

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

FAKULTA STAVEBNÍ

STUDIJNÍ PROGRAM GEODÉZIE A KARTOGRAFIE

STUDIJNÍ OBOR GEOMATIKA



DIPLOMOVÁ PRÁCE

TESTOVÁNÍ ZPŮSOBŮ KLASIFIKACE POKRYTÍ ÚZEMÍ
VYBRANÝCH EVROPSKÝCH OBLASTÍ

Vedoucí práce: prof. Ing. Lena Halounová, CSc.

Katedra geomatiky

2020

Bc. Tomáš Bouček

ZADÁNÍ DIPLOMOVÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: <u>Bc. Bouček</u>	Jméno: <u>Tomáš</u>	Osobní číslo: <u>460381</u>
Zadávající katedra: <u>katedra geomatiky</u>		
Studijní program: <u>Geodézie a kartografie</u>		
Studijní obor: <u>geomatika</u>		

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce: <u>Testování způsobů klasifikace pokrytí území vybraných evropských oblastí</u>	
Název diplomové práce anglicky: <u>Testing of land cover classification types of selected European regions</u>	
Pokyny pro vypracování: Zpracujte klasifikaci pokrytí území pro různou úroveň tématické podrobnosti. Využijte data z družic Sentinel 2A, 2B. Vyhodnoťte získané přesnosti z jednotlivých oblastí na vybraném území Švédska, Itálie a České republiky. Doporučte, jaká metoda klasifikace bude nejvhodnější i pro další evropská území.	
Seznam doporučené literatury: Halounová L., Pavelka K.: Dálkový průzkum Země. Vydavatelství ČVUT, Praha 2007. ISBN: 80-01-03124-1 Halounová, L.: Zpracování obrazových dat. ČVUT v Praze, 2008. ISBN: 978-80-01-04253-3 Active Learning Methods for Remote Sensing Image Classification https://ieeexplore.ieee.org/abstract/document/4812037 https://pythontips.com/2017/11/11/introduction-to-machine-learning-and-its-usage-in-remote-sensing/	
Jméno vedoucího diplomové práce: <u>prof. Ing. Lena Halounová, CSc.</u>	
Datum zadání diplomové práce: <u>15. 2. 2020</u>	Termín odevzdání diplomové práce: <u>23. 5. 2020</u> <i>Údaj uveďte v souladu s datem v časovém plánu příslušného ak. roku</i>
Podpis vedoucího práce	Podpis vedoucího katedry

III. PŘEVZETÍ ZADÁNÍ

<i>Beru na vědomí, že jsem povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je nutné uvést v diplomové práci a při citování postupovat v souladu s metodickou příručkou ČVUT „Jak psát vysokoškolské závěrečné práce“ a metodickým pokynem ČVUT „O dodržování etických principů při přípravě vysokoškolských závěrečných prací“.</i>	
Datum převzetí zadání	Podpis studenta(ky)

Čestné prohlášení

Prohlašuji, že jsem diplomovou práci na téma „Testování způsobů klasifikace pokrytí území vybraných evropských oblastí“ vypracoval samostatně za použití uvedené literatury a zdrojů.

Dále prohlašuji, že nemám závažný důvod proti užití tohoto školního díla ve smyslu § 60 zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne

.....

Bc. Tomáš Bouček

Poděkování

Děkuji paní prof. Ing. Leně Halounové, CSc., za odborné vedení této diplomové práce. Dále děkuji panu Ing. Lukáši Brodskému, Ph.D., za užitečné rady a konzultace, díky nimž mohla tato práce vzniknout a panu Ing. Martinu Landovi, Ph.D., za pomoc při řešení technických problémů v programech QGIS a GRASS.

Anotace

Tato diplomová práce je zaměřena na klasifikaci druhů pokrytí povrchu Země (tzv. land cover) pomocí družicových dat. Cílem je získat co nejlepší výsledky pro různé úrovně tematické podrobnosti a vyhodnotit přesnosti, s jakými byla klasifikace provedena.

Klíčová slova

Dálkový průzkum Země, řízená klasifikace, Sentinel-2, land cover, CORINE, QGIS, GRASS, strojové učení, kontingenční tabulka

Abstract

This master's thesis is focused on classification of types of the Earth surface coverage (a.k.a land cover) using satellite data. The goal is to get the best results for different levels of thematic details and evaluate accuracy with which the classification was performed.

Key words

Remote sensing, supervised classification, Sentinel-2, land cover, CORINE, QGIS, GRASS, machine learning, contingency table

Obsah

Seznam zkratk.....	8
1 Úvod	9
2 Zpracovávaná území.....	10
3 Použitá data	11
3.1 Sentinel-2	11
3.1.1 Pásma snímaná družicemi Sentinel-2.....	12
3.1.2 Produkty Sentinel-2.....	14
3.1.3 Použité scény	14
3.2 CORINE Land Cover.....	19
4 Použitý software	21
4.1 QGIS	21
4.2 GRASS GIS	22
5 Klasifikace.....	23
5.1 Řízená klasifikace	24
5.1.1 Klasifikátor maximum likelihood.....	27
5.1.2 Strojové učení (<i>machine learning</i>)	28
5.1.2.1 Random Forest klasifikátor.....	28
6 Existující land cover produkty.....	29
7 Příprava trénovacích dat a klasifikace	30
7.1 Klasifikace první úrovně CLC	30
7.1.1 Úprava dat první úrovně CLC	34
7.2 Klasifikace druhé úrovně CLC.....	46
7.3 Klasifikace třetí úrovně CLC	56
8 Ukázky výsledků klasifikace	62
9 Klasifikace pomocí strojového učení	66
9.1 Porovnání s maximum likelihood	72

10 Závěr.....	82
Reference.....	84
Seznam obrázků	86
Seznam tabulek.....	88
Seznam příloh.....	90

Seznam zkratek

CLC	CORINE Land Cover
CORINE	Coordination of Information on the Environment
ČVUT	České vysoké učení technické v Praze
DLR	German Aerospace Center
DPZ	Dálkový průzkum Země
EEA	Evropská agentura pro životní prostředí (European Environment Agency)
ESA	Evropská kosmická agentura (European Space Agency)
EU	Evropská unie (European Union)
FOSS4G	Free and Open Source Software for Geospatial
FSv	Fakulta stavební
GIS	Geografický informační systém
GMES	Global Monitoring for Environment and Security
GPL	General Public License
GRASS	Geographic Resources Analysis Support System
ISP	Instrument Source Packet
JRC	Joint Research Centre
LTR	long term release
MMU	Minimální mapovací jednotka (Minimal Mapping Unit)
MSI	MultiSpectral Instrument
NASA	Národní úřad pro letectví a kosmonautiku (National Aeronautics and Space Administration)
NDVI	Normovaný rozdílový vegetační index (Normalized difference vegetation index)
NGCC	National Geomatics Center of China
NOAA	Národní úřad pro oceán a atmosféru (National Oceanic and Atmospheric Administration)
OSM	OpenStreetMap
RGB	Červená-Zelená-Modrá (Red-Green-Blue)
S2GLC	Sentinel-2 Global Land Cover
SWIR	Krátkovlnné infračervené záření (Short-wavelength infrared)
UTM	Univerzální transversální Mercatorův systém souřadnic (Universal Transverse Mercator)
VNIR	Viditelné a blízké infračervené záření (Visible and near-infrared)
WGS84	Světový geodetický systém 1984 (World Geodetic System 1984)

1 Úvod

Tato diplomová práce byla zpracovávána v rámci projektu „*Geo-harmonizer: EU-wide automated mapping system for harmonization of Open Data based on FOSS4G and Machine Learning*“, jehož cílem je vyvinout původní, webově orientovaný, modulární systém („Geo-harmonizer“) určený pro uchovávání a sdílení geografických dat s tematickým obsahem (land cover, kvalita životního prostředí apod.) pokrývající území kontinentální Evropy. Na projektu spolupracuje více zemí Evropské unie. Českou republiku v projektu zastupuje ČVUT, které je i koordinátorem celého projektu [1].

Práce se zabývá řízenou klasifikací obrazových dat pořízených metodou dálkového průzkumu Země. Konkrétně se jedná o družicová multispektrální data získaná na principu odrazu elektromagnetického záření. Klasifikace byla provedena pro různé úrovně tematické podrobnosti, kde byla sledována přesnost, na kterou se lze v jednotlivých úrovních dostat.

Při práci byly použity dva způsoby klasifikace. Jedním z nich je způsob využívající metodu *maximum likelihood*, druhým způsobem je metoda založená na tzv. strojovém učení (*machine learning*). Samotné klasifikaci předchází příprava a úprava trénovacích ploch. Trénovací plochy nebyly zadávány interaktivně výběrem oblastí na kompozitním snímku, což může vést k nesprávnému určení jednotlivých tříd. Místo toho bylo využito volně dostupných dat (CORINE) poskytujících informace o druhu pokrytí Země (tzv. *land cover*).

Klasifikována byla tři různá evropská území reprezentující rozdílné klimatické podmínky v Evropě. Jedná se o území ve Švédsku, v České republice a v Itálii, která nejlépe vystihují různorodost evropského kontinentu.

Cílem této práce je tedy klasifikovat tyto tři vybraná území, upravit data CORINE tak, aby mohla sloužit jako trénovací plochy a zjistit, které třídy z nomenklatury CORINE je možné na základě spektrálních příznaků dostatečně přesně klasifikovat.

2 Zpracovávaná území

Jelikož se Evropa skládá z více podnebných pásů, bylo i pro klasifikaci vybráno více území zastupující jednotlivé přírodní i podnebné podmínky v Evropě. Celkově byla vybrána tři území. Jedním z nich je území okolo města Kristianstad nacházejícího se ve Švédsku a reprezentujícího oblasti severní Evropy s velkým počtem jezer a lesů. Dále je to území v České republice okolo Mladé Boleslavi reprezentující středoevropské oblasti. A třetím územím je oblast okolo italského Říma reprezentující jižní Evropu a její subtropické podnebí.



Obr. 1: Poloha zpracovávaných lokalit

3 Použitá data

Celý projekt je koncipován tak, aby se, pokud možno, vycházelo pouze z volně dostupných dat (tzv. *open data*, případně *open source data*). Jako zdroje potřebné ke zpracování této práce byla tedy použita také pouze volně dostupná data.

Pro obrazová data byly použity scény z družic Sentinel-2, které lze po vytvoření účtu stáhnout ze stránek Evropské kosmické agentury (ESA) pomocí webové aplikace Copernicus Open Access Hub dostupné na internetové adrese: <https://scihub.copernicus.eu/dhus/#/home>. Tato aplikace spadá pod program Copernicus, což je program Evropské unie pro pozorování Země a poskytuje informační služby založené nejen na družicovém pozorování Země, ale i na dalších (jiných než vesmírných) datech [2].

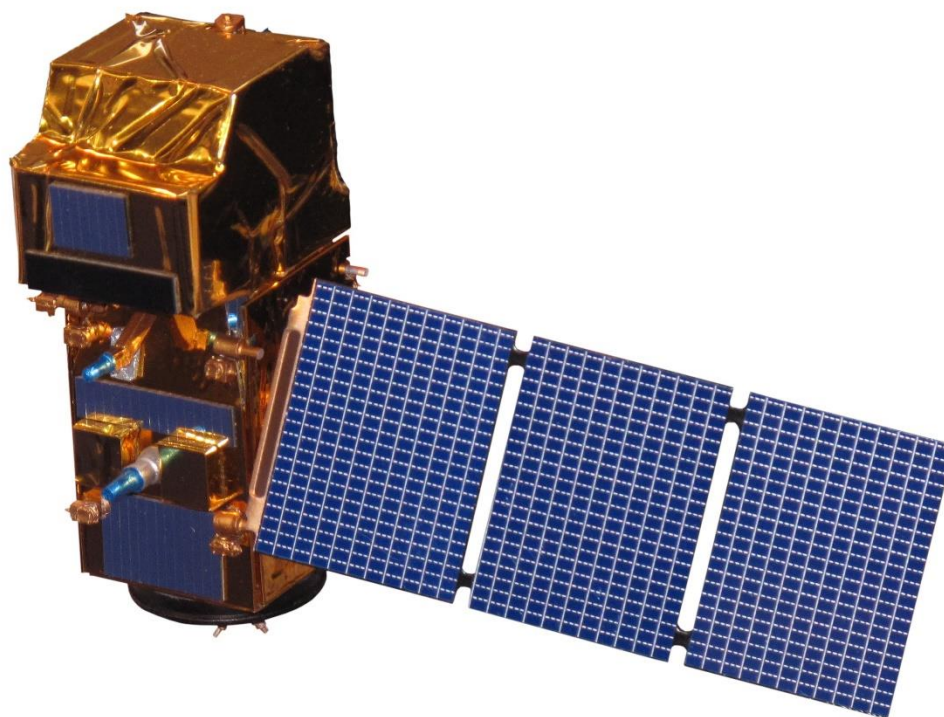
Jako základ pro trénovací plochy bylo využito dat z programu CORINE, konkrétně dat CLC. Jedná se o vektorovou či rastrovou vrstvu pokrývající většinu Evropy a poskytující informaci o druhu pokrytí povrchu Země. Data jsou po zaregistrování volně stažitelná z internetové stránky: <https://land.copernicus.eu/pan-european/corine-land-cover/clc2018?tab=download>. Program CORINE rovněž spadá pod program Copernicus.

3.1 Sentinel-2

Sentinel-2 je multispektrální operační obrazová mise v rámci projektu Copernicus (dříve známého jako GMES) společně prováděného Evropskou komisí a Evropskou vesmírnou agenturou. Projekt vznikl za účelem globálního pozorování povrchu Země ve vysokém rozlišení a s možností častého opakování pro zajištění kontinuity dat.

Mise Sentinel-2 využívá dvou družic: Sentinel-2A a Sentinel-2B, které byly vypuštěny na oběžnou dráhu Země v červnu 2015, resp. v březnu 2017 z Francouzské Guayany. Obě družice obíhají Zemi na polárních drahách ve výšce 786 km nad zemským povrchem a jsou umístěny na stejném heliosynchronním orbitu, což znamená, že družice prochází nad určeným místem povrchu Země vždy ve stejném slunečním čase. Dráhy obou družic jsou od sebe navzájem odfázovány o 180°. Měření odrazivosti zemského povrchu pokrývá území od 56° jižní šířky až po 84° severní šířky. Družice disponují širokou šířkou skenovaného řádku (290 km)

a poměrně krátkým časem opakování, který činí 10 dní na rovníku a pouhých 5 dní při započtení obou družic. Životnost družic byla navržena na dobu sedmi let, ale fungovat by mohly až dvanáct roků [3][4].



Obr. 2: Model družice Sentinel-2, Zdroj: Wikipedia

3.1.1 Pásma snímaná družicemi Sentinel-2

Obě družice Sentinel-2 nesou optický senzor MSI snímající ve třinácti pásmech od viditelného po střední infračervené pásmo. Z toho mají čtyři kanály prostorové rozlišení 10 m, šest kanálů prostorové rozlišení 20 m a tři kanály prostorové rozlišení 60 m. Senzor využívá principu *push-broom*, kdy řádky obrazu sbírají ve stejném okamžiku celou šířku skenovaného pásu a k poskytnutí nových řádků ke snímání je využito dopředného pohybu družice. Záření odražené ze Země a její atmosféry do senzoru MSI je zachycováno pomocí trojzrcadlového dalekohledu a zaostřeno rozdělovačem paprsků do dvou separátních ohniskových sestav. Jedna obsahuje deset pásem viditelného a blízkého infračerveného záření (VNIR), druhá pak tři pásma středního infračerveného záření (SWIR). Družice

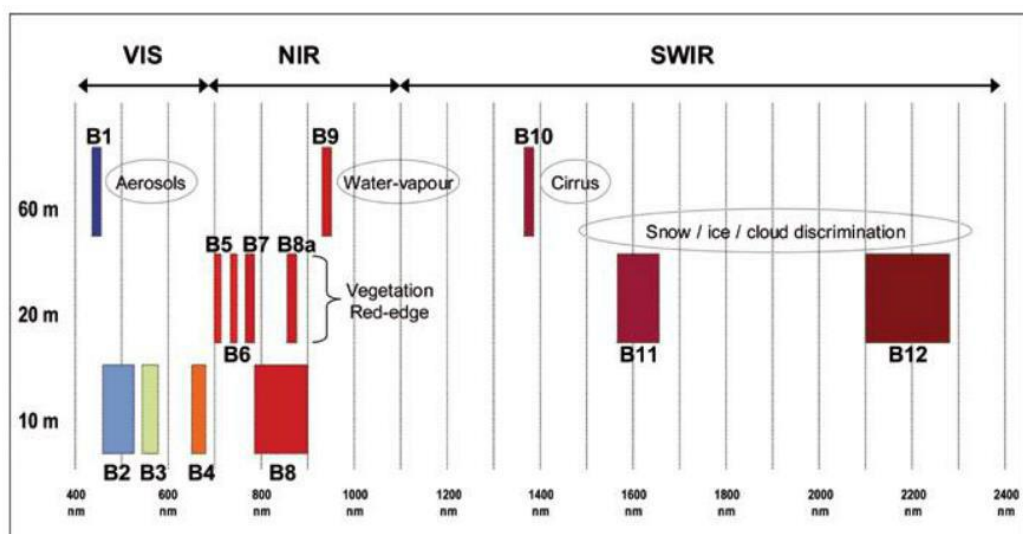
Sentinel-2 pracují s radiometrickým rozlišením 12 bitů, což umožňuje zachytit a rozlišit až 4095 různých hodnot intenzity záření [5][6].

Veškerá spektrální pásma snímaná družicemi Sentinel-2 a informace o nich jsou shrnuta v následující tabulce.

Tab. 1: Pásma Sentinel-2, zdroj: Wikipedia, [5]

Pásma	Popis pásma	Střední vlnová délka [nm]	Šířka pásma [nm]	Prostorové rozlišení [m]
B1	Ultra blue (Coastal aerosol)	443	20	60
B2	Blue	490	65	10
B3	Green	560	35	10
B4	Red	665	30	10
B5	Vegetation red edge	705	15	20
B6	Vegetation red edge	740	15	20
B7	Vegetation red edge	783	20	20
B8	NIR	842	115	10
B8a	Narrow NIR	865	20	20
B9	Water vapour	945	20	60
B10	SWIR – Cirrus	1380	30	60
B11	SWIR	1610	90	20
B12	SWIR	2190	180	20

Přehledné zobrazení jednotlivých pásem rozdělených podle prostorového rozlišení a podle druhu snímaného elektromagnetického záření ukazuje následující obrázek.



Obr. 3: Pásma Sentinel-2, zdroj ESA, [5]

3.1.2 Produkty Sentinel-2

Data z družic Sentinel-2 se zpracovávají do celkem pěti úrovní, z nichž tři jsou dostupné běžným uživatelům. Scény jsou poskytovány ve formě tzv. granulí (dlaždic), což jsou nejmenší části scén obsahující všechna spektrální pásma. Produkty jsou tedy kompilací jednotlivých granulí o pevné velikosti pořízených podél jedné oběžné dráhy.

Prvním produktem je produkt úrovně Level-0, který pro běžné uživatele není dostupný. Přestavuje komprimovaná nepředzpracovaná data ve formátu ISP a tvoří základ pro následné vytvoření produktů úrovně Level-1. Rozměry jednotlivých granulí jsou $25 \text{ km} \times 23 \text{ km}$. Dalším produktem je produkt úrovně Level-1A, který rovněž není pro běžné uživatele dostupný a je získán dekomprimací neupravených obrazových dat úrovně Level-0. Pro data této úrovně je vytvořen geometrický model umožňující lokalizaci každého pixelu v obraze. Rozměrově odpovídá produktu úrovně Level-0.

Nejnižším zpracovaným produktem dostupným pro běžné uživatele je produkt úrovně Level-1B. Rozměry rovněž odpovídají předchozímu produktu. Rozdílem ale je, že již obsahují atmosférické a geometrické korekce. Prvním volně dostupným produktem je produkt úrovně Level-1C. Rozměry dlaždic jsou $100 \text{ km} \times 100 \text{ km}$ a jedná se již o ortorektifikovaná obrazová data v projekci UTM / WGS84. Produkty z úrovně Level-1C jsou převzorkovány na prostorové rozlišení 10 m, 20 m a 60 m v závislosti na prostorovém rozlišení jednotlivých spektrálních kanálů.

Posledním produktem je produkt úrovně Level-2A rovněž volně dostupný běžným uživatelům. Vychází z produktů úrovně Level-1C, na rozdíl od této úrovně ale obsahuje i atmosférické korekce. Jednotlivé dlaždice mají rozměry stejné jako předchozí úroveň, tedy $100 \text{ km} \times 100 \text{ km}$, a jsou také v projekci UTM / WGS84. Obrazová data jsou uložena ve formátu JPEG 2000 (JP2). V této diplomové práci byly využity právě produkty této úrovně [5].

3.1.3 Použité scény

V práci byly použity pro každé z vybraných území dvě totožné scény, ale z jiného časového období. Jedná se o jarní a letní scénu z roku 2018 (důvod, proč byly použity scény právě z tohoto roku, je zmíněn v následující podkapitole) a použita byla všechna pásma s rozlišením 20 m. Takových pásem je pouze šest,

nicméně stažená data obsahují i převzorkovaná desetimetrová data (kromě pásma B8, které bylo převzorkováno ručně), takže celkový počet pásem použitých ke klasifikaci je deset pro jednu scénu (B2, B3, B4, B5, B6, B7, B8, B8a, B11 a B12). Scény byly vybírány tak, aby byly pokud možno bezoblačné, případně s maximální hodnotou oblačnosti 5 %. Jednotlivé dlaždice jsou popsány v následující tabulce obsahující název a datum pořízení. Ukázky scén na RGB kompozitních obrazových datech pro jednotlivé vybrané dlaždice jsou na obr. 4, 5 a 6.

Tab. 2: Použité scény

Území	Dlaždice	Datum pořízení	
		Jarní scéna	Letní scéna
Kristianstad	T33VVC	7. 5. 2018	26. 7. 2018
Mladá Boleslav	T33UVR	19. 4. 2018	3. 7. 2018
Řím	T33TTG	21. 4. 2018	15. 7. 2018



Obr. 4: Jarní scéna oblasti okolo města Kristianstad

Scéna je tvořena převážně pevninskou částí a největší rozlohu v ní zabírají lesní porosty. Druhou největší kategorií, co se rozlohy týče, jsou vodní plochy. To je dáno především značnou částí mořské hladiny, která do scény zasahuje. Nemalý podíl na vodní ploše, ale mají i velká jezera, která se vyskytují v téměř každé oblasti vybrané scény. Dále pak převážně v jižní části scény jsou patrné zemědělské oblasti, a malou měrou jsou ve scéně zastoupeny i oblasti zastavěné člověkem.



Obr. 5: Jarní scéna z České republiky

Původně bylo zamýšleno mít tuto scénu danou tak, aby přibližně ve středu byla umístěna Mladá Boleslav. Ta se však nachází na rozmezí jednotlivých dlaždic, takže nakonec byla vybrána dlaždice s Prahou uprostřed. Mladá Boleslav se v této scéně nachází v pravém horním rohu.



Obr. 6: Jarní scéna oblasti z okolí Říma

Scéna s okolím Říma není situována příliš výhodně, je prakticky z poloviny tvořena mořem. Nicméně samotný Řím jsme ve scéně chtěli ponechat, takže byla ponechána i samotná scéna. Důležité je si zde uvědomit, že se jedná rozměrově o poměrně velkou scénu, a to, že polovinu z toho zabírá voda, ve výsledku nehraje žádnou nebo jen malou roli.

3.2 CORINE Land Cover

Program CORINE byl zahájen Evropskou komisí v roce 1985. Jeho cílem je zajistit sběr, koordinaci a přístup ke kvalitním informacím o životním prostředí a přírodních zdrojích v rámci Evropy. Program je rozdělen do několika částí. Jsou jimi Land Cover, Biotopes a Air.

CORINE Land Cover (CLC) je vrstva pokrývající celou Evropu. První vrstva vznikla roku 1990. Dalšími referenčními roky byly roky 2000, 2006, 2012 a zatím naposledy rok 2018. Data z roku 2018 byla využita pro tuto práci, a proto i data z družic Sentinel-2 byla vzata z tohoto roku. Databáze CORINE Land Cover je vytvářena jako celoevropská vrstva a tomu odpovídá i měřítko 1 : 100 000 s nejmenší mapovací jednotkou 25 ha a s minimální šířkou 100 m pro liniové prvky. Na vrstvě z roku 2018 spolupracovalo celkem 39 zemí a data jsou volně dostupná všem uživatelům. Tematická přesnost této vrstvy je udávána jako větší nebo rovna 85 %.

Vrstva CLC je rozdělena do tří úrovní. První vrstvu tvoří pět základních tříd, které se dále dělí do tříd druhé a třetí úrovně. Druhá úroveň je tvořena patnácti třídami a třetí úroveň tvoří celkově 44 třídy. Přehled jednotlivých tříd znázorňuje následující obrázek [7][8].



