečně přiřazena třída, která tento pixel pokrývá z největší části. Se všemi vrstvami bylo ve všech operacích pracováno v souřadnicovém systému WGS84 UTM33. Výsledkem tohoto kroku byla vektorová mřížka polygonů, z nichž každý odpovídal jednomu pixelu snímku Sentinel-2 (dále označovaná jako POP – polygony odpovídající pixelům) a nesl informaci o třídě LULUCF, která se v tomto pixelu z většiny nachází.

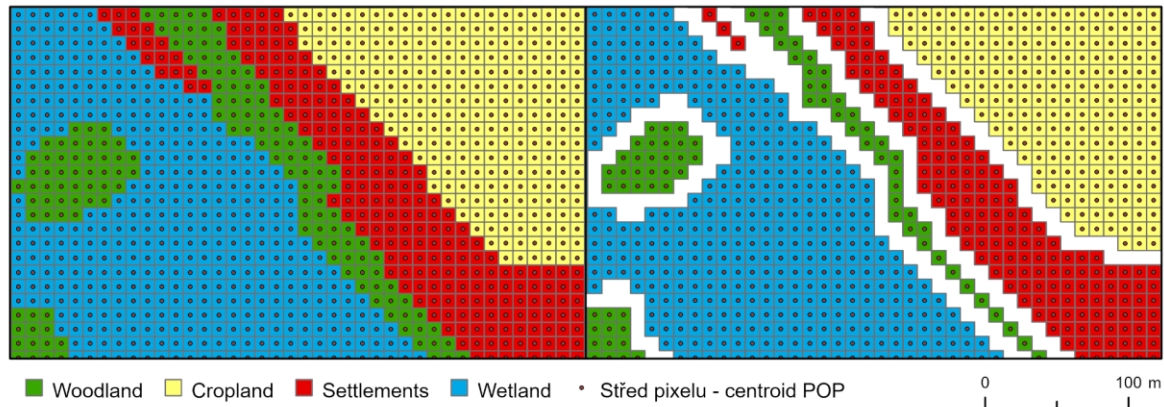
Vrstva POP v tomto provedení zahrnovala veškeré pixely v řešených k. ú., tedy i ty, které se nacházely na hranici dvou nebo více tříd LULUCF (jimž byla jedna ze tříd přiřazena výše vysvětleným postupem). Pro vytvoření vrstvy obsahující pouze čisté pixely, tedy ty, které spadají jednoznačně pouze do jedné krajinné třídy, byla využita funkce ArcGIS Pro *Select by Location* na vrstvu POP s kritériem *Completely within* vůči referenční vrstvě polygonů LULUCF. Tímto způsobem byly vybrány pouze ty pixely, které se celé nacházely uvnitř jedné třídy v referenční vrstvě LULUCF. Pro zjednodušení bude v dalším textu tato vrstva označována jako “čisté pixely”. Postup je dokumentován na Obrázku 3. Toto rozdělení později sloužilo k testování různých možností, jak co nejlépe natrénovat klasifikační algoritmus.



Obrázek 3: Postup přípravy trénovacích dat v ArcGIS Pro

Další krok tvorby trénovacího a validačního datasetu probíhal opět v prostředí GEE. Spočíval ve vytvoření bodové vrstvy centroidů POP, tedy středů jednotlivých čtverců.

Výsledkem tohoto kroku byla rovnoměrná mřížka bodů, přičemž každý bod odpovídal středu pixelu družicového snímku Sentinel-2. Každý centroid převzal z vrstvy POP informaci o referenční třídě, která se v daném čtverci nachází. Toto bylo provedeno zvlášť pro vrstvu POP obsahující veškeré pixely v území, tak pro vrstvu POP obsahující pouze čisté pixely, jak znázorňuje Obrázek 4.



Obrázek 4: Centroidy POP

V následujícím kroku byly bodovým vrstvám přiřazeny hodnoty odrazivosti ve všech pásmech S-2 a hodnoty indexů a jejich rozptylů. Všechna katastrální území se nacházejí v převážně rovinném terénu, a proto je pravděpodobnost výskytu chyb v odrazivosti způsobených sklonem svahu (tzv. slope-induced reflectance errors) velmi nízká.

K vytvoření trénovacích a validačních datasetů bylo využito pět ze šesti zpracovaných katastrálních území, jedno území bylo ponecháno jako nezávislé. Bodová vrstva pěti k. ú. byla rozdělena na trénovací a validační část v poměru 50 : 50.

K rozdělení byl využit následující způsob: Veškeré body byly rozděleny do skupin dle jednotlivých tříd LULUCF. Následně byla každému bodu v rámci skupiny přiřazena náhodná hodnota mezi 0 a 1, a provedeno rozdělení do trénovací a validační skupiny tak, že body do hodnoty 0,5 včetně byly zahrnuty do trénovacího datasetu, a s vyšší hodnotou než 0,5 do validačního datasetu. Toto rozdělení zajišťovalo, že v rámci trénovacího i validačního datasetu bude zachován poměr výskytu tříd v krajině, tedy že třídy v krajině hojnější budou v trénovacím i validačním datasetu reprezentovány více body. Zároveň zde nebylo možné přesně dodržet poměr 50 : 50, ale vzhledem k celkovým počtům pixelů tato mírná odchylka nepředstavuje problém. Následující tabulka uvádí počty trénovacích a validačních bodů, se kterými bylo pracováno. Počty jsou v řádech statisíců a tato data lze označit za tzv. big data.

Tabulka 7: Počty trénovacích + validačních bodů v jednotlivých k. ú.

Body	Veškeré	Pouze čisté
trénovací	151372	132281
validační	150088	132423

Výše popsaným způsobem bylo vytvořeno několik variant trénovacího a validačního datasetu, na základě způsobu tvorby kompozitu, časového rozsahu snímků i čistotě pixelů. Jejich přehled uvádí následující tabulka.

Tabulka 8: Varianty trénovacího + validačního datasetu

Kompozit	Časový rozsah snímků	Čistota pixelů
průměrový	duben až září	veškeré
		jen čisté
	květen až červenec	veškeré
		jen čisté
mediánový	duben až září	veškeré
		jen čisté
	květen až červenec	veškeré
		jen čisté

4.5 Použité klasifikátory

Klasifikace v této diplomové práci byly prováděny pomocí klasifikačních algoritmů Random Forest (RF) a Gradient Tree Boosting (GTB). Využití RF se ukázalo jako přínosné v mnoha studiích zabývajících se obdobným typem klasifikace, např. Svoboda a kol. (2022), Sun a Ongsomwang (2023), Yang a kol. (2021), Xiao a kol. (2021) či Rodriguez-Galiano a kol. (2012).

Random Forest

Metoda využívá princip modelu rozhodovacího stromu. Rozhodovací strom je hierarchická struktura, která simuluje proces rozhodování na základě jednoduchých,

automaticky vytvořených pravidel (kritérií). Dle daných kritérií rozřazuje data do menších podmnožin a postupně do jednotlivých tříd. Kroky v rámci rozhodovacího stromu se nazývají uzly, každý uzel představuje rozhodovací kritérium. Počáteční uzel se všemi vstupními daty je označován jako kořenový uzel (root node), dále jsou v rámci stromu vnitřní uzly (decision node) a koncové uzly stromu, které obsahují finální klasifikační třídu (v případě použití pro klasifikaci), jsou označeny jako listy (leaf node). Samostatný rozhodovací strom je však velmi náchylný k přeučení (Ho 1995). RF patří mezi tzv. ensemble learning metody, což znamená, že kombinuje více modelů za účelem dosažení vyšší přesnosti. Každý strom využívá náhodně vybranou podmnožinu vzorku dat, na které je trénován. Této metodě se říká bagging (bootstrap aggregating) a zajišťuje co nejnižší korelaci mezi jednotlivými stromy. Každý strom zařadí pixel do vybrané klasifikační třídy samostatně, následně je pixelu přiřazena ta třída, kam jej stromy nejčastěji zařadily (Breimann 2001).

Klasifikátoru RF je možno nastavit následující parametry (na tomto místě i v následujícím textu je pracováno s označením parametrů shodně s jejich označením v prostředí GEE, které je využíváno v této práci):

numberOfTrees (NT) – počet rozhodovacích stromů. Vyšší počet stromů obvykle zlepšuje výkon modelu, ale zvyšuje výpočetní náročnost.

variablesPerSplit (VPS) – počet vstupních proměnných, které budou náhodně vybrány pro hledání nejlepšího rozdělení v každém uzlu stromu. V GEE je defaultní hodnota tohoto parametru nastavena na odmocninu počtu vstupních pásem klasifikovaného obrazu.

bagFraction (BF) – procento trénovacích dat použitých pro vytvoření každého stromu. Výchozí hodnota v GEE je 0,5, tedy 50 %.

minLeafPopulation (minL) – minimální počet vzorků, které musí být v listu (konečném uzlu).

maxNodes (maxN) – maximální počet uzlů v každém stromu.

seed – nastavení náhodnosti pro replikovatelnost výsledků.

Gradient Tree Boosting

Dalším perspektivním a v této studii testovaným klasifikátorem je Gradient Tree Boosting. Na rozdíl od baggingu, způsob označovaný jako boosting trénuje modely (stromy) postupně, jeden po druhém. Každý model se učí z chyb modelu předchozího tím, že minimalizuje rezidua předchozích predikcí. Vstupy zůstávají zachovány, ale správně klasifikovaným vzorkům je přisuzována menší váha při výpočtu gradientu ztrátové funkce. Každý nový strom se snaží co nejlépe aproximovat gradient ztrátové funkce a opravit chyby předchozích stromů. GTB je pomalejší na natrénování, ale vyznačuje se vysokou přesností.

Klasifikátoru GTB je možno nastavit následující parametry (označení opět v souladu s GEE):

numberOfTrees – počet rozhodovacích stromů. Vyšší počet stromů obvykle zlepšuje výkon modelu, ale zvyšuje výpočetní náročnost.

shrinkage – rychlost učení. Parametr náleží do intervalu (0,1). Výchozí hodnota je v GEE nastavena na 0,005.

samplingRate – procento trénovacích dat, která se náhodně vybírá pro každý strom. Výchozí hodnota je v GEE nastavena na 0,7. Nepoužije se tedy celá množina trénovacích dat, ale pouze 70 % z ní.

maxNodes – maximální počet uzlů v každém stromu.

loss – typ ztrátové funkce. K dispozici jsou leastSquares (metoda nejmenších čtverců), leastAbsoluteDeviation (pro datasey s odlehlými hodnotami), huber (kombinace dvou výše uvedených). Defaultně je nastavena leastAbsoluteDeviation.

seed – nastavení náhodnosti pro replikovatelnost výsledků.

4.6 Klasifikace

V této části práce byly vytvářeny klasifikační modely na základě různého způsobu jejich natrénování. Ke tvorbě klasifikačních modelů bylo využito pět referenčních k. ú. (České Vrbné, Chvaletice, Klec, Hojovice a Prachovice). Bylo testováno jak nastavení vstupních parametrů klasifikátorů, tak různé možnosti natrénování modelu.

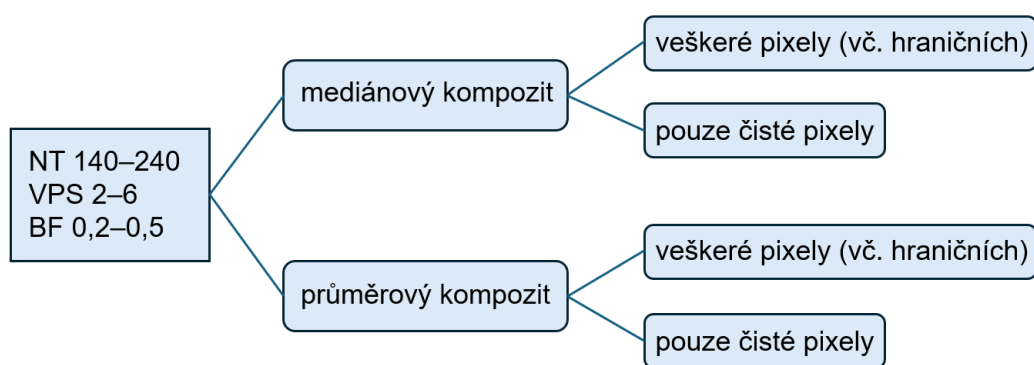
Klasifikace Random Forest

Optimalizace nastavení a trénování klasifikátoru Random Forest probíhala v několika krocích. V každé fázi byly vyhodnoceny dosažené výsledky a v další bylo pracováno pouze s nastaveními, které dosáhly nejvyšších přesností. Tímto způsobem byly tedy postupně vyřazovány méně úspěšné varianty.

V první fázi bylo testováno, zda bude klasifikace úspěšnější při použití mediánového či průměrového kompozitu, a použijí-li se pro trénování veškeré pixely, tedy včetně hraničních, či pouze pixely čisté. Klasifikace byla pro tyto čtyři možnosti spuštěna s nastavením parametrů klasifikátoru NT od 140 do 240 s krokem 20, VPS od 2 do 6 s krokem 1 a BF od 0,2 do 0,5 s krokem 0,1. Za účelem snížení výpočetní náročnosti modelu a zamezení jeho přeučení byly také omezeny parametry minL = 10, maxN = 10 000. Pro každou variantu bylo tedy testováno 120 kombinací nastavení parametrů.

