

Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Katedra geoinformatiky

**POROVNÁNÍ POUŽITELNOSTI DAT DRUŽICE
SENTINEL-2 A PLÉIADES V LOKÁLNÍM
MĚŘÍTKU PRO KLASIFIKACI OBRAZU**

Bakalárska práca

Martina FIRDOVÁ

Vedúci práce Mgr. Pavel VYVLEČKA

Olomouc 2023

Geoinformatika a kartografie

ANOTÁCIA

Táto bakalárska práca sa zaoberá porovnaním komerčných dát z družice Píiades a voľne dostupných dát Sentinel-2 na klasifikáciu záujmovej oblasti povodia Všeminky a Trnávky v lokálnej mierke. Zisťuje sa vhodnosť jednotlivých snímok v kombinácii s rôznymi nastaveniami vybraných klasifikačných algoritmov (K-means, Maximum Likelihood, Random Forest, K-Nearest Neighbors) za prítomnosti aj vegetačných indexov. Presnosť výstupov je hodnotená na základe výpočtu chybovej matice a zároveň porovnávaním klasifikácií s validačnou vrstvou v troch rôznych mierkach. Výsledky môžu byť užitočné pri výbere potrebných dát a algoritmov pri monitorovaní zmien pokryvu.

KLÚČOVÉ SLOVÁ

Diaľkový prieskum Zeme, riadená klasifikácia, neriadená klasifikácia, per-pixel, Sentinel-2, Píiades

Počet strán práce: 33

Počet príloh: 7 (z toho 6 voľné a 1 elektronická)

ANOTATION

This bachelor thesis deals with the comparison of commercial data from the Pleiades satellite and freely available Sentinel-2 data to classify the area of interest of the Všeminka and Trnava river basins at a local scale. The suitability of individual images in combination with different settings of selected classification algorithms (K-means, Maximum Likelihood, Random Forest, K-Nearest Neighbors) is investigated in the presence of vegetation indices as well. The accuracy of the outputs is evaluated by computing an error matrix and also by comparing the classifications with a validation layer at three different scales. The results can be useful in selecting the necessary data and algorithms when monitoring changes in land cover.

KEYWORDS

Remote sensing, supervised classification, unsupervised classification, per-pixel, Sentinel-2, Pléiades

Number of pages 33

Number of appendixes 7

Autorským prohlášením se student explicitně přihlašuje k výhradnímu autorství práce a současně vyjadřuje znalost práv UP Olomouc k tomuto dílu. Odevzdáním práce v digitální podobě do evidence STAG student autentizuje toto prohlášení. V případě, že autor v práci navázal na cizí autorské dílo, musí být tato skutečnost v prohlášení výslovně uvedena.

Prohlašuji, že

- bakalářskou/diplomovou práci včetně příloh, jsem vypracoval(a) samostatně a uvedl(a) jsem všechny použité podklady a literaturu.

- jsem si vědom(a), že na moji bakalářskou/diplomovou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevydělečně, ke své vnitřní potřebě, bakalářskou/diplomovou práci užívat (§ 35 odst. 3),

- souhlasím, že údaje o mé bakalářské/diplomové práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé bakalářské/diplomové práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé bakalářské/diplomové práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

Martina FIRDOVÁ

Podakovanie

Moje podakovanie patrí vedúcemu práce Mgr. Pavlovi Vyvlečkovi za ochotu, trpezlivosť, cenné rady a pripomienky pri vypracovávaní tejto bakalárskej práce.

UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta

Akademický rok: 2021/2022

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Martina FIRDOVÁ**
Osobní číslo: **R210366**
Studijní program: **B0532A330011 Geoinformatika a kartografie**
Téma práce: **Porovnání použitelnosti dat družice Sentinel-2 a Pleiades v lokálním měřítku pro klasifikaci obrazu**
Zadávající katedra: **Katedra geoinformatiky**

Zásady pro vypracování

Cílem práce je provést ve vybraném GIS softwaru dostupné metody klasifikace obrazu nad daty s velmi vysokým rozlišením (Pleiades 1) a ve srovnávací úrovni s daty vysokého rozlišení (Sentinel-2).

Porovnat použitelnost dat, která mají vyšší prostorové rozlišení s daty, které mají daleko nižší prostorové rozlišení, ale obsahují důležitá pásma v oblasti Red Edge a SWIR a jsou tak účinné v identifikaci vegetace. Aplikovanými metodami bude neřízená i řízená klasifikace (per pixel i objektová), zastoupená algoritmy jako Maximum Likelihood, Random Forest, Nearest Neighbour či Support Vector Machine. Pro klasifikaci bude vytvořeno několik vstupních sad, které budou následně klasifikovány. Výsledky klasifikací budou zpracovány a vyhodnoceny pro měřítko 1 : 10 000 – 50 000, dále budou validovány se srovnávacími daty.

Celá práce (text, přílohy, výstupy, zdrojová a vytvořená data) se odevzdá v digitální podobě na paměťovém nosiči (CD, DVD, SD karta, flash disk). Text práce s vybranými přílohami bude odevzdán ve dvou svázaných výtiscích na sekretariát katedry. O diplomové práci student vytvoří webovou stránku v souladu s pravidly dostupnými na stránkách katedry. Práce bude zpracována podle zásad dle Voženilek (2002) a závazné šablony pro diplomové práce na KGI. Povinnou přílohou práce bude poster formátu A2.

Rozsah pracovní zprávy: **max. 50 stran**
Rozsah grafických prací: **dle potřeby**
Forma zpracování bakalářské práce: **tištěná**
Jazyk zpracování: **Slovenština**

Seznam doporučené literatury:


SABINS JR, Floyd F.; ELLIS, James M, 2020. Remote sensing: Principles, interpretation, and applications. Waveland Press.
WEIH, Robert C and RIGGAN, Norman D, 2010. Object-based classification vs. pixel-based classification: Comparative importance of multi-resolution imagery.
THANH NOI, Phan and KAPPAS, Martin, 2018. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors (Basel, Switzerland). Online. 22 December 2018. Vol. 18, no. 1, pp. 18. DOI 10.3390/s18010018.
RIEDL, Martin. Klasifikace lesní vegetace pomocí dat DPZ různého prostorového rozlišení. 2016. Bakalářská práce. Univerzita Karlova, Přírodovědecká fakulta, Katedra aplikované geoinformatiky a kartografie. Vedoucí práce Štych, Přemysl.

Vedoucí bakalářské práce: **Mgr. Pavel Vyvlčka**
Katedra geoinformatiky

Datum zadání bakalářské práce: 5. května 2022
Termín odevzdání bakalářské práce: 5. května 2023

L.S.

doc. RNDr. Martin Kubala, Ph.D.
děkan



prof. RNDr. Vít Voženilek, CSc.
vedoucí katedry

V Olomouci dne 15. září 2022

OBSAH

ZOZNAM POUŽITÝCH SKRATIEK	9
ÚVOD	10
1 CIELE PRÁCE.....	11
2 SÚČASNÝ STAV RIEŠENEJ PROBLEMATIKY	12
2.1 Klasifikácie	12
2.2 Satelity	15
2.3 Spektrálne chovanie objektov.....	16
2.4 Vegetačné indexy	18
2.5 Podobné práce	20
3 METÓDY A POSTUP SPRACOVANIA	22
4 PRÍPRAVA DÁT	24
4.1 Získanie dát.....	24
4.2 Príprava snímok	24
5 KLASIFIKÁCIA	27
5.1 Riadená klasifikácia.....	27
5.2 Neriadená klasifikácia.....	28
6 VÝSLEDKY	29
6.1 Presnosť klasifikácie	29
7 DISKUSIA.....	41
8 ZÁVER	42
POUŽITÁ LITERATÚRA A INFORMAČNÉ ZDROJE	
PRÍLOHY	

ZOZNAM POUŽITÝCH SKRATIEK

Zkratka	Význam
AcATaMa	Accuracy Assessment of Thematic Maps
BOA	Bottom of Atmosphere
BSI	Bare Soil Index
CNES	Centre National D'Etudes Spatiales
ČR	Česká republika
DKV	Detailní kombinovaná vrstva
DPZ	Diaľkový prieskum Zeme
ESA	Europe Space Agency
GIS	Geografický informačný systém
GSD	Ground Sampling Distance
HDR	High Dynamic Range
ISP	Instrument Source Packet
KNN	K-Nearest Neighbor
KVES	Konsolidovaná vrstva ekosystémov
LAI	Leaf Area Index
MSI	MultiSpectral Instrument
NDMI	Normalized Difference Moisture Index
NDTI	Normalized Difference Tillage Index
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
RF	Random Forest
SAVI	Soil-Adjusted Vegetation Index
SNAP	The Sentinel Application Platform

ÚVOD

Diaľkový prieskum Zeme (DPZ) hrá dôležitú úlohu v monitorovaní, analýze a taktiež predikcii katastrof na povrchu planéty. Využíva k tomu pokročilé technológie snímania, skenovania a spracovania snímok v rôznych úrovniach. Jednou z dôležitých disciplín DPZ je klasifikácia obrazu, ktorá sa venuje triedeniu pixelov v obraze do požadovaných kategórií na základe ich podobných vlastností, a je dôležitou súčasťou analýzy povrchového pokryvu v mnohých sférach, či už v poľnohospodárstve na rozpoznávanie plodín, sledovanie zdravotného stavu vegetácie, sledovanie environmentálnej situácie, alebo v ekonomickej sfére pri mestskom plánovaní.

Cieľom tejto bakalárskej práce je porovnať výstupy klasifikácie dát z komerčných a voľne dostupných dát družíc s rozličným rozlíšením a zistiť ich použiteľnosť pri potrebe klasifikovania zemského povrchu do šiestich základných kategórií: ihličnatý les, listnatý les, trávnatá plocha, vodná plocha, zastavaná oblasť a orná pôda. Na dosiahnutie najlepších možných výsledkov bola vyskúšaná riadená aj neriadená klasifikácia, ako Maximum Likelihood, Random Forest, K-Nearest Neighbor a K-means cluster analysis s rôznymi parametrami. K zlepšeniu výsledkov je možné použiť vegetačné indexy, ktoré v tejto práci boli otestované.

1 CIELE PRÁCE

Cieľom bakalárskej práce je previesť vo vybranom GIS software dostupné metódy klasifikácie nad dátami s veľmi vysokým rozlíšením (Pléiades 1) a v porovnávanej úrovni s dátami vysokého rozlíšenia (Sentinel-2). Porovnať použiteľnosť dát, ktoré majú vyššie priestorové rozlíšenie s dátami, ktoré majú ďaleko nižšie priestorové rozlíšenie, ale obsahujú dôležité pásma v oblasti Red Edge a SWIR a sú tak účinné v identifikácii vegetácie. Aplikovanými metódami je riadená aj neriadená klasifikácia, zastúpená algoritmi Maximum Likelihood, Random Forest, K-Nearest Neighbors a K-means. Pre klasifikáciu boli vytvorené 4 vstupné sady, ktoré boli následne klasifikované. Výsledky práce sú spracované a vyhodnotené pre mierku 1 : 10 000, 1 : 20 000 a 1 : 50 000, na záver validované s porovnávacími dátami. Výstupnou formou sú tabuľky a tematické mapy.

2 SÚČASNÝ STAV RIEŠENEJ PROBLEMATIKY

Téme klasifikácia povrchu sa venovalo už množstvo prác. Pre vytvorenie prehľadu bola urobená rešerš, v rámci ktorej boli podobné práce naštudované, rovnako ako aj odborná literatúra na lepšie porozumenie danej problematiky pri nasledovnej praktickej časti.

2.1 Klasifikácie

Jednou z najdôležitejších funkcií diaľkového prieskumu je tvorba máp využitia krajiny a krajinnej pokrývky, ktorá sa tvorí procesom klasifikácie obrazu. Prebieha to na princípe zaraďovania všetkých pixelov na snímku do určitých kategórií, ktoré sa často vytvárajú podľa rozpoznávania určitých vlastností a spektrálnych vzorov, napr. pixely s podobnou spektrálnou odrazivosťou či pohlcovaním sú zoskupené a predpokladá sa, že tvoria spoločnú kategóriu – spektrálnu triedu. K spektrálnej triede sa následne priraduje informačná trieda (Campbell, Wynne 2011; Lillesand, Kiefer, Chipman 2015; Al-doski, Mansor, Shafri 2013a).

Konečným výsledkom klasifikácie obrazu je mapa s klasifikačnou schémou, čo je legenda mapy popisujúca jednotlivé kategórie, respektíve informačné triedy (Dobrovolný 1998).

Objektová klasifikácia vs. per-pixel klasifikácia

Bežne používaná per-pixel klasifikácia je založená na priamom klasifikovaní jednotlivých pixelov z celého snímku.

Objektová klasifikácia prebieha v dvoch krokoch. Najprv dochádza k segmentácii snímku do zmysluplných objektov, ktoré vznikajú zlučovaním susediacich pixelov, kým nevyformujú väčšie segmenty. Najmenší segment je definovaný pomocou prahu (threshold) užívateľom. Až potom nasleduje klasifikácia vytvorených segmentov, ktorá nepozostáva iba z klasifikovania na základe spektrálnych vlastností, ale taktiež zohľadňuje informácie o tvare, kontexte, či textúre (Jixian, Zhengjun).

V článku *Assessing object-based classification: advantages and limitations* hľadali *Desheng Liu* a *Fan Xia* nedostatky objektovej klasifikácie, kde vytýčil problém pri segmentácii, kde boli pixely viacerých objektov spojené do jedného objektu, čím nedošlo k správne mu zobrazeniu obrazu, alebo v tvaroch finálnych objektov pri tvorení menších skupín, ktoré sa nezhodujú s reálnymi objektami na zemskom povrchu. Naopak, z pozitívnych vlastností spomenul odstránenie efektu „salt-and-pepper“, ktorý vzniká pri klasifikovaní jednotlivých pixelov a branie ohľadu na vlastnosti objektu.

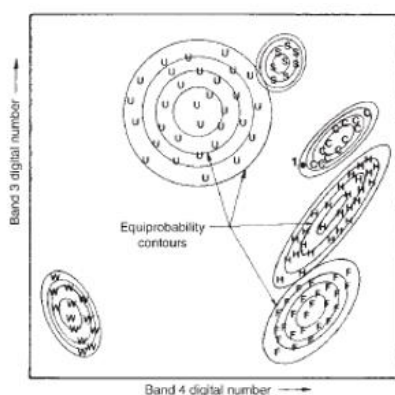
Riadená klasifikácia

Pri riadenej klasifikácii je v prvej fáze potrebný zásah spracovávateľa na vytvorenie tréningových plôch. Tie sú definované identifikovaním oblastí na snímku, ktoré je možné jednoznačne roztriediť do vybraných informačných tried. Tieto oblasti by mali mať typické spektrálne vlastnosti, na základe čoho sa vytvárajú homogénne skupiny pixelov. Pre dosiahnutie čo najkvalitnejších tréningových plôch je potrebné dbať na dostatočnú veľkosť, rozloženie po celom území pre záber rôznych možností výskytu a dostatočnú znalosť územia. Campbell uvádza ako minimálny počet 100 pixelov pre každú informačnú triedu, aby došlo k výpočtu reprezentatívnych štatistických charakteristík. Pixely z tréningových plôch by sa nemali nachádzať na rozmedzí dvoch rôznych kategórií (Campbell, Wynne 2011).

Presnosť klasifikácie následne závisí od vhodne zvoleného klasifikačného algoritmu, ktorých existuje niekoľko.

- Klasifikátor maximálnej pravdepodobnosti (Maximum likelihood)

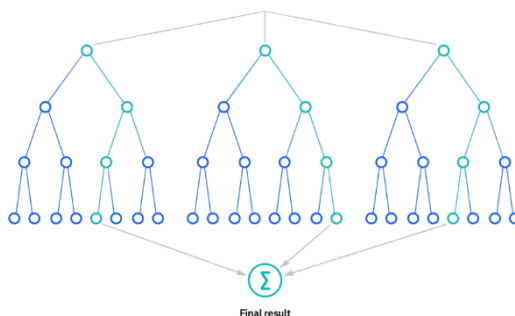
Klasifikátor pri rozrazení pixelu hodnotí rozptyl, koreláciu a kovarianciu všetkých vytvorených tried. Na to je potrebné normálne (Gaussove) rozdelenie tréningových dát. Pomocou pravdepodobnostnej funkcie sa počíta pravdepodobnosť, s ktorou neznámy pixel patrí do každej kategórie. Zaradený je do tej, v ktorej získal najvyššiu hodnotu. Spracovávateľ má možnosť nastaviť prahovú hodnotu (threshold), pod ktorú pravdepodobnosť patrenia pixelu do danej kategórie nesmie klesnúť, v prípade, že sa tak stane, pixel ostáva neklasifikovaný. (Lillesand, Kiefer, Chipman 2015) Ako nevýhoda daného klasifikátora sa dá brať výpočetná náročnosť a podmienka veľmi dobrých tréningových dát, čo môže byť problémom, nakoľko v realite nie každá informačná trieda zahŕňa iba jednu spektrálnu triedu (Sathya, Deepa 2017).



Obrázok 1 Klasifikátor Maximálnej Pravdepodobnosti (Dobrovolný, 1998)

- Random forest

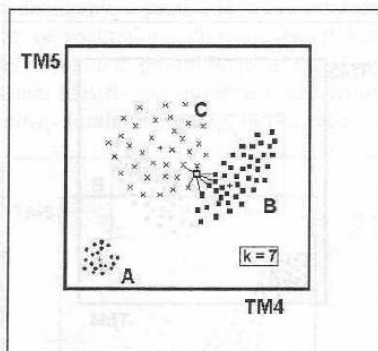
Klasifikátor súborového učenia (ensemble learning), ktorý je založený na mnohých klasifikačných rozhodnutiach pomocou „rozhodovacích stromov“ (decision trees) nezávislých na sebe. Na základe vlastností pixelov v tréningových plochách v jednotlivých kategóriách, ako napríklad spektrálna odrazivosť, nadmorská výška, či sklon, sa rozhoduje o zatriedení neklasifikovaného pixelu všetkými decision trees z celého forestu, ktorého veľkosť sa predom zadáva, kde v každom jednom dochádza k hlasovaniu zaradenia. Finálna trieda je tá, ktorá dostala najviac hlasov. Čím viac hlasov je za jednu triedu, tým je väčšia pravdepodobnosť správneho klasifikovania pixelu (Jensen 2005).



Obrázok 2 Klasifikátor Random Forest (zdroj: <https://www.ibm.com/topics/random-forest>)

- Klasifikátor K – najbližších susedov (K – nearest neighbors)

Je jednoduchý, ale efektívny klasifikátor, ktorý je v podstate modifikovaná metóda minimálnej vzdialenosti, avšak neurčuje triedu len na základe najkratšej vzdialenosti pixelu od centroidu zhluku, ale taktiež berie ohľad na pixely v okolí. V prvom rade KNN prešetruje každý neznámy pixel, identifikuje jeho k – najbližších susedov (počet predom stanovený užívateľom) a následne sa zaraďuje do triedy, z ktorej pochádza väčšia časť pixelov. Klasifikátor je výpočetne náročný pri veľkej dátovej sade (Campbell, Wynne 2011).



Obrázok 3 Klasifikátor K-Nearest Neighbor (Dorbovolný, 1998)

- Klasifikátor minimálnej vzdialenosti (Minimum distance)

Klasifikátor, ktorý funguje na základe porovnávania vzdialeností neznámeho pixelu a stredmi jednotlivých zhlukov tréningových dát, ktoré sú často definované priemernou hodnotou zhluku. Klasifikátor vždy vyberá minimálnu zhluk s najmenšou vzdialenosťou od pixelu. Výsledky nie sú vždy presné, triedy niekedy prekračujú ich reálne hranice, z tohto dôvodu je táto metóda v diaľkovom prieskume používaná minimálne (Campbell, Wynne 2011).

Neriadená klasifikácia

Na rozdiel od riadenej klasifikácie, v prípade neriadenej klasifikácie nie je v úvodnom štádiu potrebný zásah spracovávateľa, nakoľko tréningové plochy nie sú využívané. Neriadená klasifikácia spočíva v agregovaní neznámych pixelov do skupín na základe ich tendencie vytvárania zhlukov vo viacrozmernom priestore. Predpokladom pre to je spektrálne chovanie objektov, pri ktorom objekty s rovnakým povrchom majú podobné spektrálne vlastnosti, zatiaľ čo líšiace sa povrchy majú aj iné spektrálne chovanie a sú dobre separované. Tento proces zoskupovania sa vykonáva zhlukovou analýzou, ktorej výsledkom sú spektrálne triedy, ktorých množstvo definuje užívateľ. Po tomto kroku je potrebný jeho zásah, kde sa na základe referenčných dát priraduje k spektrálnym triedam informačná trieda a prípadne dochádza k zlučovaniu tried. (Dobrovolný 1998)

Medzi najčastejšie používané algoritmy neriadenej klasifikácie patria K-means a ISODATA.

- Klasifikátor K-priemeru (K-means)

Tento algoritmus po určení požadovaného množstva vytvorených zhlukov lokalizuje ich jednotlivé centroidy, a následne je každý pixel na snímku priradený k zhluku, ktorého stredný vektor sa nachádza najbližšie. Proces sa opakuje

vypočítaním nových vektorov na novo vytvorených zhlukoch a reklasifikáciou obrazu, kým neprestane dochádzať k významným zmenám. Nasleduje priradenie informačných tried (Lillesand, Kiefer, Chipman 2015).

- ISODATA (Iterative Self-Organizing Data Analysis Techniques A)

Algoritmu sa rovnako zadáva požadované množstvo zhlukov a zároveň počet iterácií. Stredy zhlukov sú distribuované rovnomerne v priestore a následne dochádza k priradeniu pixelov podľa najmenšej vzdialenosti k stredu zhluku. Proces sa opakuje kým nie je uspokojivý výsledok bez výrazných zmien. (NV5 GEOSPATIAL 2023)

2.2 Satelity

Pléiades

Zoskupenie dvoch identických družíc, Pleiades-1A a Pleiades-1B, patrí pod francúzsku vesmírnu agentúru CNES (Centre National D'Etudes Spatiales). Misia začala pre družice samostatne v decembri roku 2011 a 2012 a ponúka optické dáta veľmi vysokého priestorového rozlíšenia – 50 centimetrov pre panchromatické pásmo a dva metre pre multispektrálne pásma. Nachádzajú sa na rovnakej orbite ako satelity SPOT ako súčasť satelitnej konštelácie.

