

UNIVERZITA KARLOVA
Přírodovědecká fakulta

Katedra aplikované geoinformatiky a kartografie

Studijní program: Geografie
Studijní obor: Sociální geografie a geoinformatika



ANALÝZA ÚZEMÍ PRO REINTRODUKCI POPULACÍ KONĚ PŘEVALSKÉHO V MONGOLSKU POMOCÍ METOD DPZ

ANALYSIS OF AREA FOR REINTRODUCTION OF EQUUS PRZEWALSKII POPULATIONS IN
MONGOLIA USING REMOTE SENSING METHODS

Bakalářská práce

Petr Janoš

Vedoucí bakalářské práce: Mgr. Josef Laštovička, Ph.D.
Konzultant bakalářské práce: Mgr. Jan Svoboda

Praha 2021

Vysoká škola: Univerzita Karlova
Katedra: Katedra aplikované geoinformatiky a kartografie

Fakulta: Přírodovědecká
Školní rok: 2020/2021

Zadání bakalářské práce

pro Petra Janoše
obor Sociální geografie a geoinformatika

Název tématu: Analýza území pro reintrodukcii populací koně Převalského v Mongolsku pomocí metod DPZ

Zásady pro vypracování

Práce se bude zabývat možnostmi zjištění kvalitativních informací o krajině a jejích změn v čase pomocí metod DPZ. Konkrétně bude zaměřena na monitoring krajiny pomocí vegetačních / vodních / půdních indexů (např. NDVI, NDMI, RGI, NDSI), pozorování množství srážek a teplot (např. průměrné, medián i absolutní hodnoty – maxima a minima v jednotlivých měsících), případně dalších vhodně vybraných ukazatelů, které budou analyzovány pomocí pokročilých metod časových řad (Time Series) na volně dostupných datech (např. Landsat a ECMWF / Copernicus Climate Change Service). Ze získaných mnoholetých časových řad budou analyzovány trendy pomocí pokročilých výzkumných metod (např. pomocí klouzavých průměrů nebo specializovaných funkcí či algoritmů, jako je CUSUM nebo Bfast) a dále budou prokázány závislosti mezi jednotlivými indexy, srážkami a teplotami (např. bude užit Shapiro-Wilcoxonův test nebo ANOVA test). Analýza časových řad bude doplněna i o klasifikaci land cover zkoumaných území včetně hodnocení přesnosti s využitím volně dostupných dat Landsat nebo Sentinel-2. Mongolsko představuje velmi dynamickou krajinu, snahou bude provést komplexní zhodnocení, zda je krajina vhodná pro reintrodukcii na základě získaných informací z jednotlivých ukazatelů časových řad a klasifikací. Porovnány tak budou dvě lokality. První lokalita s již reintrodukovanou populací koně Převalského (SPA Great Gobi B) a druhé pozorované území bude představovat potenciální nová lokalita pro navrácení tohoto druhu do volné přírody (vybraná část Eastern steppe, východní Mongolsko). Práce bude konzultována s odborníky ze Zoo Praha (Mgr. Barbora Dobiášová a kolegové), bude využito znalosti řešeného území zaměstnanci Zoo. V rámci řešeršší části bude řešena problematika vhodných dat a pokročilých Time Series metod DPZ pro komplexní pozorování fyzicko-geografických charakteristik u rozsáhlejších ploch. Taktéž bude v teoretické rozpravě zpracován přehled o dostupnosti, využitelnosti a vlastnostech volně dostupných dat a jejich DPZ analýz pro porovnání v čase. Součástí empirické části práce bude provedeno detailní testování a aplikace Time Series a klasifikačních metod DPZ. Ke zpracování bude využito cloud based platformy Google Earth Engine (GEE), případně lze užít pro i klasický desktopový geoinformační a DPZ software (např. ArcGIS Pro, ENVI či obdobný komerční program). Pro statistické zpracování a zhodnocení dat bude možné využít i prostředí R či Microsoft Excel. Výstupy práce mohou být prezentovány pomocí on-line mapové aplikace (např. AGOL Story Map), případně pomocí vlastních webových stránek.

Cíle práce:

1. Představit možnosti analýzy trendů a detekce trendů
2. Představení a otestování volně dostupných dat a dostupných technologií pro tvorbu časových řad
3. Pomocí vybraných metod časových řad detailně analyzovat území
4. Pomocí vybraného klasifikátoru a dat Sentinel-2 nebo Landsat provést klasifikaci obou pozorovaných území
5. Provést detailní srovnání obou pozorovaných území

6. Prokázání závislosti / podobnosti zkoumaných území pomocí statistických metod
7. Stanovení vhodnosti vybrané potenciální lokality pro reintrodukcii koně Převalského na základě získaných výstupů a na základě konzultací s odborníky se Zoo Praha (např. v souvislosti s výskyty sucha, dzud a dalších extrémních jevů)
8. Vytvoření metodiky/algoritmu ve volně dostupné výzkumné platformě GEE, případně pomocí jiné cloud based platformy či případně pomocí desktopového softwaru

Rozsah průvodní zprávy: 40 až 50 stran

Rozsah grafických prací a příloh: dle potřeby

Seznam odborné literatury:

Kalpoma, K.A.; Kudoh, J.I. Application of snow coverage information extraction technique using NOAA satellite images in mongolian region to reduce the damage of dzud. Int. Geosci. Remote Sens. Symp. 2006, 3943–3946.

Bao, G.; Qin, Z.; Bao, Y.; Zhou, Y.; Li, W.; Sanjjav, A. NDVI-based long-term vegetation dynamics and its response to climatic change in the Mongolian plateau. Remote Sens. 2014, 6, 8337–8358.

Neteler, M. Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data. Remote Sens. 2010, 2, 333–351.

Dorjsuren, M.; Liou, Y.A.; Cheng, C.H. Time series MODIS and in situ data analysis for Mongolia drought. Remote Sens. 2016, 8, 1–16.

Noi Phan, T.; Kuch, V.; Lehnert, L.W. Land cover classification using google earth engine and random forest classifier-the role of image composition. Remote Sens. 2020, 12.

Vedoucí bakalářské práce: Mgr. Josef Laštovička, Ph.D.

Konzultanti: Mgr. Barbora Dobiášová a Mgr. Jan Svoboda

Datum zadání bakalářské práce: 12. ledna 2021

Termín odevzdání bakalářské práce: květen / červenec 2021

Platnost tohoto zadání je po dobu jednoho akademického roku.

Podpis vedoucího bakalářské práce (Mgr. Josef Laštovička, Ph.D.):

Podpis vedoucí katedry (doc. RNDr. Lucie Kupková, Ph.D.):

V Praze dne 12. ledna 2021.

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze dne 28. července 2021

Petr Janoš

Poděkování

V první řadě bych chtěl poděkovat svému vedoucímu bakalářské práce panu doktorovi Josefovi Laštovičkovi. V těžkých chvílích dokázal povzbudit a kdykoli bylo potřeba, byl mi nápomocen. Spolupráce s ním byla velmi příjemná a přínosná, a kdyby byla v budoucnu možnost další spolupráce, jistě bych neodmítnul. Dále bych chtěl poděkovat své přítelkyni Aničce za podporu nejen při psaní této práce, ale za celou dobu studia, bez které bych to jistojistě nezvládl. To platí i o mé milované rodině. V neposlední řadě musím zmínit i konzultanta Jana Svobodu, který dokázal zodpovědět snad jakoukoli otázku ohledně softwaru Google Earth Engine, nebo i Barboru Dobiášovou ze Zoo Praha, která se účastnila našich společných meetingů a poskytovala důležité informace k celému výzkumu.

Abstrakt

Bakalářská práce se zabývá analýzou časových řad na území Mongolska za účelem vybrání vhodné lokality pro reintrodukcii koně Převalského do volné přírody. V rámci rešeršní a teoretické části se práce zabývá zpracováním a analýzami časových řad, dále i cloud based technologiemi. Pro empirickou část byla využita cloud based platforma Google Earth Engine, v níž bylo na základě analýz vybráno v zájmové oblasti Numrug 10 potenciálních lokalit určených k opětovné reintrodukcii a dále byly vybrány 2 lokality v zájmové oblasti Gobi B, kde již reintrodukce proběhla. U těchto vybraných míst byly pozorovány průběhy a případně i trendy časových řad družicových optických snímků Landsat 8 mezi roky 2013 a 2021. Konkrétně bylo využito čtyř indexů NDVI, EVI, NDMI, NDSI a též hodnot srážek a teplot, pomocí kterých byly jednotlivé lokality porovnány. K zjištění rozdílů mezi lokalitami byl využit Mann-Whitney test a na základě výsledků byly doporučeny konkrétní lokace vhodné pro reintrodukcii. Celá práce byla doplněna i o klasifikaci obou zájmových oblastí v roce 2021. Pro tuto klasifikaci bylo využito metody CART a opět dat Landsat 8. V obou případech přesahovala celková přesnost klasifikace 80 % a Kappa index byl 0,6 a vyšší.

Klíčová slova: časové řady, Mongolsko, Landsat, klasifikace, Google Earth Engine

Abstract

This work is focusing on the analysis of the time series on the territory of Mongolia to identify a suitable location for reintroduction of the Przewalski's horse into the wild. The first part of the thesis deals with processing and analysis of the time series and the cloud-based technology. In the second part of the work, totally ten potential localities were chosen for reintroduction in the area of interest Numrug with use of cloud-based platform Google Earth Engine. Two other locations were chosen in the area of interest Gobi B, where the reintroduction had already been carried out. For these selected locations, the time series of Landsat 8 satellite optical images between 2013 and 2021 were observed. Individual locations were compared based on four indexes NDVI, EVI, NDMI, NDSI and measurements of precipitations and temperatures. Differences between the locations were detected using Mann-Whitney test. The results were used to recommend specific locations suitable for possible reintroduction. The thesis also contains classifications of the two areas of interest in 2021 using the CART method and Landsat 8 data. In both cases, the overall accuracy exceeded 80% and the Kappa index was 0.6 and higher.

Key words: time series, Mongolia, Landsat, classification, Google Earth Engine

Obsah

Obsah.....	6
Seznam zkratk.....	8
Seznam obrázků	10
Seznam tabulek.....	10
Seznam grafů.....	11
1 Úvod	12
2 Rešeršní část.....	14
3 Teoretická část.....	20
3.1. Časové řady	20
3.1.1. TS databáze.....	20
3.1.2. Vizualizace a reprezentace TS.....	21
3.1.3. Data Landsat z pohledu časových řad.....	22
3.1.4. Data Sentinel-2 z pohledu TS.....	23
3.2. Algoritmy časových řad.....	24
3.2.1. BFAST.....	24
3.2.2. LandTrendr.....	25
3.2.3. LandsatLinkr	26
3.3. Cloud based platformy pro DPZ	26
3.3.1. Google Eath Engine	28
3.3.2. Sentinel Hub	31
4 Metodická část.....	34
4.1. Zájmové území.....	34
4.1.1. Great Gobi B SPA (zájmová oblast A – po reintrodukci).....	34
4.1.2. Eastern Steppe – Numrug SPA, Dornod Mongol (zájmová oblast B – pro reintrodukci).....	35
4.2. Použitá data pro časové řady a klasifikace.....	36
4.2.1. GlobCover: Global Land Cover Map (ESA).....	36
4.2.2. CHIRPS.....	36
4.2.3. NASADEM	37
4.2.4. JRC Global Surface Water.....	37
4.2.5. NASA Global biomass carbon density	37
4.2.6. ERA5-Land ECMWF climate reanalysis.....	38
4.2.7. USGS Landsat 8 Level 2, Collection 2, Tier 1	38

4.2.8.	OpenStreetMap	38
4.2.9.	Data pro sběr trénovacích a validačních dat ke klasifikaci a doplňková data ..	38
4.3.	Zpracování dat	39
4.3.1.	Generování náhodných bodů	41
4.3.2.	Výpočet indexů a vybraných ukazatelů	42
4.3.3.	Statistická analýza dat	45
4.3.4.	Vytvoření map množství produkce biomasy	45
4.3.5.	Zpracování klasifikací	45
5	Výsledky	47
5.1.	Časové řady	47
5.1.1.	NDVI	47
5.1.2.	EVI	49
5.1.3.	NDMI	50
5.1.4.	NDSI	51
5.1.5.	Srážky	53
5.1.6.	Teploty	57
5.2.	Mapy množství biomasy	59
5.3.	Klasifikace	60
5.4.	Stanovení vhodných lokalit	62
6	Diskuse	63
7	Závěr	64
8	Literatura	65
9	Přílohy	70

Seznam zkratek

AG	Asymmetric Gaussian
APCA	Adaptive piecewise constant approximation
API	Application Programming Interface
ASTER	Advanced Spaceborne Thermal and Reflection Radiometer
BFAST	Breaks For Additive Seasonal and Trend
CFMASK	The C Function of Mask Algorithm
DB	DataBase
DL	Double Logistic
DPZ	Dálkový průzkum Země
ECMWF	European Centre for Medium-Range Weather Forecasts
ENVISAT	Environmental Satellite
ESA	European Space Agency
EVI	Enhanced Vegetation Index
GDEM	Global Digital Elevation Model
GEE	Google Earth Engine
GIS	Geographic Information System
GLAS	Geoscience Laser Altimeter System
HA	Harmonic Analysis
CHIPR	Climate Hazards Group InfraRed Precipitation
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data
ICESat	Ice, Cloud, and land Elevation Satellite
IDE	Integrated Development Environment
JS	Javascript
L1T	Level 1 Tier
LaSRC	Land Surface Reflectance Code
LC	Land Cover
LT	LandTrendr
LTS	Landsat Time Series
MERIS	Medium Resolution Imaging Spectrometer
MMU	Minimal Mapping Unit
NASA	National Aeronautics and Space Administration
NASADEM	National Aeronautics and Space Administration Digital Elevation Model
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NGA	National Geospatial-Intelligence Agency
NIR	Near Infra Red
NOAA	National Oceanic and Atmospheric Administration
NoSQL	No Structured Query Language
OGC	Open Geospatial Consortium
OLI/TIRS	Operational Land Imager and Thermal Infrared Sensor
PIP	Perceptually important points
PPA	Piecewise aggregate approximation
PRISM	Panchromatic Remote-sensing Instrument for Stereo Mapping

QA	Quality Assessment
SG	Savitzky-Golay filter
SQL	Structured Query Language
SR	Surface Reflectance
SRTM	Shuttle Radar Topography Mission
TAO	Top of the Atmosphere
TIR	Thermal Infrared
TS	Time Series
TSDB	Time Series DataBase
USGS	United States Geological Survey
VNIR	Visible and Near-Infrared
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service
WMTS	Web Map Tile Service
WS	Whittaker Smoothe

Seznam obrázků

Obrázek 1: Redukce pomocí metody vzorkování.....	21
Obrázek 2: Redukce pomocí PPA metody – průměr z každého segmentu	21
Obrázek 3: Metoda hlavních bodů – vlevo znázorněny PIP	22
Obrázek 4: Aplikace algoritmu BFAST – rozložení dat	24
Obrázek 5: Segmentace TS dat pomocí algoritmu LandTrendr.....	25
Obrázek 6: Kompozice snímků ze všech senzorů Landsat	26
Obrázek 7: Rozhraní webového IDE v GEE	29
Obrázek 8: Ukázka použití komunitního doplňku GEE v softwaru Q-GIS.....	30
Obrázek 9: Webové prostředí Sentinel Playground	32
Obrázek 10: Oblast A - Great Gobi B (červeně), oblast B - Eastern Steppe (oranžově)	34
Obrázek 11: Centrální oblast Great Gobi B, leden 2020	35
Obrázek 12: Travnaté pláně v centrální části oblasti Eastern Steppe, červenec 2015	36
Obrázek 13: Diagram postupu při tvorbě analýz	40
Obrázek 14: Výsledné zájmové území s 10 vygenerovanými body	41
Obrázek 15: Vizualizace funkce bufferPoints.....	43
Obrázek 16: Výběr pouze snímků zahrnující zájmové lokality (2015, červen), oblast Eastern Steppe.....	44
Obrázek 17: Roční srážkové úhrny v Great Gobi B	56
Obrázek 18: Roční srážkové úhrny, Eastern Steppe	56
Obrázek 19: Množství biomasy v zájmové oblasti Great Gobi B.....	59
Obrázek 20: Množství biomasy v zájmové oblasti Eastern Steppe	59
Obrázek 21: Výsledek klasifikace CART pro zájmovou oblast A.....	60
Obrázek 22: Výsledek klasifikace pro zájmovou oblast B	61

Seznam tabulek

Tabulka 1: Výsledky Shapiro-Wilkova testu pro jednotlivé lokality	47
Tabulka 2: NDVI p-hodnoty	48
Tabulka 3: EVI p-hodnoty	49
Tabulka 4: NDMI p-hodnoty	51
Tabulka 5: Celkový počet detekované sněhové pokrývky	52
Tabulka 6: NDSI p-hodnoty	52
Tabulka 7: Maxima srážkových úhrnů pro 5ti denní a měsíční řady	54
Tabulka 8: Teplota p-hodnoty	57
Tabulka 9: Chybová matice a Kappa index pro zájmovou oblast A	61
Tabulka 10: Chybová matice a Kappa index pro zájmovou oblast B.....	61

Seznam grafů

Graf 1: Časové řady NDVI	48
Graf 2: Časové řady EVI	50
Graf 3: Časové řady NDMI	51
Graf 4: Časové řady NDSI	53
Graf 5: Rozložení dat CHIRPS	54
Graf 6: Časové řady srážek	55
Graf 8: Časové řady teplot	58

1 Úvod

V posledních desítkách let dochází na naší planetě k masovému vymírání druhů, které je způsobeno změnami v jejich životním prostředí. Kůň Převalský, označován za posledního divokého koně na světě, žil ve volné přírodě do 60. let 20. století a dodnes přežil v zajetí díky velké snaze zoologických zahrad, které je od počátku 20. století chovají. Kůň Převalský žil v oblasti stepí a polopouští ve Střední Asii, což dnes odpovídá okolí hranice Mongolska a Číny (v těchto dvou zemích nalezneme největší rezervace). Tato území jsou velmi specifická a dynamická. Vyznačují se teplotními extrémy, kdy roční amplituda může přesahovat až 80 °C, a také velmi řídkou vegetací.

Reintrodukce na mezinárodní scéně probíhá již od 90. let minulého století, především ze stran německých, švýcarských a nizozemských organizací. Od roku 2011 se na projektu významně podílí i Zoo Praha, která podniká pravidelné přesuny koní do vybraných lokalit v Mongolsku. V roce 2019 bylo přepraveno do chráněné oblasti SPA Great Gobi B celkem 34 koní. V současnosti se hledá nová oblast a jako potenciálně vhodná byla vybrána část území Eastern Steppe nacházející se na východě Mongolska.

Cílem této práce je získat pomocí metod dálkového průzkumu země (DPZ) informace o vhodnosti nově vybrané lokality pro reintrodukcii. V práci jsou využity především metody časových řad (Time Series), které umožňují monitorovat naši krajinu pomocí vegetačních, vodních i půdních indexů (např. NDVI, NDMI, RGI, NDSI), či které umožňují pozorovat množství srážek či změny teplot. Ke zpracování je využíváno cloud based platformy Google Earth Engine (GEE), která disponuje značným množstvím volně dostupných knihoven a databází, a také výpočetním výkonem. Níže jsou uvedeny dílčí cíle této bakalářské práce:

1. Představit možnosti analýzy trendů a detekce trendů
2. Představení a otestování volně dostupných dat a dostupných technologií pro tvorbu časových řad
3. Pomocí vybraných metod časových řad detailně analyzovat území
4. Pomocí vybraného klasifikátoru a dat Sentinel-2 nebo Landsat provést klasifikaci obou pozorovaných území
5. Provést detailní srovnání obou pozorovaných území
6. Prokázání závislosti / podobnosti zkoumaných území pomocí statistických metod

7. Stanovení vhodnosti vybrané potenciální lokality pro reintrodukcii koně Převalského na základě získaných výstupů a na základě konzultací s odborníky se Zoo Praha (např. v souvislosti s výskyty sucha, dzud a dalších extrémních jevů)
8. Vytvoření metodiky/algoritmu ve volně dostupné výzkumné platformě GEE

Teoretická část práce je věnována problematice volně dostupných dat pro pozorování časových řad k získání detailních fyzicko-geografických charakteristik u rozsáhlejších ploch. V empirické části je provedeno detailní testování a jsou aplikovány pokročilé klasifikační metody a metody časových řad pomocí platformy GEE. Získané výsledky byly zhodnoceny pomocí vybraných statistických metod.

2 Rešeršní část

S nástupem moderních technologií dálkového průzkumu Země (DPZ) se objevily i nové možnosti zkoumání naší planety. Právě tyto technologie umožňují uživateli analyzovat velmi rozsáhlá území během krátkého časového úseku. Zejména pak metody Time Series (TS, česky časových řad), využívající data s vysokým prostorovým a spektrálním rozlišením, umožňují detailně sledovat vývoj určitého území (Hladky a kol. 2020). K výraznému rozšíření a vzniku nových algoritmů sloužících k detekci změn v TS došlo v roce 2008, kdy byl otevřen přístup ke všem archivovaným snímkům misí Landsat (Zhu 2017).

Lastovicka a kol. (2020) uvádějí, že pro TS pozorování je vhodné využít jednoho senzoru, avšak ideální je vybrat data s vysokým temporálním rozlišením (vzhledem k možné nehomogenitě prostředí v čase). Pro pozorování lze dle Štych a kol. (2019) využít i více senzorů z jedné mise, avšak data je pak nutno matematicky korigovat na srovnatelné podmínky měření. Ve své práci využívají dat různých senzorů mise Landsat pro vybraná území lesních ekosystémů, které z důvodů nízkého temporálního rozlišení normalizují metodou PIF, která využívá lineární regrese k porovnání dvou snímků. Data pro různé stavy lesních ekosystémů po proběhlých disturbancích testují pomocí poměrových vegetačních indexů. Hlavní hypotézou obou studií bylo prokázat možnost užití spektrálního pásma SWIR pro detekci kůrovcových kalamit, kde Hladky a kol. (2020) uvádějí, že stromy po napadnutí kůrovcem ztrácejí vláhu, což lze detekovat právě pomocí SWIR pásma. Hlavními výzkumnými otázkami bylo, zda lze detekovat i prvotní stádium disturbance pomocí vegetačních indexů, případně jaké vegetační indexy lze k pozorování kůrovcových a větrných kalamit využívat. Na rozdíl od běžně užívaných indexů, jako je NDVI, nedosahují po disturbanci SWIR indexy hodnot, jako dosahovaly před disturbancí. Typicky je to způsobeno zarůstáním dané lokality vegetací (Hladky a kol. 2020). Index NDMI se jeví jako velmi relevantní pro další testování v geograficky odlišných stanovištích, je tedy dobré zvážit využití indexu pro pozorování změn a cyklů vegetace.

Bao a kol. (2014) se zabývají dynamikou a trendy vegetačního krytu na mongolské plošině pomocí časových řad (TS, Time Series) dvou datových sad GIMMS NDVI a MODIS NDVI, jejichž hodnoty validují pomocí závislosti s klimatickými daty. Bao a kol. (2014) a Hladky a kol. (2020) uvádějí, že pro detailní měření vegetační dynamiky je vhodné použít

data z více datových zdrojů například pomocí normalizačních, harmonizačních metod či metod fúzí dat. Bao a kol. (2014) se zaměřili na metodu cross calibration pro dva zmíněné typy dat. Hlavní výzkumnou otázkou bylo, zda je možné prokázat vliv srážek či teploty na vegetační indexy, k čemuž bylo využito Pearsonova korelačního koeficientu, následně byl prokázán signifikantní vliv (Bao a kol. 2014).