Z klasifikace zde byla vyjmuta pásma snímku S-2 B1 a B9, a to z důvodu, že tato pásma mají rozlišení 60 m (ostatní pásma S-2 mají deseti či dvacetimetrové rozlišení). Jejich

odebrání bylo motivováno snahou zabránit případnému snížení prostorové přesnosti klasifikace, která by mohla vlivem vstupu pásem s nižším rozlišením nastat. Pásmo B1 je tradičně využíváno pro detekci aerosolů, pásmo B9 pro detekci vodní páry, nejde tedy o pásma vyloženě vypovídající o zemském povrchu. Do klasifikace byly zahrnuty indexy NDVI a jeho rozptyl, NDBI a NDWI. Časový rozsah snímků, ze kterých byl vytvářen kompozit, byl od května do července, ovšem rozptyl indexu NDVI byl počítán z období od dubna do září. Kratší časový rozsah pro snímky byl využit z důvodu snahy o eliminaci případných stínů, které jsou například v září již relativně dlouhé. Schéma variant v této fázi je znázorněno na Obrázku 5.



Obrázek 5: Varianty klasifikací v první fázi

Z výsledků vyplynulo, že nejvyšší přesnosti dávaly klasifikace s využitím mediánového kompozitu, trénované na veškerých pixelech. V další fázi bylo pracováno s pěti nejúspěšnějšími modely vzešlými z této varianty. Bylo testováno, jaký vliv mají některá vstupní pásma snímku S-2, a časový rozsah snímků, ze kterých je tvořen kompozit. Konkrétně byly testovány následující varianty vstupních pásem:

- všechna pásma dostupná ve snímku S-2
- vynechání šedesátimetrových pásem B1 (aerosoly) a B9 (vodní pára)
- vynechání pásem B1, B9, B7 (Vegetation Red Edge) a B8 (NIR). Hlavním důvodem pro vynechávání dalších pásem (mimo B1 a B9) byla snaha o snížení výpočetní náročnosti modelu. Vynechání právě pásem B7 a B8 bylo inspirováno studií Svoboda a kol. (2022), dle které mají tato pásma pro klasifikaci LULUCF nejmenší význam.

Testovány byly také následující nastavení časového rozsahu snímků, ze kterých jsou vytvářeny kompozity:

- snímky pro tvorbu kompozitu od května do července, rozptyl NDVI vypočten z období od dubna do září
- snímky pro tvorbu kompozitu i výpočet rozptylu NDVI od dubna do září

U všech pěti testovaných modelů vyšla nejvyšší přesnost při kombinaci využití všech pásem dostupných ve snímku S-2 a použití snímků z období od dubna do září pro tvorbu kompozitu i výpočet rozptylu NDVI.

Vzhledem k tomu, že jde o jiné nastavení, než jaké bylo využito pro klasifikace všech kombinací parametrů RF v první fázi, byly v tomto kroku opět spuštěny klasifikace pro všechny zkoumané kombinace parametrů RF, tentokrát s využitím poznatků z první i druhé fáze – tedy mediánový kompozit, trénování na veškerých pixelech, zahrnutí všech pásem snímku S-2 a použití snímků z období od dubna do září pro tvorbu kompozitu i výpočet rozptylu NDVI. Zároveň při studiu klasifikačních map a chybových matic vyšlo najevo, že jsou navzájem poměrně často zaměňovány třídy Cropland a Settlements. Do klasifikačního modelu byla tedy v tuto chvíli přidána další proměnná, a sice rozptyl indexu NDBI. Třída Cropland by měla teoreticky mít ze všech tříd rozptyl hodnot odrazivosti nejvyšší (což reflektuje již zahrnutý rozptyl NDVI), naopak třída Settlements by měla mít tento rozptyl velmi nízký (což by se mělo odrazit v proměnné rozptyl NDBI). Kombinace těchto proměnných by pak mohla pomoci tyto třídy lépe rozlišovat.

V rámci postprocessingu klasifikovaného rastru byla provedena filtrace za účelem odstranění izolovaných pixelů, tedy těch, v jejichž bezprostředním okolí 3×3 neleží žádný pixel se stejnou třídou. Třídy těchto pixelů byly nahrazeny nejčastější třídou v jejich okolí. Tím došlo, alespoň částečně, k eliminaci nežádoucího efektu “pepř a sůl”.

Nejúspěšnější model byl následně využit ke klasifikaci celého Česka.

Klasifikace Gradient Tree Boosting

V případě klasifikátoru Gradient Tree Boosting byly veškeré klasifikace prováděny s použitím mediánového kompozitu vytvořeném ze snímků od dubna do září 2023. Omezení počtu stromů (parametr maxNodes) bylo nastaveno na 10 000. V klasifikacích byl testován vliv čistoty trénovacích dat, vstupních pásem (vynechání pásem B1 a B9) a parametru samplingRate, tedy procenta trénovacích dat využívané každým stromem. Klasifikace proběhly pro modely s 8, 10, 20 a 25 stromy. Při využití 20 stromů požadovaný výpočetní výkon již narazil na limit platformy GEE a některé varianty nebylo možné provést.

Hodnocení přesnosti

Pro hodnocení přesnosti modelů byla připravena validační sada oddělená od trénovacích dat v poměru 50 : 50, přičemž trénovací i validační vzorky pocházely z pěti katastrálních území. Validace měla roli v průběžném ladění parametrů a ověřování robustnosti modelu, například včasným zachycením možného přeučení. Pro finální hodnocení

přesnosti však nebyla využita. Přesnost klasifikace byla vyhodnocována na šestém katastrálním území, které nebylo nijak zahrnuto do procesu trénování. Tento přístup umožnil objektivně posoudit přenositelnost modelu na nezávislé území s odlišnými charakteristikami, což odpovídá ambici této práce – vytvořit klasifikační model LULUCF přenositelný na širší územní celky.

Zvolený postup je odůvodněn následujícím:

- Vzhledem ke geografické blízkosti a značné spektrální podobnosti trénovacích a validačních vzorků (často sousední pixely) by validační sada neposkytla dostatečně objektivní obraz o schopnosti modelu generalizovat.
- Využití 50 % známé datové sady pro trénování modelu významným způsobem snížilo výpočetní náročnost procesu ve srovnání se situací, kdy by bylo využito 70 % či 100 % známých dat. Přesto počet trénovacích pixelů dosahoval řádů statisíců, což vzhledem k malé rozloze klasifikované oblasti poskytovalo dostatečně objemný soubor trénovacích dat. Testování navíc prokázalo, že využití 70 % vzorků by nepřineslo významné zvýšení přesnosti klasifikace.
- Byla zvažována také možnost využít čtyři katastrální území pro trénování a dvě pro validaci. Tento postup však nebyl uplatněn z důvodu snahy o zachycení co největší heterogenity krajinných typů v trénovacích datech.

Pro hodnocení přesnosti byla primárně využívána celková přesnost (overall accuracy), která sloužila k identifikaci nejúspěšnějších testovaných modelů. U těchto modelů byly následně sestaveny chybové matice (error matrix) pro podrobnější analýzu schopnosti modelu klasifikovat jednotlivé třídy LULUCF.

5 VÝSLEDKY

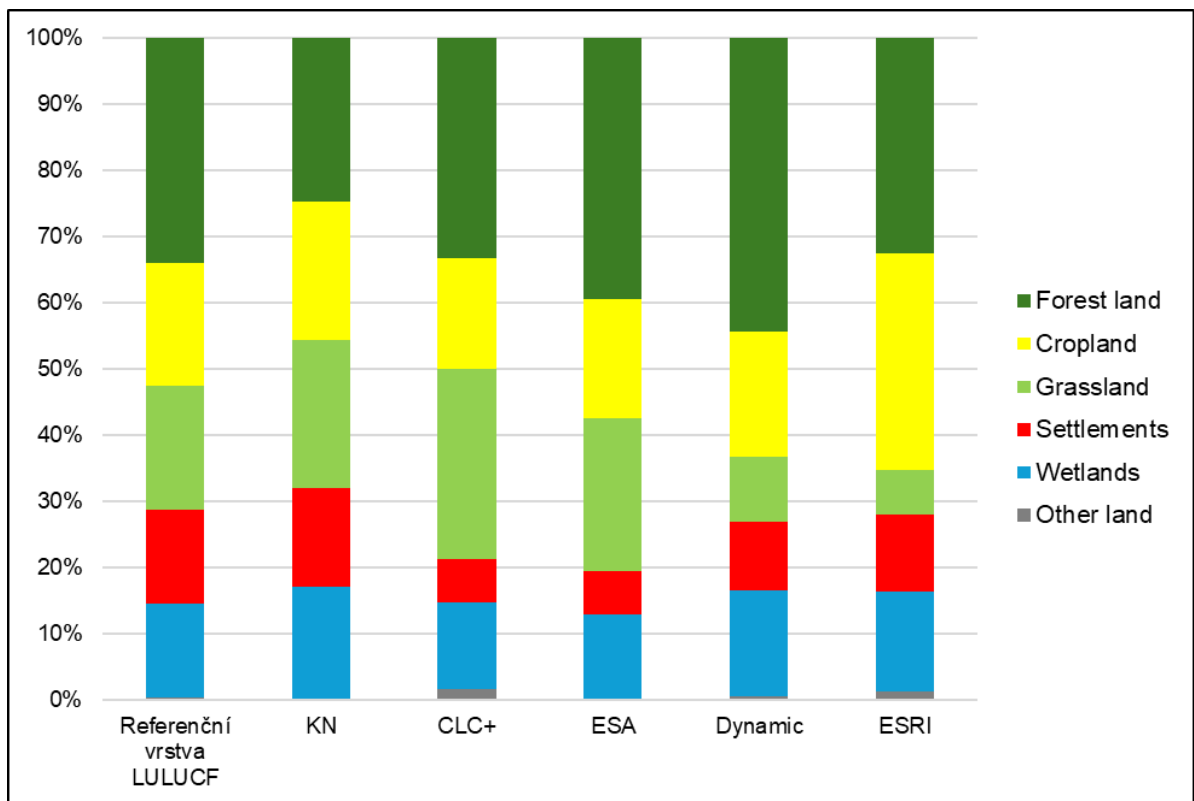
5.1 Porovnání vybraných datových zdrojů na řešených k. ú.

Tabulka 9 uvádí souhrnné údaje za 5 ze 6 řešených k. ú. V k. ú. Chvaletice nemohlo dojít k vyčíslení rozloh vrstvy Dynamic World z důvodu chyby (“díry”) ve stažených datech. Srovnávací tabulky pro jednotlivá k. ú. jsou k dispozici v Příloze 2.

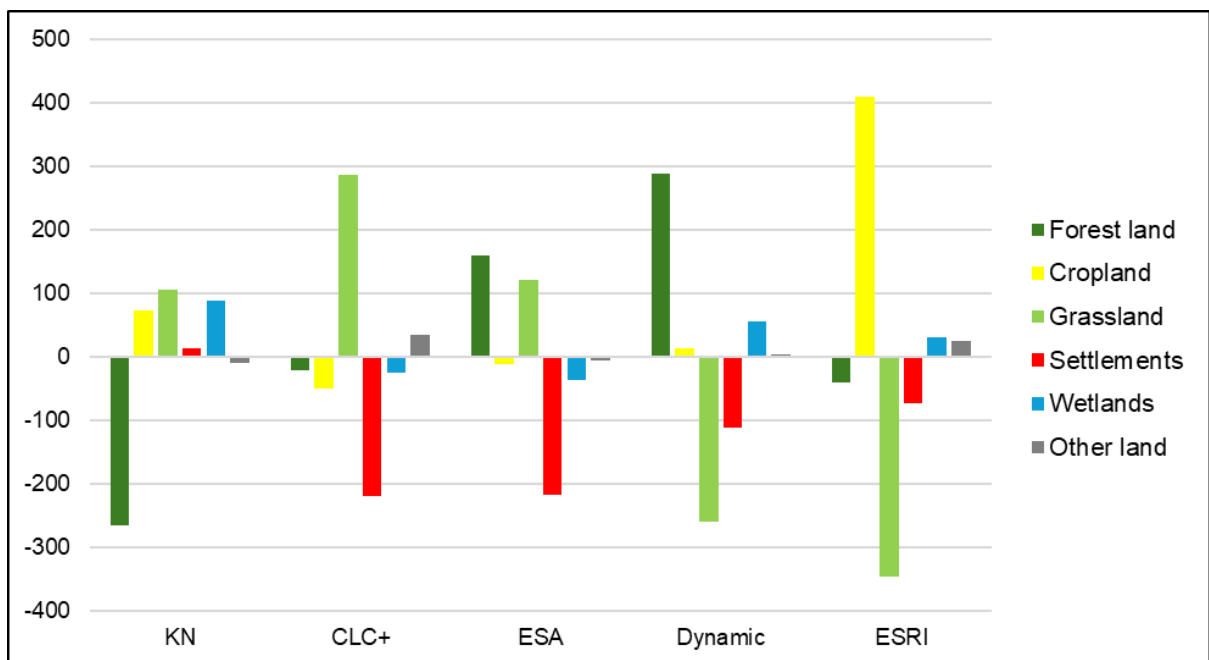
Tabulka 9: Rozlohy krajinných tříd dle různých datových zdrojů v 5 k. ú. (ha)