Pléiades sú družice synchronné so Slnkom obiehajúce vo výške 694 kilometrov a šírka snímanej plochy je 20 kilometrov. Ich cyklus trvá 26 dní. Prioritne sa z nich získavajú environmentálne dáta (monitorovanie vegetácie, hydrológia, risk a katastrofy, mapovanie,...). Spektrálne pásma:

- B0 – modré (450-530 nm)
- B1 – zelené (510-590 nm)
- B2 – červené (620-700 nm)
- B3 – blízke infračervené (775-915 nm)
- PAN – panchromatické (480-820 nm) (Earth Online – Pleiades)

Sentinel

Sentinel-2 spadá pod misiu pod programom Copernicus , ktorý je vedený Európskou vesmírnou agentúrou (ESA). Do misie sa radia:

- Sentinel-1: sprostredkováva radarové snímky dostupné bez ohľadu na počasie
- Sentinel-2: optické snímky s vysokým rozlíšením, zamerané na vegetáciu, vodu či pôdu
- Sentinel-3: optické, radarové a výškové dáta, zameriava sa na topografiu, teplotu povrchu
- Sentinel-4: monitoring zloženia atmosféry, stopové plyny a aerosoly vyskytujúcich sa hlavne nad oblasťou Európy
- Sentinel-5: monitoring atmosféry
- Sentinel-6: meranie výšky morskej hladiny

V tejto práci sú využité multispektrálne dáta misie Sentinel-2. Tvoria ju dve družice Sentinel-2A a Sentinel-2B umiestnené na rovnakej slnečno-synchronnej dráhe s posunom 180° vo výške 786 kilometrov a operujú simultánne s dĺžkou cyklu 10 dní. Vypustené na orbitu boli samostatne, Sentinel-2A v júni 2015 a Sentinel-2B v marci 2017. Družice majú vysoké priestorové rozlíšenie s 13 spektrálnymi pásmami, ktoré zaisťuje MultiSpectral Instrument (MSI), senzor nachádzajúci sa v družici. Z dát misie Sentinel-2 sa podobne ako pri družiciach Pléiades získavajú informácie o zemskom povrchu, poľnohospodárstve

a lesníctve, monitorujú sa ekosystémy, bezpečnosť a riziká katastrof a pod. v oblasti od 56° južnej šírky po 83° severnej šírky. (Sentinel-2 User Handbook 2015)

Vlnové dĺžky jednotlivých pásiem a ich priestorové rozlíšenie sa nachádza v Tabuľke 1.

Tabuľka 1 Spektrálne pásma Sentinel-2 (zdroj: Satellite Imaging Corporation, úprava: autor)

Pásma	Stredná vlnová dĺžka (nm)	Priestorové rozlíšenie (m)
B1 - Coastal aerosol	443	60
B2 - Blue	490	10
B3 - Green	560	10
B4 - Red	665	10
B5 - Vegetation Red Edge	705	20
B6 - Vegetation Red Edge	740	20
B7 - Vegetation Red Edge	783	20
B8 - NIR	842	10
B8A - Vegetation Red Edge	865	20
B9 - Water vapour	945	60
B10 - SWIR - Cirrus	1375	60
B11 - SWIR	1610	20
B12 - SWIR	2190	20

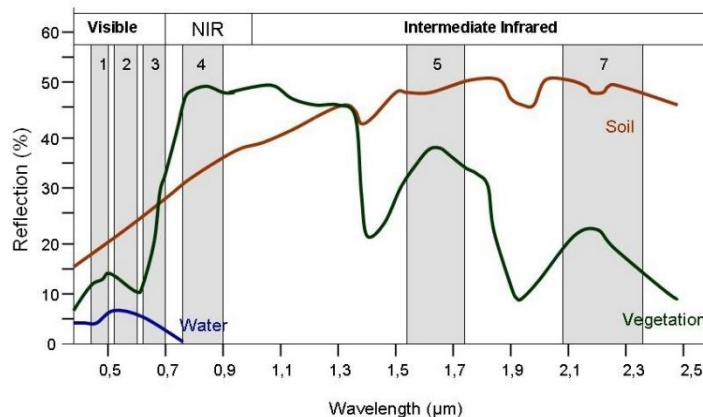
Snímky Sentinelu sa vyskytujú v piatich rôznych úrovniach, z toho sú voľne dostupné tri. Level-0 sú komprimované neupravené dáta vo formáte Instrument Source Packet (ISP). Level-1A sú dáta získané dekomprimovaním dát z Level-0, geometrický model je vyvinutý a jednotlivé pixely sú lokalizovateľné. Súradnice sú viazané na stredy pixelov. Dáta z týchto dvoch úrovní nie sú dostupné užívateľom. Level-1B poskytuje dáta po radiometrických korekciách vplyvov hornej časti atmosféry (TOA) v geometrii senzoru. Level-1C udáva dáta o odrazivosti na hornej časti atmosféry. Sú prevzorkované metódou Ground Sampling Distance (GSD) na rozlíšenie 10, 20 a 60 metrov, v závislosti od rozličných spektrálnych pásiem. Súradnice sa viažu na ľavý horný roh každého pixelu. Level-2A poskytuje dáta o odrazivosti na spodnej časti atmosféry (BOA) (Žubrietovský a kol. 2017).

2.3 Spektrálne chovanie objektov

Na základe získaných hodnôt spektrálnej intenzity emitovaného, alebo odrazeného žiarenia, je možné určiť vybrané fyzikálne vlastnosti meraného objektu. Druhovými parametrami sa vyjadruje druh látky, ako napríklad voda či les. Na určenie fyzikálneho stavu objektu (vlhkosť) sa používajú stavové parametre. Pri dostatočnej znalosti spektrálnej charakteristiky objektov a ich zodpovedajúcich stavových parametrov je možné určiť

správne spektrálne pásma na čo najlepšiu charakteristiku objektov (Kolář 1990). Ako hlavné rozdelenie sa na základe najväčších spektrálnych odlišností môže brať delenie na skupiny pevný povrch bez vegetácie, vegetačný povrch a vodu (Halounová, Pavelka, 2007).

Typické spektrálne chovanie objektov sa sleduje tzv. spektrálnou krivkou odrazivosti, ktorá má svoj špecifický priebeh pre každú skupinu povrchov (Dobrovolný 1998).



Obrázok 4 Spektrálna krivka odrazivosti pre vegetačný povrch, vodu a pôdu (zdroj: <https://seos-project.eu/remotesensing/remotesensing-c01-p06.html>).

Vegetačný povrch

Najčastejším prvkom rozlišovaným na zemskej pokrývke je vegetácia, medzi ktorú sa radí trávnatý porast, poľnohospodárske plodiny a listnaté a ihličnaté stromy (Halounová, Pavelka 2007). Vo väčšine prípadov je diaľkovým prieskumom skúmaný druh rastliny ako celok. V prípade stromov vzhľadom k ich veľkosti dochádza k pozorovaniu správania odrazivosti jeho rôznych častí, z čoho najlepšie pozorovať spektrálne chovanie umožňujú najviac listy (Kolář 1990). To je založené na obsahu ich zloženia, rovnako ako na uhle, v ktorom dochádza k odrazu, čo poukazuje na štruktúru listu a jeho vek (Halounová, Pavelka 2007).

Spektrálna krivka odrazivosti sa u listov delí na 3 hlavné časti:

- oblasť pigmentovej absorpcie (0,4 – 0,7 μm)

Pigmentové látky majú veľký vplyv na spektrálne chovanie listov vo viditeľnej časti spektra. Najvýznamnejšou je chlorofyl, ktorý môže tvoriť až 65 % listového pigmentu. V zelenom pásme má najvyššiu odrazivosť, naopak, v červenom a modrom pásme väčšinu, podľa *Campbella* 70 až 90 % dopadajúceho žiarenia absorbuje.

- oblasť bunečnej štruktúry (0,7 – 1,3 μm)

V týchto vlnových dĺžkach je hlavným faktorom pre typický nárast v odrazivosti morfológická stavba listu. Viac vrstiev listu umožňuje viacnásobný odraz žiarenia v pásme NIR, maximálny odraz je dosiahnutý pri 6-8 vrstvách listu.

- oblasť vodnej absorpcie (1,3 – 3,0 μm)

Spektrálna odrazivosť je tvorená absorpčnými pásmi vody. Dochádza tu k nepriamej úmere medzi odrazivosťou a obsahom vody v listoch. V týchto vlnových dĺžkach sa najlepšie pozoruje vodný stres rastlín (Dobrovolný 1998).

Voda

Spektrálne vlastnosti vody sa líšia od ostatných bežných povrchov, zároveň je pomerne homogénnou látkou, čo vo väčšine prípadov tvorí dobré základy pre kvalitné rozpoznanie vodných povrchov. Môže sa vyskytovať v rôznych skupenstvách, kde pri každom z nich má špecifické vlastnosti odrazivosti.

V kvapalnom skupenstve dochádza k silnej absorpcii slnečného žiarenia pri preniku vodným stĺpcom, čo sa vyznačuje nízkou odrazivosťou vo všetkých vlnových dĺžkach. Pri obsahu prímiesi biologického, či umelého pôvodu dochádza k zmene odrazivosti podľa ich zloženia.

Voda v pevnom skupenstve má opačné spektrálne chovanie. Vo viditeľnom aj NIR pásme má vysokú odrazivosť, ktorá často prevyšuje odrazivosť okolitých povrchov (Dobrovolný 1998).

Pôda

Spektrálne chovanie anorganických látok sa vyznačuje rastúcou odrazivosťou zároveň s vlnovou dĺžkou. Výkyvy môže spôsobovať vplyv faktorov ako minerálne zloženie pôdy, pôdna vlhkosť, obsah organických látok alebo textúra pôdneho povrchu.

Pôdy nasiaknuté vodou majú na základe spektrálneho chovania vody nižšiu odrazivosť. Štruktúra pôdy a nerovnosti na povrchu zapríčiňujú uhlové rozdiely v intenzite odrazeného žiarenia. S narastajúcou veľkosťou častíc v pôde sa zvyšuje jej celková odrazivosť (Dobrovolný 1998).

2.4 Vegetačné indexy

Pri súčasnej klimateckej situácii je zvýšený dopyt po presnejších informáciách o stave a fungovaní lesnej vegetácie ako doplnok k pozemnému sledovaniu. Pre sledovanie zmien v danom období je dobrým spôsobom využívanie vegetačných indexov. Vďaka spektrálnej odrazivosti povrchu je možné zaznamenávať informácie o stave vegetácie, štruktúre listov, či biogeochemickej kompozícii.

Rastliny majú v rôznych pásmach elektromagnetického žiarenia špecifickú odrazivosť. Väčšina červenej a modrej časti žiarenia sa pri fotosyntéze pohltí a časť zeleného spektra je odrážaná. V NIR spektre sa väčšina žiarenia z vegetácie na rozdiel od iných povrchov odráža, na základe čoho boli indexy vyvinuté. (Lukas, Neudert, Novák 2017)

Jeden z najpoužívanejších indexov je NDVI (Normalizovaný diferenčný vegetačný index), ktorého hodnota hovorí o zdraví a hustote vegetácie. Vďaka pigmentovým látkam, ako napr. chlorofyl, ktoré obsahujú listy, je RED pásmo pri zdravej vegetácii silno absorbované. K opaku dochádza pri pásme NIR, kde sa kvôli štruktúre listov žiarenie odráža. So zvyšujúcou sa hladinou stresu, ktorému je vegetácia vystavená, sa odrazivosť v NIRe znižuje a odrazivosť v červenom pásme stúpa vzhľadom k nízkej produkcii chlorofylu, následkom čoho dochádza k žltnutiu listov.

Vzorec pre NDVI:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

Index nadobúda hodnoty v intervale $<-1;1>$, čím je hodnota vyššia, tým obsahuje daná oblasť vyššiu koncentráciu zdravej, zelenej vegetácie. Záporná hodnota naznačuje absenciu vegetácie a zodpovedá vodným plochám. (Kiage a kol. 2007; Gaja 2013)

Najjednoduchším indexom je Simple Ratio (SR), ktorý sa počíta ako pomer medzi hodnotou odrazivosti v pásme NIR a Red. Dokáže rozlíšiť zelené listy a odhadnúť relatívnu biomasu nachádzajúcej sa v snímku. (Agricultural Imaging Solution 2021)

Vzorec pre SR:

$$SR = \frac{NIR}{RED} \quad (2)$$

Hodnoty sa pohybujú od 0 do približne 30, kde vysoké hodnoty spadajú vegetácii s vysokým LAI (leaf area index) a nízke hodnoty pre pôdu či vodu. Zdravá vegetácia sa pohybuje zvyčajne v intervale $<2;8>$. (ArcGIS Pro)

Na minimalizáciu jasu pôdy vznikajúceho pri použití vegetačných indexov zahŕňajúcich RED a NIR vlnové dĺžky bol v roku 1988 vytvorený *Huetom* index SAVI (Soil Adjusted Vegetation Index). Základný princíp je rovnaký, ako pri indexe NDVI, avšak výsledok je upravený parametrom L, ktorý má hodnoty v intervale $<0;1>$ a používa sa podľa hustoty vegetačnej pokrývky. Pri najvyššej hustote je používaná hodnota 0, najbežnejšie sa využíva hodnota 0,5 pre priemernú hustotu. Sykas (2023) vo svojom článku porovnával indexy NDVI a SAVI, kde vidieť rozdiel v saturácii, ktorá pri NDVI nastáva skôr, čím sa SAVI stáva pri správnom nastavení L parametru vhodnejšou voľbou pri hustej vegetácii.

Formula pre SAVI:

$$SAVI = \frac{NIR - RED}{NIR + RED + L} * (L + 1) \quad (3)$$

Pomocou Normalizovaného Diferenčného Indexu Vlhkosti (NDMI) sa zisťuje úroveň vlhkosti a vodného stresu v plodinách, čo je užitočné ako prevencia v poľnohospodárstve, nakoľko sa detekuje už v skorom štádiu. Pre výpočet sa používa pásma NIR a SWIR.

$$NDMI = \frac{NIR - SWIR}{NIR + SWIR} \quad (4)$$

Hodnoty sú v intervale $<-1;1>$ a čím je hodnota vyššia, tým menej vodného stresu má daná oblasť, pri hodnote 1 môže byť až presiaknutá. (EOS DATA ANALYTICS 2023; Simerská 2021; Antognelli 2018)

2.5 Podobné práce

Vozár (2018) vo svojej práci klasifikoval lesnú vegetáciu v lokalite Černý les, Groniček a Kovárův žleb na dátach vysokého (Landsat-8 a Sentinel-2) a veľmi vysokého (PlanetScope a RapidEye) rozlíšenia. Použil klasifikačnú metódu Maximum Likelihood with null class, ktorá sa zaradzuje k riadenej klasifikácii a využil software PCI Geomatica. Taktiež bol vypočítaný vegetačný index NDVI, z ktorých boli následne vytvorené rastrové mapy. Do klasifikácie index nevstupoval kvôli rozdielnosti hodnôt medzi snímkami z rôznych družíc. Vozár porovnával výsledky vykonanej klasifikácie s validačnými dátami od Corine Land Cover a ÚHUL. Najväčšiu presnosť dosiahli dáta z PlanetScope, podobne aj z RapidEye, ale vyskytol sa pri nich problém so zatienenými plochami. Pri dátach s nižším priestorovým rozlíšením dochádzalo k väčšiemu počtu prípadov s neklasifikovanými pixelmi.

V bakalárskej práci Riedla (2016) bola vytvorená klasifikácia lesnej vegetácie pomocou dát veľmi vysokého (World View – 2) a vysokého rozlíšenia (Landsat – 8) z oblasti Národného parku Šumava. V ďalšej časti práce bola vytvorená legenda. K vytvoreniu per-pixel klasifikácie využíval vybrané algoritmy riadenej (Maximum Likelihood, Minimum Distance) a neriadenej klasifikácie (ISODATA), ktoré boli spracované v software Envi 5.1., v rovnakom softvare bola zhotovená metóda klasifikačnej chybovej matice na vyhodnotenie presnosti. Najvyššiu presnosť Riedl dosiahol na dátach World View – 2 pomocou algoritmu Maximum Likelihood, a to 54,57 %. Táto nízka hodnota bola autorom odôvodnená veľkým územím a problémom pri presnom výbere tréningových plôch.

V zahraničnej práci *Object-based classification vs. Pixel-based classification: comparative importance of multi-resolution imagery* boli porovnávané výsledky klasifikácie založené na metóde per-pixel a per-object. Ako dáta boli použité 2 snímky so stredným priestorovým rozlíšením (SPOT-5) z obdobia s vegetáciou s listami a s vegetáciou bez listov a Color Infrared Digital ortophoto snímok s vysokým rozlíšením. Snímky boli zlúčené do jedného pre potrebu vytvorenia viac-časového a viac-priestorového snímku, a práve ten mal vo výsledkoch najväčšiu presnosť (82,0 %). Táto práca podporila zistenia mnohých ďalších podobne založených prác z posledných rokov tvrdiacich, že pri viac-priestorových a viac-časových snímkoch bude mať objektová klasifikácia vyššiu presnosť ako riadená a neriadená pixelová klasifikácia.

Autori práce *Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery* sa zamerali na prieskum klasifikácie rôznymi algoritmi a zároveň porovnávali výsledky algoritmov vykonávaných na snímkoch Sentinel-2 s rôznymi veľkosťami tréningových dát. Vietnamská oblasť vo veľkosti 30x30 km² bola klasifikovaná do 6 tried krajinných povrchov s použitím 14 rôznych tréningových dát v rozmedzí veľkostí 50 až 1250 px/trieda. Presnosť výsledkov sa pohybovala pri všetkých metódach medzi 90–95 %. Support Vector Machine mal najmenšie výkyvy vo výsledkoch pri zmene veľkostí tréningových plôch. Pri väčšom počte pixelov ako 750 px/trieda, alebo aspoň veľkosti 0,25% plochy skúmanej oblasti boli výsledky všetkých algoritmov najpresnejšie.

Minárčik (2015) v svojej práci porovnával algoritmy Maximum Likelihood, Support Vector Machine a z neriadenej klasifikácie ISODATA na snímkoch z WorldView – 2 a Landsat – 8, kde pri snímku s priestorovým rozlíšením 2 m (WorldView – 2) bola najpresnejšia metóda riadenej klasifikácie Maximum Likelihood s celkovou presnosťou 73,1 %. Snímok s rozlíšením 30 m (Landsat – 8) mal najlepšiu presnosť pri metóde Support Vector Machine s celkovou presnosťou 70,78 %.

Vlačihová (2018) vo svojej diplomovej práci hľadala najlepšiu metódu klasifikácie obrazu pre sledovanie zmien v povrchu krajiny. Používala riadenú i neriadenú klasifikáciu na dátach od Landsat – 8 a Sentinel – 2. K najlepším výsledkom došla pri použití algoritmu Maximum Likelihood na snímkoch Sentinelu – 2.

V článku *Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data* sa autori zaoberali porovnaním presnosti riadenej a neriadenej klasifikácie na snímkoch SPOT 5. Po vytvorení náhodných 72 vzorových bodov, ktoré tvorili 25 % celej študovanej oblasti boli vyhodnotených výsledky z chybovej matice. Riadená klasifikácia mala presnosť 90,28 %, s Kappa štatistikou 0,86 a neriadená klasifikácia vyšla s presnosťou 80,56 % s Kappa štatistikou 0,73. Z tohto záveru autori vyhodnotili, že riadená klasifikácia má presnejšie výsledky.

Štúdia *Per-pixel vs. Object-based classification of urban land cover extraction using high spatial resolution imagery* porovnáva klasifikáciu na základe zhľukovania per-pixel a per-object cez snímky s vysokým rozlíšením zo satelitu QuickBird v mestských oblastiach. Objektová klasifikácia má značne vyššiu presnosť, a to 95,2 %, zatiaľ čo najbežnejšie používaná pixelová metóda Maximum Likelihood má presnosť 87,8 %.

Zadrazil (2021) porovnával pokročilé klasifikátory Random Forest (RF) a Classification and regression tree (CART) na Etiópskej oblasti Sidama pomocou dát z družíc Landsat-8 a Sentinel-2 v cloud-based prostredí Google Earth Engine pri mierke 1 : 000 000. Vyššia celková presnosť bola dosiahnutá klasifikátorom RF, dáta zo Sentinel-2 má 82 %, z Landsatu-8 80 %.

3 METÓDY A POSTUP SPRACOVANIA

V tejto kapitole sú spomenuté potrebné kroky na dosiahnutie cieľa tejto bakalárskej práce. V nasledujúcich kapitolách dôjde k ich hlbšiemu rozboru.

Použité metódy

Pre tvorbu klasifikácií boli použité metódy per-pixel riadenej a neriadenej klasifikácie algoritmami Maximum Likelihood, Random Forest, K-Nearest Neighbors a K-Means. Následne došlo k porovnávaniu a analýze výsledkov.

Použité dáta

V práci boli použité snímky zo satelitov Sentinel-2 a Pléiades z oblasti zvoleného územia povodia Všeminky a Trnávky.

Zo satelitu Pleiades boli zakúpené katedrou geoinformatiky dva snímky (DS_PHR1A_202108140954027_FR1_PX_E017N49_1016_01230; DS_PHR1B_202109150958411_FR1_PX_E017N49_1112_01312) kvôli veľkosti povodia a potrebe pokrytia celej záujmovej oblasti. Boli nasnímané dňa 14. 8. 2021 a 15.9.2021 a získané v súradnicovom systéme WGS84. Snímky obsahujú päť spektrálnych pásiem, a to štyri v priestorovom rozlíšení 2 m (modré, zelené, červené a blízke infračervené) a panchromatické pásmo s priestorovým rozlíšením 2 m.

Dáta zo satelitu Sentinel-2 sú voľne dostupné cez program Copernicus. Rozsah vybraného snímku (S2B_MSIL2A_20211009T095029_N0301_R079_T33UYQ_20211009T120840) zaberá územie povodia a dátum je z rovnakého obdobia, ako snímok Pléiades (9. 10. 2021) z dôvodu podobného zastúpenia plôch pri klasifikácii obrazu a možnosti porovnania. Produkt je úrovne Level 2A, ktorá je najvhodnejšia vďaka už obsahujúcej atmosférickej korekcii, vďaka čomu už nie je potrebné ju robiť dodatočne.

Validačnými dátami bude Detailní kombinovaná vrstva (DKV), ktorá bola dodaná pánom profesorom RNDr. Vilémom Pechancom, Ph.D. DKV je produkt z roku 2014, ktorý rozlišuje 152 typov habitatov v Českej republike. (Procházka, 2020) Je tvorená v spolupráci s Agentúrou ochrany prírody a krajiny ČR Ústavom výzkumu globální změny AV ČR. Vznikla ako kombinácia Konsolidované vrstvy ekosystému (KVES) a Vrstvy mapování biotopu (AOPK ČR, 2014) spolu s ďalšími dostupnými zdrojmi ako ZABAGED (Základní báze geografických dat), DIBAVOD (Digitální báze vodohospodářských dat), UrbanAtlas, CORINE Land Cover a iné. (Purkyt 2022)

Použité programy

Pre spracovanie dát a tvorbu bakalárskej práce bolo potrebných niekoľko počítačových programov.

Na počítačnú prípravu dát a hlavnú časť práce, a to klasifikácia snímok a výpočet vegetačných indexov bol použitý program SNAP vo verzii 8.0.0., spracovávanie výsledkov prebiehalo v programe QGIS vo verzii 3.24.0-Tisler, tabuľky a texty boli vytvorené pomocou balíčka MS Office 365 a programom MS Word a MS Excel.