Při pozorování normalizovaného vegetačního indexu (NDVI) v rámci časových řad dochází často k narušení hodnot vlivem mraků, aerosolů či sněhovou a ledovou pokrývkou. Zhou a kol. (2016) porovnávají pět nejpoužívanějších modelů sloužících k redukci vzniklých odchylek, mezi které řadí Harmonic Analysis (HA), Double logistic (DL), Asymmetric Gaussian (AG), Whittaker smoother (WS) a Savitzky-Golay filter (SG). Cílem studie bylo porovnat zmíněné modely a vytvořit praktická doporučení pro budoucí aplikace časových řad. Zhou a kol. (2016) uvádějí, že „signál NDVI a disturbance šumu jsou vysoce závislé na regionálním podnebí a krajinném pokryvu, a proto lze očekávat, že rekonstrukční výkon bude prostorově variabilní“. Pro hodnocení výkonu jednotlivých metod byly vypočteny tři ukazatele: celková chyba rekonstrukce (ORE), modelová chyba (FRE) a normalizační hluková chyba (NNRE). Dle ORE dominoval AG model ve většině boreálních oblastí s vysokou zeměpisnou šířkou, zatímco model SG poskytoval nejlepší výsledky při rekonstrukci v tropických a subtropických oblastech (Zhou a kol. 2016). Srovnávací metoda, kterou Zhou a kol. (2016) použili ve své studii, vedla k návrhu nového adaptivního modelu sloužícího pro efektivnější rekonstrukci NDVI.

Verbesselt a kol. (2010) detekovali sezónní a trendovou složku v časových řadách. Stejně jako Zhou a kol. (2016) se shodují, že TS jsou závislé na regionálním podnebí a krajinném pokryvu, respektive detekce změn je často ovlivněna sezónními výkyvy, které mohou být místně specifické. Zatímco změny v trendové složce naznačují disturbance, změny v sezónní složce značí změny fenologické (Verbesselt a kol. 2010). Verbesselt a kol. (2010) používají obecný přístup s R balíčkem zvaný BFAST, který detekuje změny v časových řadách dle jednotlivých sezónních a trendových složek. Sledována je 16denní TS normalizovaného vegetačního indexu (NDVI). BFAST odhalil změny v TS s různou velikostí ($>0,1$ NDVI), přičemž hladina šumu se pohybovala v intervalu 0,01-0,07 a sezónní amplituda NDVI oscilovala v intervalu 0,1-0,5. BFAST se jeví jako vhodná metoda k detekci změn, jelikož lze

použit na časové řady bez nutnosti klasifikace krajinného pokryvu nebo výběru referenčního období.

Chen a kol. (2017) se věnují mangrovovým lesům na jižním pobřeží Číny. Hlavním cílem bylo detekovat změny těchto lesů za pomoci TS dat, konkrétně fúze dat mise Landsat a Sentinel-1. Jak Bao a kol. (2014), i Chen a kol. (2017) se shodují, že vegetační indexy jsou klíčovým prostředkem pro hodnocení vegetace jako takové, zejména pokud ji chceme sledovat v rámci TS. Pro sledování rozsáhlého území je však zapotřebí zpracování mnoha satelitních snímků, což lze řešit pomocí cloud based technologií.

V rámci cloud based technologií máme k dispozici vícero online platform například Google Earth Engine (GEE) nebo Sentinel Hub (Hladky a kol. 2020, Lastovicka a kol. 2020). Chen a kol. (2017) doporučují pro tvorbu časových řad platformu GEE, která nabízí obsáhlou knihovnou volně přístupných satelitních dat a zároveň poskytuje velký výpočetní výkon. Chen a kol. (2017) uvádějí, že „GEE je navržen tak, aby byli uživatelé schopni provádět analýzy na rozsáhlých územích, a přitom nebyli omezeni výkonem“.

Pro snadné vytváření TS z dat Landsat byl navrhnout algoritmus LandTrendr (Landsat-based detection of Trends in Disturbance and Recovery), který je schopen analyzovat krátkodobé i dlouhodobé trendy (Kennedy, Yang, Cohen 2010). Nevýhodou jsou požadavky na výpočetní výkon, předzpracování snímků a přístupnost v programovacím jazyce IDL, které lze však zjednodušit pomocí platformy GEE (Kennedy a kol. 2018). Kennedy a kol. (2018) ve své práci aplikují algoritmus LandTrendr v platformě GEE na šest studijních oblastí v USA, které reprezentují různé typy krajinného pokryvu (LC, Land Cover). Následně porovnávají vytvořené skripty (jazyk Javascript) pro GEE se skripty pro IDL. Výsledky prokázaly, že oba vytvořené algoritmy se shodují téměř ve všech případech až na minimální odchylky, které lze připisovat chybám způsobené překladem kódu (Kennedy a kol. 2018). Vytvořený skript *LT-GEE* tak představuje kvalitní transformaci kódu LandTrendr do platformy GEE, která je více všestranná a snadno přístupná pro širší skupinu uživatelů.

Katsanos, Retalis, Michaelides (2016) srovnávali datovou sadu CHIRPS s obdobnými klimatologickými databázemi 3B43 a E-OBS věnující se srážkám, které disponují vysokým rozlišením. V zájmovém území celého středomoří korelovala data CHIRPS se sadami 3B43 a E-OBS, přičemž korelační koeficient u sady 3B43 byl v průměru vyšší jak 0,8, u sady E-OBS nabýval hodnot nižších (Katsanos, Retalis, Michaelides 2016). V druhé části byla sada CHIRPS

porovnávána s daty ze srážkoměrných stanic na Kypru. Vyšší korelace byly pozorovány v zimních měsících a v horských oblastech, přičemž v pozorovaném období (1981-2010) docházelo k nadhodnocování srážek v posledním desetiletí, pravděpodobně z důvodů zapracování TRMM do odhadů CHIRPS (Katsanos, Retalis, Michaelides 2016, Toté a kol. 2015). Banerjee a kol. (2020) taktéž hodnotili srážkové údaje se sadou CHIRPS. Obdobně jako Katsanos, Retalis, Michaelides (2016) doporučují tuto datovou sadu především pro oblasti, pro která nejsou dostupná kvalitní, nebo dokonce žádná data. Srážky modelované pomocí CHIRPS velmi dobře odpovídají naměřeným hodnotám „in situ“. Banerjee a kol. (2020) pracovali v prostředí cloudové platformy GEE, kterou taktéž doporučují k prostorovým analýzám dálkového průzkumu země, jelikož lze získat snadno a rychle velké množství dat. Paredes-Trejo, Barbosa, Lakshmi Kumar (2017) srovnávají CHIRPS s měřeními 21 pozemních stanic v Severovýchodní Brazílii, podobně jako Katsanos, Retalis, Michaelides (2016) na Kypru. Výsledky prokazují vysokou korelaci dat CHIRPS se všemi pozemními měřeními, ale byla pozorována tendence nadhodnocovat nízké a podceňovat vysoké hodnoty srážek (>100 mm/měsíc).

Při srovnávání sady CHIRP, které oproti CHIRPS nezahrnují data ze srážkoměrných stanic), a sad African Rainfall Climatology v2 (ARC2), a Tropical Applications of Meteorology Satellite data (TAMSAT) v Africe, byla taktéž datová sada CHIRPS doporučena pro odhad srážek, nicméně stejně jak Banerjee a kol. (2020) zmiňují Dinku a kol. (2018) náchylnost modelu v pobřežních a horských oblastech.

Bai a kol. (2018) upozorňují na významný rozdíl mezi bodovým a mřížkovým hodnocením srážek. Také bylo zpozorováno, že CHIRPS prokazuje lepší výsledky pro oblasti s většími srážkami, než pro suché či polosuché oblasti. Dále sada vykazuje lepší výsledky v letních měsících než v zimních, a to především kvůli omezenosti této datové sady detekovat sněžení. K podobným výsledkům došli i (Toté a kol. 2015), tedy dochází k nadhodnocování nízkým a podceňování vysokých hodnot dešťových srážek, a v období dešťů jsou získávány opět lepší výsledky, než v suchých obdobích.

Použití sady CHIRPS se tak zdá jako velmi dobrou volbou v oblastech, kde nejsou dostupná kvalitní či žádná data, nicméně je nutné počítat s možnými odchylkami ve výsledcích, a to především v oblastech a ročních obdobích, které zmiňují Bai a kol. (2018), Toté a kol. (2015), Banerjee a kol. (2020), Dinku a kol. (2018).

Informace o sněhové pokrývce jsou dalším důležitým ukazatelem pro celou řadu výzkumných účelů. Normalizovaný rozdílový sněhový index (NDSI, Normalized Difference Snow Index) patří mezi nepoužívanější indexy, který je schopen takové informace poskytnout. Salomonson, Appel (2004) odhadovali podíl sněhové pokrývky z dat MODIS právě pomocí tohoto indexu. Cílem práce bylo zjistit, zdali existuje nějaký znak/vzorec v datech NDSI, který by mohl posloužit k odhadu podílu sněhu v jednom pixelu dat MODIS. Bylo zjištěno, že vztah mezi částečnou sněhovou pokrývkou a NDSI byl silný, nicméně za podmínky, že je aplikován lokálně a na velkých oblastech, jako je Severní Amerika.

Shimamura, Izumi, Matsuyama (2006) použili k identifikování zasněžených oblastí tři odlišné metody. První metodu označují za obecnou a jednoduchou techniku sloužící k identifikaci sněhové pokrývky, jelikož je založena na viditelné bělosti sněhu (odrazivosti). Druhá metoda je založena na NDSI indexu, která využívá viditelné a krátkovlnné infračervené odrazy. Poslední metoda je založena novém indexu sněhové pokrývky nazývaném S3. Ten obdobně jako NDSI využívá viditelné, krátkovlnné infračervené odrazy, ale navíc i blízké infračervené odrazy. Shimamura, Izumi, Matsuyama (2006) uvádí, že za dobrých podmínek může viditelná (červená) odrazivost velmi dobře identifikovat zasněžené oblasti, ale nesmí být přítomna vegetace. NDSI je schopen tyto oblasti, kde dochází ke smíšení sněhové pokrývky a vegetace rozlišit, a to za pomoci vegetačního indexu NDVI. Nově vytvořený index S3 je dle výsledků nejlepší pro hodnocení sněhové pokrývky, jelikož je schopen automaticky (bez referenčních údajů) roztrždit zasněžené oblasti, oblasti s vegetací a smíšené oblasti, a to právě syntézou prvních metod.

V platformě GEE hodnotí Li a kol. (2018) sněhovou pokrývku s využitím dat MODIS a dat Landsat TM (tu označují za „skutečná data“ díky vysoké snímkovací frekvenci). Celková průměrná přesnost dat MODIS byla téměř 89 % v porovnání s daty Landsat. GEE doporučují k rozsáhlejším, ale i menším prostorovým analýzám, především kvůli časové úspoře a menší náročnosti procesu.

Munkhjargal a kol. (2019) využívali snímky Landsat 7, 8 a Sentinel-2A se „sněhovými“ datovými sadami MODIS Terra a MODIS Aqua (MODIS, Moderate Resolution Imaging Spectroradiometer) s denní frekvencí snímkování (2000-2017). Munkhjargal a kol. (2019) zdůrazňují, že „data Landsat a Sentinel-2A sice nabízejí vysoké prostorové rozlišení, avšak při nižší frekvenci snímkování v porovnání s daty MODIS“, a proto vyzdvihují kombinaci obou

kolekcí. Cílem bylo zjistit dobu trvání sezónní sněhové pokrývky v pohoří Khentii v Severním Mongolsku. K validaci jsou využita data „in situ“ z pozemního měření. Výsledky ukázaly, že průměrný nárůst sezónní sněhové pokrývky roste s nadmořskou výškou přibližně +6 dní/100 m, přičemž nebyly nalezeny výrazné rozdíly ve výšce sněhové pokrývky v závislosti na nadmořské výšce. Jižní svahy byly vyhodnocovány jako nezasněžené, nebo jen částečně zasněžené, což je způsobeno převládajícími větry a rychlou sublimací sněhu (Munkhjargal a kol. 2019). V nižších oblastech (<960 m. n. m.) se průměrná délka sezónní sněhové pokrývky pohybuje okolo 124 dnů, na vrcholcích pohoří (>2800 m. n. m.) až 226 dnů, přičemž chyba měření je 12-13 dnů. Datování začátku sněžení či tání vykazovalo velkou meziroční variabilitu a nebyl nalezen žádný významný trend.

Klasifikace krajinného pokryvu (Land Cover, Land Use) je v dálkovém průzkumu jeden z nejvíce zkoumaných postupů, jelikož poskytuje užitečné informace o rozložení tříd krajinného pokryvu (voda, les, zástavba atd.) pro mnoho výzkumných účelů. Přestože tvorba mapy krajinného pokryvu je poměrně jednoduchá, není snadné zachovat dobrou kvalitu a přesnost (Zhu, Woodcock 2014). Použití multitemporálních dat je v tomto směru přínosné, neboť mohou přispět ke zvýšení přesnosti, a to zejména kvůli schopnosti kvalitně detekovat vegetaci. Zhu, Woodcock (2014) uvádějí dvě základní pravidla při tvorbě klasifikací: (1) ve všech použitých snímcích nesmí být oblačnost, což je pro data s nízkou frekvencí snímání, jako je Landsat, obtížné, (2) předpoklad, že v časovém intervalu, kdy je tvořena mapa krajinného pokryvu, nedojde k žádné její změně. Velikost chyby je v některých klasifikacích větší než samotná změna klasifikovaného povrchu, a proto nelze změnu identifikovat (Zhu, Woodcock 2014).

3 Teoretická část

3.1. Časové řady

TS jsou užitečná forma dat, díky kterým můžeme odhalit důležité vlastnosti zkoumaného území (Hladky a kol. 2020). V rámci statistické analýzy lze identifikovat zákonitosti a detekovat změny v jejich chování pomocí dekompozice časové řady na trendovou, cyklickou či sezónní složku (West 1997). Na základě TS dat jsme schopni sestavit model pomocí interpolace a odhadnout tak i budoucí vývoj sledovaných hodnot, který nazýváme extrapolací. TS se dělí na dvě základní skupiny: (1) spojité jsou takové řady, kde známe hodnoty v každém okamžiku pro celý sledovaný časový interval, (2) diskrétní časové řady, u kterých známe hodnoty jen pro určitý čas, kdy byla hodnota získána. V rámci DPZ jsou TS využívány především pro pozorování krajinného pokryvu, ale i různých indexů (Kennedy a kol. 2018).

3.1.1. TS databáze

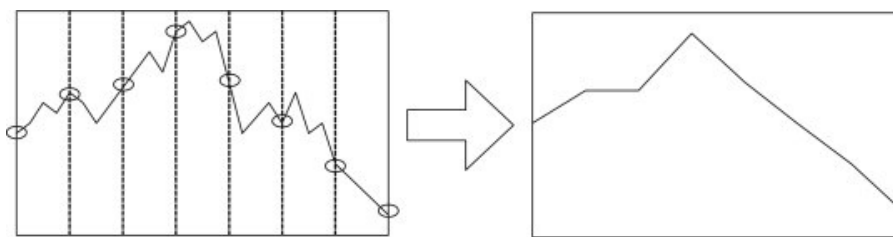
Databáze (DB) jsou definovány jako soubor uspořádaných informací, které jsou organizovány tak, aby k nim bylo možné snadno přistupovat, spravovat je a aktualizovat. Zpravidla jsou řízeny systémem pro správu dat, přičemž spolu s daty je takové spojení označováno jako databázový systém, zjednodušeně též databáze (Oracle 2019).

V rámci objektově orientovaných DB na bázi SQL (Structured Query Language) či NoSQL jsou dnes k dispozici také TSDB (Time Series Database), které jsou určeny a optimalizovány pro TS data (Mueen a kol. 2009). Jako takové je nelze zařadit do ani jedné výše zmíněné skupiny, jelikož jsou založeny jak na SQL, tak NoSQL, záleží dle jednotlivých softwarů. V současnosti se TSDB používají především k monitoringu real-time aplikací, různých procesů v celé řadě odvětví, analýze dat z webů či různých senzorů (Timescale 2018). Jednotlivé softwary se liší nabízenými funkcemi, ale obecně je největší výhodou TSDB ve snadné vizualizaci a analýze dat, které lze i statisticky testovat. Softwarů je spousta, mezi známější se řadí InfluxDB (napsán v jazyce Go), TimescaleDB (založený na PostgreSQL), Prometheus, Graphite nebo i Firebase od Googlu. Některé softwary, jako

je i poslední zmíněná platforma Firebase používaná pro vývoj mobilních a webových aplikací, nejsou TSDB v právním slova smyslu. Jedná se o tzv. real-time DB, které však nabízejí podobné funkce jako TSDB, a lze je tak do této skupiny zařadit.

3.1.2. Vizualizace a reprezentace TS

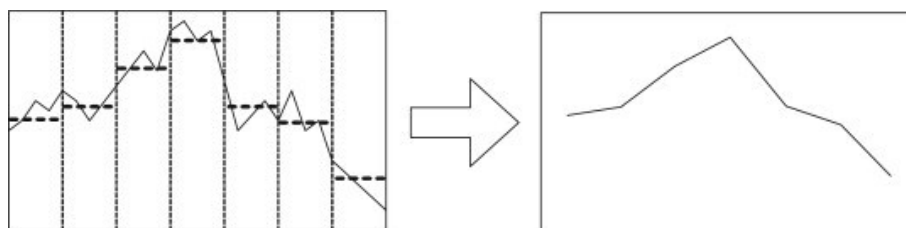
Při tvorbě obsáhlých TS, kde jsou data k dispozici ve vysoké frekvenci, či jsou pro dlouhé časové období, vzniká problém s vhodnou reprezentací a vizualizací dat. Bazální dostupná metoda sloužící k redukci dat je vzorkování, kdy se využívá vzorec m/n , kde m je délka časové řady a n počet vzorkovacích bodů (Fu 2011). Nevýhodou této metody je tendence narušovat tvar vzorkované TS (obrázek 1), a to především v případech, kdy je frekvence dat příliš nízká (Fu 2011).



Obrázek 1: Redukce pomocí metody vzorkování

(zdroj: Fu 2011)

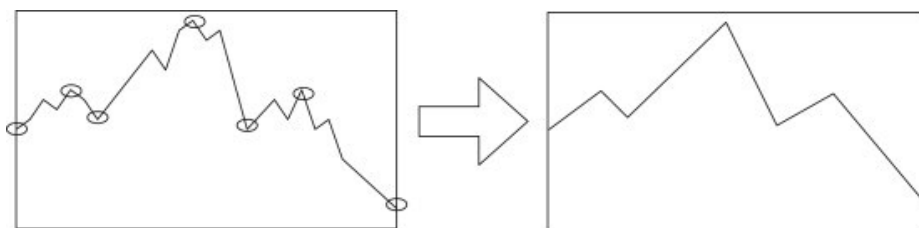
Pokročilejší a odolnější metodou je „Piecewise aggregate approximation“ (PPA). Podobně jako při vzorkování se rozdělí TS na jednotlivé segmenty (obrázek 2), z kterých se však vypočítá průměr (Fu 2011), oproti vzorkování, kde se vezme jen daná hodnota. PPA byla ještě rozpracována a vznikla pokročilejší metoda APCA (Adaptive piecewise constant approximation), ve které není šíře daného segmentu pevná, ale mění se dle tvaru TS (Fu 2011).



Obrázek 2: Redukce pomocí PPA metody – průměr z každého segmentu

(zdroj: Fu 2011)

Další metoda „hlavních bodů“ pracuje s tzv. klíčovými/lomovými body. Tyto body jsou nazývány jako percepčně důležité body (PIP, Perceptually important points). Z celého úseku TS se určí dva PIP znázorňující minimum a maximum, a vznikne tak rozdělení na jednotlivé segmenty. V těch se určí další PIP a postupuje se stejně až pro PIP_n (obrázek 3), které je určeno dle požadované přesnosti TS (Fu 2011).



Obrázek 3: Metoda hlavních bodů – vlevo znázorněny PIP

(zdroj: Fu 2011)

3.1.3. Data Landsat z pohledu časových řad

V roce 2008 byly zpřístupněny archivy s veškerými daty Landsat, což zcela změnilo pohled na práci s nimi. Při práci s daty TS, u kterých cílíme na pozorování změn v čase, je důležité brát v potaz čtyři důležité aspekty: frekvenci snímkování, preprocessing, dostupné algoritmy a následné aplikace (Zhu 2017). Na počátku se data Landsat v časových řadách používala zřídka, především kvůli vysokým nákladům na pořízení dat a také kvůli nutnému vysokému výpočetnímu výkonu, proto byl používán především bitemporální přístup, který sledoval snímky pouze ve dvou časech (Zhu 2017). S rostoucím výpočetním výkonem, který se stával dostupnější pro širší okruh uživatelů, vznikaly složitější algoritmy a metody pro sledování změn v čase, schopné poskytnout komplexnější vysvětlení, co se na zemském povrchu děje (Zhu 2017).

Mise Landsat dosahuje vysokého temporálního rozlišení, kdy může být konkrétní místo na zemském povrchu snímkováno jedním satelitem jednou za 16 dní (tzn. pokrytí celé zeměkoule trvá 16 dní). Kombinací vysokého temporálního a prostorového rozlišení 30m/pixel je umožněno provádět přesnější TS analýzy, jelikož je možné lépe zachytit meziroční sezónní změny (Zhu 2017).

Pro přesné zpracování analýz v čase je nutné provést preprocessing dat, aby nebyly výsledky ovlivněny vnějšími jevy. V rámci preprocessingu jde o radiometrické

korekce, jejichž součástí jsou i atmosférické korekce, topografické korekce, detekce mraků a mrakových stínů (Zhu 2017). Dále jsou podstatné i geometrické korekce dat, které snižují nepřesnosti ve snímku způsobené pohyby nosiče, pohyby Země atd. V současné době jsou k dispozici snímky

Level 1 (L1T), které disponují vysokou geometrickou přesností (< 30 m u více jak 99 % dat), tudíž lze předpokládat, že data budou dostatečně dobře zarovnána (Zhu, Woodcock 2014). Pro porovnávání dat v čase je nutné provést i radiometrické korekce, jejichž součástí jsou

i atmosférické, případně topografické korekce. Radiometrické korekce představují převod naměřených hodnot spektrální záře na relativní odrazivost záření na horní hranici atmosféry (TOA – Top of the Atmosphere). Radiometrické korekce lze rozdělit na dvě skupiny: relativní a absolutní korekce (Zhu 2017). V dnešní době jsou pro mnoho dat k dispozici již radiometricky korigované verze (obsahující tedy i atmosférické a topografické korekce), u dat Landsat se jedná o data Level 2 označována jako Surface Reflectance data (dříve Climate Data Record).

Atmosférické korekce u sady Landsat 8 SR (Surface Reflectance) jsou provedeny pomocí algoritmu odrazivosti zemského povrchu LaSRC (Land Surface Reflectance Code). K detekci mraků u optických družicových snímcích využívá algoritmus LaSRC vrstvu QA (Quality Assurance), která je vytvářena při procesu atmosférických korekcí algoritmem CFMask. QA obsahuje masku mraků, která detekuje vliv aerosolů na odvozenou odrazivost povrchu (Skakun a kol. 2019).

3.1.4. Data Sentinel-2 z pohledu TS

Evropská unie ve spolupráci s Evropskou kosmickou agenturou (ESA, European space agency) založila v roce 2014 program s názvem Copernicus se záměrem monitorovat zemský povrch pomocí metod DPZ (Phiri a kol. 2020). První družice Snetinel-2 byla vypuštěna v roce 2015 a nese označení Sentinel-2A, druhá družice v roce 2017 s označením Sentinel-2B (Phiri a kol. 2020).

