

Česká zemědělská univerzita v Praze

Fakulta životního prostředí

Katedra prostorových věd



**Možnosti detekce mokřadních biotopů Natura 2000
pomocí dat Sentinel-2**

Bakalářská práce

Vedoucí práce: doc. Ing. Petra Šímová, Ph.D.

Bakalant: Anna Brandejsová

© 2021 ČZU v Praze



**Česká zemědělská univerzita v Praze
Fakulta životního prostředí**

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Autorka práce:	Anna Brandejsová
Studijní program:	Geografické informační systémy a dálkový průzkum Země v životním prostředí
Vedoucí práce:	doc. Ing. Petra Šimová, Ph.D.
Garantující pracoviště:	Katedra prostorových věd
Jazyk práce:	Čeština
Název práce:	Možnosti detekce mokřadních biotopů Natura 2000 pomocí dat Sentinel-2
Název anglicky:	Possibilities of detection of Natura 2000 wetland habitats using Sentinel-2 data
Cíle práce:	Česká republika disponuje velmi prostorově i tematicky podrobným mapováním přírodě blízkých biotopů, pořízeným zejména pro účely naplňování směrnice 92/43/EEC, o ochraně přírodních stanovišť, volně žijících živočichů a planě rostoucích rostlin. Tato Vrstva mapování biotopů (VMB) je od dob svého vzniku průběžně aktualizována terénním průzkumem, přičemž aktualizace celého území ČR trvá 12 let. Nabízí se tak využití volně dostupných družicových dat pro pomoc s mapováním, na druhou stranu ale přesnost klasifikace do podrobných tříd odpovídajících VMB může být problematická. Rámcovým cílem bakalářské práce je zjistit, zda a do jaké míry lze s využitím dat družic Sentinel-2 detekovat vybrané typy mokřadních biotopů v modelovém území. Formulace dílčích cílů je úkolem řešitelky.
Metodika:	Rámcové body postupu jsou následující: (1) Výběr modelového území a explorativní analýza kategorií mokřadních biotopů VMB v něm. (2) Příprava trénovacích a validačních dat na základě VMB. (3) Řízená klasifikace pomocí alespoň dvou machine learning algoritmů, popř. doplňkově neřízená klasifikace. (4) Vyhodnocení přesnosti klasifikace (chybová matice, porovnání s VMB). (5) Interpretace a diskuse výsledků. K analýzám a prezentaci výsledků řešitelka dle vlastního uvážení využije programy ArcGIS, R, SNAP a prostředí Google Earth Engine. Postupy zautomatizuje v podobě skriptů a/nebo work-flow modelů (např. Model Builder).

Doporučený rozsah práce: 30 - 60 stran

Klíčová slova: dálkový průzkum Země, soustava Natura 2000, Sentinel-2, Vrstva mapování biotopů, mokřadní biotopy

Doporučené zdroje informací:

1. Bhatnagar, S., Gill, L., Regan, S., Naughton, O., Johnston, P., Waldren, S., & Ghosh, B. (2020). Mapping vegetation communities inside wetlands using Sentinel-2 imagery in Ireland. *International Journal of Applied Earth Observation and Geoinformation*, 88, 102083.
2. Corbane, C., Lang, S., Pipkins, K., Alleaume, S., Deshayes, M., García Millán, V. E., Strasser, T., vanden Borre, J., Toon, S., & Michael, F. (2015). Remote sensing for mapping natural habitats and their conservation status - New opportunities and challenges. *International Journal of Applied Earth Observation and Geoinformation*, 37.
3. Chytrý, M., Kučera, T., Kočí, M., Grulich, V., & Lustyk, P. (2010). Katalog biotopů České republiky. Druhé vydání.
4. Nomura, K. & Mitchard, E.T.A. 2018. More than meets the eye: using Sentinel-2 to map small plantations in complex forest landscapes. *Remote Sensing*, doi:10.3390/rs10111693.
5. Schmidt, J., Fassnacht, F. E., Neff, C., Lausch, A., Kleinschmit, B., Förster, M., & Schmidlein, S. (2017). Adapting a Natura 2000 field guideline for a remote sensing-based assessment of heathland conservation status. *International Journal of Applied Earth Observation and Geoinformation*, 60, 61–71.

Předběžný termín obhajoby: 2020/21 LS – FŽP

Elektronicky schváleno: 23. 3. 2021
doc. Ing. Petra Šimová, Ph.D.
Vedoucí katedry

Elektronicky schváleno: 24. 3. 2021
prof. RNDr. Vladimír Bejček, CSc.
Děkan

V Praze dne 30. 3. 2021

Čestné prohlášení

Prohlašuji, že jsem bakalářskou práci na téma "Možnosti detekce mokřadních biotopů Natura 2000 pomocí dat Sentinel-2" vypracovala samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce.

Jsem si vědoma, že na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů, ve znění pozdějších předpisů, především ustanovení § 35 odst. 3 tohoto zákona, tj. o užití tohoto díla.

Jsem si vědoma, že odevzdáním bakalářské práce souhlasím s jejím zveřejněním podle zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů, ve znění pozdějších předpisů, a to i bez ohledu na výsledek její obhajoby. Svým podpisem rovněž prohlašuji, že elektronická verze práce je totožná s verzí tištěnou a že s údaji uvedenými v práci bylo nakládáno v souvislosti s GDPR.

V Praze dne 31.3.2021

(podpis autora práce)

Poděkování

Ráda bych touto cestou poděkovala vedoucí práce, doc. Ing. Petře Šimové Ph.D., za vedení této práce, trpělivost, a za pravidelné a přínosné konzultace i v této distanční době.

Bakalářská práce byla řešena v rámci projektu TAČR SS01010046 – Možnosti aktualizace mapování biotopů NATURA 2000 pokročilými metodami dálkového průzkumu Země (SentiMAP).

Abstrakt

Tato bakalářská práce je zaměřena na mapování evropsky chráněných biotopů, které spadají pod soustavu Natura 2000. Předmětem výzkumu jsou mokřadní biotopy, které byly detekovány v rámci mapování biotopů a nachází se v Karlovarském kraji. Mapování biotopů pro celé území ČR v terénu trvá 12 let, proto tato bakalářská práce zjišťuje možnosti přispění dálkového průzkumu Země k urychlení vytváření Vrstvy mapování biotopů. K detekci mokřadů byly využity multispektrální snímky z družice Sentinel-2.

Pro výzkum byly zpracovány snímky z vegetačního období pro mokřadní biotopy z roku 2018. Výsledné mapování bylo provedeno za využití řízené klasifikace, konkrétně algoritmů Random forest a Classification a regression trees v programu Google Earth Engine. Ke zhodnocení přesnosti výsledků byly použity charakteristiky přesnosti, které vycházejí z výskytu biotopů dle Vrstvy mapování biotopů a z předpokládaného výskytu biotopů z klasifikace.

Z analýzy jsem ověřila předpoklad, že mapování mokřadů pomocí multispektrálních snímků je náročný, především z důvodu heterogenity mokřadních biotopů. Z klasifikace vychází, že biotopy byly často nadhodnocovány. Z klasifikací lze indikovat přibližná místa výskytu biotopů, také se podařilo odlišit mokřadní plochy od okolí. Klasifikátor Random Forest dosahoval spolehlivějších výsledků pro všechny kategorie. Zároveň mezi detekovanými biotopy nedocházelo k výrazné záměně. Výsledky této práce nenahradí mapování biotopů v terénu, ale tento výzkum by do budoucna mohl pro mapování biotopů sloužit jako podklad k určení přibližného výskytu biotopů.

Klíčová slova: dálkový průzkum Země, soustava Natura 2000, Sentinel-2, Vrstva mapování biotopů, mokřadní biotopy

Abstract

This bachelor thesis is focused on detecting habitats of the Site of Community importance which are part of the network Natura 2000. The objective of the mapping are wetlands occurred in Karlovy Vary Region. Originally habitat mapping in Czech Republic takes up to 12 years. Therefore, this bachelor thesis examines the possibilities of the contribution of remote sensing to accelerate the creation of the Habitat mapping layer. Multispectral images from the Sentinel-2 satellite were used for the detection of wetlands habitats.

For the research there were images processed from vegetation season of wetlands of the year 2018. The mapping was performed using a supervised classification, specifically the algorithms Random forest and Classification and regression trees in the Google Earth Engine. To evaluate the accuracy of the results, there were used accuracy metrics, which were based on the occurrence of habitats according to Habitat mapping layer in addition to the expected occurrence of habitats from the results of classifications.

The analysis verified an assumption that mapping wetlands using multispectral images is challenging, the main reason is heterogeneous surface of wetlands. The result of classification proves that habitats have been overestimated. From the classification it is possible to indicate approximate location of habitats also the result differs wetlands from the surroundings. The classification by Random forest achieved more reliable results for all categories. The outcomes of this thesis cannot substitute terrain mapping however this research could serve as a basis for habitats mapping in the future for determination approximate occurrence of habitats.

Keywords: remote sensing, Natura 2000, Habitat mapping layer, Sentinel-2, wetland habitats

Obsah

1. Úvod.....	12
2. Cíle práce	13
3. Literární rešerše	14
3.1 Chráněné biotopy soustavy Natura 2000	14
3.2 Multispektrální snímky	15
3.3 Klasifikace biotopů	19
3.3.1 Klasifikátory	21
4. Metodika	23
4.1 Výběr zájmového území a vhodných biotopů.....	23
4.2 Vstupní data	30
4.3 Zpracování klasifikace	32
4.4 Zhodnocení přesnosti	35
5. Výsledky	37
5.1 Random Forest	37
5.2 Classification and Regression Trees	39
5.3 Porovnání výsledků RF a CART.....	42
6. Diskuse	44
7. Závěr.....	48
8. Seznam použitých zdrojů	49
Přílohy.....	53

Seznam obrázků

- Obrázek č. 1: Sentinel-2, pásma v jednotlivých rozlišeních. Heslovitě popsany účel pro jednotlivá pásma B1 (445 nm), B2 (490 nm), B3 (560 nm), B4 (665 nm), B5 (705 nm), B6 (740 nm), B7 (783 nm), B8 (842 nm), B8A (865nm), B9 (940 nm), B10 (1375 nm), B11 (1610 nm) a B12 (219, nm) B1 (445 nm), B9 (940 nm), B10 (1375 nm). 18
- Obrázek č. 2: Ukázka chybové matice, vzorec pro výpočet celkové přesnosti (N) a přesnosti uživatele (Total na pravé straně) a zpracovatele (Total v dolním řádku matice). 21
- Obrázek č. 3: Rozložení výskytu vybraných biotopů na zájmovém území, Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, přechodová rašeliniště R2.3, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6. 23
- Obrázek č. 4: Pokrytí zájmové lokality snímkem S2. 24
- Obrázek č. 5: Lesní prameniště bez tvorby pěnovců, Paukertová, 2009,
Zdroj: <http://www.paukertova.cz/view.php?nazevclanku=biotopy-r1-4-lesni-prameniste-bez-tvorby-penovcu&cislocclanku=2009020001> 25
- Obrázek č. 6: Nevápnitá mechová slatiniště, AOPK ČR, 2012.
Zdroj: <https://www.ochranaprirody.cz/lokality/?idlokality=13958> 26
- Obrázek č. 7: Přechodová rašeliniště, AOPK ČR, 2021.
Zdroj: <https://www.ochranaprirody.cz/lokality/?idlokality=1045> 26
- Obrázek č. 8: Rákosiny eutrofních stojatých vod (M1.1), Novák, 2020.
Zdroj: <https://www.zahrada-cs.com/foto/cz/103075/> 27
- Obrázek č. 9: Vegetace vysokých ostřic (M1.7), Duchoslav, 2020,
Zdroj: <http://flora.upol.cz/fotogalerie/info/9206--Magno-Caricion-elatae/8.html>. 27
- Obrázek č. 10: Vlhké pcháčové louky (T1.5), Duchoslav, 2020.
Zdroj: <http://flora.upol.cz/vegetace/info/9237-Calthion-palustris.html> 28
- Obrázek č. 11: Vlhká tužebníková lada, Otýpková, 2005.
Zdroj: <http://www.botanickafotogalerie.cz/fotogalerie.php?lng=cz&latName=Lysimachio%20vulgaris-Filipenduletum%20ulmariae&title=Lysimachio%20vulgaris-Filipenduletum%20ulmariae%20|%20vlhk%C3%A1%20tu%C5%BEbn%C3%ADkov%C3%A> 28
- Obrázek č. 12: Mokřadní vrbiny (K1), Paukertová, 2006.
Zdroj: Paukertová, 2006, <http://www.paukertova.cz/view.php?cislocclanku=2009020010> .. 29

Obrázek 13: Údolní jasanovo-olšové luhy (2.2), Duchoslav, 2020. Zdroj: http://dendrologie.upol.cz/spolecenstva/info/19-Udolni-jasanovo-olsove-luhy.html	29
Obrázek č. 14: Náhodné body v zájmovém území vytvořené pro okolí a lokality biotopů, podklad je ortofoto ČR.....	30
Obrázek č. 15: Odlišnost spektrálních vlastností vybraných biotopů v květnu a červenci ve všech vhodných pásmech pro mapování.....	32
Obrázek č. 16: Index NDVI vytvořený v GEE, hodnoty nabývají -1 až 1, čím blíže je hodnota 1, tím více vegetace se v místě nachází, tento index byl vstupním pásmem pro klasifikace.	34
Obrázek č. 17: Výsledek klasifikace území za využití algoritmu RF, v legendě jsou všechny zobrazené biotopy včetně okolního povrchu, konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostržic M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.	37
Obrázek č. 18: Výřez území, demonstrace zobrazení biotopů klasifikovaných algoritmem RF v porovnání s reálným výskytem biotopů z VMB. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostržic M1.7, Lesní prameniště bez tvorby pěnovcůR1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.	38
Obrázek č. 19: Výsledek klasifikace území za využití algoritmu CART, v legendě jsou všechny zobrazené biotopy včetně okolního povrchu. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostržic M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.	40
Obrázek č. 17: Výřez území, kde dochází k velkému nadhodnocování biotopů klasifikátorem CART v porovnání s výskytem biotopů z VMB. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostržic M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.	41

Seznam tabulek

Tabulka č. 1: Zastoupení vybraných biotopů v porovnání s rozlohou Karlovarského kraje.	25
Tabulka č. 2: Tabulka zastoupení pro biotopy, které jsou spektrálně odlišné a byly použity do klasifikace.	32
Tabulka č. 3: Chybová matice a výpočty přesností z polygonů pro klasifikátor RF, v polygonech. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostržic M1.7, Lesní prameniště	

<i>bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.</i>	38
<i>Tabulka č. 4: Chybová matice a výpočet přesností z bodů pro klasifikátor RF. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostríc M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.</i>	38
<i>Tabulka č. 5: Chybová matice a výpočet přesností z polygonů pro algoritmus CART. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostríc M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.</i>	41
<i>Tabulka č. 6: Chybová matice a výpočet přesností z bodů pro klasifikátor CART. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostríc M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.</i>	42
<i>Tabulka č. 7: Tabulka přesností pouze pro třídy: Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostríc M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, tedy pouze kategorie mokřady (M) a rašeliniště (R) za použití RF.</i>	42
<i>Tabulka č. 8: Tabulka přesností pouze pro třídy: Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostríc M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, tedy pouze kategorie mokřady (M) a rašeliniště (R) za použití CART.</i>	43
<i>Tabulka 9: Tabulka přesností pro celé kategorie z klasifikace RF pro celé polygony, lesy (L), Mokřady a pobřežní vegetace (M), R Prameniště a rašeliniště (R), Sekundární trávníky a vřesoviště (T) a O jsou ostatní plochy</i>	43
<i>Tabulka 10: Tabulka přesností pro celé kategorie z klasifikace CART pro celé polygony, lesy (L), Mokřady a pobřežní vegetace (M), R Prameniště a rašeliniště (R), Sekundární trávníky a vřesoviště (T) a O jsou ostatní plochy</i>	43

1. Úvod

Natura 2000 je soustava chráněných území vytvářená státy Evropské unie. Lokality ochrany byly určeny směrnicí o ochraně volně žijících ptáků (147/2009/ES) a směrnicí o ochraně přírodních stanovišť, volně žijících živočichů a planě rostoucích rostlin (92/43/EHS). Evropsky Významné lokality (dále jako EVL), vznikají mapováním biotopů v terénu a jejich podkladem v ČR je Vrstva mapování biotopů (dále jako VMB), která vychází z biotopů v Katalogu biotopů (Chytrý M. et al., 2010). S využitím Dálkového průzkumu Země (dále jako DPZ) jsem ověřila možnosti, které mohou výrazně urychlit proces mapování biotopů, který v terénu trvá 12 let. Veškeré analýzy v této práci byly zaměřeny na mapování biotopů z nejpodrobnějších tříd spadajících pod kategorie mokřadů a pobřežní vegetace, pramenišť a rašelinišť a z vybraných tříd kategorií lesů a křovin. Mapování bylo soustředěno na lokalitu Karlovarského kraje. Podkladem pro mapování byly multispektrální snímky z družice Sentinel-2, která funguje od roku 2015.