SNAP

The Sentinel Application Platform (SNAP) je software vyvinutý inštitúciami Brockmann Consult, Skywatch, Sensar and C-S určený primárne na analýzu a spracovávanie Sentinel dát, dokáže však pracovať aj s dátami iných satelitov. Je to súbor spustiteľných nástrojov

a APIs (Application Programming Interfaces), ktoré boli vyvinuté na uľahčenie využívania, zobrazovania a spracovania rôznych údajov z diaľkového snímania. (*Sentinel Toolbox Help*)

QGIS

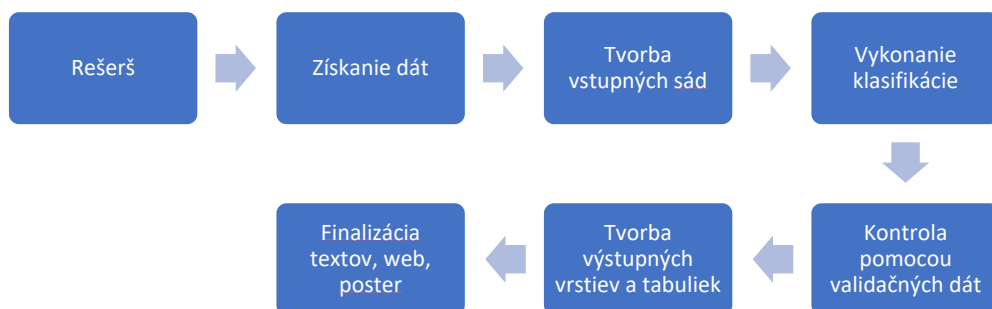
QGIS (pôvodne „Quantum GIS“), je v súčasnosti jeden z najrozšírenejších, voľne dostupných geografických informačných systémov, licencovaný pod GNU General Public Licence, Je oficiálnym projektom Open Source Geospatial Foundation (OSGeo). Podporuje množstvo vektorových, rastrových a databázových formátov a funkcií a má možnosť doinštalovania nových pluginov. Nové verzie programu vychádzajú pravidelne každé štyri mesiace. (Jenkins 2023)

Postup spracovania

Prvým krokom k spracovaniu bakalárskej práce bola potrebná rešerš a preštudovanie dostupných domácich a zahraničných zdrojov. Z tých boli vybrané poznatky, vďaka ktorým sa dalo priblížiť k výberu vhodných klasifikačných metód a nasledujúcemu postupu.

Po vykonaní dostatočnej rešerše bolo možné prejsť k praktickej časti. Ako prvé bolo potrebné získať dáta, na ktorých sa vykonáva celá práca (kapitola 4.1). Došlo k ich úprave a tvorbe štyroch vstupných sád, na ktorých prebiehali podobné procesy a na záver boli porovnávané.

Vytvorené klasifikačné mapy boli validované pomocou kontrolnej vrstvy DKV a zároveň bola zisťovaná presnosť chybovou maticou. Posledným krokom pred finalizáciou práce boli vytvorené mapové výstupy a tabuľky.



Obrázok 5 Postup práce

4 PRÍPRAVA DÁT

Pred samotným začiatkom klasifikovania je potrebné dáta získať a vhodne pripraviť na nasledujúci proces. Podstúpené kroky sú vypísané v tejto kapitole.

4.1 Získanie dát

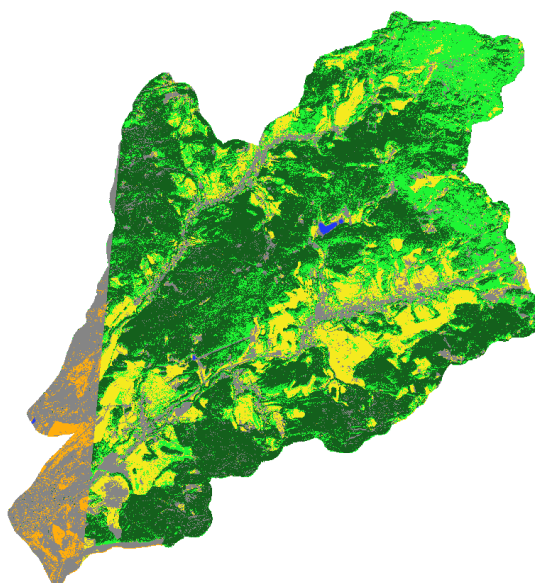
Snímok s veľmi vysokým rozlíšením zo satelitu Pleiades bol zakúpený katedrou geoinformatiky zo záujmovej oblasti z dátumu 14. 8. 2021 a 15.9.2021. Na základe kúpeného snímku bol vyberaný snímok zo Sentinelu-2 z rovnakého obdobia (9. 10. 2021), z dôvodu rovnakého zastúpenia plôch pri klasifikácii obrazu, ktorý je voľne dostupný na stránke Copernicus Open Access Hub. Po potrebnej registrácii boli zadané požadované parametre na snímok, kde ako prioritné nastavenie bol výber snímku z podobného časového obdobia ako snímok z Pleiades a oblasť bez, alebo s minimálnou oblačnosťou (nastavenie 0 – 10 % pokryvu). Produkt je úrovne Level 2A, ktorá je najvhodnejšia vďaka už obsahujúcej atmosférickej korekcii, takže nie je potrebné ju robiť dodatočne (Žubrietovský a kol. 2017).

4.2 Príprava snímok

Stiahnuté snímky bolo potrebné pripraviť na klasifikáciu, na čo bol zvolený software QGIS a SNAP. Snímok Pleiades bol otvorený v programe QGIS, nakoľko v danom formáte ho nebolo možné otvoriť priamo v SNAPe a tam boli potrebné dáta vyexportované vo formáte .tif.

Scéna z Pléiades je zložená z dvoch snímok, ktoré kvôli následnému spojeniu bolo potrebné v prvom rade georeferencovať, keďže pri pôvodných snímkach dochádza k miernemu posunu. Georeferencovanie prebehlo v QGISe pomocou funkcie *Georeferencer*, kde bol nahraný potrebný raster a následne vybraných približne 250 GCP (Ground Control Points) podľa zhodných miest z referenčnej mapy Google Satellite použitej z QGISom ponúkaných XYZ dlaždíc. Následovalo nastavenie transformácie, kde ako prevzorkovacia metóda bolo ponechané nastavenie najbližšieho suseda a ako transformačný typ bola po vyskúšaní viacerých transformácií zvolená Polynomická transformácia 3. rádu.

Po georeferencovaní bolo potrebné vytvoriť z daných snímok mozaiku, ktorá vzniká spojením vybraných snímok. Bola vytvorená v QGISe cez funkciu *Mosaic* s predvoleným nastavením, okrem harmonizačnej metódy, kde bolo zvolené nastavenie *band*, nakoľko pri iných spôsoboch mozaikovania dochádzalo k značným rozdielom medzi dvoma časťami snímku a farbami v multispektrálnom snímku v 3. pásme Pléiades, a tým pádom aj k skresleným výsledkom klasifikácie (viď obrázok 5).



Obrázok 6 Chybná klasifikácia ML v dôsledku nesprávneho mozaikovania.

Na získanie čo najlepšej kvality sa vykonal proces pansharpeningu (skratka pre panchromatický sharpening), čo je forma „image fusion“, ktorá kombinuje panchromatický snímok s vysokým priestorovým rozlíšením a multispektrálny snímok s vysokým spektrálnym rozlíšením, čím vzniká ostrejší multispektrálny snímok so zachovanými spektrálnymi informáciami (Njambi 2022).



Obrázok 7 Snímok pred (vľavo) a po (vpravo) pansharpeningu. (zdroj: <https://up42.com/blog/how-pansharpening-improves-satellite-imagery>)

Následne boli snímky presunuté do programu SNAP. Snímok Sentinelu tvorí veľkú oblasť 10980 px x 10980px, pre vymedzenie územia a rýchlejšiu manipuláciu s ním bol snímok v SNAPe orezaný pomocou funkcie *subset* a plocha bola vybraná podľa vektorovej vrstvy povodí stiahnutej z Dibavodu pomocou funkcie *land/sea mask*. V tomto procese boli ponechané iba pásma, ktoré budú využívané v klasifikácii, a to pásma s priestorovým rozlíšením 10 m a 20 m (odstránené b1, b9, b10 s priestorovým rozlíšením 60 m). Snímku bol zmenený súradnicový systém pomocou funkcie *reprojection* na S-JTSK v Křovákovom zobrazení. Finálnym krokom bolo prevzorkovanie všetkých pásiem pomocou funkcie Resampling na rozlíšenie 20 m, čo znamená, že v novo vytvorenej sieti pixelov sa zadávajú

pixelom nové hodnoty na základe vybranej metódy, ktorá v tomto prípade bola Nearest neighbor – tá vyberá novú hodnotu pixelu podľa hodnoty najbližšieho pixelu novej lokácie. Táto metóda je síce najjednoduchšia, no môže dôjsť k duplikácii či strateniu pixelov. Vďaka prevzorkovaniu pásiem na rovnaké rozlíšenie je možné vykonávať výpočty medzi pásmami pomocou indexov (Lillesand, Kiefer, Chipman 2015; Fundamentals of Remote Sensing 2019). Na mozaikovaných snímkoch Pléiades bol v SNAPe uskutočnený rovnaký postup ako pri snímkoch Sentinelu.

Týmto boli získané štyri sety pripravených snímok v programe SNAP: Sentinel-2 s pásmami s rozlíšením 10 m, Sentinel-2 so všetkými pásmami s rozlíšením 10 m a 20 m prevzorkovaných na 20 m, multispektrálne snímky Pléiades s rozlíšením 2 m a panchromatický snímok Pléiades po pansharpeningu s rozlíšením 0,5 m, na ktorých mohol nasledovať ďalší postup, a to vykonávanie klasifikácií.

5 KLASIFIKÁCIA

Následným krokom bolo vykonanie samotných klasifikácií. Kategórie, do ktorých sú snímky klasifikované, boli vybrané na základe úpravy DKV a sú nasledujúce.

- Ihličnaté lesy
- Listnaté lesy
- Trávnaté plochy/lúky
- Vodné plochy
- Zastavané oblasti
- Orná/holá pôda

5.1 Riadená klasifikácia

Na vykonanie riadenej klasifikácie bolo v prvom rade potrebné vytvoriť trénovacie plochy. U nich je dôležitá čo najlepšia kvalita kvôli reprezentácii vybraných kategórií, podľa ktorých sa bude algoritmus riadiť. Podľa *Campbella* je potrebných aspoň 100 pixelov v jednotlivých kategóriách na dosiahnutie dobrých výsledkov, ale dôležitá je ich spektrálna homogenita. Vzhľadom k rôznemu priestorovému rozlíšeniu v štyroch vytvorených setoch, aj počet pixelov v trénovacích plochách sa líši (viď Tabuľka 2), a jednotlivé polygóny museli byť prispôbené rozlíšeniu a realistickým tvarom objektov v daných snímkach, kde vidieť medzi jednotlivými snímkami mierne zmeny v povrchu vplyvom rozličného obdobia zhotovenia snímok Pléiades a Sentinel.

K vytvoreniu trénovacích plôch došlo v programe SNAP pomocou vytvorenia *New Vector Data Container*, čo je kategória pre vyznačovanú triedu a následne pomocou *Polygon Drawing Tool* boli vyznačované polygóny (zapínanie/vypínanie pomocou Mask/Layer Manager).

Kontrola správneho vyznačenia tried, a teda normálne (Gaussovo) rozdelenie, bola prevedená pomocou funkcie *Statistics (Analysis)*, kde je zobrazené spektrálne rozdelenie tried pomocou grafov. Najhoršia separabilita sa ukazovala pri kategóriách orná pôda a zastavané oblasti.

Tabuľka 2 Počet pixelov trénovacích plôch v jednotlivých triedach

	Sentinel 10 m	Sentinel 20 m	Pléiades - multispektrálne	Pléiades - pansharpening
Listnatý les	2 778	726	36 806	476 196
Ihličnatý les	5 921	1 452	38 552	498 776
Lúka, tráva	9 129	2 283	26 172	338 231
Vodná plocha	418	109	8 355	108 008
Zastavaná oblasť	4 333	1 091	13 885	191 828
Orná/holá pôda	2 795	694	36 584	473 156

V dôsledku malého zastúpenia triedy vodná plocha vo vymedzenom území došlo k vytvoreniu aj malej trénovacej plochy, v prípade Sentinelu s priestorovým rozlíšením 20 m s minimálnym počtom pixelov.

5.2 Neriadená klasifikácia

Ďalším druhom vykonanej klasifikácie bola neriadená klasifikácia, konkrétne K-means Cluster Analysis. Ako už bolo spomínané v kapitole neriadená klasifikácia, pri vykonávaní sa nepoužívajú tréningové plochy. Nastavenie bolo ponechané predvolené, okrem počtu tried, kde bolo zvolené najvyššie možné číslo v programe SNAP, a to 100 tried. Pri počte 30 dochádzalo k vytvoreniu jednej kategórie pre vodné plochy a tieňe spôsobené lesom, v prípade sady Sentinel-2 s pásmami s rozlíšením 10 m došlo k zmiešaniu v jednej kategórii vodných plôch a zastavaných oblastí. Po vytvorení 100 tried bolo potrebné manuálne rozraďovanie do kategórií podľa vytvorenej legendy a následné zlúčenie tried v programe QGIS pomocou funkcie *Reclassify*. Nakoľko tu došlo k viditeľnému roztriedeniu tieňov do samostatnej kategórie a pri rozraďení do zvolených šiestich kategórií by došlo k predom známej chybe, bola vytvorená siedma kategória - tieňe.

Vegetačné indexy

Po vyskúšaní základných pásiem boli zistené nedostatky v klasifikácii, preto nasledovalo pridanie vegetačných indexov. Pomocou funkcie *Band math* boli vypočítané vybrané vegetačné indexy a pridané do klasifikácie. Hlavným riešeným problémom boli zatienené oblasti a problém v rozlíšení pôdy a zástavby. Na všetkých sádach bol vypočítaný NDVI index vzhľadom k prítomnosti potrebných pásiem na jeho výpočet u oboch satelitoch, rovnako ako index SAVI, ktorý je identický ako NDVI index s pridaným parametrom zahustenia vegetácie. Ďalším pridaným indexom bol index NDTI, ktorý sa v štúdiu *Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery* ukázal ako najlepší index pre rozlíšenie pôdy od zastavaných oblastí. Na záver bol pridaný index BSI, ktorý podľa *Rouibah a Belabbas* v kombinácii s NDTI a NDVI najlepšie rozlišoval holú suchú pôdu od ostatných povrchov. Tieto dva indexy boli vytvorené iba pre sadu Sentinel-2 s 20 m rozlíšením z dôvodu absencie SWIR pásma u ostatných sád.

6 VÝSLEDKY

Pre zistenie úspešnosti klasifikácie bolo potrebné získané výsledky spracovať. Bola zisťovaná celková, užívateľská a spracovateľská presnosť klasifikácie pomocou chybovej matice a ručne vypočítaný Kappa koeficient. Klasifikácie s najvyššie dosiahnutými presnosťami v rámci jednotlivých sád boli porovnávané s DKV v mierke 1 : 10 000, 1 : 20 000 a 1 : 50 000. Výsledky klasifikácií sa nachádzajú v prílohe 1 – 4.

Vyhodnocovanie klasifikácií bolo vykonávané v programe QGIS. Klasifikácie vytvorené v SNAPE sú vo formáte .dim (BEAM-DIMAP format), ku ktorým sa vytvára zároveň priečinok súborov vo formátoch HDR a IMG, ktorý je použitý na zobrazenie klasifikácie v QGISe. Klasifikáciám boli zmenené hodnoty pomocou funkcie *Reclassify* a symbolológia na farby vhodné k daným kategóriám. Satelitný snímok Sentinelu sa v QGISe otvára jednotlivo prostredníctvom samostatných súborov pre všetky pásma. Tie je potrebné kombinovať v jednom rastru pre zobrazenie snímku v pravých farbách, preto ich bolo potrebné najprv pomocou funkcie *Built Virtual Raster* a zaškrtnutím funkcie na vytvorenie samostatných pásiem zlúčiť na jeden raster.

6.1 Presnosť klasifikácie

Odhad presnosti klasifikácie bol vytvorený pomocou chybovej matice, ktorá bola zhotovená v programe QGIS pomocou pluginu AcATaMa (Accuracy Assessment of Thematic Maps). Tá porovnáva vzťah medzi referenčnými dátami a korešpondujúcimi výsledkami vytvorenej klasifikácie, z ktorých je možné vyčítať niekoľko charakteristík. Celková presnosť, ktorá vyjadruje presnosť pre celú klasifikáciu, je vypočítaná delením celkového počtu správne klasifikovaných pixelov (súčet hlavnej diagonály) s celkovým počtom klasifikovaných pixelov. Bližšie výsledky pre jednotlivé kategórie sa získavajú pomocou používateľskej a spracovateľskej presnosti. Používateľská je pomerom správne klasifikovaných pixelov v každej kategórii s celkovým počtom pixelov zaradených do daných kategórií podľa referenčných dát. Spracovateľská presnosť je udávaná pomerom správne klasifikovaných pixelov v každej kategórii s celkovým počtom klasifikovaných pixelov v daných kategóriách. (Lillesand, Kiefer, Chipman 2015) Pomocou chybovej matice bol ešte dodatočne v Exceli vypočítaný Kappa koeficient, ktorý predáva informáciu o tom, aká kvalitná bola klasifikácia v porovnaní s náhodne roztriedenými pixelmi. Kappa koeficient má hodnoty v intervale $<-1;1>$, kde hodnoty blížiac sa k -1 značia značne horšie roztriedenie ako náhodné rozdelenie, hodnoty blížiac sa k 1 indikujú, že klasifikácia je signifikantne lepšia (*Accuracy Metrics 2014*).

Pre každú vytvorenú klasifikáciu bolo v prvom rade vygenerovaných dokopy 70 náhodných bodov pomocou *Stratified Random Sampling* manuálne rozložených v kategóriách strategicky podľa ich veľkostí. Počet bodov bol vybraný podľa bakalárskej práce *Vozára*, ktorý používal vzorku 30 kontrolných bodov. Ako napríklad v bakalárskej práci *Sýkoru*, je možné v chybovej matici použiť aj tréningové plochy (ROI), ktoré boli použité v riadenej klasifikácii, kde bol dosiahnutý vysoký odhad presnosti, avšak týmto spôsobom by sa zisťovala presnosť iba daných plôch, a nie celého územia.

K vygenerovaným bodom bola následne ručne priradená skutočná kategória za pomoci daného satelitného snímku a ortofotomapy. Za ideálnych predpokladov by sa súčet čísel na hlavnej diagonále matice rovnal počtu všetkých kontrolných bodov, čo by znamenalo 100% správnosť zaradených kategórií pri kontrolných bodoch.

K vygenerovaným 70 bodom boli následne ku klasifikáciám s najlepším výsledkom pre overenie vygenerované body podľa automatického rozloženia *Area based proportion*, kde boli počty v skupinách vytvorené na základe strednej chyby priemeru (SE). Celkový počet bodov sa zvýšil pri očakávanej SE 0,005 na 397 bodov, pri SE 0,007 na 203 bodov. Chybové matice sa nachádzajú v prílohe 5.

V tejto práci boli otestované 4 algoritmy s rôznymi nastaveniami parametrov, výsledky sú prezentované v nasledujúcom texte.

Sentinel-2 v rozlíšení 10 m

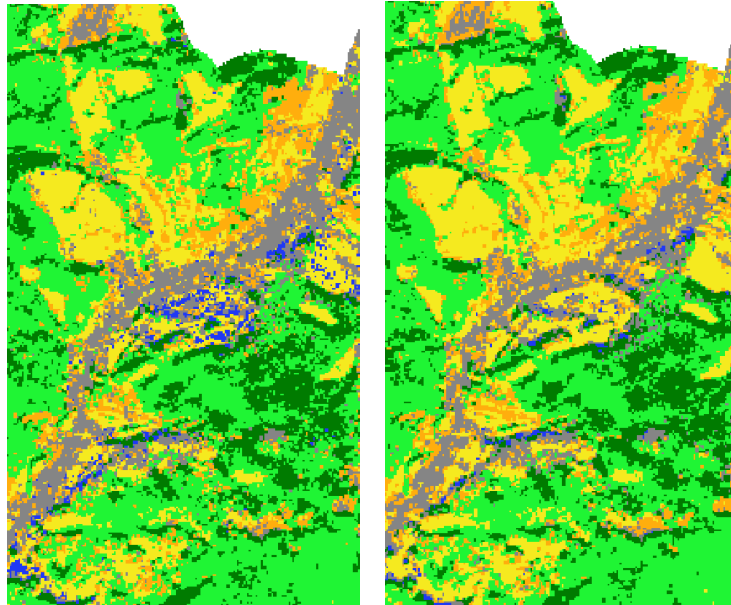
Pre Sentinel-2 set s pásmami s priestorovým rozlíšením 10 m vyšli podobné presnosti pri predvolených nastaveniach v algoritmoch. Klasifikátor Random Forest dosiahol celkovú presnosť 70 %. Pri ihličnatých lesoch došlo k zámene jedného kontrolného bodu za listnatý les, u listnatých boli tri body zaradené nesprávne ako ihličnatý les. Najnižšiu užívateľskú presnosť získali vodné plochy, kde bol správne zaradený iba jeden bod.

Tabuľka 3 Chybová matica pri klasifikátore RF so 70 kontrolnými bodmi

		Validačné data						Presnosť		
		Ihličnaté lesy	Listnaté lesy	Lúky, trávniky	Vodné plochy	Zastavané plochy	Orné/holé pôdy	Celkom	Používateľská	Kappa
Klasifikácia	Ihličnaté lesy	14	1	0	0	0	0	15	0,9333	
	Listnaté lesy	3	12	0	0	0	0	15	0,8000	
	Lúky, trávniky	0	1	11	0	1	2	15	0,7333	
	Vodné plochy	0	3	2	1	1	0	7	0,1429	
	Zastavané plochy	0	0	1	0	6	2	9	0,6667	
	Orné/holé pôdy	0	1	2	0	1	5	9	0,5556	
Presnosť	Celkom	17	18	16	1	9	9	70		0,6293
	Spracovateľská	0,8235	0,6667	0,6875	1,0000	0,6667	0,5556		0,7000	
	Kappa									

Podobnú celkovú presnosť dosiahol klasifikátor KNN (74 %) s Kappa koeficientom 0,62. Rovnako ako pri RF, najväčší problém bol s vodnými plochami, ktoré boli označené aj ako zatienené časti v zastavaných oblastiach.

V obrázku č. 5 vidieť porovnanie problémovej oblasti s oboma klasifikátormi.



Obrázok 8 Problémové plochy falošne zobrazujúce vodné plochy v zastavaných oblastiach. RF (vľavo) a KNN (vpravo).

Po zmene klasifikátora RF z pôvodného nastavenia 10 stromov na 250 stromov celková presnosť narástla na 74 %. Falošné vodné plochy boli zredukované, avšak stále z kontrolných bodov bol iba 1 správne lokalizovaný. Lesy boli rozlíšené správne so všetkými bodmi v prípade listnatých aj ihličnatých. Nízka presnosť bola aj v rozlišovaní ornjej pôdy, kde štyri body boli zaznačené ako trávnaté plochy.

Sentinel-2 v rozlíšení 20 m

V rámci danej sady dát bola algoritmom Random Forest s počtom stromov 250 a za použitia všetkých spektrálnych pásiem a indexov NDVI, SAVI, NDTI s BSI dosiahnutá celková presnosť 67 %. V rámci jednotlivých kategórií bolo zaznamenané, že algoritmus dosiahol najvyššiu užívateľskú presnosť v kategórii ihličnaté lesy a lúky. Vodné plochy dosiahli najnižšiu užívateľskú presnosť (14 %) a podobne je na tom s hodnotou 33 % holá pôda, kde boli správne klasifikované iba 3 body. V spracovateľskej presnosti majú vodné plochy 100 %, čo znamená, že iba 1 bod bol označený za vodu, ktorá aj v skutočnosti vodou je. Problém s vodnými plochami môže byť zapríčinený nedostatkom tréningových plôch. Pri porovnaní algoritmu RF bez vegetačných indexov a s vegetačným indexom vidieť z klasifikácie redukované množstvo zastavaných oblastí a správne klasifikovaná časť s ornou pôdou na spodku snímku (viď obr. ...), ktorý bol problematický vo väčšine klasifikácií, avšak aj napriek tomu pri kontrolných bodoch nedosahuje dobré výsledky.