Jejich hlavním cílem je poskytovat družicová data s vysokým prostorovým (10-60m), temporálním (5 dní se dvěma družicemi Sentinel-2A a Sentinel-2B) a spektrálním

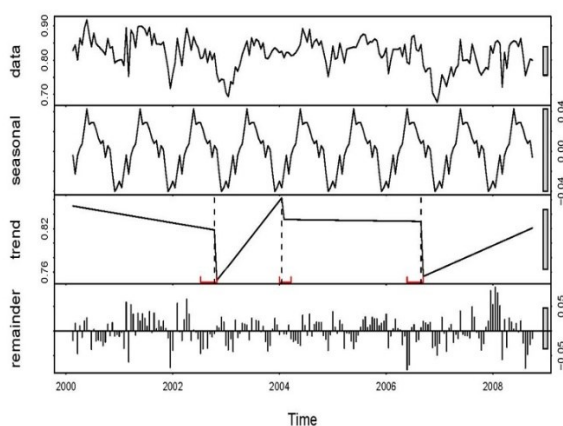
(13 pásem) rozlišením (Jönsson a kol. 2018) za účelem účinného monitorování LC, klimatických změn a přírodních katastrof (Phiri a kol. 2020).

Dalším důležitým cílem Sentinel-2 je mimo jiné také doplnit ostatní družicové mise, např. Landsat (Jönsson a kol. 2018). Data Sentinel-2 je lze navíc propojit s daty z jiných misí DPZ, čímž lze např. zvýšit celkovou přesnost klasifikací (Phiri a kol. 2020). Dle dostupných zdrojů vyplývá, že v případě analýzy LC s využitím klasifikátorů strojového učení (např. Support Vectors Machine, Random Forest) je možné dosáhnout více jak 80% přesnosti (Phiri a kol. 2020). Data Sentinel-2 jsou bezplatně přístupná na webu Copernicus Open Access Hub, podobně jako data Landsat.

3.2. Algoritmy časových řad

3.2.1. BFAST

Pro detekci a rozpoznání změn v TS je navržen algoritmus BFAST – Break For Additive Seasonal and Trend. Ze vstupních dat je schopen vytvářet TS a pomocí dekompozice je rozložit na sezónní, trendovou a rušivou složku (Verbesselt a kol. 2010). Jeho výhodou je zohlednění změn v sezónní složce a schopnost současně odhalovat případné náhlé změny v dlouhodobém časovém horizontu (Watts, Laffan 2014, Verbesselt a kol. 2010), což je znázorněno na obrázku 4, kdy byly odhaleny tři náhlé změny v trendové složce. Použit lze na různé typy dat, např. MODIS či Landsat.



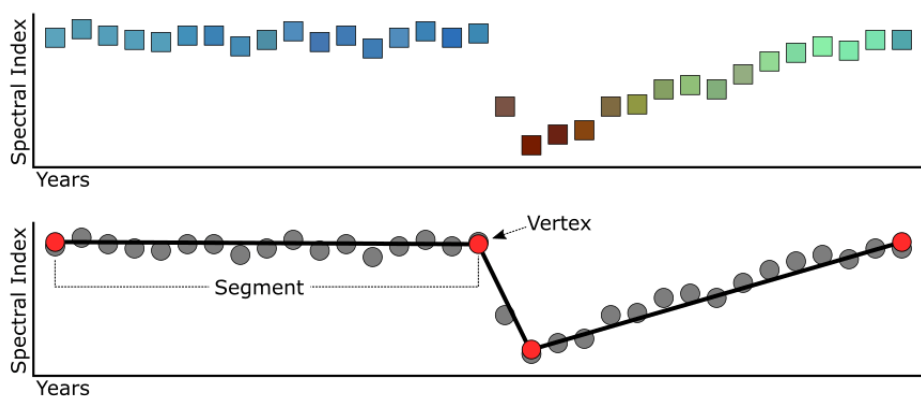
Obrázek 4: Aplikace algoritmu BFAST – rozložení dat

(zdroj: BFAST, <https://bfast.r-forge.r-project.org>)

Další balíček s označením *BFASTmonitor* poskytuje funkce pro sledování disturbancí v TS modelech. Na základě modelu vycházejícího z historických dat je schopen detekovat změny v nově získaných datech, přičemž standardně využívá model sezónního trendu (Verbesselt a kol. 2010). BFAST a BFASTmonitor jsou distribuovány v podobě balíčku pro R.

3.2.2. LandTrendr

Algoritmus LandTrendr (LT) – Landsat-based Detection of Trends in Disturbance and Recovery rozpoznává disturbance a následně vyhodnocuje změny z velkého počtu vstupních dat (Kennedy, Yang, Cohen 2010), přičemž používá segmentační metodu k detekci náhlých změn (Zhu 2017). Snímky jsou redukovány na jedno pásmo nebo jeden spektrální index, poté jsou identifikovány tzv. hraniční body, které oddělují segmenty změny a stability (obrázek 5) (Kennedy a kol. 2018).



Obrázek 5: Segmentace TS dat pomocí algoritmu LandTrendr

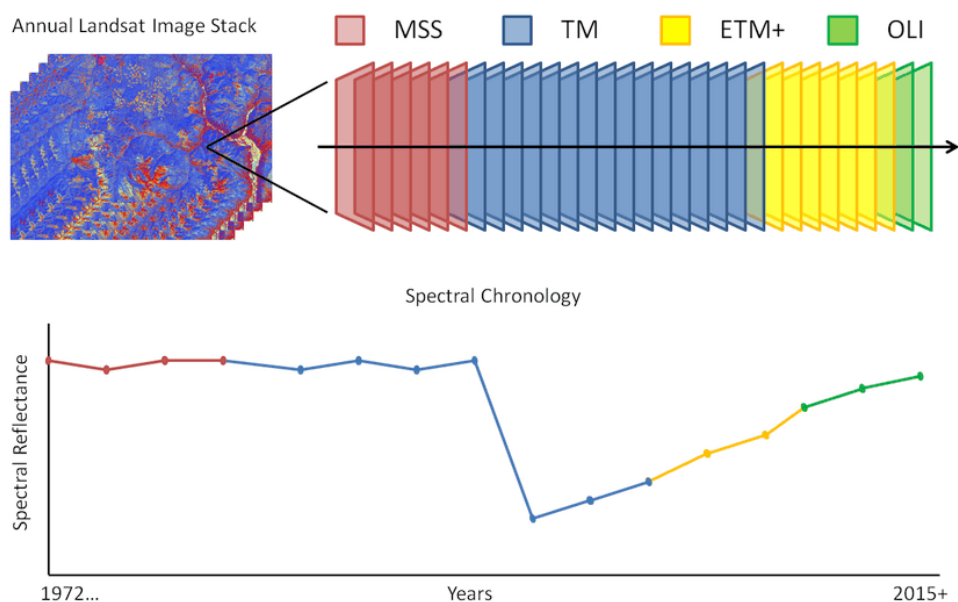
(zdroj: LT-GEE, <https://emapr.github.io/LT-GEE/index.html>)

Použití LT je náročnější, neboť vyžaduje preprocessing dat, složitější organizaci dat a výpočetní výkon. Navíc je LT uživatelům zpřístupněn pouze v programovacím jazyce IDL (Kennedy a kol. 2018), což může být pro mnoho uživatelů omezující. V rámci GEE je nabízen jako doplněk LT-GEE, který zjednodušuje kroky předzpracování a zároveň vykazuje stejně kvalitní výsledky (Kennedy a kol. 2018). Navíc je dodáván s obsáhlou dokumentací a je možno jej jednoduše spustit pomocí funkce `ee.Algorithms.TemporalSegmentation.LandTrendr`.

3.2.3. LandsatLinkr

LandsatLinkr (LLR) je automatizovaný systém zpracovávající snímky misí Landsat, který byl vyvinut byl na univerzitě v Oregonu. Propojuje snímky MSS, TM, ETM+ a OLI v čase, a to jak v prostorové, tak spektrální úrovni (Vogeler a kol. 2018). Lze tak zpracovat stovky snímků a vytvářet roční kompozice bez oblačnosti a mraků (obrázek 6). Stejně jako BFAST je dostupný v podobě R balíčku.

Složené snímky pomocí LLR lze prohlížet pomocí interaktivní webové aplikace LLR-TimeMachine založené pouze na JS, CSS a HTML (standartní pro weby). Jednoduše je lze po stažení spustit v internetovém prohlížeči jako lokální aplikaci ve webovém prohlížeči (Braaten a kol. 2016). Zobrazit lze spektrální historii vybraných pixelů jako graf TS nebo snímky oblasti v rozsahu 127x127 pixelů, které lze navíc spustit jako časoběrný obrázek. LandsatLinkr je poskytován i ve verzi LLR-LandTrendr pro snadné spuštění složených snímků LLR prostřednictvím algoritmu LandTrendr (Braaten a kol. 2016).



Obrázek 6: Kompozice snímků ze všech senzorů Landsat

(zdroj: LandsatLinkr, <https://jdbcode.github.io/LandsatLinkr/guide.html>)

3.3. Cloud based platformy pro DPZ

DPZ hraje zásadní roli v porozumění a poznávání procesů na naší planetě. Jsme schopni monitorovat vývoj životního prostředí, včasné detekovat rizika či analyzovat

rostoucí urbanizaci (Gomes, Queiroz, Ferreira 2020). Prostorová data však vyžadují velký úložný prostor i výpočetní výkon, a proto je nutné pro efektivní využívání takových dat zajistit kvalitní technickou infrastrukturu (Lee, Kang 2015). S rychlým rozvojem počítačových sítí, ukládání dat a samotné kapacity sběru dat se nyní big data rychle rozšiřují ve všech vědních a technických oblastech, včetně fyzikálních, biologických a biomedicínských věd (Sowmya, Suneetha 2014). Podle Lee, Kang (2015) prostorová „big data“ přesahují kapacitu výpočetních systémů a každým rokem roste jejich velikost až o 20 %.

V rámci cloud based technologií je užíváno několik pojmů, které je nutno zmínit a definovat. Konkrétně zmíněná big data, dále data warehouse, data mining a cloud based platformy. Jako big data se označují data, která jsou příliš velká nebo složitá pro práci s klasickými softwary. Cílem problematiky big data je najít efektivní a rychlý způsob, jak tato data analyzovat a systematicky získávat věcné informace (Sowmya, Suneetha 2014). S tím je spojeno spousta dalších problémů týkajících se samotného cyklu dat: zachycení dat, ukládání, analyzování, vyhledávání, sdílení, přesouvání, vizualizace, aktualizace či ochrana proti zneužití.

Datový sklad (data warehouse) je digitální centrální systém sloužící pro ukládání velkého množství dat (zpravidla big data) z jednoho či více různých zdrojů za účelem provádění následných efektivních analýz. Na jednom místě se ukládají data historická, ale i aktuální (Santoso, Yulia 2017).

Proces, při kterém dochází k získávání informací a vzorců z velkého objemu dat (např. z DW) se nazývá „data mining“ (tj. získávání dat). K tomu jsou často využívají metody strojového učení, statistiky a databázových systémů. Data mining zahrnuje kromě analýzy také správu dat, předzpracování i vizualizaci či aktualizaci dat (Sowmya, Suneetha 2014).

Cloud based platformy jsou zpravidla softwary umístěné na vzdáleném úložišti či serveru (tzv. cloudové úložiště), které disponují velkým výpočetním výkonem. V oboru DPZ existuje hned několik specializovaných softwarů, které zpracovávají velkoobjemová data ze satelitů. Liší se funkcemi, dostupnými databázemi i licencováním (Gomes, Queiroz, Ferreira 2020). Známe jsou Open Data Cube, System for Earth Observation Data Access, Processing and Analysis for Land Monitoring, openEO, JEODPP a pipsCloud

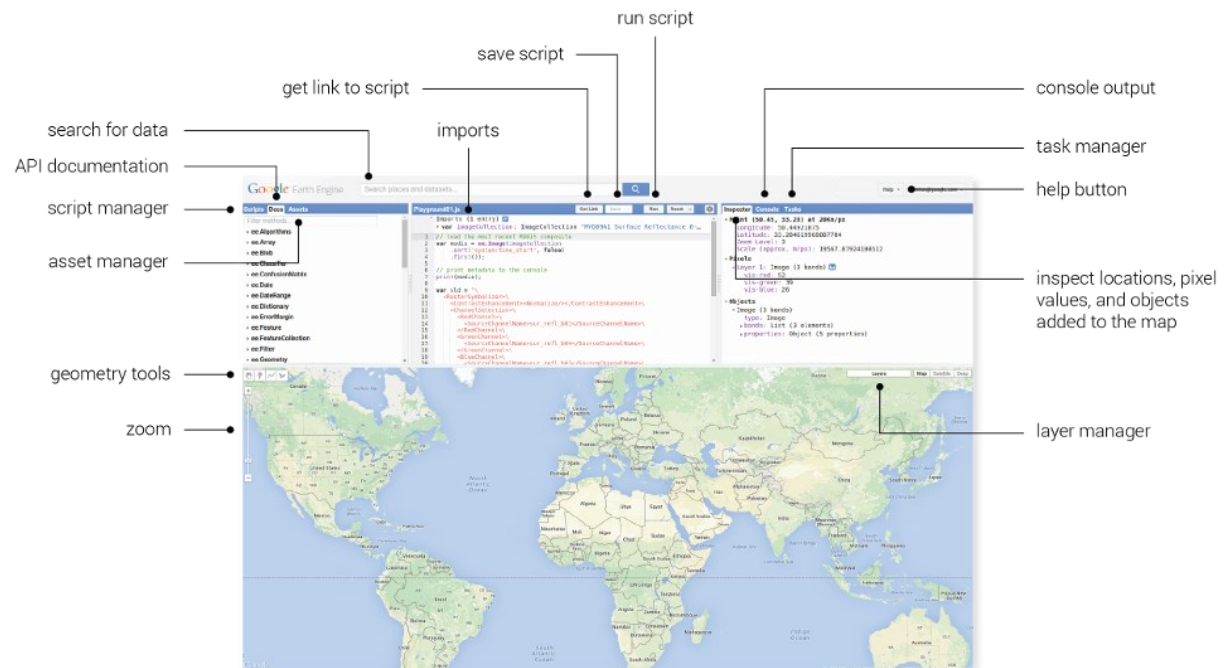
(Gomes, Queiroz, Ferreira 2020), přičemž velmi populární jsou i platformy Sentinel Hub (SH) a Google Earth Engine (GEE), které budou představeny dále.

3.3.1. Google Earth Engine

GEE je cloud based platforma kombinující multi-petabajtový katalog satelitních snímků a geoprostorových datových sad s možností tato data analyzovat a vizualizovat. Uživatelům umožňuje detekovat změny, mapovat trendy a rozdíly na povrchu Země (GEE 2020). GEE byl spuštěn v roce 2010 společností Google a v současné době je bezplatně zpřístupněn pro všechny uživatele, včetně malých a středních společností (Gomes, Queiroz, Ferreira 2020).

Primárně byl vytvořen v programovacím jazyku JavaScript (JS), který se řadí mezi nejpoužívanější jazyky současnosti (Bissyande a kol. 2013). JS běží na straně klienta, a proto je rychlejší než ostatní jazyky, které běží na straně serveru (např. PHP), proto našel uplatnění především u webových stránek a aplikací. Dalším podporovaným jazykem je Python.

Google nabízí několik možností, jak s GEE pracovat: (1) Code Editor (Editor kódu) je webová IDE (Integrated Development Environment, vývojářské prostředí), kde uživatel v horní části obrazovky píše do textového editoru JS kód formou API (Application Programming Interface), a v dolní části vidí téměř okamžitě výsledek. K dispozici má katalog dat, dokumentaci s rychlým vyhledáváním, úložiště skriptů a vlastních souborů, panel geometrických funkcí atd. (obrázek 7). Před používáním se musí uživatel přihlásit pomocí účtu Google.

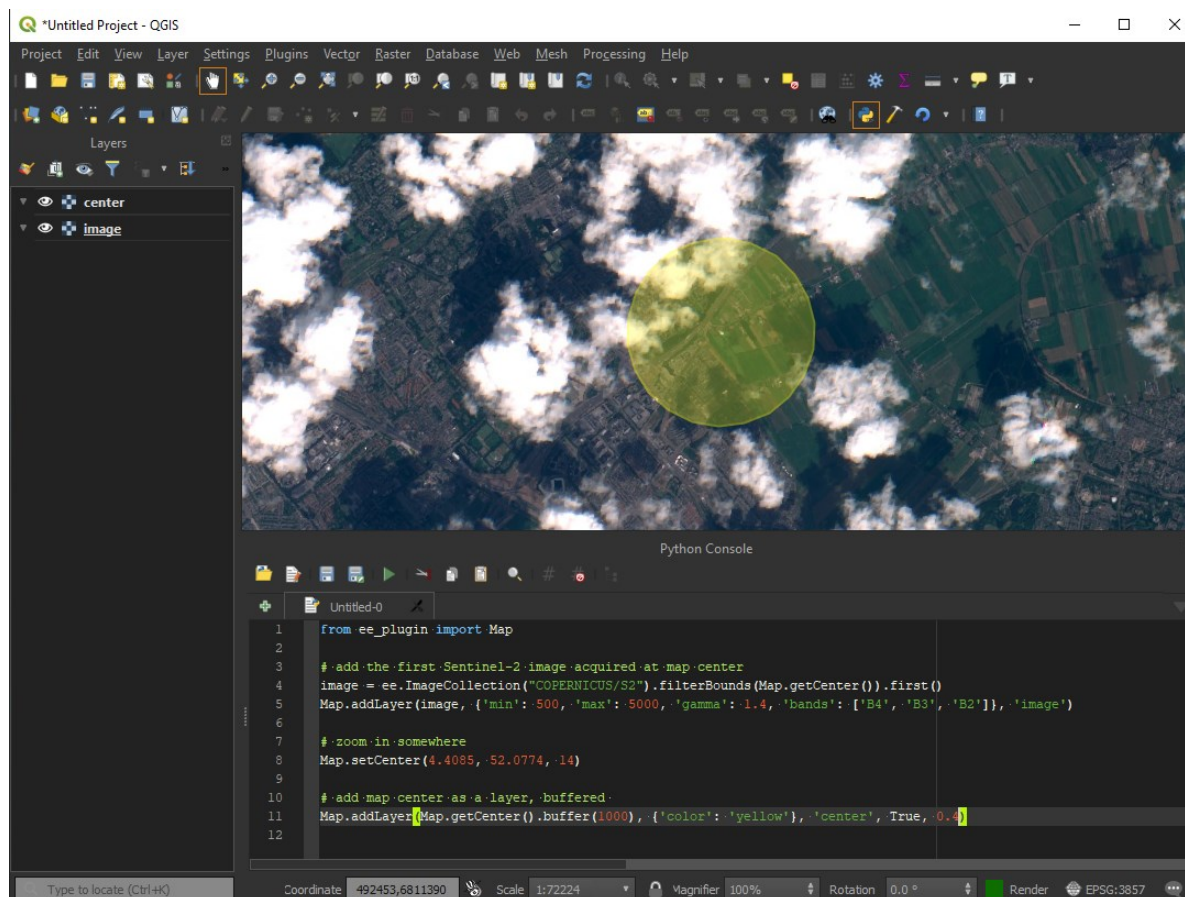


Obrázek 7: Rozhraní webového IDE v GEE

(zdroj: GEE, <https://earthengine.google.com/platform/>)

(2) Explorer (průzkumník) je odlehčená verze Code editoru, který slouží pro prohlížení dat z katalogu a umožňuje provádět jednoduché analýzy (GEE 2020). Prohlížet data může kdokoli i bez registrace, nicméně pokud chce uživatel data importovat, analyzovat, či ukládat a exportovat výsledky, musí být opět přihlášen ke Google účtu.

(3) Client libraries (uživatelské knihovny) poskytují tzv. wrapper (pro JS i Python), díky kterému je možné přistupovat vzdáleně k GEE API. Uživatel tak může vytvářet vlastní aplikace a skriptovat kód GEE na vlastním desktopu. K dispozici je na GitHub řada ukázek využití. Příkladem může být komunitní plug-in (doplňěk) *gjs-eathengine-plugin* pro GIS open-source software Q-GIS (obrázek 8), který umožňuje pracovat s GEE API přímo v programu (Qgis-gee 2019).



Obrázek 8: Ukázka použití komunitního doplňku GEE v softwaru Q-GIS

(zdroj: <https://gee-community.github.io/qgis-earthengine-plugin/>)

V katalogu dat je uloženo velké množství geoprostorových dat, včetně optických a radarových snímků z různých satelitů či leteckých systémů, předpovědí počasí a klimatu, krajinného pokryvu, socioekonomických a topografických datových souborů. Pro uživatele jsou navíc tato data před zpřístupněním předzpracována, což umožňuje velmi efektivní práci a odstraňuje mnoho překážek spojených se správou dat (Gomes, Queiroz, Ferreira 2020).

GEE používá čtyři typy objektů (reprezentující typy dat) s nimiž lze pracovat pomocí API. (1) *Image* představuje rastrová data, která se mohou skládat z jednoho nebo více pásem, jež obsahují název, datový typ, měřítko a projekci. Kolekce či časová řada jenž tvořena více snímky je reprezentována datovým typem (2) *ImageCollection*. Vektorové prvky jsou definovány pomocí (3) *Feature* a reprezentovány určitým typem geometrií a seznamem jejich atributů. Posledním typem je (4) *FeatureCollection* znázorňující skupiny vzájemně souvisejících *Features*, pomocí kterých je možné provádět třídění, filtrování či vizualizace dat (Gomes, Queiroz, Ferreira 2020).

Každý objekt patří do konkrétní třídy a každá třída má k dispozici určitou sadu funkcí. Ve skriptech se objekty a metody kombinují a odesílají na servery Earth Engine, kde se zpracovávají a následně vracejí výsledek uživateli. Ten může být vyexportován do tří služeb Googlu: (1) Google Drive, (2) Cloud Storage a (3) Assets ve zvoleném rozlišení a formátu.

Samotná knihovna GEE obsahuje více jak 800 funkcí, nicméně tyto funkce jsou pro uživatele bohužel uzavřené (Gomes, Queiroz, Ferreira 2020). Nelze je aktualizovat či rozšiřovat, což do jisté míry patří mezi největší zápory GEE. Další nevýhodou jsou i licenční podmínky stanovené Googlem, které nezaručují uživatelům duševní vlastnictví jejich skriptů a dat (Gomes, Queiroz, Ferreira 2020). Mimo jiné nelze platformu GEE využívat pro komerční využití bez povolení od samotné společnosti Google. V jistých případech může být také současná podporovaná datová sada ze dne na den nahrazena novou, a uživateli může narušit správný chod jeho skriptů.