	Referenční vrstva LULUCF	KN	CLC+	ESA World Cover	Dynamic World	ESRI Land Cover
Forest land	978,33	711,83	956,09	1136,08	1265,03	937,90
Cropland	527,97	600,61	477,85	515,77	540,86	936,39
Grassland	538,32	642,92	824,83	659,02	277,23	191,73
Settlements	409,52	422,91	188,56	190,41	296,55	336,10
Wetlands	402,46	490,30	376,44	364,40	456,81	432,12
Other land	9,81	0,00	43,84	2,30	13,51	34,17

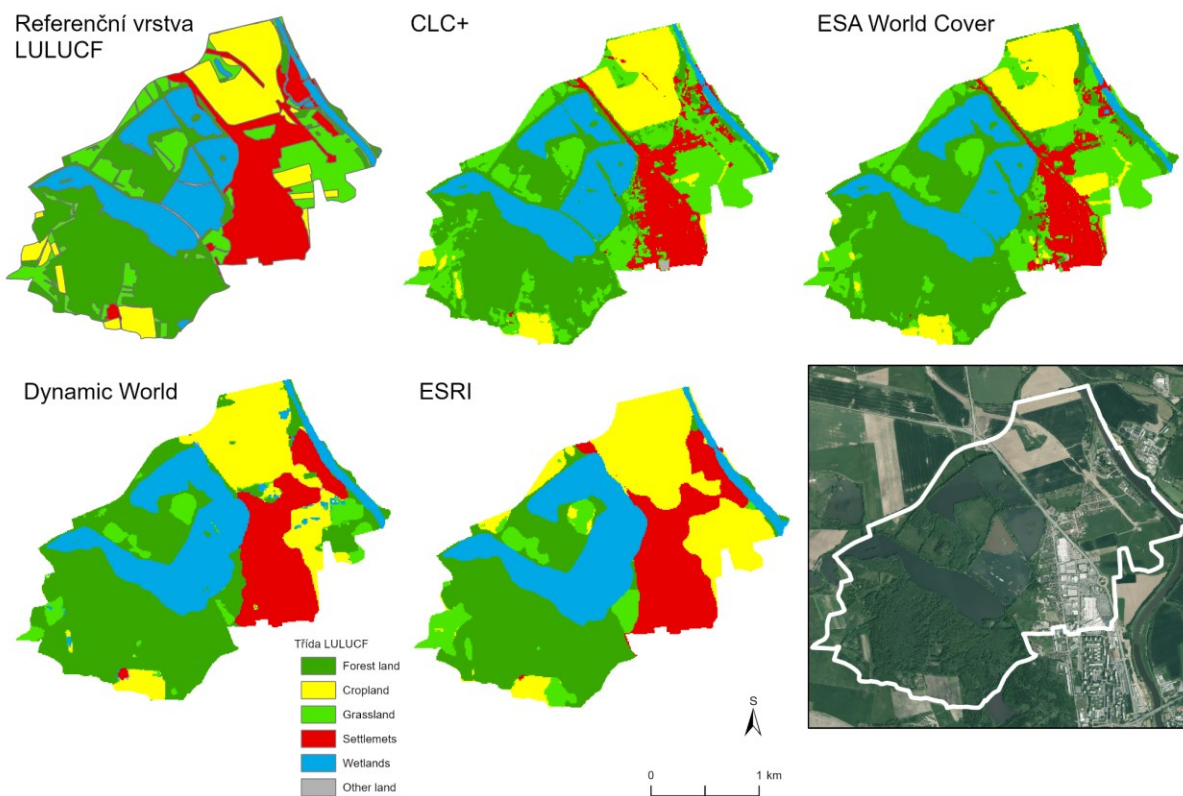
Obrázek 6 vyjadřuje podíl jednotlivých tříd v k. ú. dle jednotlivých zdrojů, Obrázek 7 pak srovnání jednotlivých databází s referenční vrstvou LULUCF vytvořenou v rámci této práce. Kladné hodnoty v Obrázku 7 znamenají, že daný zdroj kategorii oproti referenční vrstvě nadhodnocuje. Co se týče srovnání s KN, je zřejmé, že KN eviduje menší rozlohu lesů, než bylo zjištěno výzkumem aktuálního stavu v krajině. Dalším zajímavým jevem je nadhodnocení vodních ploch katastrem. Zde rozdíly vycházejí z manuálních úprav zejména v jihočeských k. ú. s velkým množstvím vodních útvarů. V rámci referenčních vrstev nebyly do Wetlands řazeny ostrovy a poměrně rozsáhlé oblasti přibřežní vegetace, které KN eviduje jako vodní plochy. Co se týče Settlements, vrstvy CLC+, ESA World Cover i Dynamic World zaznamenávají výrazně nižší rozlohy této třídy. Jedním z důvodů je fakt, že v rámci LULUCF se do této třídy zahrnují přírodní plochy ve městech, které vrstvy vycházející z dat DPZ spíše vyhodnocují jako traviny. To by odpovídalo i zjištění, že CLC+ i ESA mají oproti referenčním vrstvám traviny výrazně nadhodnocené. Dynamic World naopak třídu travin také podhodnocuje, ale naopak nejvýrazněji nadhodnocuje lesy. Z hodnocených zdrojů se všem ostatním vymyká ESRI Land Cover, kde jsou navíc rozdíly oproti referenční vrstvě LULUCF nejvyšší v absolutních rozlohách.



Obrázek 6: Srovnání zastoupení krajinných tříd dle různých datových zdrojů v 5 k. ú.



Obrázek 7: Srovnání datových zdrojů vůči referenční vrstvě LULUCF



Obrázek 8: Srovnání globálních datasetů v k. ú. České Vrbné

5.2 Výsledky klasifikace

Cílem této části práce je podat přehled výsledků jednotlivých kroků při hledání nejvhodnějšího modelu ke klasifikaci krajiny pro účely LULUCF, které jsou popsány v kapitole 4.6. Není-li uvedeno jinak, uváděné hodnoty přesnosti se vztahují k nezávislému katastrálnímu území Škrdlovice, které nebylo použito při trénování modelu.

5.2.1 Výsledky klasifikace Random Forest

V první fázi bylo hodnoceno, zda bude klasifikace úspěšnější při použití mediánového či průměrového kompozitu, a použijí-li se pro trénování veškeré pixely, tedy včetně hraničních, či pouze pixely čisté. Výsledky pěti nejlepších přesností každé varianty a jejich průměry jsou vypsány v Tabulce 10. Z hodnocení vyplývá, že nejlepších výsledků dosahovaly klasifikace při vyžití mediánového kompozitu a trénování modelu na veškerých pixelech. Tento model dosáhl přesnosti 80,25 %.

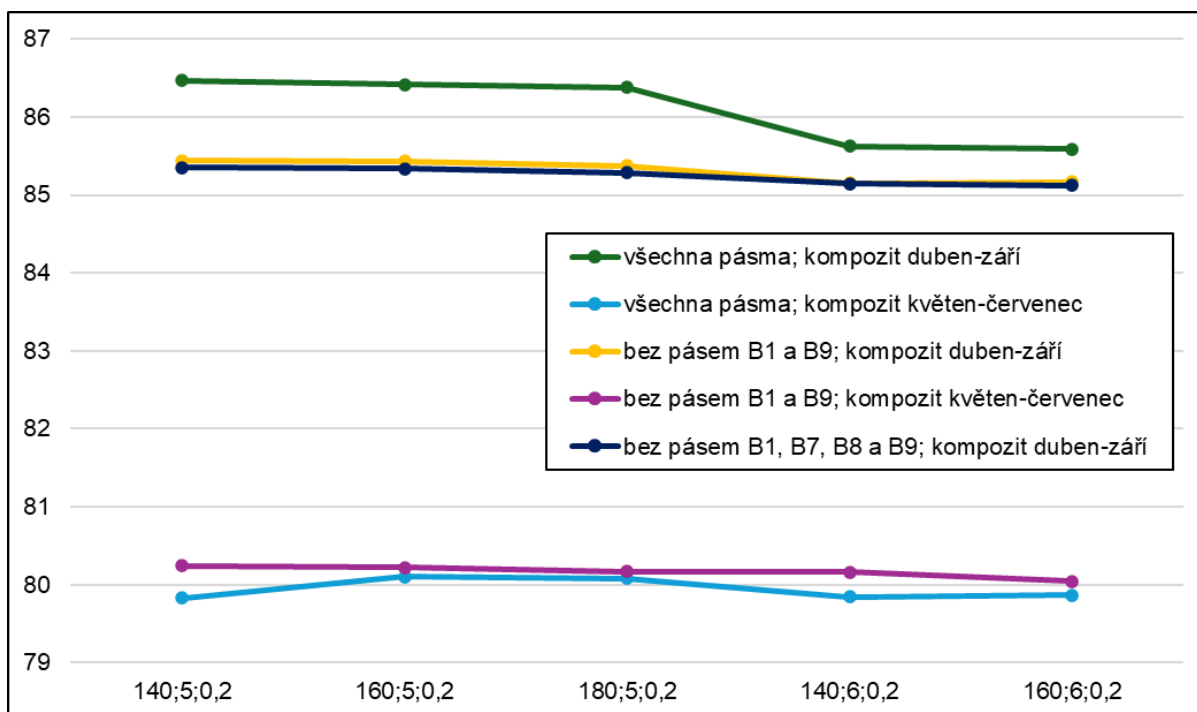
Tabulka 10: Celkové přesnosti pěti nejlepších klasifikací testovaných variant

Varianta →	průměrový, veškeré pixely	průměrový, čisté pixely	mediánový, veškeré pixely	mediánový, čisté pixely
1.	78,01	78,72	80,25	77,06
2.	77,52	78,7	80,22	77,04
3.	77,34	78,69	80,17	77,03
4.	77,33	78,69	80,16	77,03
5.	77,21	78,68	80,04	77,02

Poté byl na těchto pěti nejúspěšnějších modelech hodnocen vliv vstupních pásem a časového rozsahu snímků pro tvorbu kompozitu. Bylo testováno následujících pět variant:

- využita všechna pásma S-2, kompozit tvořen snímky od dubna do září
- odebrána pásma B1 a B9, kompozit tvořen snímky od dubna do září
- odebrána pásma B1, B7, B8 a B9, kompozit tvořen snímky od dubna do září
- využita všechna pásma S-2, kompozit tvořen snímky od května do července
- odebrána pásma B1 a B9, kompozit tvořen snímky od května do července

Výsledky tohoto testování znázorňuje Obrázek 9. Je patrné, že ve všech případech se nejlépe osvědčilo zahrnout do klasifikace všechna pásma S-2 a vytvářet kompozit z delšího časového rozsahu snímků, tedy od dubna do září.



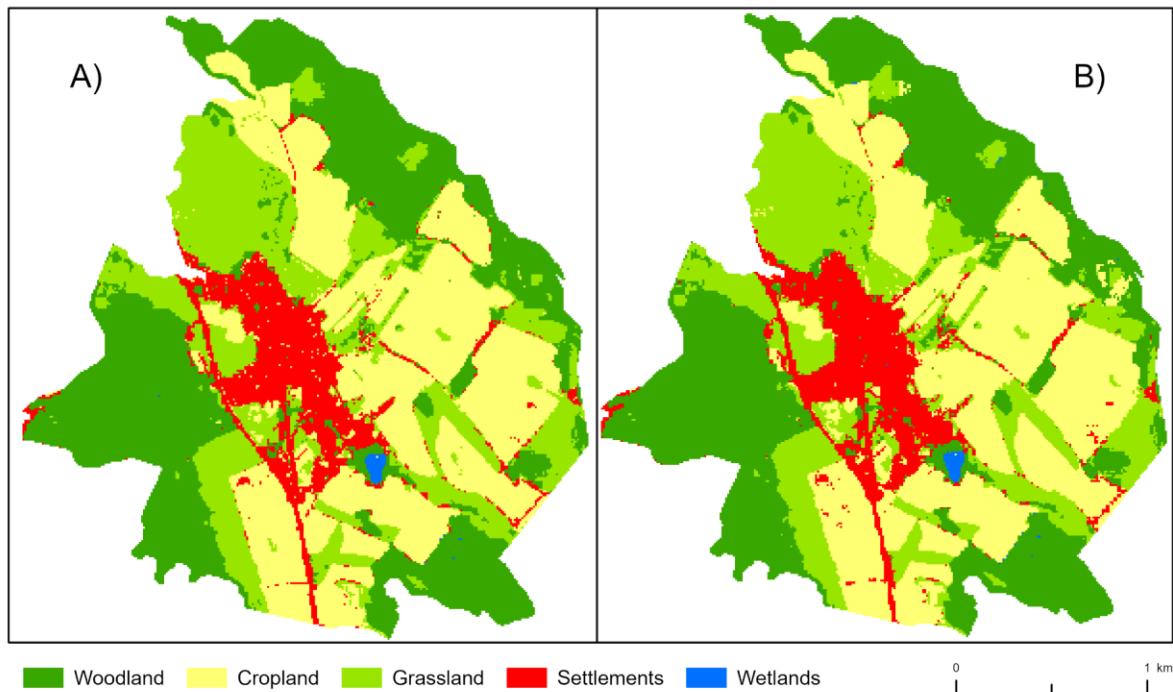
Obrázek 9: Vliv vstupních pásem a časového rozsahu snímků pro tvorbu kompozitů. Osa x: parametry NT; VPS; BF, osa y: celková přesnost klasifikace (%)

V další fázi byly opět provedeny klasifikace pro všechny zkoumané kombinace parametrů RF, ale tentokrát s nastavením “všechna pásma, kompozit duben–září”. V této fázi byla do modelu navíc přidána proměnná rozptyl NDBI a opět byly provedeny klasifikace pro všechny kombinace parametrů RF. Následující tabulka uvádí přesnosti pěti nejúspěšnějších modelů pro obě varianty.

