

## Oponentní posudek diplomové práce

Název práce: **Změna land cover v zaledněných horských oblastech**

Autorka: **Bc. Anna Brázdová**

Vedoucí práce: Ing. Lukáš Brodský, Ph.D.

Předložená diplomová práce řeší problematiku změny krajinného pokryvu ve specifickém prostředí horských ledovců s využitím časové řady dat Landsat a metod strojového učení. Jedná se o aktuální téma zejména v kontextu změny klimatu a důsledků této změny na vývoj kryosféry a hydrologických poměrů. Určování rozsahu horských ledovců a jejich změn z dat DPZ je téma v literatuře dlouhodobě řešené. Zadání a cíle práce obecně řadím mezi náročnější, zejména v oblasti rozpoznávání ledovce pokrytého sutí a odlišení reálných a dočasných či falešných změn.

Práce má standardní strukturu. Na první dojem zaujme vysokou stylistickou úrovní, logickou návazností a provázaností textu v rámci kapitol. Úvod práce stručně nastiňuje význam zvoleného tématu, podrobně seznamuje s cíli práce a přehledně shrnuje obsah dílčích kapitol. Hlavní očekávaný přínos je spatřován v nalezení „techniky detekce změn“ (str. 12), které nejsou, dle zjištění autorky, pro zaledněné oblasti dostatečně prozkoumány. Kapitola Úvod do problematiky je zpracována systematicky od popisu typů ledovců, jejich vlastností a změn, přes metody jejich zkoumání pomocí DPZ, až ke strojovému učení, kdy je největší pozornost věnována třem vybraným metodám, a to KNN, random forest (RF) a gradient boosted decision trees (GBDF). Jak v části věnované DPZ, tak strojovému učení jsou vyčleněny sekce zaměřené na asi nejkomplicovanější oblast řešené problematiky, což je rozpoznání ledovce, jehož povrch je zakrytý sutí. Samostatná sekce je věnována rešerši metod detekce změn. Úvod do problematiky je nejrozsáhlejší částí diplomové práce. Opírá se o množství (odhadem více než 50) zahraničních zdrojů a svědčí o tom, že si autorka během studia výborně osvojila práci s literaturou a je schopna zpracovat kvalitní odbornou rešerši. Z obsahového a formálního hlediska kapitola obsahuje minimum nepřesností. Např. na str. 20 jsou promíchány pojmy vztahující se k laserové altimetrii a laserovému skenování; formulace „Rozhodovací strom je sestaven z trénovací množiny, ...“ (str. 25) není nejvhodnější.

Podobně zdařilou je i první sekce kapitoly 3 Data a metodika. Seznamuje s fyzickogeografickými a socioekonomickými poměry zájmového území, tj. ostrova Disko. Postrádám geografické vymezení a charakteristiku užšího zájmového území, o němž se čtenář dozvídá až v sekci Předzpracování dat (chybně označené jako 2.3.1 místo 3.3.1). Sekce 3.2 popisuje použité datové zdroje a z nich odvozené příznaky pro klasifikaci. Jejich obecný popis je v pořádku, ale některé detaily by bylo vhodné doplnit přímo a nenechávat je na čtenáři či je vynechat vůbec. Např. by bylo vhodné v textu nebo jakou součást popisu tabulky 1 uvést, že byly použity produkty úrovně L2. Je zmíněna spektrální odrazivost ve viditelném, blízkém a středním infračerveném pásmu. Povrchové teplotě jako zdroji dat pozornost věnována prakticky není (např. jednotky, rozlišení). Co je míněno „velmi dobrou prostorovou kvalitou dat Landsat“ (str. 45, první odstavec)? Autorka uvádí spektrální index označený jako „nový poměr pásem“ (NBR). Jedná se o nově vytvořený index? Pokud ano, měl by mu být v metodice věnován větší prostor a na příkladech ukázat, že je skutečně vhodný pro odlišení ledovců pokrytých sutí. Pokud se nejedná o autorčinu invenci, pak by měla být uvedena citace. Zkratka NBR je zavádějící s ohledem na její běžné využití pro „normalized burn ratio“.

Metodika řešení směřovaná k vyhodnocení změny krajinného pokryvu zájmové oblasti o velikosti přibližně 50 km x 50 km mezi lety 1985, 2000 a 2021 je rozdělena do tří částí. Předzpracování dat se věnuje základním operacím převedení všech použitých datových vrstev, tj. scén z družic Landsat, povrchové rychlosti ledovců výškového modelu ArcticDEM do téhož zobrazení, rozlišení a jejich koregistraci s využitím funkcí softwaru QGIS. Z jakých období byly počítány průměrné rychlosti pohybu ledovců pro rok 1985 v místech chybějících dat? Uvedení, že byly počítány s využitím funkce *r.patch* není dostatečné (str. 49 a 50). Jakým způsobem bylo docíleno a ověřeno, že veškerá data byla koregistrována tak, aby se s přesností na jeden pixel překrývala? To je, jak autorka správně uvádí, pro řešení zásadní krok a obecně netriviální úkol, který by neměl být shrnut do jedné věty. Druhá část metodiky je věnována výběru a realizaci postupů detekce změn krajinného pokryvu. Prvním z postupů je klasifikace v dílčích časových horizontech a jejich následné porovnání s využitím matice změny. Autorka nejprve volí klasifikační legendu, která odpovídá zájmovému