3.3.2. Sentinel Hub

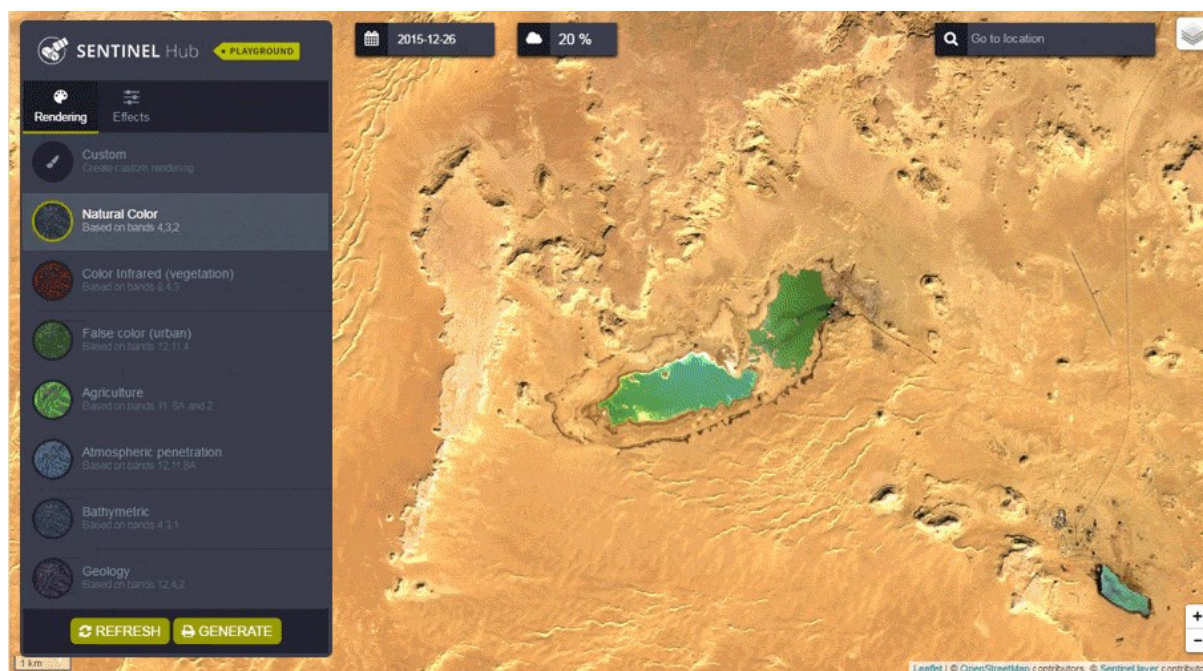
SH lze představit jako alternativu pro platformou GEE. Určena je zejména pro prohlížení dat a rychlé analýzy. Díky „user friendly interface“ (uživatelsky přívětivému rozhraní) je snadno ovladatelná i bez znalosti programovacích jazyků. Jedná o soukromou platformu s veřejným přístupem a oproti GEE je uživateli v bezplatné verzi zpřístupněna jen část funkcí (Gomes, Queiroz, Ferreira 2020). SH nabízí pět základních platebních plánů: (1) Free, (2) Exploration, (3) Basic, (4) Enterprise S a (5) Enterprise L. Bezplatný balíček umožňuje prohlížet, vybírat a stahovat pouze nezpracovaná data, a nesmí být použita pro komerční účely. Zbylé čtyři balíčky jsou odstupňovány především dle výše limitů pro přístupy a dotazy na server, možností komerčního použití, či umožnění přistupovat k datům prostřednictvím protokolů OGC a API. Cenové rozmezí se pohybuje od 25 € (Exploration) do 834 € (Enterprise L) za měsíc, přičemž je možné sestavit individuální plán. (Gomes, Queiroz, Ferreira 2020).

Nabízeny jsou dva typy dat. Jednak veřejně přístupné, tak i komerční. Mezi veřejné patří kolekce Sentinel (Sentinel-1, Sentinel-2 L1C and L2A, Sentinel-3 OLCI and SLSTR, Sentinel-5P), Landsat (Landsat 8, 7 a 5), DEM (Mapzen and Copernicus), Envisat, MODIS (MCD43A4, v6), Copernicus (Corine Land Cover, Global Land Cover, Water Bodies).

Z komerčních jsou dostupné data od Airbus, PlanetScope a Maxar. Uživatel má však možnost použít svá vlastní data (Gomes, Queiroz, Ferreira 2020).

K dispozici jsou v rámci SH tři služby: Sentinel Playground, EO Browser a Dashboard. Playground se používá především pro rychlou vizualizaci snímků Sentinel-1, Sentinel-2, Landsat 8, DEM a MODIS, a to v plném rozlišení. Jedná se o alternativu GEE Explorer.

Nabízen je denně aktualizovaný dataset Sentinel-2 vhodný pro pozorování každodenních změn na povrchu Země. Uživatel si zobrazí požadovanou lokaci, zvolí datum a maximální oblačnost, a následně mu jsou zobrazeny nabízené snímky s možností volby předvolené kombinace jednotlivých pásem pro zvolený účel (zemědělství, geologie, zobrazení v přirozených barvách apod.). Samozřejmě je k dispozici i možnost kombinovat pásma libovolně (obrázek 9).



Obrázek 9: Webové prostředí Sentinel Playground

(zdroj: <https://www.sentinel-hub.com>)

EO Browser (Earth Observation Browser) nabízí mnohem více funkcí. Pomocí JS skriptů (Evalscrip) je možné provádět analýzy, sestavovat timelapse (časosběrné snímky) nebo stahovat vybrané snímky. Připraveny jsou i vzdělávací tutoriály, které se věnují 12 tématům.

Dashboard poskytuje přístup ke konfiguračnímu nástroji, skrz který je možné přistupovat k uživatelským vytvořeným kolekcím dat. Vytvářet lze vlastní služby WMS/WMTS/WFS/WCS, přičemž každou je možné individuálně nakonfigurovat. Vstupní data mohou být použita z prostředí SH, nebo lze nahrát i data vlastní a následně je upravovat či analyzovat. Správce má možnost nastavit práva a také platební plán pro uživatele (SentinelHub 2020).

SH využívá k reprezentaci dat koncept zdroje dat, instancí a vrstev, které jsou v něm k dispozici jako služby. Zdroj dat je podobný typu *ImageCollection* v GEE. Představuje datové sady, které mají stejnou sadu pásem a metadata (Gomes, Queiroz, Ferreira 2020). K dispozici je rozsáhlá dokumentace s příklady použití webového rozhraní a SH API. Podobně jako GEE neposkytuje SH záruku na poskytování služeb, a může tak nabízené funkce a služby kdykoli změnit (Gomes, Queiroz, Ferreira 2020). Oproti GEE nenabízí SH žádné nástroje, které by usnadňovali sdílení skriptů mezi uživateli, a proto jsou pak skripty nejčastěji sdíleny pomocí GitHubu.

4 Metodická část

4.1. Zájmové území

Do práce vstupují dvě zájmové oblasti nacházející se v protilehlých částech Mongolska (obrázek 10). Národní park Great Gobi B SPA (Strictly Protected Area) nacházející se na jihozápadě, kde je již kuň Převalský reintrodukovan (1), a oblast na východě poblíž národních parků Numrug a Dornod Mongol, kde je reintrodukce plánována (2).



Obrázek 10: Oblast A - Great Gobi B (červeně), oblast B - Eastern Steppe (oranžově)

(zdroj: vlastní tvorba)

4.1.1. Great Gobi B SPA (zájmová oblast A – po reintrodukci)

Great Gobi B SPA je přísně chráněná oblast na jihozápadě Mongolska s rozlohou okolo 9000 km². Povrch tvoří převážně stepi, pouště a polopouště. Průměrné roční srážky činí přibližně 100 mm/rok, avšak ve vyšších oblastech (až 2840 m. n. m.) dosahují 180 mm/rok. Roční amplituda teplot je velmi vysoká a může dosahovat až 80 °C (extrémy jsou +40 °C v létě, -40 °C v zimě), přičemž průměrné zimní i letní teploty jsou nižší (Wehrden, Wesche, Tungalag 2006). Zima je dlouhá a chladná, léto krátké a horké. Sněžení je intenzivnější než v ostatních

oblastech Gobi a v průměru trvá období se sněhovou pokrývkou 97 dní (obrázek 11), přičemž je tato hodnota velmi variabilní (Kaczensky a kol. 2008). Na východě dominují roviny s nízkými horami a na západě zvlněné kopce. Nadmořská výška se zvyšuje ze severozápadu (1000 m. n. m.) směrem ke hranici s Čínou (2840 m. n. m.) (Kaczensky a kol. 2008). Až na hraniční přechody a několik obydlí pastevců je oblast neobydlená (Wehrden, Wesche, Tungalag 2006).



Obrázek 11: Centrální oblast Great Gobi B, leden 2020

(zdroj: Google Street View, <http://6b.cz/6PIV>)

4.1.2. Eastern Steppe – Numrug SPA, Dornod Mongol (zájmová oblast B – pro reintrodukci)

Východní stepi tvoří bezlesé rozsáhlé stepi, mírně zvlněné kopce, mokřady a pohoří Chjangan, které vede až k hranicím s Čínou (UNESCO 2014). Nadmořská výška se v oblasti pohybuje v intervalu od 500 do 1800 m. n. m. Roční srážky činí v průměru přibližně 230 mm \pm 70 mm, nicméně ve stepích a polopouštích se blíží hodnota ke 100 mm/rok a ve vyšších nadmořských výškách (např. pohoří Chjangan) až 400 mm/rok. Častým povrchem jsou travnaté pláně, které se řadí mezi největší zbývající nedotčené travnaté plochy mírného pásma na Zemi (obrázek 12), v některých oblastech i s křovinami, vzácnější jsou ojedinělé stepní lesy (Olson a kol. 2010, UNESCO 2014). Součástí ekosystému Eastern Steppe je i rozsáhlá fauna, do které patří např. gazela asijská, též Dzeren, tvořena až 2 miliony jedinců, psík mývalovitý, netopýr asijský atd. (UNESCO 2014)



Obrázek 12: Travnaté pláně v centrální části oblasti Eastern Steppe, červenec 2015

(zdroj: Google Street View, <http://6b.cz/cI2K>)

4.2. Použitá data pro časové řady a klasifikace

V práci bylo využito několik datových sad z katalogu GEE, které jsou volně dostupné. Dále jsou uvedeny jednotlivé datové zdroje a jejich úrovně předzpracování dat užitých pro výběr zájmových území v rámci zájmové oblasti B, pro zpracování časových řad a také pro klasifikace.

4.2.1. GlobCover: Global Land Cover Map (ESA)

Sada poskytující informace o krajinném pokryvu v globálním měřítku. Kompozity jsou vytvářeny na základě dat ze senzoru MERIS (Medium Resolution Imaging Spectrometer) Level 1B na palubě satelitu ENVISAT, který disponuje rozlišením 300 m/pixel (ESA 2010). Tento datový zdroj byl vybrán pro rozlišení více typů krajiny v zájmovém území. GlobCover byl vytvořen pro rok 2009, vydán byl o rok později (ESA 2010). Poskytován je společností ESA (European Space Agency, Evropská kosmická agentura).

4.2.2. CHIRPS

CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) je datová sada vyvinutá pracovníky z USGS (United States Geological Survey), která poskytuje informace o vydatnosti srážek (Funk a kol. 2015). Data jsou nabízena od roku 1981 do

současnosti s rozlišením $0,05^\circ$, přičemž model je vyvíjen od roku 1999 a až zpětně z dostupných dat dopočítán (Funk a kol. 2015). Pro tvorbu modelu jsou využívána naměřená data z meteorologických stanic, mimo jiné i odhady srážek na bázi satelitních dat od NASA (National Aeronautics and Space Administration) a NOAA (National Oceanic and Atmospheric Administration) (Funk a kol. 2015). Katalog GEE nabízí dvě sady: denní záznamy a pětidenní kolekce. K dispozici je i sada CHIPR, která však oproti sadě CHIRPS nepracuje s pozemními daty.

4.2.3. NASADEM

Sada NASADEM byla použita pro splnění požadavku na sklonitost terénu. Jedná se o přepracovaná a vylepšená data SRTM (Shuttle Radar Topography Mission), kde byla zvýšena přesnost pomocí několika datových sad: ASTER (Advanced Spaceborne Thermal and Reflection Radiometer), GDEM v2 (Global Digital Elevation Model), ICESat, GLAS (Geoscience Laser Altimeter System) a PRISM (Panchromatic Remote-Sensing Instrument for Stereo Mapping) (Crippen a kol. 2016). NASADEM je nabízen v dlaždicích o velikosti $1^\circ \times 1^\circ$ pro 60° severní z. š. a 56° jižní z. š., což je přibližně 80 % celého povrchu Země (Crippen a kol. 2016). Na vývoji se podílela především NASA, dále NGA (National Geospatial-Intelligence Agency) a vesmírné mezinárodní agentury Německa a Itálie.

4.2.4. JRC Global Surface Water

Datový soubor obsahující informaci o povrchové vodě od roku 1984 po současnost disponující 30 m rozlišením. Skládán je z téměř 4,5 milionů snímků misí Landsat 5, 7 a 8 (Pekel a kol. 2016). Každý pixel je vyhodnocen jako voda nebo povrch bez vody a následně uložen. Ze získaných informací se vytváří model detekce změn, na základě kterého se jednotlivé pixely třídí do kategorií dle sezónnosti (Pekel a kol. 2016). V práci bylo pracováno se všemi kategoriemi.

4.2.5. NASA Global biomass carbon density

Kolekce snímků poskytující informace o nadzemní a podzemní hustotě biomasy pro rok 2010. K dispozici je pro celý svět v rozlišení 300 m/pixel (Spawn a kol. 2020). Hodnota pro

nadzemní biomasy (tedy „zelenou“ složku) je získávána kombinací map nadzemní biomasy a empirických modelů pracujících s LC. Mimo konkrétních hodnot produkce biomasy (udávané v Mg C/ha) je uváděna také hodnota nejistoty měření pro každý pixel (Spawn a kol. 2020).

4.2.6. ERA5-Land ECMWF climate reanalysis

Datová sada poskytující konzistentní informace pro sledované klimatické proměnné v průběhu několika desetiletí disponující větším rozlišením ($0,1^\circ$) než sesterská sada ERA5 (Muñoz Sabater 2019). Sada je tvořena pomocí klimatické „reanalýzy“ dat ECMWF ERA5 a kombinuje modelová data s pozorováními z celého světa. Poskytuje přes 50 klimatických proměnných ve formě měsíčních průměrů, přičemž v této práci jsou využívány údaje o teplotě (Muñoz Sabater 2019).

4.2.7. USGS Landsat 8 Level 2, Collection 2, Tier 1

Pro TS byla využita kolekce Landsat 8 Surface Reflectance (odrazivost povrchu), která byla získána senzory OLI/TIRS. Snímky obsahují dohromady 9 pásem, přičemž 5 pásem zahrnuje viditelné a blízké infračervené záření (VNIR), 2 pásma krátkovlnné infračervené záření (SWIR), a jedno pásmo termální infračervené záření (TIR) (Vermote a kol. 2016). Sada je doplněna o pásmo QA. Pomocí tzv. QA pixelu (generovaného algoritmem CFMask), který udává kvalitu pixelu, se provádějí filtrace mraků, stínů, sněhu či aerosolu. Data Landsat 8 SR byla vytvořena pomocí algoritmu LaSRC (Vermote a kol. 2016).

4.2.8. OpenStreetMap

Polygonová data budov ve formátu *shapefile* byla stáhnuta z databáze OpenStreetMap, která je spravovaná a vytvářena dobrovolníky z celého světa, vytvořena byla pro rok 2013 (OpenStreetMap 2013).

4.2.9. Data pro sběr trénovacích a validačních dat ke klasifikaci a doplňková data

Pro sběr trénovacích a validačních dat bylo využito volně dostupných dat v rámci desktopového prostředí Google Earth Pro, které obsahuje různá družicová data s velmi

vysokým rozlišením. Konkrétně byla využita data CNES / Airbus (0,5 m prostorové rozlišení), Maxar Technologies (0,3 m prostorové rozlišení) a Copernicus Sentinel-2 (10 m prostorové rozlišení).

Dále byl pro doplňková data (posuny a kontroly pozic lokalit) využit software QGIS a plugin QuickMapServices, který umožňuje snadné načítání podkladových map od různých poskytovatelů pomocí standartu WMS. Mezi známější patří Bing, ESRI, Google či OSM a spoustu dalších. Výsledkem bylo několik vrstev s různými zdroji dat (snímků), jako je CNES / Airbus, Copernicus Landsat, Maxar Technologies, TerraMetrics.

4.3. Zpracování dat

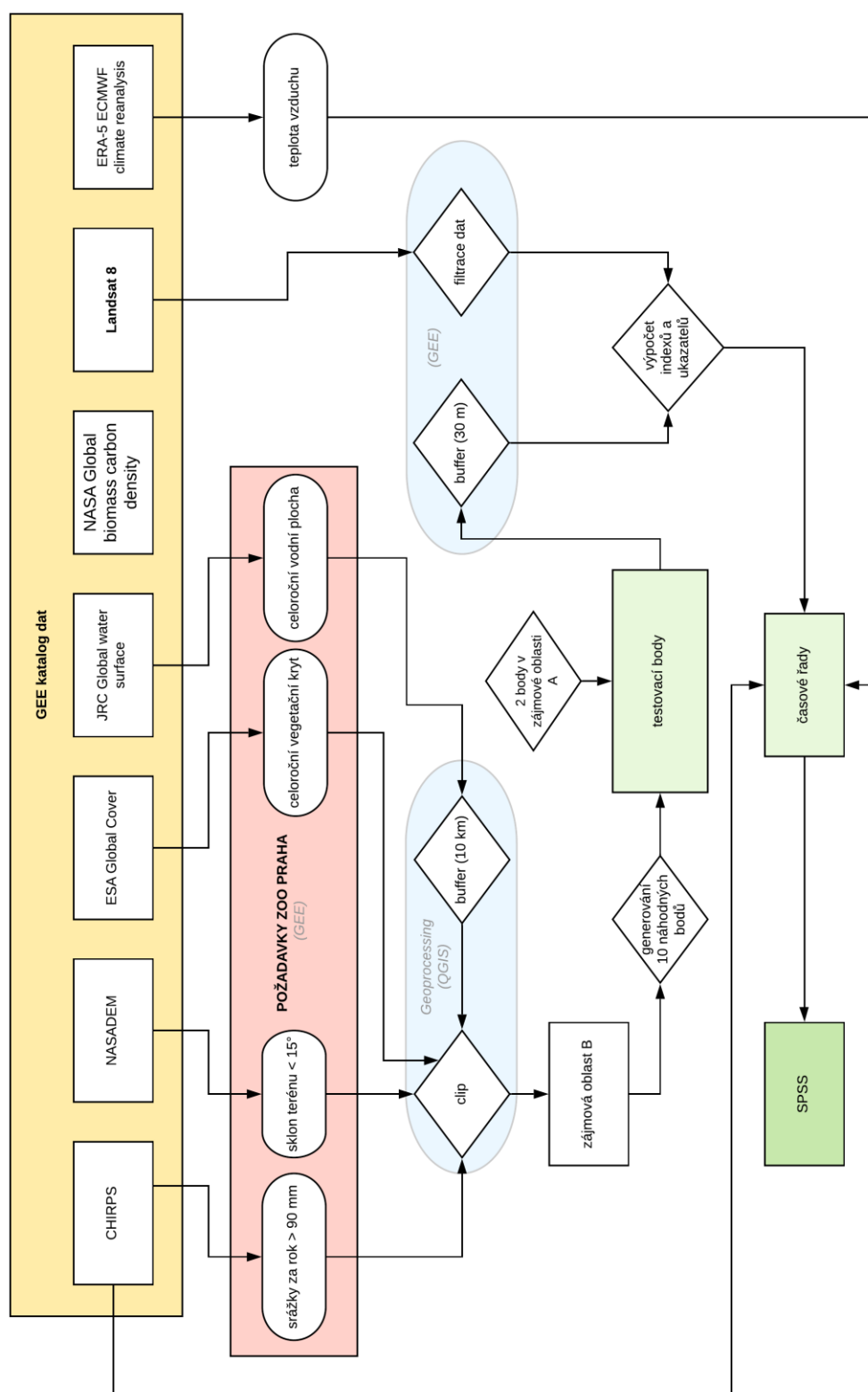
V prvním kroku bylo nutné na základě požadavků Zoo Praha vybrat ve zvolené zájmové oblasti B (pro potenciální reintrodukcii) 10 náhodně vygenerovaných lokalit, reprezentující vhodné budoucí lokality k reintrodukcii. V rámci těchto vybraných 10 lokalit bylo provedeno bližší šetření pomocí časových řad družicových snímků mise Landsat. Níže jsou uvedeny požadavky Zoo Praha pro nalezení vhodných potenciálních míst k reintrodukcii koně Převalského a k následnému dalšímu šetření v rámci vybrané zájmové oblasti B:

1. průměrné roční srážky větší jak 90 mm
2. území pouze s vegetačním krytem
3. sklonitost terénu max. 15°
4. vzdálenost od vodního útvaru max. 10 km

Z původní oblasti zájmu A Great Gobi B byly od Zoo Praha obdrženy a následně do studie přidány 2 vybrané lokality, kde se kuň Převalský často vyskytoval po již úspěšné reintrodukcii. Tyto body byly ručně posunuty do nejbližšího místa s travním pokryvem s nejvyšším množstvím biomasy pro shodu s vybranými lokalitami ze zájmové oblasti B.

Pomocí GEE skriptů se pro těchto celkem 12 vybraných lokalit sestavily časové řady vybraných vegetačních indexů a též indexu pro detekci sněhu. Následně proběhla jejich statistická analýza a získané výsledky byly mezi sebou porovnány s cílem nalezení nejvhodnější lokality v zájmové oblasti B na základě rozdílnosti s původními územími, kde reintrodukce již úspěšně proběhla. Na obrázku 13 je vidět schéma postupu při tvorbě časových řad. Časové řady následně byly statisticky zhodnoceny a doplněny i o klasifikace obou

zájmových oblastí. Vytvořeny byly i mapy reprezentující množství biomasy a srážek, opět pro obě zájmové oblasti.

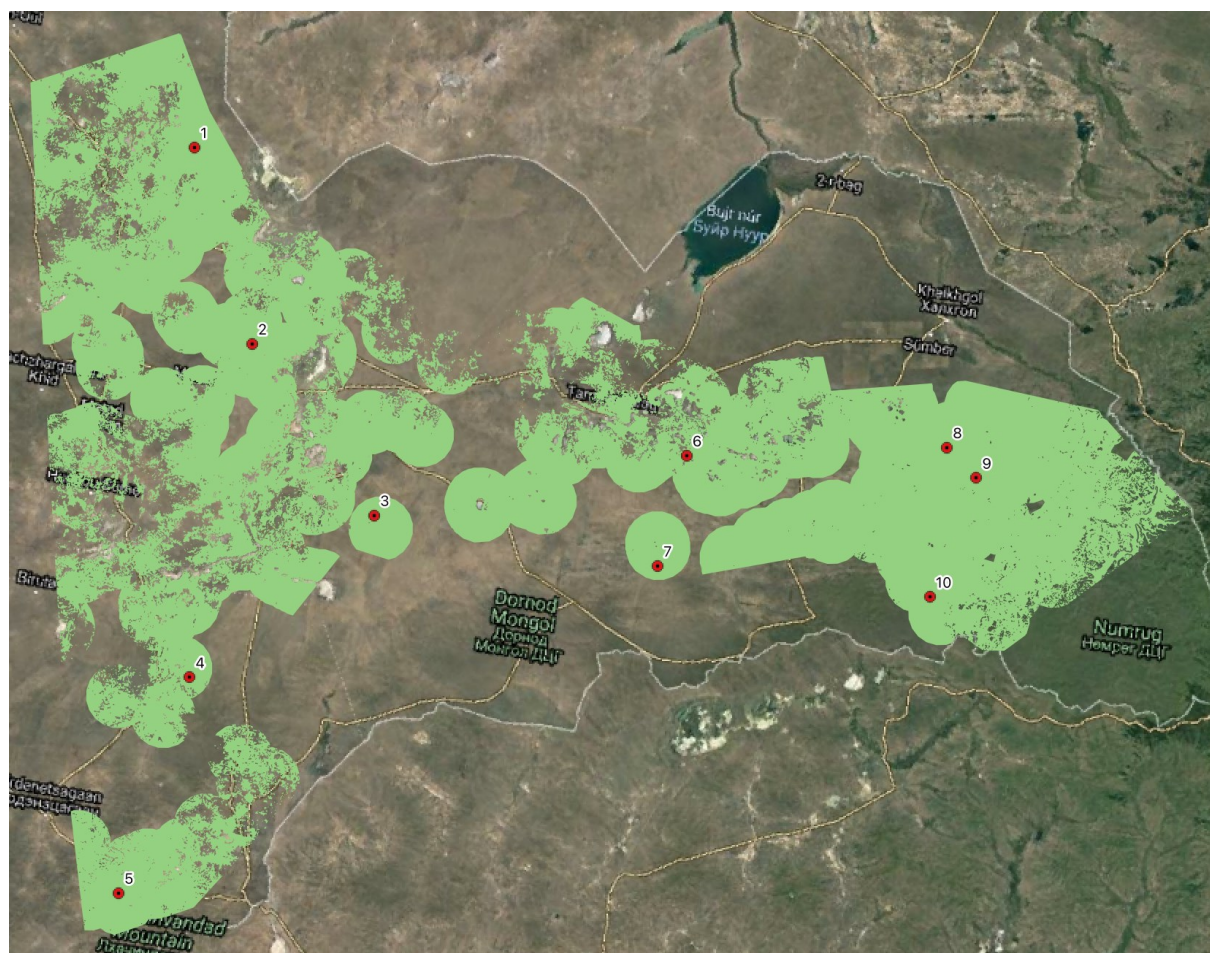


Obrázek 13: Diagram postupu při tvorbě analýz

(zdroj: vlastní tvorba)

4.3.1. Generování náhodných bodů

Na základě výše uvedených požadavků Zoo Praha bylo v rámci sledované zájmové oblasti B vybráno potenciální území, ve kterém bylo náhodně vygenerováno 10 bodů, tj. lokalit pro oblast Eastern Steppe, v rámci nichž byly splněny požadované podmínky pro výskyt jedinců koně Převalského. Body byly generovány pomocí funkce `ee.FeatureCollection.randomPoints` v prostředí GEE. V dalším kroku byly body ručně přemístěny tak, aby se nacházely pouze na jednom typu krajinného pokryvu, a zároveň nebyly na jejich hranici. Pokud docházelo k výraznému shluku bodů, byly taktéž ručně přesunuty, a to na základě konzultace se Zoo Praha. K tomu byl využit software QGIS a doplněk QuickMapServices. Výsledné území s 10 vygenerovanými a upravenými body, představující lokality pro další analýzy, lze vidět na obrázku 14.