Tabulka 11: Vliv rozptylu NDBI na výsledek klasifikace – celkové přesnosti pěti nejlepších modelů s a bez této proměnné

Varianta →	Bez zahrnutí rozptylu NDBI		Se zahrnutím rozptylu NDBI	
	NT; VPS; BF	Celková přesnost (%)	NT; VPS; BF	Celková přesnost (%)
1.	140;4;0,2	86,72	220;5;0,2	87,45
2.	160;4;0,2	86,68	200;5;0,2	87,44
3.	180;4;0,2	86,63	240;5;0,2	87,43
4.	220;4;0,2	86,62	160;5;0,2	87,42
5.	240;4;0,2	86,62	180;5;0,2	87,40

Obrázek 10 znázorňuje vliv zahrnutí rozptylu NDBI mezi vstupní proměnné na nejúspěšnějších modelech obou variant. Je patrné, že v rámci zástavby se vyskytuje výrazně méně pixelů chybně klasifikovaných jako zástavba, nicméně v rámci tříd Cropland a Grassland se stále vyskytují některé osamocené pixely chybně zařazené do třídy Settlements.



Obrázek 10: Vliv proměnné rozptylu NDBI. A) bez zahrnutí rozptylu NDBI, B) se zahrnutím rozptylu NDBI

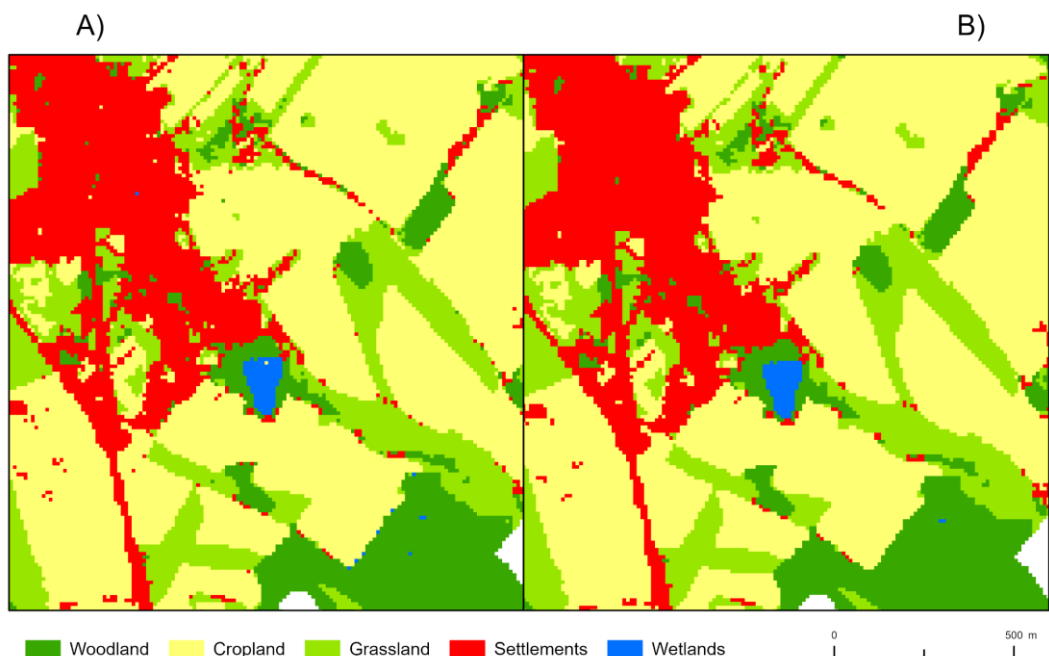
Následující tabulka podává přehled o přesnostech na trénovacích a validačních datech pěti nejúspěšnějších modelů, ve srovnání s přesností dosažené na nezávislém k. ú.

Tabulka 12: Srovnání přesností dosažených na trénovacích, validačních a nezávislých datech jako indikátor přeučení modelu

Parametry modelu	Přesnost na trénovacích datech (%)	Přesnost na validačních datech (%)	Přesnost na nezávislých datech (%)
220;5;0,2	92,72	91,93	87,45
200;5;0,2	92,75	91,92	87,44
240;5;0,2	92,75	91,89	87,43
160;5;0,2	92,75	91,89	87,42
180;5;0,2	92,75	91,88	87,40

Hodnota přesnosti na trénovacích datech je sledována především pro účely odhalení případného přeučení modelu. Vysoká přesnost na trénovacích datech a zároveň výrazně nižší přesnosti na validačních nebo nezávislých datech by naznačovala, že model se příliš přizpůsobil trénovací množině a ztrácí schopnost generalizace. V tomto případě však rozdíly mezi trénovací a validační přesností zůstaly velmi malé (pohybují se pouze kolem 0,8–0,9 %), což indikuje absenci přeučení.

Na nejúspěšnějším modelu vzešlém z proběhlých klasifikací byla následně provedena filtrace osamocených pixelů. Na Obrázku 11 je znázorněn vliv odfiltrování osamocených pixelů. Klasifikovaný obraz je na první pohled plynulejší, se zmírněním nežádoucího efektu “pepř a sůl”. Problematické oblasti se vyskytovaly jak na hranici tříd, tak i v rámci homogenní plochy jedné třídy (zejména v rámci Cropland a Settlements). V oblastech na pomezí přírodních tříd Woodland, Cropland a Grassland se nezdědky vyskytovaly pixely zařazené do Settlements či Wetlands. Takovéto chyby v klasifikovaném obrazu zůstaly i po odfiltrování osamocených pixelů, i když v mnohem omezenější míře.



Obrázek 11: Vliv odfiltrování osamocených pixelů, demonstrace problematických oblastí. A) před filtrací, B) po filtraci

Dále byla třída Woodland rozdělena na třídy Forest land a Other land. Plošky třídy Woodland s rozlohou menší než 0,05 ha, což v rastrovém klasifikovaném obrazu odpovídá shluku 5 pixelů, byly označeny jako Other land. Plošky s rozlohou od 0,05 ha byly zařazeny do třídy Forest land. Celková přesnost klasifikace k. ú. Škrdlovice se díky provedené filtraci a rozdělení tříd dostala na 87,42 %. Chybovou matici znázorňuje Tabulka 13. Hodnoty v tabulce odpovídají počtům správně či chybně klasifikovaných pixelů.

Tabulka 13: Chybová matice nejúspěšnější klasifikace RF

Klasifikovaná data →	Forest land	Cropland	Grassland	Settlements	Wetlands	Other land
Referenční data ↓						
Forest land	18266	321	524	255	15	19
Cropland	380	18165	2218	733	0	43
Grassland	838	490	10015	536	0	67
Settlements	286	121	428	4463	0	29
Wetlands	10	0	0	6	113	0
Other land	16	0	6	2	0	1

Zjištěním, které stojí za povšimnutí, je vliv vstupních pásem B1 a B9. Ve všech případech dávaly klasifikační modely lepší výsledky, byla-li tato pásma do klasifikace zahrnuta. Pásmo B1 je tradičně využíváno pro detekci aerosolů, pásmo B9 pro detekci vodní páry. Nejen, že celková přesnost byla díky těmto pásmům vyšší, ale algoritmus RF jim také ve svých rozhodovacích procesech přiřazoval nejvyšší důležitost po rozptylu NDVI a rozptylu NDBI. Jelikož jde ale o pásma s rozlišením pixelu 60 m, ve výsledném klasifikovaném obrazu se kvůli tomu v některých místech vyskytovaly nežádoucí artefakty. V některých oblastech, které byly souvisle klasifikovány jako plocha jedné třídy o rozměrech 60×60 m, byl zřejmý vliv šedesátimetrového vstupního pásma, které prostorovou přesnost klasifikace zhoršuje. Příklad takové situace je na Obrázku 12, kde je ve čtvercem 60×60 m “otupen” jižní roh zatravněné plochy.



Obrázek 12: Vliv pásem B1 a B9 na prostorovou přesnost klasifikace

5.2.2 Výsledky klasifikace Gradient Tree Boosting

Výsledky testovaných klasifikačních modelů shrnuje Tabulka 14.

Tabulka 14: Celkové přesnosti klasifikačních modelů GTB

Sampling rate	Čistota pixelů	Vstupní pásma	Celková přesnost pro NT=8 (%)	Celková přesnost pro NT=10 (%)	Celková přesnost pro NT=20 (%)	Celková přesnost pro NT=25 (%)
0,7	všechny	všechna pásma	86,12	86,14	–	–
0,7	čisté	všechna pásma	84,80	84,95	–	–
0,7	všechny	bez B1, B9	86,05	86,06	–	–
0,5	všechny	všechna pásma	86,18	86,39	–	–
0,5	všechny	bez B1, B9	86,43	86,47	–	–
0,3	všechny	všechna pásma	86,36	86,48	86,67	86,69
0,3	všechny	bez B1, B9	86,31	86,38	86,65	86,73
0,2	všechny	všechna pásma	84,72	84,78	85,86	85,99

Z výsledků uvedených v Tabulce 14 vyplývá, že nejvyšší celkové přesnosti bylo dosaženo při použití sampling rate 0,3, bez zařazení spektrálních pásem B1 a B9 a při hodnotě parametru NZ rovné 25. Tato konfigurace dosáhla celkové přesnosti 86,73 %. Obecně lze konstatovat, že mírné snížení hodnoty sampling rate (např. 0,3 oproti 0,7) nemělo negativní vliv na přesnost klasifikace. Naopak, menší množství trénovacích dat mohlo přispět ke zlepšení schopnosti modelu generalizovat. Odebrání pásem B1 a B9 mělo rovněž pozitivní vliv na přesnost modelu. Modely trénované pouze na „čistých“ pixelech dosahovaly mírně nižší přesnosti, což potvrzuje, že zařazení přechodových zón mezi třídami může zvýšit robustnost modelu.

Na klasifikovaný obraz vzešlý z nejúspěšnější klasifikace byla následně aplikována filtrace osamocených pixelů a třída Woodland byla rozdělena na Forest land a Other land. Po těchto úpravách dosáhla celková přesnost 86,7 %.

Tabulka 15 znázorňuje chybovou matici nejúspěšnější klasifikace. Hodnoty v tabulce odpovídají počtům správně či chybně klasifikovaných pixelů. Chybová matice ukazuje, že

největší chyby byly zaznamenány mezi třídami Grassland a Cropland, případně Grassland a Forest land. U třídy Settlements docházelo nejčastěji k její záměně za třídy Cropland a Grassland.

Tabulka 15: Chybová matice nejúspěšnější klasifikace GTB

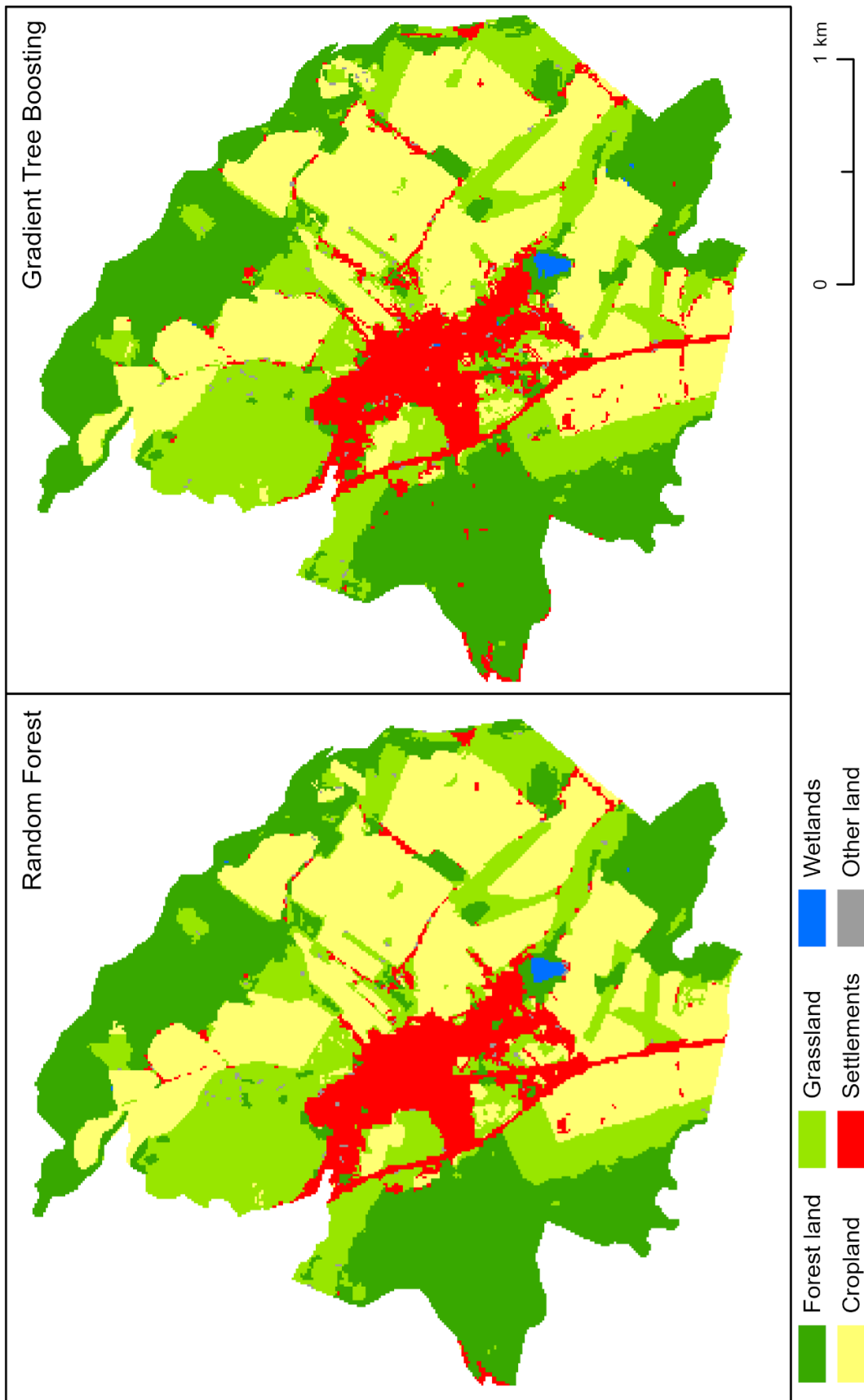
Klasifikovaná data →	Forest land	Cropland	Grassland	Settlements	Wetlands	Other land
Referenční data ↓						
Forest land	18076	340	596	334	26	28
Cropland	254	17936	2272	1017	0	60
Grassland	753	474	10047	585	0	87
Settlements	293	124	447	4399	2	62
Wetlands	14	0	0	2	113	0
Other land	11	1	7	2	0	4