Obr. 7: Třídy CLC [9]

4 Použitý software

Jak již bylo řečeno, celý projekt je založen na tom, že se budou používat pouze volně dostupná data, a to samé platí i pro použitý software. Proto se i autor této práce snažil pracovat s volně dostupným (open source) softwarem. Pro přípravu trénovacích ploch byl použit program QGIS a pro následnou klasifikaci program GRASS GIS. Výjimku tvoří komerční program PCI Geomatica, ve kterém probíhaly prvotní klasifikační experimenty.

4.1 QGIS

QGIS je multiplatformní geografický informační systém vyvíjen jako open source a publikovaný pod všeobecnou licencí GNU GPL. Jedná se o nejrozšířenější volně dostupný GIS. Projekt QGIS vznikl v roce 2002 a první verze vyšla v roce 2009. Původní název byl Quantum GIS, který byl ale v roce 2013 změněn na QGIS. Mezi hlavní výhody patří rychlost vývoje a rozšiřování funkcionality. Současným konceptem ve vývoji je pravidelné publikování nových verzí, kdy dlouhodobá stabilní verze (LTR) je doplněna dvěma krátkodobými verzemi sloužícími ke zveřejňování nových funkcionalit v kratších intervalech [10][11].

QGIS disponuje velkým množstvím nástrojů. Kromě svých nativních nástrojů využívá i nástroje z knihovny GDAL a dále pak nástroje softwarů GRASS či SAGA. Důležitou součástí programu QGIS tvoří také tzv. zásuvné moduly (*plugins*), které jsou vyvíjeny nezávislými organizacemi či vývojáři a rozšiřují tak jeho funkcionalitu a možnosti využití. Důležitým zásuvným modulem je např. GRASS plugin, který umožňuje používat právě nástroje z programu GRASS GIS v prostředí QGIS. Při práci byla použita verze 3.12.2-București.



Obr. 8: QGIS logo, zdroj: Wikipedia

4.2 GRASS GIS

GRASS (Geographic Resources Analysis Support System) je multiplatformní desktopový geografický informační systém. Stejně jako QGIS je i GRASS open source software publikovaný pod všeobecnou licenci GNU GPL. Ve světě Open Source GIS se jedná o nejkompexnější nástroj s velkou škálou funkcionality, která obtojí i při nejnáročnějších GIS analýzách. GRASS GIS disponuje počtem více než 350 různých modulů s možností doinstalovat spoustu dalších tzv. Add-ons.

Původně byl vyvíjen americkou armádou (konkrétně U.S. Army Construction Engineering Research Laboratories) jako vojenský nástroj pro správu půdy a plánování životního prostředí. Pod vojenskou správu spadal v letech 1982 až 1995. Roku 1997 byl pak převzat pod správu akademickou. V současnosti je GRASS využíván mnoha agenturami, mezi něž patří například NASA nebo NOAA. Na vývoji a udržování programu GRASS se značnou mírou podílí i České vysoké učení technické v Praze a jeho absolventi. Při práci byla použita verze 7.8.2 (později pak verze 7.8.3) [12][13].



Obr. 9: GRASS GIS logo, zdroj: Wikipedia

5 Klasifikace

Klasifikací se rozumí proces, při kterém je získán obrazový soubor, jehož digitální hodnoty pixelů udávají jejich zařazení do konkrétních tříd. Vzniká tak mapa, která jednoznačně dává informaci o druhu jednotlivých ploch vymezených pixely se stejnou digitální hodnotou. V oboru DPZ se nejčastěji používá klasifikace pixel po pixelu (*per pixel*), možná je ale i klasifikace po tzv. segmentech, které jsou tvořeny více pixely dohromady. Tento způsob se využívá především u dat velmi vysokého rozlišení, kdy je jeden objekt tvořen velmi často právě více pixely.

Pro provedení klasifikace je zapotřebí znát tzv. klasifikační pravidlo neboli nalézt vztah, který příslušné hodnotě parametrů (např. druhy povrchu Země) jednoznačně přiřadí digitální hodnoty naměřených obrazových dat, a tím je i zařadí do jednotlivých tříd. Při hledání klasifikačních pravidel se vychází z toho, jakým způsobem se mění radiometrická veličina v závislosti na změně daných parametrů. Pokud jsou tyto změny význačné a charakteristické, že je lze použít k sestavení klasifikačního pravidla, nazývají se tzv. příznaky. Existují různé druhy příznaků (prostorové, spektrální, časové či polarizační), v této práci ale bylo využito pouze příznaků spektrálních, které vyjadřují odrazivé a vyzařovací vlastnosti zkoumaného povrchu.

Klasifikační pravidlo ale ve většině úloh není předem známo a je potřeba ho vytvořit. K tomu je možno zaujmout dva postoje. Prvním z nich je tzv. neřízená klasifikace (*unsupervised classification*), při které jsou automaticky vytvářeny shluky na základě vzájemné podobnosti radiometrické veličiny, a ty jsou následně uspořádány do tříd. Úkolem zpracovatele je pak zjistit, co jednotlivé třídy představují. Druhým způsobem je klasifikace řízená (*supervised classification*), kdy je předem známo, do jakých tříd bude snaha obrazový soubor klasifikovat. Klasifikační pravidlo se zde vytváří na tzv. trénovacích plochách, které jsou zadány zpracovatelem. Oba tyto způsoby je možné i kombinovat, čímž pak vzniká klasifikace hybridní, kdy shluky, vytvořené neřízenou klasifikací a spojeny do jednotlivých tříd, jsou následně použity jako trénovací množiny pro klasifikaci řízenou. V práci byla použita pouze řízená klasifikace [14].

5.1 Řízená klasifikace

Řízenou klasifikaci obrazového souboru můžeme rozdělit do čtyř částí. První z nich je výběr trénovacích ploch. Lze tak učinit buď interaktivně výběrem ploch na kompozitním snímku, anebo je mít předem připravené z jiných dostupných zdrojů. Vhodně zvolené trénovací plochy jsou základem pro správnou klasifikaci, a je tedy důležité, aby měl zpracovatel při interaktivním zadávání trénovacích ploch potřebnou znalost zpracovávaného území. Trénovací plochy se volí na základě zvolených tříd a pro každou třídu je potřeba mít dostatečné množství těchto ploch, které jednotlivé třídy reprezentují. Při výběru tříd je lepší vybírat menší plochy z různých míst než vzít pouze jednu velkou plochu, která by ale nemusela reprezentovat celou scénu. V druhém kroku se pak z těchto trénovacích ploch vytvoří soubor příznaků, ve kterém jsou uloženy spektrální informace trénovacích ploch.

Třetím krokem je spuštění samotné klasifikace, k čemuž je potřeba zvolit vhodný klasifikátor. Klasifikátory jsou založeny na různých způsobech rozpoznávání, jako je statistické rozpoznávání vzoru (např. klasifikátor maximální pravděpodobnosti, Bayesovo kritérium, klasifikátor minimální vzdálenosti, klasifikátor rovnoběžnostěnů), neuronové sítě, syntaktické rozpoznávání vzoru, rozpoznávání na základě grafické shody, fuzzy klasifikace či za použití umělé inteligence a strojového učení. V práci byl použit statistický klasifikátor maximální pravděpodobnosti (*maximum likelihood*) a klasifikátor založený na strojovém učení.

Posledním krokem řízené klasifikace je vyhodnocení jejího výsledku. Představu o přesnosti klasifikace může dát tzv. kontingenční tabulka, která ukazuje, jak se pixely z trénovacích ploch zařadily do jednotlivých tříd. Pokud se pixely trénovacích ploch jedné třídy zařadily ve větším počtu do třídy jiné, indikuje to, že trénovací množiny nebyly vhodně zvoleny a jednotlivé třídy se vzájemně překrývají. Je tedy potřeba se vrátit k prvnímu kroku a upravit trénovací množiny. Kontrolu tříd je možné provést i pomocí histogramů jednotlivých tříd v příslušném pásmu či pomocí tzv. rozptylogramu (*scatter plot*), což je dvourozměrný graf, kde jsou na každou osu vynášeny hodnoty vybraných trénovacích ploch tříd z jednoho pásma (různé pásmo na ose x a y). Pokud se shluky pixelů z jednotlivých tříd překrývají, svědčí to o špatné separabilitě těchto tříd. Výsledek, tedy přesnost klasifikace, je pak nutno zkontrolovat pomocí tzv. testovacích ploch, které nejsou totožné

s plochami trénovacími. Porovnává se zde výsledek klasifikace s těmito testovacími plochami a kontroluje se, zda se hodnoty pixelů shodují [14][15].

Přesnost klasifikace lze posoudit i pomocí několika kritérií. Jedním z nich je celková přesnost udávající poměr všech správně klasifikovaných pixelů a všech pixelů z trénovacích a poté testovacích ploch. Vzorec pro celkovou přesnost klasifikace má tvar

$$CP = \frac{1}{n} \sum_{k=1}^N a_{kk} \quad , \quad (1)$$

kde

n je celkový počet pixelů použitých pro trénovací nebo testovací plochy,

N je celkový počet klasifikačních tříd,

a_{kk} jsou diagonální prvky kontingenční tabulky udávající počet správně zařazených pixelů z trénovacích a poté i testovacích ploch do jednotlivých tříd.

Další možností posouzení přesnosti je průměrná přesnost, která pouze zprůměruje počet správně klasifikovaných pixelů jednotlivých tříd. Vzorec má tedy podobu

$$PP = \frac{1}{N} \sum_{k=1}^N a_{kk} \quad . \quad (2)$$

Tyto hodnoty uvádí pouze přesnost zařazení všech pixelů do zvolených tříd. Přesnost jednotlivých tříd lze posoudit např. tzv. tvůrčí přesností

$$TP_i = \frac{a_{ii}}{\sum_{k=1}^N a_{ki}} \quad , \quad (3)$$

kde

a_{ii} je počet správně klasifikovaných pixelů dané třídy,

$\sum_{k=1}^N a_{ki}$ je celkový počet pixelů dané třídy použitých pro trénovací nebo testovací plochy.

Tyto přesnosti jednotlivých tříd jsou uváděny právě v kontingenční tabulce. Pokud by nás ale zajímala pravděpodobnost, že klasifikovaný pixel dané třídy opravdu patří do této třídy, tzn. že se jedná o pixel z trénovacích nebo testovacích ploch pro danou třídu, můžeme použít přesnost uživatelskou

$$UP_i = \frac{a_{ii}}{\sum_{k=1}^N a_{ik}} , \quad (4)$$

kde

$\sum_{k=1}^N a_{ik}$ je celkový počet pixelů klasifikovaných do dané třídy.

Trochu jinou míru přesnosti udává tzv. *kappa index*. Ten přesnost klasifikace zjišťuje pro dvě nezávislé klasifikace se stejnou spolehlivostí. V potaz bere náhodnou shodu, což je pravděpodobnost, že se klasifikace s referenčními daty shoduje pouze náhodně. Index kappa lze zapsat následovně:

$$\kappa = \frac{P_0 - P_c}{1 - P_c} , \quad (5)$$

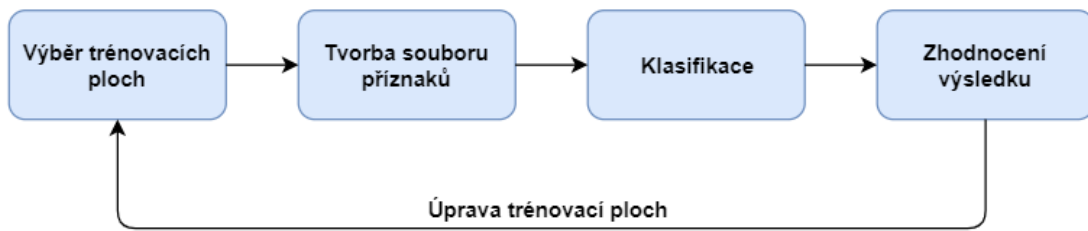
kde

P_0 vyjadřuje shodu mezi referenčními daty a klasifikací,

P_c vyjadřuje náhodnou shodu.

Kappa index nabývá hodnot nejčastěji v rozmezí $\langle 0;1 \rangle$, kdy hodnota 1 by znamenala, že se klasifikace vyhnula 100 % chyb vznikajících při náhodném procesu [14].

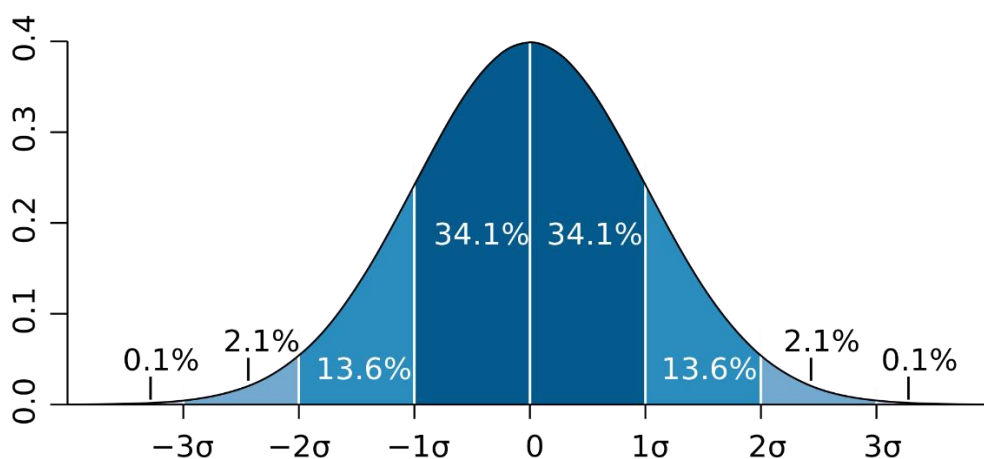
Řízená klasifikace je tedy posloupnost jednotlivých dílčích kroků a její schéma je znázorněno na následujícím obrázku.



Obr. 10: Schéma řízené klasifikace

5.1.1 Klasifikátor maximum likelihood

Maximum likelihood klasifikátor neboli klasifikátor maximální pravděpodobnosti je statistický klasifikátor založen na předpokladu, že rozdělení bodů tvořící jednu třídu trénovacích dat má normální (Gaussovo) rozdělení (viz obr. 11). V DPZ je tento předpoklad často splněn, a proto klasifikátor maximum likelihood patří mezi nejčastěji používané. Klasifikace probíhá tak, že je vypočtena pravděpodobnost, s jakou jednotlivé pixely patří do té či oné třídy. Po vypočtení pravděpodobnosti pro každou třídu se pixel zařadí tam, kde je jeho pravděpodobnost největší, případně nikam, pokud je nastavena nějaká prahová hodnota [14].



Obr. 11: Graf normálního rozdělení, zdroj: Wikipedia

5.1.2 Strojové učení (*machine learning*)

Strojové učení je obor umělé inteligence, které dává počítačům možnost učení, aniž by tak byly explicitně naprogramovány. Ty tak mohou předvídat pravděpodobný výsledek na základě vzorů. Strojové učení tak lze využít téměř ve všech možných oborech při řešení nejrůznějších problémů. Strojové učení lze rozdělit na učení řízené, neřízené a tzv. zpětnovazební. Při řízeném učení jsou počítači dány příklady vstupů a výstupů a jeho cílem je „*naucit se*“ pravidla vytvářející ze vstupů výstupy. Trénovací proces trvá tak dlouho, dokud není dosaženo požadované přesnosti. U neřízeného učení dochází k tomu, že počítač sám nachází vzory a struktury a ty pak seskupuje do různých skupin. Využití tak nachází například v hledání skrytých vzorců v datech. Řízené i neřízené učení lze kombinovat i dohromady, jako je tomu i u řízené a neřízené klasifikace. Zpětnovazební učení, jak již název napovídá, je založeno na učení pomocí zpětné vazby. Algoritmus provádí rozhodnutí v daném prostředí a postupně se zlepšuje na základě posledního rozhodnutí [16][17].

5.1.2.1 Random Forest klasifikátor

Své uplatnění nachází strojové učení i v oboru DPZ, především tedy v klasifikaci obrazových dat. Existuje mnoho algoritmů využívající strojové učení. V dálkovém průzkumu Země patří mezi nejčastěji používané algoritmy tzv. *Random Forest* klasifikátor, a to hlavně díky přesnosti, s jakou je schopen klasifikaci provést. Jedná se o souborový klasifikátor, což znamená, že využívá několik tzv. rozhodovacích stromů pod sebou. Pro zařazení prvku do určité třídy „shromažďuje hlasy“ z různých rozhodovacích stromů a následně rozhodne, do jaké třídy bude prvek zařazen. Mezi výhody tohoto klasifikátoru patří efektivita na objemných databázích, schopnost zpracovat tisíce různých vstupních proměnných bez jejich odstranění nebo efektivní metoda pro odhad chybějících dat a udržení přesnosti, i když velká část dat chybí [17][18].

6 Existující land cover produkty

Land cover produktů pokrývajících alespoň částečně území Evropy existuje celá řada. Lze je dělit na globální, kontinentální a národní. Přehled produktů ukazuje následující tabulka.

Tab. 3: Přehled land cover produktů

Název	Pokrytí	Správce
CCI Land Cover	globální	ESA
GlobLand30	globální	NGCC
GLC2000	globální	JRC
NASA Land Cover	globální	NASA
OSM	globální	komunita
CORINE Land Cover	kontinentální	EEA
HRL (High Resolution Layers)	kontinentální	EEA
OSM Land Use Land Cover	kontinentální	GIScience, Research Group, Heidelberg University
Urban Settlement	kontinentální	EEA
Urban Atlas	kontinentální	EEA
Riparian zones	kontinentální	EEA
Natura 2000 mapping	kontinentální	EEA
Land Cover Map of Europe 2017 (10m)	kontinentální	S2GLC Project
Pan-European land cover (30m) map of 2015 based on Landsat and LUCAS data	kontinentální	Pflugmacher, Rabe, Peters, Hostert
LISA	národní (Rakousko)	GeoVille
DeCOVER	národní (Německo)	DLR
SIOSE	národní (Španělsko)	National Reference Center on Land Cover and on Land Use and Spatial Planning

Jedním z cílů programu „Geo-harmonizer“ je vytvořit harmonizovaný land cover s upravenou nomenklaturou CORINE, kdy dojde k přechodu od ručně kreslených polygonů k automatizovaným pixelovým třídám.

7 Příprava trénovacích dat a klasifikace

Samotné klasifikaci musí předcházet příprava trénovacích množin. V opačném případě jsou totiž výsledky ve většině případů nekvalitní a přesnost dosažená při klasifikaci nízká. Jako trénovací množiny byla v této práci použita data CLC. Pro ukázkou zde budou uvedeny výsledky klasifikace, kdy do těchto dat nebylo nijak zasahováno, a které lze posléze porovnat s výsledky založenými na již upravených datech.

7.1 Klasifikace první úrovně CLC

První úroveň CLC obsahuje pouze pět tříd. Při překladu do češtiny se jedná o následující třídy:

- Zástavba (1)
- Zemědělské oblasti (2)
- Lesní a polopřírodní oblasti (3)
- Mokřady (4)
- Vodní plochy (5)

Nejprve tedy bylo vyzkoušeno, jak dopadne klasifikace, když se vezmou všechny tyto plochy a použijí se jako trénovací množiny. V prvním kroku bylo potřeba vzít data CLC, která jsou dostupná jako souvislá vrstva pokrývající celou Evropu, a oříznout je příslušnou scénou ze Sentinel-2. Jako první byla ke zpracování vybrána oblast okolo města Kristianstad. Do programu QGIS byla nahrána jedna z rastrových vrstev obsahující jedno pásmo ze Sentinel-2. Byla zvolena vrstva označovaná jako *TCI (true color image)*, díky které bylo možno si učinit i přibližný obraz o tom, jak dané území vypadá. Následně byla do programu nahrána i vektorová vrstva CLC. Pomocí funkce *Extract/Clip by Extent* byla vrstva CLC oříznuta tak, aby pokrývala pouze území vybrané scény. Vrstva CLC i scéna ze Sentinel-2 mají každá jiný referenční souřadnicový systém. U vrstvy CLC se jedná o systém ETRS89 (EPSG kód 3035) a u scény Sentinel-2 je to WGS84 (EPSG kód 32633). Aby se předešlo případným potížím, byla oříznutá vrstva CLC

přetransformována do systému, který používá Sentinel-2. K tomu byla použita funkce *Reproject Layer*.

Atributová tabulka vrstvy CLC kromě jiných obsahuje pouze informaci o zařazení jednotlivých polygonů do třetí úrovně.

Tab. 4: Atributová tabulka CLC

	fid	OBJECTID	Code_18	Remark	Area_Ha	ID	c18
1	5179	1782120	243	NULL	39,546322984965556	EU_1782120	243
2	5176	1780854	242	NULL	58,29467684996552	EU_1780854	242
3	5177	1782173	243	NULL	70,85170732493987	EU_1782173	243
4	5182	1786887	313	NULL	36,752374794973896	EU_1786887	313
5	5183	1788206	324	NULL	51,12874650503571	EU_1788206	324
6	5180	1788442	324	NULL	82,04902156999013	EU_1788442	324

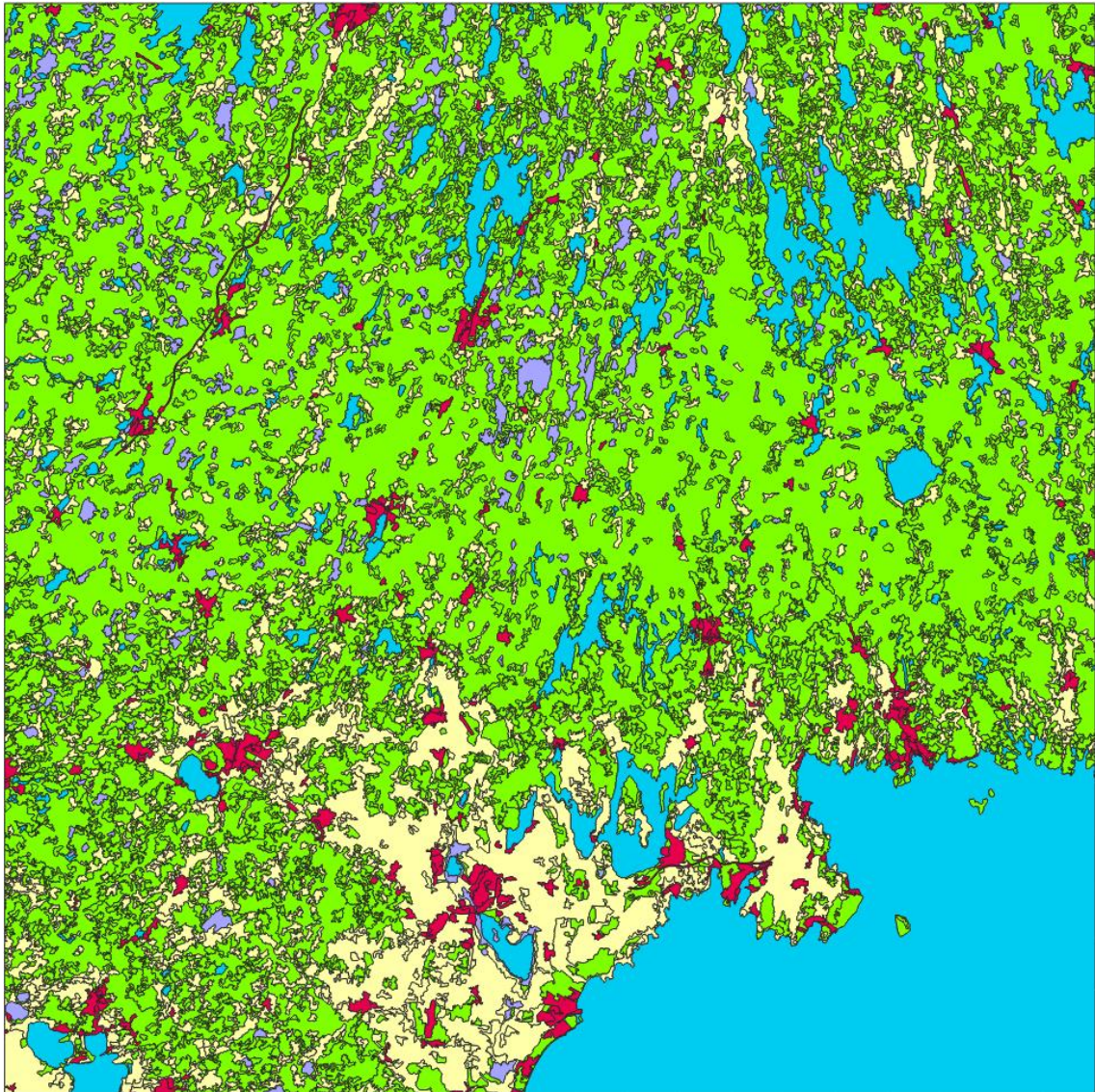
Pro vizualizaci první úrovně bylo tak vhodné vytvořit sloupec v atributové tabulce, který by dal informaci o zařazení do úrovně první. Pomocí nástroje *Field Calculator* byl vytvořen nový sloupec „*level_1*“ a hodnoty spočítány následujícím vzorcem.

$$level_1 = \frac{c18}{100} , \quad (6)$$

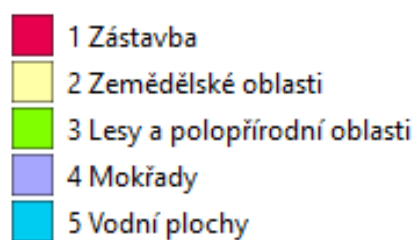
kde

c18 je sloupec v atributové tabulce CLC dávající informaci o třetí úrovni.