(Chytrý M. et al., 2010) Mokřady jsou velice významným ekosystémem, jelikož udržují vodu v krajině a tím přispívají do koloběhu vody v přírodě. Patří mezi nejohroženější ekosystémy, nejen v České republice dochází k jejich degradaci či úplnému zániku. Spadají pod ně i rašeliniště, která jsou významná především proto, že zadržují velké množství uhlíku a tím přispívají ke zpomalení klimatických změn (Natura 2000, European Commission, 2021). Pro mapování jsem se zaměřila na mokřadní biotopy a zvolila jsem území Karlovarského kraje, kde jsem vycházela z výskytu biotopů během let 2015 a 2018.

Detekce biotopů probíhala za použití řízené klasifikace na datech z VMB a družicových snímků Sentinel-2 v programu Google Earth Engine. V této práci byla ke klasifikaci biotopů využita řízená klasifikace Random forest a Classification and Regression Trees.

2. Cíle práce

Tato bakalářská práce se zabývá možnostmi, kterými se dají zmapovat biotopy ze soustavy Natura 2000 za multispektrálních snímků. V této práci byly stanoveny následující dílčí cíle:

Vytvořit literární rešerši vycházející z výzkumů zabývajících se klasifikováním vegetačního pokryvu.

Zjistit, které biotopy jsou spektrálně odlišné a jsou vhodné pro trénovací data ke klasifikaci.

Vybrat vhodné klasifikační algoritmy pro mapování mokřadů v Karlovarském kraji.

Ověřit, zda lze detekovat lépe hlavní kategorie nebo podrobné třídy.

Zhodnotit přesnosti výsledků klasifikace ve srovnání s podkladem VMB.

3. Literární rešerše

3.1 Chráněné biotopy soustavy Natura 2000

Natura 2000 je největší koordinovaná síť chráněných oblastí na světě. Jedná se o soustavu chráněných území, která obsahuje především nejvzácnější a nejohroženější evropské druhy a stanoviště (Levin et al., 2007). Společně ji vytvářejí členské státy Evropské unie. Vznikla za účelem ochrany druhů živočichů, rostlin a biotopů v roce 2000. První mapování biotopů pro tyto účely probíhalo na českém území mezi lety 2001 a 2005. Skládá se ze dvou typů chráněných území, jedná se o ptačí oblasti a evropsky významné lokality (Ministerstvo Životního Prostředí, 2021.).

Podkladem pro biotopy zahrnuté v soustavě Natura 2000 se stala Vrstva mapování biotopů (VMB). Mapování biotopů ČR je projekt, který funguje od roku 2000. Cílem bylo vytvoření odborného podkladu pro vyznačení evropsky významných lokalit (dále jako EVL) do soustavy Natura 2000. Základní VMB byla dokončena v roce 2005 a s ní vznikla i celoplošná informace o výskytu a stavu biotopů v ČR (AOPK ČR, 2021.). Ukázalo se, že tato vrstva byla pro ochranu přírody velkým přínosem, mapování biotopů tudíž pokračovalo. Od roku 2006 je celé území ČR aktualizováno. Zaktualizovat celé území České republiky trvá 12 let. Biotopy ve VMB jsou popsány v příručce Katalog biotopů (Chytrý M. et al., 2010), která byla vydaná Agenturou ochrany přírody a krajiny České republiky.

Biotopy v Katalogu biotopů jsou rozděleny do 9 základních skupin. Osm z nich je pod hlavním zájmem ochrany přírody (jedná se o vodní toky a nádrže, mokřady a pobřežní vegetaci, prameniště a rašeliniště, skály, sutě a jeskyně, alpínské bezlesí, sekundární trávníky a vřesoviště, křoviny a lesy), zbývající skupina zahrnuje biotopy, které jsou silně ovlivněné nebo vytvořené člověkem (Chytrý M. et al., 2010).

V této bakalářské práci jsem se zaměřila na detekci mokřadních biotopů. Mokřady jsou cenné přírodní zdroje, které poskytují životnímu prostředí mnoho výhod. Mezi mokřady dle definice Ramsarské úmluvy (Russi et al., 2012) patří území bažin, slatin, rašelinišť a další území pokrytá vodou, vytvořená přirozeně či uměle. Voda v mokřadech může být stojatá či tekoucí, sladká, brakická nebo slaná. Vyskytují se po celém území České republiky až na vyšší horské polohy (Chytrý M. et al., 2010). Mokřady jsou velice významným ekosystémem, jelikož udržují vodu v krajině a tím přispívají do koloběhu vody v přírodě. (Moffett & Gorelick, 2013). Rašeliniště mají důležitou roli v klimatických změnách z důvodu vysokého zadržování uhlíku.

Zadržují vodu a tím přispívají svým výparem do koloběhu vody v přírodě. Rašeliniště vznikají na svazích i rovinách, struktura povrchu se odvíjí od množství a rovnoměrnosti rašeliny (Chytrý M. et al., 2010). Pro tuto práci jsem zvolila mokřadní biotopy z kategorie mokřady a pobřežní vegetace, prameniště a rašeliniště, lesy a z kategorie křoviny.

3.2 Multispektrální snímky

K mapování mokřadních biotopů jsem využila multispektrální snímky z družice Sentinel-2 (dále jako S2), která funguje od roku 2015. Jedna družice prolétne nad stejným místem na Zemi každých 10 dní, ve vyšších zeměpisných šířkách se frekvence přelétnutí družice zkracuje. Jelikož se skládá ze dvou samostatných družic (Sentinel 2A a 2B), frekvence se snižuje o polovinu. To znamená, že v České republice jsou dostupné snímky každé 2 až 3 dny. S2 je spravovaný programem Copernicus, který je řízený Evropskou unií a data jsou k dispozici zdarma (Copernicus - Sentinel-2 2021), tím se řadí mezi open-source data. Snímky z S2 jsou multispektrální, jelikož družice nese optický multispektrální nástroj (MSI), který snímá ve 13 pásmech na škále od viditelného po střední infračervené spektrum. 3 pásma se nacházejí v červeném okraji viditelného spektra, která jsou vhodná k mapování vegetace. Pásma mají prostorové rozlišení 10 m, 20 m nebo 60 m.

Jelikož v této bakalářské práci využívám multispektrální snímky, tak přiblížím jejich využití. Multispektrální optické snímky obsahují velké množství informací, které lze použít k rozlišení prvků na zemském povrchu. Multispektrální snímání je založené na odrazivosti povrchu. Sluneční záření směřuje k Zemi a vlnové délky světla jsou absorbovány a odráženy každým povrchem jinak. Odražené sluneční záření je zachyceno senzorem neseným družicí. Údaj o odrazivosti lze použít nejen ke klasifikaci využití půdy nebo jeho pokryvu na základě spektrálních vlastností objektů (Peter, 2021). Existuje mnoho možností využití snímků k monitorování krajinného pokryvu a jeho využití. Multispektrální snímky lze použít k mapování lesů, zástavby, obsahu uhlíku a přírodních hazardů (Kovyazin et al., 2020). Další využití zahrnuje zemědělství nebo mapování vodních zdrojů. Mnoho studií kombinuje zpracovávání družicových snímků ze S2 s daty vytvořenými za použití dalších metod DPZ, například bodová mračna z Lidaru, snímky z dronů nebo z dalších multispektrálních družic například Landsat od USGS (Phiri et al., 2020)

Mapování stanovišť pomocí DPZ je z pohledu mapování krajinného pokryvu běžnou praxí a je prováděno mnoha způsoby. V posledních letech je využití DPZ pro mapování a

hodnocení biotopů stále více rozšířené po celém světě. Mapování lze aplikovat na mnoho druhů biotopů, například pro lesy, mokřady či traviny (Corbane et al., 2015). DPZ je v současné době na takové úrovni, kdy je jeho použití možné aplikovat k ochraně přírody a vytvářet periodické zprávy o současném stavu krajiny (Schmidt et al., 2017). Satelitní data mají pokrytí v pravidelné frekvenci, takže je možné stanoviště biotopů sledovat sezónně. Ze satelitních snímků je možné získat mnoho dalších informací o okolním využití půdy a jejich změnách v čase. Použití DPZ dat ke klasifikaci krajinného pokryvu je oproti sbírání dat v terénu či leteckému snímání méně nákladné a méně časově náročné nejen pro velké geografické oblasti (Ozesmi & Bauer, 2002).

V současné době již bylo provedeno mnoho výzkumů, které byly zaměřené na práci se satelitními snímky pro mapování biotopů ze soustavy Natura 2000. Anglický výzkumný tým Gavish et al. (2018) porovnával klasifikační systémy povrchu CORINE, FAO-CCS a EUNIS. Ze všech schémat zde vychází, že se změnami prostředí v závislosti na spektrální odrazivosti a vegetačních indexech (např. NDVI) z multispektrálních snímků se zvyšují hodnoty diverzity. Další mapování biotopů Natura 2000 proběhlo za využití družice QuickBird (Förster et al., 2008), v němž byly jako podkladová data použity lesnické mapy nebo typografické vegetační mapy. Výzkum byl zaměřený na výskyt a kvalitu biotopů dle německé Natura 2000 za pomoci dvou metod klasifikace. Další výzkum zaměřený na mapování biotopů (Schmidt et al., 2017) byl zaměřený na hodnocení stavu ochrany vřesovišť. Pro práci kombinovali hyperspektrální data z AISA Dual a RGB snímky z dronu (UAV) nad lokalitou. Dosud však neproběhl žádný výzkum, který by se soustředil na VMB, která je pouze pro ČR a je zároveň velmi podrobná a propracovaná.

Vzhledem k tomu, že v této bakalářské práci pracuji s mokřady, které jsou přírodním objektem, souvisí s jejich charakteristikou i změna spektrálních vlastností. Se změnou ročního období se v průběhu roku mění spektrální variabilita krajiny. Tyto změny jsou důsledkem přírodních procesů, jako jsou například změna barvy listů, která je výrazná především na jaře, vývoj listové plochy v průběhu roku, změna struktury půdy při jejím sečení či pastvě, období květu nebo stárnutí listů na podzim (Levin et al., 2007). Hypotéza spektrální variability (dále jako SVH), uvádí, že „druhovú bohatost souvisí s jakýmkoli objektivním měřítkem (např. směrodatná odchylka) změny ve spektrálních charakteristikách snímků z dálkového průzkumu.“ (Palmer, 2002). Je mnoho zastánců SVH (Rocchini et al., 2019), ale také se hypotéza často zpochybňuje. Například výzkum od Schmidtlein et al. (2017) je zaměřen na tvrzení, že krajinný pokryv (land cover) je hlavním důvodem proč SVH neplatí v aplikování na krajinu. V tomto výzkumu použili dvě měření spektrální variability, jednu na základě

euklidovské vzdálenosti z průměrných hodnot spektrálních vlastností v časové řadě snímků MODIS a druhou na základě neřízené klasifikace (k-means shluková analýza). Průměrná rozmanitost stanovišť z hlediska odrazivosti není tak indikativní pro hojnost druhů, jako spíše pro počet stanovišť v oblasti. Stanoviště je možné zjistit z počtu shluků, které vyplývají z algoritmu k-means (Schmidtlein & Fassnacht, 2017). Z těchto informací vyplývá, že k detekci mokřadů je vhodné využít spektrální variabilitu ke klasifikaci mokřadů například v podobě směrodatné odchylky indexu NDVI.

Pro přesné měření je důležité, aby sledované biotopy byly vzájemně ve spektrálních vlastnostech odlišné, proto pro detekci mokřadů je vhodné vytvořit spektrální analýzu obrazu z jednoho pixelu. Lze vytvořit automatizovaný proces, který při přiblížení umístí každý obrazový pixel do třídy (např. druh biotopu) na základě jeho spektrální vlastnosti tak, aby byl optimálně oddělitelný od ostatních tříd. Tato metoda je obecně úspěšná při identifikaci celých biotopů. Mapování menších prvků v mokřadech je kvůli jejich malé velikosti obtížnější, kontrast mezi rostlinnými druhy se tudíž nachází v nižších spektrech (Moffett & Gorelick, 2013). Při mapování mokřadů může dojít také k nesrovnalostem, které mohou být zapříčiněny charakteristikou mokřadů nebo již vstupními daty z Natura 2000. Mokřady jednoho typu, mohou být zalesněné, křovinné nebo bylinné (Mahdavi et al., 2018). Podobné druhy mokřadů, které by měly být klasifikovány samostatně, sdílejí vzájemně podobné ekologické vlastnosti a tím může dojít k záměně v DPZ prostředí.

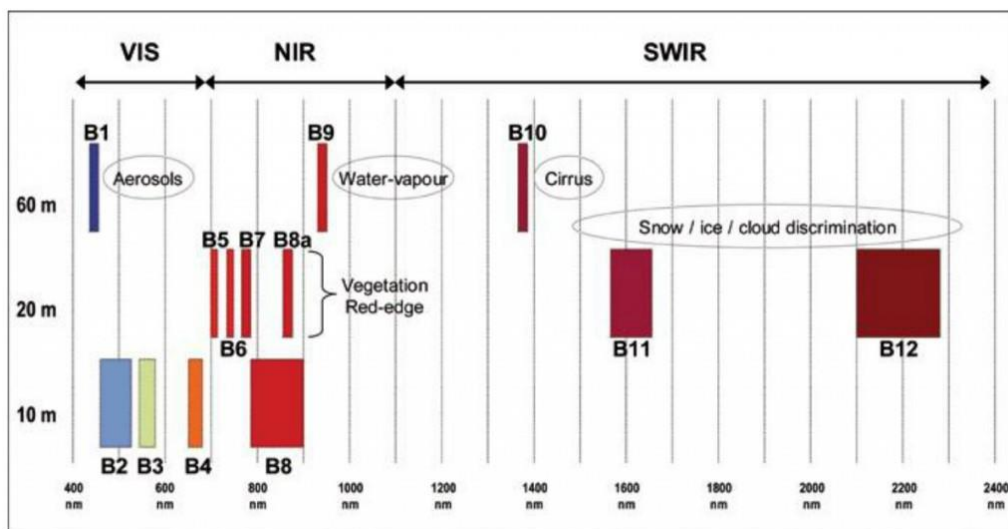
Vzhledem k tomu, že podkladem pro tuto práci jsou data, která byla sesbírána v terénu, je důležité brát ohled na možné chyby již při jejich zpracování. Při práci s těmito daty může dojít ke zvýraznění spektrálních chyb, což může být zapříčiněno subjektivním vnímáním odborníků v terénu, kteří často místně relevantní druhy nadhodnocují a místně hojně druhy podhodnocují (Förster et al., 2008). Z tohoto důvodu spolehlivost dat musí být posuzována v kontextu s cílem omezit nepřiměřenou kritiku a falešné vnímání výsledků založených na DPZ. Je důležité odvozenou přesnost chybové matice interpretovat s ohledem na tyto vlivy (Schmidt et al., 2017).

Před zpracováním satelitních snímků je důležité data prostudovat a připravit na další zpracování. Družicová data před dalším zpracováním mají hotové atmosférické korekce odrazivosti (Ruiz et al., 2021). Data ze Sentinel jsou poskytována ve dvou formách, a sice Level-1C, který reprezentuje TOA (top of atmosphere) tedy korekce odrazivosti ve vrchní atmosféře, a Level-2A, ve kterém dochází k BOA (Bottom of the atmosphere) což je korekce

odrazivosti ve spodní atmosféře (Rumora et al., 2020). Pro použití snímků ke klasifikaci vegetace je vhodnější využití Level-2A.

Pro další zpracování je důležité vybrat vhodná spektrální pásma, zde záleží na mapovaných biotopech, jak se v daných pásmech chovají. Mapování vegetace se běžně provádí s využitím pásem infračerveného spektra. Mezi tzv. vegetační infračervená pásma (dále jako VRE) se řadí spektra (Bandy) B5, B6, B7, B8A (Qiu et al., 2017). Období, ve kterém byly zpracovávány snímky zaznamenány, je vhodné vybrat ve vegetačním období, které probíhá na území ČR od dubna do září. Ze snímků s nízkým výskytem mraků je vhodné vytvořit kompozit. To znamená, že se vhodné snímky zprůměrují a vytvoří se tzv. mozaika (Rapinel et al., 2019).

Dalším důležitým krokem pro přípravu družicových dat je tzv. převzorkování. Pro využití multispektrální snímků k mapování vegetace je nejvhodnější rozlišení 10 m, avšak pouze pásma ve viditelném spektru a v blízkém infračerveném (NIR) jsou v tomto rozlišení. Ostatní pásma, která jsou vhodná pro mapování vegetace, je nutné před dalšími analýzami převzorkovat (resample) (Gudmann et al., 2020). Na obrázku č.1 je všech 13 pásem s příslušným rozlišením a vlnovou délkou, zároveň je u každého pásma stručný popis jeho významu. Převzorkování pásem funguje tak, že dojde k přetřansformování každého pixelu na více menších pixelů o požadované délce 10 m. Zachovává stejné číselné hodnoty a polohu, tudíž se původní data snímku nemění (Sánchez-Espinosa & Schröder, 2019).