5.2.3 Srovnání klasifikátorů

Pomocí metody RF se podařilo dosáhnout celkové přesnosti 87,4 %, metodou GTB pak 86,7 %. Obrázek 13 nabízí vizuální srovnání nejúspěšnějších klasifikací získaných metodami RF a GTB. V obou výstupech je patrné podobné prostorové rozložení jednotlivých tříd, rozdíly jsou nápadné zejména v kompaktnosti jednotlivých ploch. V klasifikovaných obrazech jsou nápadné izolované shluky pixelů uprostřed ploch tříd Forest land či Cropland. V některých případech se skutečně jedná o chyby, které se nepodařilo odstranit ani pomocí rozptylu NDBI, nicméně v případě GTB jsou pixely třídy Settlements v západní části území správně identifikovány jako rozptýlená zástavba – jedná se o jednotlivé domky v lese. Výstup GTB je oproti RF více ovlivněn efektem „pepř a sůl“, což vede k fragmentovanějšímu obrazu, zejména v okolí intravilánu obce. Třída Settlements se navíc v některých oblastech nevhodně mísí s třídami Cropland a Grassland. Tento jev souvisí se složením trénovací množiny – do třídy Settlements byly zařazovány přírodní plochy v obcích, které jsou spektrálně podobné ostatním přírodním třídám. RF oproti GTB častěji zařazoval plochy Cropland do třídy Forest land. Problematičnost rozlišování těchto dvou tříd mohla být způsobena zahrnutím sadů ve třídě Cropland. Oba klasifikátory často chybně zařazovaly plochy třídy Cropland do třídy Grassland. Tento jev jednoznačně souvisí s definicí třídy Cropland. Do té byly v průběhu terénního mapování zařazovány i plochy toho času ležící ladem, tedy zatravněné. Bylo tak konáno

ve snaze vytvořit trénovací dataset využitelný i pro budoucí výzkum a vycházelo se z předpokladu, že v dalších letech budou tyto plochy opět obdělávány.

Celkově lze z obrázku, chybových matic i studia mapových podkladů konstatovat, že metoda RF poskytuje kompaktnější a plynulejší klasifikovaný obraz, zatímco GTB zachycuje jemnější detaily, ale za cenu většího prostorového šumu. Metodou RF se podařilo docílit vyšší celkové přesnosti. Je však třeba zdůraznit, že tyto výsledky vycházejí z klasifikace pouze jednoho katastrálního území, což představuje příliš omezený rozsah na to, aby bylo možné učinit obecné závěry o kvalitě klasifikace jednotlivých metod. Pro komplexnější posouzení by bylo nutné provést testování na větším a rozmanitějším územním celku.



Obrázek 13: Srovnání klasifikací RF a GTB

6 DISKUSE

Hlavním cílem této diplomové práce bylo vytvoření klasifikačního modelu využitelného pro účely LULUCF reportingu v podmínkách Česka. Pro klasifikaci byla využívána data Sentinel-2 a práce probíhala zejména na platformě Google Earth Engine. Zvoleny byly klasifikátory Random Forest a Gradient Tree Boosting, které jsou v GEE dostupné a patří mezi pokročilé a perspektivní klasifikátory. V praktické části byly testovány různé varianty nastavení parametrů obou klasifikátorů a trénování modelu, konkrétně způsob tvorby kompozitu, vliv šumu v trénovacích datech a vliv výběru vstupních pásem a indexů. Významným přínosem této práce byl vlastní, podrobný trénovací dataset získaný terénním mapováním, sestavený v souladu s pravidly LULUCF reportingu. Na základě výsledků hodnocení přesnosti byl vybrán nejúspěšnější model, který byl v první řadě použit na vybrané katastrální území a následně využit pro klasifikaci širšího územního celku – celého Česka.

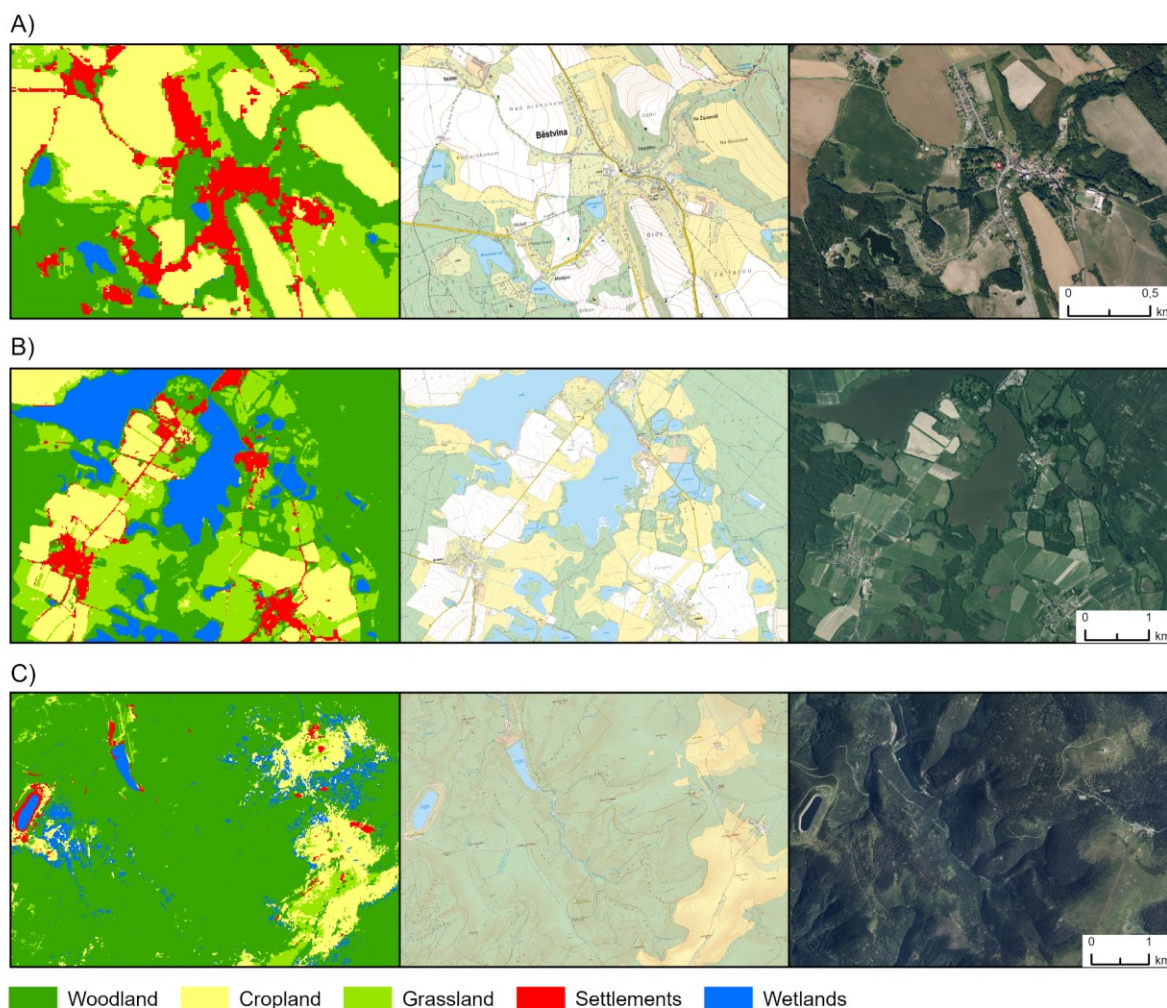
Na základě studia klasifikačních map a přesností jsou v této části textu zhodnoceny výhody a rezervy dat DPZ pro účely LULUCF. Jako nejvhodnější způsob trénování obou klasifikátorů se ukázalo využití mediánového kompozitu tvořeného snímkem Sentinel-2 pořízených v období od dubna do září 2023. Toto časové rozmezí zajistilo dostatečné množství snímků pro zachycení sezónní dynamiky vegetace v rámci různých fenologických fází. Závěry studie Phan a kol. (2020) dokládají, že využití mediánového kompozitu poskytuje srovnatelné přesnosti jako komplexnější přístup analýzy časové řady (time series), navíc s výhodou výrazně nižšího objemu dat. Pro oba klasifikátory se také pro trénování osvědčilo využití dat i s oblastí na rozhraních více LULUCF tříd, tedy s určitou mírou šumu, nikoli pouze čisté vzorky šumem nezátížené. V průběhu klasifikace se ukázal být klíčový vliv indexů NDVI a NDBI, respektive rozptylů jejich hodnot od jara do podzimu. Rozptyl NDVI zásadním způsobem přispěl ke zvýšení přesnosti klasifikace, což se ukázalo již v prvotních fázích testování. Z toho důvodu je zahrnut ve veškerém testování klasifikačních modelů v této práci. Význam rozptylu indexu NDBI se projevil až v pozdější fázi experimentální činnosti a vyplynul ze studia výsledných klasifikačních map. V nich docházelo k častému zaměňování pixelů přírodních tříd za zastavěná území a naopak. Zahrnutí rozptylu NDBI tento problém výrazným způsobem omezilo. Z výstupů algoritmu RF vyplynulo, že právě tyto dvě proměnné, tedy rozptyl NDVI a NDVI, měly z hlediska rozhodovacího procesu klasifikátoru nejvyšší důležitost. Rozptyl NDVI byl vyhodnocen jako klíčová vstupní proměnná rovněž ve studiích Svoboda a kol. (2022) a Phalke a kol. (2020).

Překvapivým zjištěním práce bylo, že RF zařadil mezi nejdůležitější vstupní proměnné pásma B1 a B9. Pásmo B1 je využíváno pro detekci aerosolů, pásmo B9 pro detekci vodní páry, navíc jde o pásma s prostorovým rozlišením 60 m, zatímco ostatní pásma mají 10 či 20 m. Jejich zahrnutí do klasifikačního modelu však v drtivé většině případů vedlo ke zlepšení

celkové přesnosti, i když občas za cenu prostorových artefaktů v klasifikovaném obrazu. K vysvětlení tohoto jevu a ověření relevance by byl zapotřebí další výzkum a testování. Z dostupných dat a interpretací není možné jednoznačně určit důvod zvýšení přesnosti klasifikace při zahrnutí těchto pásem. Ve studii Phan a kol. (2020) autoři prováděli klasifikace RF s využitím dat Landsat 8. Do klasifikací zahrnuli pásmo B1 snímku Landsat 8 jakožto jednu z pomocných proměnných. Ve všech případech dosahovaly klasifikace s pomocnými proměnnými vyšších přesností, a pásmo B1 bylo z celkových dvaceti vyhodnoceno jako čtvrté nejdůležitější pro rozhodování algoritmu. Většinou praxí ovšem je, že jsou tato dvě pásma z land cover/land use klasifikací vyjímána (Svoboda a kol. (2022), Wang a kol. (2016), Gašparović a Jogun (2017)).

V průběhu práce bylo nezbytné přizpůsobit postupy výpočetním omezením platformy GEE. Co se týče klasifikátoru RF, bylo nutné vytvořit samostatné skripty, jeden pro trénování klasifikačních modelů a jejich ukládání do složky Assets (úložiště vlastních datových sad uživatele v GEE), a druhý pro jejich následné spuštění. V obou případech bylo využito programovacích cyklů a operace se prováděla pro všechny požadované kombinace parametrů. Snaha sloučit v jednom skriptu jak trénování, tak spuštění 200 klasifikačních modelů narazila na výpočetní limity platformy – objevily se chybové hlášky o překročení výpočetní kapacity. Problém nespočíval v počtu spouštěných modelů, ale spíše ve velikosti vstupního datasetu, který lze označit za big data. V případě klasifikátoru GTB bylo trénování i spuštění prováděno v jednom skriptu, nicméně tento byl spouštěn zvlášť pro každou variantu, bez využití cyklů. Zde výpočetní kapacita GEE omezovala počet stromů, který mohl být využit pro klasifikaci (bez výraznějších omezení parametrů, než je uvedeno v textu). Klasifikaci s využitím 20 stromů již pro některé testované varianty nebylo možné provést.