Sloupec „*c18*“ (viz. tab. 4) obsahuje hodnoty jako např.: 243, což znamená, že polygon patří do třídy 2.4.3. Samotné dělení stem by samo o sobě nestačilo, abychom dostali číslo říkající, do jaké kategorie první úrovně prvek patří. Nicméně při nastavení nového sloupce na hodnoty *integer* se problém vyřeší, protože v tomto případě jsou všechny hodnoty za desetinou čárkou „odříznuty“ bez zaokrouhlování. Díky této úpravě bylo možné si jednotlivé třídy první úrovně vizualizovat. Jednotlivým třídám byly přiřazeny následující barvy dle symboliky CORINE. Jednotlivé třídy první úrovně ve scéně z oblasti okolo Kristianstadu ukazuje obr. 12.



Obr. 12: Kristianstad – první úroveň CLC



Obr. 13: Legenda první úrovně CLC

Dalším krokem bylo vyzkoušet tyto plochy jako trénovací množiny pro řízenou klasifikaci. Ta byla spouštěna v programu GRASS, kde se nejprve „nalinkovala“ (z důvodu ušetření paměti) jednotlivá pásma s rozlišením 20 m. Těch je celkem 9

pro jednu scénu (plus pásmo B8, které bylo ale potřeba převzorkovat). Klasifikace probíhala na multitemporálních datech (byly použity dvě scény dohromady – jarní a letní) a celkově bylo tedy použito 20 pásem pro jednu klasifikaci.

Převzorkování pásma B8 proběhlo v programu QGIS. Použita byla funkce *Resampling* spadající pod nástroje SAGA. K přiřazení hodnot nově vytvořeným pixelům byla použita metoda průměrné hodnoty. V tomto případě funguje tak, že vezme čtyři pixely 10 m × 10 m tvořící jeden pixel 20 m × 20 m a zprůměruje jejich hodnoty. Posledním rastrem, který byl přidán, byla rasterizovaná vrstva CLC. Rasterizace byla provedena pomocí nástroje *v.to.rast*.

Po „nalinkování“ všech potřebných rastrových vrstev do programu GRASS bylo potřeba vytvořit tzv. obrazovou skupinu tvořenou všemi pásmy použitými ke klasifikaci. K tomu slouží v programu GRASS modul *i.group*. Dále byl vytvořen soubor se spektrálními příznaky modulem *i.gensig* a nakonec byla provedena samotná klasifikace maximum likelihood modulem *i.maxlik*. Přesnost klasifikace, tj. vytvoření kontingenční tabulky, byla zjištěna pomocí modulu *r.kappa*.

Tab. 5: Kontingenční tabulka (Kristianstad) bez úprav CLC

Třída	1	2	3	4	5	Počet pixelů
1	67,54	10,52	20,30	1,30	0,33	730883
2	13,59	65,27	17,68	3,24	0,22	5046981
3	4,31	3,96	77,59	13,48	0,65	18462414
4	0,96	2,57	11,42	83,48	1,57	710945
5	1,11	0,30	3,09	1,73	93,77	5188877

Poznámka: Údaje v tabulce (i ve všech následujících kontingenčních tabulkách), kromě informace o počtu pixelů, jsou v procentech.

Celková přesnost	78,21 %
Kappa index	0,65

Jak je z kontingenční tabulky vidět, tak kromě vodních ploch výsledky nedosahují nijak závratných přesností. Především u zástavby a zemědělských oblastí je přesnost poměrně nízká a je vidět vzájemná záměna mezi těmito třídami. Obě třídy pak značnou částí zasahují i do třídy lesů a polopřirodních oblastí. Z výsledků tedy vyplývá, že je potřeba trénovací plochy jednotlivých tříd upravit.

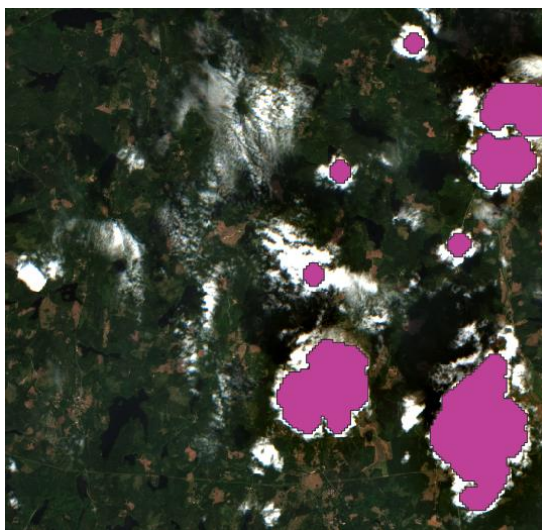
7.1.1 Úprava dat první úrovně CLC

Scény byly vybírány tak, aby byly bezoblačné. Takové scény ale nejsou vždy k dispozici a stejné to bylo i v tomto případě, kdy bylo pár míst pokrytých mraky tak, jako je tomu na následujícím obrázku.



Obr. 14: Oblačnost na části letní scény z Kristianstadu

Nejedná se o velké plochy, ale i tak mohou výsledek klasifikace negativně ovlivnit. V souborovém adresáři obsahující jednotlivá pásma ze Sentinel-2 se nachází také soubor s názvem *MSK_CLOUDS_B00.gml*. Jedná se o vektorovou masku pokrývající oblačnost ve scéně.



Obr. 15: Maska oblačnosti

Je vidět, že maska ne vždy pokryje celou oblačnost. Proto je vhodné ji pomocí nástroje *Buffer* dostatečně rozšířit, aby oblačnost byla zakrytá co nejvíce. Funkcí *Difference* byly následně tyto plochy z trénovacích dat odstraněny.

Dále bylo vhodné se podívat na jednotlivé podtřídy tvořící pět základních tříd CLC. Třída „Zástavba“ (*Artificial Surfaces*) z tab. 3 se skládá z následujících podtříd.

- Městské oblasti (1.1)
 - Souvislá zástavba (1.1.1)
 - Nesouvislá zástavba (1.1.2)
- Průmyslové, obchodní a dopravní jednotky (1.2)
 - Průmyslové nebo obchodní jednotky (1.2.1)
 - Silniční a železniční sítě a přidružené pozemky (1.2.2)
 - Přístavy (1.2.3)
 - Letiště (1.2.4)
- Doly, skládky a staveniště (1.3)
 - Místa těžby nerostů (1.3.1)
 - Skládky (1.3.2)
 - Staveniště (1.3.3)
- Umělé nezemědělské oblasti s vegetací (1.4)
 - Městská vegetace (1.4.1)
 - Sportovní a rekreační zařízení (1.4.2)

Už jen z tohoto výčtu je patrné, že ne všechny podtřídy jsou vhodné jako trénovací plochy pro třídu „Zástavba“. Např. podtřída 1.4 se jistě jako trénovací množina bude zaměřovat s třídou lesů či zemědělských oblastí. Jako trénovací množiny pro třídu „Zástavba“ byly tedy použity pouze následující podtřídy: *1.1.1*, *1.2.1*, *1.2.2*, *1.2.3* a *1.2.4*. Zbylé podtřídy byly z trénovacích ploch odstraněny.

Stejně se postupovalo i u dalších tříd první úrovně. Třída „Zemědělské oblasti“ se skládá z následujících podtříd:

- Orná půda (2.1)
 - Nezavlažovaná orná půda (2.1.1)
 - Trvale zavlažovaná půda (2.1.2)
 - Rýžová pole (2.1.3)
- Trvalé plodiny (2.2)
 - Vinice (2.2.1)
 - Ovocné stromy a bobuloviny (2.2.2)
 - Olivovníky (2.2.3)
- Pastviny (2.3)
 - Pastviny (2.3.1)
- Různorodé zemědělské oblasti (2.4)
 - Roční plodiny spojené s trvalými kulturami (2.4.1)
 - Různorodé kultivační oblasti (2.4.2)
 - Půda převážně zemědělsky obhospodařovaná, s významnými oblastmi přirozené vegetace (2.4.3)
 - Zemědělské oblasti pokryté lesním porostem (2.4.4)

Třída „Zemědělské oblasti“ je značně různorodá, a proto i ta byla dosti upravena. Pro trénovací množiny reprezentující „Zemědělské oblasti“ byly vybrány pouze podtřídy 2.1.1 a 2.3.1, tedy nezavlažovaná orná půda a pastviny. Třída „Lesy a polopřírodní oblasti“ se skládá z následujících podtříd:

- Les (3.1)
 - Listnatý les (3.1.1)
 - Jehličnatý les (3.1.2)
 - Smíšený les (3.1.3)
- Keře či bylinná vegetace (3.2)
 - Přírodní travní porosty (3.2.1)
 - Slatiny a vřesoviště (3.2.2)
 - Sklerofylní vegetace (3.2.3)
 - Přechodné lesnaté či keřovité oblasti (3.2.4)
- Otevřená místa s trochou nebo žádnou vegetací (3.3)
 - Pláže, duny, písky (3.3.1)

- Holé skály (3.3.2)
- Oblasti s řídkou vegetací (3.3.3)
- Vyhořelé oblasti (3.3.4)
- Ledovce a oblasti s trvalým sněhem (3.3.5)

Tato třída je rozdělena do tří, ne příliš podobných, podtříd. Bylo tedy rozhodnuto, že pro klasifikaci první úrovně se tato třída zúží pouze na „Lesy“, tzn. že byly vybrány pouze podtřídy „Listnatý les“ (3.1.1) a „Jehličnatý les“ (3.1.2). Dalšími třídami jsou třídy „Mokřady“ a „Vodní plochy“. Ty se také dělí do různých podtříd, nicméně jejich různorodost není tak velká, a proto byly tyto třídy ponechány se všemi svými podtřídami.

Po výběru vhodných podtříd reprezentující jednotlivé základní třídy proběhly ještě dvě menší úpravy. První z nich byl tzv. vnitřní buffer, kterým byly odstraněny přechody mezi jednotlivými polygony různých tříd. Byla tedy použita funkce *Buffer* a nastavena hodnota -20 m, což odpovídá jednomu dvacetimetrovému pixelu. Druhou úpravou bylo „rozsekání“ polygonů na menší části, a to z dříve zmíněných důvodů. Pro trénovací množiny je vhodnější používat menší než větší plochy a ve scéně z Kristianstadu se nachází např. lesní polygon zabírající téměř čtvrtinu celé scény, což není nejvhodnější. K rozdělení polygonů byla nejprve vytvořena polygonová mřížka 2 km × 2 km pomocí funkce *Create Grid*. Následně byly polygony tvořící trénovací plochy touto mřížkou rozděleny funkcí *Intersection*. Po těchto úpravách byla klasifikace spuštěna znovu s následujícím výsledkem.

Tab. 6: Kontingenční tabulka (Kristianstad) po základních úpravách

Třída	1	2	3	4	5	Počet pixelů
1	71,90	9,54	16,27	2,05	0,24	100386
2	7,62	78,77	11,95	1,61	0,05	3529294
3	2,90	3,45	86,31	6,82	0,51	15020855
4	0,97	2,20	15,50	79,89	1,44	634840
5	1,08	0,14	2,36	1,12	95,30	5033153

Celková přesnost	86,85 %
Kappa index	0,78

Z výsledků je vidět zlepšení přesnosti klasifikace (kromě mokřadů, které se poněkud překvapivě zhoršily), nicméně stále dochází k velké záměně mezi třídami. Bylo tedy zapotřebí podívat se na jednotlivé třídy hlouběji. Nejprve bylo potřeba zjistit, proč dochází k záměně mezi třídami. Na základě šetření byly zjištěny následující problémy a použita následující řešení.

Tab. 7: Úpravy trénovacích ploch

Třída	Problém	Řešení	Pravidlo
1. Zástavba	Vegetace v zástavbě	Odstranění všech pixelů vegetace	$NDVI > 0,2$
	Vodní plochy v zástavbě	Odstranění všech pixelů vody	$NDVI < 0$
2. Zemědělské oblasti	Holá půda v zemědělských oblastech	Odstranění všech pixelů holé půdy	$NDVI < 0,2$
	Možný výskyt zástavby a vodních ploch v zemědělských oblastech	Odstranění pixelů zástavby a vodních ploch (vyřešeno předchozím řešením)	$NDVI < 0,2$
	Stromová seskupení v zemědělských oblastech	Odstranění všech pixelů stromových seskupení	Smazání trénovacích ploch zemědělských oblastí klasifikovaných jako třída lesů
3. Lesy	Holá půda či komunikace v lesích	Odstranění všech pixelů holé půdy a komunikací	$NDVI < 0,2$
	Vodní plochy v lesích	Odstranění všech pixelů vodních ploch (vyřešeno předchozím řešením)	$NDVI < 0$
	Vykácená místa či paseky v lesích	Odstranění těchto nelesních pixelů	Smazání trénovacích ploch lesů klasifikovaných jako zemědělské oblasti
4. Mokřady	Záměna s třídou lesů	Odstranění pixelů lesa	Smazání trénovacích ploch mokřadů klasifikovaných jako třída lesů
5. Vodní plochy	Ostrovky menší než MMU	Odstranění pixelů neobsahující vodu	$NDVI > 0$ ($> 0,1$ pro oblast Říma)

Jak je z tabulky vidět, úpravy jsou založené převážně na tzv. NDVI neboli na normovaném rozdílovém vegetačním indexu. Jedná se o poměrový vegetační index a mezi vegetačními indexy je nejpoužívanější. Vyjadřuje nárůst hodnoty odrazivosti

mezi vlnovou délkou červeného a blízkého infračerveného záření, která je nejvyšší u zdravé zelené vegetace [14]. Nabývá hodnot $\langle -1; 1 \rangle$ a vypočte se podle vzorce

$$NDVI = \frac{NIR - R}{NIR + R}, \quad (7)$$

kde

NIR jsou hodnoty blízkého infračerveného záření,

R jsou hodnoty červeného záření.

Hodnoty NDVI pro různé druhy povrchů jsou uvedeny v následující tabulce.

Tab. 8: Hodnoty NDVI, [19]

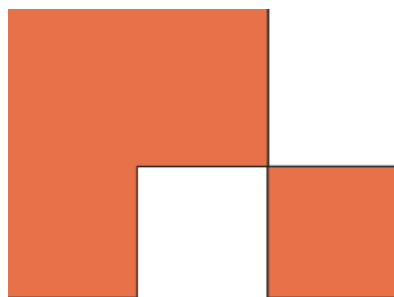
Druh povrchu	NDVI
Hustá vegetace	0,500 – 1
Střední vegetace	0,140 – 0,500
Řídká vegetace	0,090 – 0,140
Holá půda	0,025 – 0,140
Oblačnost, led a sníh	-0,046 – 0,025
Voda	-1 – -0,046

Poznámka: Hodnoty NDVI se z různých zdrojů pro různé povrchy mírně liší a je potřeba je brát orientačně. Hodnoty NDVI mohou být ovlivněny např. množstvím srážek a z toho plynoucím výskytem vodních kapek na vegetaci, čímž se hodnota NDVI snižuje. V práci bylo vycházeno z hodnot udávaných ve zdrojích [19] a [20].

Pro aplikování zmíněných úprav trénovacích ploch bylo tedy vypočteno pomocí funkce *Raster Calculator* pásmo NDVI pro jarní i letní scénu. Vypočtená pásma byla následně oříznuta příslušnou klasifikační třídou pomocí funkce *Clip Raster by Mask Layer*, a každé z nich bylo poté znovu pomocí funkce *Raster Calculator* překlasifikováno na základě zvolených pravidel z tab. 6. Pro třídu „Zástavba“ byly tedy pixely s hodnotami $\langle 0; 0,2 \rangle$ (kombinace dvou pravidel dohromady) překlasifikovány na hodnotu 0 (nesplňují pravidlo, budeme je chtít v trénovacích plochách ponechat) a zbylé na hodnotu 1 (budeme je chtít odstranit). Tím tedy byla vytvořena maska pro vymezení třídy „Zástavba“.

Překlasifikované vrstvy byly funkcí *Polygonize (Raster to Vector)* převedeny do vektorové podoby, kde byly následně smazány polygony s hodnotou 0 (zůstanou tak pouze oblasti, které chceme z trénovacích ploch odstranit). Takto upravené vektorové vrstvy byly spojeny do jedné funkcí *Merge Vector Layers* a vzniklé polygony pak byly funkcí *Difference* odstraněny z trénovacích ploch.

Poznámka: Při takto prováděné polygonizaci rastru často dochází k případům jako na obr. 16, tj. že se polygony vzájemně dotýkají pouze v jednom bodě, což při následném spuštění funkce *Difference* vyústí v zobrazení chybové hlášky ohledně neplatné geometrie (invalid geometry). Tento problém lze ale vyřešit použitím funkce *Buffer* s hodnotou bufferu nastavenou na 0.



Obr. 16: Neplatná geometrie vektorové vrstvy

Tímto způsobem byly ošetřeny všechny problémy týkající se NDVI a následně byla znovu spuštěna klasifikace.

Tab. 9: Kontingenční tabulka (Kristianstad) po úpravách NDVI

Třída	1	2	3	4	5	Počet pixelů
1	99,31	0,65	0,02	0,01	0,01	36618
2	0,81	84,87	12,61	1,71	0,00	3117390
3	0,29	5,80	87,79	6,08	0,05	14966203
4	0,58	2,63	17,92	78,40	0,48	634840
5	0,86	0,00	0,12	0,00	99,02	4009643

Celková přesnost	89,12 %
Kappa index	0,80

Úpravy založené na NDVI výrazně pomohly především třídě „Zástavba“ kde se správně klasifikovalo téměř 100 % pixelů. Zlepšení je vidět i u tříd „Zemědělské

oblasti“, „Lesy“ a „Vodní plochy“. Jediné „Mokřady“ opět mírně s přesností klesly, ale je to dáno tím, že se více pixelů z této třídy klasifikovalo do „Lesů“. Na řadu tedy přišly zbylé úpravy zmíněné v tab. 6.

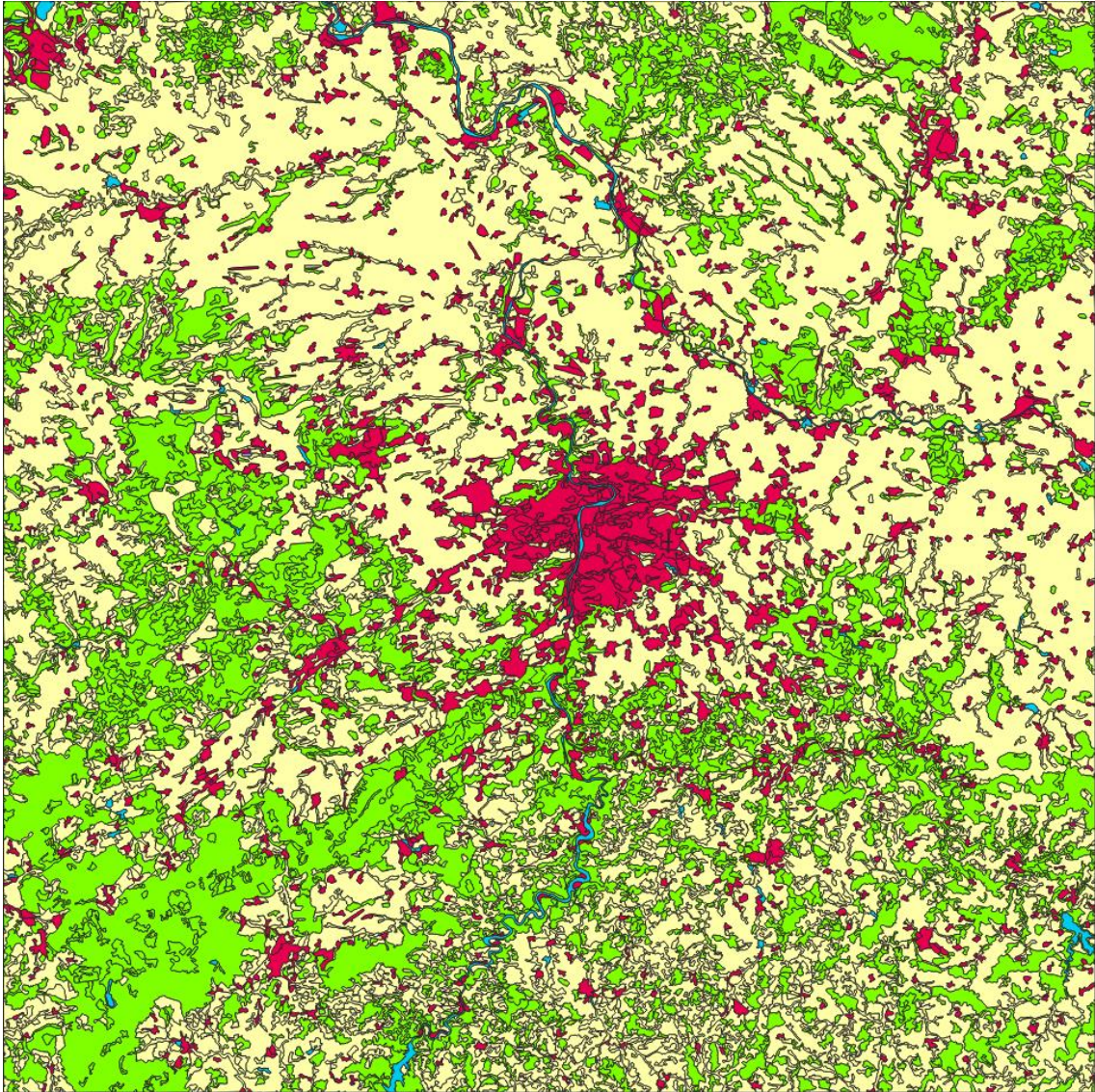
Jako příklad je zde uveden případ míchání „Zemědělských oblastí“ a „Lesů“. Nejprve byl vzat klasifikovaný rastr a funkcí *Raster Calculator* byl překlasifikován tak, že pixelům klasifikovaným jako „Les“ byla přiřazena hodnota 1 a všem ostatním pixelům hodnota 0. Dále byl tento překlasifikovaný rastr oříznut trénovacími plochami „Zemědělských oblastí“ funkcí *Clip Raster With Polygon* (funkce *Clip Raster by Mask Layer* nebyla použita, protože výpočet s ní trval příliš dlouho). Ten byl následně polygonizován funkcí *Polygonize (Raster to Vector)* a z této vektorové vrstvy byly smazány všechny polygony s hodnotou 0. Nakonec byly funkcí *Difference* odstraněny tyto plochy z trénovacích množin. Stejným způsobem byly ošetřeny i zbylé problémy uvedené v tab. 6. Výsledek klasifikace po takto upravených trénovacích množinách je uveden v následující tabulce.

Tab. 10: Kontingenční tabulka (Kristianstad) po finálních úpravách

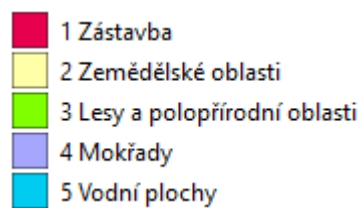
Třída	1	2	3	4	5	Počet pixelů
1	99,26	0,70	0,02	0,01	0,01	36618
2	0,84	97,16	0,09	1,90	0,00	2724293
3	0,32	2,38	95,91	1,34	0,05	13350020
4	0,61	3,32	0,19	95,25	0,63	484226
5	0,91	0,00	0,07	0,00	99,02	4009643

Celková přesnost	96,67 %
Kappa index	0,94

Z výsledků vyplývá, že díky provedeným úpravám bylo dosaženo vyšší přesnosti klasifikace. Klasifikace pro tuto scénu na první úrovni CLC mohla být tedy považována za dokončenou. Pro zbylé dvě oblasti byla klasifikace provedena stejným způsobem a budou zde ukázány výsledky před a po úpravách trénovacích ploch. Jako první ze zbylých dvou území byla zpracovávána oblast z České republiky.



Obr. 17: Mladá Boleslav – první úroveň CLC



Obr. 18: Legenda první úrovně CLC

Na scéně z České republiky se nachází pouze čtyři malé polygony zastupující třídu „Mokřady“ a proto bylo rozhodnuto, že z důvodu velmi nízkého zastoupení nebude snaha tuto třídu klasifikovat.