Obrázek č. 1: Sentinel-2, pásma v jednotlivých rozlišeních. Heslovitě popsany účel pro jednotlivá pásma B1 (445 nm), B2 (490 nm), B3 (560 nm), B4 (665 nm), B5 (705 nm), B6 (740 nm), B7 (783 nm), B8 (842 nm), B8A (865nm), B9 (940 nm), B10 (1375 nm), B11 (1610 nm) a B12 (219, nm) B1 (445 nm), B9 (940 nm), B10 (1375 nm).

3.3 Klasifikace biotopů

Vhodnými vstupními daty pro mapování vegetace bývá spojení mediánů více snímků z různých časových období, které se většinou nachází v úseku vegetačního období (Jiao et al., 2019). K dosažení spolehlivějších výsledků je možné před samotnou klasifikací vytvořit některý z vegetačních indexů. Ve studii od Bhatnagar et al. (2020), zaměřené na mapování mokřadů ze snímků S2 v Irsku, bylo využito 10 pásem ve zpracování Level-2 a tři indexů: Normalised Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI) and Normalised Difference Water Index (NDWI). Výsledkem bylo, že při spojení indexů a spektrálních pásem pro monitorování stavu mokřadů vychází spolehlivější výsledky. Zjistili také, že klasifikace je zároveň závislá na velikosti vegetačních společenství v každém mokřadu (Bhatnagar et al., 2020). Jelikož mokřady mohou mít na svém území i vyšší vegetaci, je účinnější využít pásma s vysokou vlnovou délkou. Ke spektrálním pásmům kromě vegetačních indexů lze připojit například DEM, což je digitální model terénu, či TWI (Topographic wetness index), který se zvláště hodí pro mapování mokřadů, neboť nese informaci o vlhkosti pokryvu (Ruiz et al., 2021).

Existují dva základní typy klasifikace, a sice řízená a neřízená klasifikace. Řízená klasifikace přiřazuje hodnoty v jednotlivých pixelech do předem stanovených kategorií. Při zpracování se vychází z informací z předem vytvořených trénovacích dat. Odlišnost množin je způsobena jejich odlišností ve spektrální odrazivosti. Pro klasifikaci se vybírají vhodná pásma z multispektrálních snímků, vychází se zde ze spektrálních vlastností dané kategorie. Základem je, aby kategorie byly co nejvíce odlišné. Pro ověření spolehlivosti se využívají validační data, těch je zpravidla více než trénovacích (Killeen et al., 2015). Trénovací a validační data mohou být v podobě náhodných bodů nacházejících se v lokalitě dané kategorie. Přesnost se hodnotí pomocí chybové matice (Briand & Wüst, 2002). Existuje více klasifikátorů, které jsou blíže popsány v podkapitole Klasifikátory, jedná se například o random forest, maximum likelihood, nearest neighborhood, (Gavish et al., 2018) atd. Důležité je zajistit, aby vstupní množiny byly z velké části homogenní a měly jednoznačnou spektrální charakteristiku, protože právě tyto faktory mohou nejvíce ovlivnit nesprávnost klasifikace.

Neřízená klasifikace na rozdíl od řízené klasifikace dostává reálné hodnoty o pokryvu až na samotném konci. Na začátku při použití shlukové analýzy automatizovaně vytvoří třídy na základě spektrální odrazivosti. Proces shlukování se opakuje, proto je tato metoda časově náročnější. Výhodou je, že je možné odlišit i malé spektrální třídy. Klasifikovaný výsledek je

následně porovnáván s mapami či leteckými snímky a pro vzniklé třídy se zjišťuje informace o pokryvu půdy (Dobrovolný, 1998).

Na základě mnoha studií, které se zabývají klasifikací na datech ze S2, bylo prokázáno, že použití řízené klasifikace je využíváno častěji než neřízená klasifikace (Phiri et al., 2020). Mnoho klasifikačních algoritmů vychází z metody řízené klasifikace. Řízená klasifikace založená na jednotlivých pixelech se často používá v kombinaci s objektovou analýzou obrazu (OBIA) (Moffett & Gorelick, 2013). Tato kombinace vytváří klasifikace s vysokou přesností. Klasická klasifikace založená na pixelech přiřazuje třídu krajinného pokryvu jednotlivému pixelu. Všechny pixely mají stejnou velikost, stejný tvar a nemají žádnou spojitost se svým okolím (Phiri et al., 2020). OBIA vychází ze segmentů, seskupuje malé pixely do vektorových objektů. Namísto jednotlivých pixelů segmentace automaticky digitalizuje obraz (GIS Geography, 2021). Ve výsledku je například možné rozlišit řeku od rybníku, či spektrálně stejné jevy, které jsou ve skutečnosti odlišené.

Přesnost klasifikace lze zjistit z chybové matice (confusion matrix), která je vypočítána z výsledků klasifikace a z podkladových dat, kterými mohou být letecké snímky či data z terénu (Killeen et al., 2015). Chybová matice je čtvercová, jsou zde zastoupeny počty pixelů a hodnoty z klasifikace, příklad vizualizace matice je na Obr. č. 2. Takto vizualizovaná matice je vhodná k určení větších chyb, například chybně klasifikovaná celá třída. Výsledná hodnota ukazuje na kolik procent klasifikace odpovídá skutečnosti (Kulkarni et al., 2020). Ke zjišťování podrobných přesností se využívají metody přesnosti, a sice přesnosti zpracovatele (producer's accuracy), přesnost uživatele (user's accuracy) a celková přesnost (overall accuracy). Přesnost zpracovatele je vypočítána z počtu správně klasifikovaných pixelů vydělených celkovým počtem pixelů. Tato přesnost je založená na předpokladu, že pixel byl správně klasifikován. Přesnost uživatele je naopak založená na pravděpodobnosti, s jakou byl pixel klasifikován špatně. Vypočítá se rozdílem mezi správně klasifikovanými pixely a pixelů, které byly použity jako trénovací. Celková přesnost udává poměr mezi všemi správně klasifikovanými pixely a celkovým počtem pixelů (Dobrovolný, 1998)

		Reference image			Total
		A	B	C	
Classified image	a	37	3	7	$\Sigma a = 47$
	b	9	25	5	$\Sigma b = 39$
	c	11	2	43	$\Sigma c = 56$
Total		$\Sigma A = 57$	$\Sigma B = 30$	$\Sigma C = 55$	$N = 142$

Obrázek č. 2: Ukázka chybové matice, vzorec pro výpočet celkové přesnosti (N) a přesnosti uživatele (Total na pravé straně) a zpracovatele (Total v dolním řádku matice).

3.3.1 Klasifikátory

Algoritmů řízené a neřízené klasifikace, které jsou vhodné k mapování vegetace, existuje mnoho. V této práci popíšu pouze zástupce, které jsou vhodné pro detekci mokřadních biotopů nebo pro odlišení spektrálních vlastností.

Random forest (dále jako RF) je kombinovaná metoda strojového učení (machine learning), která je využívána pro klasifikaci a regresi. Rozhodovací klasifikační proces je založený na velkém množství rozhodovacích stromů, na základě výběru náhodné podmnožiny trénovacích souborů a proměnných (Schmidtlein & Fassnacht, 2017). Základním procesem klasifikátoru je meta-algoritmus strojového učení bagging. Vytváří velké množství učících se podmnožin z tréninkových dat, které na sobě nejsou závislé (Fawagreh et al., 2014). Každá podmnožina ve stromové struktuře je zařazena do určité třídy. Poté jsou z dílčích oddílů podmnožin, na základě tzv. hlasování o klasifikaci, zařazeny do třídy s největším počtem hlasů. Zařazování se opakuje až do chvíle, kdy funkci nelze dále minimalizovat. Na rozdíl od jiných klasifikátorů, RF je vhodný pro práci s datovou sadou s mnoha proměnnými zejména proto, že zvládne velké rozdíly v krajinném pokryvu tříd (Slagter et al., 2020), proto je často využíván k mapování přírodní vegetace. Výstupem je průměrná predikce jednotlivých stromů (Rapinel et al., 2019).

Classification And Regression Trees (CART) je také založený na rozhodovacích stromech. CART je binární rozhodovací strom, který vzniká opakovaným dělením uzlu neboli kořene do dvou dceřiných uzlů, které jsou dále binárně děleny na další uzly. Kořen obsahuje hodnotu a prediktor X na jejichž základě dochází k opakovanému rozdělování na homogennější skupiny (De' Ath & Fabricius, 2000). Za účelem vytvoření optimální velikosti stromu dochází

k tzv. prořezávání, které vychází z počtu pozorování v uzlu. Každá skupina je dána typickou hodnotou proměnné odezvy, počtem výskytu ve skupině a hodnotou vysvětlujících proměnných, které skupinu definují (Oshiro et al., 2012). CART je ideální pro analýzu komplexních ekologických dat.

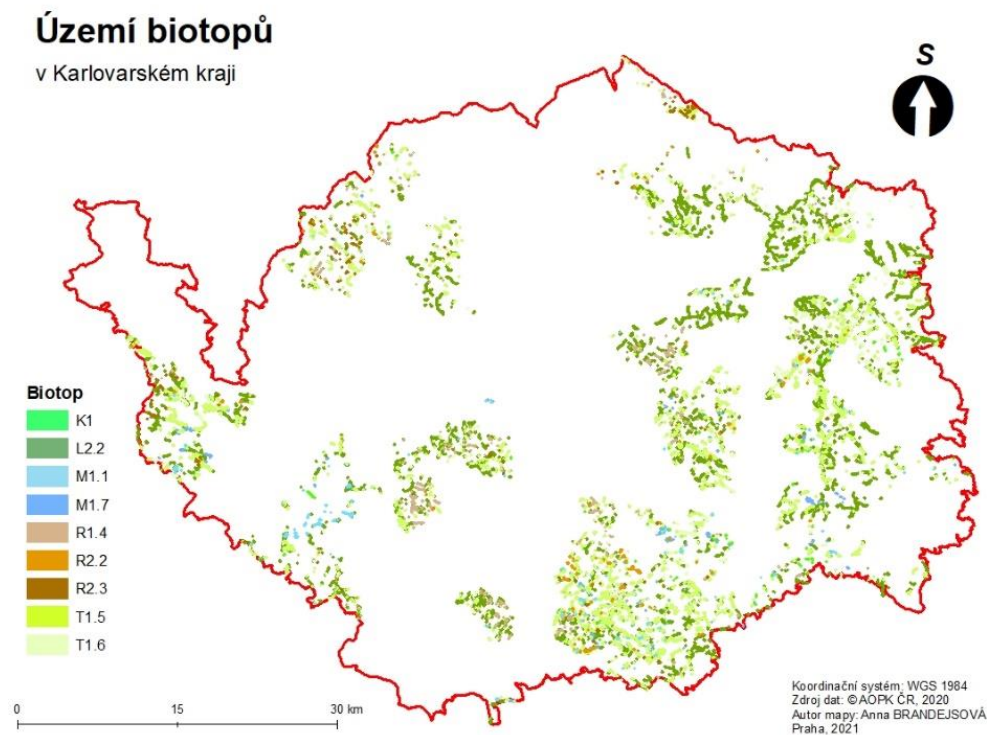
Metoda k-means je nejjednodušší algoritmus neřízené klasifikace. Algoritmus je založen na shlukování, vychází z počtu shluků a počtu iterací. Iterace udává počet opakování procesu porovnávání okolí jednotlivých pixelů (Slagter et al., 2020). Tato metoda je časově náročnější. Je vhodná k rozeznání spektrálních odlišností. Při pozorování spektrální variability je metoda k-means v průběhu roku velmi stabilní. U této metody je nízký počet tříd a může dojít k reakci, že například listnatý les, který ztratí zeleň, bude stále jednou třídou (Oshiro et al., 2012). Další algoritmus neřízené klasifikace je ISODATA. Je velmi podobný algoritmu k-means, pouze nemusí mít přesně stanovený počet shluků.

4. Metodika

4.1 Výběr zájmového území a vhodných biotopů

Lokalita

Na začátku analýzy bylo prvním úkolem najít vhodnou lokalitu s vysokým zastoupením mokřadů, které byly ve vrstvě mapování biotopů zaktualizovány po roce 2015. Pro porovnání zastoupení vhodných biotopů ve všech krajích ČR jsem z webové stránky AOPK ČR získala veškeré biotopy VMB. Data byla ve formátu shapefile, který je kompatibilní s programem ArcMap. Ze všech krajů České republiky se největší zastoupení mokřadů a rašelinišť nachází v Karlovarském kraji.

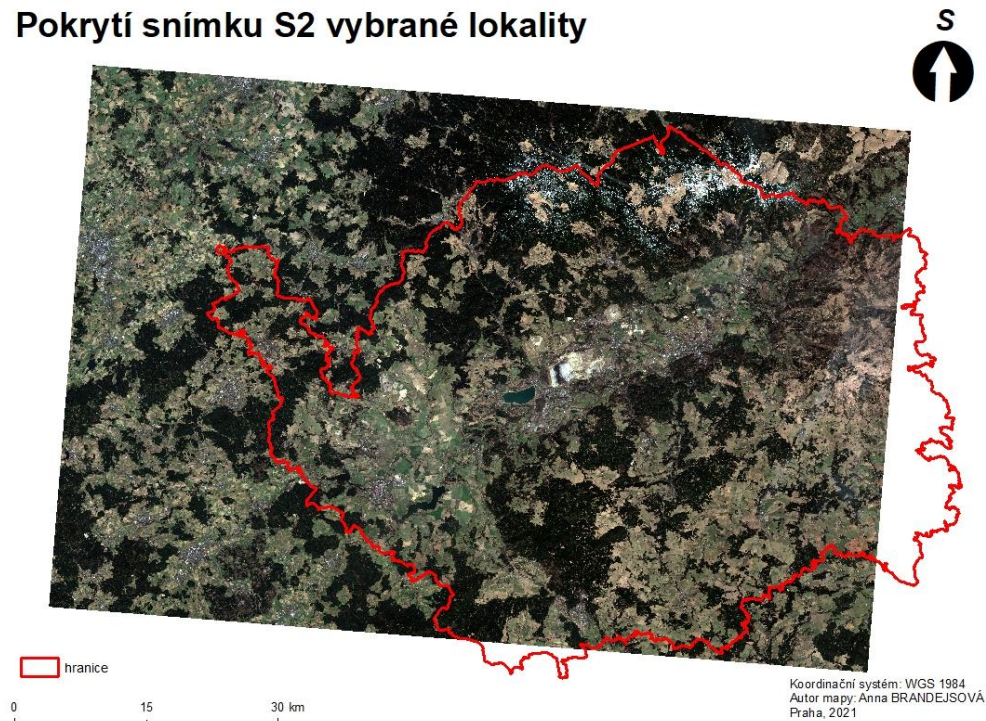


Obrázek č. 3: Rozložení výskytu vybraných biotopů na zájmovém území, Mokřadní vrby K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnovec R1.4, Nevápnitá mechová slatiniště R2.2, přechodová rašeliniště R2.3, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.

Další důvod pro výběr lokality byl, že se biotopy nachází rozmístěné po celém území kraje viz Obr. č. 3. Karlovarský kraj je z většiny pokryt snímkem ze S2, konkrétně snímkem S2A, pokrytí je zobrazeno na obrázku č. 4. Z těchto důvodů jsem pro zpracování použila celý kraj a na další analýzy jsem k ohraničování využívala jeho hranice. Karlovarský kraj se rozkládá na rozloze o 3315 km². Biotopy se zde nacházejí jak v nížinách, tak ve vyšších nadmořských

výškách v Krušných horách. Tuto skutečnost je vhodné zahrnout do analýzy, jelikož biotopy mohou s nadmořskou výškou měnit svůj spektrální charakter. Rozsah nadmořských výšek na celém území je od 320, n. m. do 1244 m n. m.

Pokrytí snímku S2 vybrané lokality



Obrázek č. 4: Pokrytí zájmové lokality snímkem S2.

Biotopy

Dalším krokem bylo vybrat vhodné biotopy s dostatečným zastoupením, které se nacházejí na vybrané lokalitě. Za využití katalogu biotopů jsem vybrala další vhodné biotopy pro zpracování celkově z 5 kategorií. Mapovanými biotopy byly primárně mokřady a rašeliniště. Jako další biotopy připadaly v úvahu některé třídy z kategorie lesů, sekundárních trávníků a vřesovišť a z kategorie křovin. Před dalším zpracováním dat z VMB jsem shapefile přetransformovala, za pomoci funkce *Project* v ArcMapu, z výchozího souřadnicového systému S-JTSK Křovák East-North na požadovaný evropský souřadnicový systém WGS-84, ve kterém jsou snímky z družice S2.