Tabuľka 4 Chybová matica pri klasifikátore RF so 70 kontrolnými bodmi

		Validačné data						Presnosť		
		Ihličnaté lesy	Listnaté lesy	Lúky, trávniky	Vodné plochy	Zastavané plochy	Orné/holé pôdy	Celkom	Používateľská	Kappa
Klasifikácia	Ihličnaté lesy	13	2	0	0	0	0	15	0,8667	
	Listnaté lesy	3	10	0	0	1	1	15	0,6667	
	Lúky, trávniky	0	1	13	0	0	1	15	0,8667	
	Vodné plochy	0	2	4	1	0	0	7	0,1429	
	Zastavané plochy	0	0	1	0	7	1	9	0,7778	
	Orné/holé pôdy	0	0	0	0	6	3	9	0,3333	
Presnosť	Celkom	16	15	18	1	14	6	70		0,5953
	Spracovateľská	0,8125	0,6667	0,7222	1,0000	0,5000	0,5000		0,6714	
	Kappa									

Ešte vyššiu presnosť docielila neriadená klasifikácia K-means s NDVI indexom (84 %). V danej klasifikácii bola pridaná kategória tieň, avšak pri kontrole nebol ani 1 bod zaradený správne. Aj napriek tomu táto klasifikácia mala najvyššie hodnotenie zo všetkých vykonaných. Kappa koeficient dosahoval hodnotu 0,85. V používateľskej presnosti došlo k 100 % presnosti spomedzi kontrolných bodov v kategórii lúky a vodné plochy, medzi ostatnými kategóriami sa hodnoty pohybovali nad úrovňou 70 %.

Tabuľka 5 Chybová matica pri klasifikátore K-means so 70 kontrolnými bodmi

		Validačné data							Presnosť		
		Ihličnaté lesy	Listnaté lesy	Lúky, trávniky	Vodné plochy	Zastavané plochy	Orné/holé pôdy	Tieň	Celkom	Používateľská	Kappa
Klasifikácia	Ihličnaté lesy	13	2	0	0	0	0	0	15	0,8667	
	Listnaté lesy	0	14	1	0	0	0	0	15	0,9333	
	Lúky, trávniky	0	0	15	0	0	0	0	15	1,0000	
	Vodné plochy	0	0	0	5	0	0	0	5	1,0000	
	Zastavané plochy	0	0	3	0	7	0	0	10	0,7000	
	Orné/holé pôdy	0	0	2	0	0	5	0	7	0,7143	
	Tieň	1	0	1	1	0	0	0	3	0,0000	
Presnosť	Celkom	14	16	22	6	7	5	0	70		0,8515
	Spracovateľská	0,9286	0,8750	0,6818	0,8333	1,0000	1,0000	0,0000		0,8429	
	Kappa										

Najnižšia celková presnosť bola získaná algoritmom Maximum Likelihood so základným nastavením, kde dosiahla hodnotu 72 % a kappu koeficient 0,68. Najväčším problémom bola kategória zastavané oblasti, kde aj priamo vo výstupe klasifikácie vidieť, že obsahuje ďaleko viac plochy ako v realite.

Pri overovaní presnosti s viacerými bodmi došlo k zhoršeniu výsledkov, avšak pri porovnaní boli stále najlepšie výsledky klasifikátora K-means s NDVI indexom a 100 zhlukmi. Problém dochádzal pri rozlišovaní ihličnatých a listnatých lesov, čo môže byť spôsobené aj nesprávnym zaradením autorom, nakoľko v danej oblasti sa nachádzajú aj zmiešané lesy, kde je problematické rozlíšiť správnu kategóriu aj za pomoci referenčných dát.

Tabuľka 6 Chybová matica pri klasifikátore K-means s 396 kontrolnými bodmi

		Validačné data							Presnosť		
		Ihličnaté lesy	Listnaté lesy	Lúky, trávniky	Vodné plochy	Zastavané plochy	Orné/holé pôdy	Tieň	Celkom	Používateľská	Kappa
Klasifikácia	Ihličnaté lesy	86	29	1	0	0	0	3	119	0,7227	
	Listnaté lesy	4	95	10	0	1	0	2	112	0,8482	
	Lúky, trávniky	1	22	74	0	5	9	0	111	0,6667	
	Vodné plochy	0	0	0	10	0	0	0	10	1,0000	
	Zastavané plochy	0	2	0	0	12	0	0	14	0,8571	
	Orné/holé pôdy	0	2	3	0	1	14	0	20	0,7000	
	Tieň	1	1	0	0	0	1	7	10	0,7000	
Presnosť	Celkom	92	151	88	10	19	24	12	396		
	Spracovateľská	0,9348	0,6291	0,8409	1,0000	0,6316	0,5833	0,5833		0,7525	
	Kappa										0,6810

Pléiades – multispektrálne pásma

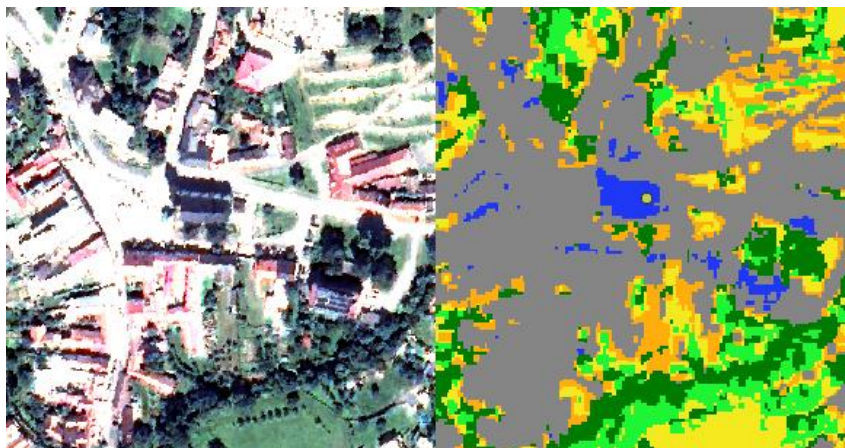
Výsledky klasifikácie multispektrálnych pásiem v rozlíšení 2 m boli dosiahnuté najlepšie pomocou algoritmu Random Forest, s počtom stromov 250, zvýšeným počtom trénovacích vzoriek na 50 000 a zároveň s vegetačným indexom NDVI a SAVI. Bola dosiahnutá celková presnosť 86 % a Kappa koeficient 0,83. Vo všetkých kategóriách bola dosiahnutá užívateľská presnosť nad 50 %, 100 % bolo v dvoch kategóriách, a to lúky, resp. trávnaté oblasti a vodné plochy. Vzhľadom na problematickú kategóriu holých/orných pôd, v tejto klasifikácii došlo k pomerne dobrým výsledkom, kde iba jeden bod bol nesprávne zaradený do kategórie lúky.

Tabuľka 7 Chybová matica pri klasifikátore RF so 70 kontrolnými bodmi

		Validačné data						Presnosť		
		Ihličnaté lesy	Listnaté lesy	Lúky, trávniky	Vodné plochy	Zastavané plochy	Orné/holé pôdy	Celkom	Používateľská	Kappa
Klasifikácia	Ihličnaté lesy	11	1	3	0	0	0	15	0,7333	
	Listnaté lesy	1	13	1	0	0	0	15	0,8667	
	Lúky, trávniky	0	0	15	0	0	0	15	1,0000	
	Vodné plochy	0	0	0	7	0	0	7	1,0000	
	Zastavané plochy	0	0	0	0	6	3	9	0,6667	
	Orné/holé pôdy	0	0	1	0	0	8	9	0,8889	
Presnosť	Celkom	12	14	20	7	6	11	70		
	Spracovateľská	0,9167	0,9286	0,7500	1,0000	1,0000	0,7273		0,8571	
	Kappa									0,8253

Podobná presnosť bola dosiahnutá klasifikátorom Maximum Likelihood, kde pri zmene z pôvodného nastavenia parametrov na veľkosti trénovacích vzoriek na 50 000 a pridaní NDVI indexu sa zvýšila presnosť o 30 %, a to na 84 % s Kappa koeficientom 0,8. Všetky body boli správne zaradené v rovnakých kategóriách lúk a vodných plôch. Horšie výsledky boli pri ihličnatom lese, kde boli 4 body zaradené ako listnatý les a dva body ako trávnatá plocha.

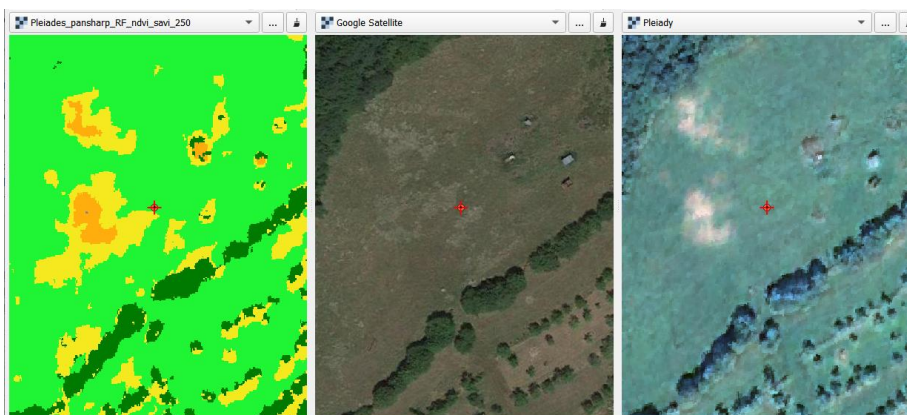
Klasifikátorom KNN bola získaná celková presnosť 61 %, kde ani pri zmene parametrov nedošlo k zlepšeniu. Po pridaní NDVI indexu a zmene počtu najbližších susedov na 20 boli správne roztriedené všetky body označené za kategóriu orná pôda a lúka, avšak roztriedenie lesov nebolo presné. Došlo aj k problému pri vyznačovaní vodných plôch, za ktoré algoritmus označoval aj tieň (viď obrázok 8). Tento problém sa vyskytuje hlavne na snímkach Pléiades vzhľadom k vysokému rozlíšeniu.



Obrázok 9 Zatižené oblasti označené ako vodné plochy

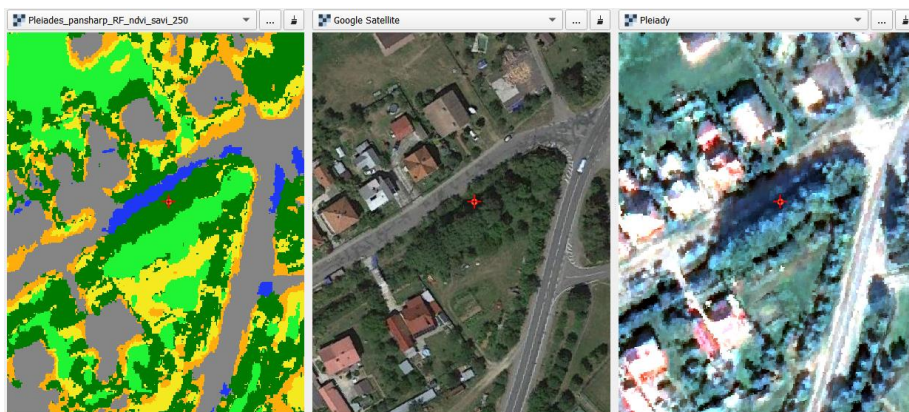
Pléiades – po pansharpeningu

Snímok s najlepším rozlíšením nevyhnikal pri žiadnom algoritme klasifikácie. Najlepší výsledok dosiahol použitím klasifikátora riadenej klasifikácie Random Forest pri nastavení 250 stromov s vegetačným indexom NDVI a SAVI – celková presnosť 78,6 % a Kappa koeficient 0,63. Pri klasifikácii došlo k správne označeniu všetkých bodov označujúcich trávnaté plochy, avšak spracovateľská presnosť je chybná. Prevažná časť územia je označená ako listnatý les, 6 bodov z 15 je v realite trávnatou plochou (viď Tabuľka 7).



Obrázok 10 Trávnaté plochy nesprávne označené ako lesy

Oblasť s tieňom od stromov je chybné označená ako ihličnatý les a zároveň vodná plocha. Po overení pomocou Google Street View je potvrdené, že sú to iba listnaté stromy.



Obrázok 11 Zatiernená oblasť vyznačená ako ihličnatý les a vodná plocha



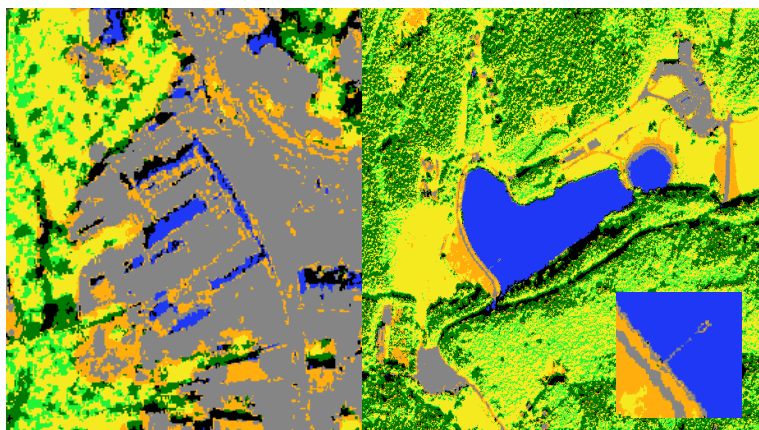
Obrázok 12 Záber z Google Street View na reálnu scénu

Tabuľka 8 Chybová matica pri klasifikátore RF so 70 kontrolnými bodmi

		Validačné data						Presnosť		
		Ihličnaté lesy	Listnaté lesy	Lúky, trávniky	Vodné plochy	Zastavané plochy	Orné/holé pôdy	Celkom	Používateľská	Kappa
Klasifikácia	Ihličnaté lesy	12	2	1	0	0	0	15	0,8000	
	Listnaté lesy	0	9	6	0	0	0	15	0,6000	
	Lúky, trávniky	0	0	15	0	0	0	15	1,0000	
	Vodné plochy	0	0	0	5	2	0	7	0,7143	
	Zastavané plochy	0	0	0	0	7	2	9	0,7778	
	Orné/holé pôdy	0	0	2	0	0	7	9	0,7778	
Presnosť	Celkom	12	11	24	5	9	9	70		
	Spracovateľská	1,0000	0,8182	0,6250	1,0000	0,7778	0,7778		0,7857	
	Kappa									0,6293

Pri tomto nastavení došlo k viacerým nedostatkom. Prevažná časť plochy je označená ako listnatý les, z kontrolných bodov je v realite 6 z nich trávnatá plocha.

Najslabší výsledok (55 %) bol dosiahnutý použitím neriadenej klasifikácie K-means pri vytvorení 100 zhlukov. Rovnako ako pri RF, aj tu na prvý pohľad vidieť, že v klasifikácii prevláda kategória trávnatých plôch, ktorá nesedí s reálnym snímkom. Pri kategorizovaní do 6 hlavných kategórií bolo pri jednotlivých skupinách prihliadané na väčšie zhluky pixelov, následkom čoho pri opakovaných situáciách, kde bola za trávnatú plochu označená celá lúka a zároveň roztrúsené pixely v lese vznikla veľká plocha trávnatých porastov. Napriek detailnému označeniu vodných plôch, kde sa podarilo podrobne vyznačiť aj mólo, došlo v rovnakej kategórii aj k označeniu tieňov zo zástavieb. Pri označovaní ihličnatých lesov je diskutabilné, či išlo o správne označenie ihličnatých stromov, alebo boli vyznačované tiene medzi stromami.



Obrázok 13 Označenie tieňov zo zástavieb ako vodná plocha (vľavo). Detailne označená vodná plocha s detailom na mólo (vpravo)

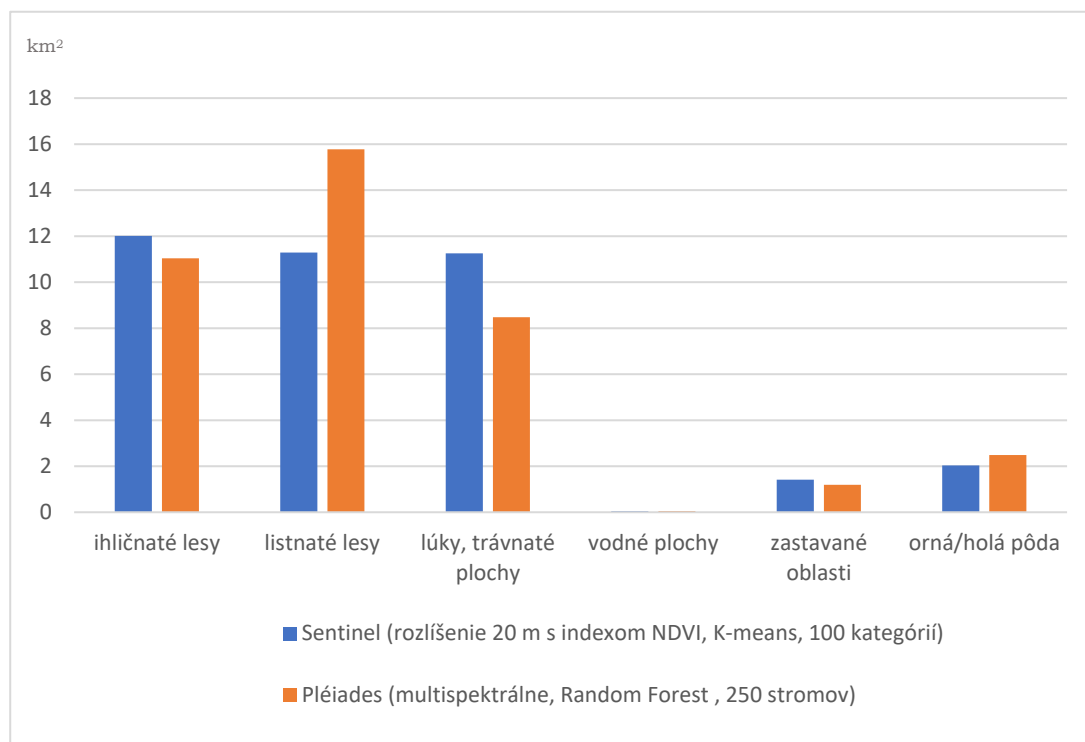
Pri overení najlepšieho výsledku použitím 204 kontrolných bodov pri algoritme Random Forest s 250 stromami došlo k zníženiu celkovej presnosti o 25 % (viď Tabuľka 9). K problémom dochádzalo pri rozlišovaní typov lesov, kde sa užívateľská presnosť pohybuje na 50 % pri oboch typoch, väčším problémom bolo rozlišovanie vodných plôch, kde boli správne zaradené iba 3 kontrolné body. Zlý výsledok značí aj Kappa koeficient, ktorý má hodnotu 0,37.

Tabuľka 9 Chybová matica pri klasifikátore RF s 204 kontrolnými bodmi

		Validačné data						Presnosť		
		Ihličnaté lesy	Listnaté lesy	Lúky, trávniky	Vodné plochy	Zastavané plochy	Orné/holé pôdy	Celkom	Používateľská	Kappa
Klasifikácia	Ihličnaté lesy	48	25	17	0	1	1	92	0,5217	
	Listnaté lesy	9	36	21	0	0	0	66	0,5455	
	Lúky, trávniky	2	3	14	0	0	2	21	0,6667	
	Vodné plochy	0	0	0	3	7	0	10	0,3000	
	Zastavané plochy	0	1	0	0	5	1	7	0,7143	
	Orné/holé pôdy	0	1	2	0	1	4	8	0,5000	
Presnosť	Celkom	59	66	54	3	14	8	204		
	Spracovateľská	0,8136	0,5455	0,2593	1,0000	0,3571	0,5000		0,5392	
	Kappa									0,37

Porovnanie rozlôh

V jednotlivých klasifikáciách došlo k menším či väčším rozdielom v rozložení a veľkosti plôch vybraných kategórií. V nasledujúcom grafe sú vyjadrené veľkostné zastúpenia plôch v klasifikáciách s najlepším výsledkom spomedzi satelitu Pléiades (multispektrálne, Random Forest s 250 stromami) a Sentinel (rozlíšenie 20 m s indexom NDVI, K-means, 100 kategórií). Plocha vodných plôch je v porovnaní s ostatnými kategóriami príliš malá, a rozdiel je v grafe nepozorovateľný – rozloha 0,0434 km² pri Sentineli, 0,0523 km² pri Pléiades.



Graf 1 Porovnanie rozlôh na snímku Pléiades a Sentinel

Porovnanie výsledkov klasifikácie s validačnými dátami

Pôvodné kategórie v DKV boli zlúčené do 6 kategórií, ktoré boli použité pri klasifikácii v tejto bakalárskej práci. Zvyšné, nezaradené kategórie boli pri porovnávaní z vrstvy vymazané.

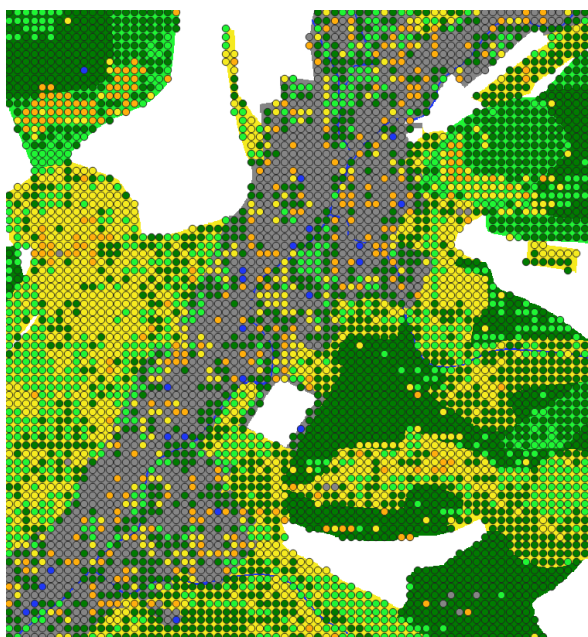
Pomocou funkcie *Create grid* boli v QGISe vytvorené pre každú klasifikáciu tri pravidelné mriežky bodov so vzdialenosťami medzi sebou 10, 20 a 50 metrov, ktorá sa nachádza iba na ploche orezanej DKV. Ich účelom je porovnávať zhodu klasifikácie s validačnými dátami v troch rôznych mierkach. V prvom rade bola z pretriedenej DKV polygónovej vrstvy vytvorená rastrová vrstva s veľkosťou pixelu pol metra kvôli požadovanej presnosti hraníc vyznačených oblastí. V ďalšom kroku boli hodnoty rastru DKV extrahované do bodovej mriežky funkciou *Sample raster values*. Rovnaký postup bol aplikovaný na porovnávané klasifikácie. Z extrahovaných bodov boli pomocou funkcie *Statistics by categories* zistené počty bodov v jednotlivých kategóriách, ktoré boli vložené do Excel tabuľky. V nej došlo k porovnaniu počtu bodov v jednotlivých klasifikáciách, snímkach, mierkach s porovnávanou vrstvou DKV. Zo získaných hodnôt bolo vypočítané percentuálne zastúpenie v kategóriách.

Iba vizuálnym porovnaním lokality z klasifikovaného snímku a DKV vidieť podobnosť v rozložení jednotlivých kategórií. Najväčším rozdielom vyzerá byť rozdiel v kategórii holá pôda, ktorá podľa DKV je vyznačená iba na miestach polí, zatiaľ čo pri tvorbe tréningových množín boli vyznačené aj iné netravnaté plochy iba s hlinou.