Obrázek 14: Výsledné zájmové území s 10 vygenerovanými body

(zdroj: vlastní tvorba)

4.3.2. Výpočet indexů a vybraných ukazatelů

Pro sledování trendů v zájmových oblastech a jejich lokalitách byly stanoveny 4 hlavní indexy vypočítané z dat Landsat 8: (1) NDVI – Normalized Difference Vegetation Index, (2) EVI – Enhanced Vegetation Index, (3) NDMI – Normalized Difference Moisture Index, (4) NDSI – Normalized Difference Snow Index. A dále pak surové hodnoty teplot a srážek z dat CHRIPS a hodnoty teplot z dat ERA5-Land.

NDVI je poměrně často využívaným indexem pro sledování zdravotního stavu vegetace. Využíváno je blízké infračervené (NIR, Near Infra Red) a červené pásmo (Red). Zdravá vegetace má většinou vysoký podíl chlorofylu (zeleného barviva) v listech, který způsobuje pohlcení záření v červené části spektra. Zároveň má vegetace v oblasti NIR velmi vysokou odrazivost díky buněčné struktuře. U poškozené vegetace jsou hodnoty odrazivosti přesně naopak, zejména kvůli poškozené buněčné struktuře a absenci chlorofylu. Hodnoty indexu se pohybují v intervalu od -1 do +1, přičemž záporné a nízké hodnoty indexu reprezentují poškozenou a neživou vegetaci či jiné povrchy. Pro zelenou vegetaci jsou běžné hodnoty u dat Landsat typicky v rozmezí od 0,3 do 0,9 (v závislosti na typu vegetace a zdravotním stavu) (Rouse a kol. 1974).

$$\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})}$$

Původně byl vyvinut pro data MODIS s cílem vylepšit NDVI pomocí optimalizace signálu v oblastech s vysokým indexem listové plochy LAI (Leaf Area Index). EVI je nejvíce užitečný v oblastech s vysokým LAI, kde může docházet k nasycení NDVI. Využívána je modrá část (Blue) odrazivosti ke korekci povrchu a k redukci atmosférických vlivů, včetně aerosolů (Sims a kol. 2006).

$$\text{EVI} = 2,5 * \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + 6 * \text{Red} - 7,5 * \text{Blue} + 1)}$$

Hodnoty se pohybují v rozmezí od 0 do 1. Světlé prvky, jako jsou mraky a bílé zástavba spolu s tmavými prvky jako je voda, mohou vést k výraznému zkreslování hodnot, proto by měli být vždy odmaskováni (Huete a kol. 2002).

Úroveň vodního stresu vegetace reprezentuje NDMI. Hodnoty se pohybují stejně jako u všech normalizovaných indexů od -1 do +1. Používá se především ke sledování sucha, ale v některých případech také k monitorování množství paliva v oblastech, které jsou náchylné

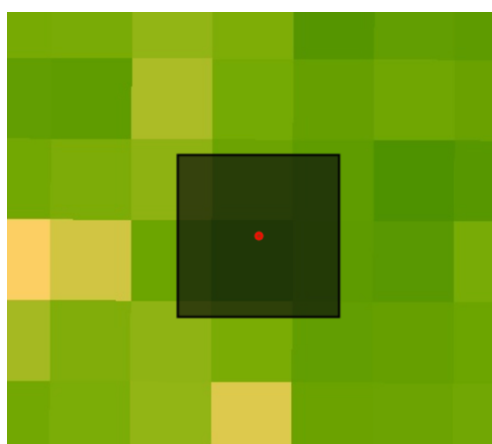
na požáry (Wilson, Sader 2002, Skakun, Wulder, Franklin 2003). K výpočtu se používají pásma NIR a pásmo z krátké infračervené části spektra (SWIR).

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

Využívá se k detekci přítomnosti sněhu či ledu. Jelikož se jedná opět o standardizovaný index, hodnoty se tedy pohybují od -1 do +1. Při výpočtu se používají pásma zeleného (Green) a SWIR záření. Na vlnových délkách 0,66 a 1,6 mm je atmosféra průhledná a zároveň sníh vykazuje malou (1,6 mm) a vysokou (0,66 mm) odrazivost (Salomonson, Appel 2004).

$$\text{NDSI} = \frac{(\text{Green} - \text{SWIR})}{(\text{Green} + \text{SWIR})}$$

K výpočtu byla využita opět platforma GEE. Pro výpočet indexů a ukazatelů srážek / teplot byl vytvořen kolem bodů *buffer*, aby se zabránilo možnému zkreslení výsledků, protože jeden konkrétní pixel nemusí být pro dané zájmové území dostatečně reprezentativní. Vytvořená funkce *bufferPoints* má dva parametry: (1) *radius* (number) – poloměr bufferu v metrech, (2) *bounds* (boolean) – tvar bufferu, čtverec či kruh. V práci se pracovalo se čtvercovým buffrem o velikosti 30 m, aby byly pokryty všechny okolní pixely, jelikož jeden pixel má rozlišení právě 30 m (obrázek 15). Všechny časové řady pro daný bod byly tvořeny váženým průměrem hodnot ze všech zasažených pixelů, kde váhou byla plocha daného pixelu.

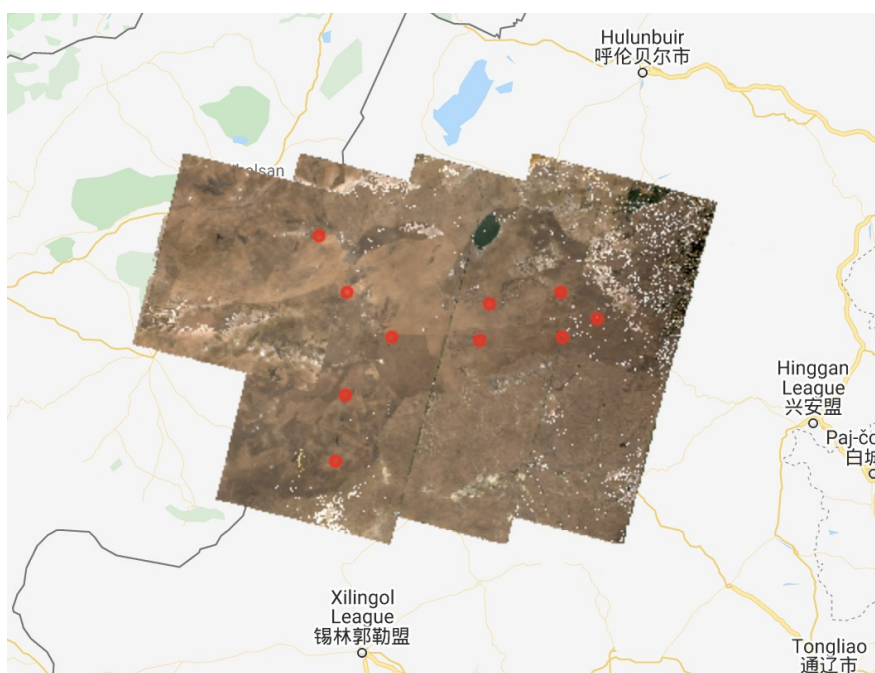


Obrázek 15: Vizualizace funkce *bufferPoints*

(zdroj: vlastní tvorba)

Následně bylo potřeba u dat Landsat 8 odmaskovat mraky, sníh a stíny, což zajistila funkce *fmask*. V případě výpočtu sněhového indexu se sníh nemaskoval. Jelikož Landsat 8

nabízí snímky až od roku 2013, nemusel se vybírat konkrétní časový interval a použily se téměř všechny dostupné (do konce roku 2020), přičemž pro NDVI, EVI a NDMI bylo vybráno pouze vegetační období v roce, tedy červen-září. Pro zachování výkonu výpočtu se kolekce ořezala na území pouze v okolí bodů, tj. pracovalo se pouze se snímky, které zahrnovaly body znázorňující středy zájmových území (obrázek 16). V dalším kroku se pro každou lokalitu dle vzorců vypočítalo nové pásmo, jenž představovalo hodnoty daného indexu a podle něj bylo také pojmenováno.



Obrázek 16: Výběr pouze snímků zahrnující zájmové lokality (2015, červen), oblast Eastern Steppe

(zdroj: vlastní tvorba)

Některé lokality se nacházely na rozmezí více dlaždic, což způsobovalo chybu zdvojených hodnot pro stejný den, a proto se z původní kolekce vypočítala nová, tvořena průměry pro každý den.

U dat teplot a srážek se jednalo o surové hodnoty získané z dostupných archivů v GEE. Z dostupných dat 5ti denních průměrů byly vytvořeny měsíční průměry. Data byla opět vygenerována za jednotlivé lokality ve vybraných obdobích.

Hodnoty za jednotlivá období a lokality se vyexportovaly z GEE do formátu *csv* na Google Drive, z kterého bylo možné soubory stáhnout a dále provádět analýzy, tedy vytvořit časové řady v prostředí Excel. U výsledných časových řad bylo využito klouzavého průměru pro možnost vizuálně analyzovat porovnat trendy jednotlivých křivek.

4.3.3. Statistická analýza dat

Výsledné časové řady byly analyzovány v prostředí SPSS, ve kterém byly porovnány měření ze všech 12 lokalit vybranými statistickými metodami. Před samotným testováním byl využit Shapirův-Wilk test pro analýzu normality dat. Některé lokality nevykazovaly normální rozdělení a z tohoto důvodu byl pro testování využit neparametrický Mann–Whitney test, též nazýván jako Wilcoxon rank-sum test, jehož snahou bylo prokázat rozdíly mezi odlišnými lokalitami, a naopak nalézt nejmenší rozdíly mezi podobnými zájmovými územími, ideálně pak mezi lokalitami s již reintrodukovanou populací a vhodně vybranými body pro budoucí reintrodukcí. Stanovená hladina významnosti byla $\alpha = 0,05$.

U dat srážek bylo zjištěno velké množství odlehlých hodnot, proto u tohoto ukazatele nebylo provedeno testování a bylo nahrazeno analýzou formou boxplot grafů a také tabulkou souhrnných hodnot (pro 5ti denní a měsíční úhrny), dále i mapami průměrných srážek pro obě oblasti (z roku 2010 pro porovnání s jinými ukazateli).

4.3.4. Vytvoření map množství produkce biomasy

Vyinterpolovaná data NASA Global biomass carbon density byla využita pro vytvoření produkce množství nadzemní biomasy. Data jsou dostupná pouze za rok 2010, proto nebylo možné využít tento typ dat k časovým řadám a ze získaných dat byly vytvořeny pro obě oblasti pouze mapy.

4.3.5. Zpracování klasifikací

Pro doplnění celkového obrazu obou porovnávaných oblastí byla doplněna i řízená klasifikace. Jako vhodný klasifikátor byl vybrán klasifikátor CART. Pro samotnou klasifikaci byl v prostředí GEE vytvořen kompozit z dat Landsat 8 metodou mediánu z období květe-září 2020. Pomocí dat z OpenStreetMap, které byly naimportovány do GEE, bylo provedeno odmaskování zástavby. Následně byla stanovena klasifikační legenda pro 4 vybrané typy land cover: (1) voda

(2) holá půda

(3) holá půda / traviny

(4) traviny / hustá vegetace

Do třídy „voda“ spadají veškeré vodní plochy, které byly dostatečně velké, aby mohly být pomocí dat Landsat 8 rozpoznány. Třidu „holá půda“ tvoří především písek a skály, přičemž může obsahovat i suché a velmi řídké traviny. Třetí třída reprezentuje přechod mezi travinami a holými půdami. Tvořena je hustším travnatým porostem, který je tvořen trsy či většími souvislými plochami, přičemž se mohou vyskytovat i jednotlivé nízké křoviny. Čtvrtá třída je tvořena přechodem mezi travinami a hustší vegetací. Náleží zde husté traviny, často doplněné křovinami s většími listy. V konkrétních místech může být tvořena i stromy.

Trénovací data byla vybrána pro každou oblast zvlášť, vybrána byla pomocí softwaru QGIS. V prostředí GEE byly převedeny na polygony pomocí bufferu o šířce 90 m. Vstupní data byla v prostředí GEE analyzována a odlehlé hodnoty byly odstraněny. Po provedené klasifikaci byly provedeny posklasifikační úpravy, konkrétně byl využit majoritní filtr a minimální mapovací jednotka (MMU) byla stanovena na území zahrnující 3 pixelů, což představuje 0,27 ha. Území menší, než MMU byla rozpuštěna do okolí.

Následně byly nasbírána validační data a provedeno hodnocení přesnosti. Výsledný raster klasifikace byl vyexportován z GEE do ArcMap Desktop, kde bylo vytvořeno metodou Stratified Sampling (Manakos a kol. 2018) 500 bodů, které byly validovány opět pomocí Google Earth Pro. Následně byly body naimportovány zpět do GEE, kde bylo provedeno hodnocení přesnosti pomocí chybové matice (Stehman 1997) a Kappa indexu (Foody 2002).

5 Výsledky

5.1. Časové řady

Lokality s reintrodukovanou populací koně Převalského v zájmové oblasti A (Great Gobi B) jsou ve výsledcích označovány jako AA a AB. Pro všechny potenciální lokality k reintrodukcii bylo provedeno porovnání výsledků časových řad pomocí statistického testování. Před samotným testováním byla zjištěna normalita dat. Pro tyto účely byl využit Shapiro-Wilk test, který odhalil normální rozdělení pouze v některých případech (viz tabulka 1).

Tabulka 1: Výsledky Shapiro-Wilkova testu pro jednotlivé lokality

	p-hodnota			
	NDVI	EVI	NDMI	NDSI
1	0,554*	0,001	0,387*	1,000E-06
2	0,797*	0,003	0,455*	0,115*
3	0,281*	0,017	0,247*	0,081*
4	0,007	0,001	0,05*	3,220E-04
5	0,342*	0,016	0,276*	4,430E-04
6	0,791*	0,119*	0,306*	2,92E-07
7	0,264*	0,021	0,164*	7,46E-10
8	0,041	0,001	0,452*	0,001
9	0,036	0,001	0,691*	7,500E-05
10	0,334*	0,011	0,085*	3,59E-07
AA	0,001	0,006	0,03	4,900E-05
AB	0,003	0,031	0,088*	1,21E-07

*tučně: signifikantní hodnoty označeny **

(zdroj: vlastní zpracování)

5.1.1. NDVI

Bod AA z oblasti v Great Gobi B vykazoval v rámci indexu NDVI v průměru oproti ostatním sledovaným lokalitám nejmenší hodnoty. Body 8 a 9 pomocí provedeného testu vyšly signifikantně (tabulka 2), lze je považovat za místa s nejslabším vegetačním krytem na východě Mongolska, nicméně v porovnání s původní lokalitou AA jako stejné a tedy dostačující. Naopak bod AB měl v průměru jedny z nevyšších hodnot, a proto lze bod 1, 6 a 7

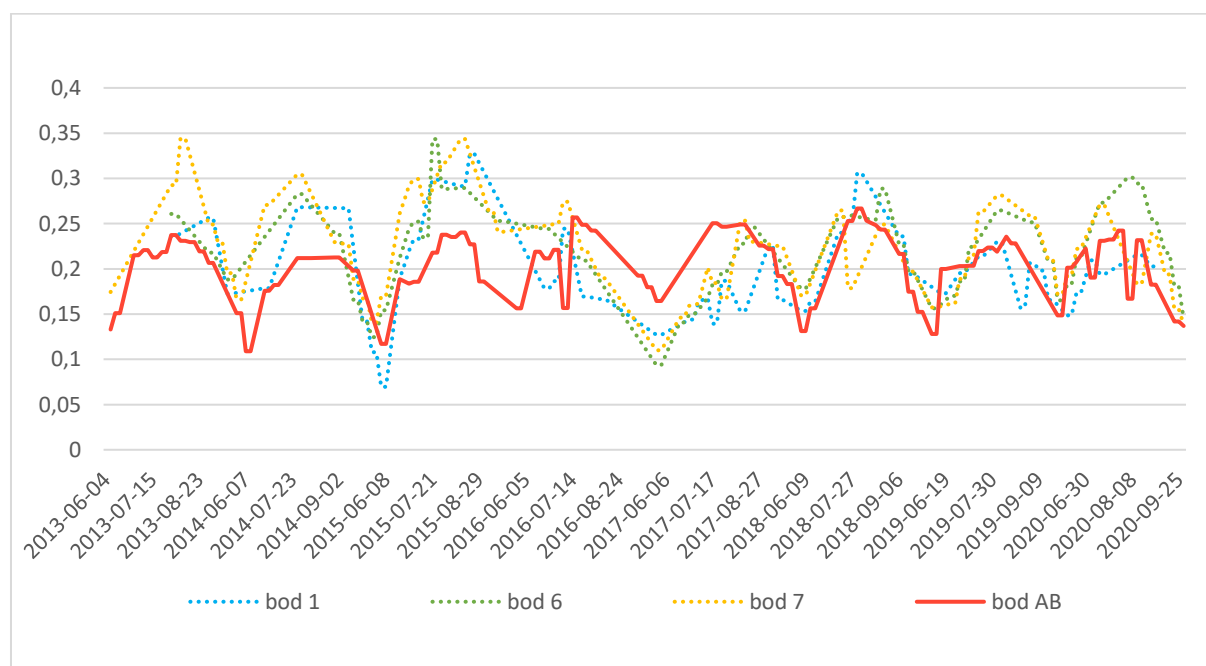
považovat za lokality s nevhodnějším vegetačním krytem (tabulka 2) a stejné, jako je původní lokalita AB. Graf 1 znázorňuje klouzavé průměry hodnot NDVI pro lokalitu AB ze zájmové oblasti A a lokality ze zájmové oblasti B prezentující vybrané lokality vhodné k reintrodukci (lokality byly vybrány na základě statistického testování a průběhu hodnot všech pozorovaných lokalit).

Tabulka 2: NDVI p-hodnoty

id	p-hodnota	
	AA	AB
1	3,149E-07	0,328*
2	0,007	3,200E-05
3	0,001	1,110E-04
4	0,005	2,300E-05
5	0,005	1,30E-05
6	1,704E-12	0,124*
7	6,086E-15	0,751*
8	0,184*	4,495E-13
9	0,099*	6,600E-05
10	0,002	0,001

tučně: signifikantní hodnoty označeny *

Graf 1: Časové řady NDVI



(zdroj: Landsat 8 / vlastní zpracování)

5.1.2. EVI

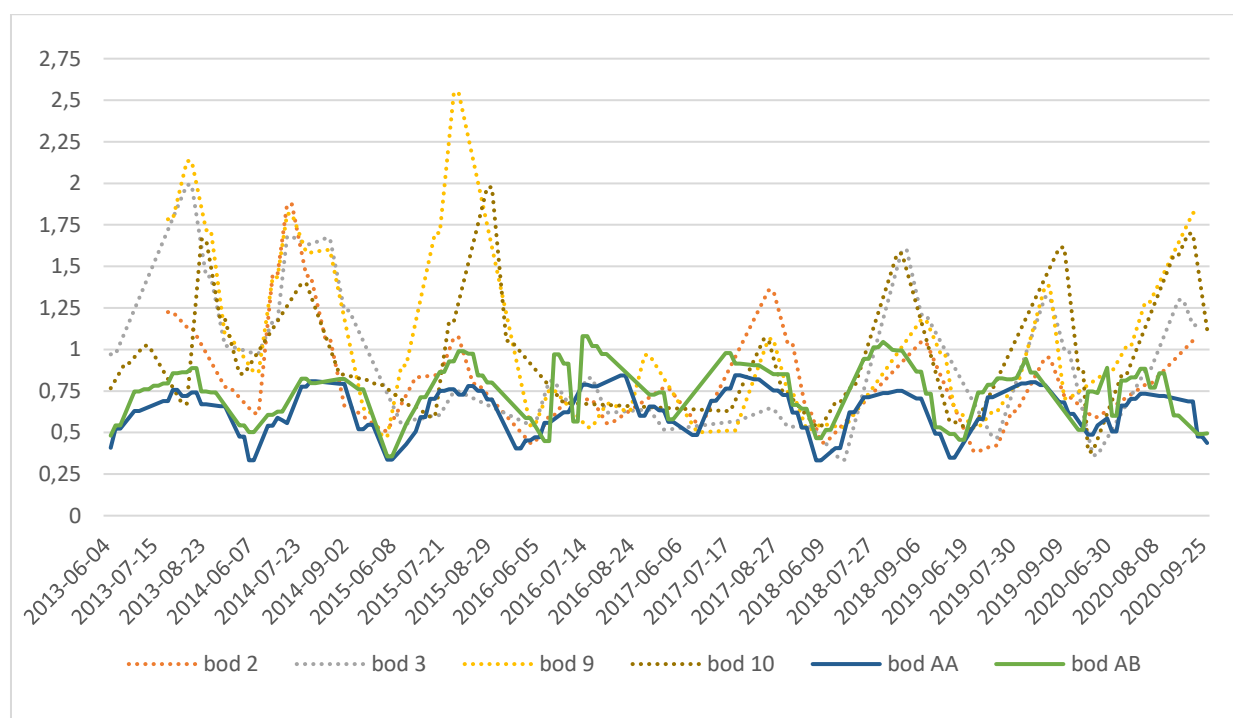
Hodnota EVI pro lokalitu AA byla ve všech případech v průměru menší. Pouze lokalita 8 vykazovala podobné hodnoty. AB se podobala mnohem více sledovaným lokalitám ze zájmové oblasti B, konkrétně bodům 2, 3, 8, 9 a 10 (tabulka 3). V grafu 2 jsou znázorněny klouzavé průměry hodnot EVI původních lokalit ze zájmové oblasti A spolu s lokalitami s ze zájmové oblasti B vhodnými k reintrodukci (lokality byly vybrány na základě statistického testování a průběhu hodnot všech pozorovaných lokalit).