Biotop		Počet výskytů	km ²	Zastoupení biotopů %
Mokřadní vrbiny	K1	416	0.98	0.03%
Údolní jasanovo-olšové luhy	L2.2	2177	22.73	0.69%
Rákosiny eutrofních stojatých vod	M1.1	121	0.28	0.01%
Vegetace vysokých ostřic	M1.7	325	0.99	0.03%
Lesní prameniště bez tvorby pěnovců	R1.4	307	0.19	0.01%
Nevápnitá mechová slatiniště	R2.2	144	0.61	0.02%
přechodová rašeliniště	R2.3	168	0.58	0.02%
Vlhké pcháčové louky	T1.5	1670	13.06	0.39%
Vlhká tužebníková lada	T1.6	667	3.41	0.10%
			42.8	1.29%
Rozloha Karlovarského kraje km ²	3 314			

Tabulka č. 1: Zastoupení vybraných biotopů v porovnání s rozlohou Karlovarského kraje.

Kategorie biotopů jsem roztrídila na třídy, které jsem utvořila z nejpodrobnějších úrovní z Katalogu biotopů. Z jednotlivých tříd jsem výběr zúžila o záznamy biotopů, které byly zaktualizovány po roce 2015, jelikož družice S2 funguje od tohoto roku. Z tohoto zúženého výběru jsem vybrala pouze třídy, které byly na území dostatečně zastoupené. Zastoupení biotopů je zobrazené v tabulce č.1. Z toho vychází, že z kategorie mokřady a rašeliniště se na území hojně vyskytovaly třídy lesní prameniště bez tvorby pěnovců (na Obr. č. 5), nevápnitá mechová slatiniště (na Obr. č. 6), přechodová rašeliniště (na Obr. č. 7), rákosiny eutrofních stojatých vod (na Obr. č. 8) a vegetace vysokých ostřic (Obr. č. 9). Z kategorie sekundárních trávníků a vřesovišť pro analýzy připadaly v úvahu vlhké pcháčové louky (Obr. č. 10) a vlhká tužebníková lada (Obr. č. 11). U křovin byly vhodné pouze mokřadní vrbiny (Obr. č. 12). Z kategorie lesů na vybrané lokalitě byla výrazně zastoupená pouze třída údolní jasanovo-olšové luhy (Obr. č. 13).



Obrázek č. 5: Lesní prameniště bez tvorby pěnovců, Paukertová, 2009



Obrázek č. 6: Nevápnitá mechová slatiniště, AOPK ČR, 2012.



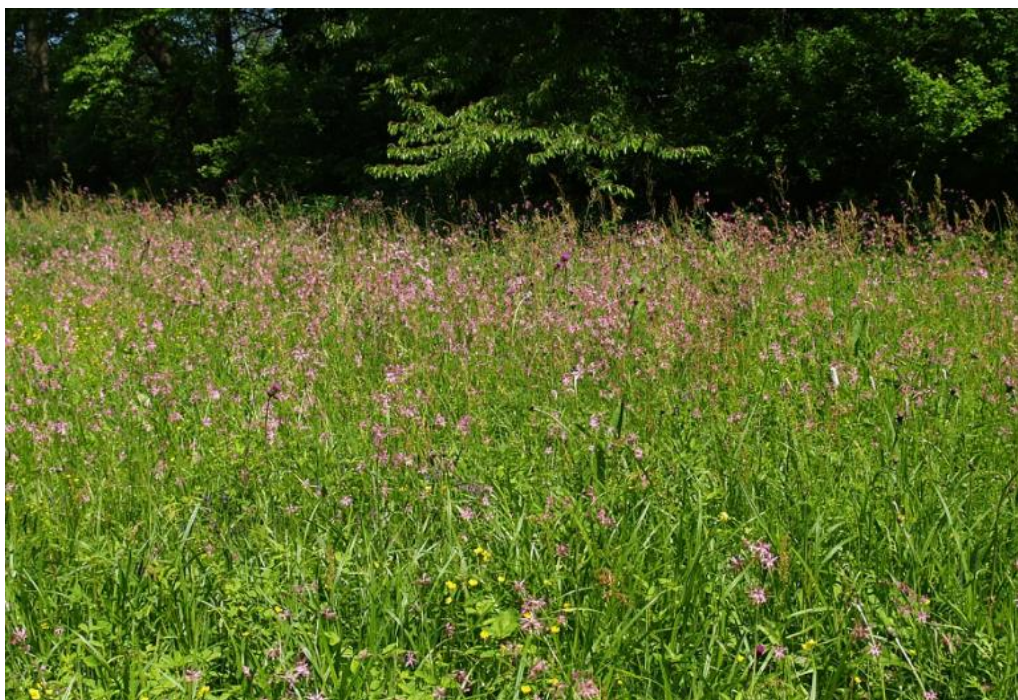
Obrázek č. 7: Přejíhodová rašeliniště, AOPK ČR, 2021



Obrázek č. 8: Rákosiny eutrofních stojatých vod (M1.1), Novák, 2020.



Obrázek č. 9: Vegetace vysokých ostřic (M1.7), Duchoslav, 2020,



Obrázek č. 10: Vlhké pcháčové louky (T1.5), Duchoslav, 2020.



Obrázek č. 11: Vlhká tužebníková lada, Otýpková, 2005.



Obrázek č. 12: Mokřadní vrbyny (K1), Paukertova, 2006.

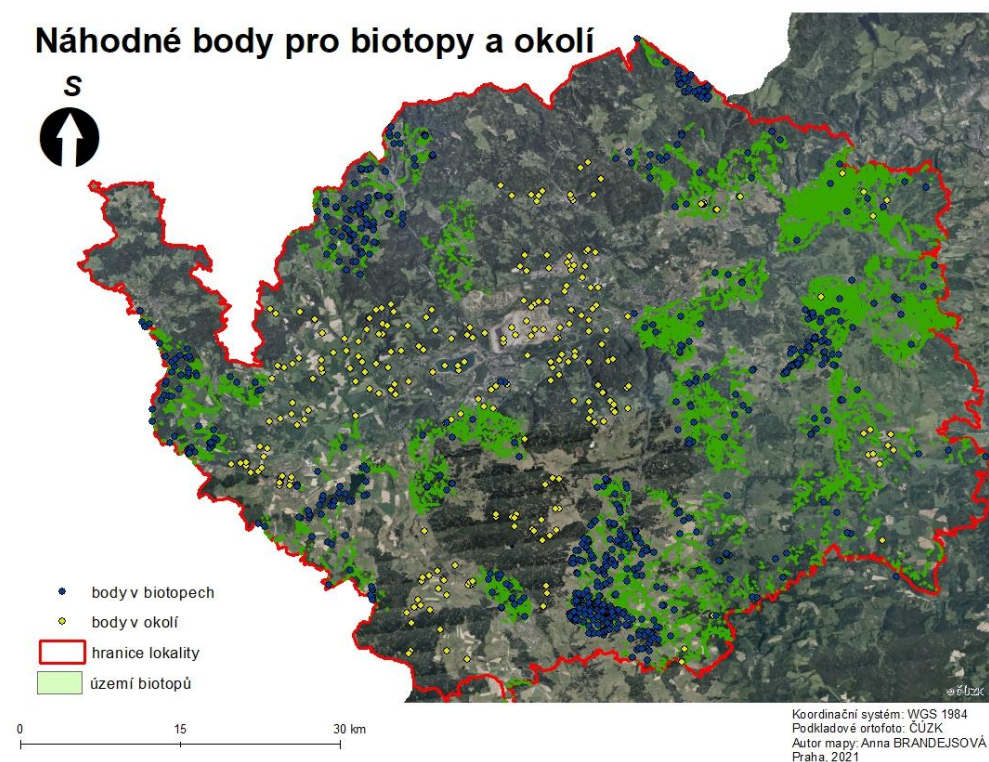


Obrázek 13: Údolní jasonovo-olšové luhy (2.2), Duchoslav, 2020.

4.2 Vstupní data

Přípravu vstupních dat pro klasifikaci jsem zpracovala v programu ArcMap. Zvolila jsem klasifikaci založenou na výběru pixelů a tudíž jsem vytvářela náhodné body. Nejprve jsem zajistila, že okrajové pixely polygonů nebudou spadat do jiné kategorie, protože analýza probíhala s daty v 10 m rozlišení. U všech vybraných biotopů z VMB jsem vytvořila *buffer* o velikosti -10 metrů, to znamená, že všechny biotopy byly oříznuty o desetimetrový okraj. Jelikož výskyt biotopů se často rozkládal na velmi malých plochách, při ořezu biotopů došlo ke snížení počtu záznamů. Zbylo tedy méně polygonů s větší rozlohou, ve kterých se jednotlivé pixely ze snímků S2 budou nacházet pouze v zájmovém území.

Náhodné body jsem generovala do připravených polygonů, pro každou třídu jsem zvolila stejný počet bodů. V každém polygonu byl vytvořen maximálně 1 bod s tím, že každá třída obsahovala 115 bodů nesoucích informace v atributové tabulce s názvem biotopu, ve kterém se nachází. Pro zjednodušení přepisování informací do bodů jsem použila funkci *Spatial join*. Každý bod obsahoval sloupec s názvem příslušné třídy biotopu. Tyto připravené body sloužily jako trénovací a validační body pro řízenou klasifikaci (viz Obr. č. 14)



Obrázek č. 14: Náhodné body v zájmovém území vytvořené pro okolí a lokality biotopů, podklad je ortofoto ČR

Do klasifikace vstupovaly body pro biotopy a okolí, podkladem pro vytváření okolních bodů byl snímek S2 a z něj neřízená klasifikace k-means, která byla rozdělena do tříd podle

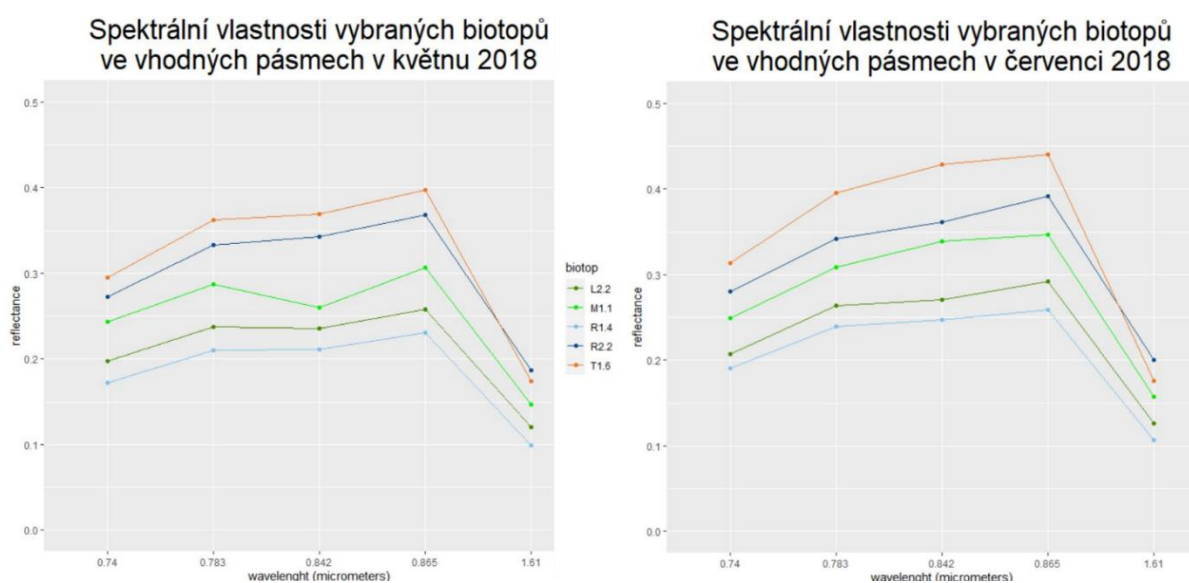
spektrálních vlastností povrchu. Z důvodu toho, aby se nestalo, že se náhodný bod nachází v neaktualizovaném biotopu, jsem pracovala s podkladovou vrstvou VMB. Podle spektrálních odlišností jsem vytvářela body pro zástavbu, silnice, vodu, pole a lesy, které nespádaly do chráněných biotopů. Pro vytvoření okolních bodů jsem vytvořila více variant a porovnávala jsem jejich spolehlivost při detekci. Do bodů jsem zakomponovala jevy, které musí být od biotopů odlišné. Zkoušela jsem například náhodné body vytvořené pouze z tříd z výsledku klasifikace k-means, dále body vytvořené ručně na území, kde se nacházela zástavba, silnice a pole. Také jsem zkoušela vytvořit náhodné body v polygonu ve středu lokality, kde se nenacházel žádný mapovaný biotop. Nejlepší vrstvou pro okolní jevy byla kombinace výše uvedených metod. Pro body jsem zajistila dostatečný rozestup, tím že jsem jako vstupní parametr pro generování zvolila minimální rozestup 10 metrů. Vytvořených bodů pro celé území je 300, protože okolí je na větší ploše nežli biotopy. Na obrázku č. 14 jsou zobrazeny všechny vytvořené náhodné body.

Spektrální charakteristiky

Jelikož klasifikace obrazu je založená na spektrálních charakteristikách, vstupní data vstupující do klasifikace se musí vzájemně odlišovat ve spektrálních vlastnostech. Pro zjištění spektrálních vlastností biotopů, které se skládají z informací o odrazivosti ve vlnové délce daného pásma z multispektrálního snímku, jsem provedla analýzu v Google Earth Engine (dále jako GEE). Výsledné spektrální hodnoty z GEE jsem zpracovala v programu R. Pro jednotlivé hodnoty odrazivosti biotopů a dané vlnové délky jsem vytvořila grafy za pomoci pluginu ggplot2. Každý graf je pro jeden měsíc a obsahuje informaci o odrazivosti v daném pásmu pro každý biotop. V příloze č. 1 jsou příslušné grafy, tzv. spektra in area, vytvořené pro všechny měsíce z vegetačního období.

Z grafů spektrálních vlastností pro celé vegetační období vychází, že k mapování vybraných biotopů jsou nejvíce vhodná pásma ve vlnových délkách 705 nm, 740 nm, 783 nm, 842 nm, 865 nm, 1610 nm a eventuelně i 2190 nm. Tyto vlnové délky odpovídají pásmům B5, B6, B7, B8, B8A, B11, B12, které se převážně nachází v infračerveném spektru. Pásma B5, B6, B7 a B11 je nutné převzorkovat na rozlišení 10 m, toto je patrné z obrázku č. 1 v kapitole 3.1. Pro zpracování klasifikace byla vytvořena kombinace více vhodných spekter ze dvou měsíců, především z pásem, která ovlivňují mapované biotopy, za účelem vytvoření optimálního rastru. Vhodnými měsíci pro kompozit jsou květen a červenec, jelikož hodnoty spektrálních vlastností v dubnu jsou ve vyšších nadmořských výškách ovlivněny výskytem sněhu a snímky v kompozitu z červnu mají v roce 2018 vyšší oblačnost.

Z předešlých analýz byly vybrány potenciálně odlišitelné biotopy. Vhodnými biotopy se stalo pouze 5 tříd ze 4 kategorií, které byly dostatečně odlišné ve svých spektrálních vlastnostech. Tyto biotopy byly použity pro následnou klasifikaci. Vybranými biotopy pro klasifikaci obrazu se staly: Lesní prameniště bez tvorby pěnvců (R1.4), Nevápnitá mechová slatiniště (R2.2), Rákosiny eutrofních stojatých vod (M1.1), Vlhká tužebníková lada (T1.6) a Údolní jasanovo-olšové luhy (L2.2). Obrázek č. 15 zobrazuje rozdíly hodnot pro všechna pásma, ve kterých jsou výrazné spektrální odlišnosti. V tabulce č. 2 je zobrazeno zastoupení vybraných biotopů v porovnání s rozlohou celého kraje.



Obrázek č. 15: Odlišnost spektrálních vlastností vybraných biotopů v květnu a červenci ve všech vhodných pásmech pro mapování.

Biotope		Počet výskytů	km ²	zastoupení biotopů %
Údolní jasanovo-olšové luhy	L2.2	2177	23.0	0.69%
Rákosiny eutrofních stojatých vod	M1.1	121	0.28	0.01%
Lesní prameniště bez tvorby pěnvců	R1.4	307	0.19	0.01%
Nevápnitá mechová slatiniště	R2.2	144	0.61	0.02%
Vlhká tužebníková lada	T1.6	667	3.41	0.10%
			27.5	1%

Tabulka č. 2: Tabulka zastoupení pro biotopy, které jsou spektrálně odlišné a byly použity do klasifikace.