území. Čtenáři by v této fázi pomohl obrázek či alespoň odkaz na obrázky zájmového území uvedené dále. Následně je s ohledem na řízené metody klasifikace pozornost věnována výběru trénovacích množin. Nad jakými daty byly trénovací množiny vybírány? Pouze nad daty Landsat nebo i dalšími referenčními daty uvedenými v sekci 3.1? Kladem je, že se autorka věnuje otázce vyváženosti datové sady. Uvažovala autorka kromě vizualizací (viz obrázek 4) i o jiných způsobech zjištění separability zvolených tříd vdaném příznakovém prostoru? Pro vlastní klasifikaci byly zvoleny 3 výše uvedené metody (KNN, RF, GBDT) u nichž byla zjišťována přesnost, doba trénování a predikce s defaultním nastavením. Na základě získaných výsledků uvedených v tabulce přílohy 1 byly dále laděny pouze hyperparametry algoritmu RF (počet stromů a proměnných). Dále byla testována metoda detekce změn z rozdílu obrazových dat. Autorka použila k této analýze zelené pásmo, index sněhu a povrchovou teplotu. Chybí zdůvodnění, proč právě tyto příznaky. Jako další byla použita multitemporální klasifikace nad původním a transformovaným příznakovým prostorem souboru obsahujícího datové vrstvy dvou časových horizontů. Ačkoliv se jedná o obecně známý postup, měla by být uveden alespoň jeden zdroj odkazující na metodu hlavních komponent. U všech klasifikačních postupů je vyhodnocena přesnost klasifikace na základě testovací množiny manuálně sbíraných bodů. Validita změny je navíc testována vyhodnocením kódu charakterizujícím posloupnost změny. Jak bylo nakládáno s pixely, kde změna byla vyhodnocena jako nepřípustná?

Kapitola výsledky přehledně seznamuje s dosaženými přesnostmi klasifikace, hodnocením důležitosti jednotlivých příznaků pro klasifikaci. Změny krajinného pokryvu jsou rozděleny na „skutečné/pravé“ a „nepravé“, kvantifikovány a vizualizovány ve formě mapových výstupů. V diskusi se autorka kriticky vrací k dílčím výsledkům a použitým metodám a v rámci možností je porovnává s literaturou. V této části práce se opět projevil silná stránka autorky jednak systematicky analyzovat každý z dílčích kroků, ale i vyvodit obecnější závěry a doporučení pro další práci. Závěr stručně shrnuje obsah práce a dosažené výsledky.

Jak z výše uvedeného vyplývá, předložená diplomová práce je kvalitně zpracovaná. Autorka maximálně využila znalosti z oblasti DPZ získané v průběhu studia, dokázala je aplikovat k vyřešení zadaného úkolu a implementovat v podobě skriptů v jazyce Python. Jak vyplývá z otázek uvedených výše, některým dílčím krokům zpracování mohla být věnována větší pozornost. Cíle práce byly splněny s využitím známých a v literatuře publikovaných postupů. Práce splňuje formální náležitosti, text obsahuje minimum chyb a překlepů.

Diplomovou práci Bc. Anny Brázdové **doporučuji** k obhajobě a navrhuji hodnocení stupněm **výborně až velmi dobře**.

#### Doplňující otázky

1. Na str. 12 se odkazujete na práci Chughtai, Abbasi, Karas (2021): „Zároveň většina studií zkoumajících změny krajinného pokryvu využívá postklasifikační porovnání, které však zahrnuje mnoho kroků, a může trpět falešně pozitivními změnami způsobenými chybnou klasifikací.“ Jak je to s Vámi navrženým postupem? Hrozí u něj stejné úskalí? Viz také konstatování v diskusi ohledně posklasifikačního porovnání, které bylo nakonec vyhodnoceno jako jeden z nejlepších postupů: „Přestože průměrná přesnost jednotlivých klasifikací byla větší než 95 %, průměrná přesnost této techniky pro detekci změn byla pouze 84,45 %. Tuto nižší přesnost však lze přiřadit i tomu, že technika byla validována na jiných datech, než na kterých byla natrénována.“ (str. 69)
2. Na str. 29 uvádíte: „V současné době neexistuje žádná zcela jedinečná klasifikační metodika, která by umožňovala konzistentně mapovat zaledněné terény. Co dobře funguje v jednom regionu nebo při použití určité datové sady, nemusí mít stejný úspěch v jiné části kryosféry či při použití jiného zdroje dat.“ Jak Vaše práce přispěla k řešení nastíněného problému?