Tabulka 3: EVI p-hodnoty

id	p-hodnota	
	AA	BB
1	6,648E-20	1,759E-16
2	0,001	0,426*
3	0,001	0,217*
4	8,102E-10	2,200E-05
5	8,258E-08	0,001
6	4,550E-21	1,733E-16
7	1,437E-22	3,521E-20
8	0,124*	0,107*
9	0,001	0,095*
10	0,001	0,106*

tučně: signifikantní hodnoty označeny *

Graf 2: Časové řady EVI



(zdroj: Landsat 8 / vlastní zpracování)

5.1.3. NDMI

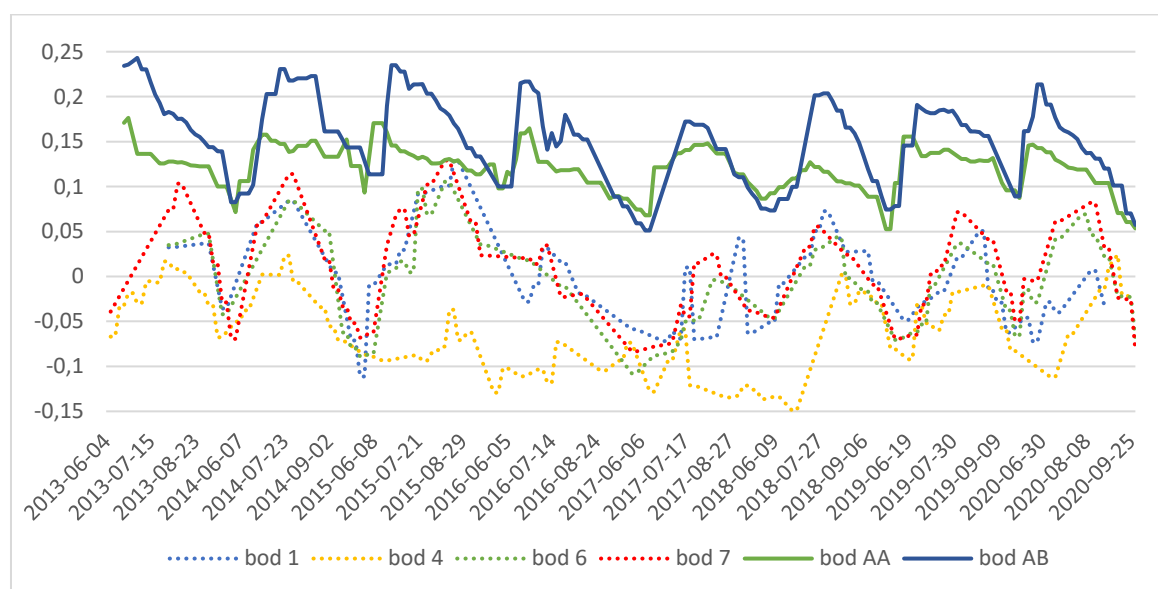
Sledování vlhkosti, respektive úrovně vodního stresu rostlin pomocí NDMI ukázalo, že všechny nové potenciální lokality jsou sušší, než obě původní AA a AB (tabulka 4), vybrané lokality AA a AB se totiž nacházejí v přímé blízkosti vodních toků. To přímo ovlivnilo výsledky měření, nicméně se lokalita AA i AB nacházejí v okolí polopouště, v blízkém okolí byly hodnoty průměrně nižší než v lokalitách oblasti B. Důležité je zmínit, že naměřené hodnoty pro všechny body (oblast A i B) se pohybovaly v intervalu přibližně (-0,2; +0,3). Graf 3 znázorňuje klouzavé průměry hodnot NDMI pro vybrané lokality vhodné k reintrodce v porovnání s již reintrodukovanými oblastmi (lokality byly vybrány na základě statistického testování a průběhu hodnot všech lokalit).

Tabulka 4: NDMI p-hodnoty

id	p-hodnota	
	AA	AB
1	7,440E-36	1,822E-32
2	3,998E-30	6,793E-41
3	7,012E-31	1,725E-41
4	1,313E-62	9,314E-59
5	9,123E-35	5,880E-46
6	2,591E-37	1,504E-36
7	1,574E-33	6,761E-35
8	1,238E-72	9,208E-57
9	2,710E-28	5,054E-42
10	1,142E-29	2,467E-40

tučně: signifikantní hodnoty označeny *

Graf 3: Časové řady NDMI



(zdroj: Landsat 8 / vlastní zpracování)

5.1.4. NDSI

Nejvíce detekované sněhové pokrývky bylo v lokalitě AB s počtem 93 případů. Naopak nejméně případů (23) bylo pozorováno v potenciální lokalitě č. 3 (tabulka 5). Dle výsledků provedených testů se opět podobá potenciálním oblastem lokalita více AB, nízké rozdíly panují zejména v porovnání s lokalitami 2, 3, 4 a 8 (tabulka 6). Graf 4 pak znázorňuje klouzavé průměry vybraných lokalit z obou zájmových oblastí, přičemž kladné hodnoty znázorňují detekci sněhové pokrývky. Jako vhodná lokalita se jeví bod č. 4, u kterého je

detekován malý počet případů sněhové pokrývky. Graf 4 znázorňuje klouzavé průměry hodnot NDSI pro vybrané lokality vhodné k reintrodce v porovnání s již reintrodukovanými oblastmi (lokality byly vybrány na základě statistického testování a průběhu hodnot všech lokalit).

Tabulka 5: Celkový počet detekované sněhové pokrývky

id	N	medián
1	43	0,612
2	26*	0,534
3	23*	0,446
4	30*	0,515
5	40	0,567
6	66	0,599
7	89	0,566
8	68	0,504
9	42	0,606
10	35*	0,606
AA	76	0,281
AB	93	0,509

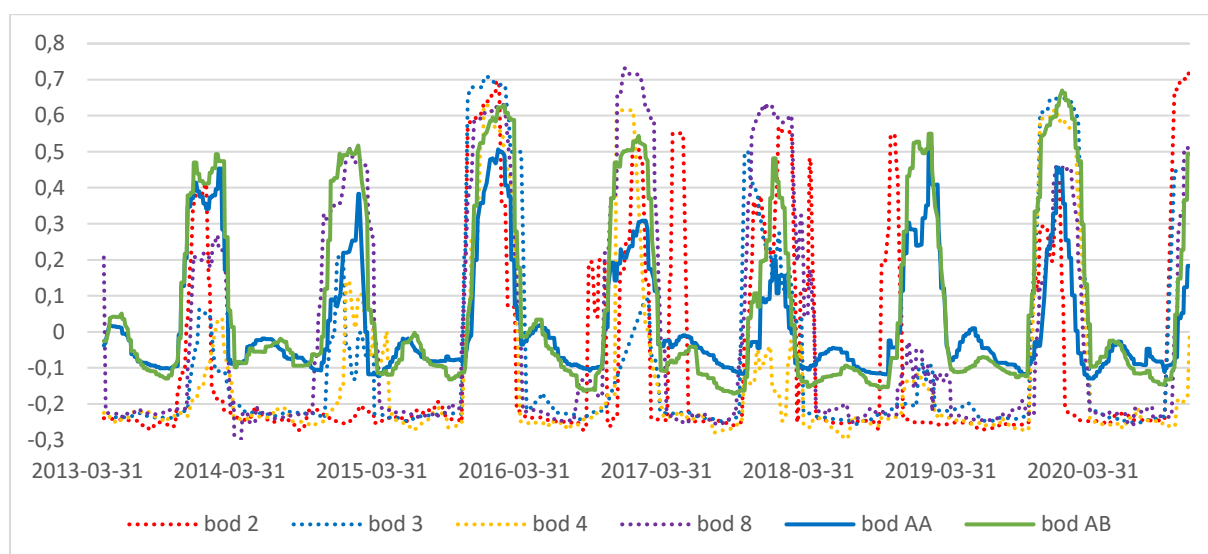
tučně: nejmenší naměřené hodnoty označené *

Tabulka 6: NDSI p-hodnoty

id	p-hodnota	
	AA	AB
1	1,405E-13	1,100E-05
2	2,000E-06	0,354*
3	9,736E-03	0,486*
4	1,700E-04	0,654*
5	8,367E-11	0,013
6	3,468E-16	2,500E-05
7	3,324E-12	0,048
8	2,307E-09	0,468*
9	8,363E-13	8,100E-05
10	4,943E-14	2,000E-06

tučně: signifikantní hodnoty označeny *

Graf 4: Časové řady NDSI

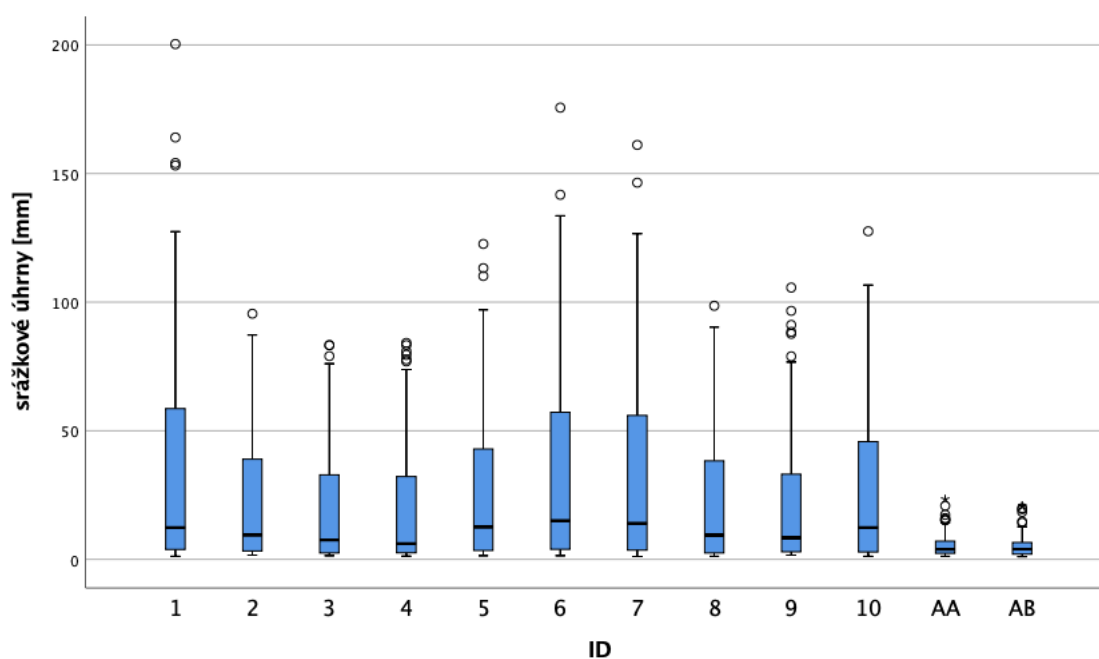


(zdroj: Landsat 8 / vlastní zpracování)

5.1.5. Srážky

Data průměrných srážkových úhrnů vykazovala velký počet odlehlých hodnot (především pro lokality AA, AB), a proto bylo upuštěno od statistických testů. Dle boxplot grafů (graf 5) doplněné o naměřené maxima a graf časových řad je však patrné, že dvě lokality ze zájmové oblasti A disponují mnohem menšími hodnotami, než všechny potenciální lokality v oblasti B. V tabulce 7 lze vidět zmíněná maxima, a to pro 5ti denní a měsíční časové řady, přičemž tyto údaje jsou tvořeny sečtením srážkových úhrnů za dané období, ze kterého jsou vyneseny maximální hodnoty pro danou lokalitu. V grafu 6 je pak možné vidět hodnoty srážek za všechny lokality.

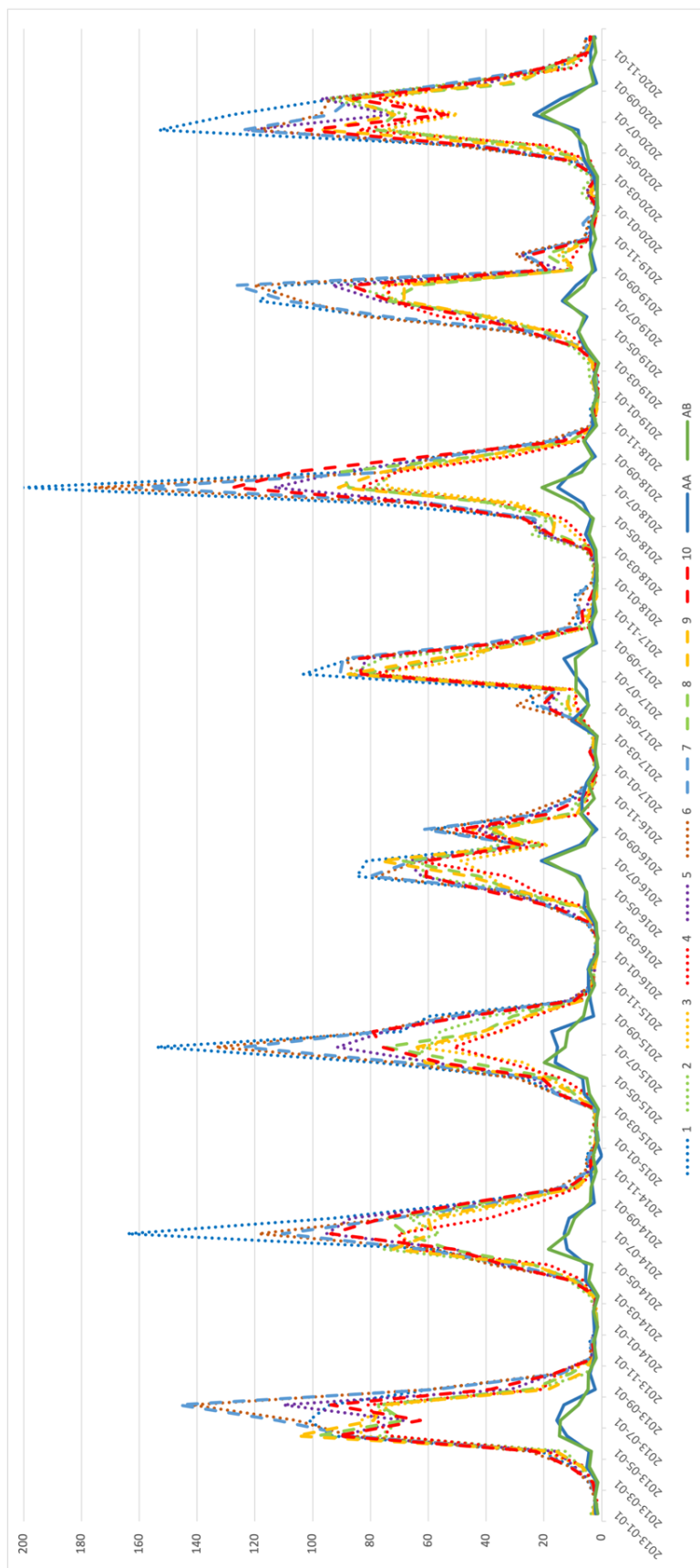
Graf 5: Rozložení dat CHIRPS



Tabulka 7: Maxima srážkových úhrnů pro 5ti denní a měsíční řady

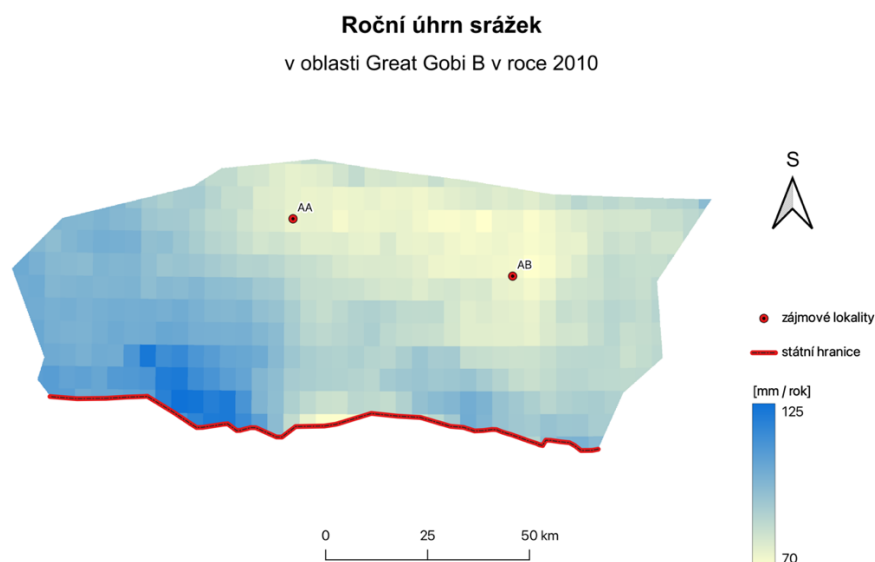
id	5 dní	měsíc
1	53,4	200,3
2	38,7	95,5
3	48,2	83,3
4	33,5	84,2
5	54,5	122,6
6	53,9	175,6
7	55,5	161,1
8	44,8	98,6
9	55,8	105,7
10	47,0	127,5
AA	9,2	23,4
AB	7,8	20,8

Graf 6: Časové řady srážek



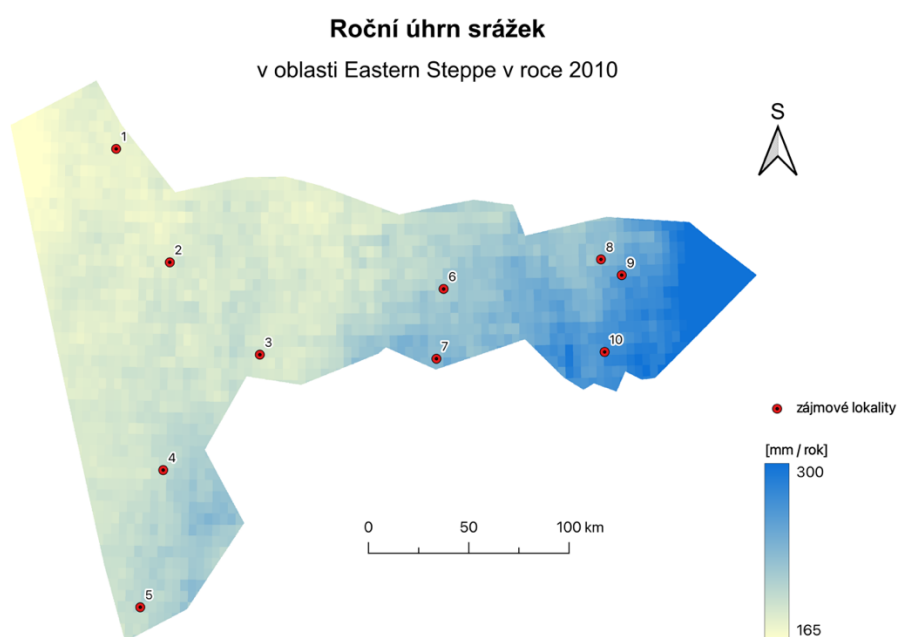
(zdroj: CHIRPS / vlastní zpracování)

V rámci sledování srážkových úhrnů byly sestrojeny i dvě mapy reprezentující roční úhrn srážek pro rok 2010, a to obě zájmové oblasti (obrázek 17, obrázek 18). Z mapy pro zájmovou oblast A je vidět značnější úhrn srážek v jihozápadní oblasti. Obecně však lze říct, že se jedná o oblast velmi suchou. U zájmové oblasti B pak vidíme značnější úhrny srážek, zejména pak ve východní části.



Obrázek 17: Roční srážkové úhrny v Great Gobi B

(zdroj: CHIRPS / vlastní zpracování)



Obrázek 18: Roční srážkové úhrny, Eastern Steppe

(zdroj: CHIRPS / vlastní zpracování)

5.1.6. Teploty

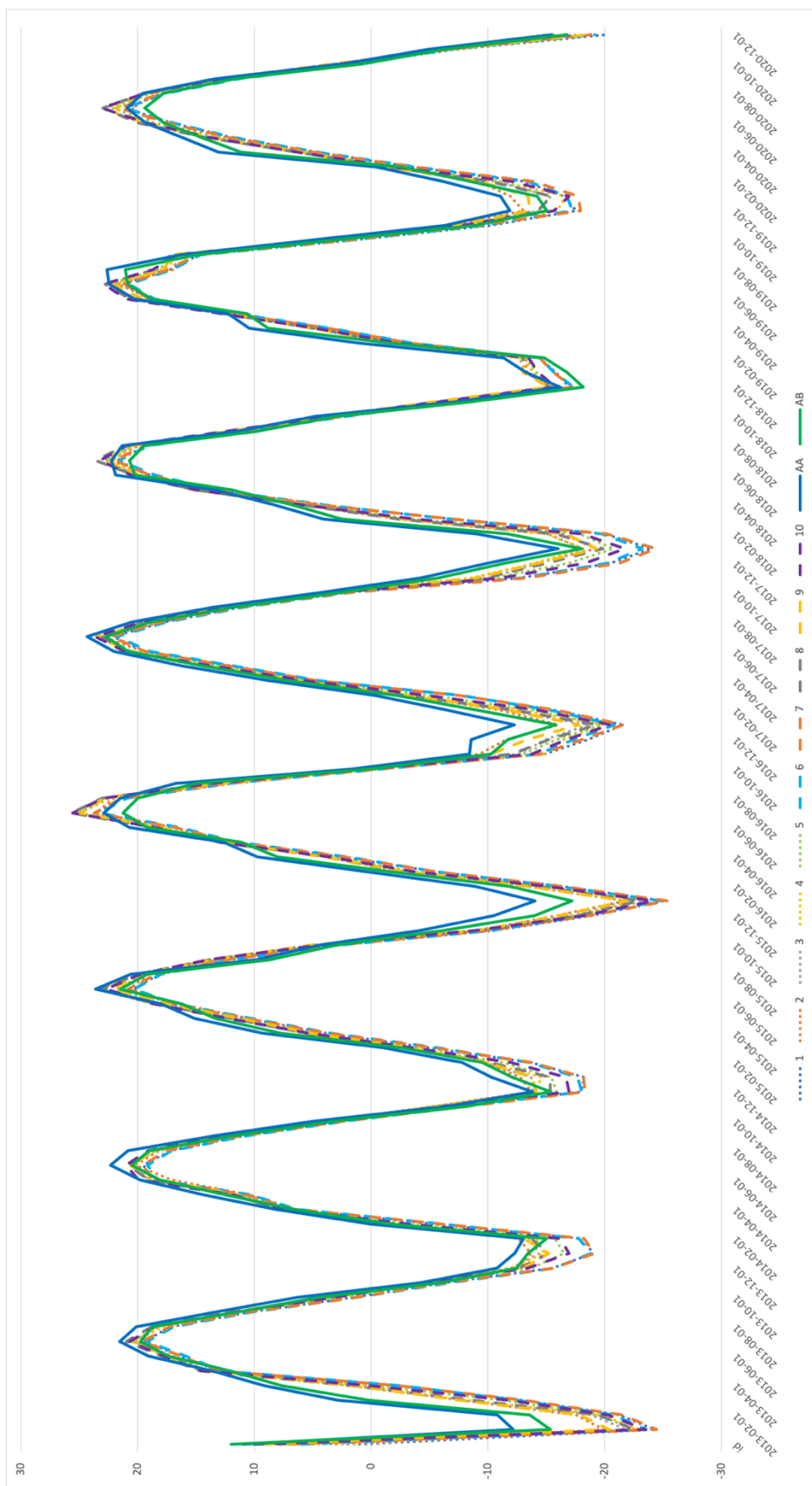
Teplota je téměř ve všech lokalitách velmi podobná. Pouze tři lokality v porovnání s lokalitou AA byly vyhodnoceny jako odlišné (1, 6, 7), přičemž všechny tři zmíněné lokality v zájmové oblasti B vykazovaly hodnoty nižší (tabulka 8). Hodnoty v grafu 8 jsou uvedeny ve °C.