Pri vizuálnom porovnaní obsahuje vrstva DKV vodné plochy naprieč celým územím, objavuje sa tam viacero jazierok, ktoré sú nevšimnuteľné na používaných snímkach, a zároveň potoky tečúce cez zastavané oblasti. Pri klasifikáciách sa objavujú modré pixely na podobných miestach, avšak vo väčšine prípadov ide o pomýlenie algoritmu so zástavbou, málokedy dochádza k prieniku s plochou reálnej vodnej plochy.

Bodovo boli s DKV porovnávané nastavenia klasifikácie pri jednotlivých sadách snímok, ktoré v celkovej presnosti získali najviac percent. Celková tabuľka s porovnaním v počtoch bodov a percentuálnym zastúpením jednotlivých kategórií sa nachádza v prílohe č. 6.

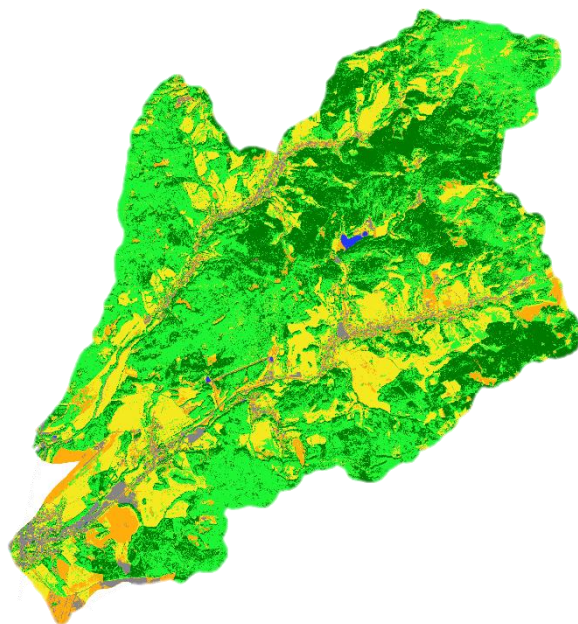
Najväčšie problémy sa vyskytujú pri snímkach Pléiades po pansharpeningu. Vysoké rozlíšenie zapríčiňuje menšiu homogenitu dát, čím spôsobuje nižšiu zhodu s vrstvou DKV, kde sú polygóny jednotvárne. Tomu by sa dalo predísť postklasifikačnou úpravou a použitím *majority filter*. Ten odstraňuje „salt-and-pepper“ efekt, ktorý odstráni izolované pixely a ich hodnotu mení na väčšinovú, predpokladanú správnu hodnotu (Lillesand, Kiefer, Chipman 2015). Pri riadenej klasifikácii s metódou Maximum Likelihood dochádza k najväčšej zmene v zastúpení ihličnatého lesa, oproti referenčným dátam DKV so zastúpením 30,9% sa jeho zastúpenie zvýšilo na 53,6%. Podobné výsledky sú vo všetkých troch testovaných mierkach.



Obrázok 14 Body so symbolológiou podľa klasifikácie RF s nastavením 250 stromov a 50 000 tréningových vzoriek na podklade DKV

Pri porovnaní DKV a klasifikácie, ktorá mala podľa chybovej matice najvyššiu presnosť, a to multispektrálny snímok Pléiades pri klasifikátore RF s nastavením 250 stromov a 50 000 tréningových plôch dochádza k výrazným percentuálnym zmenám v rozložení krajiny. Podľa klasifikátora sa v záujmovom území nachádza o 10 % viac listnatých lesov ako pri tvorbe DKV, čo vzhľadom k 10 ročnému rozostupu je nepravdepodobné.

Klasifikácia ukazuje aj menšiu plochu zastavaných oblastí, za čím môžu byť aj presnejšie obrysy pri klasifikácii Pléiades, na čo má vplyv vysoké rozlíšenie snímku.

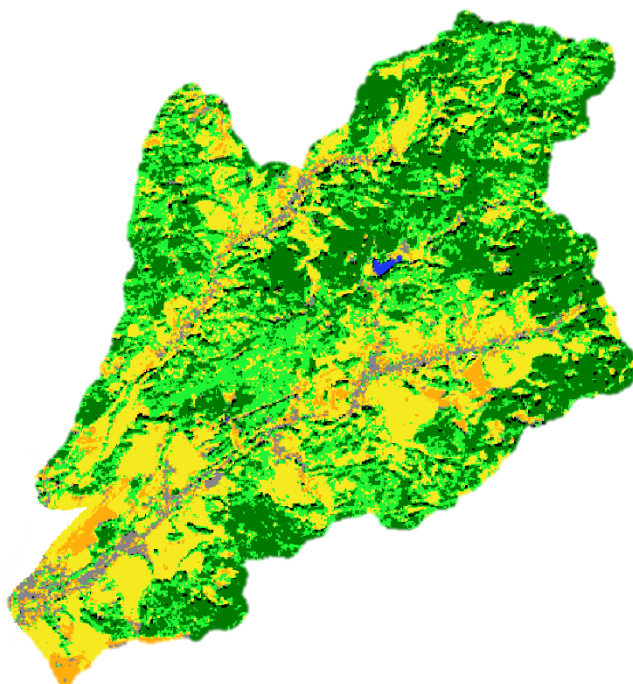


Obrázok 15 Klasifikácia Random Forest pri multispektrálnom snímku Pléiades

Tabuľka 10 Porovnanie zastúpenia plôch DKV a RF pri snímku z Pléiades

	DKV	RF
Ihličnaté lesy	30,85%	27,51%
Listnaté lesy	26,03%	38,04%
Trávnaté plochy	34,05%	23,93%
Vodné plochy	0,76%	0,16%
Zastavané plochy	6,88%	3,39%
Holé pôdy	1,43%	6,96%

Klasifikácia Sentinel-2 snímku s priestorovým rozlíšením 20 m s indexom NDVI a použitým klasifikátorom K-means sa percentuálne najviac podobá DKV. Najväčší rozdiel je v zastúpení orných pôd, ktorá je v DKV striktne vymedzená iba na plochy orných pôd, zatiaľ čo pri klasifikácii je na základe trénovacích plôch označovaná každá holá pôda.



Obrázok 16 Klasifikácia K-means pri snímku Sentinel-2 s rozlíšením 20 m

Tabuľka 11 Porovnanie zastúpenia plôch DKV a K-means pri snímku zo Sentinel-2

	DKV	K-means
Ihličnaté lesy	30,85%	30,12%
Listnaté lesy	26,03%	24,82%
Trávnaté plochy	34,05%	27,99%
Vodné plochy	0,76%	0,15%
Zastavané plochy	6,88%	9,25%
Holé pôdy	1,43%	7,67%

7 DISKUSIA

Počas spracovávanía práce došlo k niekoľkým problémom, ktoré by pre kvalitnejšie výsledky bolo potrebné v budúcnosti vyriešiť.

K prvému problému došlo pri počítačnom spracovávaní snímkov. Nakoľko scéna bola pri satelite Pléiades rozdelená v dvoch snímkach a bolo potrebné ich spojiť, rozdielne dátumy v období zbierania úrody, to znamená jeden snímok z augusta a druhý zo septembra, tvoria problém pri realistickom vyobrazení vegetačného povrchu. Pre samotnú klasifikáciu to neznamená problém, avšak získané dáta sa netýkajú jedného termínu. Rovnako by bol ideálnejší prípad, kde by snímok Sentinelu pochádzal z rovnakého obdobia.

Rozdielne dátumy sú už problémom pri validácii klasifikácií. Pre porovnanie presnosti dáta nie sú relevantné, nakoľko za 9 rokov mohlo dôjsť k výrazným zmenám vo všetkých klasifikovaných kategóriách. Preto táto vrstva poslúžila skôr ako porovnanie zmeny za dané obdobie.

Pri viacerých klasifikáciách došlo k problému pri klasifikovaní oblastí s tieňmi. Tieto situácie sa riešili viac pri snímkach Pléiades kvôli ich vysokému priestorovému rozlíšeniu, zatiaľ čo pri nižšom rozlíšení Sentinelu tieň neboli tak viditeľné. Podľa metadat snímkov Pléiades boli zhotovené okolo 10 hodiny ráno, čo v danom období spôsobuje dlhší tieň ako je samotný objekt. V prípade o niečo vyššej výšky Slnka by boli tieň menšie, avšak problém s ich klasifikáciou by sa stále nevyriešil. Možnosť správnejšej klasifikácie je vytvoriť kategóriu tieňov, rovnako ako bola v tejto práci pri neriadenej klasifikácii a pridať ich do tréningových plôch, aby mal algoritmus šancu ich rozpoznať a zaradiť správne.

Ďalším nedostatkom bolo množstvo „salt-and-pepper“ efektu, ktorý sa vyskytoval hlavne pri snímkach s veľmi vysokým rozlíšením. Pri tvorbe chybovej matice niekedy dochádzalo k ich označeniu, čo mohlo skresliť výsledky, rovnako ako pri porovnávaní s DKV. Dalo by sa tomu predísť napríklad pri post-klasifikačnej úprave použitím majority filtra, ktorý by malé plochy odstránil, avšak kvôli testovaniu práce čistého algoritmu filter nebol v práci použitý.

Lepšie výsledky bez „salt-and-pepper“ efektu by mohli byť dosiahnuté použitím objektivej klasifikácie, ktorá v zadaní práce bola spomenutá ako možnosť, avšak kvôli výberu programu SNAP, ktorý objektívnu klasifikáciu neposkytuje, nebola vo finále vykonaná.

Rovnako, ako bolo v práci spomenuté, tréningové plochy sú dôležitou súčasťou a majú významný vplyv na výsledky klasifikácie. Pokiaľ plochy neboli správne vybrané, mohlo dôjsť k zhoršenej klasifikácii. Vo výsledkoch bolo vidieť veľký nedostatok pri rozlišovaní zastavaných plôch a holých pôd, aj po skúšaní rôznych indexov. Pri kontrole spektrálnych vlastností tried bolo zistené, že dané triedy majú podobné spektrálne chovanie a vyskúšané algoritmy mali problém ich rozlíšiť.

8 ZÁVER

Cieľom bakalárskej práce bolo vytvoriť a porovnať klasifikácie rôznych algoritmov v oblasti riadenej a neriadenej klasifikácie, aplikovaných na 4 sadách snímok satelitov s veľmi vysokým rozlíšením (Pléiades) a vysokým rozlíšením (Sentinel). S cieľom dosiahnuť presnejšie výsledky, boli použité aj vegetačné indexy. Vytvorené klasifikácie boli zhodnotené prostredníctvom chybových matíc, z ktorých boli získané hodnoty presnosti klasifikácie. Ďalším cieľom práce bolo porovnať klasifikácie s DKV v troch rôznych mierkach, ktoré znázorňujú tri rôzne detaily povrchu krajiny. Následne boli vytvorené výstupy vo forme tabuliek a máp.

V teoretickej časti práce bola urobená rešerš danej problematiky, kde boli preštudované práce s podobným, alebo rovnakým zámerom. Z týchto štúdií vyplynulo, že v mnohých prípadoch dosiahli najlepšie výsledky klasifikácie pri použití algoritmu Maximum Likelihood v kombinácii so snímkom vysokého rozlíšenia.

Pre vykonanie klasifikácie boli vybrané štyri algoritmy dostupné v programe SNAP: Maximum Likelihood, Random Forest, K-nearest neighbors a K-means cluster analysis. Podľa potreby boli skúšané rozdielne parametre algoritmov a kombinácie pásiem, kým nebol získaný najlepší výsledok. Po vykonaní klasifikácie boli výstupy vyexportované do softwaru QGIS, kde boli dodatočne upravené a následne boli porovnávané.

Aj keď očakávaným výsledkom bol najlepší výstup pri snímku z Pléiades vzhľadom k najlepšiemu rozlíšeniu, v praxi bola väčšia presnosť pre danú mierku pri snímku Sentinel-2 s najnižším rozlíšením, ale spektrálnymi pásmami navyše. Pre podrobnú klasifikáciu sú snímky od Pléiades vyhovujúce, avšak príliš veľký detail vytváral problematické miesta vo vybraných klasifikovaných kategóriách. V zvolenej mierke je prospešnejšia prítomnosť viacerých spektrálnych pásiem, ako veľmi vysoké rozlíšenie.

Výsledky môžu byť užitočné pri výbere potrebných dát a algoritmov pri monitorovaní zmien pokryvu.

POUŽITÁ LITERATÚRA A INFORMAČNÉ ZDROJE

A CANADA CENTRE FOR REMOTE SENSING REMOTE SENSING TUTORIAL, 2019. *Fundamentals of Remote Sensing* [online]. Natural Resources Canada. Cit z : https://natural-resources.canada.ca/sites/www.nrcan.gc.ca/files/earthsciences/pdf/resource/tutor/fundam/pdf/fundamentals_e.pdf

Accuracy Metrics, no date [online]. Cit z : https://gsp.humboldt.edu/olm/Courses/GSP_216/lessons/accuracy/metrics.html [cit 10 august 2023].

AL-DOSKI, Jwan, MANSOR, Shattri B a SHAFRI, Helmi Zulhaidi Mohd, 2013b. Image Classification in Remote Sensing. [online]. Vol. 3. Cit z : <https://core.ac.uk/download/pdf/234663192.pdf>

ANTOGNELLI, Sara, 2018. NDVI and NDMI vegetation indices: instructions for use. *Agricolus* [online]. 28 máj 2018. Cit z : <https://www.agricolus.com/en/vegetation-indices-ndvi-ndmi/> [cit 10 august 2023].

CAMPBELL, James B. a WYNNE, Randolph H., 2011. *Introduction to Remote Sensing, Fifth Edition*. Guilford Press. ISBN 978-1-60918-177-2. Google-Books-ID: NkLmDjSS8TsC

DELWART, Steven, 2015. ESA Standard Document. [online]. Číslo 1. Cit z : https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook

DOBROVOLNÝ, Petr, 1998. *Dálkový průzkum Země, Digitální zpracování obrazu* [online]. MU Brno. ISBN 978-80-210-1812-9. Cit z : <https://is.muni.cz/publication/205712/cs/Dalkovy-pruzkum-Zeme-Digitalni-zpracovani-obrazu/Dobrovolny> [cit 10 august 2023].

ETTEHADI OSGOUEI, Paria et al., 2019. Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. *Remote Sensing*. Vol. 11, číslo 3, s. 345. DOI 10.3390/rs11030345.

GAJA, Pavel, 2013. *Analýza pudní vlhkosti a charakteristik sucha prostředky DPZ* [online]. Cit z : https://is.muni.cz/th/hyj9g/DP_Gaja.pdf

Introduction to Remote Sensing [online]. Cit z : <https://seos-project.eu/remotesensing/remotesensing-c01-p06.html> [cit 10 august 2023].

ISMAIL, Mohd Hasmadi, PAKHRIAZAD, HZ a SHAHRIN, MF, 2009. Evaluating supervised and unsupervised techniques for land cover mapping using remote sensing data. *Geografia : Malaysian Journal of Society and Space*.

ISODATA [online]. Cit z : <https://www.nv5geospatialsoftware.com/docs/isodataclassification.html> [cit 3 august 2023].

JENKINS, Keith, no date. LibGuides: GIS: Geographic Information Systems: QGIS. [online]. Cit z : <https://guides.library.cornell.edu/gis/qgis> [cit 15 august 2023].

JENSEN, John R., 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Prentice Hall. ISBN 978-0-13-145361-6. Google-Books-ID: F84PAQAIAAJ

JIXIAN, Sun Xiaoxia Zhang a ZHENGJUN, Liu, no date. A COMPARISON OF OBJECT-ORIENTED AND PIXEL-BASED CLASSIFICATION APPROACHS USING QUICKBIRD

IMAGERY. [online]. Cit z : https://www.isprs.org/proceedings/xxxvi/2-W25/source/A_COMPARISON_OF_OBJECT-ORIENTED_AND_PIXEL-BASED_CLASSIFICATION_APPROACHS_USING_QUICKBIRD.pdf

KIAGE, Lawrence et al., 2007. Recent land-cover/use change associated with land degradation in the Lake Baringo catchment, Kenya, East Africa: Evidence from Landsat TM and ETM+. *International Journal of Remote Sensing - INT J REMOTE SENS*. Vol. 28, s. 4285–4309. DOI 10.1080/01431160701241753.

LILLESAND, Thomas, KIEFER, Ralph W. a CHIPMAN, Jonathan, 2015. *Remote Sensing and Image Interpretation*. John Wiley & Sons. ISBN 978-1-118-34328-9. Google-Books-ID: AFHDCAAAQBAJ

LIU, Desheng a XIA, Fan, 2010. Assessing object-based classification: advantages and limitations. *Remote Sensing Letters*. Vol. 1, číslo 4, s. 187–194. DOI 10.1080/01431161003743173.

Metody dálkového průzkumu v precizním zemědělství, no date www.agroportal24h.cz [online]. Cit z : <https://www.agrojournal.cz/clanky/metody-dalkoveho-pruzkumu-v-preciznim-zemedelstvi-263> [cit 10 august 2023].

MINÁRČIK, Miroslav, 2015. KLASIFIKÁCIA KRAJINNÉHO POKRYVU HORSKÝCH EKOSYSTÉMOV S VYUŽITÍM OBRAZOVÝCH DÁT RÔZNEHO PRIESTOROVÉHO A SPEKTRÁLNEHO ROZLIŠENIA. [online]. Cit z : https://dspace.cuni.cz/bitstream/handle/20.500.11956/65689/BPTX_2014_1_11310_0_385844_0_159968.pdf?sequence=1&isAllowed=y

MYINT, Soe W. et al., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*. Vol. 115, číslo 5, s. 1145–1161. DOI 10.1016/j.rse.2010.12.017.

Naše družice | Copernicus [online]. Cit z : <https://www.copernicus.eu/cs/o-programu-copernicus/infrastruktura/nase-druzice> [cit 10 august 2023].

NDMI: Vegetation Index Equation And Values Interpretation, 2022 [online]. Cit z : <https://eos.com/make-an-analysis/ndmi/> [cit 10 august 2023].

NJAMBI, Rose, 2022. How pansharpening improves satellite imagery. *UP42 Official Website* [online]. 26 júl 2022. Cit z : <https://up42.com/blog/how-pansharpening-improves-satellite-imagery> [cit 15 júl 2023].

Pleiades - Earth Online, no date [online]. Cit z : <https://earth.esa.int/eogateway/missions/pleiades> [cit 15 júl 2023].

PROCHÁZKA, Bc Jan, 2020. STANOVENÍ ESAI INDEXU PRO ČESKOU REPUBLIKU. [online]. Cit z : https://www.geoinformatics.upol.cz/dprace/magisterske/prochazka20/data/dp_prochazka_web.pdf

PURKYT, Ing Jan, 2022. Produkční a evapotranspirační funkce jednotlivých kategorií využití krajiny v měnících se klimatických podmínkách. [online]. Cit z : https://theses.cz/id/jn10gr/DSP_Purkyt.pdf?zpet=%2Fvyhledavani%2F%3Fsearch%3Ddetailni%20kombinovana%20vrstva%26start%3D1

RCOMBEAU, 2021. Blog Hiphen - What is a Vegetation Index? *Hiphen* [online]. 19 apríl 2021. Cit z : <https://www.hiphen-plant.com/vegetation-index/3582/> [cit 10 august 2023].

RIEDL, Martin, 2016. *Klasifikace lesní vegetace pomocí dat DPZ vysokého a velmi vysokého prostorového rozlišení* [online]. Cit z : https://dspace.cuni.cz/bitstream/handle/20.500.11956/74369/BPTX_2015_1_11310_0_412626_0_173911.pdf?sequence=1&isAllowed=y

ROUIBAH, K. a BELABBAS, M., 2020. Applying Multi-Index approach from Sentinel-2 Imagery to Extract Urban Area in dry season (Semi-Arid Land in North East Algeria). *Revista de Teledetección*. Číslo 56, s. 89–101. DOI 10.4995/raet.2020.13787.

SATHYA, P a DEEPA, V Baby, 2017. Analysis of Supervised Image Classification Method for Satellite Images. *Open Access*. Vol. 5, číslo 2.

SAVI and NDVI spectral index comparison, no date *GIS and Earth Observation University* [online]. Cit z : <https://www.geo.university/pages/savi-and-ndvi-spectral-index-comparison> [cit 10 august 2023].

Sentinel Toolbox Help, no date [online]. Cit z : <https://seadas.gsfc.nasa.gov/help-8.0.0/general/overview/SnapOverview.html> [cit 10 august 2023].

Sentinel-2A Satellite Sensor | Satellite Imaging Corp, no date [online]. Cit z : <https://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/sentinel-2a/> [cit 10 august 2023].

SIMERSKÁ, Denisa, 2021. Aplikační možnosti DPZ v monitoringu kulturní zemědělské krajiny. [online]. Cit z : https://is.muni.cz/th/rue00/Diplomova_prace.pdf?zpet=https:%2F%2Ftheses.cz%2Fvyhledavani%2F%3Fsearch%3Dvegeta%C4%8Dn%C3%AD%20indexy%20rvi%26start%3D1

SR—ArcGIS Pro | Documentation, [online]. Cit z : <https://pro.arcgis.com/en/pro-app/latest/arcpy/image-analyst/sr.htm> [cit 10 august 2023].

THANH NOI, Phan a KAPPAS, Martin, 2018. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors*. Vol. 18, číslo 1, s. 18. DOI 10.3390/s18010018.

VLAČIHOVÁ, Martina, 2018. *Využití automatické klasifikace družicových snímků pro sledování změn v krajině* [online]. Cit z : <https://portal.upce.cz/portal/studium/prohlizeni.html>

VOZÁR, Tomáš, 2020. *Klasifikace lesní vegetace pomocí družicových dat vysokého a velmi vysokého prostorového rozlišení* [online]. Cit z : https://dspace.vsb.cz/bitstream/handle/10084/141593/VOZ0051_HGF_B1316_3646R006_2020.pdf?sequence=1&isAllowed=y

WEIH, Robert a RIGGAN, Norman. Object-based classification vs. pixel-based classification: Comparative importance of multi-resolution imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* [online]. Vol. Vol. XXXVIII-4/C7. Cit z : https://www.isprs.org/proceedings/xxxviii/4-c7/pdf/Weih_81.pdf

ZADRAŽIL, Filip, 2021. *Klasifikace land cover change v Etiopii pomocí dat Landsat a Sentinel-2* [online]. Cit z : <https://dspace.cuni.cz/bitstream/handle/20.500.11956/147703/130317397.pdf?sequence=1&isAllowed=y>

ŽUBRIETOVSKÝ, Lukáš, Lenka ŠVÁBOVÁ a Ondřej ŠVÁB. Sentinel 2 – Datové specifikace. CollGS: Data z družic Sentinel v ČR [online]. Praha, 2017 [cit. 2020-03-30].