4.3 Zpracování klasifikace

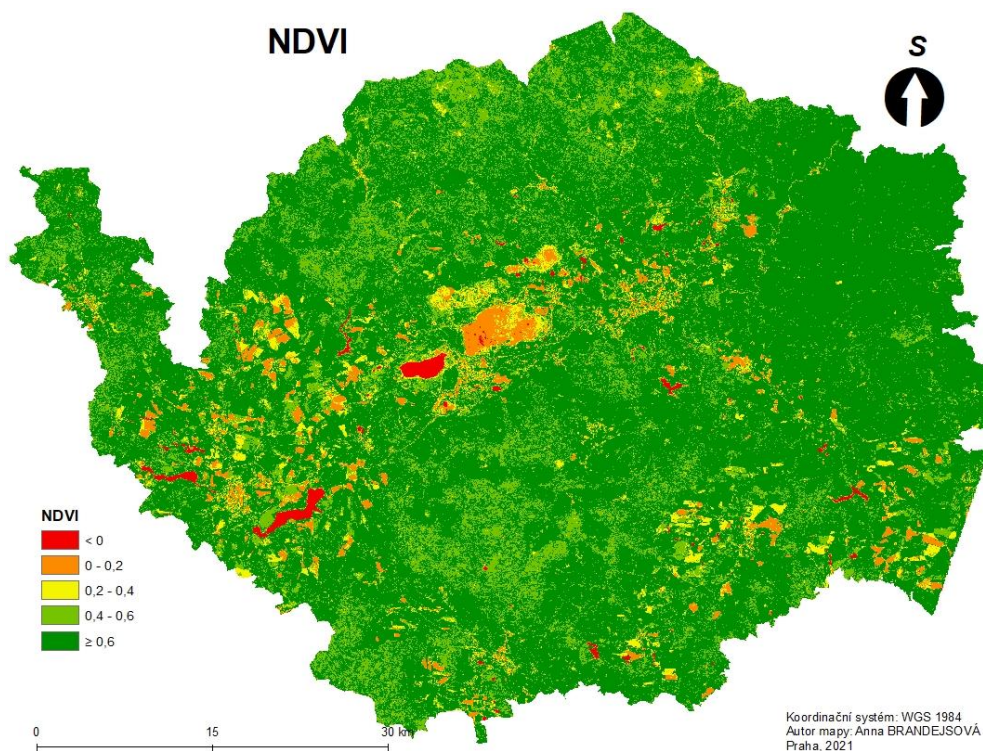
Řízenou klasifikaci obrazu jsem vytvářela na platformě Google Earth Engine (dále jako GEE), kde je možné zpracovávat satelitní data z více zdrojů. Tato platforma je především vhodná zpracování území o větší rozloze, další výhodou je, že na jednom místě je dostupné velké množství dat a informací. V tomto prostředí se příkazy zadávají v programovacím jazyku

JavaScript. GEE jsem vybrala na zpracování klasifikace, především protože na domácí podmínky je toto nejvhodnější alternativa, především pro náročné výpočetní operace, jako je klasifikace obrazu (Bourgeau-Chavez et al., 2017). Zároveň jsem detekci prováděla v prostředí bez licence, tudíž je tento výzkum dostupnější více zájemcům o tuto problematiku. Jen v GEE například lze snímky S2 zpracovat jako kompozity nejvhodnějších pásem, lze vytvářet vegetační indexy, spektrální charakteristiky a klasifikace obrazu. Dále je toto prostředí vhodné k výběru snímků s nízkou oblačností.

Za účelem získání snímků s nízkou či žádnou oblačností je vhodné vytvořit mediánový kompozit ze snímků z vybrané časové řady. Pro maskování mraků se používají prahové hodnoty pro Band 2, tím lze odstranit husté mraky. K odstranění mraků typu cirrus se využívá Band 10 (Oshiro et al., 2012). Pro co nejpřesnější analýzy bylo důležité vybrat snímky s nízkou oblačností. Pracovala jsem se snímky z vegetačního období, které je vhodné při práci s vegetací a zároveň v ČR má nízkou oblačnost.

Jelikož existuje mnoho návodů a příkladů příkazů (skriptů), inspirovala jsem se výzkumem (Nomura & Mitchard, 2018), který byl zaměřený na klasifikaci krajinného pokryvu ze snímků S2. Zabýval se klasifikací na menších území v Jižním Myanmaru, detekovali zde palmové plantáže. Cílem tohoto výzkumu bylo odlišení druhů palm. Ke klasifikaci využívali klasifikátor RF a CART za použití všech 13 pásem S2. Použitý a upravený skript je uveden v příloze č. 2. Tento skript jsem upravila pro své potřeby. Především jsem využívala mediánový kompozit namísto jednoho nejméně oblačného snímku a využívala jen vybraná pásma z více kompozitů.

Vzhledem k faktu, že přidáním indexů mezi vstupní informace se může zvýšit přesnost (Bhatnagar et al., 2020), jsem vytvořila z kompozitu mediánů pásem pro květen normalizovaný vegetační index (NDVI). Pro zpřesnění výsledků klasifikace vegetace jsem NDVI zahrнула do vstupních pásem, který je na obr. č. 16. Z NDVI jsem vytvořila směrodatnou odchylku, kterou jsem také přidala mezi pásma. Směrodatná odchylka udává variabilitu krajiny, což souvisí se spektrální heterogenitou vegetace.



Obrázek č. 16: Index NDVI vytvořený v GEE, hodnoty nabývají -1 až 1, čím blíže je hodnota 1, tím více vegetace se v místě nachází, tento index byl vstupním pásmem pro klasifikace.

Pro vstupní snímky jsem zvolila podmínku 3 % oblačnosti. Vytvořila jsem mediánový kompozit nejméně oblačných snímků pro květen a červenec zvlášť. Z kompozitů vytvořených ze všech pásem z července a z května jsem podle spektrálních charakteristik z obrázků č. 7 a č. 8 vybrala pásma, která výrazně ovlivňují vybrané biotopy a jsou v nich dostatečně odlišené. Pro květen jsou to pásma B6 ve vlnové délce 740 nm, B7 v 783 nm a B8A v 865 nm. V červenci se vybranými pásmy staly B8 ve vlnové délce 842 a B11 v 1610 nm. Klasifikace byla zpracována z RGB pásem, 5 vybraných pásem z května a července, indexu NDVI a jeho směrodatné odchylky.

Rozdělení dat na trénovací a validační je předpokladem pro klasifikaci a následné zhodnocení přesnosti pomocí chybové matice. Trénovací a validační data vychází z bodů ve vybraných biotopech a v jejich okolí. Zvolila jsem klasifikaci založenou na pixelech, počet náhodných bodů pro okolí bylo přibližně 300, zatímco v každém známém biotopu jejich počet byl přesně 115, to znamená, že pro všechny biotopy bylo náhodných bodů 575. Data na trénovací a validační jsem rozdělila přímo v GEE. Poměr trénovacích a validačních byl 20:80. *Random seed*, tedy náhodný výběr hodnot, měl možnou škálu od 0 do 999, zvolila jsem hodnotu 100. To znamená, že se náhodný výběr každých bodů do klasifikace vytvoří v požadovaném množství a tím i klasifikace, ve výsledku je klasifikace průměrem hodnot.

Klasifikace Random Forest vycházela z trénovacích dat, kde je nutné zadat sloupec, ve kterém se nachází informace s názvem biotopu a také vychází z vybraných pásem. Dle Oshiro et al. (2012) by optimální počet stromů měl být na škále od 64 do 128, tudíž jsem počet náhodných stromů zvolila na 100. Klasifikace CART vycházela ze stejných trénovacích dat a pásem jako RF, dalšími vstupními parametry mohou být maximální nebo minimální počet listových uzlů na každý strom, já jsem však další parametry nezadávala a nechala jsem tyto hodnoty bez limitů. Obě klasifikace byly přidány do mapy jako vrstvy.

Poslední položkou byl export výsledků, data z GEE lze ukládat na Google drive, na Cloud Storage nebo přímo jako asset v GEE. Zvolila jsem uložení do Google drive, odkud lze rastr snadno stáhnout.

Přesnosti klasifikace vytvořené skriptem (Nomura & Mitchard, 2018) v GEE mi nepřipadaly věrohodné. Již z vizuálního porovnání výsledných klasifikovaných vrstev s VMB bylo jasné, že hodnoty přesnosti vykazované skriptem nabývají podezřele vysokých hodnot. I při výrazných změnách během klasifikace byl výsledek chybové matice GEE vždy nad 90%. Proto jsem přesnosti klasifikace ověřila v Excelu za využití vzorců pro charakteristiky přesnosti a mé pochybnosti se potvrdily. Veškeré zhodnocení přesnosti jsem tedy zpracovala v Excelu.

4.4 Zhodnocení přesnosti

Ke zhodnocení přesnosti výsledků jsem použila chybovou matici, která vyjadřuje vztah mezi výsledky klasifikace a realitou. Porovnávala jsem výsledky vytvořené algoritmy RF a CART ve dvou možnostech. V první jsem použila validační body a v druhé jsem prováděla validaci z celých polygonů.

Jelikož jsem při klasifikaci v GEE zvolila výběr trénovacích a validačních dat automaticky, v ArcMapu jsem vytvořila nová validační data pro provedení validace z pixelů. Ze všech dat jsem za pomoci python skriptu vybrala 80 %. Z těchto dat jsem vytvořila *Extract values to point*, to znamená, že se do atributové tabulky validačních bodů přepsaly výsledky z klasifikace. Z těchto bodů jsem vytvořila tabulku *Frequency*, která udává hodnoty kolikrát se daný jev vyskytl v každé třídě. Posledním krokem byla funkce *Pivot table*. Tato funkce vytvoří tabulku, která již má podobu chybové matice. Je čtvercová a udává jak reálné hodnoty, tak hodnoty získané klasifikací. Z vizualizované chybové matice je patrné, jak odpovídají reálné třídy klasifikovaným a naopak. Čísla mimo hlavní diagonálu znamenají záměnu prvků, tedy

prvky, které byly špatně klasifikovány. Pro zjištění procentuálních výsledků jsem použila charakteristiky přesnosti.

Při validaci výsledků za využití polygonů jsem vytvořila z rastru klasifikace polygon. Přepsala jsem hodnoty s informacemi o biotopech do nového vektorového shapefile za pomoci *Intersect* z polygonové vrstvy a z vrstvy klasifikace. Pro vytvoření chybové matice jsem postupovala stejně jako u zjišťování přesnosti z bodů a tj. vytvoření tabulky *Frequency* a z ní následně *Pivot table*, která má již výslednou podobu chybové matice.

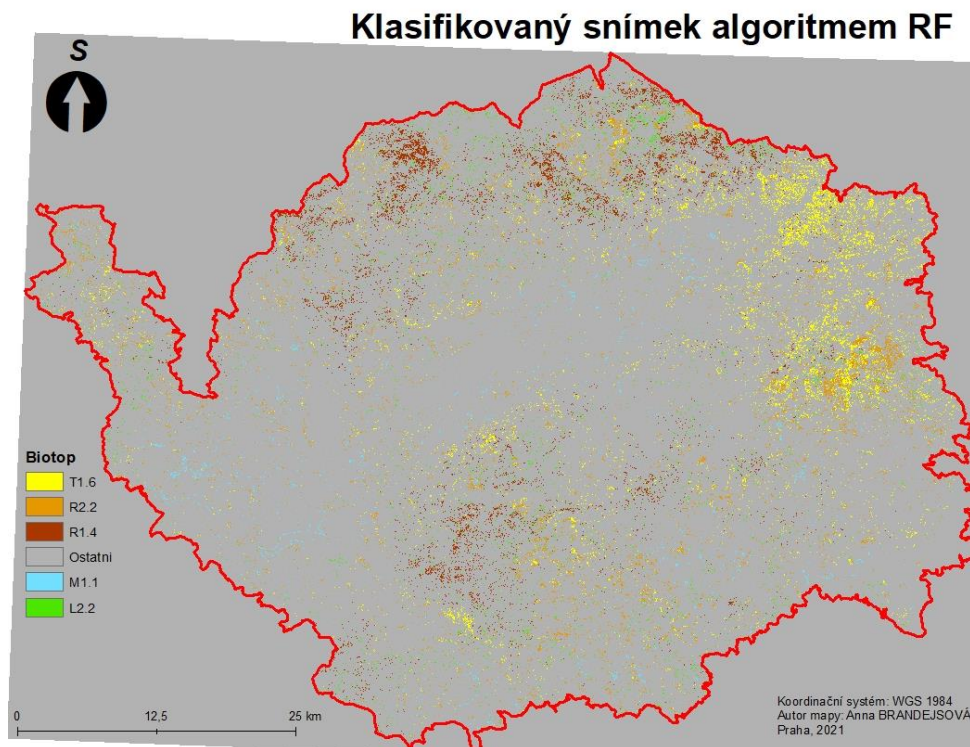
Pro zjištění přesnosti klasifikace z chybové matice se používá více charakteristik, pro tuto práci jsem zvolila následující tři metody. První byla *user's accuracy*, která vyjadřuje pravděpodobnost, jakou pixel ve skutečnosti odpovídá klasifikované třídě. Vypočítá se z čísla na hlavní diagonále a součtem klasifikovaných pixelů. Druhou je charakteristika přesnosti *producer's accuracy*, která na rozdíl od té předchozí vychází z pravděpodobnosti, že reálná třída je klasifikovaná správně. Výsledek vychází z poměru mezi počtem pixelů správně klasifikovaných, tedy hodnotou na hlavní diagonále a celkovým počtem pixelů v třídách pro realitu. Poslední přesností je celková přesnost klasifikace, která vychází ze součtu všech výsledků na hlavní diagonále v poměru s celkovým počtem pixelů. Přesnosti jsem počítala z pixelů a polygonů pro klasifikaci algoritmu RF a CART. Výsledkem jsou tedy čtyři tabulky přesností.

5. Výsledky

Výsledkem této bakalářské práce je klasifikace biotopů v Karlovarském kraji vytvořená dvěma klasifikátory, a to RF a CART. V porovnání použitých klasifikátorů RF dosáhl spolehlivějších výsledků nežli algoritmus CART. V obou případech se podařilo rozpoznat ostatní povrch od vybraných biotopů, ale i přesto docházelo k časté záměně pixelů jak mezi biotopy vzájemně, tak i mezi biotopy a ostatním povrchem.

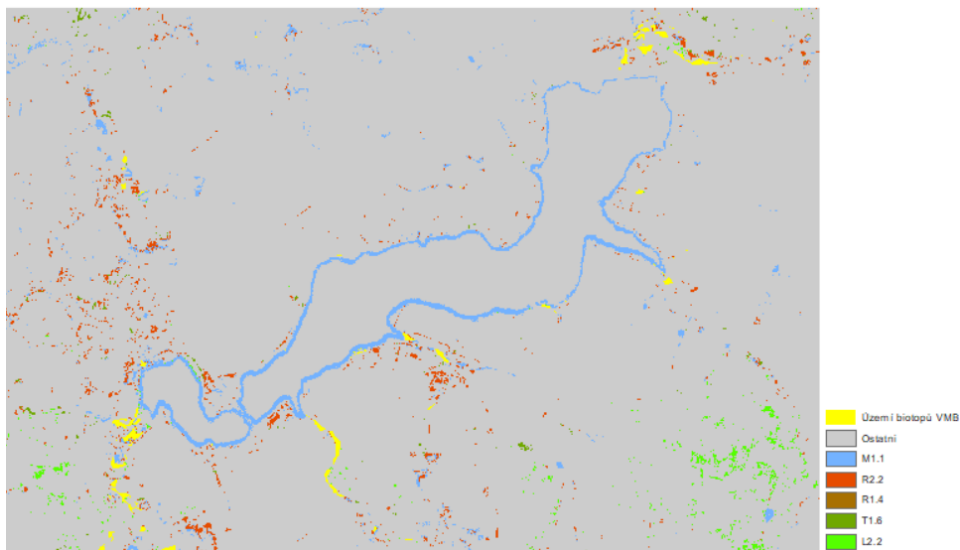
5.1 Random Forest

Na obrázku č. 17 jsou vizualizované výsledky vytvořené klasifikací algoritmem RF. Biotopy jsou detekovány na přibližných lokalitách reálného výskytu biotopů dle VMB. Z těchto výsledků vychází, že mapované biotopy jsou velice heterogenní.



Obrázek č. 17: Výsledek klasifikace území za využití algoritmu RF, v legendě jsou všechny zobrazené biotopy včetně okolního povrchu, konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostříc M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčkové louky T1.5, Vlhká tužebníková lada T1.6.

Pro přesnější porovnání reálného výskytu biotopů a jejich detekci jsem využila více detailních výřezů území, na obrázku č. 18 je příklad výřezu v okolí vodní nádrže Jesenice. Zde je vidět, že okolí kolem vodní plochy bylo klasifikováno jako rákosiny eutrofních vod. Na tomto místě se přitom biotop M1.1 ve skutečnosti se sice nachází, při klasifikaci byl ale výrazně nadhodnocen.



Obrázek č. 18: Výřez území, demonstrace zobrazení biotopů klasifikovaných algoritmem RF v porovnání s reálným výskytem biotopů z VMB. Konkrétně se jedná o Mokřadní vrby K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnovec R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.