Tabulka 8: Teplota p-hodnoty

id	p-hodnota	
	AA	AB
1	0,030	0,228*
2	0,149*	0,781*
3	0,245*	0,969*
4	0,183*	0,907*
5	0,133*	0,716*
6	0,024	0,189*
7	0,032	0,242*
8	0,249*	0,992*
9	0,222*	0,980*
10	0,184*	0,801*

tučně: signifikantní hodnoty označeny *

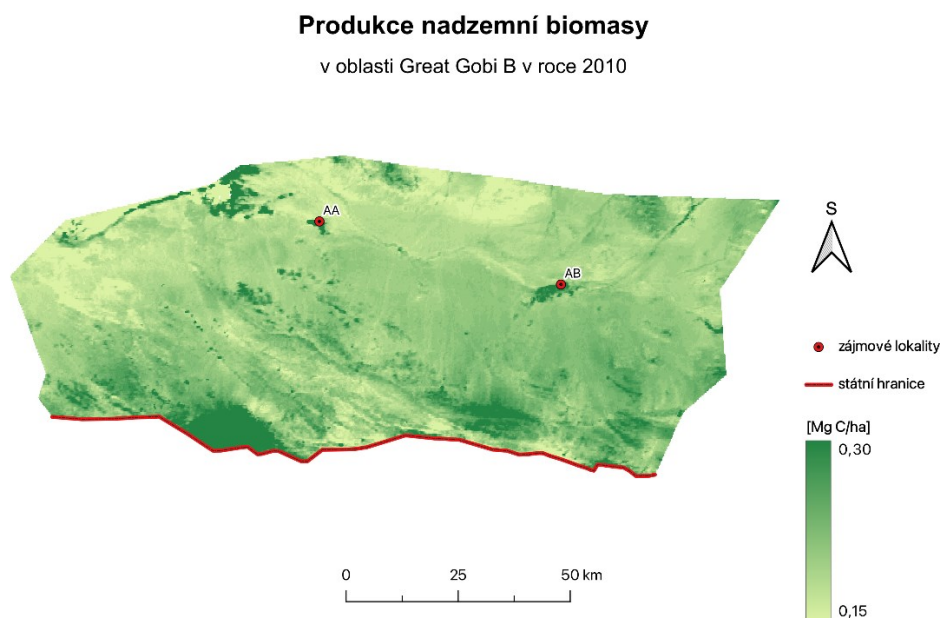
Graf 7: Časové řady teplot



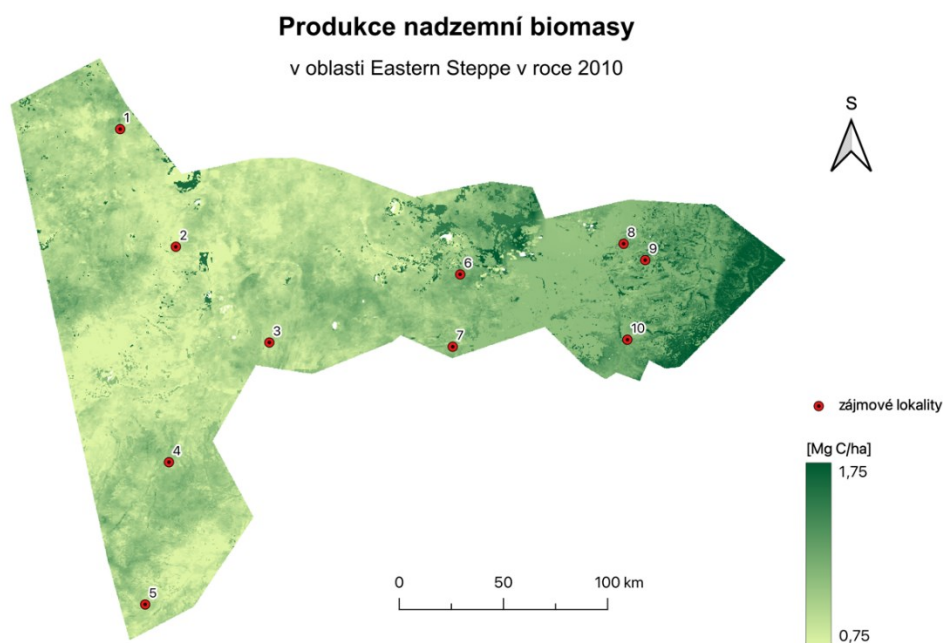
(zdroj: ERA5 / vlastní zpracování)

5.2. Mapy množství biomasy

Na obrázku 19 a 20 je možné vidět produkci nadzemní biomasy. Snímky byly vytvořeny pomocí NASA Global biomass carbon density. Z obou obrázků je vidět řádově nižší hodnoty v zájmové oblasti A oproti zájmové oblasti B.



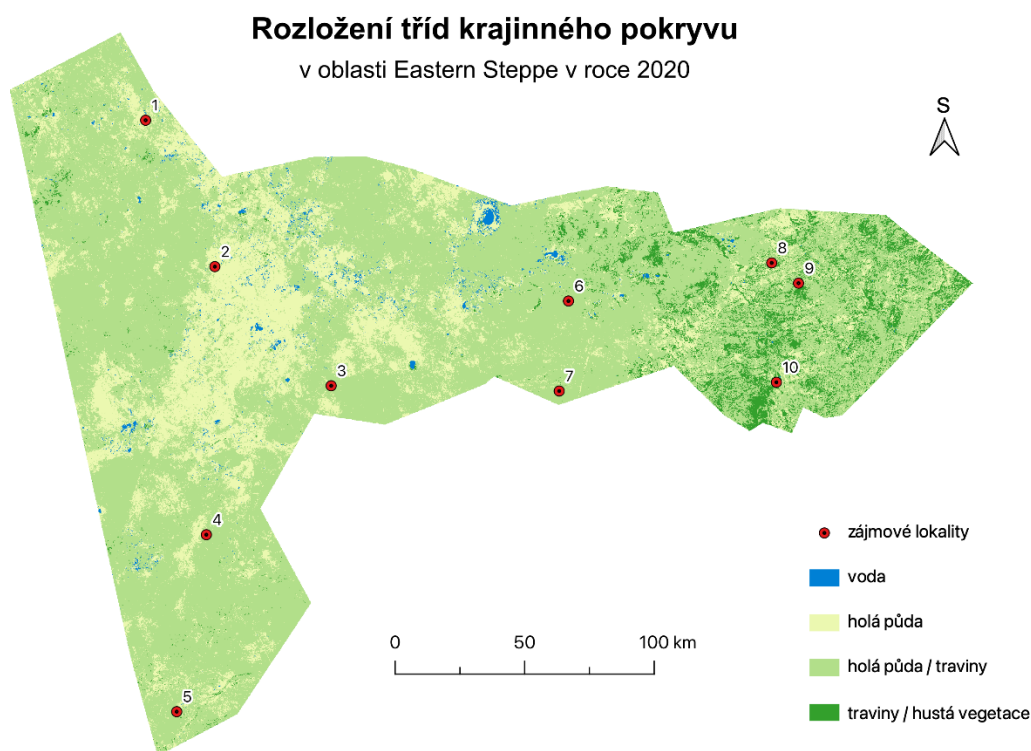
Obrázek 19: Množství biomasy v zájmové oblasti Great Gobi B
(zdroj: NASA Global biomass carbon density / vlastní zpracování)



Obrázek 20: Množství biomasy v zájmové oblasti Eastern Steppe
(zdroj: NASA Global biomass carbon density / vlastní zpracování)

5.3. Klasifikace

Na obrázku 21 a 22 jsou vytvořené klasifikace metodou CART. Na snímcích je vidět že obě oblasti jsou mezi sebou porovnatelné z hlediska land cover. Celková přesnost klasifikace dosahovala pro oblast A 82 % a pro zájmovou oblast B dosahovala 88 %, což lze považovat za uspokojivý výsledek (viz tabulka 9 a 10). Většinu obou území tvoří přechod holé půdy a travin.

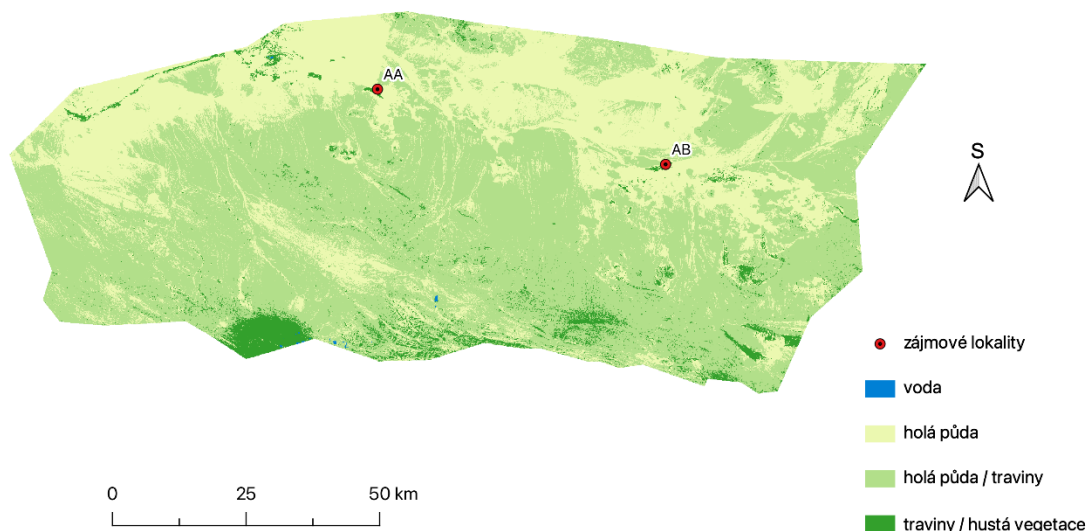


Obrázek 21: Výsledek klasifikace CART pro zájmovou oblast A

(zdroj: Lndsat 8 / vlastní zpracování)

Rozložení tříd krajinného pokryvu

v oblasti Great Gobi B v roce 2020



Obrázek 22: Výsledek klasifikace pro zájmovou oblast B

(zdroj: Landsat 8 / vlastní zpracování)

Tabulka 9: Chybová matice a Kappa index pro zájmovou oblast A

voda	3	0	0	0
holá půda	2	82	23	0
holá půda / traviny	5	37	313	12
traviny / hustá vegetace	0	1	10	16
Kappa index:	0,60	Celková přesnost:		82 %

(zdroj: vlastní zpracování)

Tabulka 10: Chybová matice a Kappa index pro zájmovou oblast B

voda	3	0	0	0
holá půda	0	99	23	0
holá půda / traviny	0	7	153	4
traviny / hustá vegetace	0	0	3	9
Kappa index:	0,76	Celková přesnost:		88 %

(zdroj: vlastní zpracování)

5.4. Stanovení vhodných lokalit

Všechny lokality v zájmové oblasti Eastern Steppe lze pro reintrodukci doporučit, neboť u sledovaných indexů a ostatních parametrů bylo prokázáno, že lokality vykazují minimálně takové hodnoty, jako lokality z původní oblasti Great Gobi B. Často je však převyšují.

Ojedinělé se zdají být lokality č. 1, 6 a 7. V rámci NDVI a NDMI vykazovaly vysoké hodnoty, navíc disponují velkými srážkovými úhrny. Sněhová pokrývka byla v těchto lokalitách detekována oproti lokalitám 2, 3, 4 a 10 častěji, nicméně vzhledem původním lokalitám AA a AB srovnatelně, či méně. V blízkosti lokality 6 i 7 se nachází vodní plochy a nadprůměrná produkce biomasy, což by bylo pro koně Převalského jistě přínosem.

Vybrané datové vrstvy a vytvořená mapová aplikace jsou dostupné na webu www.mongolsko.maweb.eu.

6 Diskuze

U časových řad je často problémem nedostatek dat, respektive nedostatek snímků bez rušivých vnějších vlivů. V případě odmaskování oblačnosti, stínů a sněhu může být v datech i chybějící období mezi hodnotami dosahující až 3 měsíce (např. pro oblast Eastern Steppe). Shen a kol. (2014) uvádí metodu, kdy se chybějící hodnoty nahradí střední hodnotou sousedních hodnot.

Za účelem získání více informací u konkrétních lokací by bylo vhodné zvážit rozšíření indexů (NDVI, EVI, NDMI a NDSI) o další vhodně vybrané. Dále by bylo vhodné otestovat např. pomocí korelační analýzy souvislost těchto indexů, protože např. lokality 1, 6 a 7 stanovené jako nejvhodnější pro reintrodukci, reprezentují se vysokými hodnotami pro NDVI, NDMI a srážky. Pokud by byla prokázána korelace, bylo by vhodné vybrat nejvíce reprezentativní indexy a odstraněné nahradit nově zvolenými. Další cenné informace z TS dat lze získat pomocí pokročilých algoritmů (např. BFAST, LT), které umožňují detekovat pravidelnosti v datech. Vzhledem k charakteru území Mongolska se nabízí i použití většího bufferu pro vybrané lokality, z nichž se počítají indexy.

Provedená klasifikace by mohla být modifikována hned několika způsoby. Vzhledem k nedostatku kvalitních snímků pro dané období, které by nabízely velmi vysoké rozlišení pro celé zájmové území, bylo poměrně obtížné vytvářet trénovací body, ale i vyhodnocovat body validační. Data *in-situ* by pomohla zvýšit celkovou přesnost a poskytla by více informací o krajinných typech. Druhým způsobem, jak vylepšit současné klasifikace území, je použít syntézy dvou typů dat, teda např. Landsat 8 a Sentinel-2.

Cloudová platforma GEE se ukázala jako vhodný nástroj pro zpracování velkého objemu dat. Skripty jsou ukládány pomocí verzovacího systému, díky čemuž lze v případě nutnosti obnovovat jednotlivé pracovní postupy. Nevýhodou pak je uzavřená knihovna funkcí, která se nemůže rozšiřovat o funkce vytvořené samotnými uživateli. Chybí také podrobné popisy některých algoritmů, které by objasňovaly jejich chod.

7 Závěr

Bakalářská práce byla zaměřena pokročilé metody DPZ, konkrétně pak na analýzy časových řad. Hlavním cílem práce bylo nalezení vhodné metodiky pro posouzení vhodnosti vybraných míst v Mongolsku k reintrodukci koně Převalského.

V rámci teoretické části byl představen úvod do problematiky časových řad a také do problematiky cloud based technologií. Značná část práce byla věnována pokročilým metodám časových řad, zejména v rámci volně dostupných balíčků a knihoven do předních programovacích jazyků pro GIS. Část teoretické části byla věnována i analýzám a detekci trendů, především ve spojení s volně dostupnými daty Sentinel-2 a Landsat, které představují vhodná a volně dostupná multispektrální data s vysokým prostorovým, temporálním a radiometrickým rozlišením pro časové řady.

V rámci empirické části se na základě předem definovaných parametrů od Zoo Praha (např. sklonitost, vzdálenost od vody, srážky nebo typ land cover) podařilo vygenerovat potenciální vhodné oblasti, u kterých byly pomocí časových řad sledovány vybrané charakteristiky. Celkem tak bylo vybráno 10 lokalit vhodných k budoucí reintrodukci ve východním Mongolsku, konkrétně v oblasti zvané Numrug. Těchto 10 lokalit pak bylo porovnáno s dvěma místy v západní části Mongolska, kde již v minulosti reintrodukce proběhla. Pro práci byla využita volně dostupná cloud based technologie GEE a také volně dostupná data.

Pro všech 12 lokalit byly vygenerované časové řady vegetačních indexů, srážek, teplot. Pro vzájemné porovnání všech lokalit bylo využito statistického testování. U lokalit byla ověřena normalita dat a vzhledem k tomu, že normální rozdělení bylo dosaženo pouze u některých bodů, bylo nutno provést neparametrické testování metdou Mann Whitney. Na základě výsledků byly stanoveny 3 lokality (konkrétně 1, 6 a 7), které byly vyhodnoceny pro koně Převalského jako nevhodnější.

Pro doplnění pohledu na obě pozorované oblasti byla práce doplněna o klasifikace obou území a také map znázorňující množství biomasy a srážek. Pro klasifikace bylo využito klasifikační metody CART a volně dostupných dat Landsat 8. Celková přesnost klasifikace převyšovala u obou území 80 % a Kappa index 0,6. Všechny vytvořené skripty v prostředí GEE byly uvedeny jako přílohy této práce a mohou posloužit k dalšímu výzkumu oblastí.

8 Literatura

- BAI, L., SHI, C., LI, L., YANG, Y., WU, J. (2018): Accuracy of CHIRPS satellite-rainfall products over mainland China. *Remote Sensing*, 3, 10.
- BANERJEE, A., CHEN, R., MEADOWS, M. E., SINGH, R. B., MAL, S., SENGUPTA, D. (2020): An analysis of long-term rainfall trends and variability in the uttarakhand himalaya using google earth engine. *Remote Sensing*, 4, 12.
- BAO, G., QIN, Z., BAO, Y., ZHOU, Y., LI, W., SANJJAV, A. (2014): NDVI-based long-term vegetation dynamics and its response to climatic change in the Mongolian plateau. *Remote Sensing*, 9, 6, 8337–8358.
- BISSYANDE, T. F., THUNG, F., LO, D., JIANG, L., REVEILLERE, L. (2013): Popularity, interoperability, and impact of programming languages in 100,000 open source projects. *Proceedings - International Computer Software and Applications Conference*, 303–312.
- BRAATEN, J., COHEN, W. B., YANG, Z., ERIK, H. (2016): LandsatLinkr, <https://doi.org/10.5281/zenodo.807733> (11. 7. 2021).
- CHEN, B., XIAO, X., LI, X., PAN, L., DOUGHTY, R., MA, J., DONG, J., QIN, Y., ZHAO, B., WU, Z., SUN, R., LAN, G., XIE, G., CLINTON, N., GIRI, C. (2017): A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*, 131, 104–120.
- CRIPPEN, R., BUCKLEY, S., AGRAM, P., BELZ, E., GURROLA, E., HENSLEY, S., KOBRICK, M., LAVALLE, M., MARTIN, J., NEUMANN, M., NGUYEN, Q., ROSEN, P., SHIMADA, J., SIMARD, M., TUNG, W. (2016): Nasadem global elevation model: Methods and progress. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41, 125–128.
- DINKU, T., FUNK, C., PETERSON, P., MAIDMENT, R., TADESSE, T., GADAIN, H., CECCATO, P. (2018): Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society*, June 2017, 144, 292–312.
- ESA (2010): ESA Data User Element, http://dup.esrin.esa.int/page_globcover.php (18. 7. 2021).
- FOODY, G. M. (2002): Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 1, 80, 185–201.
- FU, T. C. (2011): A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 1, 24, 164–181.
- FUNK, C., PETERSON, P., LANDSFELD, M., PEDREROS, D., VERDIN, J., SHUKLA, S., HUSAK, G., ROWLAND, J., HARRISON, L., HOELL, A., MICHAELSEN, J. (2015): The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, 2, 1–21.
- GEE (2020): Platform – Google Earth Engine, <https://earthengine.google.com/platform/> (13. 7.

- 2021).
- GOMES, V. C. F., QUEIROZ, G. R., FERREIRA, K. R. (2020): An overview of platforms for big earth observation data management and analysis. *Remote Sensing*, 8, 12, 1–25.
- HLADKY, R., LASTOVICKA, J., HOLMAN, L., STYCH, P. (2020): Evaluation of the influence of disturbances on forest vegetation using Landsat time series; a case study of the Low Tatras National Park. *European Journal of Remote Sensing*, 1, 53, 40–66.
- HUETE, A., DIDAN, K., MIURA, T., RODRIGUEZ, E. P., GAO, X., FERREIRA, L. G. (2002): Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 1–2, 83, 195–213.
- JÖNSSON, P., CAI, Z., MELAAS, E., FRIEDL, M. A., EKLUNDH, L. (2018): A method for robust estimation of vegetation seasonality from Landsat and Sentinel-2 time series data. *Remote Sensing*, 4, 10.
- KACZENSKY, P., ENKHTSAIKHAN, N., GANBAATAR, O., WALZER, C. (2008): The great gobi b strictly protected area in mongolia - refuge or sink for wolves cards lupus in the Gobi? *Wildlife Biology*, 4, 14, 444–456.
- KATSANOS, D., RETALIS, A., MICHAELIDES, S. (2016): Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. *Atmospheric Research*, 169, 459–464.
- KENNEDY, R. E., YANG, Z., COHEN, W. B. (2010): Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. *Remote Sensing of Environment*, 12, 114, 2897–2910.
- KENNEDY, R. E., YANG, Z., GORELICK, N., BRAATEN, J., CAVALCANTE, L., COHEN, W. B., HEALEY, S. (2018): Implementation of the LandTrendr algorithm on Google Earth Engine. *Remote Sensing*, 5, 10, 1–10.
- LASTOVICKA, J., SVEC, P., PALUBA, D., KOBLIUK, N., SVOBODA, J., HLADKY, R., STYCH, P. (2020): Sentinel-2 data in an evaluation of the impact of the disturbances on forest vegetation. *Remote Sensing*, 12, 12.
- LEE, J. G., KANG, M. (2015): Geospatial Big Data: Challenges and Opportunities. *Big Data Research*, 2, 2, 74–81.
- LI, Z., LIU, C., ZHANG, P., TIAN, B. (2018): Assessment of snow cover product using google earth engine cloud computing platform. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018-July, 5203–5205.
- MANAKOS, I., TOMASZEWSKA, M., GKINIS, I., BROVKINA, O., FILCHEV, L., GENÇ, L., GITAS, I. Z., HALABUK, A., INALPULAT, M., IRIMESCU, A., JELEV, G., KARANTZALOS, K., KATAGIS, T., KUPKOVÁ, L., LAVRENIUK, M., MESAROŠ, M., MIHAILESCU, D., NITA, M., RUSNAK, T., STYCH, P., ZEMEK, F., ALBRECHTOVÁ, J., CAMPBELL, P. (2018): Comparison of global and continental land cover products for selected study areas in South Central and Eastern European Region. *Remote Sensing*, 12, 10.
- MUEEN, A., KEOGH, E., ZHU, Q., CASH, S., WESTOVER, B. (2009): Exact discovery of time series motifs. *Society for Industrial and Applied Mathematics - 9th SIAM International*

- Conference on Data Mining 2009, Proceedings in Applied Mathematics, 1, 469–480.
- MUNKHJARGAL, M., GROOS, S., PAN, C. G., YADAMSUREN, G., YAMKIN, J., MENZEL, L. (2019): Multi-source based spatio-temporal distribution of snow in a semi-arid headwater catchment of Northern Mongolia. *Geosciences (Switzerland)*, 1, 9.
- MUÑOZ SABATER, J. (2019): ERA5-Land monthly averaged data from 1981 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), <https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.68d2bb30?tab=overview> (20. 7. 2021).
- OLSON, K. A., FULLER, T. K., MUELLER, T., MURRAY, M. G., NICOLSON, C., ODONKHUU, D., BOLORTSETSEG, S., SCHALLER, G. B. (2010): Annual movements of Mongolian gazelles: Nomads in the Eastern Steppe. *Journal of Arid Environments*, 11, 74, 1435–1442.
- OPENSTREETMAP (2013): OpenStreetMap: Mongolia buildings, https://wiki.openstreetmap.org/wiki/WikiProject_Mongolia (29. 7. 2021).
- ORACLE (2019): What is a database?, <https://www.oracle.com/database/what-is-database/> (20. 7. 2021).
- PAREDES-TREJO, F. J., BARBOSA, H. A., LAKSHMI KUMAR, T. V. (2017): Validating CHIRPS-based satellite precipitation estimates in Northeast Brazil. *Journal of Arid Environments*, 139, 26–40.
- PEKEL, J. F., COTTAM, A., GORELICK, N., BELWARD, A. S. (2016): High-resolution mapping of global surface water and its long-term changes. *Nature*, 7633, 540, 418–422.
- PHIRI, D., SIMWANDA, M., SALEKIN, S., RYIRENDA, V. R., MURAYAMA, Y., RANAGALAGE, M., OKTAVIANI, N., KUSUMA, H. A., ZHANG, T., SU, J., LIU, C., CHEN, W. H., LIU, H., LIU, G., CAVUR, M., DUZGUN, H. S., KEMEC, S., DEMIRKAN, D. C., CHAIRET, R., BEN SALEM, Y., AOUN, M., KIALA, Z., MUTANGA, O., ODINDI, J., PEERBHAY, K. (2020): remote sensing Sentinel-2 Data for Land Cover / Use Mapping : A Review. *Remote Sensing*, 3, 42, 14.
- QGIS-GEE (2019): QGIS earthengine plugin - integrates Google Earth Engine and QGIS using Python API, <https://gee-community.github.io/qgis-earthengine-plugin/> (20. 7. 2021).
- ROUSE, J. W., HAAS, R. H., SCHELL, J. A., DEERING, D. W. (1974): Monitoring Vegetation Systems in the Great Plains with ERTS. In: *Third Earth Resources Technology Satellite-1 Symposium- Volume I: Technical Presentations*. NASA, Washington D.C., 309–319.
- SALOMONSON, V. V., APPEL, I. (2004): Estimating fractional snow cover from MODIS using the normalized difference snow index. *Remote Sensing of Environment*, 3, 89, 351–360.
- SANTOSO, L. W., YULIA (2017): Data Warehouse with Big Data Technology for Higher Education. *Procedia Computer Science*, 124, 93–99.
- SENTINELHUB (2020): SentinelHub Dashboard, <https://www.sentinel-hub.com/develop/dashboard/> (19. 7. 2021).
- SHEN, M., ZHANG, G., CONG, N., WANG, S., KONG, W., PIAO, S. (2014): Increasing altitudinal gradient of spring vegetation phenology during the last decade on the