Dostupné z: <https://collgs.czechspaceportal.cz/wp-content/uploads/2020/08/sentinel2czfinal.pdf>

PRÍLOHY

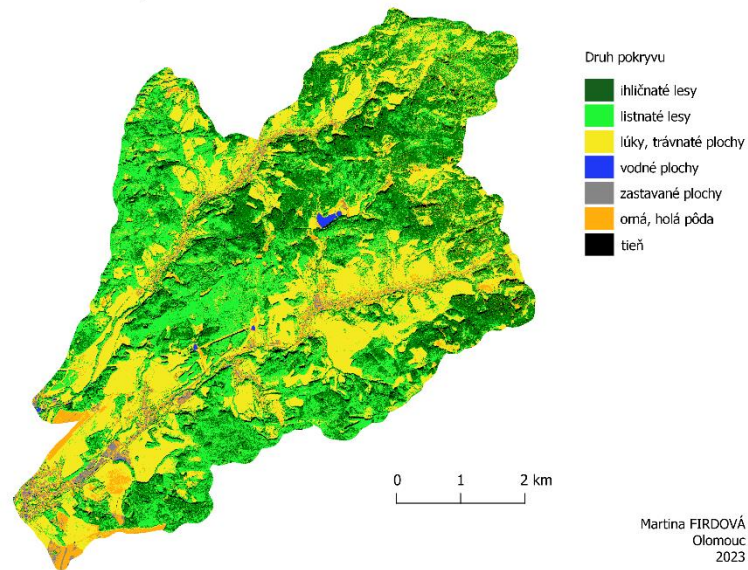
ZOZNAM PRÍLOH

Viazané prílohy:

- Príloha 1 Výstupy klasifikátora K-means
- Príloha 2 Výstupy klasifikátora Maximum Likelihood
- Príloha 3 Výstupy klasifikátora Random Forests
- Príloha 4 Výstupy klasifikátora K-Nearest Neighbor
- Príloha 5 Chybové matice
- Príloha 6 Porovnanie klasifikácií s DKV

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

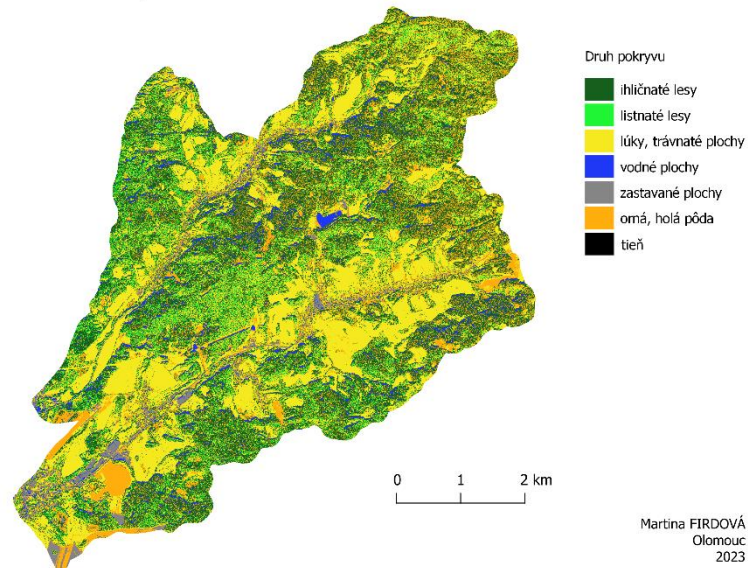
K-means cluster analysis



Obrázok 17 Neriadená klasifikácia K-means po úprave 100 tried na multispektrálnom snímku Pléiades

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

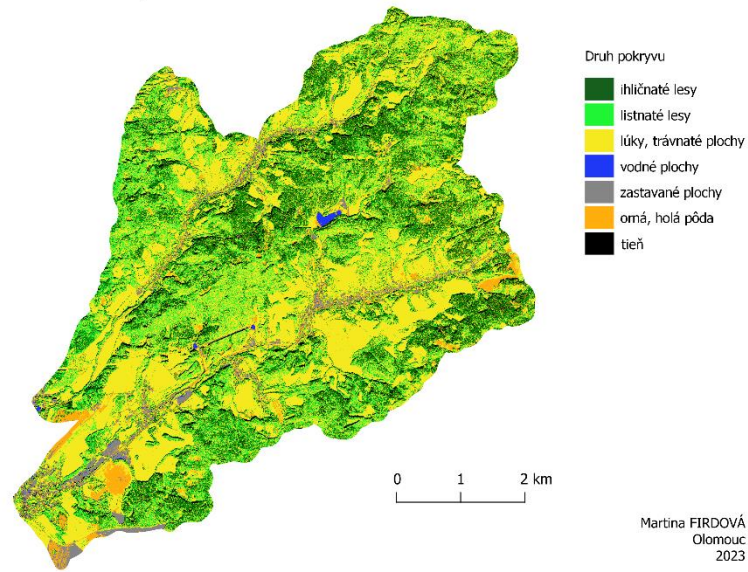
K-means cluster analysis



Obrázok 18 Neriadená klasifikácia K-means po úprave 30 tried na snímku Pléiades po pansharpeningu.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

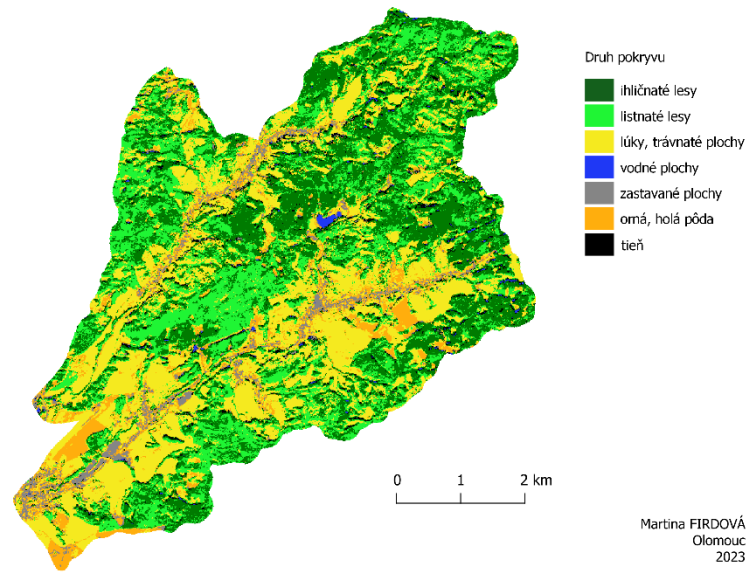
K-means cluster analysis



Obrázok 19 Neriadená klasifikácia K-means po úprave 100 tried na snímku Pléiades po Pansharpeningu

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

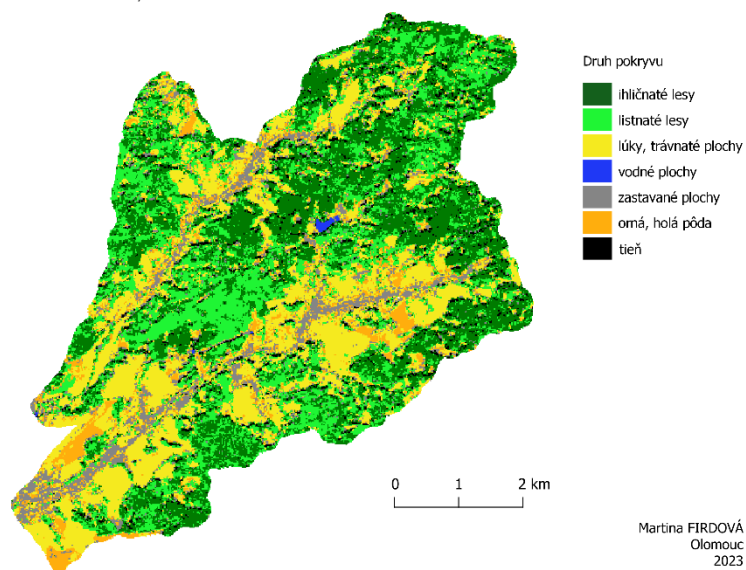
K-means cluster analysis



Obrázok 20 Neriadená klasifikácia K-means po úprave 100 tried na snímku Sentinel-2 s rozlíšením 10 m.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

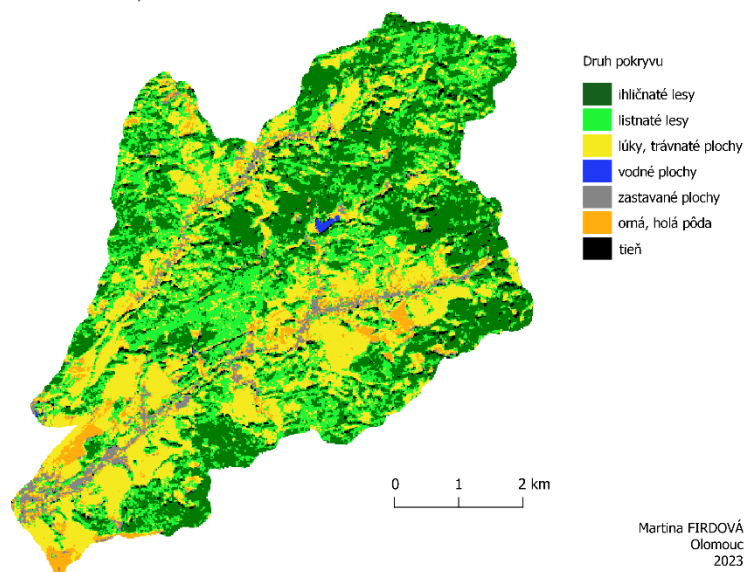
K-means cluster analysis



Obrázok 21 Neriadená klasifikácia K-means po úprave 100 tried na snímku Sentinel-2 s rozlíšením 20 m.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

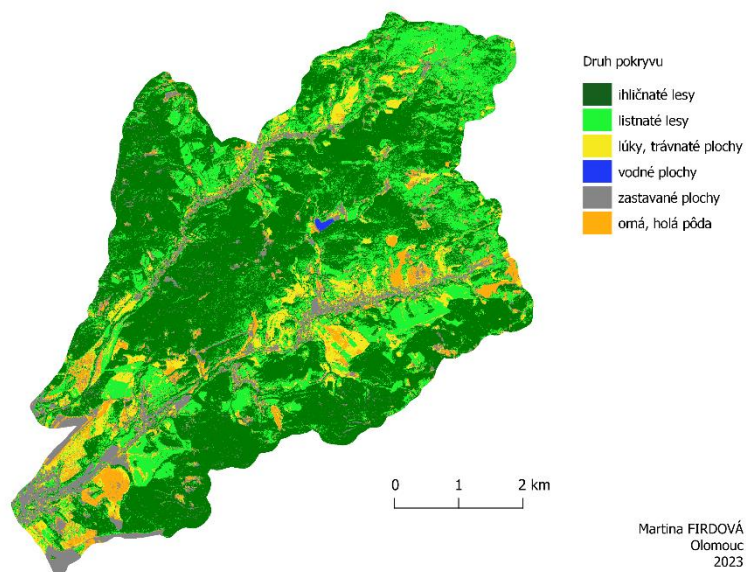
K-means cluster analysis



Obrázok 22 Neriadená klasifikácia K-means po úprave 100 tried na snímku Sentinel-2 s rozlíšením 20 m a indexom NDVI.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

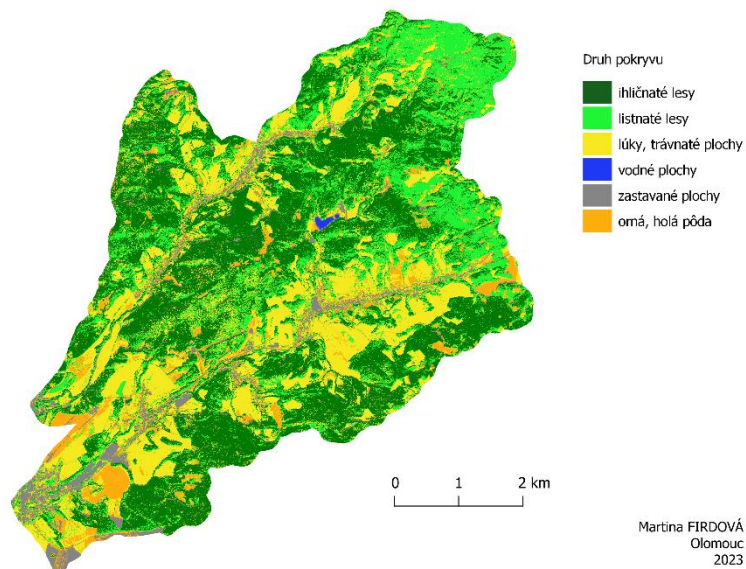
Maximum Likelihood



Obrázok 23 Riadená klasifikácia Maximum Likelihood na multispektrálnom snímku Pléiades.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

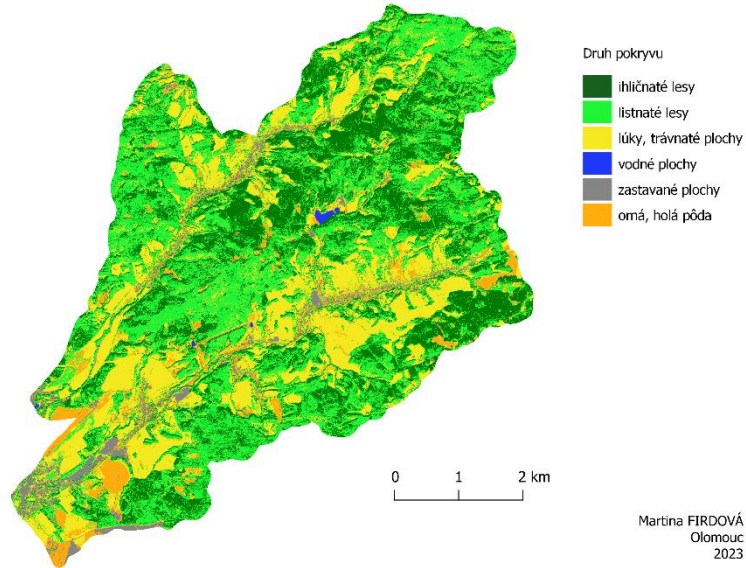
Maximum Likelihood



Obrázok 24 Riadená klasifikácia Maximum Likelihood na multispektrálnom snímku Pléiades s počtom tréningových vzoriek 20 000.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

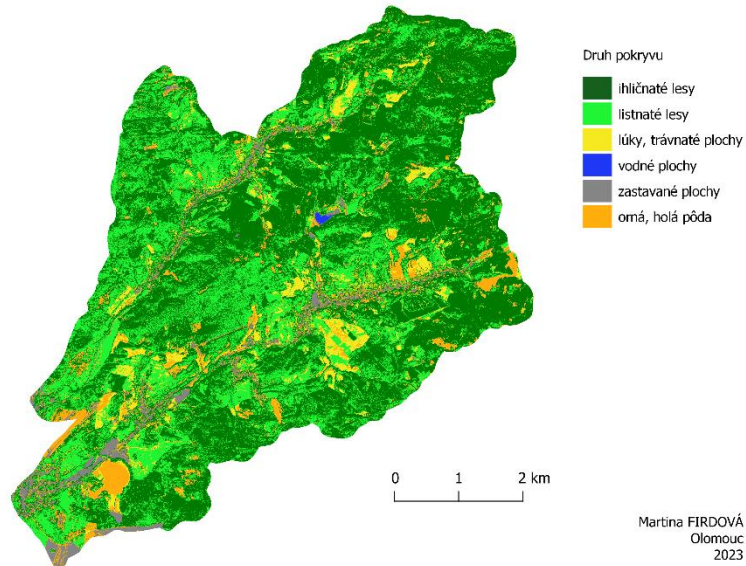
Maximum Likelihood



Obrázok 25 Riadená klasifikácia Maximum Likelihood na multispektrálnom snímku Pléiades s počtom tréningových vzoriek 50 000.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

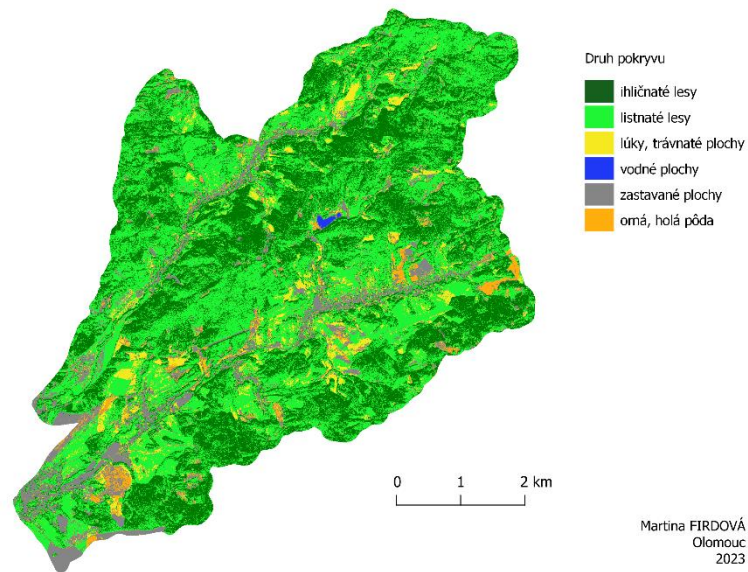
Maximum Likelihood



Obrázok 26 Riadená klasifikácia Maximum Likelihood na snímku Pléiades po pansharpeningu

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

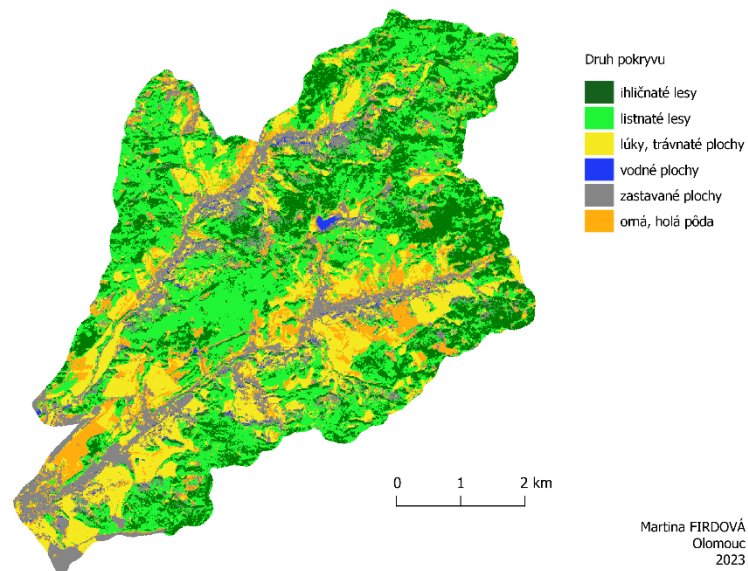
Maximum Likelihood



Obrázok 27 Riadená klasifikácia Maximum Likelihood na snímku Pléiades po pansharpeningu s počtom tréningových vzoriek 10 000.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

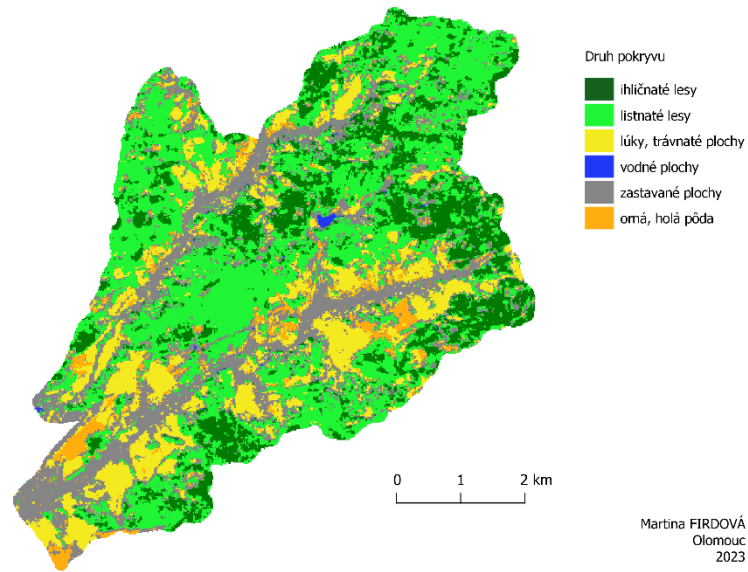
Maximum Likelihood



Obrázok 28 Riadená klasifikácia Maximum Likelihood na snímku Sentinel-2 s priestorovým rozlíšením 10 m.

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

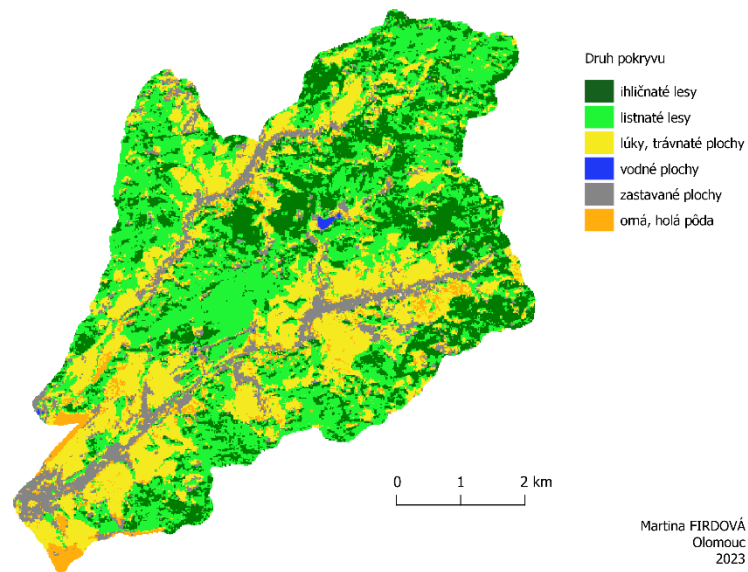
Maximum Likelihood



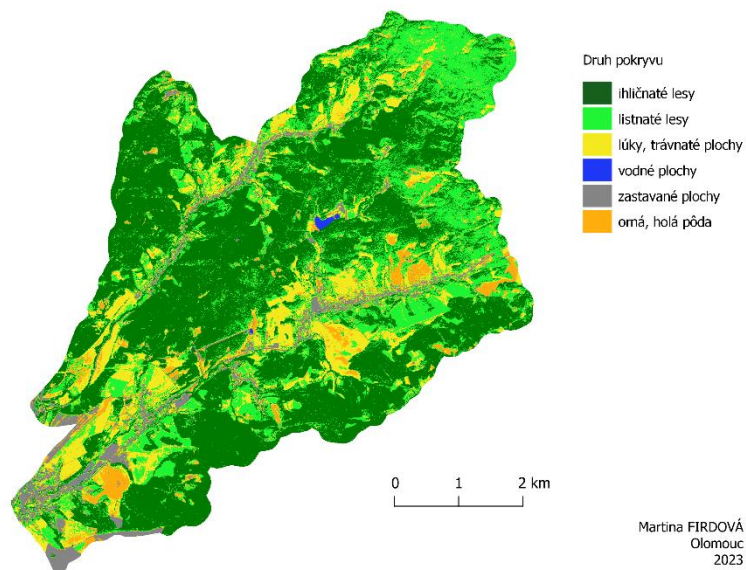
Obrázok 29 Riadená klasifikácia Maximum Likelihood na snímku Sentinel-2 s rozlíšením 20 m

KLASIFIKÁCIA POKRYVU KRAJINY NA ÚZEMÍ POVODIA VŠEMINKY A TRNÁVKY

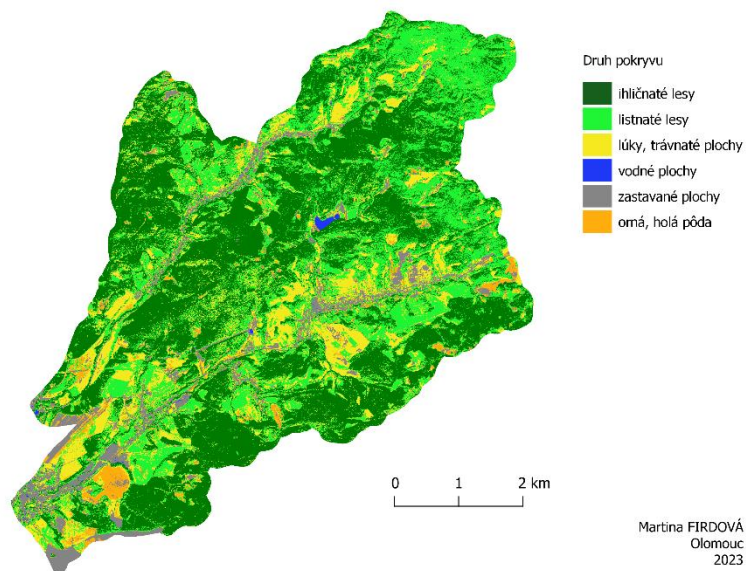
Maximum Likelihood



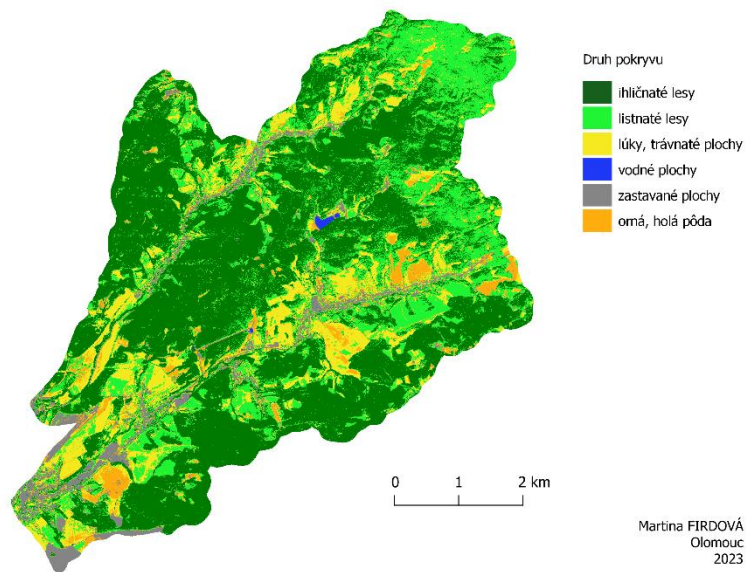
Obrázok 30 Riadená klasifikácia Maximum Likelihood na snímku Sentinel-2 s rozlíšením 20 m s počtom tréningových vzoriek 5 000 a indexami NDVI, SAVI, NDTI.