Tab. 11: Kontingenční tabulka (Ml. Boleslav) bez úprav CLC

Třída	1	2	3	5	Počet pixelů
1	65,60	14,59	12,85	6,96	2768028
2	8,33	79,99	10,60	1,08	18697470
3	3,09	3,63	92,22	1,07	8478948
5	6,31	1,92	5,71	86,06	193863

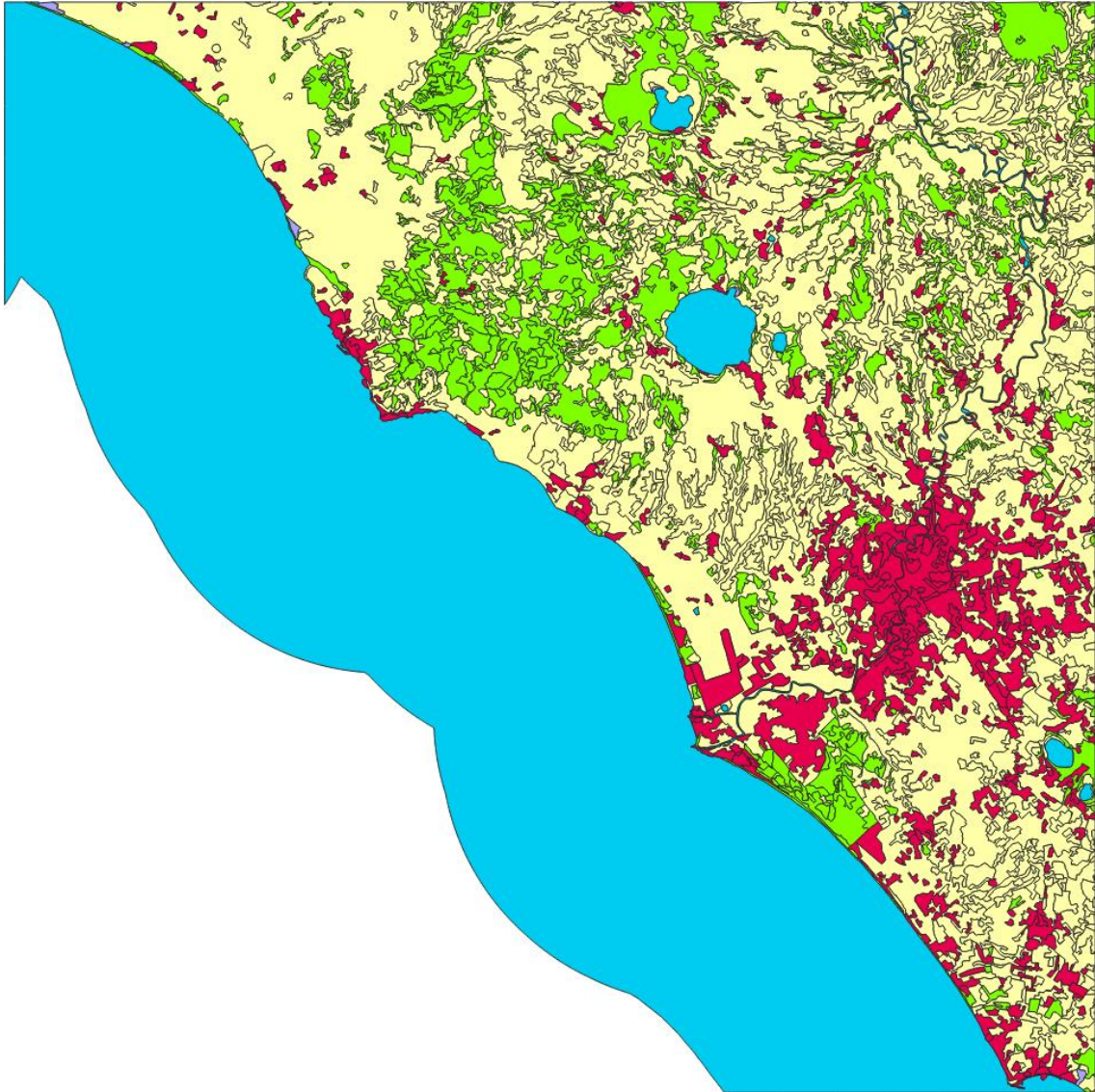
Celková přesnost	82,15 %
Kappa index	0,69

Tab. 12: Kontingenční tabulka (Ml. Boleslav) s úpravami NDVI

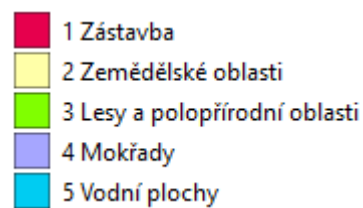
Třída	1	2	3	5	Počet pixelů
1	99,91	0,01	0,00	0,08	135348
2	0,77	95,31	3,92	0,00	11589343
3	0,53	3,75	95,72	0,00	4835052
5	2,01	0,00	0,00	97,99	70718

Celková přesnost	95,48 %
Kappa index	0,90

Při klasifikování scény z České republiky stačilo v úpravách trénovacích ploch zahrnout pouze úpravy na základě NDVI. Úpravy typu „Smazání ploch z trénovacích množin klasifikovaných jako jiná třída“ nebylo nutné zavádět, protože nedocházelo k tak významné záměně tříd, jako tomu bylo ve scéně ze Švédska. Posledním zpracovávaným územím byla oblast okolo Říma.



Obr. 19: Řím – první úroveň CLC



Obr. 20: Legenda první úrovně CLC

Stejně jako u scény z České republiky, tak i scéna z okolí Říma má velmi nízké zastoupení třídy mokřadů (7 polygonů), a proto ani zde nebyla snaha tuto třídu klasifikovat. Pro vodní plochy byla zvednuta hodnota hranice NDVI na 0,1 z důvodu, že

většina vnitrozemských vodních ploch neměla zápornou hodnotu NDVI. Důvodem může být znečištění případně výskyt vodních rostlin ovlivňující hodnotu NDVI.

Tab. 13: Kontingenční tabulka (Řím) bez úprav CLC

Třída	1	2	3	5	Počet pixelů
1	71,17	17,26	11,56	0,01	1813797
2	10,66	61,12	28,22	0,00	9944164
3	2,35	5,92	91,71	0,01	2482870
5	0,81	0,07	0,06	99,06	9081939

Celková přesnost	79,93 %
Kappa index	0,71

Tab. 14: Kontingenční tabulka (Řím) s úpravami NDVI

Třída	1	2	3	5	Počet pixelů
1	99,99	0,01	0,00	0,00	217757
2	1,18	89,85	8,97	0,00	4202334
3	0,15	5,55	94,29	0,00	1759446
5	0,54	0,00	0,00	99,46	8960214

Celková přesnost	96,20 %
Kappa index	0,93

Tab. 15: Kontingenční tabulka (Řím) po finálních úpravách

Třída	1	2	3	5	Počet pixelů
1	99,99	0,01	0,00	0,00	217757
2	1,23	98,72	0,04	0,00	3825558
3	0,15	2,30	97,55	0,00	1661716
5	0,54	0,00	0,00	99,46	8960214

Celková přesnost	99,06 %
Kappa index	0,98

Klasifikace scény z okolí Říma dopadla velmi dobře. Byly zavedeny jak úpravy na základě NDVI, tak i úpravy založené na vyhledání překryvů mezi třídami.

Jediným nedostatkem bylo, že se celá Tibera protékající Římem klasifikovala jako zástavba. Řešením by bylo zavést samostatnou třídu pro řeky, čímž by se Tibera správně klasifikovala. (Bude to ukázáno při klasifikaci druhé úrovně CLC.)



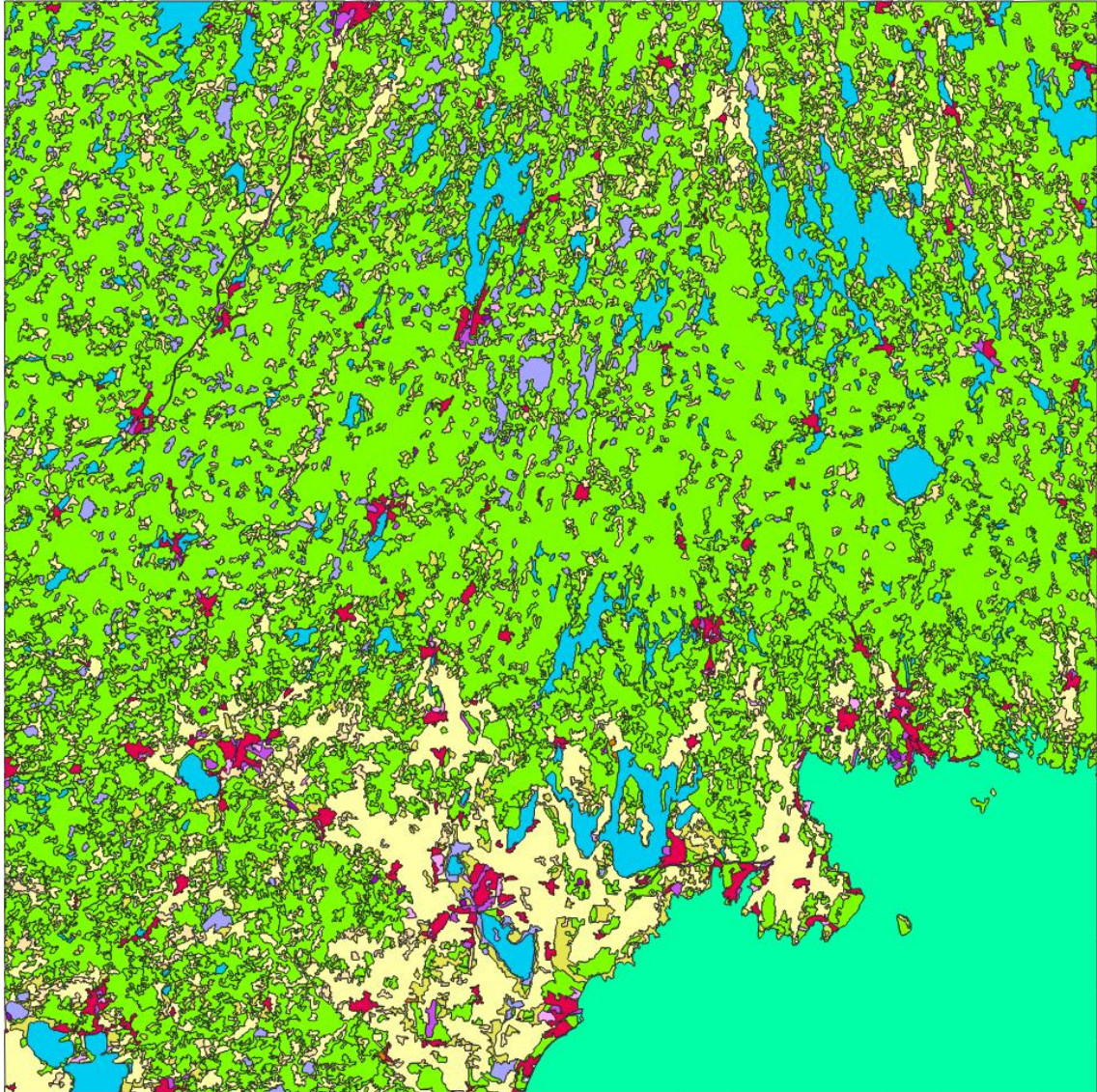
Obr. 21: Tibera klasifikovaná jako zástavba

7.2 Klasifikace druhé úrovně CLC

Druhou úroveň CLC tvoří celkem 15 následujících tříd.

- Městské oblasti (1.1)
- Průmyslové, obchodní a dopravní jednotky (1.2)
- Doly, skládky a staveniště (1.3)
- Umělé nezemědělské oblasti s vegetací (1.4)
- Orná půda (2.1)
- Trvalé plodiny (2.2)
- Pastviny (2.3)
- Různorodé zemědělské oblasti (2.4)
- Les (3.1)
- Keře či bylinná vegetace (3.2)
- Otevřená místa s trochou nebo žádnou vegetací (3.3)
- Vnitrozemské mokřady (4.1)
- Přímořské mokřady (4.2)
- Vnitrozemské vodní plochy (5.1)
- Oceány a moře (5.2)

Nejprve bylo tedy vyzkoušeno, jaký bude výsledek klasifikace druhé úrovně CLC bez úprav trénovacích ploch. Postup byl stejný jako při klasifikaci první úrovně CLC a jako první byla znovu vzata scéna z okolí města Kristianstad.



Obr. 22: Kristianstad – druhá úroveň CLC

 11 Městské oblasti	 31 Les
 12 Průmyslové, obchodní a dopravní jednotky	 32 Keře či bylinná vegetace
 13 Doly, skládky a staveniště	 33 Otevřená místa s trochou nebo žádnou vegetací
 14 Umělé nezemědělské oblasti s vegetací	 41 Vnitrozemské mokřady
 21 Orná půda	 42 Přímořské mokřady
 22 Trvalé plodiny	 51 Vnitrozemské vodní plochy
 23 Pastviny	 52 Oceány a moře
 24 Různorodé zemědělské oblasti	

Obr. 23: Legenda druhé úrovně CLC

Pro vizualizaci druhé úrovně CLC byl vytvořen podobně jako u první úrovně pomocný sloupec v atributové tabulce s názvem *level_2*. Jeho hodnoty byly vypočteny následovně:

$$level_2 = \frac{c18}{10}, \quad (8)$$

kde

c18 je informace o třetí úrovni CLC (viz. tab. 4).

Tab. 16: Kontingenční tabulka (Kristianstad, 2. úroveň) bez úprav CLC

Třída	11	12	13	14	21	22	23	24	31	32	33	41	42	51	52	Počet pixelů
11	35,69	9,15	2,57	11,13	1,09	8,41	6,95	1,77	9,41	3,70	6,63	0,40	2,79	0,30	0,02	526915
12	12,82	32,14	10,40	3,66	1,40	3,46	5,88	1,27	6,09	8,49	8,47	1,34	4,24	0,25	0,09	116793
13	5,37	12,43	31,77	3,37	1,78	4,13	6,88	1,43	8,52	6,87	10,82	2,01	2,16	2,29	0,16	22049
14	9,49	2,46	1,82	27,82	1,70	9,18	16,95	2,80	15,32	5,62	4,03	0,63	1,94	0,23	0,00	65165
21	2,82	1,10	1,49	4,52	43,41	6,98	25,70	3,23	4,61	2,00	1,09	0,91	2,01	0,13	0,00	3376638
22	2,12	0,90	0,54	5,25	4,55	66,84	5,85	0,80	8,86	1,50	0,90	0,96	0,42	0,52	0,00	5009
23	3,89	0,91	0,81	7,46	5,86	8,61	45,66	4,81	9,44	4,75	1,82	3,52	2,21	0,26	0,00	662557
24	3,49	0,85	0,50	6,72	6,38	4,45	29,51	9,27	15,29	11,53	1,01	5,78	4,22	1,00	0,00	1002472
31	1,03	0,40	0,22	2,54	0,55	3,16	2,95	1,75	46,01	29,93	0,92	8,64	0,85	1,03	0,01	17136287
32	0,49	0,20	0,25	0,65	0,43	0,54	1,66	1,20	9,80	76,64	1,01	6,27	0,75	0,13	0,00	1322858
33	1,40	1,40	5,52	1,38	0,07	1,37	2,09	1,09	13,66	8,99	61,12	0,63	0,39	0,68	0,20	5416
41	0,11	0,29	0,25	0,37	0,84	0,36	1,40	1,09	5,84	6,15	0,15	80,63	0,34	2,18	0,00	710231
42	2,06	4,90	0,49	0,98	1,37	0,49	8,33	2,16	16,86	13,24	0,78	6,08	42,16	0,10	0,00	1020
51	0,11	0,08	0,75	0,29	0,13	0,18	0,14	0,57	4,00	0,61	0,11	3,02	0,05	85,35	4,60	1820469
52	0,04	0,07	0,87	0,06	0,01	0,02	0,04	0,04	0,07	0,04	0,59	0,03	0,02	0,42	97,66	3366221

Celková přesnost	54,52 %
Kappa index	0,43

Z tabulky vyplývá, že kromě posledních dvou tříd („Vnitrozemské vodní plochy“, „Oceány a moře“) dochází ke značné záměně v podstatě u každé třídy. Byly tedy zavedeny stejné úpravy jako u klasifikace první úrovně. Zároveň bylo také vhodné podívat se na jednotlivé třídy a případně rozhodnout, zda má smysl pokoušet se o jejich klasifikaci. Takto byly vybrány třídy 1.4 „Umělé nezemědělské oblasti s vegetací“ a 2.4 „Různorodé zemědělské oblasti“, které se v podstatě nedají na základě spektrálních příznaků rozpoznat od zbylých tříd. Dále bylo rozhodnuto, že

nebude snaha klasifikovat ani třídu 4.2 „Přímořské mokřady“, kterou v této scéně reprezentoval jediný polygon a docházelo u ní k významné záměně s jinými třídami. Kromě těchto tří tříd byly z trénovacích množin odstraněny i podtřídy 1.1.2 „Nesouvislá zástavba“ a 2.1.2 „Trvale zavlažovaná půda“, které nebyly brány v úvahu ani při klasifikaci první úrovně CLC.

Nejprve byly provedeny úpravy na základě NDVI. Ty jsou shrnuty v následující tabulce.

Tab. 17: Pravidla pro hodnoty NDVI pro 2. úroveň CLC

Třída(y)	Odstranění pixelů s hodnotou NDVI
1.1, 1.2, 1.3	NDVI < 0 a NDVI > 0,2
2.1, 2.2, 2.3	NDVI < 0,2
3.1, 3.2	NDVI < 0,2
3.3	NDVI < 0 a NDVI > 0,3
5.1, 5.2	NDVI > 0 (> 0,1 pro oblast Říma)

Po těchto úpravách následovalo vyhledávání překryvů mezi jednotlivými třídami a jejich odstraňování z trénovacích ploch. Bylo tak činěno postupně s tím, že po několika úpravách byla klasifikace spuštěna znovu. Konečný výsledek ukazuje následující kontingenční tabulka.

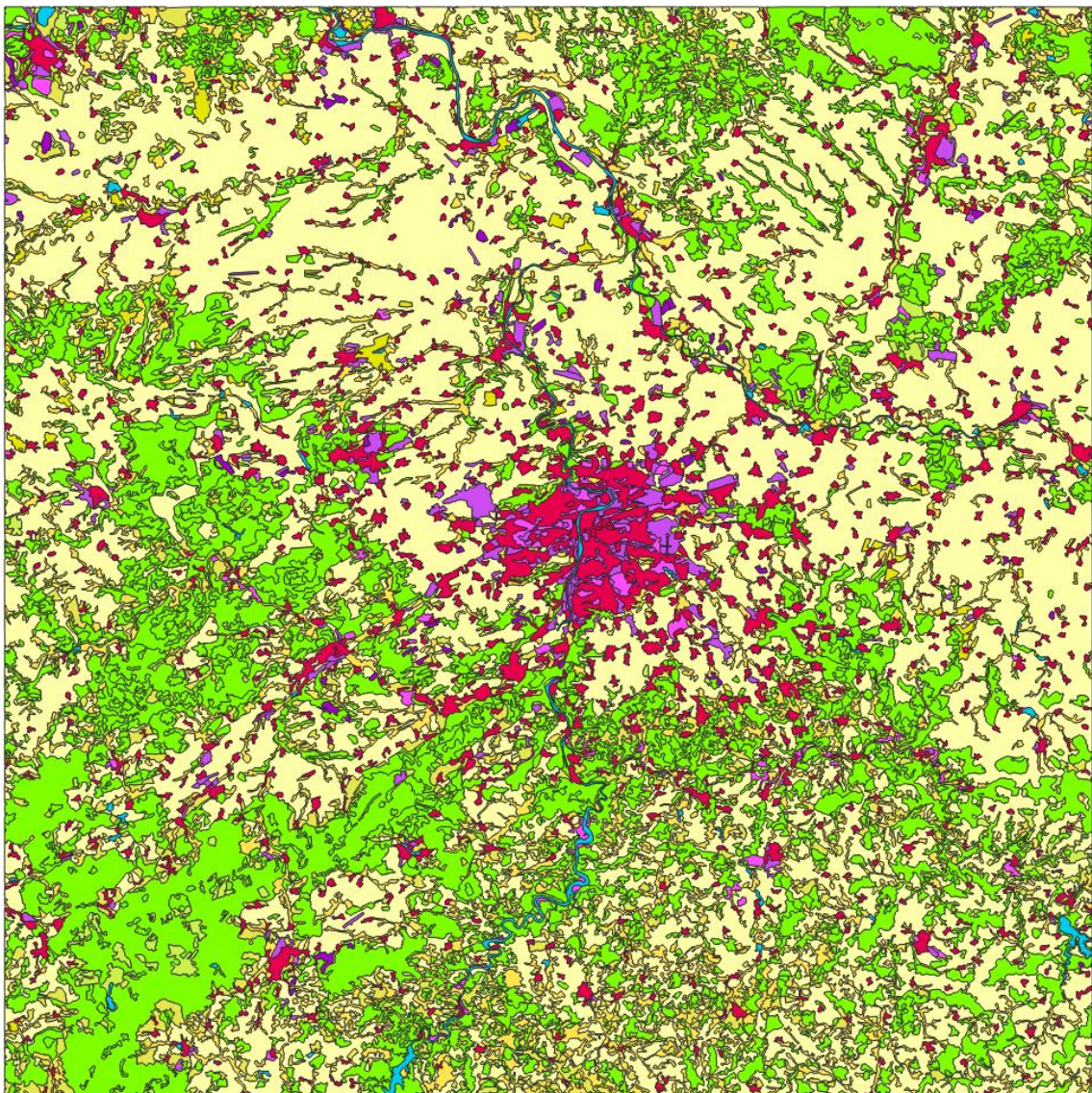
Tab. 18: Kontingenční tabulka (Kristianstad, 2. úroveň) s úpravami CLC

Třída	11	12	13	21	22	23	31	32	33	41	51	52	Počet pixelů
11	93,22	4,95	1,69	0,00	0,00	0,00	0,00	0,00	0,13	0,00	0,00	0,00	767
12	12,62	59,91	25,71	0,43	0,00	0,00	0,00	0,00	1,32	0,00	0,00	0,02	15608
13	0,41	3,75	90,64	0,12	0,00	0,00	0,00	0,00	5,05	0,00	0,02	0,00	4135
21	0,04	0,46	0,27	88,59	0,47	4,33	0,70	2,35	1,50	1,29	0,00	0,00	1292486
22	0,00	0,10	0,00	2,67	87,17	5,64	1,98	1,45	0,73	0,26	0,00	0,00	3032
23	0,03	0,26	0,17	1,37	0,42	88,60	0,85	3,69	0,92	3,67	0,00	0,00	406812
31	0,01	0,15	0,30	1,69	0,61	2,22	92,38	1,85	0,42	0,25	0,11	0,00	10153714
32	0,00	0,09	0,03	1,25	0,18	3,92	3,66	87,12	0,26	3,50	0,00	0,00	972663
33	0,48	1,20	1,93	0,00	0,00	0,00	0,00	0,00	96,39	0,00	0,00	0,00	415
41	0,01	0,10	0,71	1,31	0,33	3,20	0,30	0,89	0,16	92,23	0,77	0,00	501073
51	0,00	0,19	0,90	0,00	0,00	0,00	0,03	0,00	0,00	0,00	98,86	0,02	783827
52	0,00	0,06	0,69	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	98,72	3225816

Celková přesnost	93,15 %
Kappa index	0,89

Z výsledků vyplývá, že i klasifikaci druhé úrovně CLC pro území okolo Kristianstadu lze provést s poměrně vysokou přesností. Jedinou třídou, kterou se nepovedlo dostatečně dobře klasifikovat, je třída 1.2 „Průmyslové, obchodní a dopravní jednotky“. Třidu by proto bylo vhodné spojit s jednou ze tříd 1.1 či 1.3, případně ji z klasifikace vynechat úplně. S tím by následně vzrostla i celková přesnost klasifikace.

Dalším klasifikovaným územím byla oblast z České republiky.



Obr. 24: Mladá Boleslav – druhá úroveň CLC

11	Městské oblasti
12	Průmyslové, obchodní a dopravní jednotky
13	Doly, skládky a staveniště
14	Umělé nezemědělské oblasti s vegetací
21	Orná půda
22	Trvalé plodiny
23	Pastviny
24	Různorodé zemědělské oblasti
31	Les
32	Keře či bylinná vegetace
41	Vnitrozemské mokřady
51	Vnitrozemské vodní plochy

Obr. 25: Legenda druhé úrovně CLC (oblast z ČR)

Oproti scéně ze Švédska, která obsahovala všech 15 tříd druhé úrovně CLC, scéna z České republiky obsahuje tříd o tři méně, tedy 12. Výsledek bez úprav trénovacích ploch ukazuje následující tabulka.