Model, který dosáhl nejvyšší úspěšnosti při klasifikaci nezávislého katastrálního území, byl následně aplikován na území celého Česka. Tento výsledek je v DP uveden jako důkaz funkčnosti vytvořeného algoritmu, ale v rámci rozsahu této DP již nebylo možné detailně vyhodnotit dílčí i celkovou přesnost pomocí chybové matice a interpretovat výsledky. Nicméně byla provedena vizuální zhodnocení řady lokalit a typů krajiny a vybrané výsledky jsou uvedeny v této DP. Klasifikační mapa Česka k roku 2023 je v Příloze 3. Následující obrázek ukazuje detailní pohledy na některé lokality.



Obrázek 14: Přiblížení klasifikované vrstvy Česka. A) Běstvina v okrese Chrudim, B) Opatovický rybník a Svět u Třeboně, C) Hrubý Jeseník – Dlouhé Stráně, Praděd, Ovčárna

Co se týče lokalit A) a B), z klasifikovaného obrazu je dobře patrná reálná struktura krajiny, i když v některých místech dochází k záměnám bloků třídy Cropland za třídu Grassland, zároveň v plochách těchto tříd se občas vyskytují pixely chybně klasifikované jako Settlements. Na druhou stranu, lokalita C) zde slouží jako ukázka problematičnosti klasifikace v horských oblastech. Subalpínské louky v Hrubém Jeseníku byly klasifikovány jako mix tříd Cropland, Grassland, Wetlands i Settlements. Zároveň se v ploše Woodland hojně vyskytují pixely zařazené do Wetlands, nejedná se však o lokalitu rašelinišť, které by v tomto směru mohly způsobovat jisté zkreslení. V západní části výřezu jsou však patrné dobře vyklasifikované nádrže PVE Dlouhé stráně.

Celkově vzato, nejlepší výsledky model poskytoval v oblastech podobným referenčním k. ú., tedy menším obcím v zemědělské krajině. Naopak ve větších městech docházelo k výraznějším chybám, zejména k nesprávnému zařazování přírodních ploch do třídy Settlements. Nejproblematictější se ukázala být klasifikace horských oblastí. Je zřejmý vliv

absence trénovacích dat právě z horských lokalit, které by zahrnovaly například husté svažité lesy, bezlesí s kosodřevinami a kamennými moři, či subalpínské louky. Zahrnutí těchto ploch do trénovací množiny by s velkou pravděpodobností zlepšilo výsledky klasifikace v horských lokalitách. Další možností, jak zvýšit přesnost klasifikace v těchto oblastech, je využití digitálního modelu terénu (DMT), jak s úspěchem prováděli například Phan a kol. (2020) nebo Svoboda a kol. (2022). Vzhledem k rovinatému reliéfu mapovaných území nebyl DMT do modelu přidáván – v těchto podmínkách by měl pro učení modelu pouze omezený přínos.

Zatímco data DPZ poskytují informaci spíše o fyzických vlastnostech povrchu, tedy land cover, třídy LULUCF odpovídají spíše využití území, tedy land use. Každá LULUCF třída zpravidla zahrnuje širší spektrum typů povrchů s odlišnou spektrální charakteristikou, například ve třídě Forest land jde o porosty stromů i dočasně vykácené pozemky, ve třídě Cropland o ornou půdu i sady, a do třídy Settlements jsou zahrnovány jak umělé povrchy (zástavba, silnice, lomy), tak i přírodní plochy ve městech. Z uvedeného vyplývá potenciální zdroj chyb prováděných klasifikací z dat DPZ, a také jisté omezení v jejich využívání pro tento účel.

Dalším limitem dat DPZ, zde konkrétně dat Sentinel-2, je jejich relativně nízké prostorové rozlišení ve srovnání s daty KN. Nižší rozlišení nejen že samo o sobě snižuje prostorovou přesnost klasifikace, ale zároveň zvyšuje citlivost výsledků na šum a efekt „pepř a sůl“ v klasifikační vrstvě. Každý jednotlivý chybný pixel totiž odpovídá ploše 100 m², což může při plošných výpočtech způsobovat významné zkreslení. Dále je nutné zohlednit skutečnost, že data z DPZ jsou rastrového charakteru, což samo o sobě přesné výpočty rozloh komplikuje. Na druhou stranu, ani KN nelze považovat za bezchybný referenční zdroj. Z hlediska správnosti a aktuálnosti záznamů má svá omezení, jak poukazují např. studie Míček a kol. (2020) či Pazúr a kol. (2015). Zastaralé nebo záměrně neaktualizované informace – například kvůli zachování nároku na dotační podporu – mohou být významným zdrojem nepřesností při porovnávání s aktuálními satelitními daty.

Data Sentinel-2 představují z hlediska využití pro účely LULUCF velký potenciál, avšak výsledky této práce ukazují, že samotná rastrová klasifikace není zcela dostačující. Jako vhodnější přístup se nabízí například kombinace dat KN a DPZ. KN se jeví jako přesnější zdroj pro třídu Settlements, kde by mohl spolehlivě fungovat například na dopravní infrastrukturu, která bývá v rastrových klasifikacích často přerušovaná nebo nesouvislá. Naopak DPZ lépe zachycuje přírodní třídy, neboť poskytuje objektivní informace o aktuálním stavu krajiny. Určitým problémem zůstávají plochy Cropland ležící ladem, které v některých případech bývají chybně klasifikovány jako travní porosty. V tomto směru by mohlo pomoci využití časových řad za víceletá období nebo integrace dalších datových zdrojů, například LPIS. Významným předpokladem pro zvýšení přesnosti by zároveň bylo vytvoření

heterogennějšího trénovacího datasetu, než jaký nabízí tato diplomová práce, zejména s ohledem na zastoupení horských oblastí. Problematické je také slučování tříd Forest land a Other land do jedné třídy pro účely klasifikace, a jejich následné rozdělování na základě rozlohy v rastrovém obrazu. Z tohoto hlediska se jeví jako perspektivní využití objektivě orientované klasifikace. Příkladem již existujícího takového přístupu je databáze CLC+, založená na objektivě orientované klasifikaci, jedním z jejíž výstupů (tzv. CLC+ Instances) mají být vrstvy odpovídající požadavkům LULUCF.

Reporting v rámci LULUCF slouží jako důležitý vstup do klimatických modelů, které jsou základem pro plánování adaptačních a mitigačních opatření souvisejících se změnou klimatu. Jsou-li vstupní data nepřesná, tato nepřesnost se promítne i do výstupů modelů. Vzhledem k tomu, že realizace adaptačních či mitigačních opatření je často finančně nákladná, je žádoucí, aby byla tato opatření co nejpřesněji cílená a efektivní. Z tohoto důvodu je kladen důraz na kvalitu a spolehlivost dat využívaných v rámci LULUCF reportingu.

Součástí práce bylo také představení a praktické porovnání vybraných volně dostupných globálních land cover/land use datasetů, konkrétně CLC+, ESA World Cover, Dynamic World a ESRI Land Cover. Smyslem bylo nabídnout další, alternativní zdroje informací o krajině, potenciálně využitelných pro účely LULUCF. Zmíněné datasety byly porovnány s referenční vrstvou k. ú. jak vizuálně, tak na základě rozloh jednotlivých tříd. Z analýzy tabulek, grafů a klasifikačních map vyplynulo, že datasety ESA World Cover a CLC+ mají tendenci nadceňovat travní porosty, zatímco podhodnocují zástavbu a ornou půdu. Dynamic World vykazuje oproti referenčním datům vyšší zastoupení lesních porostů, ale současně podhodnocuje travní plochy a zástavbu. ESRI Land Cover jako jediný naopak nadhodnocuje ornou půdu a podhodnocuje travní porosty. Tato pozorování jsou v souladu s několika nedávnými studiemi. Duarte a kol. (2023) porovnávali ESA World Cover s národním land cover/land use datasetem Portugalska a došli k obdobným závěrům: přeceňování travních ploch a podhodnocování zástavby a orné půdy datasetem ESA World Cover. Venter a kol. (2022), kteří se věnovali vzájemnému porovnání ESA World Cover, ESRI Land Cover a Dynamic World s referenčním datasetem LUCAS, rovněž zaznamenali přeceňování travních porostů v datasetu ESA World Cover. Jejich studie dále upozorňuje na tendenci Dynamic World nadhodnocovat třídu Snow and ice, což se potvrdilo i v rámci této práce – konkrétně šlo o malou plošku orné půdy v katastrálním území Škrdlovice, který byl chybně klasifikován právě jako Snow and ice. ESA World Cover byl v uvedených studiích obecně hodnocen jako dataset s nejvyšší prostorovou přesností. Nejlepší prostorový detail datasetu ESA World Cover a nadceňování třídy Snow and ice u Dynamic World potvrzují i Xu a kol. (2024), kteří srovnávali ESA World Cover, Dynamic World a ESRI Land Cover v celosvětovém měřítku.

Předložené srovnání je nutno interpretovat s jistou mírou opatrnosti, byť se některé závěry shodují napříč více autory. Vzájemné srovnávání datasetů je velmi komplexní úlohou.

Mezi faktory, na které je třeba dbát a které komplikují prosté srovnání rozloh jednotlivých tříd v různých datasetech patří například použitá MMU, referenční rok či způsob pořizování dat. Jisté prostorové zkreslení také může pramenit z převodů dat mezi rastrovým a vektorovým formátem. Významnou roli hraje rovněž rozdílnost klasifikačních systémů. Při jejich převodu do jednotné nomenklatury může docházet ke zkreslení či ztrátě specifických významů jednotlivých tříd.

Veškeré GEE kódy vytvořené v rámci této práce jsou k dispozici [na sdíleném úložišti](#).

7 ZÁVĚR

Členské státy EU každoročně vykazují emise skleníkových plynů z několika sektorů lidské činnosti. Jedním z těchto sektorů je LULUCF, jemuž byla věnována předkládaná diplomová práce. V rámci LULUCF je krajina členěna do šesti tříd (Forest land, Cropland, Grassland, Settlements, Wetlands, Other land). V Česku jsou v současnosti pro tento účel využívána data KN. Tato DP se podrobně zabývá možnostmi využívání dat DPZ pro účel LULUCF.

Součástí práce bylo představení a praktické porovnání vybraných volně dostupných globálních datasetů land cover/land use založených na datech dálkového průzkumu Země, konkrétně CLC+, ESA World Cover, Dynamic World a ESRI Land Cover. Cílem bylo nabídnout přehled alternativních zdrojů informací o krajině, které by mohly být potenciálně využitelné pro účely LULUCF reportingu. Jako velmi perspektivní se jeví databáze CLC+, respektive její výstupy CLC+ LULUCF Instances.

Práce se také zabývala tvorbou vlastního klasifikačního modelu. Využita byla satelitní data Sentinel-2 a výpočty probíhaly především v prostředí Google Earth Engine. V rámci práce byly testovány dva klasifikační algoritmy, Random Forest a Gradient Tree Boosting. Hodnocen byl vliv vstupních pásem Sentinel-2, způsobu tvorby kompozitu, čistoty dat pro trénování modelu a nastavení parametrů klasifikátorů. Významnou přidanou hodnotou této práce byl vlastní, podrobný trénovací dataset získaný terénním mapováním, sestavený v souladu s pravidly LULUCF reportingu. Část tohoto datasetu sloužila pro trénování modelů, část byla použita pro ověření přesnosti a přenositelnosti v nezávislém území. Nejlepších výsledků bylo dosaženo při využití mediánového kompozitu ze snímků z období od dubna do září a začlenění trénovacích vzorků z hraničních oblastí mezi třídami. Pro vylepšení výsledků klasifikace se také ukázal být klíčový vliv indexů NDVI a NDBI, respektive jejich rozptylů. Klasifikátor RF poskytoval více prostorově soudržné výstupy, zatímco Gradient Tree Boosting lépe identifikoval drobné objekty, ale zároveň vykazoval více šumu a fragmentace. Metodou RF se podařilo dosáhnout celkové přesnosti 87,4 %, metodou GTB pak 86,7 %. Nejúspěšnější model RF byl následně aplikován na celé území Česka jako ukázka funkčnosti tohoto algoritmu.

Zjištění této práce potvrdila, že prostá rastrová klasifikace pro účely LULUCF reportingu má svá omezení. Jako nejproblematičtější se ukázalo samotné pojetí LULUCF tříd, které vycházejí z principu land use, zatímco satelitní data popisují primárně land cover. Jednotlivé třídy LULUCF často slučují spektrálně odlišné typy povrchu, což komplikuje klasifikaci. Dále je limitujícím faktorem relativně nízké prostorové rozlišení dat Sentinel-2 a jejich rastrový charakter, který z hlediska počítání rozloh jednotlivých tříd zvyšuje dopad i drobných klasifikačních chyb. I přes uvedená omezení je zřejmé, že data Sentinel-2 mají pro využití

v LULUCF reporting značný potenciál. Předkládaná diplomová práce navazuje na předchozí výzkum v této oblasti a dále jej rozvíjí prostřednictvím komplexního testování možností DPZ pro klasifikaci LULUCF tříd na území Česka, přičemž využívá vlastní podrobný trénovací dataset získaný terénním mapováním, vytvořený v souladu s požadavky LULUCF.