Accuracy z polygonů		SKUTEČNOST					celkem	UA
RF	L2.2	M1.1	R1.4	R2.2	T1.6			
KLASIFIKÁTOR	L2.2	5390	18	112	59	356	5935	91%
	M1.1	853	144	1	111	352	1461	10%
	R1.4	986	0	197	9	83	1275	15%
	R2.2	1703	166	13	240	1174	3296	7%
	T1.6	2600	11	53	88	1178	3930	30%
	O1	2530	139	314	220	941	4144	
	celkem	11532	339	376	507	3143	20041	
PA	47%	42%	52%	47%	37%			
OCA	36%							

Tabulka č. 3: Chybová matice a výpočty přesností z polygonů pro klasifikátor RF, v polygonech. Konkrétně se jedná o Mokřadní vrby K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnovec R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.

Accuracy z bodů		SKUTEČNOST					celkem	UA	
RF	O1	M1.1	R2.2	R1.4	T1.6	L2.2			
KLASIFIKÁTOR	O1	239	42	41	59	48	42	471	51%
	M1.1	0	23	2	0	1	0	26	88%
	R2.2	0	6	27	0	3	1	37	73%
	R1.4	0	0	0	16	0	0	16	100%
	T1.6	0	0	5	4	23	0	32	72%
	L2.2	0	0	0	2	2	21	25	84%
	celkem	239	71	75	81	77	64	607	
PA	100%	32%	36%	20%	30%	33%			
OCA	57%								

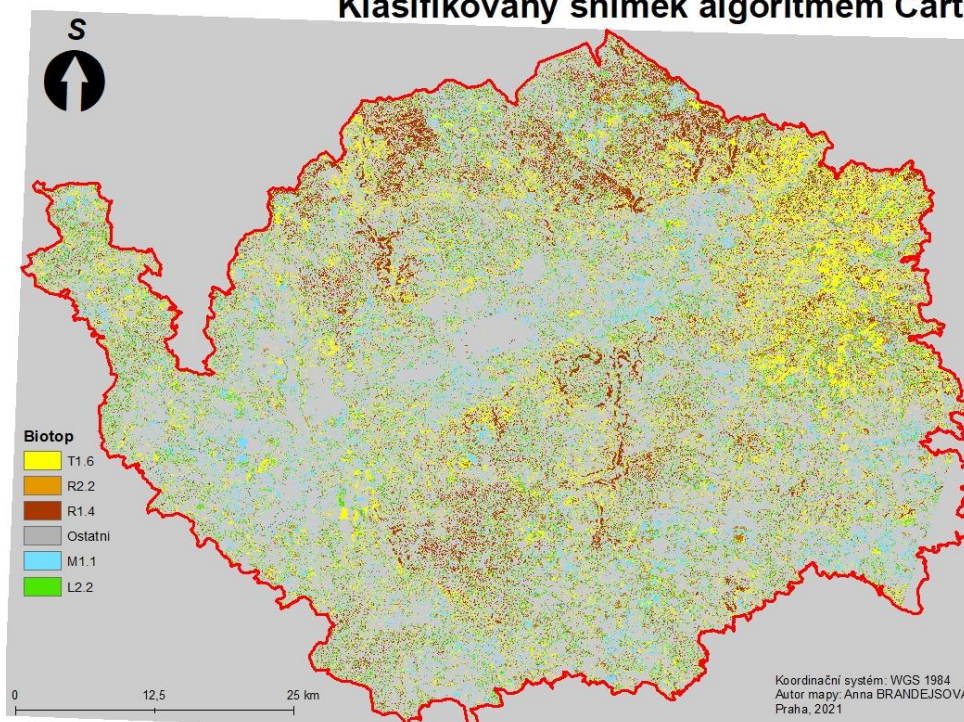
Tabulka č. 4: Chybová matice a výpočet přesností z bodů pro klasifikátor RF. Konkrétně se jedná o Mokřadní vrby K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnovec R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.

Z chybové matice je patrné, jak se biotopy vzájemně zaměňují. Toto je zapříčiněno jejich spektrálními vlastnostmi. Záměna mezi biotopy lesní prameniště bez tvorby pěnovců (R1.4) a rákosiny eutrofních stojatých vod (M1.1) a mezi lesní prameniště bez tvorby pěnovců (R1.4) a Nevápnitá mechová slatiniště (R2.2), byla zanedbatelná, jelikož mají odlišné spektrální vlastnosti. Záměna biotopů byla výrazná především se třídou Údolní jasanovo-olšové luhy (L2.2). Pro zjištění procentuálních výsledků jsem použila charakteristiku přesnosti uživatele (dále jako UA) a zpracovatele (dále jako PA). Celková přesnost klasifikace (dále jako OCA) Random Forest vyšla 57 %. Tento výsledek je ovlivněn tím, že je přesnost prováděna z bodů, která nemusí být vždy úplně směrodatná především z důvodu nerovnoměrnosti bodů. Proto jsem pro ověření přesnosti vytvořila také přesnost vycházející z polygonů, které vyšla OCA 36 %, ovšem přesnější jsou hodnoty přesnosti pro jednotlivé biotopy, které z výpočtů UA a PA dosahují vyšších hodnot. Tyto hodnoty a záměny lze vyčíst z chybové matice v tabulce č. 3 a č. 4.

5.2 Classification and Regression Trees

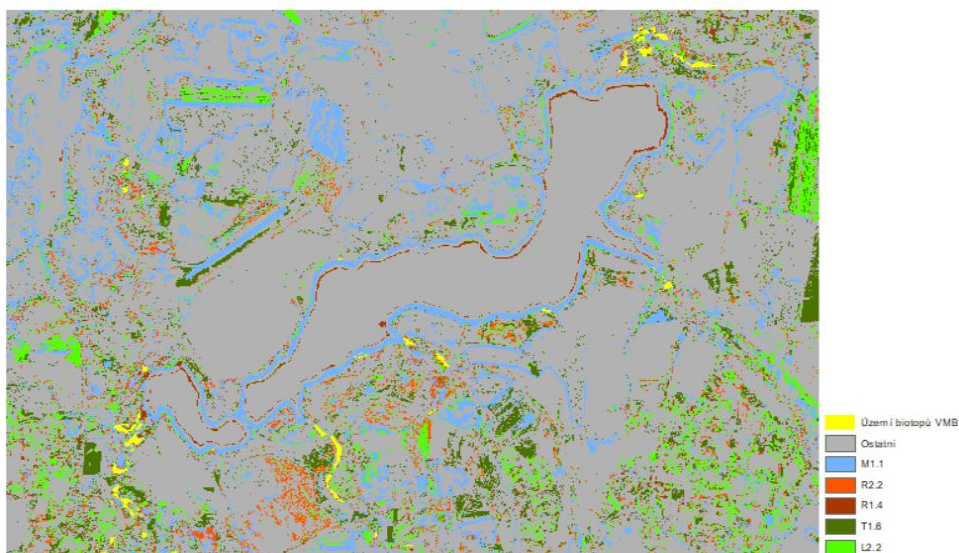
Klasifikace za použití algoritmu CART je již na první pohled rozdílná oproti klasifikaci předešlé. Biotopy byly detekovány na většině území, především v jeho středu, kde se ve skutečnosti žádné chráněné biotopy nenacházejí. Z obrázku č. 19 je patrná větší četnost jednotlivých biotopů nežli na obrázku č. 17 vytvořeného algoritmem RF.

Klasifikovaný snímek algoritmem Cart



Obrázek č. 19: Výsledek klasifikace území za využití algoritmu CART, v legendě jsou všechny zobrazené biotopy včetně okolního povrchu. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnvců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.

V porovnání klasifikátorů dochází v klasifikaci za využití algoritmu CART k výraznému nadhodnocování všech biotopů. Pro vizualizaci porovnání klasifikátorů jsem použila výřez ze stejného místa (viz obr. č. 20). V okolí nádrže se nachází ve velké míře všechny biotopy, což neodpovídá realitě. Oproti předešlé klasifikaci se na vodní ploše objevuje biotop lesní prameniště bez tvorby pěnvců (R1.4). Tato klasifikace vycházela ze stejných předpokladů jako RF, ale dochází zde k větší vzájemné záměně mezi biotopy a méně přesné detekci okolí od biotopů.



Obrázek č. 20: Výřez území, kde dochází k velkému nadhodnocování biotopů klasifikátorem CART v porovnání s výskytem biotopů z VMB. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostříc M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.

Procentuální výsledky přesnosti za využití charakteristiky UA vychází, že se s vysokou přesností podařilo zmapovat třídu údolní jasanovo-olšové luhy (L2.2), zato nejhůře dopadla třída nevápnitá mechová slatiniště (R2.2) s pouhými 7 % přesnosti. V tabulce č. 5 a č. 6 jsou chybové matice pro algoritmus CART vypočítané z polygonů a bodů. Opět se potvrdilo, že přesnost za využití polygonů dosahuje nižších hodnot. Tabulka č. 6 uvádí výsledky pro zhodnocení přesnosti z validačních pixelů. Z této chybové matice je patrné, že nejvíce chybových pixelů spadá do klasifikovaného okolí.

Accuracy z polygonů			SKUTEČNOST				celkem	UA
	CART	L2.2	M1.1	R1.4	R2.2	T1.6		
KLASIFIKÁTOR	L2.2	18048	198	383	355	2336	21320	85%
	M1.1	7759	261	74	401	2081	10576	2%
	R1.4	8646	226	175	347	2095	11489	2%
	R2.2	11648	66	435	103	1072	13324	1%
	T1.6	6461	228	92	436	1973	9190	21%
	O1	16814	156	350	408	2676	20404	
	celkem	69376	1135	1509	2050	12233	86303	
	PA	26%	23%	12%	5%	16%		
	OCA	24%						

Tabulka č. 5: Chybová matice a výpočet přesností z polygonů pro algoritmus CART. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostříc M1.7, Lesní prameniště bez tvorby pěnovců R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.

Accuracy z bodů		SKUTEČNOST							
KLASIFIKÁTOR	CART	O1	M1.1	R2.2	R1.4	T1.6	L2.2	celkem	UA
	O1	228	29	27	43	33	29	389	95%
	M1.1	6	33	10	1	6	3	59	56%
	R2.2	0	2	27	1	4	1	35	77%
	R1.4	3	0	0	18	4	1	26	69%
	T1.6	2	3	3	8	21	1	38	55%
	L2.2	0	4	8	10	9	29	60	48%
	celkem	239	71	75	81	77	64	607	
PA	41%	46%	36%	22%	27%	45%			
OCA	59%								

Tabulka č. 6: Chybová matice a výpočet přesností z bodů pro klasifikátor CART. Konkrétně se jedná o Mokřadní vrbiny K1, Údolní jasanovo-olšové luhy L2.2, Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnovec R1.4, Nevápnitá mechová slatiniště R2.2, Vlhké pcháčové louky T1.5, Vlhká tužebníková lada T1.6.

5.3 Porovnání výsledků RF a CART

Z celkového porovnání přesností pro obě klasifikace plyne, že výsledky dosahují podobných hodnot. K záměně biotopů v obou klasifikacích nejčastěji dochází v kategoriích), toto je především ovlivněno jejich nízkou četností, proto u zhodnocení dochází k nízkým hodnotám jednotlivých přesností. Zatímco ve třídě L2.2 dochází k častým záměnám, přesnost výsledků tím není výrazně ovlivněna, jelikož zastoupení třídy má vysokou četnost.

Při porovnání výsledků pouze pro podrobné třídy z kategorie mokřadů (M) a rašelinišť (R) pro zjištění, zda je výsledek ovlivněn podobností biotopů v dané kategorii a zda je detekce přesnější klasifikaci podrobných tříd nežli celých kategorií. Je z výsledků pro jednotlivé biotopy patrná vyšší přesnost pro jednotlivé biotopy (viz tabulka č. 7 a č. 8). Výsledek je tedy ovlivněn i biotopy z kategorie lesy a sekundární trávníky a vřesoviště.

Accuracy z polygonů		SKUTEČNOST				
KLASIFIKÁTOR	RF	M1.1	R1.4	R2.2	celkem	UA
	M1.1	144	1	111	256	56%
	R1.4	0	197	9	206	96%
	R2.2	166	13	240	419	57%
	O1	139	314	220	419	
	celkem	310	211	360	1554	
	PA	46%	93%	67%		
	OCA	37.39%				

Tabulka č. 7: Tabulka přesností pouze pro třídy: Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnovec R1.4, Nevápnitá mechová slatiniště R2.2, tedy pouze kategorie mokřady (M) a rašeliniště (R) za použití RF.

Accuracy z polygonů		SKUTEČNOST				
	CART	M1.1	R1.4	R2.2	celkem	UA
KLASIFIKÁTOR	M1.1	261	74	401	736	35%
	R1.4	226	175	347	748	23%
	R2.2	66	435	103	604	17%
	O1	156	350	408	914	
	celkem	709	1034	1259	3002	
	PA	37%	17%	8%		
	OCA	18%				

Tabulka č. 8: Tabulka přesností pouze pro třídy: Rákosiny eutrofních stojatých vod M1.1, Vegetace vysokých ostřic M1.7, Lesní prameniště bez tvorby pěnoveců R1.4, Nevápnitá mechová slatiniště R2.2, tedy pouze kategorie mokřady (M) a rašeliniště (R) za použití CART

Při porovnání přesností klasifikace z polygonů pro hrubé kategorie biotopů za předpokladu, že hrubé kategorie mohou být snadněji detekovatelné především z důvodu, že jsou méně podrobné. Jedná se o kategorie lesy (L), Mokřady a pobřežní vegetace (M), Prameniště a rašeliniště (R), Sekundární trávníky a vřesoviště (T). Z tabulky č. 9 a č. 10 lze pozorovat záměny hlavních kategorií. Velmi častá záměna dochází u kategorie lesy se všemi kategoriemi. Další častá záměna je mezi kategorií mokřadů (M) a rašelinišť (R), která výrazně ovlivňuje výsledné přesnosti daných kategorií. Z těchto porovnání vychází, že klasifikace z podrobných tříd vyšla pro jednotlivé biotopy přesněji.

Accuracy z polygonů		SKUTEČNOST					
	RF	L	M	R	T	celkem	UA
KLASIFIKÁTOR	L	5390	18	171	356	5935	91%
	M	853	144	223	352	1572	9%
	R	2689	1152	1338	2160	7339	18%
	T	2600	11	200	1178	3989	30%
	O	2530	139	332	941	3942	24%
	celkem	14062	1464	2264	4987	22777	
	PA	38%	10%	59%	24%		
OCA	36%						

Tabulka 9: Tabulka přesností pro celé kategorie z klasifikace RF pro celé polygony, lesy (L), Mokřady a pobřežní vegetace (M), Prameniště a rašeliniště (R), Sekundární trávníky a vřesoviště (T) a ostatní plochy (O).

Accuracy z polygonů		SKUTEČNOST					
	CART	L	M	R	T	celkem	UA
KLASIFIKÁTOR	L	18048	198	738	2336	21320	85%
	M	7759	261	784	2081	10885	2%
	R	20294	8712	9132	9718	47856	19%
	T	6461	228	819	1973	9481	21%
	O	16814	156	791	2676	20437	
	celkem	69376	9555	12264	18784	109979	
	PA	26%	3%	74%	11%		
OCA	27%						

Tabulka 10: Tabulka přesností pro celé kategorie z klasifikace CART pro celé polygony, lesy (L), Mokřady a pobřežní vegetace (M), Prameniště a rašeliniště (R), Sekundární trávníky a vřesoviště (T) a ostatní plochy (O).

6. Diskuse

Tato bakalářská práce je jedním z prvních výzkumů, který se zabývá klasifikací mokřadních biotopů v podrobnosti definované Katalogem biotopů. Účelem práce bylo zjistit, jakými metodami DPZ lze co nejpodrobněji detekovat třídy biotopů dle výskytu z VMB. Za využití klasifikátorů RF a CART jsem docílila detekci přibližného výskytu zvolených biotopů. Nejvíce problematickou částí výzkumu bylo odlišit biotopy od okolí. Docházelo zde k nadhodnocování biotopů a tím i k záměně biotopů s okolím.

Z tohoto výzkumu plyne, že klasifikace mokřadních biotopů s sebou nese mnoho nepřesností především proto, že se jedná o homogenní biotopy. Celou klasifikaci biotopů jsem prováděla z nejpodrobnějších tříd dané kategorie, ale ověřovala jsem i možnost zvolení méně podrobné kategorie pro vstupní data, jelikož by se zvětšila plocha biotopů, protože z výzkumu od Bhatnagar et al. (2020) plyne závislost spolehlivějších výsledků na velikosti společenstev.

Při zjišťování, zda lze lépe rozlišit podrobné třídy nebo celé kategorie biotopů z Katalogu, jsem předpokládala, že hlavní kategorie mohou být detekovány lépe, jelikož nezasahují do tematické podrobnosti. Z analýzy vychází, že detekce podrobnějších kategorií vychází s vyšší přesností, a i když jsou podrobnější biotopy příbuzné z fytoecologického pohledu vzájemně se spektrálně liší. U klasifikace mokřadních biotopů z podrobných tříd docházelo k záměně mokřadů (M) a rašelinišť (R) s kategoriemi lesy (L) a traviny (T), proto jsem zjišťovala přesnost klasifikace pouze pro mokřady (M) a rašeliniště (R) a ve výsledku byla přesnost pro samostatné biotopy vyšší.