Tab. 19: Kontingenční tabulka (Ml. Boleslav, 2. úroveň) bez úprav CLC

Třída	11	12	13	14	21	22	23	24	31	32	41	51	Počet pixelů
11	27,68	5,55	5,88	32,74	1,82	7,22	7,47	2,17	3,84	1,65	1,42	2,56	2015457
12	11,38	21,21	16,50	22,96	1,51	5,68	8,21	1,24	3,87	3,79	1,12	2,53	447975
13	2,84	6,95	40,75	9,12	2,59	10,71	7,92	1,31	5,77	4,90	2,26	4,88	85951
14	4,81	1,60	2,28	26,97	1,02	5,60	19,16	1,69	27,20	2,19	3,21	4,28	218510
21	1,29	0,64	2,13	2,75	59,55	18,03	8,89	2,27	2,21	1,03	1,01	0,21	14843162
22	0,95	0,26	1,72	4,04	8,79	54,10	16,47	1,31	6,89	2,93	1,92	0,61	200284
23	1,15	0,38	1,12	5,29	3,84	6,05	66,46	1,96	6,79	2,96	3,67	0,34	1219478
24	4,13	0,73	1,45	13,51	5,84	9,20	24,63	4,87	22,22	3,56	6,65	3,22	2434718
31	0,43	0,17	0,20	3,14	0,39	1,19	1,76	0,91	82,92	2,67	5,54	0,69	8183859
32	0,76	0,72	1,58	4,33	0,99	6,54	9,93	1,22	42,58	17,49	12,76	1,09	295125
41	0,49	2,17	1,08	1,63	0,64	1,87	7,09	0,84	13,00	0,69	65,81	4,68	2030
51	1,08	0,67	1,58	4,88	0,23	0,68	0,77	0,99	2,94	0,72	3,54	81,93	193551

Celková přesnost	58,46 %
Kappa index	0,46

Výsledek je obdobný jako u scény ze Švédska. Na řadu tedy přišly úpravy trénovacích ploch. Stejně jako v předchozím případě byly i zde odstraněny třídy 1.4 a 2.4. Spolu s nimi byla smazána i třída 4.1 „Vnitrozemské mokřady“, jež má ve

scéně z České republiky jen minimální zastoupení. Po úpravách byl dosažen následující výsledek.

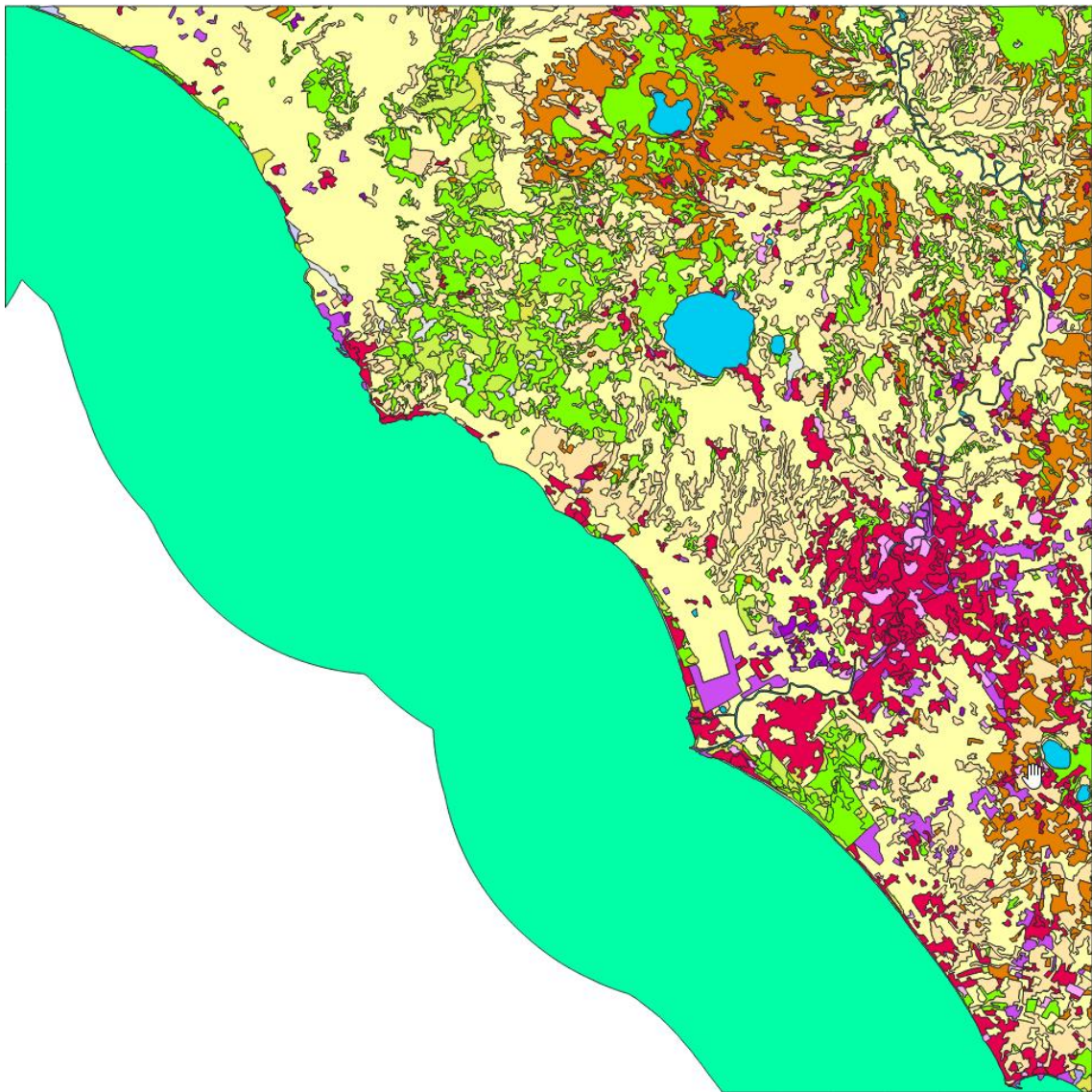
Tab. 20: Kontingenční tabulka (Ml. Boleslav, 2. úroveň) s úpravami CLC

Třída	11	12	13	21	22	23	31	32	51	Počet pixelů
11	94,30	1,86	3,84	0,00	0,00	0,00	0,00	0,00	0,00	10313
12	1,82	79,64	18,31	0,00	0,01	0,00	0,00	0,06	0,16	52845
13	0,04	3,45	95,33	0,05	0,02	0,00	0,00	0,01	1,08	20212
21	0,00	0,54	0,43	93,52	1,73	0,91	2,56	0,31	0,00	7707489
22	0,00	0,57	0,54	3,54	86,92	7,17	0,42	0,84	0,00	74662
23	0,00	0,41	0,15	0,50	1,23	96,12	0,70	0,90	0,00	763934
31	0,00	0,09	0,05	0,36	1,24	1,41	94,50	2,34	0,00	7033552
32	0,00	0,81	0,97	0,18	4,39	0,98	2,32	90,36	0,00	75788
51	0,06	0,60	1,34	0,00	0,00	0,00	0,00	0,00	98,00	70718

Celková přesnost	94,01 %
Kappa index	0,90

Výsledek je obdobný jako u předchozí scény. Lépe se klasifikovala třída 1.2, kterou by ale i tak bylo nejspíše vhodnější spojit s jednou z tříd 1.1 či 1.3.

Poslední scénou byla scéna z okolí Říma. Stejně jako u scény ze Švédska se i zde vyskytuje všech 15 tříd druhé úrovně CORINE. Při klasifikaci první úrovně nebyly brány v potaz třídy spadající pod kategorii 4, tedy „Mokřady“. Pro druhou úroveň byly ale ponechány.



Obr. 26: Řím – druhá úroveň CLC

	11 Městské oblasti		31 Les
	12 Průmyslové, obchodní a dopravní jednotky		32 Keře či bylinná vegetace
	13 Doly, skládky a staveniště		33 Otevřená místa s trochou nebo žádnou vegetací
	14 Umělé nezemědělské oblasti s vegetací		41 Vnitrozemské mokřady
	21 Orná půda		42 Přímořské mokřady
	22 Trvalé plodiny		51 Vnitrozemské vodní plochy
	23 Pastviny		52 Oceány a moře
	24 Různorodé zemědělské oblasti		

Obr. 27: Legenda druhé úrovně CLC (Řím)

Tab. 21: Kontingenční tabulka (Řím, 2. úroveň) bez úprav CLC

Třída	11	12	13	14	21	22	23	24	31	32	33	41	42	51	52	Počet pixelů
11	55,56	5,89	1,33	6,24	1,72	6,65	4,25	3,63	2,51	2,02	4,31	4,62	0,98	0,28	0,00	1324562
12	22,87	26,46	5,07	6,83	4,29	4,03	10,57	2,18	1,31	3,07	7,88	4,05	1,02	0,36	0,01	335845
13	6,49	8,58	24,90	2,55	6,63	5,13	17,47	2,91	3,06	11,79	4,14	4,12	1,40	0,84	0,00	46675
14	11,47	4,38	2,00	16,90	4,74	9,88	9,47	3,70	16,81	5,46	2,25	10,13	1,88	0,95	0,00	106706
21	3,06	1,82	3,00	1,94	42,49	9,68	14,83	4,72	3,66	6,14	1,62	4,53	2,46	0,06	0,00	5673730
22	3,76	0,64	0,74	2,15	4,75	41,22	5,66	3,36	22,10	9,44	0,59	4,87	0,70	0,03	0,00	1359897
23	2,28	1,45	1,49	1,77	5,18	4,47	24,19	2,02	7,98	37,70	2,88	7,27	1,20	0,14	0,00	107818
24	6,48	1,58	1,47	2,60	11,21	18,56	11,48	7,14	18,23	11,24	1,34	7,28	1,34	0,06	0,00	2802449
31	0,41	0,17	0,34	0,56	0,66	2,67	1,34	1,04	78,89	11,08	0,15	2,50	0,13	0,07	0,00	1982258
32	0,41	0,22	0,72	0,49	0,54	1,27	5,25	0,64	24,87	62,86	0,82	1,59	0,28	0,04	0,00	427240
33	3,91	3,71	1,17	0,40	3,40	1,47	15,77	1,18	5,06	43,49	14,71	3,51	0,81	1,38	0,01	73372
41	2,64	0,79	0,59	1,59	4,07	2,61	3,59	1,41	16,20	5,40	1,44	55,56	2,88	1,24	0,00	5409
42	1,50	1,65	0,48	1,40	2,91	2,12	8,04	1,40	1,14	1,82	1,87	3,18	71,32	1,15	0,00	5984
51	0,50	0,28	1,39	0,96	0,71	0,43	0,09	0,28	0,81	0,07	0,16	3,76	1,30	41,39	47,87	251730
52	0,01	0,04	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,02	0,49	99,35	8829954

Celková přesnost	63,34 %
Kappa index	0,54

Byly provedeny stejné úpravy a smazány třídy 1.4 a 2.4 jako v předchozích případech. Po úpravách bylo dosaženo následujících výsledků.

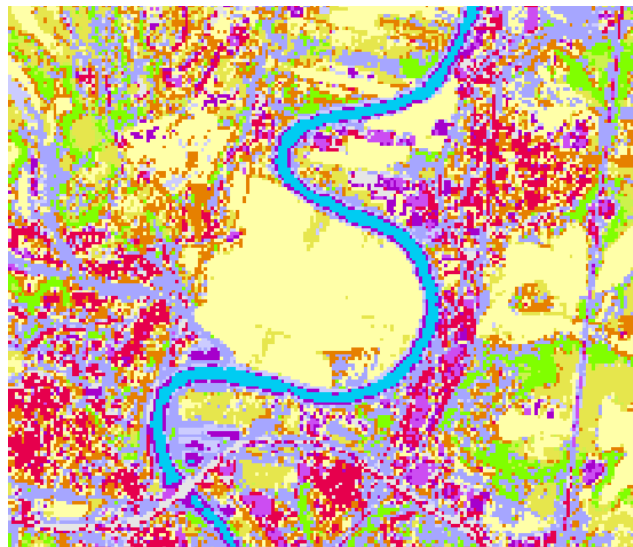
Tab. 22: Kontingenční tabulka (Řím, 2. úroveň) s úpravami CLC

Třída	11	12	13	21	22	23	31	32	33	41	42	51	52	Počet pixelů
11	93,44	1,30	1,66	0,00	0,00	0,00	0,00	0,00	3,59	0,00	0,00	0,00	0,00	98840
12	10,13	80,54	3,27	0,01	0,02	0,00	0,00	0,00	5,94	0,04	0,05	0,01	0,00	52340
13	0,26	2,85	94,03	0,04	0,07	0,00	0,00	0,00	0,50	0,02	0,28	1,94	0,00	4591
21	0,27	0,30	1,20	88,75	1,37	1,74	0,07	0,04	0,45	1,26	4,56	0,00	0,00	1939791
22	0,34	0,12	0,44	2,10	89,16	2,65	1,60	0,29	0,21	1,30	1,80	0,00	0,00	621159
23	0,34	0,24	0,64	1,22	1,05	87,29	0,56	3,14	0,87	1,97	2,68	0,00	0,00	35559
31	0,02	0,02	0,05	1,12	2,92	1,67	89,92	0,78	0,02	3,28	0,19	0,00	0,00	1606272
32	0,02	0,02	0,08	1,81	2,04	1,87	3,99	86,49	0,10	3,03	0,56	0,00	0,00	260011
33	2,41	2,69	4,61	0,26	0,02	0,22	0,00	0,00	89,59	0,10	0,11	0,00	0,00	8226
41	0,27	0,21	0,73	0,76	3,67	2,25	1,94	4,67	1,06	79,31	5,07	0,06	0,00	3296
42	0,34	1,47	1,39	3,00	2,33	3,80	0,84	0,99	0,29	3,59	81,60	0,36	0,00	4760
51	0,03	0,06	3,56	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,77	95,34	0,20	160174
52	0,00	0,01	0,16	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,02	0,86	98,93	98840

Celková přesnost	95,53 %
Kappa index	0,92

I zde jsou výsledky srovnatelné s předchozími scénami. Stejně jako na předchozích scénách, tak i v této scéně by bylo vhodnější spojit třídu 1.2 „Průmyslové, obchodní a dopravní jednotky“ s třídou 1.1 či 1.3. Třídy 4.1 „Vnitrozemské mokřady“ a 4.2 „Přímořské mokřady“ by buď bylo potřeba dále upravovat (ale za cenu dalšího snížení již tak nízkého počtu pixelů těchto tříd), nebo je jako u klasifikace první úrovně CLC úplně vynechat.

Při klasifikaci první úrovně CLC se nepovedla klasifikovat Tibera protékající Římem. Při zavedení třídy pro vnitrozemské vodní plochy se ale již klasifikovala poměrně úspěšně.



Obr. 28: Správně klasifikovaná Tibera

7.3 Klasifikace třetí úrovně CLC

Třetí úroveň CLC tvoří celkově 44 tříd. V překladu do češtiny se jedná o následující třídy.

- Souvislá zástavba (1.1.1)
- Nesouvislá zástavba (1.1.2)
- Průmyslové nebo obchodní jednotky (1.2.1)
- Silniční a železniční sítě a přidružené pozemky (1.2.2)
- Přístavy (1.2.3)
- Letiště (1.2.4)
- Místa těžby nerostů (1.3.1)
- Skládky (1.3.2)
- Staveniště (1.3.3)
- Městská vegetace (1.4.1)
- Sportovní a rekreační zařízení (1.4.2)
- Nezavlažovaná orná půda (2.1.1)
- Trvale zavlažovaná půda (2.1.2)
- Rýžová pole (2.1.3)
- Vinice (2.2.1)
- Ovocné stromy a bobuloviny (2.2.2)
- Olivovníky (2.2.3)
- Pastviny (2.3.1)
- Roční plodiny spojené s trvalými kulturami (2.4.1)
- Různorodé kultivační oblasti (2.4.2)
- Půda převážně zemědělsky obhospodařovaná, s významnými oblastmi přirozené vegetace (2.4.3)
- Zemědělské oblasti pokryté lesním porostem (2.4.4)
- Listnatý les (3.1.1)
- Jehličnatý les (3.1.2)
- Smíšený les (3.1.3)
- Přírodní travní porosty (3.2.1)
- Slatiny a vřesoviště (3.2.2)
- Sklerofylní vegetace (3.2.3)
- Přejídné lesnaté či keřovité oblasti (3.2.4)
- Pláže, duny, písky (3.3.1)
- Holé skály (3.3.2)
- Oblasti s řídkou vegetací (3.3.3)
- Vyhořelé oblasti (3.3.4)
- Ledovce a oblasti s trvalým sněhem (3.3.5)
- Vnitrozemské bažiny (4.1.1)
- Rašeliniště (4.1.2)
- Solné bažiny (4.2.1)
- Saliniště (4.2.2)
- Přílivové oblasti (4.2.3)
- Vodní toky (5.1.1)
- Vodní plochy (5.1.2)
- Pobřežní laguny (5.2.1)
- Estuáry (typ říčního ústí do moře) (5.2.2)
- Moře a oceány (5.2.3)

Na scéně z oblasti s okolím města Kristianstad se nachází celkem 30 tříd třetí úrovně CLC. Mezi nimi i třídy 1.1.2, 1.4.1, 1.4.2, 2.4.2 a 2.4.3, které byly z klasifikace rovnou vyloučeny, jelikož se jedná o nehomogenní třídy skládající se z různých prvků, které na základě spektrálních příznaků nelze rozlišit od ostatních tříd. Dále se zde vyskytuje i pár polygonů třídy 3.3.2 „Holé skály“, u kterých se ale ukázalo, že jsou pokryty vegetací a při úpravách na základě NDVI z trénovacích ploch kompletně zmizely. Do klasifikace tedy vstupovalo celkově 24 tříd.

Jejich úprava byla obdobná jako u první a druhé úrovně CLC. Nejprve byly zavedeny úpravy týkající se NDVI odpovídající hodnotám z tab. 17, které bylo možno aplikovat i na třetí úroveň. Následovalo vyhledávání oblastí, kde docházelo k záměně tříd.

Stejným způsobem byla klasifikována i scéna z České republiky. V ní se nachází celkem 26 tříd třetí úrovně CLC. Stejně jako v předchozí scéně byly i zde odstraněny třídy 1.1.2, 1.4.1, 1.4.2, 2.4.2 a 2.4.3. Spolu s nimi byla odstraněna i třída 3.1.3 „Smíšené lesy“, u které se ukázalo (viz tab. 23 na následující straně), že přesnost klasifikace ji nižší než 50 %. Klasifikováno bylo tedy 20 tříd.

Scénu z okolí Říma tvoří 35 tříd. Smazány byly třídy 1.1.2, 1.4.1, 1.4.2, 2.4.1, 2.4.2, 2.4.3 a 3.1.3 z výše zmíněných důvodů. Zbylo tedy 28 tříd pro klasifikaci. Postup úpravy dat byl obdobný. Nejprve byly provedeny úpravy pomocí NDVI a dále pak odstraněny plochy se záměnou tříd. Ty ale z důvodu vyššího počtu tříd nebyly prováděny třídu po třídě. Místo toho byl rastr obsahující výsledek klasifikace oříznut trénovacími plochami. Následně byl vypočten rozdíl dvou rastrů

$$R_{\text{rozdíl}} = R_{\text{klasifikace}} - R_{\text{trénovací}} \quad (9)$$

kde

$R_{\text{klasifikace}}$ je rastr obsahující výsledek klasifikace,

$R_{\text{trénovací}}$ je rastr trénovacích ploch.

Hodnota 0 pak ve výsledném rastru vyjadřovala správně klasifikované pixely trénovacích ploch a zbylé hodnoty ukazovaly pixely, kde docházelo k záměně. Ty byly následně z trénovacích ploch smazány. (V principu se jedná o stejný postup, který byl použit na ostatních scénách, pouze byl aplikován pro veškeré třídy a na všechny jejich pixely najednou.)

Tab. 23: Kontingenční tabulka (Kristianstad, úroveň 3)

Třída	111	121	122	123	124	131	132	133	211	222	231	311	312	313	322	324	331	411	412	423	511	512	522	523	Počet pixelů
111	95,01	2,03	1,09	0,31	0,62	0,00	0,94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	641
121	3,20	62,92	0,75	7,65	0,37	7,63	11,90	1,51	0,39	0,00	0,02	0,00	0,00	0,02	0,00	0,00	1,89	0,00	0,00	1,74	0,00	0,00	0,02	0,00	5623
122	0,55	2,16	90,48	0,22	0,44	1,72	1,88	1,72	0,06	0,00	0,00	0,00	0,06	0,06	0,00	0,00	0,44	0,00	0,00	0,22	0,00	0,00	0,00	0,00	1806
123	7,70	7,48	0,66	75,58	1,76	3,08	1,76	0,44	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,11	0,00	0,00	0,22	0,00	0,00	0,99	0,22	909
124	0,86	1,15	2,59	1,73	89,91	0,29	2,59	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,86	0,00	0,00	0,00	0,00	0,00	0,00	0,00	347
131	0,14	3,35	0,05	0,45	0,23	91,80	0,91	1,18	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,31	0,00	0,00	0,00	0,00	0,00	0,59	0,00	2208
132	0,26	4,47	0,26	0,00	0,00	0,79	91,58	0,00	0,26	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2,11	0,00	0,00	0,26	0,00	0,00	0,00	0,00	380
133	0,27	3,23	0,40	0,81	0,00	3,36	2,02	88,96	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,81	0,00	0,00	0,13	0,00	0,00	0,00	0,00	743
211	0,01	0,29	0,07	0,03	0,01	0,06	0,56	0,02	88,32	0,07	4,91	0,05	0,03	0,32	0,03	1,01	1,15	1,73	0,32	1,02	0,00	0,00	0,00	0,00	1210356
222	0,00	0,10	0,03	0,00	0,00	0,00	0,03	0,00	1,34	87,59	5,15	1,99	0,00	1,24	0,00	0,31	0,48	0,24	0,10	1,37	0,00	0,00	0,00	0,00	2910
231	0,01	0,19	0,07	0,06	0,01	0,08	0,07	0,01	0,64	0,17	90,45	1,43	0,04	1,99	0,25	1,95	0,69	0,45	0,62	0,84	0,00	0,00	0,01	0,00	308595
311	0,00	0,06	0,05	0,03	0,00	0,10	0,01	0,00	1,37	0,06	1,40	90,40	0,09	3,28	0,23	0,10	0,23	0,06	1,01	1,44	0,00	0,02	0,04	0,00	1328595
312	0,00	0,16	0,04	0,04	0,00	0,19	0,01	0,00	0,76	0,25	0,72	0,21	88,09	3,40	0,02	2,44	0,16	1,84	0,13	1,35	0,00	0,13	0,07	0,00	7816405
313	0,01	0,37	0,14	0,09	0,00	0,27	0,03	0,01	1,82	1,25	1,99	17,63	10,35	47,11	0,36	7,63	0,54	0,93	6,28	2,98	0,00	0,03	0,18	0,00	122988
322	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,11	0,00	0,54	0,11	0,00	1,71	96,15	0,11	0,21	0,00	0,00	1,07	0,00	0,00	0,00	0,00	934
324	0,00	0,09	0,02	0,00	0,00	0,01	0,01	0,01	0,53	0,14	0,84	0,24	3,22	2,11	0,08	90,25	0,12	1,06	0,39	0,88	0,00	0,00	0,00	0,00	862170
331	0,24	1,45	0,00	0,00	0,00	0,72	0,24	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	97,35	0,00	0,00	0,00	0,00	0,00	0,00	0,00	415
411	0,01	0,07	0,01	0,15	0,01	0,79	0,04	0,00	0,72	0,10	0,59	0,45	0,23	1,05	0,00	1,13	0,03	92,91	0,07	0,26	0,00	1,29	0,10	0,00	56616
412	0,00	0,09	0,00	0,07	0,01	0,47	0,06	0,05	0,75	0,02	0,70	1,19	2,84	1,40	0,00	0,16	0,18	1,68	89,00	0,54	0,14	0,63	0,03	0,00	407761
423	0,65	4,53	0,43	1,72	0,22	0,00	1,51	0,00	1,51	0,65	10,34	2,16	0,43	9,05	0,00	0,22	0,65	3,02	0,22	62,72	0,00	0,00	0,00	0,00	464
511	0,00	0,04	0,00	0,09	0,00	0,75	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,00	98,81	0,26	0,00	0,00	2275	
512	0,00	0,24	0,00	0,01	0,00	0,83	0,02	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,05	0,00	0,00	2,06	96,61	0,15	0,02	700071	
522	0,00	0,34	0,00	2,20	0,00	0,68	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,34	96,44	0,00	590
523	0,00	0,20	0,00	0,03	0,00	0,38	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,42	98,49	3225145	

Celková přesnost	90,64 %
Kappa index	0,87

Tab. 24: Kontingenční tabulka (Ml. Boleslav, úroveň 3)