V rámci dalšího výzkumu by bylo vhodné rozšířit trénovací dataset o další typy české krajiny. Dále se jako perspektivní jeví využití digitálního modelu terénu, kombinace více datových zdrojů (např. KN pro antropogenní třídy vyžadující vyšší prostorový detail a LPIS pro identifikaci úhorů) či přechod na objektově orientované klasifikační přístupy.

ZDROJE

Literatura

Blujdea, V., Abad Vinas, R., Federici, S., Grassi, G. (2016): The EU greenhouse gas inventory for LULUCF sector: I. Overview and comparative analysis of methods used by EU member states. *Carbon management* 6 (5); 247-259.

Breimann, L. (2001): Random Forests. *Machine Learning* 45(1), 5-32.

Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Brooks Hyde, S., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., tait, A. M. (2022): Dynamic World, Near real-time global 10 m land use land cover mapping. *Sci Data* 9, 251.

European Environment Agency (2020): CLC+ Backbone Product Specification And User Manual for CLC+ Backbone Raster Product 2021. Issue 1.2.

Duarte, D., Fonte, C., Costa, H., Caetano, M. (2023): Thematic Comparison between ESA WorldCover 2020 land Cover product and a National Land Use Land Cover Map. *Land*, 12, 490.

Esch, T., Metz, A., Marconcini, M., Keil, M. (2014): Combined use of multi-seasonal high and medium resolution satellite imagery for parcel-related mapping of cropland and grassland. *International Journal of Applied Earth Observation and Geoinformation*, 28, 230-237.

Gao, B.-C. (1996): NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 58, 257–266.

Gašparović, M., Jogun, T. (2017): The effect of fusing Sentinel-2 bands on land-cover classification. *International Journal of Remote Sensing*, 39(3), 822–841.

Latta, G. S., Baker, J. S., Ohrel, S. (2018): A Land use and Resource Allocation (LURA) modeling system for projecting localized forest CO₂ effects of alternative macroeconomic futures. *Forest and Policy Economics*, 87, 35-48.

Hendriks, C., Arets, E., Huijstee, J. (2021): WOt special 3, Statutory Research Tasks Unit for Nature & the Environment (WOT Natuur & Milieu).

Ho, T. K. (1995): Random Decision forests. Proceeding of the Third International Conference on Document Analysis and Recognition. Montreal, Kanada, 278-282.

Klouček, T., Moravec, D., Komárek, J., Lagner, O., Štych, P. (2018): Selecting appropriate variables for detecting grassland to cropland changes using high resolution satellite data. PeerJ, 6, 1-20.

Míček, O., Feránek, J., Štych, P. (2020): Land use/land cover data of the urban atlas and the cadastre of real estate: An evaluation study in the Prague metropolitan region. Land, 9, 153.

Pazúr, R.; Feranec, J.; Štych, P.; Kopecká, M.; Holman, L. (2015): Changes of urbanised landscape identified and assessed by the urban atlas data: Case study of Prague and Bratislava. Land Use Policy, 61, 135–146.

Phalke, A. R., Ozdogan, M., Thenkabail, P. S., Erickson, T., Gorelick, N., Yadav, K., Congalton, R. G. (2020): Mapping croplands of Europe, Middle East, Russia, and Central Asia using Landsat, Random Forest, and Google Earth Engine. ISPRS Journal of Photogrammetry and Remote Sensing, 167, 104–122.

Phan, T. N., Kuch, V., Lehnert, L. W. (2020): Land Cover Classification using Google Earth Engine and Random Forest Classifier—The Role of Image Composition. Remote Sens. 12(15), 2411.

Penman, J.; Gytarsky, M.; Hiraishi, T.; Krug, T.; Kruger, D.; Pipatti, R.; Buendia, L.; Miwa, K.; Ngara, T.; Tanabe, K.; et al. (2003): Good Practice Guidance for Land Use, Land-Use Change and Forestry; Institute for Global Environmental Strategies (IGES) for the IPCC: Kanagawa, Japan, 590p.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J. P. (2012): An assessment of the effectiveness of a random forest classifier for land-cover classification, ISPRS Journal of Photogrammetry and Remote Sensing, 67: 93–104.

Savaresi, A., Perugini, L., Chiriaco, M. V. (2020): Making sense of the LULUCF Regulation: Much ado about nothing? RECIEL, 29, 212-220.

Schils, R. L. M., Buffe, C., Rhymer, C. M., Francksen, R. M., Klaus, V., K., Abdalla, M., Milazzo, F., Lellei-Kovács, E., ten Berge, H., Bertora, C., Chodkiewicz, A., Damatirca, C., Feigenwinter, I., Fernández-Rebollo, P., Ghiasi, S., Hejduk, S., Hiron, M., Janicka, M., Pellaton, R., Smith, K. E., Price, P. N. (2022): Permanent grasslands in Europe: Land use change and intensification decrease their multifunctionality. *Agriculture, Ecosystems & Environment* 330, 107891.

Sperry, K.P.; Hilfer, H.; Lane, I.; Petersen, J.; Dixon, P.M.; Sullivan, L.L. Species diversity and dispersal traits alter biodiversity spillover in reconstructed grasslands. *J. Appl. Ecol.* 2019, 56, 2216–2224.

Sun, J., Ongsomwang, S. (2023): Optimal parameters of random forest for land cover classification with suitable data type and dataset on Google Earth Engine. *Front. Earth Sci.* 11:1188093.

Svoboda, J.; Štych, P.; Laštovička, J.; Paluba, D.; Kobliuk, N. (2022): Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia. *Remote Sens.* 14, 1189.

Šandera, J.; Štych, P. (2020): Selecting relevant biological variables derived from Sentinel-2 data for mapping changes from grassland to arable land using random forest classifier. *Land*, 9, 420.

Tucker, C. J. (1979): Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150.

Venter, Z. S., Barton, D. N., Chakraborty, T., Siemensen, T., Singh, G. (2022): Global 10 m Land Use land Cover Datasets: A Comparison of Dynamic World, World Cover and Esri Land Cover. *Remote Sens*, 14, 4101.

Wang, Q., Shi, W., Li, Z., Atkinson P. M. (2016): Fusion of Sentinel-2 images. *Remote Sens. Environ.*, 187, 241–252.

Xiao, W., Xu, S., and He, T. (2021): Mapping paddy rice with Sentinel-1/2 and phenology-object-based algorithm—an implementation in hangjiahu plain in China using GEE platform. *Remote Sens.* 13, 990.

Xu, P., Tsendbazar, N.-E., Herold, M., Bruin, S., Koopmans, M., Birch, T., Carter, S., Fritz, S., Lesiv, M., Mazur, E., Pickens, A., Potapov, P., Stolle, F., Tyukavina, A., Kerchove, R., Zanaga, D. (2024): Comparative validation of recent 10 m-resolution global land cover maps. *Remote Sens. Environ.*, 311, 114316.

Yang, Y., Yang, D., Wang, X., Zhang, Z., Nawaz, Z. (2021): Testing accuracy of land cover classification algorithms in the Qilian mountains based on GEE cloud platform. *Remote Sens.* 13, 5064.

Zha, Y., Gao, J., Ni, S. (2003): Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* 24, str. 583–594.

Oficiální dokumenty

EEA (2022): Annual European Union Greenhouse Gas Inventory 1990–2020 and Inventory Report 2023. Submission to the UNFCCC Secretariat. Dostupné z:

<https://www.eea.europa.eu/en/analysis/publications/annual-european-union-greenhouse-gas-2>

IPCC (2019), Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems, Summary for Policy Makers. Dostupné z:

<https://www.ipcc.ch/srccl/chapter/summary-for-policymakers/>

Nařízení Evropského parlamentu a Rady (EU) 2018/841 ze dne 30. května 2018 o zahrnutí emisí skleníkových plynů a jejich pohlcování v důsledku využívání půdy, změn ve využívání půdy a lesnictví do rámce politiky v oblasti klimatu a energetiky do roku 2030 a o změně nařízení (EU) č. 525/2013 a rozhodnutí č. 529/2013/EU. Dostupné z: [https://eur-](https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=CELEX%3A32018R0841)

[lex.europa.eu/legal-content/CS/TXT/?uri=CELEX%3A32018R0841](https://eur-lex.europa.eu/legal-content/CS/TXT/?uri=CELEX%3A32018R0841)

Vyhláška č. 357/2013 Sb. (Vyhláška o katastru nemovitostí)

Vyhláška č. 84/1996 Sb. (Vyhláška Ministerstva zemědělství o lesním hospodářském plánování)

Zákon č. 289/1995 Sb. (Zákon o lesích a o změně některých zákonů)

Internetové zdroje

Biologické centrum AV ČR: *Kůrovcová kalamita* (online). 2024 (cit. 14. 4. 2024). Dostupné z: <https://www.podkurou.cz/seznamte-se-s-kurovcem/kurovcova-kalamita>

Copernicus: *CLCplus: a framework for harmonized and tailored land cover information of Europe* (online). 2025 (cit. 22. 4. 2025). Dostupné z: <https://land.copernicus.eu/en/products/clc-a-new-generation-land-information-system-for-europe>

ČSÚ: *Farm Structure Survey results* (online). 2014 (cit. 14. 4. 2024). Dostupné z: https://www.czso.cz/csu/czso/2127-06-2005_analysis-uvod

ČSÚ: Veřejná databáze, vlastní výběr (online). 2025 (cit. 8.3.2025). Dostupné z: <https://vdb.czso.cz/vdbvo2/faces/cs/index.jsf?page=uziv-dotaz#>

CzechTerra (online). 2015 (cit. 14. 5. 2024). Dostupné z: <https://www.czechterra.cz/>

ČHMÚ (2023): *National greenhouse gas inventory report of the Czech Republic. Submission under UNFCCC*. Dostupné z: https://www.chmi.cz/files/portal/docs/uoco/oez/nis/NIR/CZE_NIR-2023-2021_UNFCCC_allinone_ISBN.pdf

ČHMÚ (2015): *National greenhouse gas inventory report of the Czech Republic. Submission under UNFCCC*. Dostupné z: https://www.chmi.cz/files/portal/docs/uoco/oez/nis/NIR/CZE_NIR-2015-2013_final_UNFCCC_final.pdf

ČHMÚ: *Skleníkový efekt. Studie pro MŽP* (online). 2024 (cit. 14. 4. 2024). Dostupné z: https://www.chmi.cz/files/portal/docs/meteo/ok/klimazmena/files/cc_chap05.pdf

Encyklopedie Českých Budějovic (online). 2024 (cit. 15. 8. 2024). Dostupné z: <https://encyklopedie.c-budejovice.cz/>

Eurostat: *Farm Structure Survey* (online). 2023 (cit 15. 5. 2024). Dostupné z: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Farm_structure_survey_\(FSS\)](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Farm_structure_survey_(FSS))

Geoportál ČÚZK: *Katastrální mapa ČR ve formátu SHP distribuovaná po katastrálních územích* (online). 2023 (cit. 2. 4. 2023). Dostupné z: [https://geoportal.cuzk.cz/\(S\(h1lymgvzwhezavikf3syrurz\)\)/Default.aspx?mode=TextMeta&side=katastr_map&metadataID=CZ-00025712-CUZK_SERIES-MD_KM-KU-SHP&menu=211](https://geoportal.cuzk.cz/(S(h1lymgvzwhezavikf3syrurz))/Default.aspx?mode=TextMeta&side=katastr_map&metadataID=CZ-00025712-CUZK_SERIES-MD_KM-KU-SHP&menu=211)

Geoprohlížeč ČÚZK: *Analýzy Výškopisu* (online). 2025 (cit. 8.3.2025). Dostupné z: <https://ags.cuzk.cz/av/>

Obec Klec (online). 2024 (cit. 15. 8. 2024). Dostupné z: <https://obecklec.cz/>

Obec Prachovice (online). 2024 (cit. 15. 8. 2024). Dostupné z: <https://www.obecprachovice.cz/>

Ministerstvo životního prostředí: *Pařížská dohoda* (online). 2023 (cit. 14. 4. 2024). Dostupné z: https://www.mzp.cz/cz/zmena_klimatu

Ministerstvo životního prostředí: *Změna klimatu* (online). 2025 (cit. 12. 3. 2025). Dostupné z: https://www.mzp.cz/cz/parizska_dohoda

ČÚZK: Prohlížeč služby - WMS (online). Dostupné z: [https://geoportal.cuzk.cz/\(S\(abavoqgimhe12egfghluru2f\)\)/Default.aspx?mode=TextMeta&side=wms.verejne&text=WMS.verejne.uvod](https://geoportal.cuzk.cz/(S(abavoqgimhe12egfghluru2f))/Default.aspx?mode=TextMeta&side=wms.verejne&text=WMS.verejne.uvod) V této práci byly využívány WMS ZM10, WMS MČR1M a Ortofoto.