Pracovala jsem s vybranými mediánovými kompozity pásem, které spektrálně odlišují mokřadní biotopy z července a května. Výslednou kombinací pro detekci mokřadních biotopů bylo 5 pásem VRE z celého května a července, jelikož jsem pracovala s vegetací, která se ve vegetačním období výrazně mění ve spektrálních vlastnostech. Pro vytváření kompozitů je možné zpracovávat snímky z více časových úseků ve vegetačním období. Výběr vhodného období pro klasifikaci vegetace souvisí s hypotézou spektrální variability (Palmer, 2020).

Výzkumů zaměřených na detekci chráněných území, které spadají pod soustavu Natura 2000, za využití DPZ, již bylo zpracováno pro celou Evropu mnoho. V této práci jsem vycházela z více výzkumů. Múcher, (2013) využil pro detekci vřesovišť ze soustavy Natura 2000 100 trénovacích bodů a hyperspektrální data. Z této studie vychází, že při detekci vřesovišť z hyperspektrálních dat lze traviny odlišit ve velké přesnosti a do

budoucná lze pomoci eliminovat častou záměnu travin. Výzkum od Regos (2018) vycházel z porovnávání 10 klasifikačních metod, přičemž RF nabýval velmi přesných výsledků.

Výsledné nadhodnocování biotopů mohlo být ovlivněno skutečností, že se mokřadní biotopy vyskytují na malých územích, při práci s družicovými snímky může docházet k záměně hraničních pixelů. Nejvhodnější pásma S2 (Red Edge a Near Infrared) pro klasifikaci vegetace, mají 20 m rozlišení. Jelikož polygony detekovaných biotopů jsou převážně malé, použila jsem *buffer* pouze -10 metrů, s větším *bufferem* by se eliminoval počet trénovacích dat pro ořez hraničních pixelů.

Nejvhodnější rozdělení náhodných bodů bylo pro každou třídu biotopu stejný počet. Ve studii (Thanh Noi & Kappas, 2017), zaměřené na řízenou klasifikaci založené na pixelech z náhodných bodů, bylo v každé třídě vytvořeno 135 bodů. U prováděné klasifikace Random forest v této studii vyšly výsledky pro mapování krajinného pokryvu s vysokou spolehlivostí, přibližně 90 %. Proto jsem pro trénovací zvolila stejný počet bodů (115 random bodů) pro každou třídu. V úvahu připadala možnost zvolit počet náhodných bodů podle zastoupení jednotlivých biotopů, ale třídy měly podobný počet polygonů, a proto jsem tuto možnost nezvolila. Z náhodných bodů jsem vytvářela trénovací a testovací body, zvolila jsem poměr 20:80 a proces výběru se 100 opakoval, toto vycházelo z nastavení *random seed* v GEE.

Zpracování klasifikace jsem vytvářela v GEE na základě skriptu z vědeckého článku od Nomura (2018), který je zaměřený na klasifikaci biotopů na menších územích ze snímků S2. Tento skript jsem pro mé potřeby pozměnila. Pro vstupní snímky jsem zvolila mediánový kompozit všech snímků s oblačností menší než 3 %, namísto původního jednoho snímku s nejnižší oblačností. Vypočítaný Confusion matrix ze skriptu vycházel ve nereálně vysokých hodnotách a při ověření přesnosti klasifikace v jiném programu jsem zjistila, že hodnoty neodpovídají realitě, proto jsem všechny chybové matice vytvářela v Excelu za pomoci vzorců pro charakteristiky přesnosti. Z tohoto článku jsem použila také skript pro vytvoření spektrální charakteristiky pro detekované biotopy z jednotlivých snímků S2, z výsledných hodnot jsem zpracovala příslušné grafy ve statistickém programu R.

Detekci biotopů jsem zpracovávala v GEE především, protože je to nejvhodnější alternativa v domácím prostředí. Zároveň toto byla vhodná příležitost ukázat možnost detekce biotopů v prostředí mimo licenční programy, je to vhodné především pro budoucímu použití, jelikož by tato práce mohla být přínosná nejen pro AOPK ČR, ale i pro zájemce o tuto problematiku, kteří nemají přístup například licenčním programům (ENVI).

Z výsledků této práce plyne, že mapování mokřadních biotopů s sebou nese mnoho nepřesností. Při klasifikaci mokřadů se projevilo velké nadhodnocování. Výsledky klasifikace mohou být zapříčiněny heterogenitou mapovaných biotopů, nebo podkladovými daty z VMB, které jsou mapovány v terénu a může zde dojít k subjektivnímu ovlivnění vstupních dat a jistých nepřesností. Pracovala jsem pouze s biotopy, které byly aktualizované po r. 2015, nadhodnocování s tím může souviset, jelikož je šance, že byl biotop správně klasifikován, ale zatím nebyl aktualizován. Ale při použití všech zaznamenaných biotopů by nebylo zaručené přesné geografické určení, jelikož na začátku mapování v roce 2000 nebyly běžně rozšířené GPS.

Mokřad (M) byl klasifikován na místech, která jsou vlhká, ale neřadí se mezi chráněné území (například se jednalo o břehy okolo vodních ploch). Velmi častým jevem klasifikace biotopů bylo nadhodnocování jejich výskytu nejen v okolí vody. Mapování mokřadů je v tomto ohledu složité, jelikož z družicového snímku je těžké rozeznat, zda se jedná o mokřad nebo pouze o mokrou půdu či břeh. Častým jevem byla také častá záměna za křoviny či les, jelikož existuje více druhů mokřadních biotopů a některé mají vysoký pokryv.

Největší problém klasifikace byla záměna docházející mezi biotopy a okolím. Snažila jsem se přijít na alternativu, kde by šlo zanedbat okolí, proto jsem využila pro klasifikaci lokalitu pouze z území biotopů. V této klasifikaci došlo k velké záměně biotopů, v každém polygonu bylo obsaženo více tříd, takže bylo jisté, že je lepší zvolit celé velké území nežli lokality výskytu.

Jelikož se na vybrané lokalitě nachází velký rozsah nadmořských výšek, pro zpřesnění klasifikace by bylo vhodné zahrnout digitální model terénu (DTM). Protože jsem tento výzkum zpracovávala v prostředí GEE, kde doposud DTM není zpracované pro Evropu, do přesnosti klasifikace jsem rozdíl nadmořských výšek nezahrnula. Je pravděpodobné, že výsledky byly ovlivněny spektrálními změnami, které vznikaly změnou nadmořské výšky. Další možnost zlepšení výsledků při detekci mokřadních biotopů je přidání topografického indexu vlhkosti (dále jako TWI). Výsledná klasifikace byla ovlivněna vlhkostí půdy a TWI je vhodný pro mapování hydrologických procesů.

Kromě použité řízené klasifikace založené na pixelech RF a CART, lze k mapování mokřadních biotopů využít klasifikaci s objektovou analýzou obrazu (OBIA). U mapování biotopů se často kombinuje OBIA s řízenou klasifikací založenou na pixelech, toto bylo aplikováno ve výzkumu (Moffett & Gorelick, 2013). Další zaměření na detekci vegetace je od

Förster (2006), které porovnává objektovou a pixelovou klasifikaci. Z čehož OBIA vychází s lepšími hodnotami. Proto je možné tento výzkum zkombinovat s metodou OBIA, která vychází z tvaru, velikosti a textury klasifikovaných objektů, a to by mohlo zlepšit přesnost detekce homogenních biotopů.

Výsledky klasifikace pro detekci mokřadních biotopů z této bakalářské práce prozatím nenahradí mapování VMB v terénu, jelikož nebylo možné spolehlivě odlišit jednotlivé třídy biotopů. Tento výzkum by sloužil k mírnému urychlení procesu, jelikož se podařilo detekovat přibližná místa výskytu biotopů, jako podkladová vrstva nebo jako indikátor náhlé změny výskytu biotopů.

7. Závěr

Cílem této práce bylo ověřit, jaké jsou možnosti pro detekci mokřadních biotopů a zajistit co nejpodrobnější mapování. Analýzy byly vytvářeny pro Karlovarský kraj a mokřadní biotopy z VMB, které byly aktualizovány po roce 2015. Veškeré klasifikace byly vytvářeny z vegetačního období roku 2018 ze snímků z S2. Na základě předešlých výzkumů jsem zvolila algoritmus RF, který při mapování mokřadů vykazoval výsledky s vysokou přesností. Pro porovnání více algoritmů řízené klasifikace jsem využila klasifikátor CART and RF. V obou případech docházelo k vysokému nadhodnocování biotopů, ale RF pro jednotlivé biotopy dosáhl vyšších hodnot v charakteristikách přesnosti.

Podařilo se detekovat přibližnou lokalitu výskytu mokřadních biotopů z VMB. Biotopy byly odlišeny od okolní plochy, ale jednotlivé třídy se nepodařilo přesně identifikovat. Celkové výsledky jsou negativně ovlivněny výskytem biotopů na malých plochách.

Při interpretaci výsledků jsem vycházela z klasifikování podrobných tříd a klasifikování celých kategorií vycházejících z Katalogu biotopů. Detekce podrobných tříd vykazala přesnější výsledky pro jednotlivé biotopy, z toho vyhází, že se jedná o homogenní biotopy, a i podobné třídy se liší svými spektrálními vlastnostmi.

V závěru tato práce potvrdila předpoklad, že mapování mokřadních biotopů je velmi náročný úkol, podařilo se za využití podkladu biotopů z VMB detekovat přibližnou lokalitu výskytu. I když mezi podtřídami nedocházelo k vysoké záměně, biotopy nebyly za pomoci klasifikací detekovány s vysokou přesností, jelikož dochází k výraznému nadhodnocování.

8. Seznam použitých zdrojů

- Bhatnagar, S., Gill, L., Regan, S., Naughton, O., Johnston, P., Waldren, S., & Ghosh, B. (2020). MAPPING VEGETATION COMMUNITIES INSIDE WETLANDS USING SENTINEL-2 IMAGERY IN IRELAND. *International Journal of Applied Earth Observation and Geoinformation*, 88, 102083. <https://doi.org/10.1016/j.jag.2020.102083>
- Bourgeau-Chavez, L. L., Endres, S., Powell, R., Battaglia, M. J., Benscoter, B., Turetsky, M., Kasischke, E. S., & Banda, E. (2017). Mapping boreal peatland ecosystem types from multitemporal radar and optical satellite imagery. *Canadian Journal of Forest Research*, 47(4). <https://doi.org/10.1139/cjfr-2016-0192>
- Briand, L. C., & Wüst, J. (2002). Empirical studies of quality models in object-oriented systems. In *Advances in Computers* (Vol. 56, Issue C, pp. 97–166). Academic Press Inc. [https://doi.org/10.1016/S0065-2458\(02\)80005-5](https://doi.org/10.1016/S0065-2458(02)80005-5)
- Chytrý M., Kučera T., Kočí M., Grulich V., & Lustyk P. (2010). Katalog biotopů České republiky, Milan Chytrý, Tomáš Kučera, Martin Kočí, Vít Grulich, Pavel Lustyk (editoři). Druhé vydání. Praha: AOPK ČR, 2010. 445 S. 21.
- Co je Natura 2000? - Ministerstvo životního prostředí. (2021), Retrieved March 5, 2021, from https://www.mzp.cz/cz/natura_definice
- Corbane, C., Lang, S., Pipkins, K., Alleaume, S., Deshayes, M., García Millán, V. E., Strasser, T., vanden Borre, J., Toon, S., & Michael, F. (2015). Remote sensing for mapping natural habitats and their conservation status - New opportunities and challenges. *International Journal of Applied Earth Observation and Geoinformation*, 37. <https://doi.org/10.1016/j.jag.2014.11.005>
- De’Ath, G., & Fabricius, K. E. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*, 81(11). [https://doi.org/10.1890/0012-9658\(2000\)081\[3178:CARTAP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2)
- Dobrovolný, CSc. (1998). Dálkový průzkum Země, Digitální zpracování obrazu. Přírodovědecká fakulta Masarykovy univerzity, Katedra geografie.
- Fawagreh, K., Gaber, M. M., & Elyan, E. (2014). Random forests: From early developments to recent advancements. *Systems Science and Control Engineering*, 2(1). <https://doi.org/10.1080/21642583.2014.956265>
- Förster, M., Frick, A., Walentowski, H., & Kleinschmit, B. (2008). Approaches to utilising QuickBird data for the monitoring of NATURA 2000 habitats. *Community Ecology*, 9(2). <https://doi.org/10.1556/ComEc.9.2008.2.4>
- Gavish, Y., O’Connell, J., Marsh, C. J., Tarantino, C., Blonda, P., Tomaselli, V., & Kunin, W. E. (2018). Comparing the performance of flat and hierarchical Habitat/Land-Cover classification models in a NATURA 2000 site. *ISPRS Journal of Photogrammetry and Remote Sensing*, 136, 1–12. <https://doi.org/10.1016/j.isprsjprs.2017.12.002>
- GIS Geography.(2021) *OBIA - Object-Based Image Analysis (GEOBIA) - GIS Geography*. Retrieved March 7, 2021, from <https://gisgeography.com/obia-object-based-image-analysis-geobia/>

- Gudmann, A., Csikós, N., Szilassi, P., & Mucsi, L. (2020). Improvement in satellite image-based land cover classification with landscape metrics. *Remote Sensing*, *12*(21). <https://doi.org/10.3390/rs12213580>
- Jiao, L., Sun, W., Yang, G., Ren, G., & Liu, Y. (2019). A hierarchical classification framework of satellite multispectral/hyperspectral images for mapping coastal wetlands. *Remote Sensing*, *11*(19). <https://doi.org/10.3390/rs11192238>
- Killeen, P. G., Mwenifumbo, C. J., & Ford, K. L. (2015). Tools and Techniques: Radiometric Methods. In *Treatise on Geophysics: Second Edition* (Vol. 11, pp. 447–524). Elsevier Inc. <https://doi.org/10.1016/B978-0-444-53802-4.00209-8>
- Kovyazin, V. F., Romanchikov, A. Y., Lan Anh, D. T., Lan Anh, D. T., & Hung, D. V. (2020). Monitoring of Forest Land Use/Cover Change in Cat Tien National Park, Dong Nai Province, Vietnam Using Remote Sensing and GIS techniques. *IOP Conference Series: Materials Science and Engineering*, *817*(1). <https://doi.org/10.1088/1757-899X/817/1/012018>
- Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). Foundations of data imbalance and solutions for a data democracy. In *Data Democracy: At the Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering* (pp. 83–106). Elsevier. <https://doi.org/10.1016/B978-0-12-818366-3.00005-8>
- Levin, N., Shmida, A., Levanoni, O., Tamari, H., & Kark, S. (2007). Predicting mountain plant richness and rarity from space using satellite-derived vegetation indices. *Diversity and Distributions*, *13*(6). <https://doi.org/10.1111/j.1472-4642.2007.00372.x>
- Mahdavi, S., Salehi, B., Granger, J., Amani, M., Brisco, B., & Huang, W. (2018). Remote sensing for wetland classification: a comprehensive review. In *GIScience and Remote Sensing* (Vol. 55, Issue 5, pp. 623–658). Taylor and Francis Inc. <https://doi.org/10.1080/15481603.2017.1419602>
- Moffett, K. B., & Gorelick, S. M. (2013). Distinguishing wetland vegetation and channel features with object-based image segmentation. *International Journal of Remote Sensing*, *34*(4). <https://doi.org/10.1080/01431161.2012.718463>
- Mücher, C. A., Kooistra, L., Vermeulen, M., Borre, J. vanden, Haest, B., & Haveman, R. (2013). Quantifying structure of Natura 2000 heathland habitats using spectral mixture analysis and segmentation techniques on hyperspectral imagery. *Ecological Indicators*, *33*, 71–81. <https://doi.org/10.1016/j.ecolind.2012.09.013>
- Natura 2000 - Environment - European Commission*. (2021). Retrieved March 5, 2021, from https://ec.europa.eu/environment/nature/natura2000/index_en.htm
- Nomura, K., & Mitchard, E. T. A. (2018). More than meets the eye: Using Sentinel-2 to map small plantations in complex forest landscapes. *Remote Sensing*, *10*(11). <https://doi.org/10.3390/rs10111693>
- Oshiro, T. M., Perez, P. S., & Baranauskas, J. A. (2012). How many trees in a random forest? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7376 LNAI*. https://doi.org/10.1007/978-3-642-31537-4_13