Třída	111	121	122	123	124	131	132	133	211	221	222	231	311	312	321	324	411	412	511	512	Počet pixelů	
111	95,15	1,92	2,13	0,52	0,01	0,21	0,00	0,05	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	7275
121	1,23	82,36	5,08	1,83	1,60	7,45	0,00	0,05	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,13	0,01	0,00	0,24	0,01	0,00	26642
122	2,02	0,74	84,48	7,59	3,84	0,59	0,00	0,39	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,34	0,00	0,00	0,00	0,00	0,00	2030
132	0,00	3,51	2,73	86,74	2,14	1,56	0,00	2,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,17	0,00	0,00	513
124	0,54	2,23	3,27	0,96	91,99	0,96	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	2598
131	0,01	2,30	0,25	0,83	0,44	93,45	0,00	1,94	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,75	0,00	7685
132	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	105
133	0,16	2,23	0,32	0,24	0,48	4,38	0,00	92,20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1257
211	0,00	0,33	0,04	0,15	0,18	0,15	0,00	0,03	94,58	0,01	1,13	0,74	0,00	0,50	0,89	0,56	0,47	0,24	0,00	0,00	0,00	7441376
221	0,00	0,12	0,00	0,04	0,14	0,02	0,00	0,07	2,52	91,52	2,25	1,25	0,39	0,04	1,15	0,42	0,08	0,00	0,00	0,00	0,00	8538
222	0,00	0,46	0,15	0,35	0,32	0,22	0,00	0,03	1,81	1,33	89,43	1,80	0,56	1,05	1,23	0,27	0,08	0,89	0,00	0,00	0,00	48517
231	0,00	0,25	0,02	0,06	0,10	0,08	0,00	0,01	0,34	0,26	0,69	95,55	0,22	0,04	0,55	1,04	0,14	0,66	0,00	0,00	0,00	708748
311	0,00	0,07	0,05	0,13	0,02	0,04	0,00	0,00	0,02	0,02	0,40	0,23	97,50	0,04	0,09	0,91	0,34	0,15	0,00	0,00	0,00	638278
312	0,00	0,04	0,01	0,02	0,01	0,01	0,00	0,00	0,18	0,07	0,53	0,08	2,13	94,39	0,18	1,14	1,06	0,16	0,00	0,00	0,00	3334991
321	0,00	0,20	0,02	0,05	0,09	0,23	0,00	0,01	0,57	0,82	1,34	2,93	0,64	0,00	85,81	4,31	2,91	0,07	0,00	0,00	0,00	35188
324	0,00	0,82	0,33	2,31	0,89	0,71	0,00	0,03	0,20	0,42	2,39	1,38	1,35	0,75	1,87	83,43	1,47	1,65	0,00	0,00	0,00	18269
411	0,00	0,00	0,00	0,39	0,00	0,00	0,00	0,00	0,00	0,65	0,13	5,05	0,91	0,78	0,52	0,65	90,80	0,13	0,00	0,00	0,00	772
412	0,00	1,10	0,22	0,00	0,22	0,66	0,00	0,00	0,66	0,00	0,88	0,44	14,10	0,00	0,00	2,20	1,98	77,53	0,00	0,00	0,00	454
511	0,01	0,12	0,16	3,18	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	95,05	1,45	0,00	15164
512	0,01	0,50	0,06	1,05	0,00	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,05	0,00	0,94	97,21	0,00	55554

Celková přesnost	94,66 %
Kappa index	0,91

Tab. 25: Kontingenční tabulka (Řím, úroveň 3)

Třída	111	121	122	123	124	131	132	133	211	212	221	222	223	231	311	312	321	323	324	331	333	411	421	422	511	512	521	523	Počet pixelů
111	96,23	1,13	0,16	2,02	0,00	0,35	0,01	0,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	68347
121	2,21	80,37	0,05	2,47	0,16	2,89	1,33	9,87	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,65	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	18529
122	4,41	0,28	90,50	1,30	1,18	1,34	0,06	0,75	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	3220
123	2,05	3,19	0,00	93,61	0,00	0,61	0,38	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1315
124	0,70	0,80	2,52	2,91	87,39	1,83	1,16	2,32	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,37	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	13139
131	0,04	1,26	0,00	0,50	0,00	94,91	2,36	0,84	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,08	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2377
132	0,42	3,61	0,11	1,17	0,00	1,48	91,41	1,80	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	943
133	0,83	1,93	0,28	0,83	0,28	2,49	1,66	91,71	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	362
211	0,01	0,09	0,00	0,00	0,00	0,06	0,10	0,01	91,63	0,04	1,37	0,02	1,62	3,37	0,00	0,00	0,00	0,00	0,00	0,00	0,65	0,72	0,30	0,00	0,00	0,00	0,00	0,00	1554955
212	0,00	0,00	0,00	0,00	0,00	0,08	0,00	0,00	4,02	94,59	0,54	0,00	0,39	0,39	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1293
221	0,09	0,04	0,00	0,00	0,00	0,03	0,03	0,01	2,18	0,00	91,74	0,00	3,12	2,03	0,00	0,00	0,00	0,00	0,00	0,03	0,02	0,51	0,17	0,00	0,00	0,00	0,00	0,00	69857
222	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,15	0,00	2,89	90,76	3,95	0,90	0,18	0,00	0,00	0,00	0,04	0,00	0,02	1,10	0,01	0,00	0,00	0,00	0,00	0,00	369901
223	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	1,04	0,01	3,36	0,03	90,49	3,27	0,00	0,00	0,00	0,00	0,00	0,04	0,00	1,68	0,05	0,00	0,00	0,00	0,00	0,00	167896
231	0,00	0,00	0,00	0,00	0,00	0,02	0,01	0,01	2,07	0,01	1,38	0,01	1,27	89,98	0,00	0,00	0,13	0,00	0,02	0,00	2,20	2,69	0,19	0,00	0,00	0,00	0,00	0,00	14522
311	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,34	1,36	0,18	0,37	94,17	0,02	0,00	0,66	1,40	0,00	0,00	1,48	0,00	0,00	0,00	0,00	0,00	0,00	1166894
312	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	2,07	0,00	0,65	0,75	0,23	92,33	0,00	1,14	0,00	0,00	0,00	2,81	0,00	0,00	0,00	0,00	0,00	0,00	43724
321	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,76	0,00	0,01	0,07	0,16	3,39	0,02	0,00	89,20	0,00	2,38	0,00	3,74	0,11	0,16	0,00	0,00	0,00	0,00	0,00	50314
323	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,79	0,09	0,71	1,40	1,86	1,94	0,00	90,03	0,49	0,02	0,18	1,47	0,02	0,01	0,00	0,00	0,00	0,00	39727
324	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,05	0,00	0,08	0,61	0,58	2,55	2,04	0,00	0,97	0,15	91,16	0,00	1,43	0,36	0,01	0,00	0,00	0,00	0,00	0,00	126526
331	1,74	2,37	0,16	0,66	0,00	1,54	0,64	0,62	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	92,26	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	7082
333	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	1,43	0,01	0,20	0,10	1,68	4,78	0,01	0,00	0,38	0,00	0,47	0,00	89,47	1,32	0,13	0,00	0,00	0,00	0,00	0,00	13816
411	0,00	0,35	0,04	0,04	0,00	0,19	0,04	0,00	0,54	0,00	0,82	0,00	1,09	0,93	0,19	0,00	0,00	0,00	0,00	0,00	0,16	95,41	0,16	0,04	0,00	0,00	0,00	0,00	2572
421	0,00	0,00	0,00	0,00	0,00	0,08	0,00	0,00	1,02	0,00	0,53	0,00	0,23	0,68	0,00	0,00	0,00	0,00	0,00	0,00	0,34	0,34	96,67	0,11	0,00	0,00	0,00	0,00	2640
422	0,00	0,80	0,00	0,30	0,00	0,75	0,05	0,00	0,10	0,00	1,11	0,00	0,05	1,11	0,00	0,00	0,00	0,00	0,00	0,05	0,05	0,45	0,00	95,17	0,00	0,00	0,00	0,00	1989
511	0,00	0,00	0,01	1,24	0,00	0,58	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,53	97,64	0,00	0,00	0,00	18121
512	0,00	0,00	0,06	0,01	0,08	0,67	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,89	0,13	98,14	0,00	0,02	138041
521	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	182
523	0,00	0,00	0,00	0,00	0,03	0,30	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	99,66	8712670

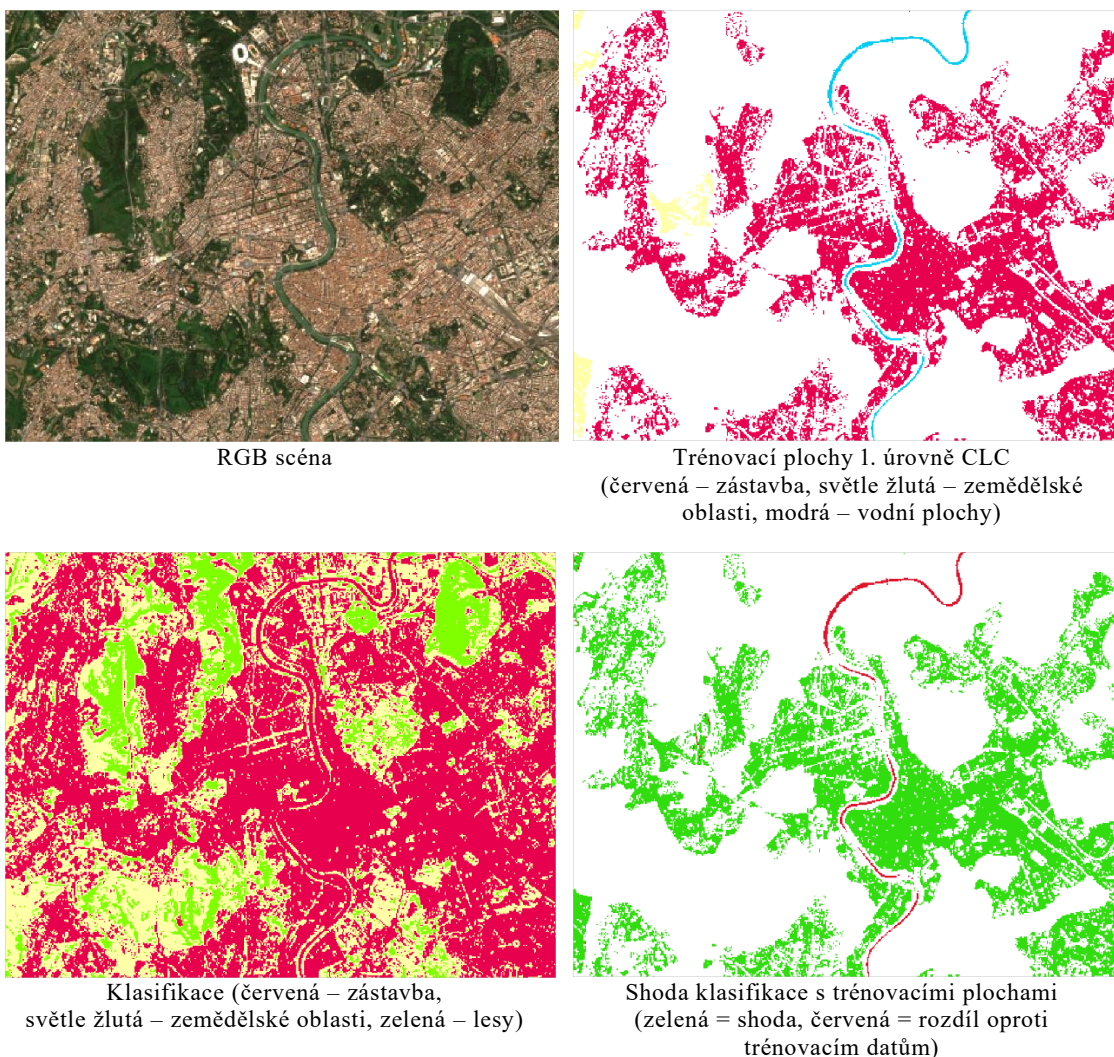
Celková přesnost	97,44 %
Kappa index	0,95

Výsledky ukazují, že i většina tříd na třetí úrovni se dá úspěšně klasifikovat. Menší rozpory v úspěšnosti klasifikace se vyskytují u tříd spadající pod zástavbu. Zde tedy záleží na konkrétní scéně a na tom, jaké prvky se v daném území vyskytují.

Na scéně s okolím Kristianstadu se nepodařilo dostatečně přesně klasifikovat, kromě dvou tříd spadajících pod zástavbu a již dříve zmíněných smíšených lesů, také třídu 4.2.3 „Přilivové oblasti“. V daném území se jedná o jediný polygon a otázkou je, zdali se vůbec jedná o správně zakreslenou či zařazenou plochu, jelikož se nachází ve vnitrozemí. U scény z České republiky jsou hůře klasifikované třídy 3.2.4 „Přechodné lesnaté či keřovité oblasti“, 4.1.2 „Rašeliniště“ a dále pak dvě třídy se zástavbou. Třídy ve scéně s okolím Říma byla až na třídu 1.2.1 „Průmyslové nebo obchodní jednotky“ klasifikovány všechny s vysokou přesností. To ale může být dáno způsobem, jakým byly jednotlivé třídy „pročištěny“.

8 Ukázky výsledků klasifikace

Následující obrázek ukazuje výsledek klasifikace města Říma pro první úroveň CLC. Je vidět, že město bylo klasifikováno správně (kromě řeky Tibery, o které byla zmínka dříve). Městské parky a jiná prostranství s vegetací jsou klasifikovány buď jako lesy (v případě výskytu stromů) anebo jako zemědělské oblasti.



Obr. 29: Klasifikace Říma (1. úroveň CLC)

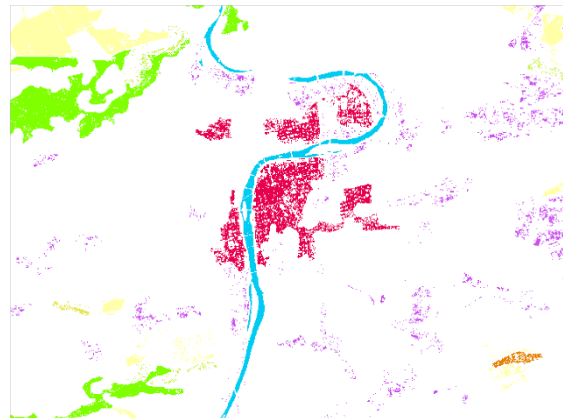
Zajímavé bylo také sledovat vývoj klasifikace městských oblastí pro různé úrovně CLC. Jak se klasifikace pro Prahu postupně měnila, znázorňuje obr. 30. Z důvodu přehlednosti zde není uveden výsledek klasifikace 3. úrovně CLC, kde jednotlivé třídy barevně splývají.



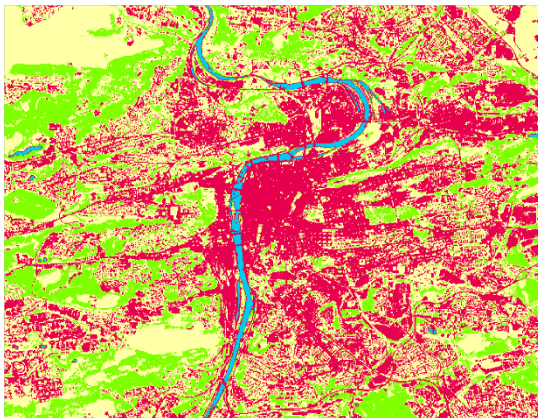
RGB scéna



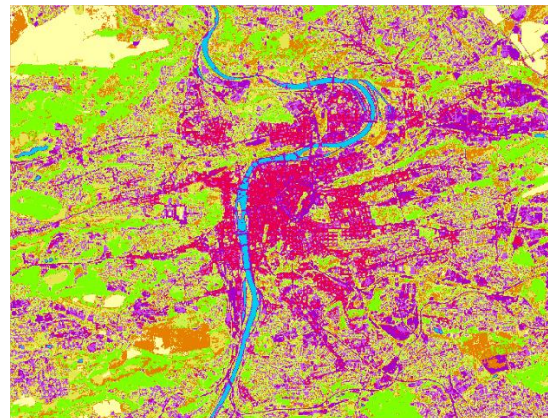
Trénovací plochy, 1. úroveň CLC
 (červená – zástavba, světle žlutá – zemědělské oblasti, zelená – lesy, modrá – vodní plochy)



Trénovací plochy, 2. úroveň CLC
 (červená – zástavba, fialová – průmyslové, obchodní a dopravní oblasti či staveniště, světle žlutá – orná půda, béžová – trvalé plodiny, zelená – lesy, modrá – vodní plochy)



Klasifikace, 1. úroveň CLC
 (červená – zástavba, světle žlutá – zemědělské oblasti, zelená – lesy, modrá – vodní plochy)



Klasifikace, 2. úroveň CLC
 (červená – zástavba, fialová – průmyslové, obchodní a dopravní oblasti či staveniště, světle žlutá – orná půda, béžová – trvalé plodiny, zelená – lesy, modrá – vodní plochy)

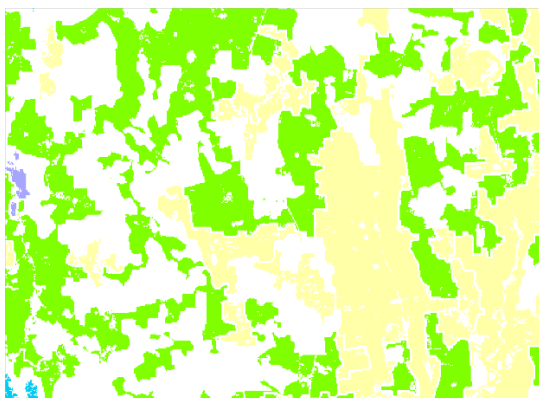
Obr. 30: Vývoj klasifikace Prahy

Z výsledků vyplývá, že klasifikovat samotnou zástavbu není problém. Problém nastává ale při snaze rozeznat různý druh zástavby. U druhé úrovně CLC se již značná část Prahy klasifikovala do tříd 1.2 „Průmyslové, obchodní a dopravní jednotky“ či 1.3 „Doly, skládky a staveniště“ (obě třídy na obr. 30 vpravo dole fialové plochy), což zřejmě neodpovídá úplně skutečnosti. Ukázalo se tak, že nejspíš byla chyba nepoužít pro trénovací plochy podtřídu 1.1.2 „Nesouvislá zástavba“, která by se po úpravách NDVI dala použít pro třídu 1.1 „Zástavba“, čímž by se pravděpodobně následně klasifikovalo do této třídy více prvků.

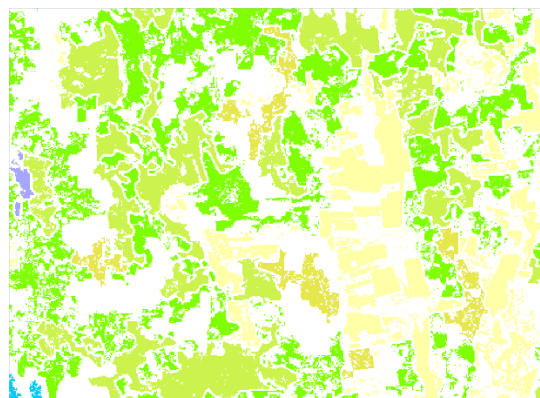
U klasifikace zemědělských a přírodních oblastí k problémům nedocházelo. Příklad klasifikace těchto oblastí je na obr. 32. Klasifikace se vztahuje k následující oblasti na scéně ze Švédska.



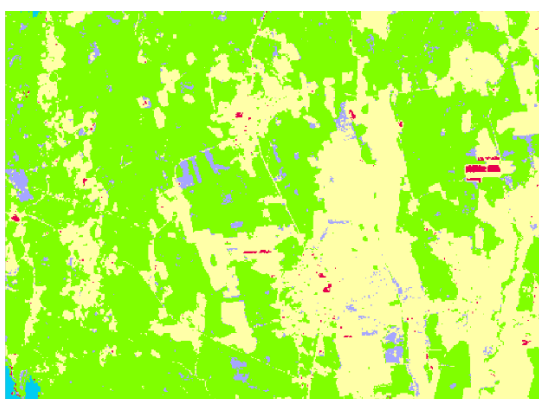
Obr. 31: Scéna se zemědělskými a přírodními oblastmi



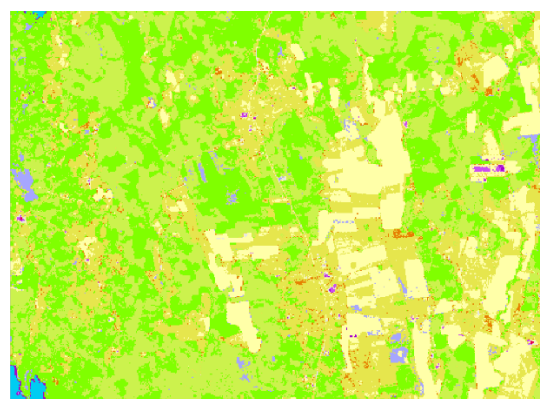
Trénovací plochy (1. úroveň CLC)
 (světle žlutá – zemědělské oblasti, zelená – lesy,
 modrošedá – mokřady, modrá – vodní plochy)



Trénovací plochy (2. úroveň CLC)
 (světle žlutá – orná půda,
 béžová – trvalé plodiny, světle zelené – pastviny,
 zelená – lesy, modrošedá – mokřady,
 modrá – vodní plochy)



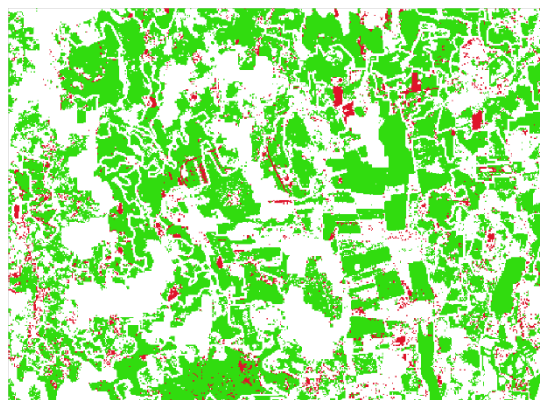
Klasifikace (1. úroveň CLC)
 (červená – zástavba, světle žlutá – zemědělské
 oblasti, zelená – lesy, modrošedá – mokřady,
 modrá – vodní plochy)



Klasifikace (1. úroveň CLC)
 (fialová – průmyslové, obchodní a dopravní oblasti
 či staveniště, světle žlutá – orná půda,
 béžová – trvalé plodiny, světle zelené – pastviny,
 zelená – lesy, modrošedá – mokřady,
 modrá – vodní plochy)



Shoda klasifikace s trénovacími plochami,
 1. úroveň CLC (zelená = shoda, červená = rozdíl
 oproti trénovacím datům)



Shoda klasifikace s trénovacími plochami,
 2. úroveň CLC (zelená = shoda, červená = rozdíl
 oproti trénovacím datům)

Obr. 32: Klasifikace zemědělských a přírodních oblastí

9 Klasifikace pomocí strojového učení

Doposud byla pro všechny klasifikace použita metoda maximum likelihood. Bylo tomu tak proto, protože výpočet maximum likelihood trval pro každou scénu „pouze“ několik minut, zatímco u strojového učení se jednalo o hodiny. V projektu „Geo-harmonizer“ bude ale pro klasifikaci celé Evropy použita právě metoda strojového učení, a proto i v této práci byla vyzkoušena. Výpočty byly ale již prováděny pouze na upravených trénovacích datech použitých pro finální klasifikace jednotlivých scén metodou maximum likelihood.

Ke klasifikaci pomocí strojového učení byl znovu použit program GRASS GIS, do kterého bylo potřeba doinstalovat modul *r.learn.ml* pomocí příkazu *g.extension*. Aby mohl tento modul (patřící mezi tzv. Add-ons) fungovat, je potřeba mít nainstalované i python balíčky „*scikit-learn*“ a „*pandas*“.

Po prvním spuštění se ukázalo, že nástroj není schopen klasifikaci spočítat při použití všech trénovacích ploch jako u maximum likelihood (výpočet skončil chybou, že se nástroj pokoušel alokovat příliš mnoho paměti). Bylo tedy potřeba z trénovacích ploch vybrat jen určitý vzorek, který by byl následně použit pro trénování. K tomu byl použit další Add-on a to modul *r.sample.category*, který umožňuje z rastrové mapy vybrat náhodně určitý zvolený počet pixelů od každé třídy tak, aby jejich zastoupení odpovídalo rozdělení jednotlivých tříd v původních trénovacích datech. Jako celkový počet pixelů určených pro trénovací data klasifikace strojového učení byla zvolena hodnota 2 500 000. Výsledky klasifikace pomocí strojového učení s použitím klasifikátoru *Random Forest* pro všechny úrovně CLC jsou následující.