- Qinghai-Tibetan Plateau. *Agricultural and Forest Meteorology*, 189–190, 71–80.
- SHIMAMURA, Y., IZUMI, T., MATSUYAMA, H. (2006): Evaluation of a useful method to identify snow-covered areas under vegetation - Comparisons among a newly proposed snow index, normalized difference snow index, and visible reflectance. *International Journal of Remote Sensing*, 21, 27, 4867–4884.
- SIMS, D. A., RAHMAN, A. F., CORDOVA, V. D., EL-MASRI, B. Z., BALDOCCHI, D. D., FLANAGAN, L. B., GOLDSTEIN, A. H., HOLLINGER, D. Y., MISSON, L., MONSON, R. K., OECHEL, W. C., SCHMID, H. P., WOFYSY, S. C., XU, L. (2006): On the use of MODIS EVI to assess gross primary productivity of North American ecosystems. *Journal of Geophysical Research: Biogeosciences*, 4, 111, 1–16.
- SKAKUN, R. S., WULDER, M. A., FRANKLIN, S. E. (2003): Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sensing of Environment*, 4, 86, 433–443.
- SKAKUN, S., VERMOTE, E. F., ROGER, J. C., JUSTICE, C. O., MASEK, J. G. (2019): Validation of the lasrc cloud detection algorithm for landsat 8 images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7, 12, 2439–2446.
- SOWMYA, R., SUNEETHA, K. R. (2014): Data Mining with Big Data. *Proceedings of 2017 11th International Conference on Intelligent Systems and Control, ISCO 2017*, 1, 26, 246–250.
- SPAWN, S. A., SULLIVAN, C. C., LARK, T. J., GIBBS, H. K. (2020): Harmonized global maps of above and belowground biomass carbon density in the year 2010. *Scientific Data*, 1, 7, 1–22.
- STEHMAN, S. V. (1997): Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 1, 62, 77–89.
- ŠTYCH, P., LAŠTOVICKA, J., HLADKÝ, R., PALUBA, D. (2019): Evaluation of the influence of disturbances on forest vegetation using the time series of Landsat data: A comparison study of the Low Tatras and Sumava national parks. *ISPRS International Journal of Geo-Information*, 2, 8.
- TIMESCALE (2018): Time-series data simplified, <https://www.timescale.com/> (20. 7. 2021).
- TOTÉ, C., PATRICIO, D., BOOGAARD, H., VAN DER WIJNGAART, R., TARNAVSKY, E., FUNK, C. (2015): Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique. *Remote Sensing*, 2, 7, 1758–1776.
- UNESCO (2014): Eastern Mongolian Steppes - UNESCO World Heritage Centre, <https://whc.unesco.org/en/tentativelists/5946/> (29. 7. 2021).
- VERBESSELT, J., HYNDMAN, R., NEWNHAM, G., CULVENOR, D. (2010): Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 1, 114, 106–115.
- VERMOTE, E., JUSTICE, C., CLAVERIE, M., FRANCH, B. (2016): Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, 185, 46–56.
- VOGELER, J. C., BRAATEN, J. D., SLESAK, R. A., FALKOWSKI, M. J. (2018): Extracting the

- full value of the Landsat archive: Inter-sensor harmonization for the mapping of Minnesota forest canopy cover (1973–2015). *Remote Sensing of Environment*, January, 209, 363–374.
- WATTS, L. M., LAFFAN, S. W. (2014): Effectiveness of the BFAST algorithm for detecting vegetation response patterns in a semi-arid region. *Remote Sensing of Environment*, 1, 154, 234–245.
- WEHRDEN, H. Von, WESCHE, K., TUNGALAG, R. (2006): Plant communities of the Great Gobi B Strictly Protected Area, Mongolia. *Mongolian Journal of Biological Sciences*, 1, 4, 63–66.
- WEST, M. (1997): Time series decomposition. *Biometrika*, 2, 84, 489–494.
- WILSON, E. H., SADER, S. A. (2002): Detection of forest harvest type using multiple dates of Landsat TM imagery. *Remote Sensing of Environment*, 3, 80, 385–396.
- ZHOU, J., JIA, L., MENENTI, M., GORTE, B. (2016): On the performance of remote sensing time series reconstruction methods – A spatial comparison. *Remote Sensing of Environment*, 187, 367–384.
- ZHU, Z. (2017): Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 370–384.
- ZHU, Z., WOODCOCK, C. E. (2014): Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144, 152–171.

9 Přílohy

Příloha 1: Skript - klasifikace CART	71
Příloha 2: Skript - výpočet indexů.....	73
Příloha 3: Skript - produkce biomasy	76
Příloha 4: Skript - ERA5, teploty	77
Příloha 5: Skript - CHIRPS, srážky.....	79

Příloha 1: Skript - klasifikace CART

```
// nahrání vrstev
/*
var east = ee.FeatureCollection("users/petrjanos/hlavni/mongolsko_2021"),
    zastav = ee.FeatureCollection("users/petrjanos/klasifikace/EAST/budovy_EAST"),
    train = ee.FeatureCollection("users/petrjanos/klasifikace/EAST/testovaci_east3"),
    valid = ee.FeatureCollection("users/petrjanos/klasifikace/EAST/valid_east");
*/

var gobi = ee.FeatureCollection("users/petrjanos/hlavni/gobi_b_land"),
    train = ee.FeatureCollection("users/petrjanos/klasifikace/GOBI/trenovaci_GOBI2"),
    valid = ee.FeatureCollection("users/petrjanos/klasifikace/GOBI/prenost_GOBI2");

// vytvoření bufferu (metry)
function bufferPoints(radius, bounds) {
  return function(pt) {
    pt = ee.Feature(pt);
    return bounds ? pt.buffer(radius).bounds() : pt.buffer(radius);
  }
}
// filtrace mračen, stínů a sněhu: fmask (qa)
function fmask(image) {
  var cloudsBitMask = (1 << 3);
  var cloudShadowBitMask = (1 << 4);
  var snowBitMask = (1 << 5); // filtrace sněhu!
  var qa = image.select('QA_PIXEL');
  var mask = qa.bitwiseAnd(cloudShadowBitMask).eq(0)
    .and(qa.bitwiseAnd(cloudsBitMask).eq(0))
    .and(qa.bitwiseAnd(snowBitMask).eq(0));
  return image.updateMask(mask);
}
// výpočet indexů (jako nová pásma)
function addNDVI(image) {
  var ndvi = image.normalizedDifference(['SR_B5', 'SR_B4']).rename('ndvi')
  return image.addBands([ndvi])
}

// =====

var L8 = ee.ImageCollection('LANDSAT/LC08/C02/T1_L2')
  .filterDate('2020-05-01', '2020-09-30')
  .filterBounds(gobi)
  .map(fmask)
  .map(addNDVI)
  .median()
  .clip(gobi)

var vis = {
  bands: ['SR_B4', 'SR_B3', 'SR_B2'],
```

```

    min: 0.0,
    max: 30000,
  };

Map.addLayer(L8, vis, 'Landsat 8 kompozit');

// Klasifikace =====

// trénovací body s buffrem
var trainGobi = train.map(bufferPoints(30, true));
// použitá pásma
var bands = ['ndvi','SR_B1','SR_B2','SR_B3','SR_B4','SR_B5','SR_B6','SR_B7','ST_B10'];
// třídy ze vstupních dat
var label = 'class';

// Trénink CART
var training = L8.select(bands).sampleRegions({
  collection: trainGobi,
  properties: [label],
  scale: 30
});
var trained = ee.Classifier.smileCart().train(training, label, bands);

// Klasifikace
var classified = L8.select(bands).classify(trained);

// =====

//// aplikace majority filtru, zdroj: Daniel Paluba (@danielp) 2021
// výpočet patch sizes
var patchsize = classified.connectedPixelCount(4, false);
// spuštění majority filtru
var filtered = classified.focal_mode(3, "square");
// updated image with majority filter where patch size is small
var final = classified.where(patchsize.lt(4),filtered);
// výpočet patch sizes
var patchsize2 = final.connectedPixelCount(4, false);
// spuštění majority filtru
var filtered2 = final.focal_mode(1, "square");
// updated image with majority filter where patch size is small
var klasifikace = final.where(patchsize2.lt(4),filtered2);

// hodnocení přesnosti =====

// výstupní klasifikace + kontrolní body
var validation = klasifikace.sampleRegions({
  collection: valid,
  scale: 30,
});

```

```

// spočítání přesnosti a matice
var ConfusionMatrix = validation.errorMatrix('actual', 'classified')
print('Chybová matice', ConfusionMatrix);
print('Celková přesnost', ConfusionMatrix.accuracy());
print('Kappa index: ', ConfusionMatrix.kappa());
// =====

Map.addLayer(klasifikace, {min: 0, max: 3, palette: ['blue', 'lightgreen', 'green',
'purple']}, 'classification_filt');

// =====

Export.image.toDrive({
  image: klasifikace,
  description: 'klasifikace',
  scale: 30,
  region: gobi,
  maxPixels: 1e9
});
/*
Export.table.toDrive({
  collection: ConfusionMatrix,
  folder: 'GEE',
  description: 'ConfusionMatrix_EAST',
  fileNamePrefix: 'ConfusionMatrix_EAST',
  fileFormat: 'CSV'
})
*/

```

Příloha 2: Skript - výpočet indexů

```

// == (1) ZAVEDENÍ FUNKCÍ =====

// vytvoření bufferu (metry)
function bufferPoints(radius, bounds) {
  return function(pt) {
    pt = ee.Feature(pt);
    return bounds ? pt.buffer(radius).bounds() : pt.buffer(radius);
  }
}

// filtrace mračen, stínů a sněhu: fmask (qa)
function fmask(image) {
  var cloudsBitMask = (1 << 3);
  var cloudShadowBitMask = (1 << 4);
  var snowBitMask = (1 << 5); // filtrace sněhu!
  var qa = image.select('QA_PIXEL');
  var mask = qa.bitwiseAnd(cloudShadowBitMask).eq(0)
    .and(qa.bitwiseAnd(cloudsBitMask).eq(0))
    .and(qa.bitwiseAnd(snowBitMask).eq(0));
  return image.updateMask(mask);
}

```

```

}
// výpočet indexů (jako nová pásma)
function addNDVI(image) {
  var ndvi = image.normalizedDifference(['SR_B5', 'SR_B4']).rename('ndvi')
  return image.addBands([ndvi])
}
function addNDMI(image) {
  var ndmi = image.normalizedDifference(['SR_B5', 'SR_B6']).rename('ndmi')
  return image.addBands([ndmi])
}
function addNDSI(image) {
  var ndsi = image.normalizedDifference(['SR_B3', 'SR_B6']).rename('ndsi')
  return image.addBands([ndsi])
}
function addEVI(image) {
  var evi = image.expression(
    '2.5 * ((NIR - RED) / (NIR + 6 * RED - 7.5 * BLUE + 1))', {
      'NIR': image.select('SR_B5'),
      'RED': image.select('SR_B4'),
      'BLUE': image.select('SR_B2')}).rename('evi');
  return image.addBands([evi])
}

// == (2) NAČTENÍ DAT =====

// testovací body
var ptsEast = ee.FeatureCollection("users/petrjanos/hlavni/east_pts2")
var ptsGobi = ee.FeatureCollection("users/petrjanos/hlavni/gobi_fix3")

// spojení bodů, vytvoření bufferu a přiřazení id
var ptsMerge = ptsEast.merge(ptsGobi)
var ptsBuff = ptsMerge.map(bufferPoints(30, true)).map(function(feature) {
  return ee.Feature(feature.geometry(), {'id': feature.id()})
})

// Landsat 8 + filtrace (datum, oblačnos, oblast)
var startDate = '2013-01-01'
var endDate = '2020-12-31'

var L8 = ee.ImageCollection('LANDSAT/LC08/C02/T1_L2')
  .filter(ee.Filter.calendarRange(6,9, 'month'))
  .filterDate(startDate, endDate)
  .filterBounds(ptsBuff)
  .map(fmask)
  .map(addNDVI)
  .map(addNDMI)
  .map(addNDSI)
  .map(addEVI);

print ('Počet snímků původní kolekce L8:', L8.size());

```

```
// == (3) PŘÍPRAVA K EXPORTU =====

// překrývání snímků - průměr hodnot ze stejných dnů
var withDates = L8.map(function(img) {
  return img.set('date', img.date().format('YYYY-MM-dd'))
})
var mosaicList = withDates.aggregate_array('date')
  .distinct()
  .map(function(date) {
    var d = ee.Date(date)
    return L8.filterDate(d, d.advance(1, 'day')).mosaic().set('system:index', date)
  })
// vytvoření mozaiky
var mosaics = ee.ImageCollection.fromImages(mosaicList)

print ('Počet snímků redukované kolekce mosaic:', mosaics.size());

// vytvoření tzv. trojčat (příprava tabulek)
var triplets = mosaics.map(function(image) {
  return image.select('ndvi').reduceRegions({
    collection: ptsBuff,
    reducer: ee.Reducer.mean().setOutputs(['ndvi']),
    scale: 10,
  })// pokud je hodnota maskovaná, dosadí -9999
  .map(function(feature) {
    var index = ee.List([feature.get('ndvi'), -9999])
    .reduce(ee.Reducer.firstNonNull())
    return feature.set({'ndvi': index, 'imageID': image.id()})
  })
}).flatten();
// formátování výsledné tabulky
var format = function(table, rowId, colId) {
  var rows = table.distinct(rowId);
  var joined = ee.Join.saveAll('matches').apply({
    primary: rows,
    secondary: table,
    condition: ee.Filter.equals({
      leftField: rowId,
      rightField: rowId
    })
  });
// spojení na základě ID
return joined.map(function(row) {
  var values = ee.List(row.get('matches'))
  .map(function(feature) {
    feature = ee.Feature(feature);
    return [feature.get(colId), feature.get('ndvi')];
  });
  return row.select([rowId]).set(ee.Dictionary(values.flatten()));
});
```

```

    });
};
// výsledná tabulka
var landsatResults = format(triplets, 'id', 'imageID');

print('Výsledky Landsat 8 SR:', landsatResults.first())

// (3) EXPORT HODNOT =====
Export.table.toDrive({
  collection: landsatResults,
  folder: 'GEE',
  description: 'NDVI',
  fileNamePrefix: 'NDVI',
  fileFormat: 'CSV'
})

// (4) ZOBRAZENÍ BODŮ A VYCENTROVÁNÍ MAPY

Map.addLayer(ptsBuff.draw({color: 'red', strokeWidth: 10}), {}, 'ptsBuff');
Map.centerObject(ptsBuff);

```

Příloha 3: Skript - produkce biomasy

```

var east = ee.FeatureCollection("users/petrjanos/hlavni/mongolsko_2021"),
    gobi = ee.FeatureCollection("users/petrjanos/hlavni/gobi_b_land");

// načtení datové sady
var origo = ee.ImageCollection("NASA/ORNL/biomass_carbon_density/v1");
// výběr nadzemní biomasy
var image = origo.select(['agb'], ['agb'])

// redukce snímku
var med = image.reduce(ee.Reducer.median())

// ořezání území
var clip_east = med.clip(east)

// parametry k vizualizaci - EAST
var vis_east = {
  bands: ['agb_median'],
  min: 0.5,
  max: 2.5,
  palette: ['d9f0a3', 'add8e', '78c679', '41ab5d', '238443', '005a32']
}

// parametry k vizualizaci - GOBI
var clip_gobi = med.clip(gobi)

```

```

var vis_gobi = {
  bands: ['agb_median'],
  min: 0,
  max: 1,
  palette: ['d9f0a3', 'addd8e', '78c679', '41ab5d', '238443', '005a32']
}

```

// načtení v mapě

```

Map.addLayer(clip_east, vis_east,'EAST: carbon_tonnes_per_ha');
Map.addLayer(clip_gobi, vis_gobi,'GOBI: carbon_tonnes_per_ha');
Map.centerObject(east);

```

// příprava exportu

```

Export.image.toDrive({
  image: clip_east,
  description: 'biomass_east',
  scale: 300,
  region: east,
  maxPixels: 1e9
});

```

```

Export.image.toDrive({
  image: clip_gobi,
  description: 'biomass_gobi',
  scale: 300,
  region: gobi,
  maxPixels: 1e9
});

```

Příloha 4: Skript - ERA5, teploty

// == (1) ZAVEDENÍ FUNKCÍ =====

// vytvoření bufferu (metry)

```

function bufferPoints(radius, bounds) {
  return function(pt) {
    pt = ee.Feature(pt);
    return bounds ? pt.buffer(radius).bounds() : pt.buffer(radius);
  }
}

```

// převod na stupně Celsia

```

function addCelsius(image) {
  var temp_2m = image.expression(
    'T - T0', {
      T: image.select('temperature_2m'),
      T0: 273.15}).rename('temp_2m').float();
  return image.addBands([temp_2m])
}

```

// == (2) NAČTENÍ DAT =====

```

// testovací body
var ptsEast = ee.FeatureCollection("users/petrjanos/hlavni/east_pts2")
var ptsGobi = ee.FeatureCollection("users/petrjanos/hlavni/gobi_fix3")

// spojení bodů, vytvoření bufferu a přiřazení id
var ptsMerge = ptsEast.merge(ptsGobi)
var ptsBuff = ptsMerge.map(bufferPoints(300, true)).map(function(feature) {
  return ee.Feature(feature.geometry(), {'id': feature.id()})
})

// Landsat 8 + filtrace (datum, oblačnos, oblast)
var startDate = '2013-01-01'
var endDate = '2020-12-31'

var dataset = ee.ImageCollection("ECMWF/ERA5_LAND/MONTHLY")
  .filterDate(startDate, endDate)
  .filterBounds(ptsBuff)
  .map(addCelsius)

// == (3) PŘÍPRAVA K EXPORTU =====

// překrývání snímků - průměr hodnot ze stejných dnů
var withDates = dataset.map(function(img) {
  return img.set('date', img.date().format('YYYY-MM-dd'))
})
var mosaicList = withDates.aggregate_array('date')
  .distinct()
  .map(function(date) {
    var d = ee.Date(date)
    return dataset.filterDate(d, d.advance(1, 'day')).mosaic().set('system:index', date)
  })
// vytvoření mozaiky
var mosaics = ee.ImageCollection.fromImages(mosaicList)

print ('Počet snímků redukované kolekce mosaic:', mosaics.size());

// vytvoření tzv. trojčat (příprava tabulek)
var triplets = mosaics.map(function(image) {
  return image.select('temp_2m').reduceRegions({
    collection: ptsBuff,
    reducer: ee.Reducer.mean().setOutputs(['temp_2m']),
    scale: 10,
  })// pokud je hodnota maskovaná, dosadí -9999
  .map(function(feature) {
    var index = ee.List([feature.get('temp_2m'), -9999])
      .reduce(ee.Reducer.firstNonNull())
    return feature.set({'temp_2m': index, 'imageID': image.id()})
  })
}

```



```

    })
  }).flatten();
// formátování výsledné tabulky
var format = function(table, rowId, colId) {
  var rows = table.distinct(rowId);
  var joined = ee.Join.saveAll('matches').apply({
    primary: rows,
    secondary: table,
    condition: ee.Filter.equals({
      leftField: rowId,
      rightField: rowId
    })
  });
// spojení na základě ID
return joined.map(function(row) {
  var values = ee.List(row.get('matches'))
    .map(function(feature) {
      feature = ee.Feature(feature);
      return [feature.get(colId), feature.get('temp_2m')];
    });
  return row.select([rowId]).set(ee.Dictionary(values.flatten()));
});
// výsledná tabulka
var tempResults = format(triplets, 'id', 'imageID');

print('Výsledky ERA5', tempResults.first())

// (3) EXPORT HODNOT =====
Export.table.toDrive({
  collection: tempResults,
  folder: 'GEE',
  description: 'temp_2m',
  fileNamePrefix: 'temp_2m',
  fileFormat: 'CSV'
})

// (4) ZOBRAZENÍ BODŮ A VYCENTROVÁNÍ MAPY

Map.addLayer(ptsBuff.draw({color: 'red', strokeWidth: 10}), {}, 'ptsBuff');
Map.centerObject(ptsBuff);

```

Příloha 5: Skript - CHIRPS, srážky

```

var chirps = ee.ImageCollection("UCSB-CHG/CHIRPS/PENTAD");

// vytvoření bufferu (metry)
function bufferPoints(radius, bounds) {

```

```
return function(pt) {
  pt = ee.Feature(pt);
  return bounds ? pt.buffer(radius).bounds() : pt.buffer(radius);
}
}

// testovací body
var ptsEast = ee.FeatureCollection("users/petrjanos/hlavni/east_pts2")
var ptsGobi = ee.FeatureCollection("users/petrjanos/hlavni/gobi_fix3")

// spojení bodů, vytvoření bufferu a přiřazení id
var ptsMerge = ptsEast.merge(ptsGobi)
var ptsBuff = ptsMerge.map(bufferPoints(300, true)).map(function(feature) {
  return ee.Feature(feature.geometry(), {'id': feature.id()})
})

// set start and end year
var startyear = 2013;
var endyear = 2020;

// make a date object
var startdate = ee.Date.fromYMD(startyear, 1, 1);
var enddate = ee.Date.fromYMD(endyear + 1, 1, 1);

// make a list with years
var years = ee.List.sequence(startyear, endyear);
print(years)
// make a list with months
var months = ee.List.sequence(1, 12);
print(months)

var monthlyPrecip = ee.ImageCollection.fromImages(
  years.map(function (y) {
    return months.map(function(m) {
      var w = chirps.filter(ee.Filter.calendarRange(y, y, 'year'))
        .filter(ee.Filter.calendarRange(m, m, 'month'))
        .sum();
      return w.set('year', y)
        .set('month', m)
        .set('system:time_start', ee.Date.fromYMD(y, m, 1));
    });
  }).flatten()
);

print(monthlyPrecip)

var meanMonthlyP = ee.ImageCollection.fromImages(
  months.map(function (m) {
```

```
var w = monthlyPrecip.filter(ee.Filter.eq('month', m)).mean();
return w.set('month', m)
      .set('system:time_start', ee.Date.fromYMD(2000, m, 1));
}).flatten()
);

var title = {
  title: 'Monthly precipitation',
  hAxis: {title: 'Time'},
  vAxis: {title: 'Precipitation (mm)'},
};

var meanMonthlyP = ui.Chart.image.seriesByRegion({
  imageCollection: meanMonthlyP,
  regions: ptsBuff,
  reducer: ee.Reducer.mean(),
  band: 'precipitation',
  scale: 2500,
  seriesProperty: 'id'
}).setOptions(title)
.setChartType('ColumnChart');

var chartMonthly = ui.Chart.image.seriesByRegion({
  imageCollection: monthlyPrecip,
  regions: ptsBuff,
  reducer: ee.Reducer.mean(),
  band: 'precipitation',
  scale: 2500,
  seriesProperty: 'id'
}).setOptions(title)
.setChartType('ColumnChart');

print(chartMonthly);
print(meanMonthlyP);
```