- Ozesmi, S. L., & Bauer, M. E. (2002). Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, 10(5). <https://doi.org/10.1023/A:1020908432489>
- Palmer, T. N. (2002). Predicting uncertainty in numerical weather forecasts. *International Geophysics*, 83(C), 3–13. [https://doi.org/10.1016/S0074-6142\(02\)80152-8](https://doi.org/10.1016/S0074-6142(02)80152-8)
- Peter. (2021, March). *CartoScience — CartoScience Lab*. <https://www.cartoscience.com/bgp>
- Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 data for land cover/use mapping: A review. In *Remote Sensing* (Vol. 12, Issue 14). <https://doi.org/10.3390/rs12142291>
- Portál AOPK ČR. (2021). Retrieved March 5, 2021, from https://portal.nature.cz/publik_syst/ctihtmlpage.php?what=1035
- Qiu, S., He, B., Yin, C., & Liao, Z. (2017). Assessments of Sentinel-2 vegetation red-edge spectral bands for improving land cover classification. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(2W7), 871–874. <https://doi.org/10.5194/isprs-archives-XLII-2-W7-871-2017>
- Rapinel, S., Fabre, E., Dufour, S., Arvor, D., Mony, C., & Hubert-Moy, L. (2019). Mapping potential, existing and efficient wetlands using free remote sensing data. *Journal of Environmental Management*, 247, 829–839. <https://doi.org/10.1016/j.jenvman.2019.06.098>
- Rocchini, D., Marcantonio, M., da Re, D., Chirici, G., Galluzzi, M., Lenoir, J., Ricotta, C., Torresani, M., & Ziv, G. (2019). Time-lapsing biodiversity: An open source method for measuring diversity changes by remote sensing. *Remote Sensing of Environment*, 231, 111192. <https://doi.org/10.1016/j.rse.2019.05.011>
- Ruiz, L. F. C., Simioni, J. P. D., Belloli, T. F., Barros Fernandes, P. C., & Guasselli, L. A. (2021). Object-based classification of vegetation species in a subtropical wetland using Sentinel-1 and Sentinel-2A images. *Science of Remote Sensing*, 3, 100017. <https://doi.org/10.1016/j.srs.2021.100017>
- Rumora, L., Miler, M., & Medak, D. (2020). Impact of various atmospheric corrections on sentinel-2 land cover classification accuracy using machine learning classifiers. *ISPRS International Journal of Geo-Information*, 9(4). <https://doi.org/10.3390/ijgi9040277>
- Russi, D., ten Brink, P., Farmer, A., Bandura, T., Coates, D., Dorster, J., Kumar, R., & Davidson, N. (2012). The Economics of Ecosystems and Biodiversity for Water and Wetlands: A final Consultation Draft. *IEEP, London and Brussels; Ramsar Secretaria*, 119.
- Sánchez-Espinosa, A., & Schröder, C. (2019). Land use and land cover mapping in wetlands one step closer to the ground: Sentinel-2 versus landsat 8. *Journal of Environmental Management*, 247, 484–498. <https://doi.org/10.1016/j.jenvman.2019.06.084>
- Schmidt, J., Fassnacht, F. E., Neff, C., Lausch, A., Kleinschmit, B., Förster, M., & Schmidlein, S. (2017). Adapting a Natura 2000 field guideline for a remote

sensing-based assessment of heathland conservation status. *International Journal of Applied Earth Observation and Geoinformation*, 60, 61–71.
<https://doi.org/10.1016/j.jag.2017.04.005>

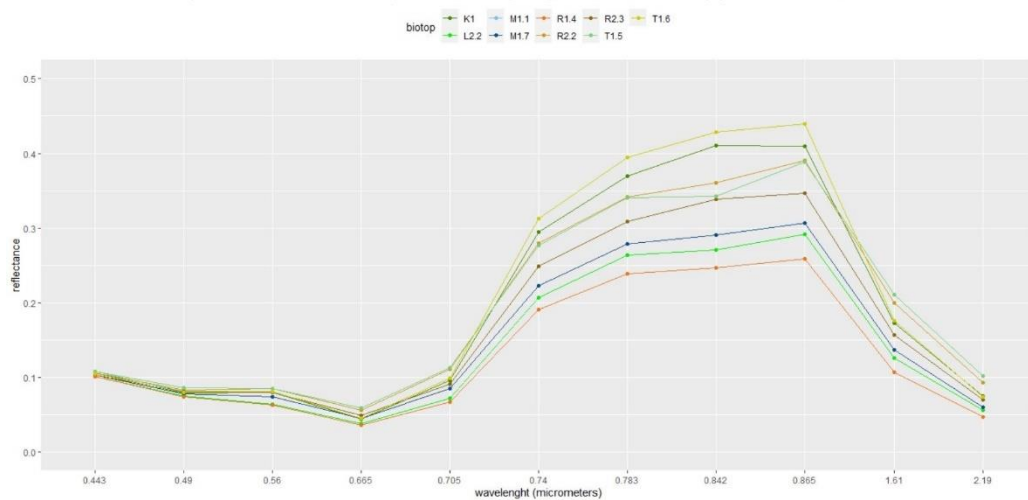
Schmidtlein, S., & Fassnacht, F. E. (2017). The spectral variability hypothesis does not hold across landscapes. *Remote Sensing of Environment*, 192.
<https://doi.org/10.1016/j.rse.2017.01.036>

Slagter, B., Tsendbazar, N.-E., Vollrath, A., & Reiche, J. (2020). Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: A case study in the St. Lucia wetlands, South Africa. *International Journal of Applied Earth Observation and Geoinformation*, 86, 102009.
<https://doi.org/10.1016/j.jag.2019.102009>

Thanh Noi, P., & Kappas, M. (2017). Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors (Basel, Switzerland)*, 18(1).
<https://doi.org/10.3390/s18010018>

User Guides - Sentinel-2 MSI - Overview - Sentinel Online - Sentinel. (2021). Retrieved March 5, 2021, from <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/overview>

Spektrální vlastnosti pásem S2 pro jednotlivé biotopy v červenci 2018



Příloha č. 2: Skript na klasifikaci biotopu za pomoci algoritmů RF a CART v GEE

```
var roi = ee.FeatureCollection("users/annabrandejsova3/hranice/hranice"), //cely kraj
    L2_2 = ee.FeatureCollection("users/annabrandejsova3/body/p_L2_2"),
    M1_1 = ee.FeatureCollection("users/annabrandejsova3/body/p_M1"),
    R1_4 = ee.FeatureCollection("users/annabrandejsova3/body/p_R1_4"),
    R2_2 = ee.FeatureCollection("users/annabrandejsova3/body/p_R2_2"),
    T1_6 = ee.FeatureCollection("users/annabrandejsova3/body/p_T1_6"),
    other = ee.FeatureCollection("users/annabrandejsova3/ostatni_mix");

/*//////////////////////////////////////
Zdroje:
- tutorial https://developers.google.com/earth-engine/tutorials/community/classify-maizeland-ng
- paper "More than meets the eye: using Sentinel-2 to map small plantations
in complex forest landscapes" by Keiko Nomura and Edward TA Mitchard.
//////////////////////////////////////

var areaName = 'Karlovarsky kraj';
var year = '2018';
var startDate = '2018-05-01';
var endDate = '2018-05-30';
var randomSeed = 100;
var startDate2 = '2018-07-01';
var endDate2 = '2018-07-30';

// Vrstvy zobrazené v konzoli
print('Area' + " " + areaName + ' ' + year);
var checkbox1 = ui.Checkbox('S2 composit', true);
var checkbox2 = ui.Checkbox('NDVI', false);
var checkbox3 = ui.Checkbox('Texture', false);
var checkbox4 = ui.Checkbox('ClassifiedRF', true);
var checkbox5 = ui.Checkbox('ClassifiedCart', true);

checkbox1.onChange(function(checked) {Map.layers().get(0).setShown(checked);});
checkbox2.onChange(function(checked) {Map.layers().get(1).setShown(checked);});
checkbox3.onChange(function(checked) {Map.layers().get(2).setShown(checked);});
checkbox4.onChange(function(checked) {Map.layers().get(3).setShown(checked);});
checkbox5.onChange(function(checked) {Map.layers().get(4).setShown(checked);});

print(checkbox5);
print(checkbox4);
print(checkbox3);
print(checkbox2);
print(checkbox1);

//// Sentinel-2 data ////
var s2 = ee.ImageCollection('COPERNICUS/S2_SR')
    .filterDate(startDate, endDate)
    .filterBounds(roi);
var s2_07 = ee.ImageCollection('COPERNICUS/S2_SR')
    .filterDate(startDate2, endDate2)
    .filterBounds(roi);

// Zobrazení vhodných období pro data
var list = ee.List(s2.aggregate_array("system:time_start")).map(function(d) { return ee.Date(d)});
print('S2 images of the area during the study period',list);

var list = ee.List(s2_07.aggregate_array("system:time_start")).map(function(d) { return ee.Date(d)});
print('S2 images of the area during the study period',list);

// údaj o oblačnosti ve snímcích v daném období (květen a červenec)
var getCloudScores = function(img){
    var value = ee.Image(img).get('CLOUDY_PIXEL_PERCENTAGE');
    return ee.Feature(null, {'score': value});};

var s2clouds = s2.map(getCloudScores);
print ('cloud score', ui.Chart.feature.byFeature(s2clouds));
var s2clouds = s2_07.map(getCloudScores);
print ('cloud score', ui.Chart.feature.byFeature(s2clouds));

// nejméně olačný kompozit, výběr a přejmenování bandů, resample
var s2image = ee.ImageCollection('COPERNICUS/S2')
    .filterDate(startDate, endDate)
    .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 2))
```

```

.sort('CLOUDY_PIXEL_PERCENTAGE')
.filterBounds(roi)
.map(function(img){
  var t = img.select(['B1','B2','B3','B4','B5','B6','B7','B8','B8A','B9','B10','B11','B12']).divide(10000);//Rescale to 0-1
  var out = t.copyProperties(img).copyProperties(img,['system:time_start']);
  return out;});
var s2image_07 = ee.ImageCollection('COPERNICUS/S2')
  .filterDate(startDate2, endDate2)
  .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 2))
  .sort('CLOUDY_PIXEL_PERCENTAGE')
  .filterBounds(roi)
  .map(function(img){
    var t = img.select(['B1','B2','B3','B4','B5','B6','B7','B8','B8A','B9','B10','B11','B12']).divide(10000);//Rescale to 0-1
    var out = t.copyProperties(img).copyProperties(img,['system:time_start']);
    return out;});

// Kompozit z mediánu nejméně oblačných snímků namísto původního jednoho nejméně oblačného snímku
var s2composit =
s2image.reduce(ee.Reducer.median()).select(['B1_median','B2_median','B3_median','B4_median','B5_median','B6_median','B7_median','
B8_median','B8A_median','B9_median','B10_median','B11_median','B12_median'],['aerosol','blue','green','red',
'red1','red2','red3','nir','red4','h2o','cirrus','swir1','swir2']);
var s2ROI = s2composit.clip(roi);
var s2composit_07 =
s2image_07.reduce(ee.Reducer.median()).select(['B1_median','B2_median','B3_median','B4_median','B5_median','B6_median','B7_med
ian','B8_median','B8A_median','B9_median','B10_median','B11_median','B12_median'],['aerosol2','blue_2','green_2','red_2',
'red12','red22','red32','nir2','red42','h2o2','cirrus2','swir12','swir22']);
var s2ROI_07= s2composit_07.clip(roi);

var vizParams = {bands: ['red', 'green', 'blue'], min: 0, max: 0.3};
var vizParams2 = {bands: ['red_2', 'green_2', 'blue_2'], min: 0, max: 0.3};
Map.addLayer(s2ROI, vizParams, 'S2 composit');
Map.addLayer(s2ROI_07, vizParams2, 'cervenec', false);

// NDVI
var red = s2ROI.select('red');
var nir = s2ROI.select('nir');
var ndvi = nir.subtract(red).divide(nir.add(red)).rename('NDVI');

var ndviParams = {min: -1, max: 1, palette: ['blue', 'white', 'green']};
Map.addLayer(ndvi, ndviParams, 'NDVI', false);

// Směrodatná odchylka z NDVI jako vrstva texture
var texture = ndvi.reduceNeighborhood({
  reducer: ee.Reducer.stdDev(),
  kernel: ee.Kernel.square(5),});

Map.addLayer(texture, {min: 0, max: 0.25},'Texture', false);

// Přidání vhodných pásem, a sice dva Sentinel snímky, ndvi a jeho odchylka
var s2final = ee.Image(s2ROI).addBands(ndvi).addBands(texture).addBands(s2ROI_07);

//všechny vhodné bandy, udělá klasifikaci z vybraných bandů obou kompozitů plus NDVI a jeho směrodatná odchylka a svazitost slope ;
var bands = ['red','blue','green','red2','red3','red4','NDVI','NDVI_stdDev','nir2','swir12'];
var s2classification = s2final.select(bands);

//// Vytvoření trénovacích a verifikačních dat ////
// Vytvoření náhodného sloupce do vzorku
var n = randomSeed;
var rnd_other = other.randomColumn('random', n);
var rnd_M1_1 = M1_1.randomColumn('random', n);
var rnd_R2_2 = R2_2.randomColumn('random', n);
var rnd_R1_4 = R1_4.randomColumn('random', n);
var rnd_T1_6 = T1_6.randomColumn('random', n);
var rnd_L2_2 = L2_2.randomColumn('random', n);

// 20:80 pro trénovací a validační body
var split1 = 0.2; // trénovací data se rozdělí na 20%a zbylých 80% je validačních

var trainingSample = rnd_other.merge(rnd_M1_1.filter(ee.Filter.lt('random', split1)))
  .merge(rnd_R2_2.filter(ee.Filter.lt('random', split1)))
  .merge(rnd_R1_4.filter(ee.Filter.lt('random', split1)))
  .merge(rnd_T1_6.filter(ee.Filter.lt('random', split1)))
  .merge(rnd_L2_2.filter(ee.Filter.lt('random', split1)));
Map.addLayer(trainingSample,false);

```



```

var testingSample = rnd_other.merge(rnd_M1_1.filter(ee.Filter.gte('random', split1)))
    .merge(rnd_R2_2.filter(ee.Filter.lt('random', split1)))
    .merge(rnd_R1_4.filter(ee.Filter.lt('random', split1)))
    .merge(rnd_T1_6.filter(ee.Filter.lt('random', split1)))
    .merge(rnd_L2_2.filter(ee.Filter.lt('random', split1)));

// Remap sloupce class, který nese informaci o třídě na číselnou hodnotu
var trainingSampleClass = trainingSample.remap(['O1','M1.1','R2.2','R1.4','T1.6','L2.2'], [0, 1, 2, 3, 4, 5], 'class');
var testingSampleClass = testingSample.remap(['O1','M1.1','R2.2','R1.4','T1.6','L2.2'], [0, 1, 2, 3, 4, 5], 'class');

print (trainingSample)
print (trainingSampleClass)
print (testingSample)
print (testingSampleClass)

// Trénovací hodnoty
var training = s2classification.sampleRegions({
  collection: trainingSampleClass,
  properties: ['class'],
  scale: 10, });

///// klasifikace and verifikace /////
// Random Forest
var trainedRF = ee.Classifier.smileRandomForest({numberOfTrees: 100,})
    .train(training, 'class', bands);

// Cart
var trainedCart = ee.Classifier.smileCart().train(training, 'class', bands);

// Provedení klasifikace vstupních dat
var classifiedRF = s2classification.classify(trainedRF);
var classifiedCart = s2classification.classify(trainedCart);

// paleta barev pro biotopy
var palette = [ 'C0C0C0', // 'O1'
    '63B8FF', // 'M1.1'
    'EE7621', // 'R2.2'
    '8B6914', // 'R1.4'
    'EEC900', // 'T1.6'
    '00EE00', // 'L2.2' ];

// Zobrazení klasifikované mapy
Map.addLayer(classifiedRF, {min: 0, max: 5, palette: palette}, 'ClassifiedRF');
Map.addLayer(classifiedCart, {min: 0, max: 5, palette: palette}, 'ClassifiedCart');
Map.centerObject(roi, 10);

// Legenda
var labels = ['O1','M1.1','R2.2','R1.4','T1.6','L2.2'];
var add_legend = function(title, lbl, pal) {
  var legend = ui.Panel({style: {position: 'bottom-left'}});
  legend.add(ui.Label({value: title, style: {fontWeight: 'bold', fontSize: '18px', margin: '0 0 4px 0', padding: '0px' }}));
  for (var x = 0; x < lbl.length; x++){
    entry = [ ui.Label({style: {color: pal[x], border: '1px solid black', margin: '0 0 4px 0', value: '■'}),
      ui.Label({ value: labels[x], style: { margin: '0 0 4px 4px' }}) ];
    legend.add(ui.Panel(entry, ui.Panel.Layout.Flow('horizontal')));
  } Map.add(legend); };

add_legend('Legend', labels, palette);

// export RF
Export.image.toDrive({
  image: classifiedRF,
  region: roi,
  description: 'RF_classification',
  scale: 10
});
// export Cart
Export.image.toDrive({
  image: classifiedCart,
  region: roi,
  description: 'Cart_classification',
  scale: 10})

```