Tab. 26: Kontingenční tabulka (Kristianstad, 1. úroveň CLC, strojové učení)

Třída	1	2	3	4	5	Počet pixelů
1	100	0,00	0,00	0,00	0,00	4443
2	0,00	100	0,00	0,00	0,00	330541
3	0,00	0,00	100	0,00	0,00	1619771
4	0,00	0,00	0,00	100	0,00	58752
5	0,00	0,00	0,00	0,00	100	486494

Celková přesnost	100 %
Kappa index	1,00

Tab. 27: Kontingenční tabulka (Ml. Boleslav, 1. úroveň CLC, strojové učení)

Třída	1	2	3	5	Počet pixelů
1	100	0,00	0,00	0,00	20346
2	0,00	100	0,00	0,00	1742186
3	0,00	0,00	100	0,00	726837
5	0,00	0,00	0,00	100	10631

Celková přesnost	100 %
Kappa index	1,00

Tab. 28: Kontingenční tabulka (Řím, 1. úroveň CLC, strojové učení)

Třída	1	2	3	5	Počet pixelů
1	100	0,00	0,00	0,00	37121
2	0,00	100	0,00	0,00	652147
3	0,00	0,00	100	0,00	283275
5	0,00	0,00	0,00	100	1527457

Celková přesnost	100 %
Kappa index	1

Tab. 29: Kontingenční tabulka (Kristianstad, 2. úroveň CLC, strojové učení)

Třída	11	12	13	21	22	23	31	32	33	41	51	52	Počet pixelů
11	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	110
12	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2248
13	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	595
21	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	186126
22	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	437
23	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	58584
31	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	1462199
32	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	140070
33	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	60
41	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	72158
51	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	112876
52	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	464538

Celková přesnost	100 %
Kappa index	1,00

Tab. 30: Kontingenční tabulka (Ml. Boleslav, 2. úroveň CLC, strojové učení)

Třída	11	12	13	21	22	23	31	32	51	Počet pixelů
11	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1631
12	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	8357
13	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	3196
21	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	1218806
22	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	11806
23	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	120803
31	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	1112234
32	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	11985
51	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	11183

Celková přesnost	100 %
Kappa index	1,00

Tab. 31: Kontingenční tabulka (Rome, 2. úroveň CLC, strojové učení)

Třída	11	12	13	21	22	23	31	32	33	41	42	51	52	Počet pixelů
11	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	18176
12	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	9625
13	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	844
21	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	356709
22	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	114225
23	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	6539
31	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	295378
32	0,00	0,00	0,00	0,01	0,00	0,00	0,01	99,99	0,00	0,00	0,00	0,00	0,00	47814
33	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	1513
41	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	606
42	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	875
51	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	29454
52	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	1618242

Celková přesnost	100,00 %
Kappa index	1,00

Tab. 32: Kontingenční tabulka (Kristianstad, 3. úroveň CLC, strojové učení)

Třída	111	121	122	123	124	131	132	133	211	222	231	311	312	313	322	324	331	411	412	423	511	512	522	523	Počet pixelů
111	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100
121	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	875
122	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	281
123	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	142
124	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	54
131	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	344
132	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	59
133	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	116
211	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	188424
222	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	453
231	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	48041
311	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	206831
312	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1216830
313	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,02	99,97	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	19146
322	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	145
324	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	134220
331	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	65
411	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	99,99	0,00	0,00	0,00	0,00	0,00	0,00	8814
412	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	63479
423	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	72
511	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	354
512	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	108985
522	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	92
523	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	502079

Celková přesnost	100,00 %
Kappa index	1,00

Tab. 33: Kontingenční tabulka (Ml. Boleslav, 3. úroveň CLC, strojové učení)

Třída	111	121	122	123	124	131	132	133	211	221	222	231	311	312	321	324	411	412	511	512	Počet pixelů		
111	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1472	
121	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	5391
122	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	411
132	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	104
124	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	526
131	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1555
132	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	21
133	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	254
211	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1505869
221	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1728
222	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	99,98	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	9818
231	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	143425
311	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	129165
312	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	674883
321	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	99,99	0,00	0,00	0,00	0,00	0,00	0,00	0,00	7121
324	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	3697
411	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	156
412	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	92
511	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	3069
512	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	11242

Celková přesnost	100,00 %
Kappa index	1,00

Tab. 34: Kontingenční tabulka (Řím, 3. úroveň CLC, strojové učení)

Třída	111	121	122	123	124	131	132	133	211	212	221	222	223	231	311	312	321	323	324	331	333	411	421	422	511	512	521	523	Počet pixelů
111	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	13549
121	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	3673
122	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	638
123	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	261
124	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2605
131	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	471
132	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	187
133	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	72
211	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	308255
212	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	256
221	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	13848
222	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	73329
223	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	99,99	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	33284
231	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2879
311	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	231325
312	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	8668
321	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	9974
323	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	7875
324	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	25083
331	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1404
333	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2739
411	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	510
421	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	523
422	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	394
511	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	3592
512	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	27365
521	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	36
523	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	1727203

Celková přesnost	100 %
Kappa index	1,00

Při klasifikaci pomocí strojového učení se téměř u všech scén a všech úrovní CLC klasifikovalo 100 % trénovacích pixelů do správných tříd. Otázkou zůstává, jak moc jsou tyto výsledky směrodatné a jak by se výsledné přesnosti změnily, kdyby byly použity všechny trénovací plochy a nejen jejich část.

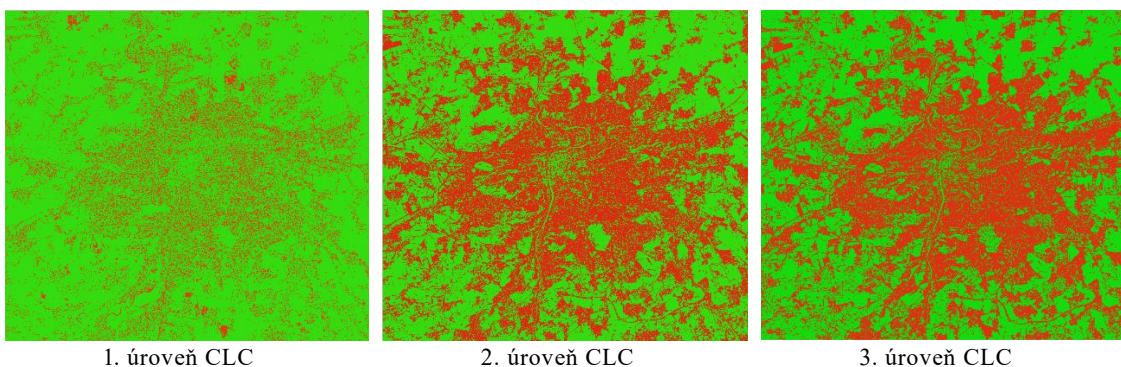
9.1 Porovnání s maximum likelihood

Porovnávat dosažené přesnosti pomocí kontingenčních tabulek není v tomto případě vhodné vzhledem k tomu, že pouze metoda maximum likelihood pracovala s celou množinou trénovacích dat a strojové učení jen s její částí. Proto byly porovnávány pouze výsledné rastry s klasifikovanými třídami. Konkrétně byla zjišťována jejich shoda.

Tab. 35: Shoda rastrů obsahující klasifikované třídy

Scéna	Shoda rastrů [%]		
	1. úroveň CLC	2. úroveň CLC	3. úroveň CLC
Kristianstad	91,23	80,94	75,36
Mladá Boleslav	92,05	76,44	72,16
Řím	96,30	81,00	78,42

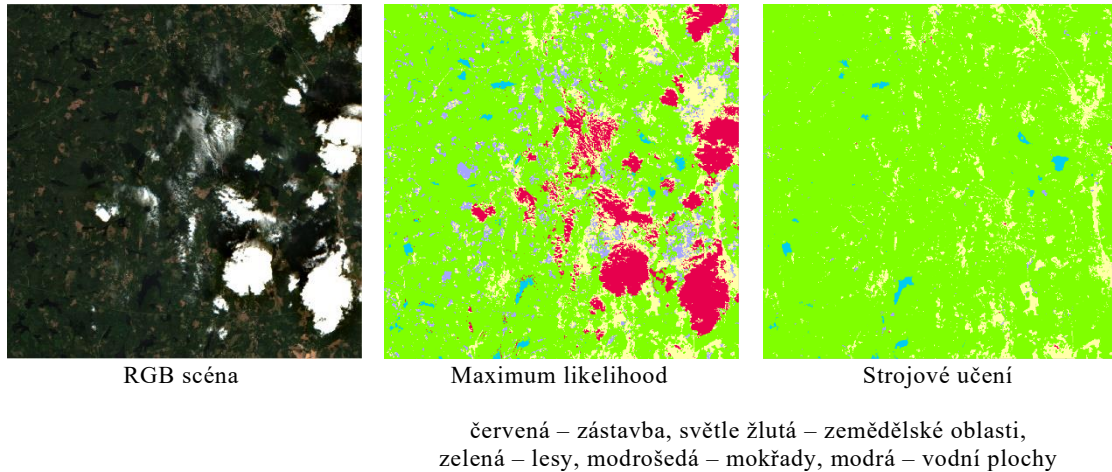
Ukázalo se, že největší rozdíly mezi výslednými rastry se nachází v zastavěných oblastech. Například klasifikace Prahy se již u druhé úrovně CLC shoduje téměř pouze v jejím centru.



zelená = shoda klasifikace maximum likelihood se strojovým učení
 červená = rozdíl mezi oběmi klasifikacemi

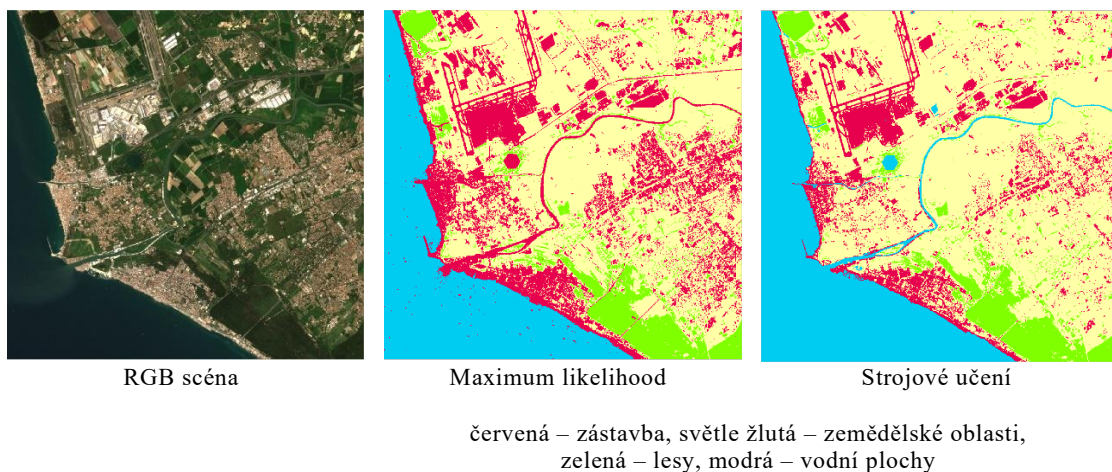
Obr. 33: Rozdílně klasifikovaná Praha

Druhou metodou porovnání byla vizuální kontrola výsledků klasifikace. Metoda strojového učení si dokázala lépe poradit s oblačností ve scéně, a i přes její výskyt dokázala dané území pokryté oblačností správně klasifikovat (výhoda použití dvou scén z různého časového období). Oproti tomu metoda maximum likelihood klasifikovala tyto oblasti jako zástavbu.



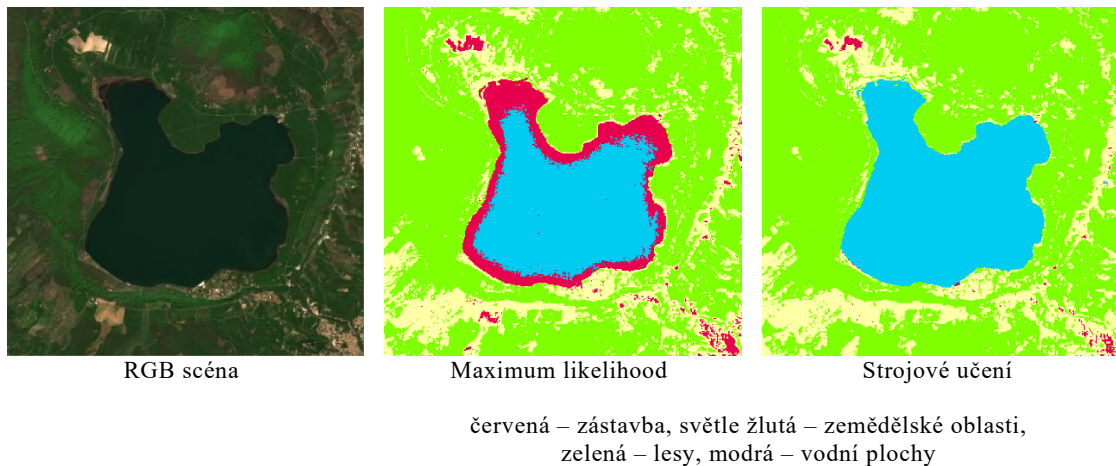
Obr. 34: Porovnání klasifikace s oblačností (Kristianstad, 1. úroveň CLC)

Lepších výsledků dosáhlo strojové učení i při klasifikaci u řeky Tibery, kterou se oproti maximum likelihood podařilo na první úrovni CLC správně klasifikovat jako vodní plochu.



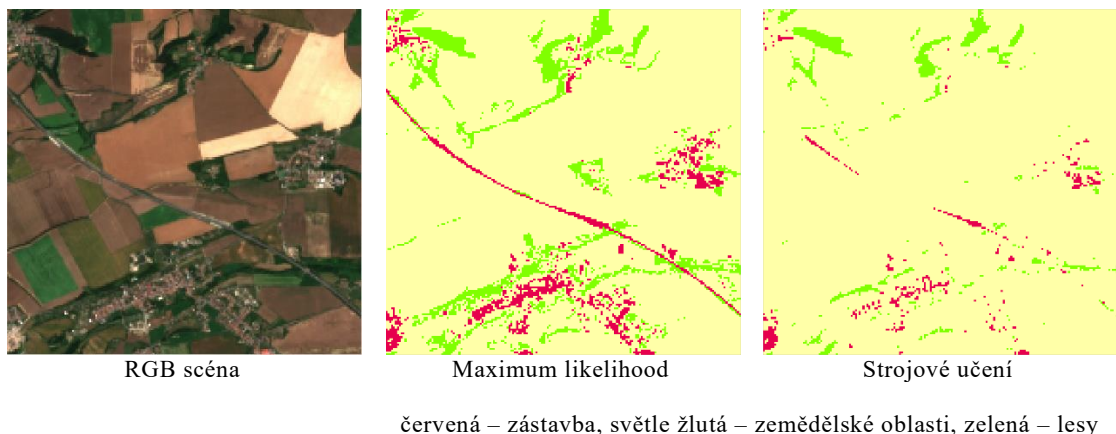
Obr. 35: Porovnání klasifikace Tibery při první úrovni CLC

Lépe si strojové učení poradilo i při klasifikování jezer na první úrovni CLC, u kterých docházelo k nějakému znečištění, zákalu vody nebo většímu výskytu vodních rostlin či řas. V tomto případě se ale nejpravděpodobněji jedná o místa s mělkou vodou, kde naměřené hodnoty nejsou pouze výsledkem odrazivosti vody, ale i dna.



Obr. 36: Porovnání klasifikace jezer (Řím, 1. úroveň CLC)

Na druhé úrovni CLC již k takovýmto chybám nedochází a řeky i jezera jsou správně klasifikovány i metodou maximum likelihood. Kde maximum likelihood poskytovalo lepší výsledky, jsou dopravní komunikace a zástavba. To bude ale nejspíš dáno počtem trénovacích ploch, který byl u těchto tříd pro strojové učení poměrně nízký, a nebyl tak dostatečně reprezentující.



Obr. 37: Porovnání klasifikace dopravní komunikace (ČR, 1. úroveň CLC)

Aby bylo možno oba způsoby klasifikace porovnat i dle přesnosti klasifikace trénovacích ploch, byly náhodně vybrané pixely použity i pro klasifikaci maximum likelihood.

Tab. 36: Kontingenční tab. max. likelihood (2,5 mil. bodů, Kristianstad, 1. úroveň CLC)

Třída	1	2	3	4	5	Počet pixelů
1	99,35	0,63	0,00	0,02	0,00	4443
2	0,91	97,16	0,10	1,84	0,00	330541
3	0,33	2,38	95,93	1,31	0,05	1619771
4	0,71	3,24	0,14	95,26	0,64	58752
5	0,92	0,00	0,07	0,00	99,01	486494

Celková přesnost	96,68 %
Kappa index	0,94

Tab. 37: Kontingenční tab. max. likelihood (2,5 mil. bodů, Ml. Boleslav, 1. úroveň CLC)

Třída	1	2	3	5	Počet pixelů
1	99,91	0,01	0,00	0,08	20346
2	0,78	95,29	3,93	0,00	1742186
3	0,54	3,74	95,72	0,00	726837
5	2,04	0,00	0,00	97,96	10631

Celková přesnost	95,46 %
Kappa index	0,90

Tab. 38: Kontingenční tab. max. likelihood (2,5 mil. bodů, Řím, 1. úroveň CLC)

Třída	1	2	3	5	Počet pixelů
1	99,99	0,01	0,00	0,00	37121
2	1,23	98,73	0,04	0,00	652147
3	0,16	2,27	97,57	0,00	283275
5	0,54	0,00	0,00	99,46	1527457

Celková přesnost	99,06 %
Kappa index	0,98

Tab. 39: Kontingenční tab. max. likelihood (2,5 mil. bodů, Kristianstad, 2. úroveň CLC)

Třída	11	12	13	21	22	23	31	32	33	41	51	52	Počet pixelů
11	97,27	0,91	1,82	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	110
12	6,41	64,46	28,20	0,44	0,00	0,00	0,00	0,00	0,44	0,00	0,00	0,04	2248
13	1,34	3,53	91,76	0,34	0,00	0,00	0,00	0,00	3,03	0,00	0,00	0,00	595
21	0,02	0,65	0,32	89,38	0,59	4,36	0,70	2,36	0,35	1,27	0,00	0,00	186126
22	0,00	0,23	0,00	2,29	89,70	4,58	1,83	1,37	0,00	0,00	0,00	0,00	437
23	0,02	0,41	0,21	1,76	0,57	88,56	0,82	3,60	0,26	3,79	0,00	0,00	58584
31	0,00	0,18	0,31	1,82	0,57	2,31	92,34	1,87	0,22	0,26	0,11	0,00	1462199
32	0,00	0,10	0,04	1,31	0,14	3,98	3,66	87,12	0,15	3,51	0,00	0,00	140070
33	0,00	1,67	0,00	0,00	0,00	0,00	0,00	0,00	98,33	0,00	0,00	0,00	60
41	0,00	0,15	0,73	1,44	0,22	3,27	0,31	0,94	0,03	92,14	0,78	0,00	72158
51	0,00	0,20	0,89	0,00	0,00	0,00	0,04	0,00	0,00	0,00	98,85	0,03	112876
52	0,00	0,04	0,72	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,52	98,72	464538

Celková přesnost	93,19 %
Kappa index	0,89

Tab. 40: Kontingenční tab. max. likelihood (2,5 mil. bodů, Ml. Boleslav, 2. úroveň CLC)

Třída	11	12	13	21	22	23	31	32	51	Počet pixelů
11	97,00	2,08	0,92	0,00	0,00	0,00	0,00	0,00	0,00	1631
12	2,08	78,63	19,12	0,00	0,01	0,00	0,00	0,06	0,10	8357
13	0,72	3,60	94,49	0,09	0,03	0,00	0,00	0,00	1,06	3196
21	0,00	0,52	0,44	93,50	1,72	0,90	2,57	0,35	0,00	1218806
22	0,00	0,45	0,53	3,43	87,09	7,03	0,50	0,97	0,00	11806
23	0,00	0,41	0,15	0,54	1,25	95,98	0,69	0,98	0,00	120803
31	0,00	0,09	0,05	0,36	1,23	1,39	94,49	2,39	0,00	1112234
32	0,00	0,73	0,98	0,13	3,97	1,16	2,31	90,71	0,00	11985
51	0,06	0,64	1,21	0,00	0,00	0,00	0,00	0,00	98,09	11183

Celková přesnost	94,00 %
Kappa index	0,90

Tab. 41: Kontingenční tab. max. likelihood (2,5 mil. bodů, Řím, 2. úroveň CLC)

Třída	11	12	13	21	22	23	31	32	33	41	42	51	52	Počet pixelů
11	93,21	1,21	1,80	0,00	0,01	0,00	0,00	0,00	3,77	0,00	0,00	0,00	0,00	18176
12	7,28	85,65	2,51	0,00	0,00	0,00	0,00	0,00	4,50	0,02	0,03	0,00	0,00	9625
13	0,24	2,84	93,60	0,12	0,00	0,00	0,00	0,00	0,95	0,00	0,12	2,13	0,00	844
21	0,27	0,26	1,32	89,10	1,41	1,97	0,12	0,06	0,45	0,91	4,13	0,00	0,00	356709
22	0,34	0,10	0,50	2,32	88,79	2,60	1,67	0,36	0,22	1,55	1,55	0,00	0,00	114225
23	0,29	0,24	0,57	1,48	1,24	87,37	0,61	3,14	0,81	2,00	2,25	0,00	0,00	6539
31	0,02	0,02	0,06	1,17	2,96	1,66	89,78	0,76	0,02	3,34	0,20	0,00	0,00	295378
32	0,03	0,01	0,11	1,96	2,01	2,01	3,96	86,30	0,13	2,86	0,61	0,00	0,00	47814
33	2,25	2,38	5,02	0,13	0,00	0,40	0,00	0,00	89,62	0,07	0,13	0,00	0,00	1513
41	0,33	0,00	0,17	0,99	3,47	1,16	0,83	2,81	0,83	85,48	3,96	0,00	0,00	606
42	0,11	0,91	1,60	3,31	2,86	2,74	0,91	0,69	0,23	3,43	82,86	0,34	0,00	875
51	0,04	0,04	3,70	0,00	0,00	0,00	0,00	0,00	0,03	0,01	0,61	95,34	0,22	29454
52	0,00	0,01	0,16	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,01	0,89	98,91	1618242

Celková přesnost	95,55 %
Kappa index	0,92

Tab. 42: Kontingenční tab. max. likelihood (2,5 mil. bodů, Kristianstad, 3. úroveň CLC)

Třída	111	121	122	123	124	131	132	133	211	222	231	311	312	313	322	324	331	411	412	423	511	512	522	523	Počet pixelů	
111	98,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100
121	1,83	74,17	0,69	6,63	0,23	8,46	3,09	1,94	0,23	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,03	0,00	0,00	1,71	0,00	0,00	0,00	0,00	0,00	875
122	0,36	1,42	94,66	0,36	0,00	1,07	0,00	2,14	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	281
123	4,23	2,11	0,70	87,32	0,70	3,52	0,70	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,70	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	142
124	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	54
131	0,29	2,62	0,00	0,87	0,29	94,48	0,00	0,29	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,16	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	344
132	0,00	0,00	0,00	0,00	0,00	0,00	96,61	1,69	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,69	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	59
133	0,00	2,59	0,00	0,00	0,00	0,86	0,00	95,69	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,86	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	116
211	0,02	0,43	0,05	0,04	0,00	0,06	0,15	0,04	89,58	0,05	4,67	0,05	0,04	0,33	0,01	1,00	0,59	1,73	0,32	0,85	0,00	0,00	0,00	0,00	0,00	188424
222	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,32	89,40	4,86	1,32	0,00	1,10	0,00	0,22	0,44	0,00	0,00	1,32	0,00	0,00	0,00	0,00	0,00	453
231	0,02	0,27	0,05	0,04	0,01	0,07	0,01	0,02	0,87	0,13	90,73	1,47	0,06	2,02	0,12	1,95	0,37	0,46	0,62	0,71	0,00	0,00	0,00	0,00	0,00	48041
311	0,01	0,09	0,04	0,04	0,00	0,11	0,00	0,00	1,72	0,07	1,50	90,41	0,11	3,53	0,19	0,12	0,19	0,05	1,00	0,81	0,00	0,02	0,00	0,00	0,00	206831
312	0,00	0,17	0,04	0,04	0,00	0,22	0,00	0,00	1,03	0,20	0,85	0,22	87,97	3,73	0,01	2,53	0,11	1,83	0,15	0,74	0,00	0,14	0,00	0,00	0,00	1216830
313	0,01	0,44	0,12	0,10	0,00	0,34	0,00	0,01	2,42	0,85	2,36	17,48	10,39	48,21	0,34	7,56	0,42	0,82	6,33	1,72	0,00	0,06	0,02	0,00	0,00	19146
322	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2,07	97,24	0,00	0,00	0,00	0,00	0,69	0,00	0,00	0,00	0,00	0,00	145
324	0,00	0,08	0,02	0,00	0,00	0,01	0,00	0,01	0,74	0,10	0,88	0,22	3,05	2,26	0,07	90,54	0,08	1,06	0,38	0,50	0,00	0,00	0,00	0,00	0,00	134220
331	0,00	1,54	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	98,46	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	65
411	0,00	0,15	0,01	0,16	0,00	0,75	0,00	0,00	0,94	0,08	0,57	0,41	0,20	1,20	0,00	1,18	0,00	92,98	0,03	0,08	0,00	1,26	0,00	0,00	0,00	8814
412	0,00	0,14	0,00	0,06	0,00	0,52	0,01	0,03	0,86	0,01	0,73	1,14	2,90	1,54	0,00	0,18	0,06	1,71	89,10	0,19	0,06	0,76	0,00	0,00	0,00	63479
423	0,00	1,39	0,00	0,00	0,00	0,00	0,00	0,00	1,39	0,00	1,39	1,39	0,00	5,56	0,00	0,00	0,00	0,00	0,00	1,39	87,50	0,00	0,00	0,00	0,00	72
511	0,00	0,00	0,00	0,00	0,00	0,28	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	99,72	0,00	0,00	0,00	0,00	354
512	0,00	0,23	0,00	0,00	0,00	0,88	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,71	97,03	0,02	0,02	0,00	108985
522	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,09	98,91	0,00	92
523	0,00	0,19	0,00	0,06	0,00	0,47	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,50	0,18	98,59	0,00	502079