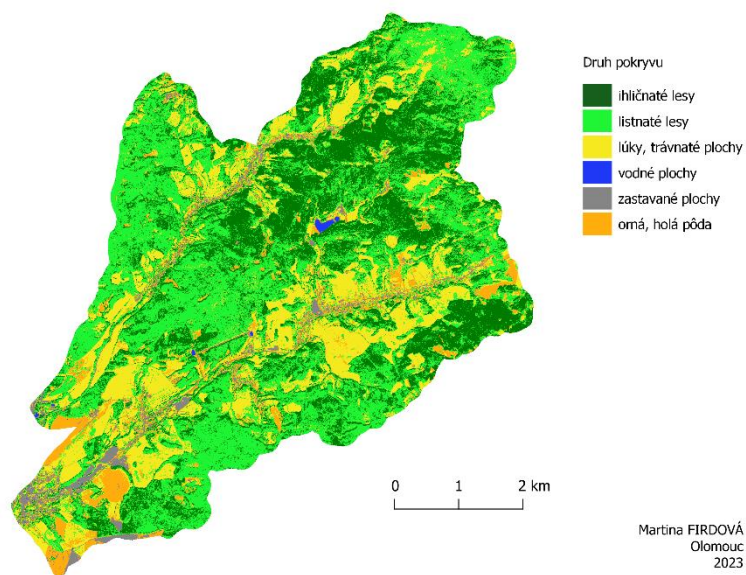
Obrázok 31 Riadená klasifikácia Random Forest na multispektrálnom snímku Pléiades so 100 stromami.



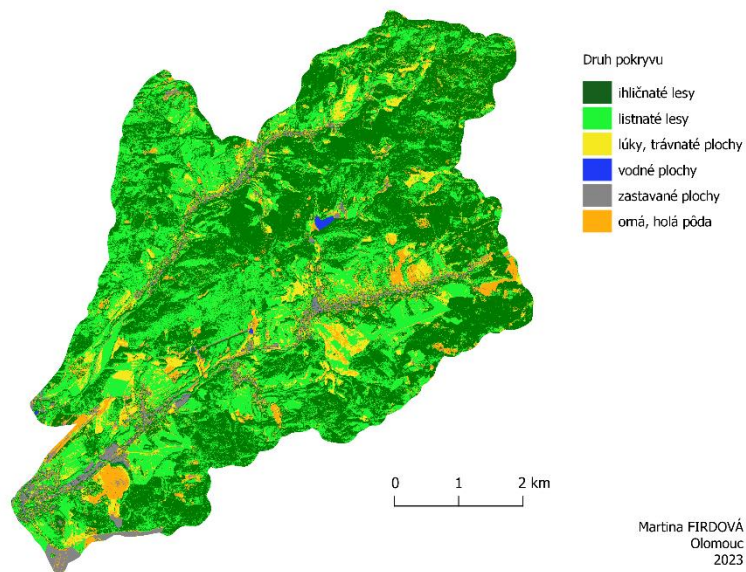
Obrázok 32 Riadená klasifikácia Random Forest na multispektrálnom snímku Pléiades s nastavením 250 stromov.



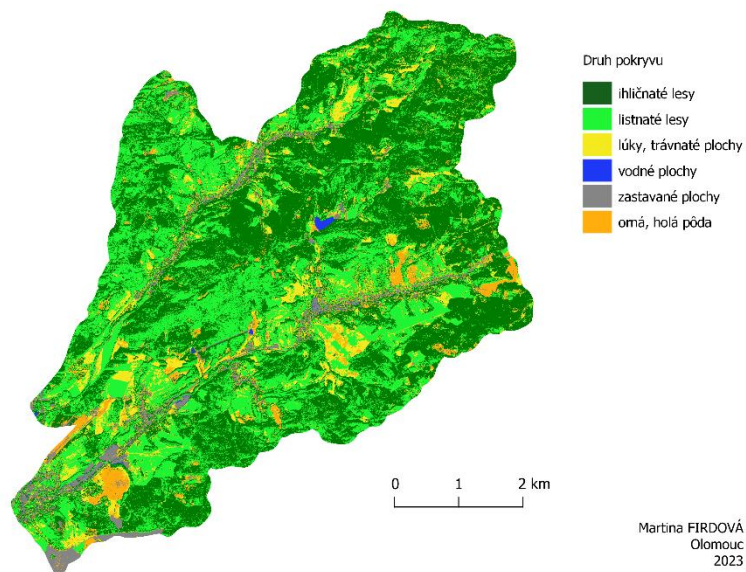
Obrázok 33 Riadená klasifikácia Random Forest na multispektrálnom snímku Pléiades.



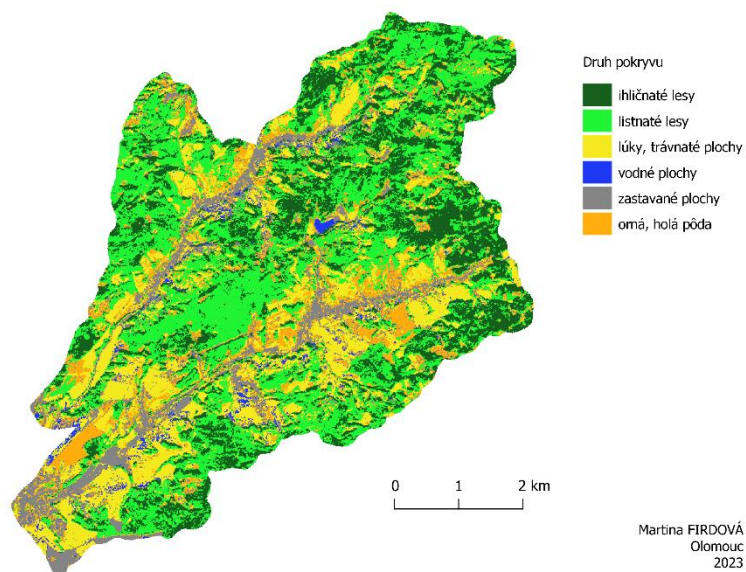
Obrázok 34 Riadená klasifikácia Random Forest na multispektrálnom snímku Pléiades s počtom stromov 250 a tréningových vzoriek 50 000.



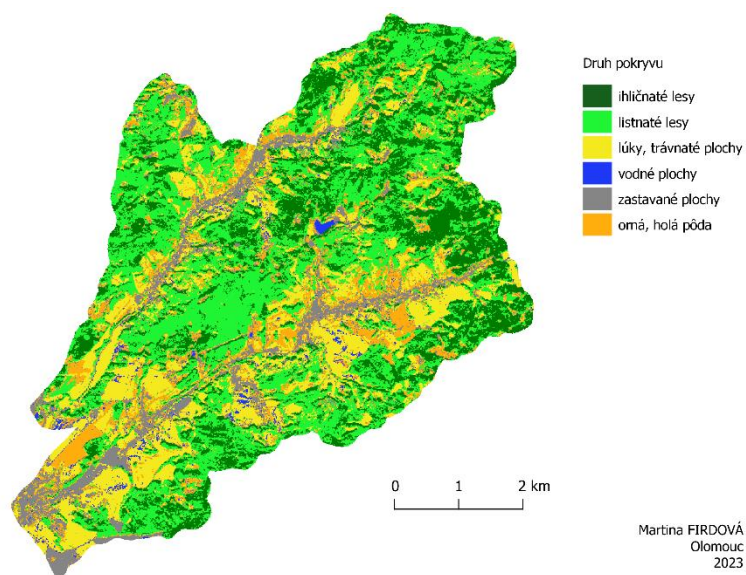
Obrázok 35 Riadená klasifikácia Random Forest na snímku Pléiades po pansharpeningu.



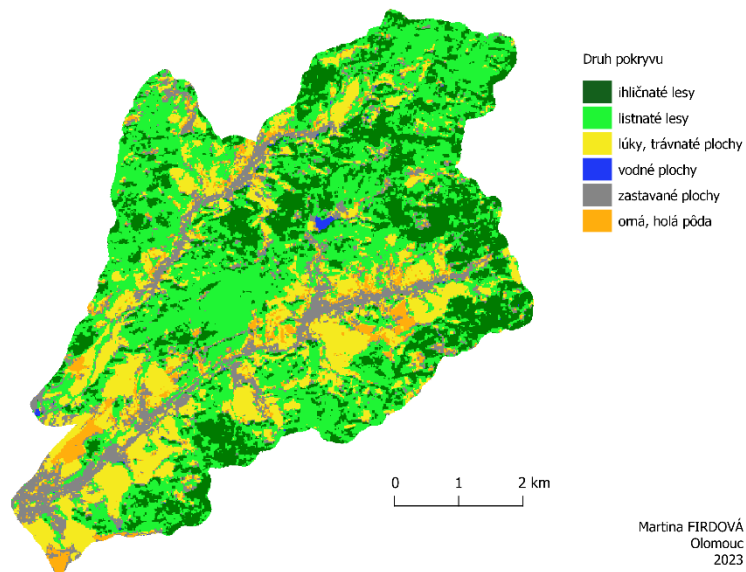
Obrázok 36 Riadená klasifikácia Random Forest na snímku Pléiades po pansharpeningu s počtom stromov 250 a indexom NDVI a SAVI.



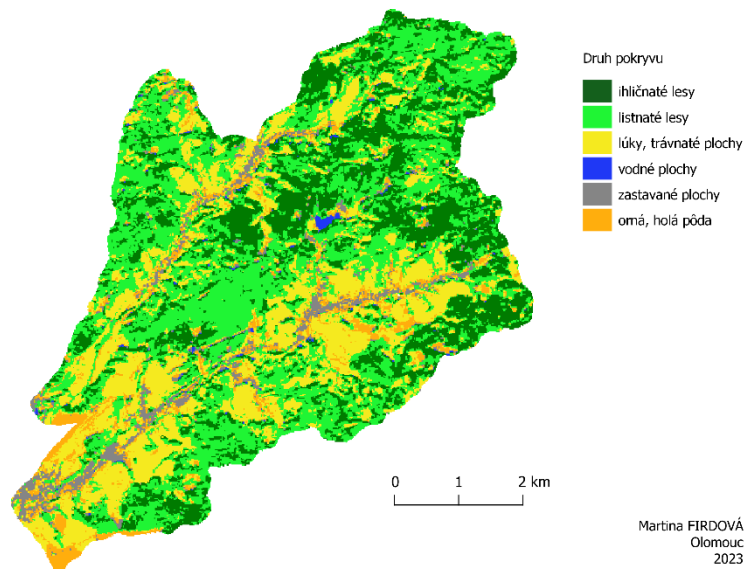
Obrázok 37 Riadená klasifikácia Random Forest na snímku Sentinel-2 s priestorovým rozlíšením 10 m.



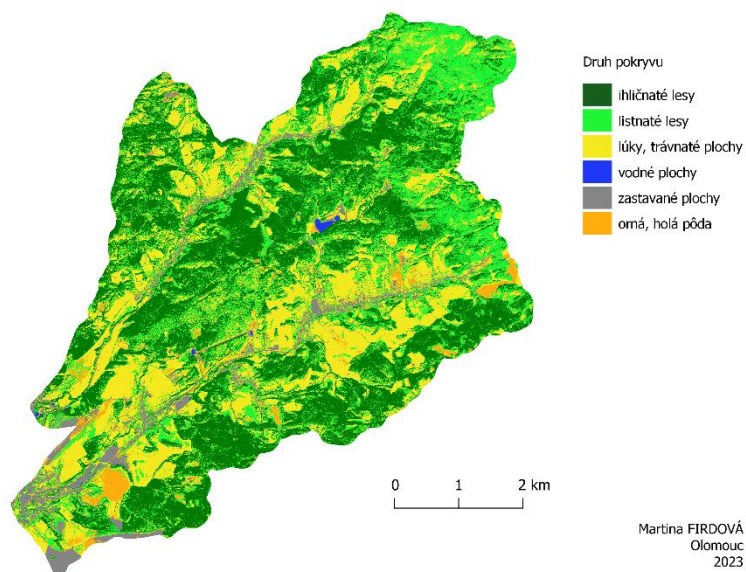
Obrázok 38 Riadená klasifikácia Random Forest na snímku Sentinel-2 s priestorovým rozlíšením 10 m s počtom stromov 250 a indexom NDVI, SAVI.



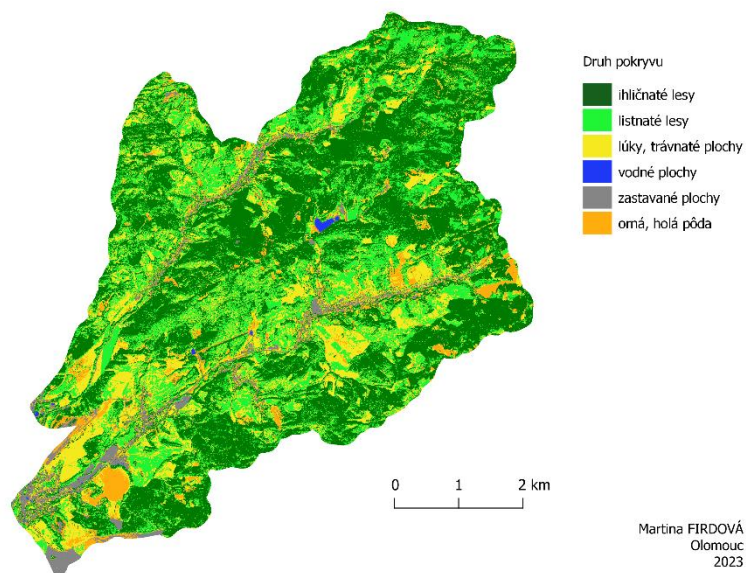
Obrázok 39 Riadená klasifikácia Random Forest na snímku Sentinel-2 s rozlíšením 20 m a počtom stromov 10.



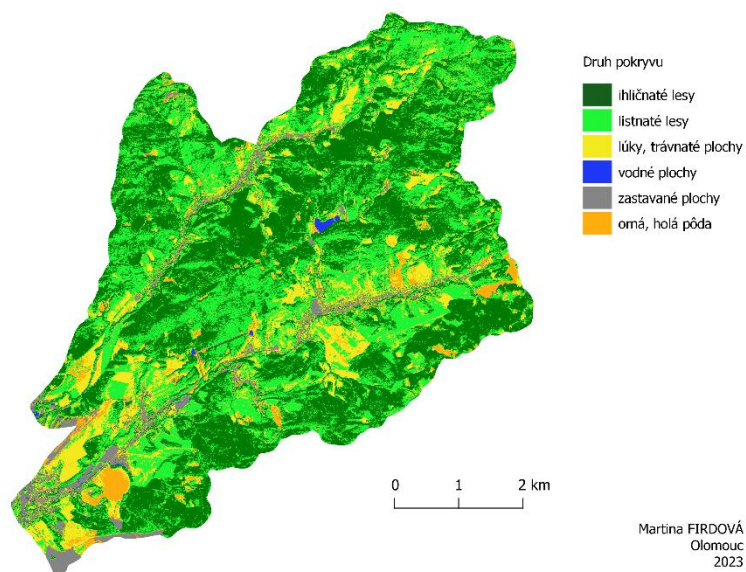
Obrázok 40 Riadená klasifikácia Random Forest na snímku Sentinel-2 s rozlíšením 20 m, s počtom stromov 250 a indexom NDVI, NDTI, SAVI, BSI



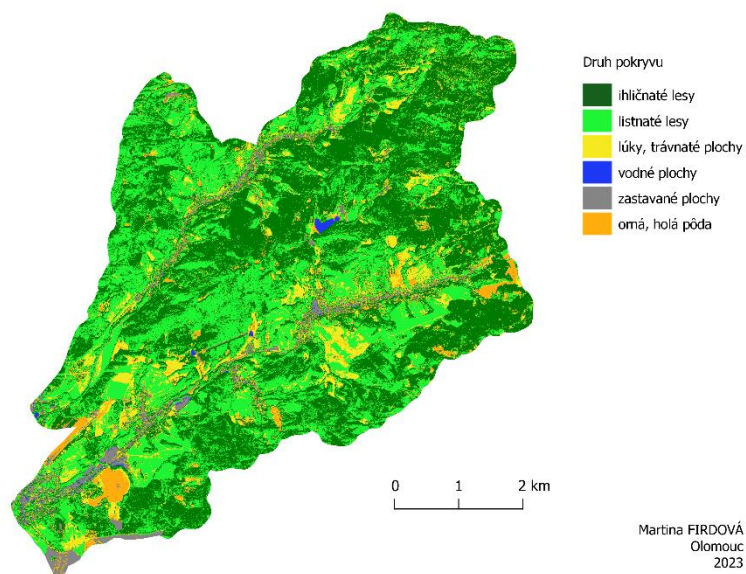
Obrázok 41 Riadená klasifikácia K-nearest neighbor na multispektrálnom snímku Pléiades



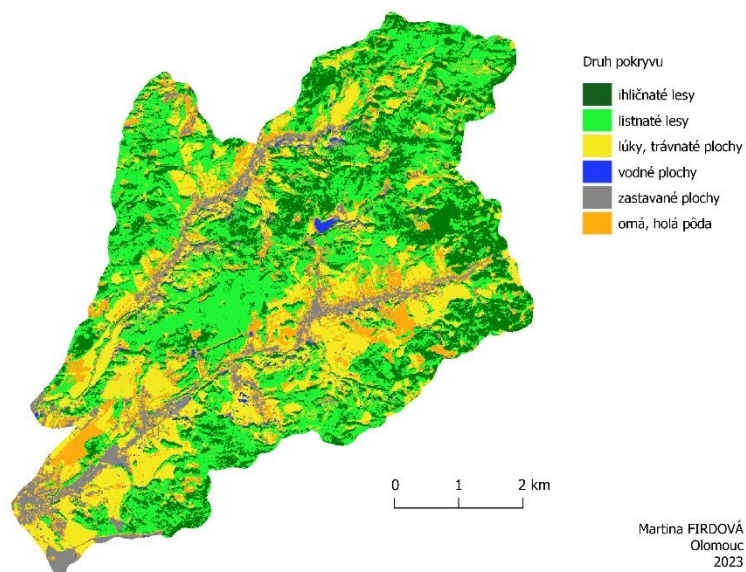
Obrázok 42 Riadená klasifikácia K-Nearest Neighbors na multispektrálnom snímku Pléiades s počtom susedov 20 a indexami NDVI a SAVI..



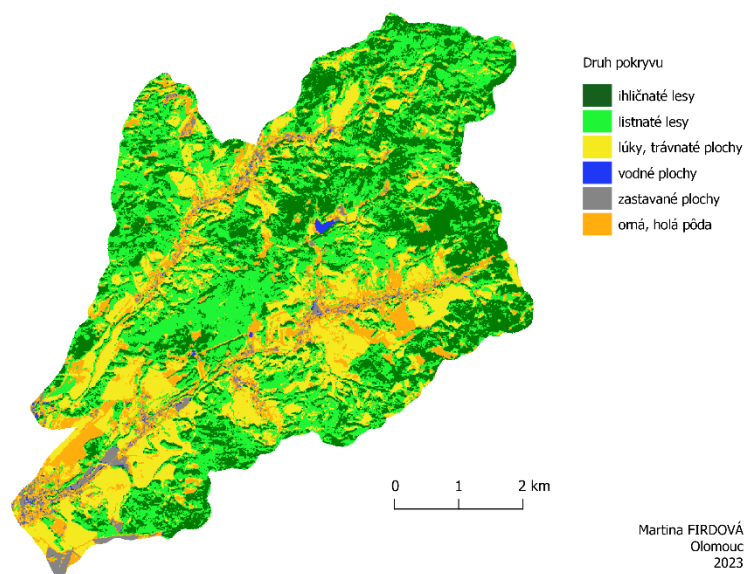
Obrázok 43 Riadená klasifikácia K-Nearest Neighbors na multispektrálnom snímku Pléiades s počtom susedov 20.



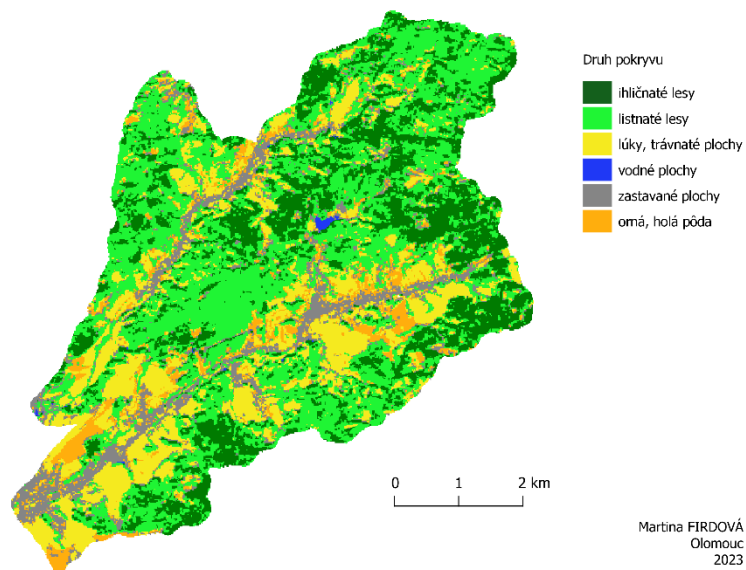
Obrázok 44 Riadená klasifikácia K-nearest neighbors na snímku Pléiades po pansharpeningu.