ÚHÚL: *Národní inventarizace lesů* (online). 2023 (cit. 15. 4. 2023) Dostupné z: <https://www.uhul.cz/portfolio/nil/>

Veřejný registr půdy - LPIS (online). 2025 (cit. 10. 6. 2023) Dostupné z: <https://mze.gov.cz/public/app/lpisext/lpis/verejny2/plpis/>

Wikipedia: *Elektrárna Chvaletice* (online). 2024 (cit. 15. 8. 2024). Dostupné z: https://cs.wikipedia.org/wiki/Elekt%C3%A1rna_Chvaletice

ZÚ, ČSÚ, ARCDATA Praha: *ArcČR 500 verze 4.3 (2024)* (online) Dostupné z: https://www.arcdata.cz/cs-cz/produkty/data/arccr_(cit. 1. 12. 2023)

SEZNAM PŘÍLOH

Příloha 1: Klasifikační nomenklatury vybraných land cover/land use databází

Příloha 2: Srovnání land cover/land use databází v řešených katastrálních územích

Příloha 3: Klasifikace LULUCF pro území Česka

Příloha 1: Klasifikační nomenklatury vybraných land cover/land use databází

ESA World Cover

Tree Cover – oblasti pokryté stromy z více než 10 %, pod jejichž korunami může být bylinný pokryv, dále plantáže a hospodářské lesy.

Shrubland – oblasti pokryté keřovou vegetací nižší než 5 m z více než 10 %.

Grassland – oblasti pokryty bylinnou vegetací z více než 10 %, tedy travní porosty, prairie, stepi, savany, či pastviny. Stromy a keře nepokrývají víc než 10 % povrchu. Může zde být zahrnuta také zemědělská půda v daném roce neobhospodařovaná.

Cropland – obhospodařovaná zemědělská půda, možno i se skupinami stromů.

Built-up – zastavěné oblasti, tedy budovy, dopravní infrastruktura. Není zde zahrnuta městská zeleň a sportoviště s přírodním povrchem.

Bare / sparse vegetation – oblasti pokryté holou půdou, pískem, skalnaté oblasti atd., kde vegetační pokryv je maximálně 10 %, dále skládky odpadu a lomy.

Snow and Ice – oblasti trvale pokryté sněhem či ledem.

Permanent water bodies – oblasti, které jsou více než 9 měsíců v roce pokryty vodou.

Herbaceous wetland – oblasti, kde bylinný pokryv tvoří více než 10 %, které jsou permanentně či pravidelně zaplavovány vodou.

Mangroves – mangrovové porosty.

Moss and lichen – oblasti pokryté mechy a/nebo lišejníky.

(Zdroj: Kerchova a kol. 2022)

Dynamic World

Trees – primární a druhotné lesy, velkoplošné plantáže.

Grass – travní porosty, pastviny, parky.

Crops – zemědělská půda, včetně rýžovišť.

Shrub and Scrub – řídká či hustá keřovitá vegetace.

Built Area – budovy, dopravní infrastruktura.

Flooded Vegetation – mangrovové lesy a jiné zaplavené ekosystémy.

Water – trvalé či sezónní vodní útvary.

Bare Ground – pouště a skály.

Snow and Ice – trvalý nebo sezónní pokryv sněhem či ledem.

(Zdroj: Brown a kol. 2022)

ESRI Land Cover

Trees – jakékoli shluky alespoň 4,5 m vysoké husté vegetace, typicky s hustými korunami, například lesy, shluky stromů v savanách, plantáže, mangrovové lesy (kde voda pod nimi není detekovatelná).

Crops – zemědělsky obhospodařované plochy s plodinami či travinami.

Rangeland – otevřené plochy porostlé zejména travinami, například přírodní louky, savany bez či s málo stromy, parkové trávníky, golfové hřiště, pastviny.

Flooded vegetation – oblasti s libovolným vegetačním krytem, které jsou většinu roku viditelně zaplavené/podmáčené, nebo sezónně zaplavované, například mangrovové lesy s viditelnou vodní hladinou, rýžová pole, intenzivně zaplavované zemědělské oblasti.

Built Area – zastavěné oblasti: budovy, dopravní infrastruktura, nepropustné povrchy.

Water – oblasti, které většinu roku tvoří vodní plochy; například řeky, rybníky, moře.

Bare ground – oblasti pokryté holou půdou či skalami, s řídkou či žádnou vegetací, například pouště, zasolené pánve, lomy.

Snow/Ice – plochy s trvale ležícím sněhem či ledem.

Clouds – oblasti, kde o land cover nejsou informace, kvůli trvalému výskytu oblačnosti.

(Zdroj: Karra a kol. 2021)

Corine Land Cover+

Rastrový produkt

- 1 Sealed – nepropustné povrchy
- 2 Woody: needle leaved trees – dřeviny: jehličnaté stromy
- 3 Woody: broadleaved deciduous trees – dřeviny: listnaté opadavé stromy
- 4 Woody: broadleaved evergreen trees – dřeviny: listnaté stálezelené stromy
- 5 Low-growing woody plants: bushes, shrubs – nízké dřeviny
- 6 Permanent herbaceous – trvalé bylinné porosty
- 7 Periodically herbaceous – sezónní bylinné porosty
- 8 Lichens and mosses – lišejníky a mechy
- 9 Non- and sparsely- vegetated – povrchy bez vegetace či s řídkou vegetací
- 10 Water – voda
- 11 Snow and ice – sníh a led

Vektorový produkt

- 11 Very High Sealing Degree – plochy s velmi vysokým podílem nepropustného povrchu
- 12 High sealing Degree – plochy s vysokým podílem nepropustného povrchu
- 21 Pure needle leaved – pouze jehličnaté dřeviny
- 22 Dominantly needle leaved – převážně jehličnaté dřeviny
- 31 Pure broadleaved deciduous – pouze listnaté opadavé dřeviny
- 32 Pure broadleaved evergreen – pouze listnaté stálezelené dřeviny
- 33 Pure broadleaved – převážně listnaté dřeviny
- 40 Shrubland – keřové porosty

- 51 Permanent herbaceous without trees – trvalé bylinné porosty bez stromů
- 52 Permanent herbaceous with few trees – trvalé bylinné porosty s malým zastoupením stromů
- 53 Permanent herbaceous with many trees – trvalé bylinné porosty s velkým zastoupením stromů
- 60 Periodically herbaceous –sezónní bylinné porosty
- 70 Lichens and mosses – lišejníky a mechy
- 81 Partly vegetated land: low vegetation cover – plochy částečně pokryté vegetací s nízkým podílem vegetace
- 82 Partly vegetated land: intermediate vegetation cover – plochy částečně pokryté vegetací se středním podílem vegetace
- 90 Non-vegetated land – plochy bez vegetace
- 100 Water –voda

(Zdroj: EEA 2024, vlastní překlad)

Příloha 2: Srovnání land cover/land use databází v řešených katastrálních územích

F = Forest land

C = Cropland

G = Grassland

S = Settlements

W = Wetlands

O = Other land

České Vrbné

	LULUCF	ESA	Dynamic	ESRI	CLC+	KN
F	172,85	210,66	222,76	171,01	181,66	89,27
C	78,47	71,31	96,23	131,18	68,36	91,10
G	89,18	117,83	28,52	37,23	133,58	108,88
S	99,86	53,08	84,89	91,54	64,31	119,20
W	109,38	98,91	118,80	121,29	103,33	143,85
O	2,57	0,28	1,01	0,00	0,87	

Klec

	LULUCF	ESA	Dynamic	ESRI	CLC+	KN
F	195,54	239,44	230,52	179,89	204,83	98,08
C	19,80	8,46	3,53	15,18	19,59	27,88
G	88,65	115,31	76,46	93,55	123,19	116,69
S	34,69	7,26	25,95	38,85	12,19	52,19
W	288,25	261,49	295,32	304,55	270,59	337,19
O	5,11	0,00	0,36	0,00	1,50	0,00

Chvaletice

	LULUCF	ESA	Dynamic	ESRI	CLC+	KN
F	288,46	325,09	-	277,68	265,64	138,26
C	134,69	133,00	-	177,13	129,76	154,66
G	22,65	93,50	-	69,37	136,52	13,85
S	228,07	73,28	-	129,83	67,20	351,77

W	6,63	9,56	-	11,24	7,64	4,38
O	0,83	0,00	-	15,85	74,23	0,00

Prachovice

	LULUCF	ESA	Dynamic	ESRI	CLC+	KN
F	184,50	237,95	218,40	196,43	185,60	153,57
C	51,24	49,01	87,21	136,15	43,91	69,22
G	115,29	136,33	69,61	38,02	186,23	144,57
S	179,39	104,13	112,86	122,49	73,61	162,79
W	1,16	2,33	24,78	4,43	1,01	1,67
O	0,25	2,02	18,91	34,17	41,18	0,00

Hojovice

	LULUCF	ESA	Dynamic	ESRI	CLC+	KN
F	251,57	267,35	333,67	235,39	232,73	227,50
C	171,77	160,77	154,20	314,54	154,06	188,18
G	119,51	142,95	48,40	1,90	179,02	128,57
S	29,52	5,00	17,23	24,17	9,81	28,41
W	2,03	0,54	16,85	0,67	0,72	3,93
O	0,02	0,00	6,12	0,00	0,07	0,00

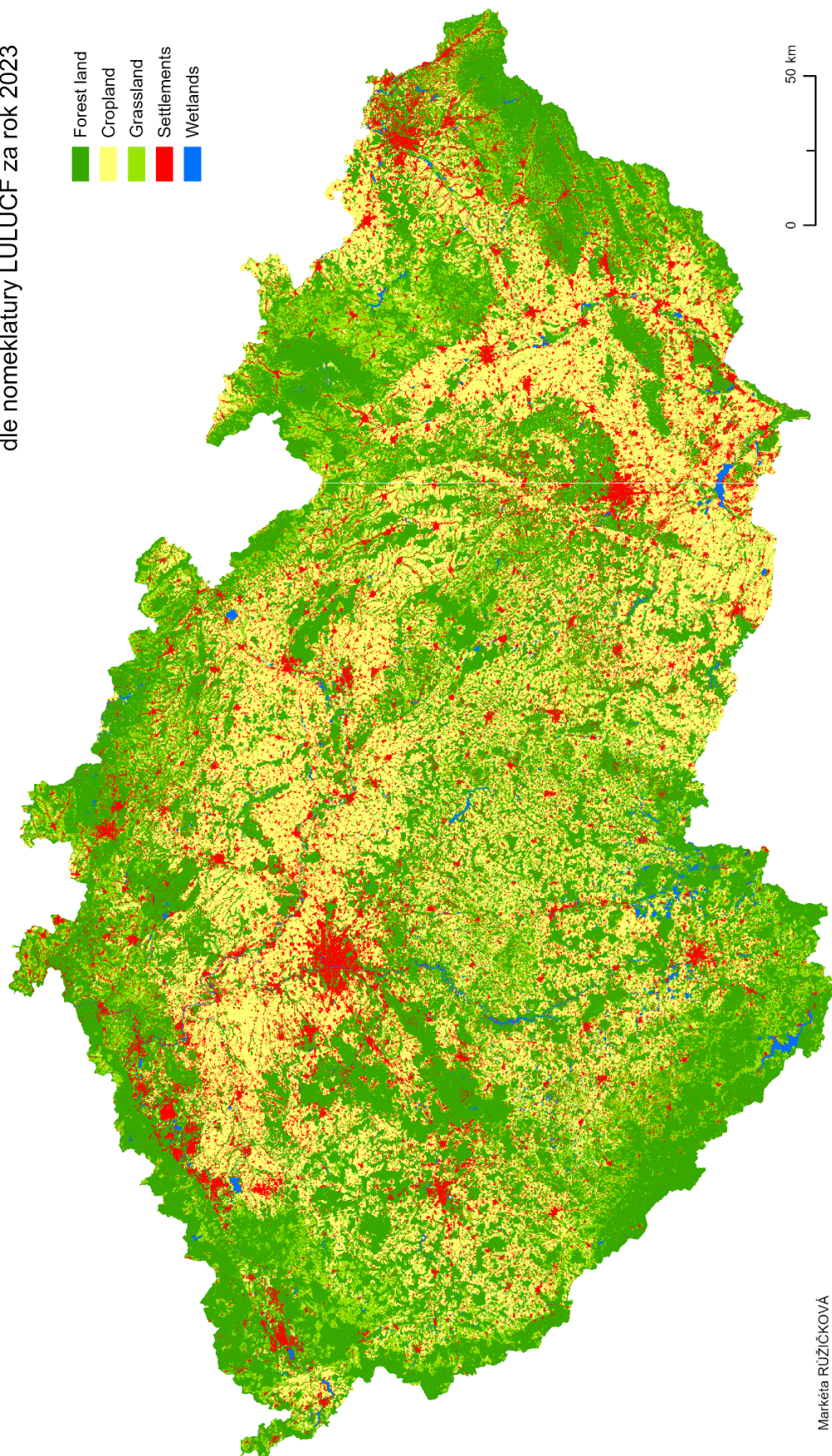
Škrdlovice

	LULUCF	ESA	Dynamic	ESRI	CLC+	KN
F	173,87	180,69	259,68	155,19	151,27	143,41
C	206,69	226,23	199,69	339,34	191,92	224,23
G	125,69	146,60	54,24	21,02	202,81	144,21
S	66,06	20,93	55,63	59,05	28,64	60,32
W	1,65	1,13	1,05	1,18	0,79	3,66
O	1,85	0,00	5,46	0,00	0,22	0,00

Příloha 3: Klasifikace LULUCF pro území Česka

Využití krajiny na území Česka

dle nomenklatury LULUCF za rok 2023



Markéta RŮŽIČKOVÁ
Praha 2025