Celková přesnost	90,75 %
Kappa index	0,87

Tab. 43: Kontingenční tab. max. likelihood (2,5 mil. bodů, Ml. Boleslav, 3. úroveň CLC)

Třída	111	121	122	123	124	131	132	133	211	221	222	231	311	312	321	324	411	412	511	512	Počet pixelů	
111	94,63	2,17	2,24	0,82	0,00	0,07	0,00	0,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1472
121	1,09	83,81	5,14	0,85	2,08	6,68	0,00	0,06	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,09	0,00	0,00	0,17	0,00	0,00	5391
122	0,73	0,00	85,89	7,06	4,38	0,73	0,00	0,73	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,49	0,00	0,00	0,00	0,00	0,00	411
132	0,00	1,92	2,88	92,31	0,00	0,96	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,92	0,00	0,00	104
124	0,57	3,23	3,23	0,00	92,40	0,57	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	526
131	0,00	2,77	0,19	0,71	0,39	94,15	0,00	1,29	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,51	1555
132	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	21
133	0,39	1,18	0,00	0,00	0,79	3,15	0,00	94,49	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	254
211	0,00	0,34	0,05	0,07	0,19	0,15	0,00	0,02	94,49	0,01	1,29	0,76	0,10	0,50	0,90	0,72	0,31	0,09	0,00	0,00	0,00	1505869
221	0,00	0,06	0,00	0,00	0,12	0,06	0,00	0,00	2,20	92,59	2,20	0,87	0,46	0,00	1,16	0,23	0,06	0,00	0,00	0,00	0,00	1728
222	0,00	0,49	0,12	0,17	0,29	0,21	0,00	0,04	1,77	1,31	89,37	2,29	0,68	1,08	1,29	0,54	0,02	0,32	0,00	0,00	0,00	9818
231	0,00	0,26	0,03	0,04	0,09	0,08	0,00	0,01	0,35	0,24	0,73	95,70	0,42	0,07	0,57	1,11	0,08	0,22	0,00	0,00	0,00	143425
311	0,00	0,07	0,05	0,10	0,01	0,04	0,00	0,00	0,02	0,02	0,39	0,27	97,59	0,05	0,08	0,92	0,27	0,12	0,00	0,00	0,00	129165
312	0,00	0,04	0,01	0,01	0,01	0,01	0,00	0,00	0,17	0,06	0,55	0,09	2,15	94,77	0,20	1,18	0,67	0,08	0,00	0,00	0,00	674883
321	0,00	0,29	0,01	0,03	0,07	0,22	0,00	0,00	0,51	0,77	1,39	2,79	1,00	0,01	86,93	4,41	1,52	0,04	0,00	0,00	0,00	7121
324	0,03	0,81	0,32	1,24	1,41	0,89	0,00	0,00	0,14	0,38	3,33	2,11	3,03	0,73	1,92	82,77	0,51	0,38	0,00	0,00	0,00	3697
411	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	4,49	0,00	1,28	0,00	0,00	94,23	0,00	0,00	0,00	0,00	156
412	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,09	10,87	0,00	0,00	3,26	0,00	84,78	0,00	0,00	0,00	92
511	0,00	0,07	0,23	3,23	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	95,24	1,24	3069
512	0,00	0,56	0,07	1,18	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,01	98,12	11242	

Celková přesnost	94,72 %
Kappa index	0,91

Tab. 44: Kontingenční tab. max. likelihood (2,5 mil. bodů, Řím, 3. úroveň CLC)

Třída	111	121	122	123	124	131	132	133	211	212	221	222	223	231	311	312	321	323	324	331	333	411	421	422	511	512	521	523	Počet pixelů
111	96,58	1,10	0,09	1,81	0,00	0,30	0,00	0,11	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	13549
121	2,04	89,60	0,05	2,61	0,35	2,29	0,65	1,96	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,44	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	3673
122	4,08	0,31	92,48	0,78	0,63	1,10	0,00	0,31	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,31	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	638
123	1,15	2,30	0,00	96,17	0,00	0,00	0,38	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	261
124	0,88	1,61	2,53	2,65	88,56	1,46	0,54	1,27	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,50	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	2605
131	0,00	0,64	0,00	0,21	0,00	95,97	2,97	0,21	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	471
132	0,53	2,67	0,00	0,53	0,00	1,07	94,12	1,07	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	187
133	1,39	0,00	0,00	1,39	0,00	2,78	2,78	91,67	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	72
211	0,00	0,09	0,00	0,00	0,00	0,07	0,10	0,00	91,57	0,02	1,50	0,02	1,63	3,39	0,00	0,00	0,00	0,00	0,00	0,01	0,65	0,73	0,22	0,00	0,00	0,00	0,00	0,00	308255
212	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,95	97,27	0,39	0,00	0,00	0,39	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	256
221	0,07	0,06	0,01	0,00	0,00	0,06	0,04	0,00	2,09	0,00	91,75	0,00	3,08	2,12	0,00	0,00	0,00	0,00	0,00	0,04	0,04	0,54	0,09	0,00	0,00	0,00	0,00	0,00	13848
222	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,14	0,01	2,95	90,79	4,05	0,80	0,20	0,00	0,00	0,00	0,04	0,00	0,03	0,99	0,00	0,00	0,00	0,00	0,00	0,00	73329
223	0,02	0,01	0,00	0,00	0,00	0,01	0,00	0,00	0,97	0,00	3,51	0,01	90,15	3,30	0,01	0,00	0,00	0,00	0,00	0,00	0,04	1,93	0,03	0,00	0,00	0,00	0,00	0,00	33284
231	0,00	0,00	0,00	0,00	0,00	0,03	0,07	0,00	2,12	0,00	1,46	0,00	1,22	90,55	0,00	0,00	0,21	0,00	0,00	0,00	2,33	1,98	0,03	0,00	0,00	0,00	0,00	0,00	2879
311	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,34	1,34	0,20	0,38	94,43	0,02	0,00	0,65	1,40	0,00	0,01	1,21	0,00	0,00	0,00	0,00	0,00	0,00	231325
312	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	2,20	0,01	0,57	0,74	0,28	92,73	0,00	1,20	0,01	0,00	0,00	2,24	0,00	0,01	0,00	0,00	0,00	0,00	8668
321	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,64	0,00	0,01	0,05	0,17	3,28	0,02	0,00	89,18	0,00	2,53	0,00	3,93	0,03	0,16	0,00	0,00	0,00	0,00	0,00	9974
323	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,68	0,09	0,71	1,54	1,64	1,77	0,00	90,83	0,55	0,01	0,08	1,10	0,00	0,01	0,00	0,00	0,00	0,00	7875
324	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	0,09	0,64	0,53	2,58	2,12	0,00	0,88	0,14	91,23	0,00	1,46	0,29	0,01	0,00	0,00	0,00	0,00	0,00	25083
331	2,14	2,64	0,14	0,64	0,07	1,57	0,50	0,43	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	91,88	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1404
333	0,00	0,00	0,00	0,00	0,00	0,00	0,04	0,00	1,42	0,00	0,15	0,11	1,57	4,38	0,00	0,00	0,47	0,00	0,22	0,00	90,51	0,95	0,18	0,00	0,00	0,00	0,00	0,00	2739
411	0,20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,78	0,00	0,98	0,00	0,59	0,59	0,20	0,00	0,00	0,00	0,00	0,00	0,39	95,88	0,20	0,20	0,00	0,00	0,00	0,00	510
421	0,00	0,00	0,00	0,00	0,00	0,19	0,00	0,00	0,76	0,00	0,38	0,00	0,57	0,38	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	97,71	0,00	0,00	0,00	0,00	0,00	523
422	0,00	0,76	0,00	0,00	0,00	0,76	0,00	0,00	0,25	0,00	1,27	0,00	0,25	1,27	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	95,43	0,00	0,00	0,00	0,00	394
511	0,00	0,00	0,00	1,11	0,00	0,56	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,50	97,83	0,00	0,00	0,00	3592
512	0,00	0,00	0,05	0,02	0,06	0,63	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,95	0,15	98,11	0,00	0,02	27365
521	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00	36	
523	0,00	0,00	0,00	0,00	0,03	0,28	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	99,67	1727203

Celková přesnost	97,48 %
Kappa index	0,95

Výsledky klasifikace maximum likelihood pro náhodně vybraných 2 500 000 bodů použitých jako trénovací množiny se vesměs shodují s výsledky, kdy pro trénovací množiny byly použity všechny plochy vzešlé z jejich úprav. V některých případech byla přesnost klasifikace dokonce vyšší. V porovnání se strojovým učením se ale i zde nacházejí třídy, jejichž přesnost klasifikace byla nižší a v tomto ohledu se strojové učení ukázalo jako lepší varianta.

10 Závěr

Cílem práce bylo klasifikovat tři vybraná evropská území s použitím dat CORINE jako trénovací množiny, upravit tyto data tak, aby je bylo možno pro trénovací množiny použít a zjistit, které třídy z nomenklatury CORINE lze úspěšně klasifikovat na základě spektrálních příznaků. Všechny tyto cíle byly úspěšně splněny. Klasifikace pro všechny lokality byla provedena na třech úrovních tematické podrobnosti, kdy první úroveň obsahovala 4–5 tříd a třetí úroveň až 28. Pro klasifikaci byla použita metoda maximum likelihood.

Všechny třídy první úrovně byly klasifikovány s přesností vyšší než 95 %. Bylo tomu i u třídy 4 „Mokřady“, u které se předpokládaly s klasifikací problémy. Úspěšná byla i klasifikace druhé úrovně, ze které byly nejprve odebrány dvě nehomogenní třídy. Jednalo se o třídy 1.4 „Umělé nezemědělské oblasti s vegetací“ a 2.4 „Různorodé zemědělské oblasti“. Ze zbylých tříd se následně hůře klasifikovala pouze třída 1.2 „Průmyslové, obchodní a dopravní jednotky“. U oblasti z okolí Říma se s nižší přesností klasifikovaly i třídy 4.1 „Vnitrozemské mokřady“ a 4.2 „Přímořské mokřady“. Těchto tříd se ale v dané scéně nachází pouze minimum.

Nejnáročnější částí byla úprava dat třetí úrovně CLC a následná klasifikace daných oblastí na této úrovni. Ukázalo se, že i na třetí úrovni CLC je možné většinu tříd klasifikovat s přesností vyšší než 85 %. Třídy s nižší přesností se pro různé oblasti liší. Pro oblast ve Švédsku to jsou třídy 4.2.3 „Přílivové oblasti“, kterou ale tvoří jediný polygon, a 3.1.3 „Smíšený les“, u které byla přesnost klasifikace nižší než 50 %, a nebyla tedy snaha ji následně klasifikovat ve zbylých dvou scénách. Nižších přesností dosáhly také třídy 1.2.1 „Průmyslové nebo obchodní jednotky“ a 1.2.3 „Přístavy“. Třída 1.2.1 byla klasifikována s nižší přesností na všech scénách. Na oblasti z České republiky dosáhly kromě třídy 1.2.1 nižší přesnosti třídy 1.2.2 „Silniční a železniční sítě a přidružené pozemky“, 3.2.4 Přechodné lesnaté či keřovité oblasti a 4.1.2 „Rašeliniště“. Oblast s okolím Říma dosáhla nižší přesnosti pouze u třídy 1.2.1.

Kromě metody maximum likelihood byla vyzkoušena metoda založená na strojovém učení. Klasifikaci bohužel nebylo možno spustit na celé množině trénovacích dat, a proto bylo náhodně vybráno celkem 2 500 000 pixelů. Jednotlivé třídy v nich byly úměrně zastoupeny tak, aby jejich zastoupení odpovídalo rozdělení tříd v původních trénovacích datech. Tímto způsobem byly následně obdrženy

zajímavé výsledky a to takové, že prakticky u všech tříd bylo dosaženo 100% přesnosti klasifikace. Výsledky by bylo ale nejspíš vhodné dále zkoumat, případně pro klasifikaci použít větší množství trénovacích dat a sledovat, zda se budou výsledky klasifikace měnit. Těchto náhodně vybraných 2 500 000 pixelů bylo následně použito i pro klasifikaci maximum likelihood. Výsledky se ale nijak významně nelišily od výsledků, u kterých byly použity veškeré trénovací množiny. Pokud tedy lze porovnat metodu maximum likelihood a strojové učení z obdržených výsledků, lépe vyšla klasifikace strojového učení.

Dalším krokem, který bylo zamýšleno udělat, byla validace dosažených výsledků na testovacích množinách. Pro validaci se chtěly použít data OSM (OpenStreetMap). Ukázalo se ale, že k tomu, aby bylo možné tyto data k validaci použít, je potřeba data OSM dosti upravit (přeložení nomenklatury OSM do CORINE, buffer liniových prvků, výběr vhodných prvků pro validování rastrových dat o velikosti pixelu 20 m × 20 m a další), což bohužel nebylo v silách v čas stihnout.

Klasifikována byla celkem tři evropská území s rozdílnými klimatickými podmínkami a biotopy. Evropu tvoří ale více takovýchto oblastí a pro úspěšné klasifikování celé Evropy bude nutno zpracovat další scény. Jako lepší metoda se ukazuje metoda strojového učení, která na vybraných testovacích datech dosáhla 100% úspěšnosti. Pro úspěšnou klasifikaci je potřeba převzaté trénovací množiny z dat CORINE náležitě upravit, jelikož data CLC jsou omezeny minimální mapovací jednotkou 25 ha, čímž dochází k významné agregaci spektrálně odlišných prvků. Při zavedení úprav na základě NDVI a odstranění záměny mezi třídami tak přesnost klasifikace narostla o desítky procent.

Reference

- [1] Iniciační fond fakulty stavební ČVUT - Aktuality | FSv ČVUT [online]. [cit. 2020-04-13]. Dostupné z: <https://web.fsv.cvut.cz/aktuality/490/>
- [2] O programu Copernicus | Copernicus [online]. [cit. 2020-04-17]. Dostupné z: <https://www.copernicus.eu/cs/o-programu-copernicus>
- [3] Copernicus: Sentinel-2 - Satellite Missions - eoPortal Directory [online]. [cit. 2020-04-17]. Dostupné z: <https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/copernicus-sentinel-2>
- [4] Sentinel-2 - Missions - Sentinel Online [online]. [cit. 2020-04-17]. Dostupné z: <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>
- [5] Sentinel 2 – Datové specifikace - CollGS - Sentinel Collaborative Ground Segment - CollGS [online]. [cit. 2020-04-18]. Dostupné z: <https://collgs.czechspaceportal.cz/sentinel-2-datove-specifikace/>
- [6] Sentinel-2 - Missions - Instrument Payload - Sentinel Handbook [online]. [cit. 2020-04-18]. Dostupné z: <https://sentinel.esa.int/web/sentinel/missions/sentinel-2/instrument-payload>
- [7] LUKA, Václav et al. Vývoj krajinného pokryvu dle CORINE Land Cover na území ČR v letech 1990-2012. [s.l.]: [s.n.], 2017. ISBN 978-80-87770-28-3.
- [8] CORINE Land Cover — Copernicus Land Monitoring Service [online]. [cit. 2020-04-19]. Dostupné z: <https://land.copernicus.eu/pan-european/corine-land-cover>
- [9] Corine Land Cover Classes. In: [cit. 2020-05-01]. Dostupné z: https://land.copernicus.eu/Corinelandcoverclasses.eps.75dpi.png/image_view_fullscreen
- [10] QGIS - GISMentors [online]. [cit. 2020-04-21]. Dostupné z: <https://gismentors.cz/skoleni/qgis/>
- [11] Obsah — Školení QGIS pro začátečníky [online]. [cit. 2020-04-21]. Dostupné z: <https://training.gismentors.eu/qgis-zacatecnik/>
- [12] GRASS GIS - GISMentors [online]. [cit. 2020-04-21]. Dostupné z: <https://gismentors.cz/skoleni/grass-gis/>

- [13] GRASS GIS - General overview [online]. [cit. 2020-04-21]. Dostupné z: <https://grass.osgeo.org/documentation/general-overview/>
- [14] HALOUNOVÁ, L.: Zpracování obrazových dat. ČVUT v Praze, 2008. ISBN: 978-80-01-04253-3
- [15] Image Classification Techniques in Remote Sensing [online]. [cit. 2020-04-25]. Dostupné z: <https://gisgeography.com/image-classification-techniques-remote-sensing/>
- [16] Machine Learning - GeeksforGeeks [online]. [cit. 2020-04-26]. Dostupné z: <https://www.geeksforgeeks.org/machine-learning/#int>
- [17] Introduction to Machine Learning and its Usage in Remote Sensing – Python Tips [online]. [cit. 2020-04-26]. Dostupné z: <https://pythontips.com/2017/11/11/introduction-to-machine-learning-and-its-usage-in-remote-sensing/>
- [18] Random Forest Classification - Towards Data Science [online]. [cit. 2020-04-26]. Dostupné z: <https://towardsdatascience.com/random-forest-classification-and-its-implementation-d5d840d8ead0>
- [19] Kafira, Vasiliki & Albanakis, K. & Oikonomidis, Dimitrios. (2015). Flood Susceptibility Assessment using G.I.S. An example from Kassandra Peninsula, Halkidiki, Greece.
- [20] HASHIM, H., Z. ABD LATIF a N. A. ADNAN. Urban vegetation classification with NDVI threshold value method with very high resolution (vhr) pleiades imagery. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* [online]. 2019, roč. XLII-4/W16, s. 237-240. ISSN 2194-9034. DOI: 10.5194/isprs-archives-XLII-4-W16-237-2019

Seznam obrázků

Obr. 1: Poloha zpracovávaných lokalit	10
Obr. 2: Model družice Sentinel-2, Zdroj: Wikipedia	12
Obr. 3: Pásma Sentinel-2, zdroj ESA, [5]	13
Obr. 4: Jarní scéna oblasti okolo města Kristianstad.....	16
Obr. 5: Jarní scéna z České republiky	17
Obr. 6: Jarní scéna oblasti z okolí Říma.....	18
Obr. 7: Třídy CLC [9]	20
Obr. 8: QGIS logo, zdroj: Wikipedia	21
Obr. 9: GRASS GIS logo, zdroj: Wikipedia	22
Obr. 10: Schéma řízené klasifikace.....	27
Obr. 11: Graf normálního rozdělení, zdroj: Wikipedia.....	27
Obr. 12: Kristianstad – první úroveň CLC.....	32
Obr. 13: Legenda první úrovně CLC	32
Obr. 14: Oblačnost na části letní scény z Kristianstadu	34
Obr. 15: Masky oblačnosti	34
Obr. 16: Neplatná geometrie vektorové vrstvy	40
Obr. 17: Mladá Boleslav – první úroveň CLC.....	42
Obr. 18: Legenda první úrovně CLC	42
Obr. 19: Řím – první úroveň CLC	44
Obr. 20: Legenda první úrovně CLC	44
Obr. 21: Tibera klasifikovaná jako zástavba	46
Obr. 22: Kristianstad – druhá úroveň CLC	47
Obr. 23: Legenda druhé úrovně CLC.....	47
Obr. 24: Mladá Boleslav – druhá úroveň CLC	50
Obr. 25: Legenda druhé úrovně CLC (oblast z ČR).....	51
Obr. 26: Řím – druhá úroveň CLC.....	53
Obr. 27: Legenda druhé úrovně CLC (Řím)	53
Obr. 28: Správně klasifikovaná Tibera	55
Obr. 29: Klasifikace Říma (1. úroveň CLC)	62
Obr. 30: Vývoj klasifikace Prahy.....	63
Obr. 31: Scéna se zemědělskými a přírodními oblastmi	64
Obr. 32: Klasifikace zemědělských a přírodních oblastí.....	65
Obr. 33: Rozdílně klasifikovaná Praha	72
Obr. 34: Porovnání klasifikace s oblačností (Kristianstad, 1. úroveň CLC)	73
Obr. 35: Porovnání klasifikace Tibery při první úrovni CLC	73

Obr. 36: Porovnání klasifikace jezer (Řím, 1. úroveň CLC).....	74
Obr. 37: Porovnání klasifikace dopravní komunikace (ČR, 1. úroveň CLC)	74

Seznam tabulek

Tab. 1: Pásma Sentinel-2, zdroj: Wikipedia, [5]	13
Tab. 2: Použité scény.....	15
Tab. 3: Přehled land cover produktů	29
Tab. 4: Atributová tabulka CLC	31
Tab. 5: Kontingenční tabulka (Kristianstad) bez úprav CLC	33
Tab. 6: Kontingenční tabulka (Kristianstad) po základních úpravách	37
Tab. 7: Úpravy trénovacích ploch	38
Tab. 8: Hodnoty NDVI, [19]	39
Tab. 9: Kontingenční tabulka (Kristianstad) po úpravách NDVI.....	40
Tab. 10: Kontingenční tabulka (Kristianstad) po finálních úpravách.....	41
Tab. 11: Kontingenční tabulka (Ml. Boleslav) bez úprav CLC.....	43
Tab. 12: Kontingenční tabulka (Ml. Boleslav) s úpravami NDVI	43
Tab. 13: Kontingenční tabulka (Řím) bez úprav CLC	45
Tab. 14: Kontingenční tabulka (Řím) s úpravami NDVI	45
Tab. 15: Kontingenční tabulka (Řím) po finálních úpravách.....	45
Tab. 16: Kontingenční tabulka (Kristianstad, 2. úroveň) bez úprav CLC.....	48
Tab. 17: Pravidla pro hodnoty NDVI pro 2. úroveň CLC	49
Tab. 18: Kontingenční tabulka (Kristianstad, 2. úroveň) s úpravami CLC.....	49
Tab. 19: Kontingenční tabulka (Ml. Boleslav, 2. úroveň) bez úprav CLC.....	51
Tab. 20: Kontingenční tabulka (Ml. Boleslav, 2. úroveň) s úpravami CLC.....	52
Tab. 21: Kontingenční tabulka (Řím, 2. úroveň) bez úprav CLC	54
Tab. 22: Kontingenční tabulka (Řím, 2. úroveň) s úpravami CLC	54
Tab. 23: Kontingenční tabulka (Kristianstad, úroveň 3)	58
Tab. 24: Kontingenční tabulka (Ml. Boleslav, úroveň 3).....	59
Tab. 25: Kontingenční tabulka (Řím, úroveň 3).....	60
Tab. 26: Kontingenční tabulka (Kristianstad, 1. úroveň CLC, strojové učení)	66
Tab. 27: Kontingenční tabulka (Ml. Boleslav, 1. úroveň CLC, strojové učení).....	67
Tab. 28: Kontingenční tabulka (Řím, 1. úroveň CLC, strojové učení)	67
Tab. 29: Kontingenční tabulka (Kristianstad, 2. úroveň CLC, strojové učení)	67
Tab. 30: Kontingenční tabulka (Ml. Boleslav, 2. úroveň CLC, strojové učení).....	68
Tab. 31: Kontingenční tabulka (Rome, 2. úroveň CLC, strojové učení).....	68
Tab. 32: Kontingenční tabulka (Kristianstad, 3. úroveň CLC, strojové učení)	69
Tab. 33: Kontingenční tabulka (Ml. Boleslav, 3. úroveň CLC, strojové učení).....	70
Tab. 34: Kontingenční tabulka (Řím, 3. úroveň CLC, strojové učení)	71
Tab. 35: Shoda rastrů obsahující klasifikované třídy	72

Tab. 36: Kontingenční tab. max. likelihood (2,5 mil. bodů, Kristianstad, 1. úroveň CLC)	75
Tab. 37: Kontingenční tab. max. likelihood (2,5 mil. bodů, Ml. Boleslav, 1. úroveň CLC).....	75
Tab. 38: Kontingenční tab. max. likelihood (2,5 mil. bodů, Řím, 1. úroveň CLC)	75
Tab. 39: Kontingenční tab. max. likelihood (2,5 mil. bodů, Kristianstad, 2. úroveň CLC)	76
Tab. 40: Kontingenční tab. max. likelihood (2,5 mil. bodů, Ml. Boleslav, 2. úroveň CLC).....	76
Tab. 41: Kontingenční tab. max. likelihood (2,5 mil. bodů, Řím, 2. úroveň CLC)	77
Tab. 42: Kontingenční tab. max. likelihood (2,5 mil. bodů, Kristianstad, 3. úroveň CLC)	78
Tab. 43: Kontingenční tab. max. likelihood (2,5 mil. bodů, Ml. Boleslav, 3. úroveň CLC).....	79
Tab. 44: Kontingenční tab. max. likelihood (2,5 mil. bodů, Řím, 3. úroveň CLC)	80

Seznam příloh

Příloha č. 1: Trénovací data pro klasifikaci maximum likelihood ve formátu .tif
(v elektronické podobě)

Příloha č. 2: Trénovací data pro klasifikaci metodou strojového učení ve formátu .tif
(v elektronické podobě)

Příloha č. 3: Rastrová data obsahující výsledky klasifikace maximum likelihood ve formátu .tif (v elektronické podobě)

Příloha č. 4: Rastrová data obsahující výsledky klasifikace metodou strojového učení ve formátu .tif (v elektronické podobě)