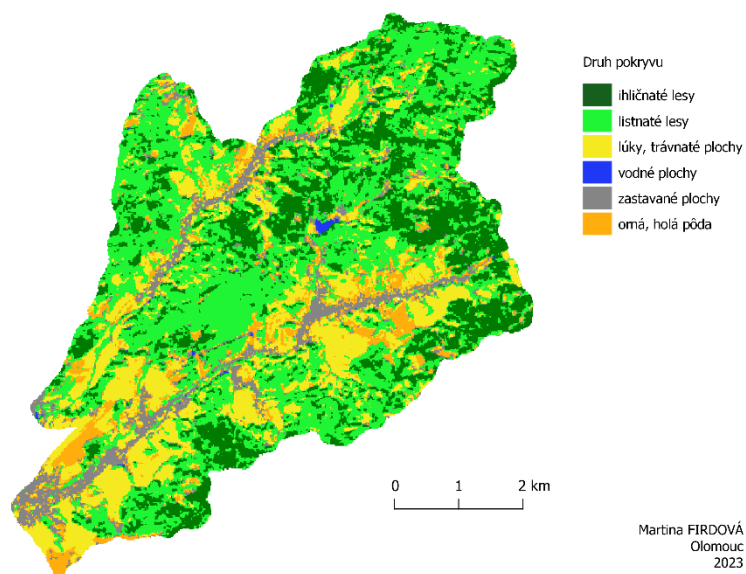
Obrázok 45 Riadená klasifikácia K-Nearest Neighbor na snímku Sentinel-2 s priestorovým rozlíšením 10 m



Obrázok 46 Riadená klasifikácia K-Nearest Neighbor na snímku Sentinel-2 s priestorovým rozlíšením 10 m a počtom susedov 250.



Obrázok 47 Riadená klasifikácia K-nearest neighbor na snímku Sentinel-2 s priestorovým rozlíšením 20 m



Obrázok 48 Riadená klasifikácia K-nearest neighbor na snímku Sentinel-2 s priestorovým rozlíšením 20 m s počtom susedov 20 a indexami NDVI, NDTI, BSI, SAVI

Priloha 4

Analysis - Accuracy assessment results

Thematic map: OUTPUT.tif

Sampling file: Pleiades_pansharp_rf_250.gpkg

Estimator: Simple/systematic estimator

Response design state: 204/204 samples labeled

1) Error matrix:

		Validation						Total	User accuracy	Total class area (km ²)	Wi
		1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)				
Thematic map classes	1	48	25	17	0	1	1	92	0.52174	18.48452	0.47374
	2	9	36	21	0	0	0	66	0.54545	13.3514	0.34218
	3	2	3	14	0	0	2	21	0.66667	4.24717	0.10885
	4	0	0	0	3	7	0	10	0.3	0.13367	0.00343
	5	0	1	0	0	5	1	7	0.71429	1.30959	0.03356
	6	0	1	2	0	1	4	8	0.5	1.49183	0.03823
Total		59	66	54	3	14	8	204		39.01818	
Producer accuracy		0.81356	0.54545	0.25926	1.0	0.35714	0.5				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.53922	0.03499

Analysis - Accuracy assessment results

Thematic map: OUTPUT.tif

Sampling file: Pleiades_multispektral_rd_50000_250.gpkg

Estimator: Simple/systematic estimator

Response design state: 202/202 samples labeled

1) Error matrix:

		Validation						Total	User accuracy	Total class area (m ²)	Wi
		1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)				
Thematic map classes	1	29	18	5	0	1	1	54	0.53704	11039731.03877	0.28292
	2	11	53	13	0	0	1	78	0.67949	15775159.41341	0.40428
	3	0	6	35	0	0	1	42	0.83333	8475858.00815	0.21722
	4	0	0	0	7	3	0	10	0.7	52251.9129	0.00134
	5	0	0	0	0	5	1	6	0.83333	1190640.52699	0.03051
	6	0	1	4	0	1	6	12	0.5	2486575.02119	0.06373
Total		40	78	57	7	10	10	202		39020215.92141	
Producer accuracy		0.725	0.67949	0.61404	1.0	0.5	0.6				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.66832	0.03321

Analysis - Accuracy assessment results
 Thematic map: OUTPUT.tif
 Sampling file: Pleiades_multispektral_ml.gpkg
 Estimator: Simple/systematic estimator
 Response design state: 200/202 samples labeled

1) Error matrix:

		Validation									
		1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	Total	User accuracy	Total class area (m ²)	Wi
Thematic map classes	1	48	31	21	0	1	1	102	0.47059	20812954.11294	0.53339
	2	4	22	19	0	1	0	46	0.47826	9445550.04472	0.24207
	3	0	2	14	0	0	0	16	0.875	3213405.3122	0.08235
	4	0	0	0	10	0	0	10	1.0	29698.22736	0.00076
	5	1	0	5	0	6	5	17	0.35294	3519137.36226	0.09019
	6	0	0	5	0	0	4	9	0.44444	1999470.86193	0.05124
Total		53	55	64	10	8	10	200		39020215.92141	
Producer accuracy		0.90566	0.4	0.21875	1.0	0.75	0.4				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.52	0.03542

Analysis - Accuracy assessment results
 Thematic map: OUTPUT.tif
 Sampling file: Pleiades_pansharp_ml.gpkg
 Estimator: Simple/systematic estimator
 Response design state: 396/396 samples labeled

1) Error matrix:

		Validation									
		1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	Total	User accuracy	Total class area (m ²)	Wi
Thematic map classes	1	77	79	44	1	3	7	211	0.36493	21319035.92715	0.54642
	2	2	63	50	0	0	0	115	0.54783	11620998.54655	0.29785
	3	0	1	23	0	0	1	25	0.92	2588240.45958	0.06634
	4	0	1	1	8	0	0	10	0.8	25402.60103	0.00065
	5	0	0	0	0	10	4	14	0.71429	1331228.14841	0.03412
	6	0	2	3	0	2	14	21	0.66667	2130800.1028	0.05461
Total		79	146	121	9	15	26	396		39015705.78551	
Producer accuracy		0.97468	0.43151	0.19008	0.88889	0.66667	0.53846				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.49242	0.02515

Pleiades_multispektral_ml_50000_ndvi_savi
Analysis - Accuracy assessment results
Thematic map: OUTPUT.tif
Sampling file: Pleiades_multispektral_ml_50000.gpkg
Estimator: Simple/systematic estimator
Response design state: 397/397 samples labeled

1) Error matrix:

		Validation						Total	User accuracy	Total class area (km ²)	Wi
		1 (Ihličnatý)	2 (Listnatý)	3 (Lúky)	4 (Vodné plochy)	5 (Zastavané oblasti)	6 (Orné pôdy)				
Thematic map classes	1	46	43	14	0	1	2	106	0.43396	10.60627	0.27184
	2	16	112	17	0	0	0	145	0.77241	14.62958	0.37496
	3	1	19	79	0	0	0	99	0.79798	9.95894	0.25525
	4	1	1	0	7	1	0	10	0.7	0.08356	0.00214
	5	0	0	0	0	8	4	12	0.66667	1.18746	0.03043
	6	0	1	6	0	4	14	25	0.56	2.55091	0.06538
Total		64	176	116	7	14	20	397		39.01671	
Producer accuracy		0.71875	0.63636	0.68103	1.0	0.57143	0.7				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.67003	0.02363

: Pleiades_pansharp_100tried
Analysis - Accuracy assessment results
Thematic map: OUTPUT.tif
Sampling file: Pleiades_pansharp_100tried.gpkg
Estimator: Simple/systematic estimator
Response design state: 398/398 samples labeled

1) Error matrix:

		Validation							Total	User accuracy	Total class area (km ²)	Wi
		1 (Ihličnatý)	2 (Listnatý)	3 (Lúky)	4 (Vodná plocha)	5 (Zastavané oblasti)	6 (Orné pôdy)	7 (Tieň)				
Thematic map classes	1	41	37	3	0	1	0	11	93	0.44086	9.36895	0.24012
	2	7	49	1	0	0	0	0	57	0.85965	5.76067	0.14764
	3	18	67	99	0	1	6	0	191	0.51832	19.22108	0.49262
	4	0	0	0	7	3	0	0	10	0.7	0.08543	0.00219
	5	0	0	0	0	9	2	0	11	0.81818	1.05418	0.02702
	6	3	0	3	0	4	10	0	20	0.5	1.9535	0.05007
	7	6	6	0	0	0	1	3	16	0.1875	1.57437	0.04035
Total		75	159	106	7	18	19	14	398		39.01818	
Producer accuracy		0.54667	0.30818	0.93396	1.0	0.5	0.52632	0.21429				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.54774	0.02498

Pleiades_multispektral_100tried
Analysis - Accuracy assessment results
Thematic map: OUTPUT.tif
Sampling file: Pleiades_multispektral_100tried.gpkg
Estimator: Simple/systematic estimator
Response design state: 397/397 samples labeled

1) Error matrix:

		Validation							Total	User accuracy	Total class area (km ²)	Wi
		2 (Ihličnatý)	4 (Listnatý)	6 (Lúky)	8 (Vodné plochy)	10 (Zastavané oblasti)	12 (Orná pôda)	14 (Tieň)				
Thematic map classes	2	58	53	3	0	0	0	2	116	0.5	11.49189	0.29451
	4	15	85	5	0	0	1	0	106	0.80189	10.55816	0.27058
	6	2	24	92	0	2	7	1	128	0.71875	12.72217	0.32604
	8	0	0	0	1	2	0	2	5	0.2	0.10326	0.00265
	10	0	0	0	0	6	0	0	6	1.0	0.56047	0.01436
	12	0	2	1	0	6	12	0	21	0.57143	2.08031	0.05331
Total		79	165	101	1	16	21	14	397		39.02022	
Producer accuracy		0.73418	0.51515	0.91089	1.0	0.375	0.57143	0.64286				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.66247	0.02376

Analysis - Accuracy assessment results
 Thematic map: OUTPUT.tif
 Sampling file: Sentinel_10m_100tried.gpkg
 Estimator: Simple/systematic.estimator
 Response design state: 397/397 samples labeled

1) Error matrix:

		Validation										
		2 (Ihličnaté)	4 (Listnaté)	6 (Lúky)	8 (Vodné plochy)	10 (Zastavané oblasti)	12 (Orná pôda)	14 (Tieň)	Total	User accuracy	Total class area (km ²)	Wi
Thematic map classes	2	71	41	2	0	0	0	4	118	0.60169	11.81086	0.3027
	4	11	83	5	0	0	0	0	99	0.83838	9.9877	0.25597
	6	0	11	101	0	1	4	0	117	0.86325	11.75455	0.30126
	8	1	0	1	2	0	0	6	10	0.2	0.15543	0.00398
	10	0	0	0	0	10	2	0	12	0.83333	1.21995	0.03127
	12	0	2	5	0	7	17	0	31	0.54839	3.12219	0.08002
	14	1	1	0	0	0	1	7	10	0.7	0.9678	0.0248
Total		84	138	114	2	18	24	17	397		39.01848	
Producer accuracy		0.84524	0.60145	0.88596	1.0	0.55556	0.70833	0.41176				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.733	0.02223

Sentinel_20m_100tried

Analysis - Accuracy assessment results
 Thematic map: OUTPUT.tif
 Sampling file: Sentinel_20m_100tried.gpkg
 Estimator: Simple/systematic.estimator
 Response design state: 394/394 samples labeled

1) Error matrix:

		Validation										
		2 (Ihličnaté)	4 (Listnaté)	6 (Lúky)	8 (Vodné plochy)	10 (Zastavané plochy)	12 (Orné pôdy)	14 (Tieň)	Total	User accuracy	Total class area (km ²)	Wi
Thematic map classes	2	75	34	4	0	0	1	2	116	0.64655	11.70316	0.29999
	4	6	82	9	0	1	0	0	98	0.83673	9.90413	0.25387
	6	0	8	81	0	3	4	0	96	0.84375	9.79415	0.25106
	8	0	0	0	10	0	0	0	10	1.0	0.04557	0.00117
	10	0	8	4	0	15	5	1	33	0.45455	3.39609	0.08705
	12	0	4	3	0	7	13	0	27	0.48148	2.7404	0.07025
	14	2	1	0	0	0	0	11	14	0.78571	1.42836	0.03661
Total		83	137	101	10	26	23	14	394		39.01186	
Producer accuracy		0.90361	0.59854	0.80198	1.0	0.57692	0.56522	0.78571				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.72843	0.02244

Sentinel_20m_100tried_ndvi

Analysis - Accuracy assessment results
 Thematic map: OUTPUT.tif
 Sampling file: Sentinel_20m_100tried_ndvi.gpkg
 Estimator: Simple/systematic.estimator
 Response design state: 396/396 samples labeled

1) Error matrix:

		Validation										
		2 (Ihličnaté)	4 (Listnaté)	6 (Lúka)	8 (Vodné plochy)	10 (Zastavané plochy)	12 (Orná pôda)	14 (Tieň)	Total	User accuracy	Total class area (km ²)	Wi
Thematic map classes	2	86	29	1	0	0	0	3	119	0.72269	12.00855	0.30782
	4	4	95	10	0	1	0	2	112	0.84821	11.28714	0.28933
	6	1	22	74	0	5	9	0	111	0.66667	11.25099	0.2884
	8	0	0	0	10	0	0	0	10	1.0	0.04338	0.00111
	10	0	2	0	0	12	0	0	14	0.85714	1.41785	0.03634
	12	0	2	3	0	1	14	0	20	0.7	2.03191	0.05208
	14	1	1	0	0	0	1	7	10	0.7	0.97203	0.02492
Total		92	151	88	10	19	24	12	396		39.01186	
Producer accuracy		0.93478	0.62914	0.84091	1.0	0.63158	0.58333	0.58333				

2) Accuracy:

Overall:

Overall Accuracy	Standard deviation
0.75253	0.02171

Príloha 6

Tabuľka 12 Porovnanie rozloženia bodov DKV a najlepšie vyhodnotených klasifikácií a porovnanie percentuálneho zastúpenia

		DKV	Neriadená klasifikácia	Riadená klasifikácia			DKV	Neriadená klasifikácia	Riadená klasifikácia		
				ML	RF	KNN			ML	RF	KNN
Sentinel 10m	1 : 10 000	97670	92860	65477	71368	92793	30,85%	30,07%	20,71%	22,57%	30,08%
	1 : 20 000	24388	23303	16381	17912	23147	30,83%	30,20%	20,73%	22,67%	30,02%
	1 : 50 000	3910	3692	2626	2872	3743	30,85%	29,83%	20,77%	22,72%	30,27%
Ihličnaté lesy	1 : 10 000	82412	75930	103756	105235	75942	26,03%	24,59%	32,82%	33,29%	24,61%
	1 : 20 000	20559	18910	25847	26207	18983	25,99%	24,51%	32,72%	33,17%	24,62%
	1 : 50 000	3310	3066	4106	4127	2996	26,11%	24,77%	32,48%	32,65%	24,23%
Listnaté lesy	1 : 10 000	107823	100965	63476	64732	100802	34,05%	32,69%	20,08%	20,48%	32,67%
	1 : 20 000	26946	25211	15937	16184	25108	34,06%	32,67%	20,17%	20,49%	32,57%
	1 : 50 000	4310	4091	2536	2578	4020	34,00%	33,05%	20,06%	20,39%	32,51%
Lúky, trávniky	1 : 10 000	2391	1334	2061	4602	1289	0,76%	0,43%	0,65%	1,46%	0,42%
	1 : 20 000	605	333	535	1114	330	0,76%	0,43%	0,68%	1,41%	0,43%
	1 : 50 000	96	60	74	177	52	0,76%	0,48%	0,59%	1,40%	0,42%
Vodné plochy	1 : 10 000	21791	10851	45954	32436	10819	6,88%	3,51%	14,54%	10,26%	3,51%
	1 : 20 000	5474	2680	11451	8168	2750	6,92%	3,47%	14,49%	10,34%	3,57%
	1 : 50 000	865	441	1845	1330	433	6,82%	3,56%	14,60%	10,52%	3,50%
Zastavané plochy	1 : 10 000	4541	26890	35420	37771	26881	1,43%	8,71%	11,20%	11,95%	8,71%
	1 : 20 000	1132	6727	8851	9417	6777	1,43%	8,72%	11,20%	11,92%	8,79%
	1 : 50 000	185	1027	1454	1557	1121	1,46%	8,30%	11,50%	12,32%	9,07%
Poľnohospodárske, holé plochy	1 : 10 000	97670	91907	67202	67472	75545	30,85%	30,12%	21,46%	21,37%	33,79%
	1 : 20 000	24388	23014	16772	16873	22971	30,83%	30,19%	21,45%	21,39%	30,19%
	1 : 50 000	3910	3654	2693	2716	3641	30,85%	29,85%	21,50%	21,51%	29,90%
Ihličnaté lesy	1 : 10 000	82412	75734	109708	124464	85370	26,03%	24,82%	35,04%	39,41%	38,19%
	1 : 20 000	20559	18798	27378	31023	18826	25,99%	24,66%	35,01%	39,33%	24,75%
	1 : 50 000	3310	3064	4392	4972	3051	26,11%	25,03%	35,07%	39,38%	25,05%
Listnaté lesy	1 : 10 000	107823	85382	92022	67376	458	34,05%	27,99%	29,39%	21,34%	0,20%
	1 : 20 000	26946	21431	23056	16848	21441	34,06%	28,11%	29,48%	21,36%	28,18%
	1 : 50 000	4310	3449	3684	2660	3428	34,00%	28,18%	29,42%	21,07%	28,15%
Lúky, trávniky	1 : 10 000	2391	467	430	679	27981	0,76%	0,15%	0,14%	0,22%	12,52%
	1 : 20 000	605	116	107	170	115	0,76%	0,15%	0,14%	0,22%	0,15%
	1 : 50 000	96	20	17	27	20	0,76%	0,16%	0,14%	0,21%	0,16%
Vodné plochy	1 : 10 000	21791	28206	28783	38330	23263	6,88%	9,25%	9,19%	12,14%	10,41%
	1 : 20 000	5474	6994	7185	9604	6940	6,92%	9,17%	9,19%	12,17%	9,12%
	1 : 50 000	865	1124	1166	1547	1093	6,82%	9,18%	9,31%	12,25%	8,97%
Zastavané plochy	1 : 10 000	4541	23394	14950	17464	10937	1,43%	7,67%	4,77%	5,53%	4,89%
	1 : 20 000	1132	5877	3704	4367	5787	1,43%	7,71%	4,74%	5,54%	7,61%
	1 : 50 000	185	930	571	704	946	1,46%	7,60%	4,56%	5,58%	7,77%
Poľnohospodárske, holé plochy	1 : 10 000	97670	90798	82795	87058	110523	30,85%	29,73%	26,17%	27,51%	34,93%
	1 : 20 000	24388	22630	20578	21808	27603	30,83%	29,67%	26,02%	27,58%	34,90%
	1 : 50 000	3910	3630	3356	3534	4456	30,85%	29,71%	26,53%	27,94%	35,23%
Ihličnaté lesy	1 : 10 000	82412	78121	112311	120387	83909	26,03%	25,58%	35,49%	38,04%	26,52%
	1 : 20 000	20559	19455	28225	30015	21027	25,99%	25,51%	35,70%	37,95%	26,59%
	1 : 50 000	3310	3176	4429	4733	3345	26,11%	25,99%	35,02%	37,42%	26,44%
Listnaté lesy	1 : 10 000	107823	111639	87533	75741	92526	34,05%	36,55%	27,66%	23,93%	29,24%
	1 : 20 000	26946	28028	21796	18944	23085	34,06%	36,75%	27,57%	23,95%	29,19%
	1 : 50 000	4310	4455	3523	3058	3685	34,00%	36,46%	27,85%	24,17%	29,13%
Lúky, trávniky	1 : 10 000	2391	995	756	505	566	0,76%	0,33%	0,24%	0,16%	0,18%
	1 : 20 000	605	240	181	128	139	0,76%	0,31%	0,23%	0,16%	0,18%
	1 : 50 000	96	40	26	22	23	0,76%	0,33%	0,21%	0,17%	0,18%
Vodné plochy	1 : 10 000	21791	5155	10649	10728	16549	6,88%	1,69%	3,37%	3,39%	5,23%
	1 : 20 000	5474	1282	2667	2658	4076	6,92%	1,68%	3,37%	3,36%	5,15%
	1 : 50 000	865	186	438	439	663	6,82%	1,52%	3,46%	3,47%	5,24%
Zastavané plochy	1 : 10 000	4541	18706	22375	22032	12378	1,43%	6,12%	7,07%	6,96%	3,91%
	1 : 20 000	1132	4629	5624	5529	3152	1,43%	6,07%	7,11%	6,99%	3,99%
	1 : 50 000	185	732	876	864	478	1,46%	5,99%	6,93%	6,83%	3,78%
Poľnohospodárske, holé plochy	1 : 10 000	97670	73308	169555	146143	73649	30,85%	24,06%	53,58%	46,18%	24,19%
	1 : 20 000	24388	18290	42405	36578	18401	30,83%	24,02%	53,63%	46,25%	24,18%
	1 : 50 000	3910	2981	6818	5858	2968	30,85%	24,38%	53,90%	46,30%	24,37%
Ihličnaté lesy	1 : 10 000	82412	43282	93599	107338	43221	26,03%	14,20%	29,58%	33,92%	14,19%
	1 : 20 000	20559	10905	23342	26852	10821	25,99%	14,32%	29,52%	33,96%	14,22%
	1 : 50 000	3310	1721	3694	4273	1739	26,11%	14,07%	29,20%	33,77%	14,28%
Listnaté lesy	1 : 10 000	107823	160452	22363	36743	160095	34,05%	52,65%	7,07%	11,61%	52,58%
	1 : 20 000	26946	40072	5593	9074	39990	34,06%	52,63%	7,07%	11,47%	52,56%
	1 : 50 000	4310	6475	903	1445	6375	34,00%	52,95%	7,14%	11,42%	52,35%
Lúky, trávniky	1 : 10 000	2391	841	256	1295	856	0,76%	0,28%	0,08%	0,41%	0,28%
	1 : 20 000	605	209	65	339	219	0,76%	0,27%	0,08%	0,43%	0,29%
	1 : 50 000	96	34	11	57	34	0,76%	0,28%	0,09%	0,45%	0,28%
Vodné plochy	1 : 10 000	21791	9613	12164	11896	9542	6,88%	3,15%	3,84%	3,76%	3,13%
	1 : 20 000	5474	2383	3002	2928	2390	6,92%	3,13%	3,80%	3,70%	3,14%
	1 : 50 000	865	362	498	484	382	6,82%	2,96%	3,94%	3,83%	3,14%
Zastavané plochy	1 : 10 000	4541	17229	18514	13055	17138	1,43%	5,65%	5,85%	4,13%	5,63%
	1 : 20 000	1132	4284	4668	3309	4264	1,43%	5,63%	5,90%	4,18%	5,60%
	1 : 50 000	185	655	726	535	680	1,46%	5,36%	5,74%	4,23%	5,58%
Poľnohospodárske, holé plochy	1 : 10 000	97670	73308	169555	146143	73649	30,85%	24,06%	53,58%	46,18%	24,19%
	1 : 20 000	24388	18290	42405	36578	18401	30,83%	24,02%	53,63%	46,25%	24,18%
	1 : 50 000	3910	2981	6818	5858	2968	30,85%	24,38%	53,90%	46,30%	24,37%